

# The connectivity degree controls the difficulty of RBN reservoir design

Emmanuel Calvet<sup>1, 2</sup>, Bertrand Reulet<sup>1, 3</sup> & Jean Rouat<sup>1, 2</sup>

<sup>1</sup> Université de Sherbrooke <sup>2</sup> Faculté de Génie Électrique et Informatique <sup>3</sup> Faculté des sciences

## Abstract

Reservoir Computing (RC) is a paradigm in artificial intelligence where a recurrent neural network (RNN) is used to process temporal data, leveraging the inherent dynamical properties of the reservoir to perform complex computations. In the realm of RC, the excitatory-inhibitory balance  $b$  has been shown to be pivotal for driving the dynamics and performance of Echo State Networks (ESN) and, more recently, Random Boolean Network (RBN). However, the relationship between  $b$  and other parameters of the network is still poorly understood. This article explores how the interplay of the balance  $b$ , the connectivity degree  $K$  (i.e., the number of synapses per neuron) and the size of the network (i.e., the number of neurons  $N$ ) influences the dynamics and performance (memory and prediction) of an RBN reservoir. Our findings reveal that  $K$  and  $b$  are strongly tied in optimal reservoirs. Reservoirs with high  $K$  have two optimal balances, one for globally inhibitory networks ( $b < 0$ ), and the other one for excitatory networks ( $b > 0$ ). Both show asymmetric performances about a zero balance. In contrast, for moderate  $K$ , the optimal value being  $K = 4$ , best reservoirs are obtained when excitation and inhibition almost, but not exactly, balance each other. For almost all  $K$ , the influence of the size is such that increasing  $N$  leads to better performance, even with very large values of  $N$ . Our investigation provides clear directions to generate optimal reservoirs or reservoirs with constraints on size or connectivity.

## 1 Introduction

Reservoir computing (RC) is a promising approach that could drastically reduce the cost of learning as the input gets projected into a higher dimensional space, *the reservoir*, read out by a single output layer. As such, when the reservoir is adequately designed, a simple linear fitting can be used to train the weights of the readout layer (Maass et al. 2002), alleviating the computational burden of other traditional machine learning methods. The Echo State Network (ESN) developed by (Jaeger 2005) comprises reservoirs with continuous activation functions, while Liquid State Machine (LSM) (Maass et al. 2002) typically includes discontinuous activation functions, among which we find the Random Boolean Network (RBN) (Glass & Hill 1998).

The connectivity degree has been extensively studied in RBN reservoirs, and in contrast to ESN, it displays desirable dynamics for very sparse matrices (Luque &

Solé 2000; Bertschinger & Natschläger 2004; Büsing et al. 2010; Snyder et al. 2012; Echlin et al. 2018), while ESN was shown to be less sensitive to this parameter (Hajnal & Lörincz 2006; Büsing et al. 2010; Krauss et al. 2019a; Metzner & Krauss 2022). On the other hand, it is well known that increasing the number of neurons improves performance (Bertschinger & Natschläger 2004; Snyder et al. 2012; Cherupally 2018; Cramer et al. 2020; Steiner et al. 2023). However, most literature on RBN compared reservoirs with rather small sizes around 1000 neurons (Bertschinger & Natschläger 2004; Natschläger et al. 2005; Büsing et al. 2010; Snyder et al. 2013; Burkow & Tufte 2016), while studies on the ESN compared reservoirs from 500, up to 20000 neurons (Triefenbach et al. 2010).

In this article, we want to study the effect of these topology parameters ( $N$  and  $K$ ), with another control parameter, which is the excitatory-inhibitory balance  $b$ , controlling the proportion of positive and negative synaptic weights (Krauss et al. 2019a; Metzner & Krauss 2022; Calvet et al. 2023). The balance is equal to  $b = (S_+ - S_-)/S$ , with  $S = KN$  the total number of synapses and  $S_{\pm}$  the number of positive and negative synapses. For a positive balance, the network has a majority of excitatory synapses and reverse, and when it is zero, the network has a perfect balance between the two,  $S_+ = S_-$ . Previous work on the ESN (Krauss et al. 2019b,a; Metzner & Krauss 2022) has studied the influence of density  $d = K/N$  and balance on the dynamics of reservoirs, showing that  $b$  was a key parameter controlling phase transitions. In particular, the edge of chaos, a dynamical phase transition between order and chaos, is believed to be fundamental in reservoir design, for a reservoir to acquire computational capabilities (for a comprehensive introduction, see (Legenstein & Maass 2006; Nowshin et al. 2020)). Nevertheless, Metzner *et al.* suggested a more complex picture than previously thought, exposing two critical points, each for a positive and negative balance, while for higher densities, an asymmetry could arise in the reservoir responses to inputs, and as a result, only the edge of chaos occurring for positive  $b$  was optimal for information propagation inside the reservoir (Metzner & Krauss 2022).

In line with Krauss and Metzner, recent work on RBN reservoirs demonstrated that the excitatory-inhibitory balance  $b$  was also key in driving dynamics and performance (Calvet et al. 2023). In particular, it was shown that the weight statistics, typically used in RBN literature (Bertschinger & Natschläger 2004; Natschläger et al. 2005; Büsing et al. 2010) are related to the balance. More striking, the RBN reservoirs also displayed an asymmetry around  $b = 0$ . The two signs of the balance produced distinct relations to performance in tasks and a reduced reservoir-to-reservoir variability for a majority of inhibition. However, this occurred for a network with extremely low density as  $d = K/N = 16/10000 = 0.0016$ , in contrast with studies on ESN.

As far as the authors are aware, the influence of the excitatory-inhibitory balance for different connectivity has yet to be studied, except for the single value of  $K = 16$  previously mentioned (Calvet et al. 2023). This article aims to explore the combined effect of connectivity ( $K$ ,  $N$ ) and the balance on the dynamics and performance of the RBN. The article is organized as follows: in the first section (Sec. 3.1) the effect of  $K$  and  $b$  is studied, both on the dynamics of free-evolving reservoirs (Sec. 3.1.1), and their performance in a memory and prediction task (Sec. 3.1.2), showing that the asymmetry in fact vanishes for very small  $K$ . In the second section (Sec. 3.2) we perform a similar analysis (dynamics in Sec. 3.2.1, and performance in Sec. 3.2.2),

but this time, we vary both  $K$  and  $N$  conjointly, and explore the relationship with  $b$ . This reveals a complex interplay between parameters and suggests that  $K$  is, in fact, governing it. Finally, in Sec. 4, we discuss our results and their implication for RBN reservoir design, revealing that in contrast with  $ESN$ , the careful selection of  $K$  leads to a significant simplification of the fine-tuning of the other topology parameters in the tested tasks.

## 2 Methodology

### 2.1 The model

Our model is an ensemble of three parts (Fig. 1), the input node  $u(t)$ , which is projecting to half of the neurons of the recurrently connected reservoir  $\vec{x}$ , among which the other half is projecting to the output node  $y(t)$ , this way, the output node never directly sees the input, and information must propagate inside the reservoir for the readout to accomplish the task at hand:

$$u_i(t) = w_i^{in} u(t) \quad (1)$$

$$y(t) = f(W^{out} \vec{x} + c) \quad (2)$$

With  $u_i(t)$  the input of the neuron  $i$ , the input weights  $w_i^{in}$  form a vector, projecting to half of the reservoir, while the other half of the weights are zeros, and reserve for the output weight matrix  $W^{out}$ . This way, a neuron in the reservoir is never connected to both the input and output. The activation function  $f$  of the output node is the sigmoid, with a bias  $c$ . Each component  $x_i(t)$  of  $\vec{x}(t)$  corresponds to the state of the neuron  $i$  inside the reservoir. It is given by:

$$x_i(t) = \theta \left( u_i(t) + \sum_{j=1}^N w_{ij} x_j(t-1) \right) \quad (3)$$

Where each neuron is connected to  $K$  other neurons, and  $w_{ij}$  is the synaptic weight connecting neuron  $j$  to neuron  $i$ , drawn in a normal distribution  $\mathcal{N}(\mu, \sigma)$ , with parameters  $\mu$  (mean) and  $\sigma$  (standard deviation). The activation function  $\theta$  is a Heaviside, thus  $x_i$  is binary.  $t \in \mathbf{N}$  and corresponds to a time step. Remark that if the input is zero, the state of a given neuron only depends on the states of its neighbours at the previous time step. Such neurons are thus said to be "memoryless", and for such a system, to sustain memory, information needs to cascade via the propagation of spikes inside the reservoir. The attractive feature of the reservoir framework is that only the output weight and bias are trainable parameters, as all other parameters are usually kept fixed, including the reservoir weights.

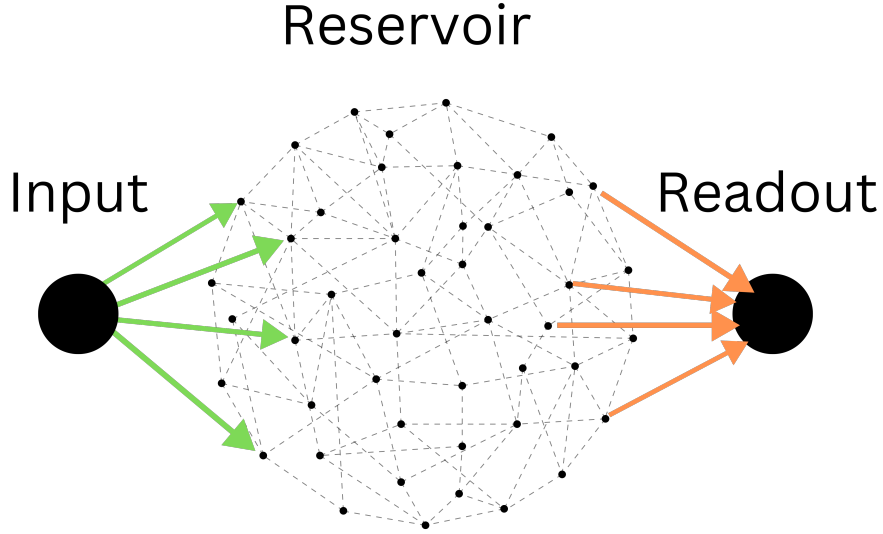


Figure 1: The model consists of an input node (left), connected by input weights (green arrows), to the reservoir (center), itself connecting via output weights (orange arrows) to the output node (right). As illustrated by the dotted black arrows, the reservoir is recurrently connected, forming a random graph. The illustrated graph has  $K = 2$  and  $N = 22$ . Note that in practice, half of the neurons (blue circles) connect to the input, and the other half to the readout.

We use a mean square error (MSE) loss function for the training process. For training the readout weights, we opted for the ADAM optimizer (Kingma & Ba 2015), providing superior results in our testings, superseding the commonly utilized Ridge regression (Burkow & Tufte 2016) in most literature. The execution is facilitated through the PyTorch library, with parameters set at  $\alpha = 0.001$  and 4000 epochs (Supplementary Materials 6.1 for additional information).

## 2.2 The control parameters

The three control parameters used in this study are  $\sigma^*$ ,  $K$ , and  $N$ . Among these,  $\sigma^*$  represents the coefficient of variation of the weight distribution within the reservoir, defined as  $\sigma^* = \sigma/\mu$ . This parameter is linked to  $b$ , the excitatory/inhibitory balance, as  $b = \text{Erf}[1/(\sqrt{2}\sigma^*)]$  (Calvet et al. 2023). The balance is also equal to  $b = (S_+ - S_-)/S$ , with  $S$  the total number of synapses, and  $S_{\pm}$  the number of positive and negative synapses, respectively. We display in Fig. 2 the relationship between the two, noting that when  $\sigma^*$  is positive, we have a majority of excitatory synapses  $b > 0$ , and when  $\sigma^*$  is negative, we have a majority of inhibitory synapses  $b < 0$ . In all experiments, we play with values of  $\sigma^*$  that allow our reservoirs to span the full range of  $b$ , corresponding to  $\sigma^* \in [10^{-2}, 10^3]$ .

Since recent work showed that the dynamics and performance of reservoirs were asymmetric about  $b = 0$  (Metzner & Krauss 2022; Calvet et al. 2023), we study the influence of two other control parameters with respect to the sign of  $b$ . These parameters are captured by the density  $d = K/N$ , following the work of (Hajnal & Lőrincz 2006; Krauss et al. 2019a; Metzner & Krauss 2022) on ESN. However, we

show in supplementary material 6.2 that the density  $d$  is not a control parameter for the RBN, since, at a fixed density, reservoirs can possess very different dynamics as  $K$  and  $N$  are concurrently varied. As such, we consider them as independent control parameters in this article. Following work in RBN (Büsing et al. 2010; Calvet et al. 2023), the connectivity degree is chosen between 1 and 16. In addition, to compare the more recent results ( $N = 10000$ ) (Calvet et al. 2023) with older literature ( $N \leq 1000$ ) (Bertschinger & Natschläger 2004; Natschläger et al. 2005; Büsing et al. 2010; Snyder et al. 2013; Burkow & Tufte 2016), we study three values of  $N = \{100, 1000, 10000\}$ .

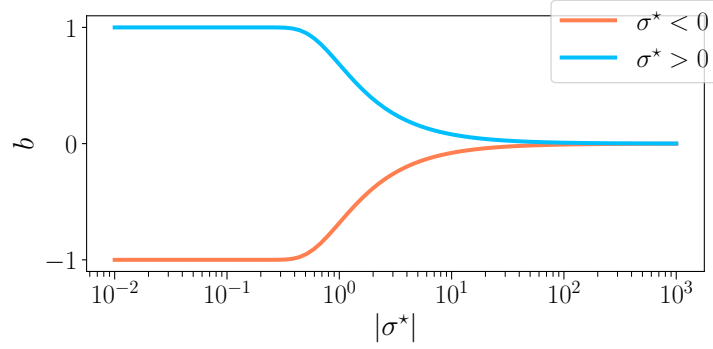


Figure 2: The excitation-inhibition balance  $b$  as a function of the synaptic weight parameter  $\sigma^*$ . For  $\sigma^* < 0$  (—) and  $\sigma^* > 0$  (—).  $\sigma^*$  is the coefficient of variation ( $\sigma/\mu$ ) of the weight distribution, which is why  $b$  is of the sign of  $\sigma^*$ . For low values of  $|\sigma^*|$ , only  $\mu$  controls the synaptic balance, meaning that for a positive mean, the weights are all excitatory, and reverse. On the other hand, when  $|\sigma^*| \rightarrow \infty$ , the mean becomes irrelevant, and  $b$  is at a perfect balance between excitation and inhibition.

## 2.3 The experiments

We perform two types of tasks: the first to probe the intrinsic dynamics of reservoirs, while they are freely evolving, and the second to test the ability to process inputs while performing memory and prediction tasks.

### 2.3.1 Free-running

Each reservoir is freely running without input for a duration of  $D = 2000$  time steps, with a random initial state with 20% of neurons to one. During a run, the activity signal  $A(t)$  (the average of states  $x_i$  at a given time step  $t$ ) is recorded. Afterward, we compute the BiEntropy ( $H_b$ ) (Croll 2014) of the binarized activity signals, which quantify the degree of order and disorder of a bit string,  $H_b = 0$  for completely periodic, and  $H_b = 1$  for totally irregular. For each triplet  $(N, K, \sigma^*)$ , we randomly generate 100 reservoirs, and we then compute the average and variance over reservoirs having the same control parameters.

Next, we classify the steady-state activity  $A(t)$ , for  $t > 1000$  time steps, into four distinct attractor categories. For each triplet  $(N, K, \sigma^*)$ , we then compute the histograms over the 100 reservoirs. The attractors are defined according to (Calvet et al. 2023) :

- **Extinguished:** The activity has died out, and the steady activity is zero at all time steps.
- **Fixed attractor:** The steady activity is non-zero, but its derivative is zero at all time steps.
- **Cyclic:** The steady activity repeats, with a period larger than one time-step.
- **Irregular:** If none of the above categories apply, the signal is irregular. Note that our model is deterministic and discrete, as such, all attractors are in theory, cyclic; however, since the duration  $D = 2000$  is extremely small compared to the maximal period of  $2^N$ , in practice, we find a statistically significant proportion of attractors in that category.

### 2.3.2 Performance in tasks

To test the computational capabilities of our reservoirs, we perform two distinct tasks. The first one consists of memorizing white-noise input received  $|\delta|$  time steps in the past. We test our reservoirs with various difficulties for  $\delta = \{-18, -14, -10, -6, -2\}$ . The higher in absolute value, the more difficult the task, since it demands the reservoir of memoryless neurons to integrate and reverberate input information through spikes cascade for longer time scales (Metzner & Krauss 2022; Calvet et al. 2023). The second task consists of predicting Mackey-Glass time series,  $\delta = 10$  time steps in the future. Mackey-Glass is a common benchmark in reservoir computing (Hajnal & Lörincz 2006; Bianchi et al. 2016; Zhu et al. 2021). We use  $\tau$ , the time constant parameter of Mackey-Glass, which controls the signal dynamics, ranging from  $\tau = 5$  (periodic),  $\tau = 15$ , to  $\tau = 28$  (chaotic). To evaluate the performance of our reservoir, we compute the correlation coefficient  $Corr(y, T)$  between the target  $T$ , and the output  $y$  (implementation details are identical to this study (Calvet et al. 2023)).

## 3 Results

### 3.1 The connectivity degree controls the optimal balance

In this section, we fix the size of the reservoir to its largest value  $N = 10000$ . We study the effect of  $K$  and  $b$  on the dynamics of free-running reservoirs (Sec. 3.1.1). Then, we study the performance in two demanding tasks (Sec. 3.1.2). We show that the asymmetry about  $b = 0$  is strongly  $K$  dependent and vanishes for low  $K$ , while the optimal balance  $b_{opt}$  is entirely controlled by  $K$ .

Additionally, we exhibit the shift of control parameters from the more natural weight distribution statistics ( $\sigma^*$ ) (Calvet et al. 2023) to the excitatory-inhibitory balance ( $b$ ). To do so, we begin by exhibiting the dynamics over  $\sigma^*$ , to then display the attractor statistics over the excitatory balance  $b$ , revealing insights into the reservoir design.

### 3.1.1 Impact of the connectivity degree and balance on dynamics

In Fig. 3, we display the average over reservoirs of the BiEntropy of the steady activities for reservoirs as a function of  $|\sigma^*|$  (lower x-axis), both with a negative (left) or positive (right) balance  $b$  (the upper x-axis displays the corresponding  $b$  values). In Fig. 3.A and B, blue regions represent an ordered phase with low BiEntropy, and red regions represent a disordered phase with a BiEntropy close to one. The regions are separated by a phase transition where the BiEntropy is intermediate, also captured by the variance of the BiEntropy (Fig. 3.C and D). The scenario is similar for both signs of  $b$  but differs in the details. The transition (abrupt for  $b < 0$ , wider for  $b > 0$ ) occurs at a value of  $\sigma^*$  that depends on  $K$  (strongly for  $b < 0$ , weakly for  $b > 0$ ). The transition widens when  $K$  decreases (strongly for  $b > 0$ ). At high  $K$ , i.e. when each neuron is connected with many, there seems to be an asymptotic value for  $\sigma^*$  (or  $b$ , indicated on the upper part of the plots), which is different for  $b > 0$  and  $b < 0$  (Calvet et al. 2023). For  $K = 2$ , the disordered phase never reaches a BiEntropy of 1, and for  $K = 1$  the reservoir is always in its ordered phase (Bertschinger & Natschl ger 2004).

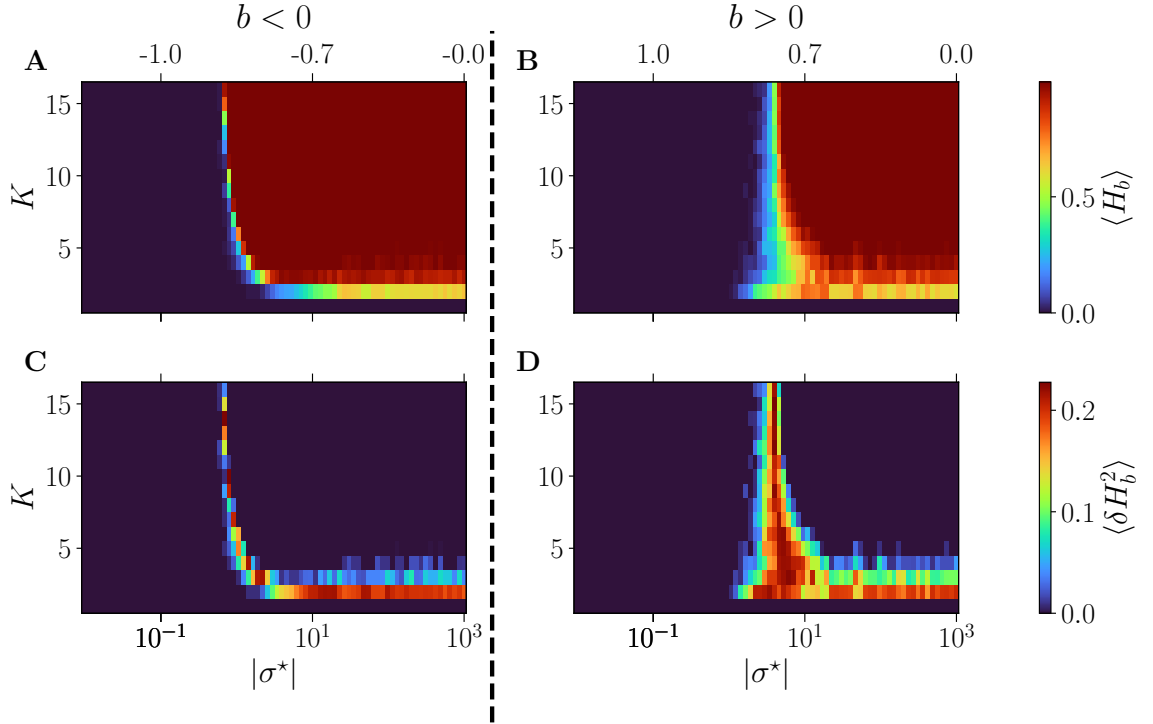


Figure 3: Dynamics of free evolving reservoirs as controlled by the connectivity degree  $K$  (y-axis) and  $|\sigma^*|$  (x-axis). The upper x-axis displays the corresponding  $b$  values, for  $b < 0$  (A and C), and  $b > 0$  (B and D). The BiEntropy is computed on the steady activities of 100 reservoirs per couple ( $K$ ,  $\sigma^*$ ). (A and B): The upper row displays the average BiEntropy of the steady state activities (upper left colormap). (C and D): The lower row shows the variance of BiEntropy over reservoirs (bottom left colormap).

In Fig. 4, we plot the statistics of attractors for reservoirs with  $K = 16$  (upper panel),  $K = 8$  (middle), and  $K = 4$  (lower), as a function of the balance  $b$ . This time  $|\sigma^*|$  is reported in the upper x-axis. The left column shows the results for  $b < 0$  and the right column for  $b > 0$ . The phase transition is characterized by going from attractors

with essentially no ( $b < 0$ ) or fixed ( $b > 0$ ) activity in the ordered phase, to attractors being all irregulars in the disordered phase, with cyclic attractors showing up at the transition. In all plots, we report the non-zero BiEntropy variance (highlighted by light-grey hatching) to indicate the critical region (Calvet et al. 2023). This transition region is clearly defined for  $K = 16$ , widens for  $K = 8$  and becomes very different for  $K = 4$ . When  $b < 0$ , there is a transition region around  $b \sim -0.7$  (gray hashed region) and a re-entrance of the critical region (orange hatching in Fig. 4.E). Indeed, for  $b$  between  $-0.7$  and  $-0.08$  ( $\sigma^*$  between  $-10$  and  $-2$ ) all attractors are irregular, and cyclic ones reappear for a balance closer to zero. For  $K = 4$  and  $b > 0$  the phase transition is never complete, there is no fully disordered phase. Lastly, near  $b = 0$ , the attractor statistics are very close from one sign to the other. For example, with  $K = 16$  and  $K = 8$  we observe an horizontal line for chaotic attractors, while for  $K = 4$ , the statistics of cyclic and irregular attractors closely match on both sides, a fact that is even more visible in the results of Sec. 3.2.1 when varying  $N$ .



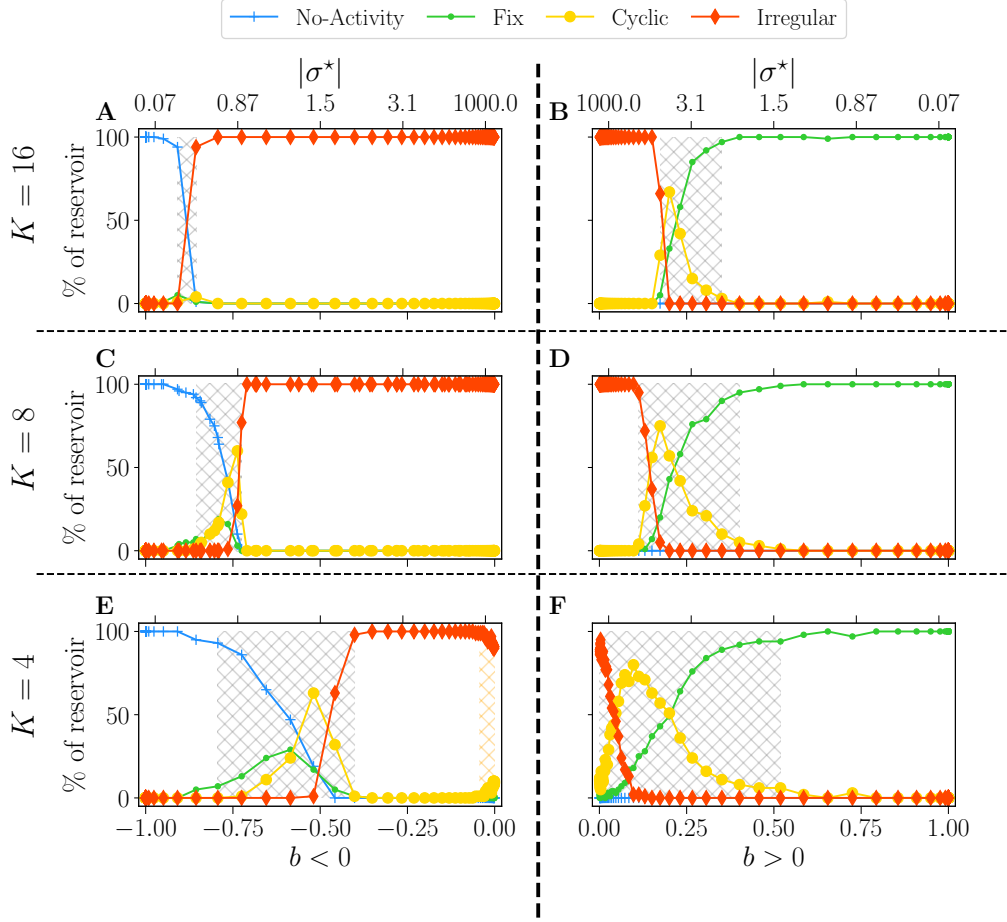


Figure 4: Attractor statistics of free-evolving RBN reservoirs, controlled by  $K$  (rows), and the balance  $b$  (x-axis). The upper x-axis represents the corresponding  $|\sigma^*|$  values, for  $b < 0$  (A, C, E), and  $b > 0$  (B, D, F). All reservoirs are of size  $N = 10000$ . Each steady activity signal is classified into one of the four categories of attractors: no-activity (+), fix (•), cyclic (•), irregular (♦). The statistics of attractors are computed over 100 reservoirs run once (y-axis). Results are shown for  $K = 16$  (A and B),  $K = 8$  (C and D), and  $K = 4$  (E and F). The light-gray hatched areas represent the critical regions (Calvet et al. 2023), defined as the region of non-zero BiEntropy variance; the threshold is chosen to 0.0001. In E, the orange hatched area represents a region of re-entrance of criticality with non-zero BiEntropy variance, distinct from the critical region. All hatched areas are computed from the data shown in Fig. 3.C and D.

Regarding the control parameter shift from  $\sigma^*$  to  $b$ , the phase transition appears inflated in  $b$ , as indicated by the dot positions, particularly for  $b < 0$ . These positions are generated on an evenly spaced logarithmic scale in  $\sigma^*$ . The irregular regime is notably compressed, demonstrated by the re-entrant critical region (refer to Fig. 4.E), spanning from  $2 \cdot 10^1$  to  $10^3$ . This observation suggests that the dynamics remain relatively consistent despite significant variations in the weight distribution parameter. In line with (Metzner & Krauss 2022; Calvet et al. 2023), we make the case that underlying  $b$  is what is driving the dynamics of these reservoirs. As such, in the rest of the article, we use  $b$  as a reference for all further investigations.

In conclusion,  $K$  has a strong influence on the dynamics of the network. For large values of  $K$ , a variety of attractors can be found only in a narrow region of  $b$  ( $\sigma^*$ ), which is different for both signs of the balance. In contrast, for lower values of  $K$ , the co-existence of several attractors is found over a very wide range of  $\sigma^*$  which

corresponds to the region where  $b$  is small, positive or negative.

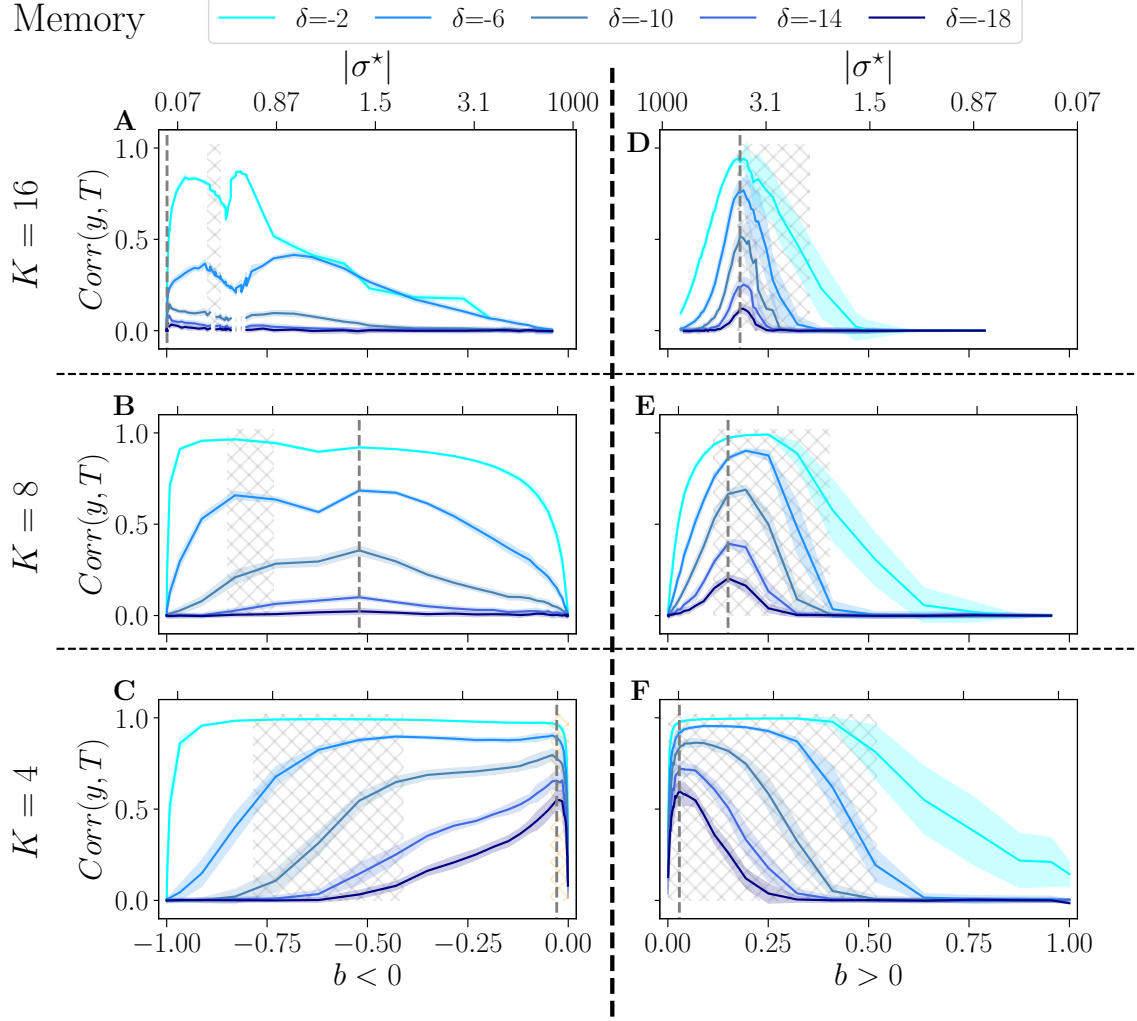


Figure 5: Performance of RBN reservoirs in the memory task of white-noise signals, for various  $\delta$ , the higher in absolute, the more difficult the task. The correlation between the target and the output (y-axis), is plotted as a function of the control parameter  $b$  (x-axis), for a positive balance (A, B, C), and a negative balance (D, E, F). The upper x-axis represents the corresponding  $|\sigma^*|$  values. The solid lines represent the average over 20 reservoirs, higher values signify better performance, while the shaded area represents one standard deviation. (A and E): the upper row displays  $K = 16$ , the middle row (C and F)  $K = 8$ , and bottom row (D and G)  $K = 4$ . The light-gray hatched areas represent the critical regions of BiEntropy variance above a threshold of 0.0001, and the dotted gray lines represent the optimal balance  $b_{opt}$  in the most difficult task.

### 3.1.2 Impact of the connectivity degree and balance on performance

In Fig. 5, we show the performance of the reservoirs for memory tasks as a function of the control parameter  $b$  ( $|\sigma^*|$  upper x-axis). Five difficulties are operated, with  $\delta$  varying from  $-2$  to  $-18$ . The left column comprises reservoirs with a negative balance and the right column with a positive one. We show the results for  $K = 16$  (upper row),  $K = 8$  (middle), and  $K = 4$  (bottom).

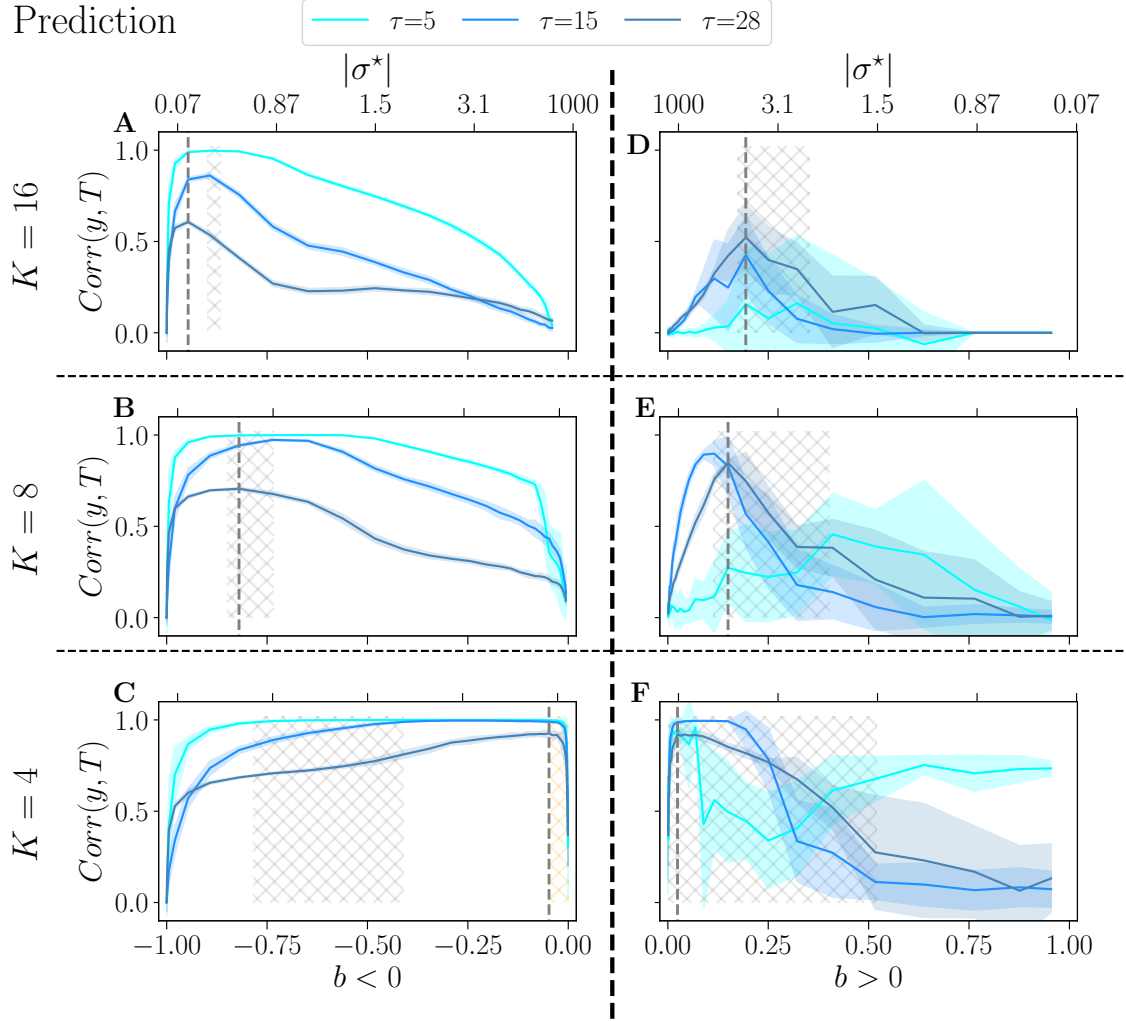


Figure 6: Performance of RBN reservoirs in the prediction task of Mackey-Glass time series, for various  $\tau$ , the higher, the more complex the signal. The correlation between the target and the output ( $y$ -axis), is plotted as a function of the control parameter  $b$  ( $x$ -axis), for a positive balance (A, B, C), and a negative balance (D, E, F). The upper  $x$ -axis represents the corresponding  $|\sigma^*|$  values. The solid lines represent the average over 20 reservoirs, higher values signify better performance, while the shaded area represents one standard deviation. (A and E)  $K = 16$ , with similar result to (Calvet et al. 2023), (C and F)  $K = 8$ , and (D and G)  $K = 4$ . As in the previous figure, the light-gray hashed areas represent the phase transition region, and the dotted gray lines represent the optimal balance  $b_{opt}$  in the most difficult task.

For each value of the delay, reservoirs perform better at low  $K$ , and show good performance over a broader range of  $b$ . Similar observations have been reported for other tasks (Büsing et al. 2010). The balance for which performance is best  $b_{opt}$  (dotted gray line) strongly depends on  $K$ : this is the most visible for  $b < 0$  and  $\delta = -18$  (the most difficult task), where  $b_{opt}$  goes from almost  $-1$  for  $K = 16$ , to almost  $0$  for  $K = 4$  (see Tab. S1). For other values of  $\delta$  the effect is less pronounced but clearly always present. For  $b > 0$  the same phenomenon appears and  $b_{opt}$  shifts from  $\sim 0.2$  for  $K = 16$ , to  $\sim 0$  for  $K = 4$ . Thus, the asymmetry between  $b > 0$  and  $b < 0$  fades as  $K$  decreases. For  $K = 4$ , the optimal balance, whether positive or negative, is almost zero, i.e., it corresponds to an almost perfect balance between excitation and inhibition. However, notes that performance drops abruptly for  $b = 0$ : the unbalance,

even very small, is essential.

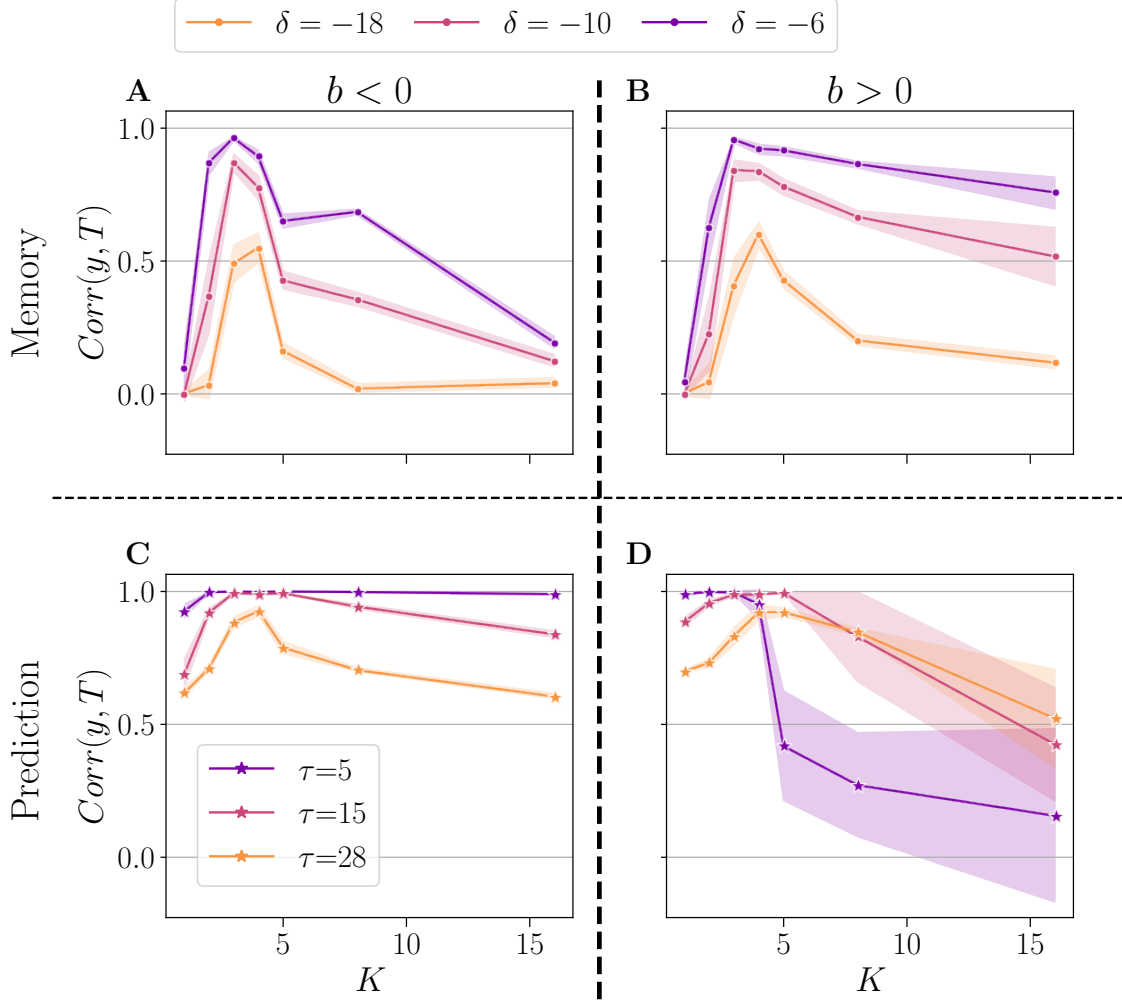


Figure 7: Summary of performance for various connectivity degrees  $K$ , in the memory (A and B) and the prediction (C and D) tasks. For both  $b < 0$  (left panel), and  $b > 0$  (right panel). For each value of  $K$ , we selected the  $b_{opt}$  value giving the highest average performance, in the most difficult task ( $\delta = -18$  for memory, and  $\tau = 28$  for prediction). We plot the performance (higher is better) of reservoirs  $Corr(y, T)$  (y-axis), plotted as a function of  $K$  (x-axis). The solid line represents the average over 20 reservoirs (generated with the same  $b_{opt}$  and  $K$  value), and the shaded area represents one standard deviation. Performance is shown for various  $\delta$  in the memory task (A and B), and  $\tau$  in the prediction (A and B).

In the prediction task (Fig. 6), a similar trend is observed: as  $K$  decreases, the high-performing region shifts towards  $b$  values close to zero. Furthermore, the range of  $b$  values within the high-performing region is also broader. Still, for  $K = 4$ , our task may not be sufficiently challenging for the reservoirs, since at  $b_{opt}$ , the three values of  $\tau$  give very close results. When  $b < 0$ , the critical region (gray hashed area) does not align well with the performance peaks, and this discrepancy is even more pronounced for lower  $K = 4$ . The peak of performance is still within the orange-hashed region, indicative of re-entrant criticality. In the case where  $b > 0$ , in line with previous work (Calvet et al. 2023), the variance is exceptionally high, especially for simpler signals  $\tau = 5$  and  $\tau = 15$ . Surprisingly, for  $K = 16$  and  $K = 4$ , reservoirs perform better at the complex task than at the simpler task  $\tau = 5$ .

Trying to relate criticality with peak performance, we observe that if there is a link between the two, it is rather loose. For  $b < 0$  the region of best performance is much broader than the critical region, indicated as hatched gray areas. In many cases,  $b_{opt}$  does not lie within the critical region. For  $b > 0$ , criticality and optimal performance seem more correlated, as optimal performance is usually obtained within the critical region. However, focusing on  $K = 4$ ,  $b < 0$  and the hardest memory task (Fig. 5C), there is a striking difference between criticality and optimal performance: performance is almost zero in the critical region while it peaks in the region of re-entrance observed in the dynamics of the free running reservoirs, indicated in Fig. 5C as an orange hatched area. Both regions show a variety of attractors, but only one corresponds to good performance.

To conclude, in Fig. 7, we show a summary of the best performance in the memory (upper panel), and prediction (lower panel). In the plot, each dot represents the average over 20 reservoirs obtained with the same connectivity parameters ( $N$ ,  $K$ ,  $b_{opt}$ ), where  $b_{opt}$  is the value that maximizes the average performance at the most difficult setting of each task ( $\delta = -18$  and  $\tau = 28$ ), see Tab. S1 and Tab. S2. As previously, we separated the case  $b < 0$  (left panel) and  $b > 0$  (right panel). We compare the performance for  $K = 1$  up to 16.

For all tasks, we note that the highest performances are consistently achieved with  $K = 3$  and  $K = 4$ , irrespective of whether  $b$  is positive or negative. However, the optimal value of  $K$  exhibits some task dependency. In the memory task, for the more challenging task ( $\delta = -18$ ),  $K = 4$  yields the best performance, despite  $K = 3$  occasionally outperforming less demanding tasks. This suggests that the optimal  $K$  may depend on the complexity of the task at hand. The sign of  $b$  has no discernible impact on the optimal  $K$ , however, it is observed that the performance for higher  $K$  values is superior when  $b > 0$ , in line with (Calvet et al. 2023).

In the prediction task, again, the most challenging setup ( $\tau = 28$ ) shows  $K = 4$  as the optimal value, irrespective of the sign of  $b$ . In general, the reservoir-to-reservoir variance is very small for  $b < 0$ . As previously observed, for higher  $K$ , we observe a significant reservoir variability, and this time, the performance is higher when  $b < 0$ .

Taken together, these findings suggest that once an optimal value for  $K$  is selected, the system's performance becomes mainly insensitive to the sign of the balance  $b$ , even though the optimal  $K$  can be dependent on the task at hand.

### 3.1.3 Discussion

In line with (Calvet et al. 2023), for a positive balance, the critical region is reasonably aligned with the performing region, for all tested  $K$ . Yet our findings somewhat challenge the idea that the edge of chaos is always optimal for computation, as it does not necessarily overlap with the region of best performance. This is especially visible in the memory tasks and reservoirs with a negative balance. Indeed, for  $K = 4$ , the re-entrant region provides the best reservoirs, while being very far from criticality.

By looking at dynamics, one might wonder if this re-entrant region of attractor diversity ( $b < 0$ ) does not belong to the critical region of the positive side, which, by shifting towards the left, overlaps on the negative sign. On the other hand, we

observe a drastic dip in performance with both signs around  $b = 0$ . This suggests that a breaking of symmetry is at play, acting as a crucial driver for performance while being surprisingly imperceptible in the dynamic.

Regarding reservoir design, we show that the optimal excitatory/inhibitory balance is intricately tied to the number of connections. For a high number of connections, a pronounced asymmetry is observed depending on whether there is a majority of inhibition or excitation.

However, when  $K = 4$ , the optimal  $b$  value is almost identical and closely balanced between excitation and inhibition, regardless of whether  $b$  is positive or negative. Consequently, the dynamics of reservoirs are nearly identical for both positive and negative  $b$ , resulting in similar performance outcomes. The task of choosing the optimal  $b_{opt}$  becomes much simpler, as the asymmetry fades away.

### 3.2 The interplay between reservoir size and connectivity degree

This section studies the joint effect of the reservoir size  $N$  ( $=100, 1000, 10000$ ) and  $K$ , in relation to  $b$ . We show that  $N$  has a comparable impact on the dynamics as  $K$ , but also impacts asymmetrically around  $b$  the performance in tasks.

#### 3.2.1 Impact of reservoir size and connectivity degree on dynamics

In Fig. 8, we set  $K = 4$  and present the attractor statistics over  $b$  for three different values of  $N$ :  $N = 10000$  (upper panel),  $N = 1000$  (middle panel), and  $N = 100$  (lower panel). We analyze these values in two cases,  $b < 0$  (left panel) and  $b > 0$  (right panel).

From our observations, it is evident that reducing  $N$  leads to a decrease in the complexity of the attractors, as indicated by the reduction of irregular attractors. In the case of  $b < 0$  and as  $N$  decreases, the re-entrant region (orange hashed area) observed with  $N = 10000$  (Fig. 8.A) merges with the critical one (gray hashed area) for  $N = 1000$  (Fig. 8.C), resulting in a spike of irregular attractors and eventually leaving room for predominantly cyclic ones as  $N = 100$  (Fig. 8.C).

Contrarily, for  $b > 0$  and  $N = 1000$ , this spike or irregular attractor is missing, and the critical phase is largely dominated by cyclic attractors, with only a few fixed and irregular ones. Interestingly, when  $N = 100$ , both signs yield very similar results, with no irregular attractors at all. This observation underscores the impact of  $N$  on the nature and complexity of the attractors.

Lastly, when discussing Fig. 4, we briefly mentioned the continuity in attractor statistics as going from a negative to a positive balance. This fact is even more salient in Fig. 8. Statistics of attractors closely match on both sides, reinforcing the picture that the critical region can span both signs, at least from the dynamic lens.

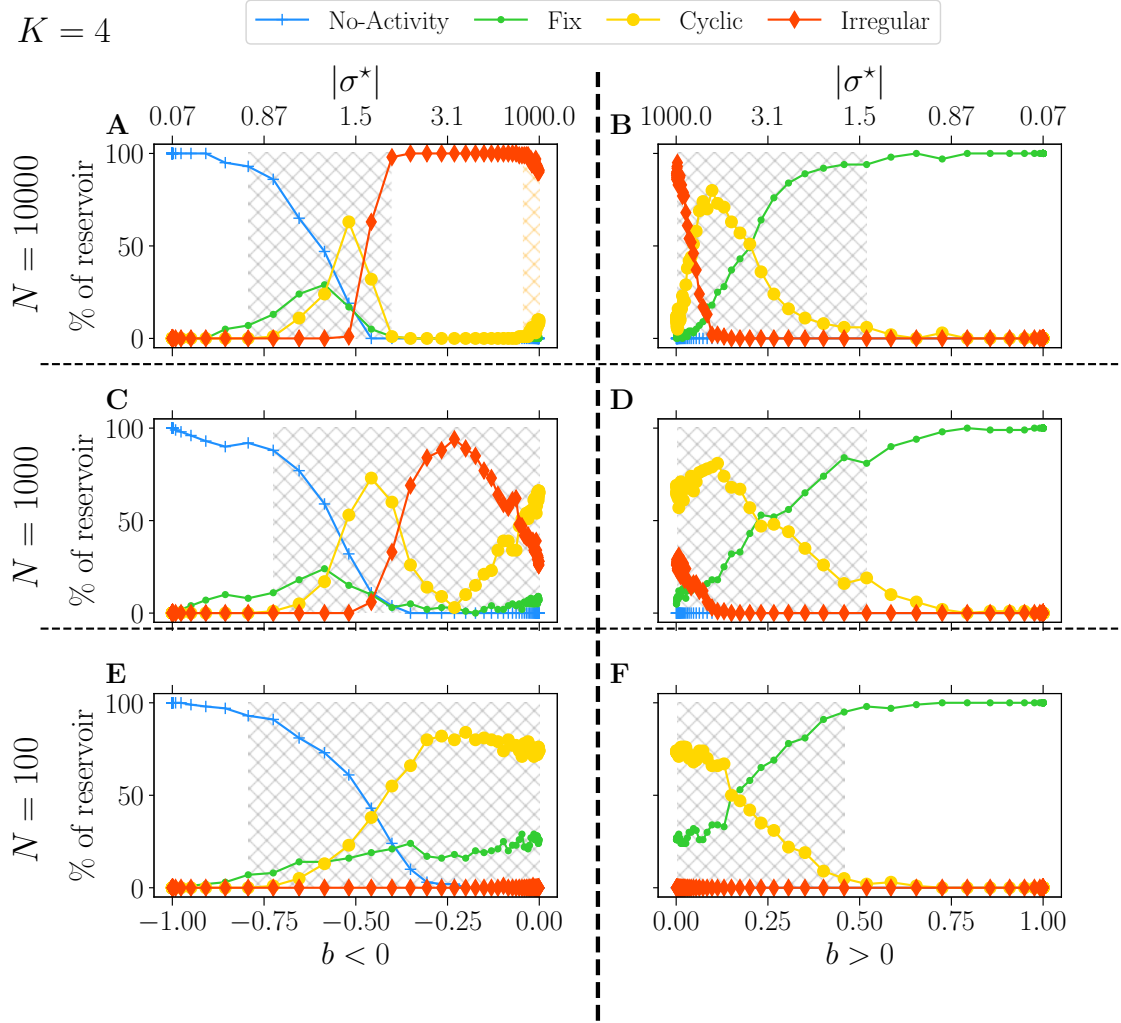


Figure 8: Attractor statistics of free-evolving RBN reservoirs for fixed  $K = 4$ , with  $N = 10000$  (A and B),  $N = 1000$  (D and D), and  $N = 100$  (E and F). Statistics of attractors over 100 reservoirs run once (y-axis) versus  $b$  (x-axis). The upper x-axis displays the corresponding  $|\sigma^*|$ , both for  $b < 0$  (left panel) and  $b > 0$  (right panel). Each activity signal is classified into one of the six categories of attractors: extinguished, fixed, cyclic, and irregular, defined in methodology Sec. 2.3. The light-gray hatched areas represent the critical regions defined; the threshold is chosen to 0.0001. In A, the orange hatched area represents the region of re-entrance of criticality with non-zero BiEntropy variance.

### 3.2.2 Impact of reservoir size and connectivity degree on performance

Results for the memory task and prediction are respectively displayed in Fig 9 and Fig. 10. We tested the performance for  $K = 4$  (upper panel),  $K = 8$  (middle) and  $K = 16$  (bottom). Reservoirs with  $b < 0$  are displayed in the left panel and  $b > 0$  in the right panel. We compare the performance for three distinct values of  $N$ :  $N = 10000$  (green curves),  $N = 1000$  (orange curves), and  $N = 100$  (blue curves). As in the previous Sec. 3.1.2, performance is shown for  $b_{opt}$ , established for the most difficult setting in each task ( $\delta = -18$  and  $\tau = 28$ ).



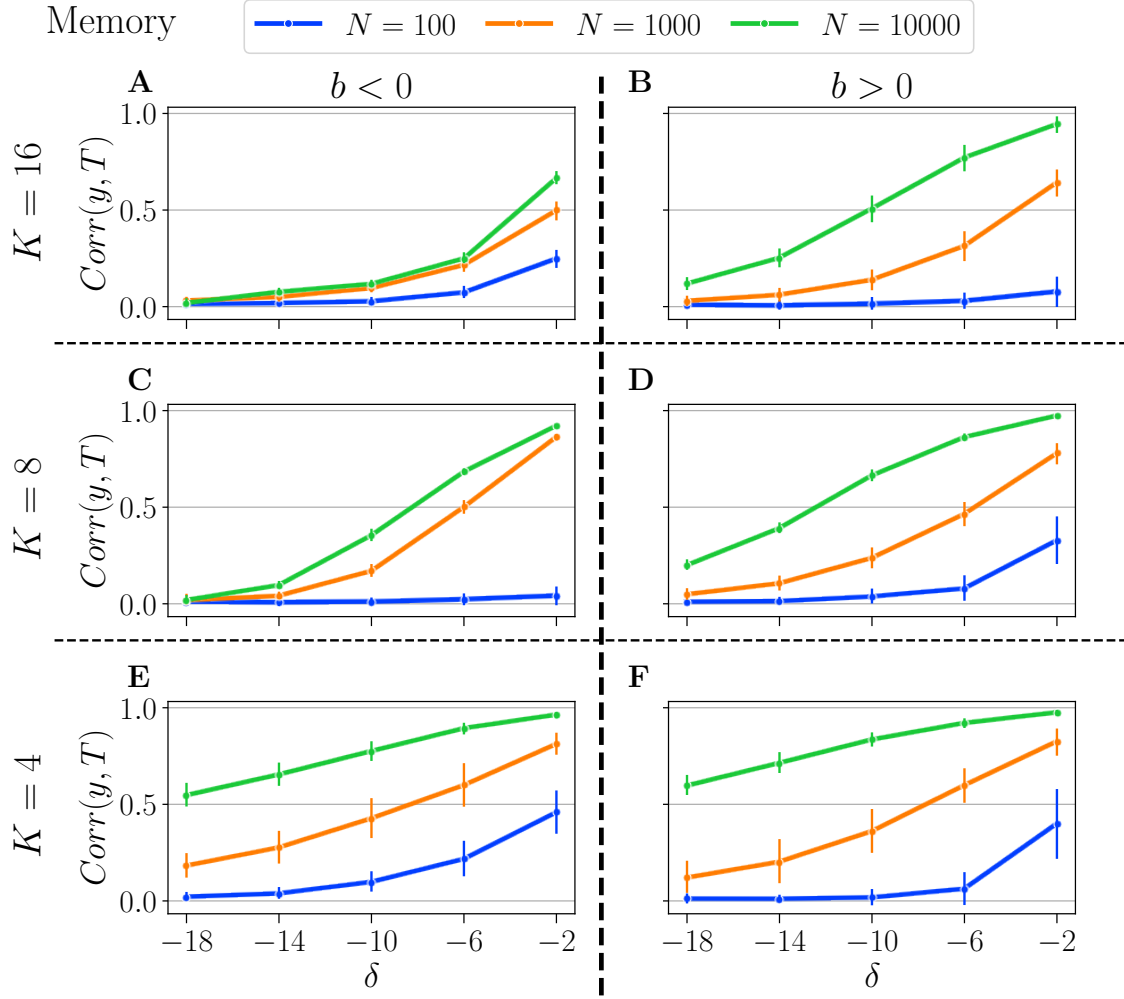


Figure 9: Summary of performance in the memory tasks, for various connectivity degrees  $K$ , and size of the reservoirs  $N$ : for  $N = 10000$  (green curves),  $N = 1000$  (orange curves), and  $N = 100$  (blue curves).  $K = 16$  (A and B),  $K = 8$  (C and D),  $K = 4$  (E and F).  $b < 0$  (left column), and  $b > 0$  (right column). Solid lines represent the average over all reservoirs generated with the same reservoir ( $N$ ,  $K$ ,  $b_{opt}$ ), and the error bar represents one standard deviation. As explained in Sec. 3.1.2,  $b_{opt}$  is obtained by selecting the balance that gives the best average performance at the most difficult setting in each respective task.

In the memory task, as expected, reducing the number of neurons diminishes the reservoir's memory capacity, and the more difficult the task, the lower the performance. In addition, decreasing the reservoir size generally increases the reservoir-to-reservoir variance, as indicated by the larger error bars, even though this is not always the case, especially when performance is already low.

The number of neurons exerts a greater influence when  $K$  is lower. Indeed, for  $K = 4$ , we observe a significant disparity between all three  $N$  values across all difficulty levels ( $\delta$ ). Surprisingly, for higher  $K$ , and especially when  $b < 0$ , performances for  $N = 1000$  and  $N = 10000$  are relatively comparable, and increasing the reservoir size is not improving performance, especially for tasks requiring longer memory. As previously noted, when  $K = 4$ , performance is similar regardless of whether  $b$  is positive or negative, a finding that is now corroborated across all tested  $N$  values.



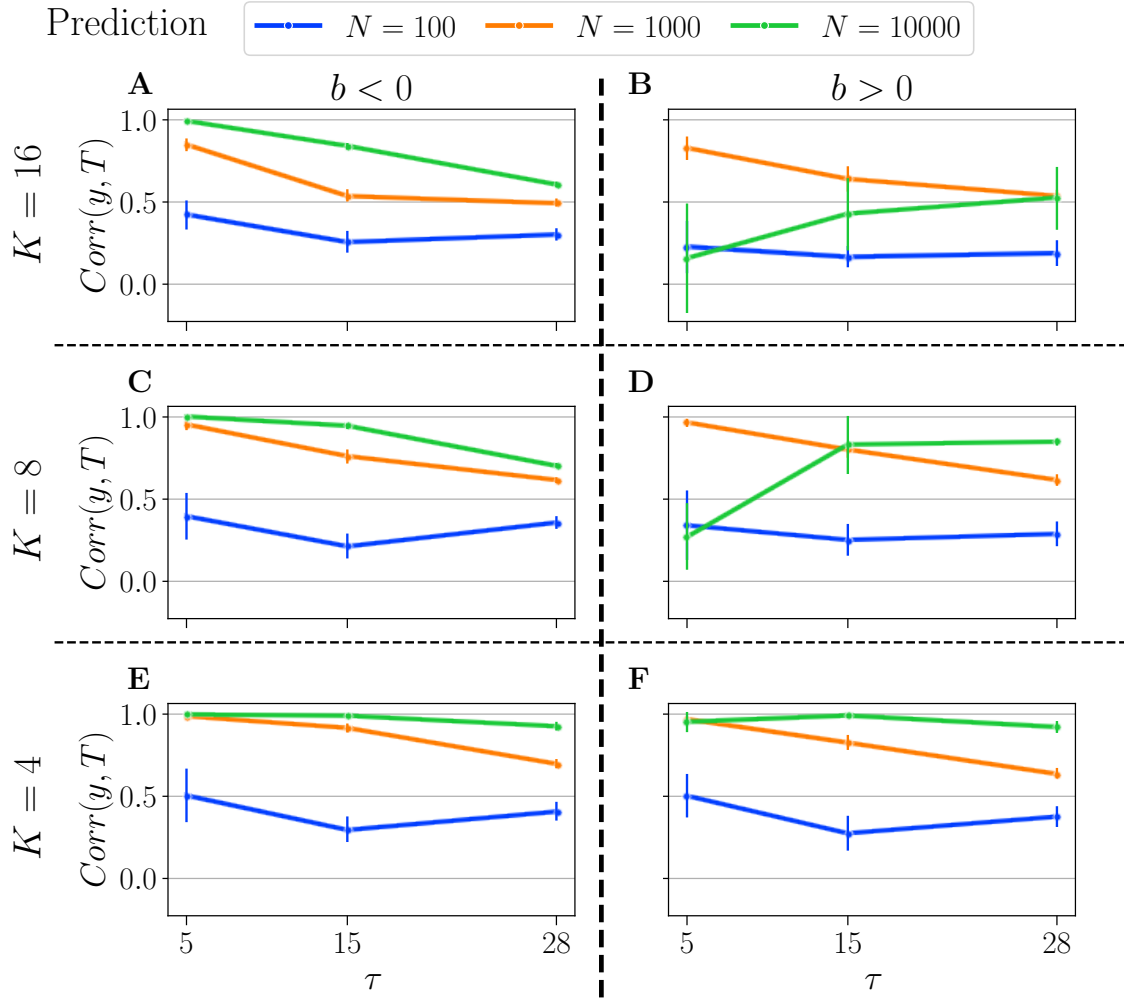


Figure 10: Summary of performance in the prediction tasks, for various connectivity degrees  $K$ , and size of the reservoirs  $N$ : for  $N = 10000$  (green curves),  $N = 1000$  (orange curves), and  $N = 100$  (blue curves).  $K = 16$  (A and B),  $K = 8$  (C and D),  $K = 4$  (E and F).  $b < 0$  (left column), and  $b > 0$  (right column). For more information on the plots, see the caption of Fig. 9.

It appears that  $K$  has minimal influence when  $N = 100$ , as reservoirs perform similarly regardless of  $K$ . The same holds true for both positive and negative  $b$ , which their identical dynamics profiles might explain. This suggests that, for low neuron counts, the system's dynamics and performance are more strongly influenced by the balance parameter  $b$  than by the number of connections  $K$ .

In the prediction task, we observe some surprising trends. Notably, having a higher  $N$  is not always advantageous, as the optimal  $N$  appears to depend on both the task and the control parameter.

Firstly, for  $b < 0$ , the performance profile is similar to that in the memory task: higher  $N$  yields better performance, and performance decreases with increasing task difficulty ( $\tau$ ). However, the performances of  $N = 10000$  and  $N = 1000$  are closer to each other and significantly higher than that of  $N = 100$ , which again remains unaffected by  $K$ .

Secondly, for  $b > 0$ , the value of  $K$  strongly influences the relationship between performance and reservoir size. With  $K = 4$ , the performance profile is similar to

that for  $b < 0$ : performance decreases monotonically with  $\tau$  and  $N$ . However, for  $K = 8$  and especially for  $K = 16$ , we observe some unexpected results. Smaller reservoirs ( $N = 1000$ ) can outperform larger ones ( $N = 10000$ ) in some tasks. This phenomenon is even more pronounced for higher  $K$ , as the orange line (representing  $N = 1000$ ) consistently outperforms the green line (representing  $N = 10000$ ) across all tested tasks.

## 4 Discussion

Our study reveals that the edge of chaos, or the critical region, does not consistently align with the peak performance region (Gallicchio 2020), and this alignment is contingent upon the sign of excitatory-inhibitory balance  $b$ . For  $b > 0$ , as previously observed (Calvet et al. 2023), the critical region coincides with the highest performance. However, for  $b < 0$ , the region of optimal performance does not coincide with the critical region when the connectivity degree  $K$  is optimally selected. Instead, supplanting the disordered phase, a re-entrance of the critical region is observed, indicated by an increased attractor diversity, which surprisingly aligns with the best-performing region. This insight suggests that the attractor dynamics can be utilized to identify the region of interest for the design or reservoirs, and this also holds for  $b > 0$  and its identified critical region (Calvet et al. 2023).

In terms of the interplay between  $b$  and the connectivity degree  $K$ , our research shows that a carefully selected  $K$  ( $K = 4$ ) renders the sign of  $b$  irrelevant, as the optimal  $b$  becomes  $\pm\epsilon$  with  $\epsilon$  very small. This suggests that the optimal balance is near, but not at, perfect symmetry, even though  $b \rightarrow 0$  results in zero performance. In statistical physics, it is well known that symmetry breaking induces critical phase transitions (Goldenfeld 2018), and our findings suggest that symmetry breaking in the balance of excitatory-inhibitory synapses is crucial for achieving optimal performance. Refining initial literature (Bertschinger & Natschlager 2004; Snyder et al. 2013; Burkow & Tufte 2016; Echlin et al. 2018), the highest-performing region is characterized by a preponderance of irregular attractors within the disordered region.

To understand this, one can consider what happens when  $\sigma^*$  tends to infinity. This can be achieved in two ways: first, when the standard deviation of the weight  $\sigma$  is fixed while the mean weights  $\mu \rightarrow 0$ , and second, when  $\mu$  is fixed while  $\sigma \rightarrow \infty$ . The first case has been covered in other works (Bertschinger & Natschlager 2004; Busing et al. 2010) and shows the importance of tuning the scaling of the input weights with the recurrent weight statistics (Burkow & Tufte 2016). In the present work, however, the second option is considered, as the mean weights is fixed, and  $\sigma$  increases to higher values. As such,  $b$  approaches zero, which results in a symmetry between excitation and inhibition but with increasingly higher synaptic weights (in absolute value). Consequently, each neuron receives equal excitatory and inhibitory recurrent inputs, and since the input weights are kept constant, the external input becomes insignificant. Finally, since neurons have a zero threshold, they have a 50% probability of spiking, leading to a random spike train. Therefore, it is not surprising to observe a performance dip as  $\sigma \rightarrow \infty$  ( $b \rightarrow 0$ ) since the reservoir activity becomes independent of the input. However, what requires further investigation is the unexpected drastic per-

formance increase when this symmetry is slightly broken as  $b = \pm\epsilon \sim 0.03$  (roughly corresponding to a 6% difference between excitatory and inhibitory synapses).

These findings highlight the critical role of  $K$  in determining other control parameters. Firstly, the optimal number of connections ( $K = 4$ ) eliminates the performance asymmetry, significantly simplifying the parameter  $b$  selection. Secondly, consistent with previous studies,  $N$  generally enhances performance, but this is only true for optimal  $K = 4$  values, particularly in the prediction task, where smaller reservoirs occasionally outperform larger ones. Additionally, the performance gain obtained by  $K$  is significant only when the reservoir size is sufficiently large. For instance, with reservoirs of size  $N = 100$ ,  $K$  had close to no effect on the best performance. However, optimally choosing  $K$  becomes key to obtaining a gain in performance when increasing the size.

Our work reveals a complex interplay between the topology and weights parameters, but assuming a reservoir of sufficient size ( $N \geq 1000$ ),  $K$  acts as a pivotal control parameter by greatly simplifying the way parameters interact with each other. When  $K$  is optimal, then  $N$  must be maximized, and  $b$  can be chosen very close to zero but finite, and of any sign.

## 5 Future work

Understanding the relationship between dynamics and performance is crucial for simplifying reservoir design (Bertschinger & Natschläger 2004; Krauss et al. 2019b,a; Metzner & Krauss 2022; Calvet et al. 2023). Our study advances this understanding but reveals a more intricate relationship than anticipated. Specifically, performance was found to be highly sensitive to symmetry breaking in the excitation-inhibition balance, while the metrics used to probe the dynamics were completely unaware of the symmetry. Future research could investigate the relationship between the optimal balance and other dynamic-probing metrics, including spatial and temporal correlation (Metzner & Krauss 2022), and possibly topology (Kinoshita et al. 2009; Masulli & Villa 2015). For instance, it could be hypothesized that the longest neural pathways in the random graph become available for information transmission only at  $b = \pm\epsilon$ , which could explain why optimal performance necessitates a breaking of symmetry in the balance.

## References

- Bertschinger, Nils & Thomas Natschläger. 2004. Real-Time Computation at the Edge of Chaos in Recurrent Neural Networks. *Neural Computation* 16(7). 1413–1436. doi:10.1162/089976604323057443. <https://direct.mit.edu/neco/article/16/7/1413-1436/6890>.
- Bianchi, Filippo Maria, Lorenzo Livi & Cesare Alippi. 2016. Investigating echo state networks dynamics by means of recurrence analysis. *IEEE transactions on neural networks and learning systems* 29(2). 427–439. doi:10.1109/TNNLS.2016.2630802. <http://www.ncbi.nlm.nih.gov/pubmed/28114039><http://arxiv.org/abs/1601.07381><http://dx.doi.org/10.1109/TNNLS.2016.2630802>.

- Burkow, Aleksander Vognild & Gunnar Tufte. 2016. *Exploring Physical Reservoir Computing using Random Boolean Networks*: Norwegian University of Science and Technology - NTNU dissertation. <http://hdl.handle.net/11250/2417596>.
- Büsing, Lars, Benjamin Schrauwen & Robert Legenstein. 2010. Connectivity, dynamics, and memory in reservoir computing with binary and analog neurons. *Neural Computation* 22(5). 1272–1311. doi:10.1162/neco.2009.01-09-947.
- Calvet, Emmanuel, Jean Rouat & Bertrand Reulet. 2023. Excitatory/inhibitory balance emerges as a key factor for RBN performance, overriding attractor dynamics. *Frontiers in Computational Neuroscience* 17. doi:10.3389/fncom.2023.1223258. <https://www.frontiersin.org/articles/10.3389/fncom.2023.1223258/full>.
- Cherupally, Sai Kiran. 2018. Hierarchical Random Boolean Network Reservoirs. Tech. rep. Portland State University Portland, OR. doi:10.15760/etd.6238. <https://archives.pdx.edu/ds/psu/25510>.
- Cramer, Benjamin, David Stöckel, Markus Kreft, Michael Wibral, Johannes Schemmel, Karlheinz Meier & Viola Priesemann. 2020. Control of criticality and computation in spiking neuromorphic networks with plasticity. *Nature Communications* 11(1). doi:10.1038/s41467-020-16548-3. <https://doi.org/10.1038/s41467-020-16548-3>.
- Croll, Grenville J. 2014. BiEntropy – the Measurement and Algebras of Order and Disorder in Finite Binary Strings. In UK) John C Amson (University of St Andrews & USA) Louis H Kauffman (University of Illinois at Chicago (eds.), *Scientific essays in honor of h pierre noyes on the occasion of his 90th birthday*, 48–64. Word Scientific series on edn. doi:10.1142/9789814579377{\\_}0004. [http://www.worldscientific.com/doi/abs/10.1142/9789814579377\\\_0004](http://www.worldscientific.com/doi/abs/10.1142/9789814579377\_0004).
- Echlin, Moriah, Boris Aguilar, Max Notarangelo, David Gibbs & Ilya Shmulevich. 2018. Flexibility of Boolean Network Reservoir Computers in Approximating Arbitrary Recursive and Non-Recursive Binary Filters. *Entropy* 20(12). 954. doi:10.3390/e20120954. <http://www.mdpi.com/1099-4300/20/12/954>.
- Gallicchio, Claudio. 2020. Sparsity in Reservoir Computing Neural Networks. In *2020 international conference on innovations in intelligent systems and applications (inista)*, 1–7. IEEE. doi:10.1109/INISTA49547.2020.9194611. <https://ieeexplore.ieee.org/document/9194611/>.
- Glass, L. & C. Hill. 1998. Ordered and disordered dynamics in random networks. *Europhysics Letters (EPL)* 41(6). 599–604. doi:10.1209/epl/i1998-00199-0. <https://iopscience.iop.org/article/10.1209/epl/i1998-00199-0>.
- Goldenfeld, Nigel. 2018. *Lectures on Phase Transitions and the Renormalization Group*. CRC Press. doi:10.1201/9780429493492. <https://www.taylorfrancis.com/books/9780429962042>.
- Hajnal, Márton Albert & András Lörincz. 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. doi:10.1007/11840817{\\_}69.
- Jaeger, Herbert. 2005. A tutorial on training recurrent neural networks , covering BPPT , RTRL , EKF and the ” echo state network ” approach. *ReVision* 2002. 1–46. <http://www.mendeley.com/catalog/>

- tutorial-training-recurrent-neural-networks-covering-bppt-rtrl-ekf-echo-state-ne
- Kingma, Diederik P. & Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *3rd international conference for learning representations, san diego*, ICLR. <http://arxiv.org/abs/1412.6980>.
- Kinoshita, Shu-ichi, Kazumoto Iguchi & Hiroaki S. Yamada. 2009. Intrinsic properties of Boolean dynamics in complex networks. *Journal of Theoretical Biology* 256(3). 351–369. doi:10.1016/j.jtbi.2008.10.014. <https://linkinghub.elsevier.com/retrieve/pii/S0022519308005456>.
- Krauss, Patrick, Marc Schuster, Verena Dietrich, Achim Schilling, Holger Schulze & Claus Metzner. 2019a. Weight statistics controls dynamics in recurrent neural networks. *PLOS ONE* 14(4). e0214541. doi:10.1371/journal.pone.0214541. <https://dx.plos.org/10.1371/journal.pone.0214541>.
- Krauss, Patrick, Alexandra Zankl, Achim Schilling, Holger Schulze & Claus Metzner. 2019b. Analysis of Structure and Dynamics in Three-Neuron Motifs. *Frontiers in Computational Neuroscience* 13(February). 1–6. doi:10.3389/fncom.2019.00005. <https://www.frontiersin.org/article/10.3389/fncom.2019.00005/full>.
- Legenstein, Robert & Wolfgang Maass. 2006. What Makes a Dynamical System Computationally Powerful? In Terrence J. Sejnowski John McWhirter Simon Haykin, Jose C. Principe (ed.), *New directions in statistical signal processing*, 127–154. The MIT Press. doi:10.7551/mitpress/4977.003.0008. <https://direct.mit.edu/books/book/3805/chapter/124942/6-what-makes-a-dynamical-system-computationally>.
- Luque, Bartolo & Ricard V. Solé. 2000. Lyapunov exponents in random Boolean networks. *Physica A: Statistical Mechanics and its Applications* 284(1-4). 33–45. doi:10.1016/S0378-4371(00)00184-9. <https://linkinghub.elsevier.com/retrieve/pii/S0378437100001849>.
- Maass, Wolfgang, Thomas Natschläger, Henry Markram, Maass, Natschläger & Markram. 2002. Real-time Computing Without Stable States: A New Framework for Neural Computation Based on Perturbations. *NEURCOMP: Neural Computation* 14(11). 2531–2560. doi:10.1162/089976602760407955. <http://www.ncbi.nlm.nih.gov/pubmed/12433288>.
- Masulli, Paolo & Alessandro E. P. Villa. 2015. The topology of the directed clique complex as a network invariant. *SpringerPlus* 5(1). 388. doi:10.1186/s40064-016-2022-y. <http://arxiv.org/abs/1510.00660><http://dx.doi.org/10.1186/s40064-016-2022-y><http://springerplus.springeropen.com/articles/10.1186/s40064-016-2022-y>.
- Metzner, Claus & Patrick Krauss. 2022. Dynamics and Information Import in Recurrent Neural Networks. *Frontiers in Computational Neuroscience* 16(April). 1–15. doi:10.3389/fncom.2022.876315. <https://www.frontiersin.org/articles/10.3389/fncom.2022.876315/full>.
- Natschläger, Thomas, Nils Bertschinger & Robert Legenstein. 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*, NeurIPS.
- Nowshin, Fahiha, Yuhao Zhang, Lingjia Liu & Yang Yi. 2020. Recent Advances in Reservoir Computing With A Focus on Electronic Reservoirs. In *2020 11th international green and sustainable computing workshops (igsc)*, 1–8. IEEE. doi:10.

- 1109/IGSC51522.2020.9290858. <https://ieeexplore.ieee.org/document/9290858/>.
- Snyder, David, Alireza Goudarzi & Christof Teuscher. 2012. Finding Optimal Random Boolean Networks for Reservoir Computing. In *Artificial life 13*, 259–266. MIT Press. doi:10.7551/978-0-262-31050-5-ch035. <https://www.mitpressjournals.org/doi/abs/10.1162/978-0-262-31050-5-ch035>.
- Snyder, David, Alireza Goudarzi & Christof Teuscher. 2013. Computational capabilities of random automata networks for reservoir computing. *Physical Review E* 87(4). 042808. doi:10.1103/PhysRevE.87.042808. <https://link.aps.org/doi/10.1103/PhysRevE.87.042808>.
- Steiner, Peter, Azarakhsh Jalalvand & Peter Birkholz. 2023. Cluster-Based Input Weight Initialization for Echo State Networks. *IEEE Transactions on Neural Networks and Learning Systems* 34(10). 7648–7659. doi:10.1109/TNNLS.2022.3145565. <https://ieeexplore.ieee.org/document/9705114/>.
- Triefenbach, Fabian, Azarakhsh Jalalvand, Benjamin Schrauwen & Jean Pierre Martens. 2010. Phoneme recognition with large hierarchical reservoirs. In *Advances in neural information processing systems 23: 24th annual conference on neural information processing systems 2010, nips 2010*, NIPS.
- Zhu, Ruomin, Joel Hochstetter, Alon Loeffler, Adrian Diaz-Alvarez, Tomonobu Nakayama, Joseph T. Lizier & Zdenka Kuncic. 2021. Information dynamics in neuromorphic nanowire networks. *Scientific Reports* 11(1). 1–15. doi:10.1038/s41598-021-92170-7. <https://doi.org/10.1038/s41598-021-92170-7>.

label0

## 6 Supplementary Material

### 6.1 Training

Both tasks follow the same protocol for training the readout weights:

- Reservoirs receive the input  $u$  for  $D = 2000$  time steps.
- We discard the first 500 time steps.
- The training is then performed on the subsequent 1500 time steps. We concatenate the reservoir outputs in time, and use the optimization procedure defined in 2.1.

Each experiment consists in 40 values of  $\sigma^*$  per sign of  $b$ , and is performed for three values of  $K = \{4, 8, 16\}$  and three values of  $N = \{100, 1000, 10000\}$ . For each value  $(\sigma^*, K, N)$ , 20 reservoirs are randomly generated, and each network is run 5 times with different randomly tossed inputs (i.e., 14, 400 simulations). Each training is performed for 4000 epochs (with a total of 57, 600, 000 training epochs).

### 6.2 The density is not a control parameter

Historically, the literature on RBN studied the connectivity degree  $K$ , along with the number of neurons in the reservoir  $N$ . It has been shown that RBN reservoirs possess interesting computational properties for very sparse matrices, with at most  $K = 25$  connections per neuron, according to (Büsing et al. 2010). On the other hand, studies in ESN have explored the impact of another parameter, the density  $d$  (Hajnal & Lörincz 2006; Krauss et al. 2019a; Metzner & Krauss 2022), which captures how many zeros there are in the adjacency matrix, ranging from zero for no connections, to one for fully connected reservoirs.

In the context of RBN reservoirs, the density is equal to  $d = K/N$ . It is worth noting that for the values of  $K$  and  $N$  used in this study, the corresponding values of  $d$  are ridiculously small, ranging from 0, 0004 to 0.16. This stresses how much the RBN behaves differently since phase transitions occur at a fraction of what is observed in an ESN. It also suggests that  $K$  is a more natural choice for the control parameter of the RBN.

Nonetheless, this is not sufficient to rule out  $d$  as a good candidate for control parameters. Recalling that it is in the same fashion that  $\mu$  and  $\sigma$  were shown to be dependent parameters and  $\sigma^* = \mu/\sigma$  was built in (Calvet et al. 2023). If  $d$  is a control parameter, however, then all combinations of  $K$  and  $N$  leading to the same values of  $d$  should provide the exact same dynamics and performance in tasks, as is the case with  $\sigma^*$ . To find if  $d$  is indeed a control parameter, we study the dynamics and performance of three couples of  $K$  and  $N$  giving the same value  $d$ :

$$d = K/N = 16/10000 = 8/5000 = 4/2500 = 0.0016$$

In Fig. S1, we show the attractor statistics in the three conditions, as a function of  $b$ . The first thing we notice is that all three tested conditions give very different attractor



profiles. The width and the location of the critical regions are not the same, and the proportion of attractors also greatly varies, with lower  $K$  and  $N$  giving much more varied dynamics and wider critical regions. It is intriguing that in Fig. 4 and Fig. 8, diminishing  $K$  and  $N$  results in a somewhat comparable increase in the dynamics variability as well, but if  $K$  and  $N$  seem related, it is not by their ratio as proposed in the density parameter.

In addition, we show in Fig. S2 the performance in memory and prediction of 20 reservoirs selected with the same  $b_{opt}$  values, for each couple  $(N, K)$  (for more information on how these plots are obtained see Sec. 3.1.2). Following the dynamics analysis, even though each couple  $(N, K)$  possesses the same density, the performance varies.

When considering  $b < 0$ , smaller  $K$  consistently provides higher performance in all considered tasks. This is interesting because it is in spite of the fact that reservoirs are smaller. On the other hand, for  $b > 0$ , we have a different behaviour, first in the memory task, performances are close even though  $(K = 4, N = 2500)$  performs slightly better in the most difficult ( $\tau = -14$ ) set-up. Second, in the prediction task,  $(K = 4, N = 2500)$  and  $(K = 8, N = 5000)$  give very similar performance, while the latter have reservoirs of twice the size. Surprisingly, even though  $(K = 16, N = 10000)$  has the biggest network, its performance is way below the two others.

As a consequence, for the same given density value obtained with various combinations of  $K$  and  $N$ , we obtain distinct dynamics statistics and performance. We conclude that the density is not a control parameter for RBN reservoirs, and one must study the effect of  $K$  and  $N$  separately, as two independent control parameters.

Regarding the objective of the paper, these results suggest that  $K$  has a stronger influence on the performance than  $N$ , as increasing the size of the network does not necessarily guarantee higher performance, but reducing  $K$  consistently improves them.



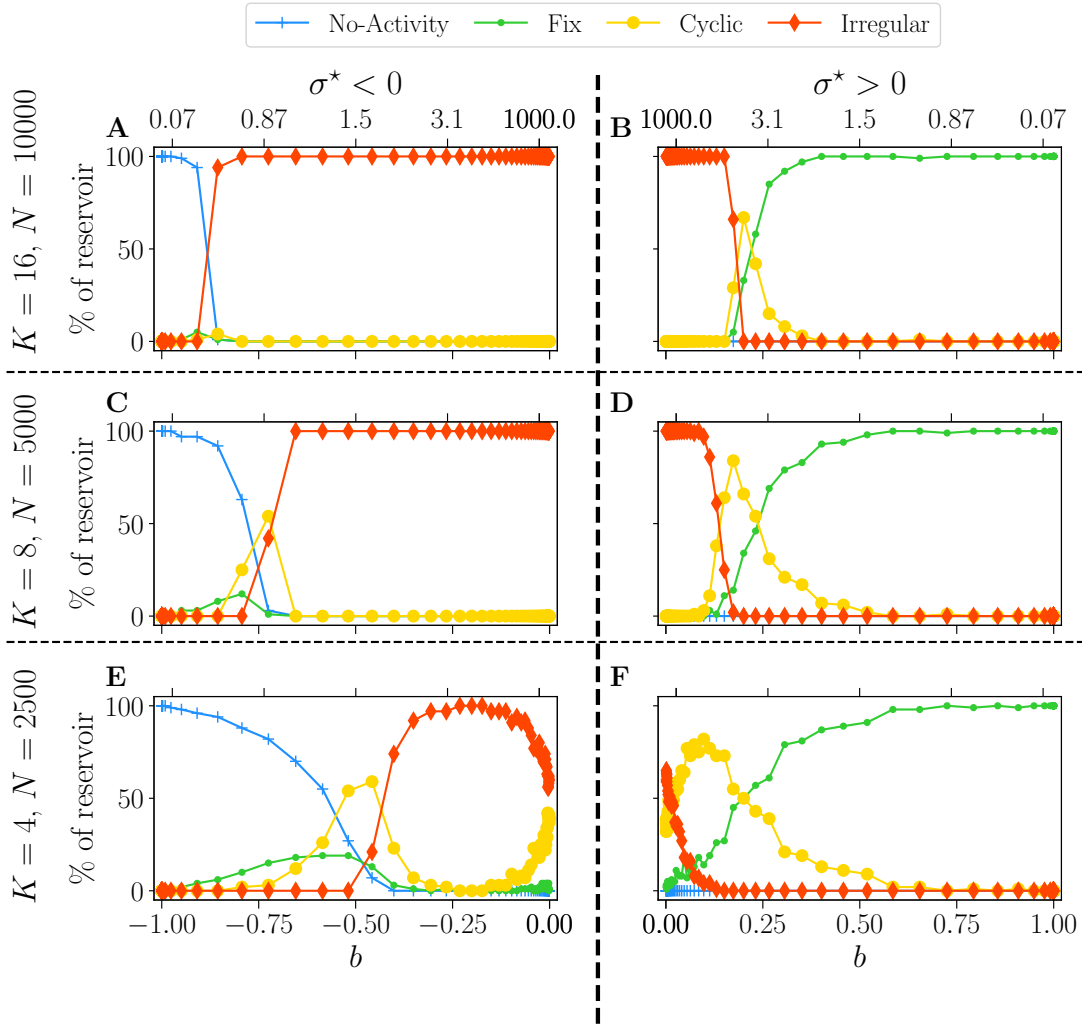


Figure S1: Attractor statistics of free-evolving RBN reservoirs, controlled by  $K$  and  $N$  with a constant ratio (rows), versus the balance  $b$  (x-axis). The lower x-axis represents the corresponding  $|\sigma^*|$ , for  $b < 0$  (A, C, E), and  $b > 0$  (B, D, F). Each steady activity signal is classified into one of the four categories of attractors: no-activity (+), fix (•), cyclic (•), irregular (♦). The statistics of attractors are computed over 100 reservoirs run once (y-axis). Results are shown for  $K = 16$ ,  $N = 10000$  (A and B),  $K = 8$ ,  $N = 5000$  (C and D), and  $K = 4$ ,  $N = 2500$  (E and F).

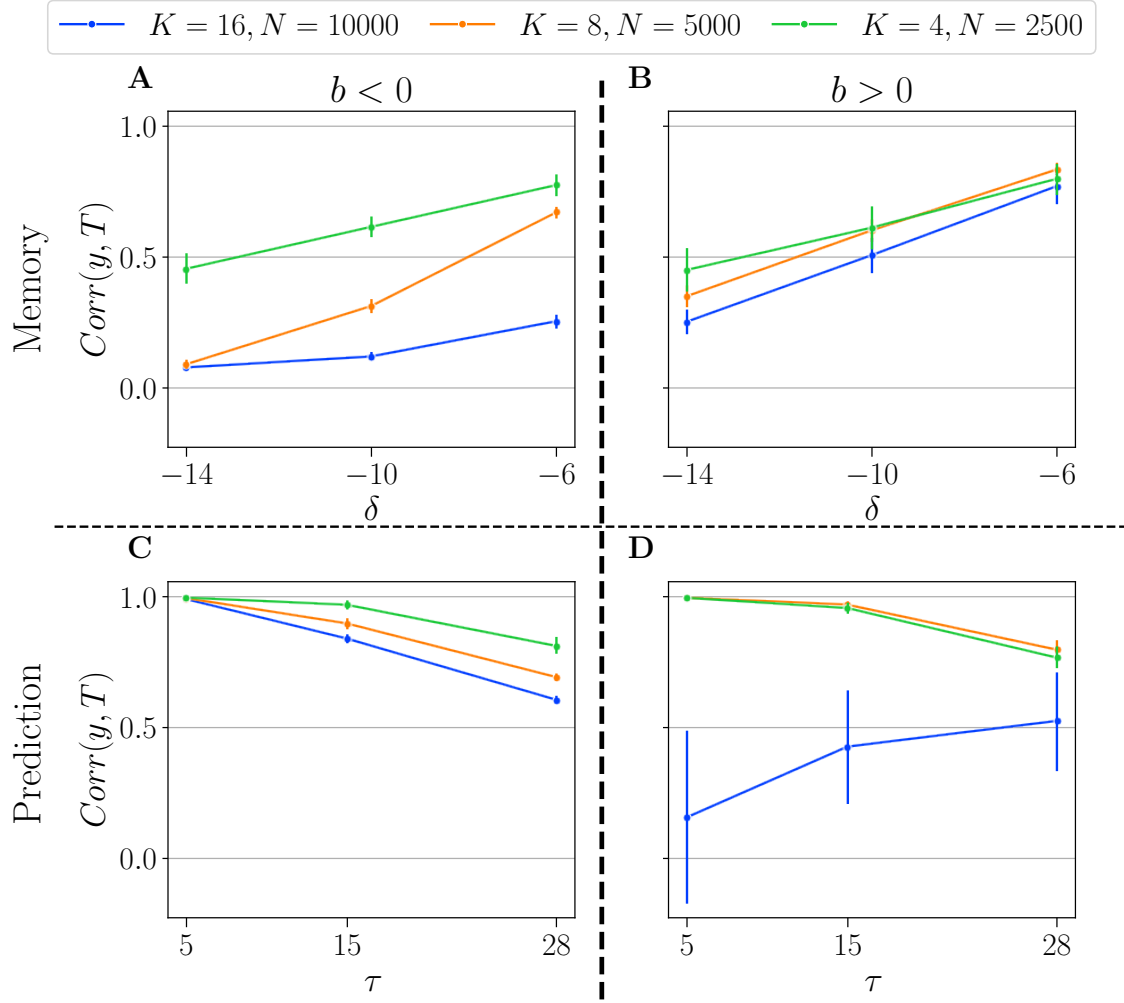


Figure S2: Summary of performance in the memory (A and B) and the prediction (C and D) tasks, for three couples of  $K$  and  $N$  resulting in the same density:  $K = 4, N = 2500$ ;  $K = 8, N = 5000$ ; and  $K = 16, N = 10000$ . Results are shown for both  $b < 0$  (left panel), and  $b > 0$  (right panel). For each value of  $(K, N)$ , we selected the  $b_{opt}$  giving the highest average performance, in the most difficult task ( $\delta = -14$  for memory, and  $\tau = 28$  for prediction). We plot the performance (higher is better) of reservoirs  $Corr(y, T)$  (y-axis), plotted as a function of their respective task parameters (x-axis):  $\delta$  (memory) and  $\tau$  (prediction). The solid line represents the average over 20 reservoirs (generated with the same  $b_{opt}$  and  $(K, N)$  values). The error bars represent one standard deviation.

### 6.3 Optimal balances for performance

$K$	1	2	3	4	5	8	16
$b > 0$	0.024	0.024	0.052	0.029	0.077	0.149	0.179
$b < 0$	-0.099	-0.069	-0.057	-0.029	-0.049	-0.52	-0.99

Table S1: Table of the balance  $b_{opt}$  giving the best average performance at white-noise memory for  $\delta = -18$  ( $N = 10000$ ). The green highlight represents the optimal  $K$  value. Lower values of  $K$  typically give the lowest  $b_{opt}$  values, while increasing  $K$  tends to increase  $b_{opt}$  in absolute value.

$K$	1	2	3	4	5	8	16
$b > 0$	0.95	0.19	0.068	0.024	0.098	0.15	0.19
$b < 0$	-0.24	-0.20	-0.12	-0.047	-0.42	-0.82	-0.95

Table S2: Table of the balance  $b_{opt}$  giving the best average performance at Mackey-Glass prediction for  $\tau = 28$  ( $N = 10000$ ). The green highlight represents the optimal  $K$  value, which typically gives the lowest  $b_{opt}$  values, while increasing  $K$  tends to increase  $b_{opt}$  in absolute value.