Universitat
de les Illes Balears

MASTER THESIS

The role of coherence in Quantum Reservoir Computing

Ana Palacios de Luis

Master's Degree in Physics of Complex Systems

Centre for Postgraduate Studies

Academic Year 2019/2020

The role of coherence in Quantum Reservoir Computing

Ana Palacios de Luis

Master Thesis
Centre for Postgraduate Studies
University of the Balearic Islands

Academic Year 2019/2020

Keywords:
Quantum Machine Learning, unconventional computing, Quantum Reservoir Computing, quantum coherence, Information Processing Capacity, spin network, robustness to dissipation

Thesis supervised by Roberta Zambrini and Gian Luca Giorgi

Abstract

The study of quantum systems from the perspective of Reservoir Computing is a promising angle that has been increasingly raising interest in the past few years, since our current technology allowed for the experimental realisation of such systems. This work explores the relation between the coherence present in the quantum reservoir and its ability to process information in the context of the resource theory of coherence. For this we take the transverse-field Ising model as our reservoir, which we study in its different dynamical regimes as a function of the intensity of the magnetic field and the disorder present within. We are able to link the operational regimes of the reservoir with higher values of the coherence, and relate the dynamics of each phase to its processing capacity. In addition, we perform an analysis of the robustness of these reservoirs to phase and bit flip noise, which shows that the latter is more destructive than the former for the system's ability to process information. Our results also establish a link between quantum correlations and high-order degrees of nonlinear processing capacity.

Acknowledgements

In the first place, I would like to thank Roberta and Gian Luca for their guidance, under which I am happy to say I have learnt a lot about the job and proceedings of a researcher. A particular thanks goes to Rodrigo Martínez for all his help and support with the numerical part of the project and valuable insights about the system, without which frustration would have most certainly taken over on multiple occasions. For the emotional support (and some useful programming tips) I am specially grateful to my quarantine companions, Javi, Jorge and Medi, as well as to Laura, Alex and the rest of my friends and peers. Finally, I would like to thank my family and friends overall for bearing with my ubiquitous spin networks for the past couple of months.

Contents

Acronyms

ANN	Artificial Neural Network
ESN	Echo State Network
ETH	Eigenstate Thermalisation Hypothesis
FFNN	Feed-Forward Neural Network
IPC	Information Processing Capacity
LSM	Liquid State Machine
MBL	Many-Body Localisation
ML	Machine Learning
NMR	Nuclear Magnetic Resonance
RC	Reservoir Computing
RNN	Recurrent Neural Network
SG	Spin Glass

Chapter 1

Introduction

In the past couple of decades there has been a remarkable burst of new technologies brought about by the advent of a Machine Learning (ML) approach to a great variety of fields, both inside physics and outside of it [1]. Versatile tasks like pattern recognition, image and speech processing and temporal series prediction are among the most cost-effective applications. The burgeoning enthusiasm about the possibilities that may be unlocked with ML techniques has turned them into an extremely active field of research.

As a part of this scenario, Reservoir Computing (RC) [2, 3, 4] is increasingly gaining attention because of its potential for fast, real-time data processing. A RC processor which is embedded in a larger machine, say by being a part of it already, facilitates the implementation problem and avoids the possible losses that may take place during the transport of information away from the main device. This fact, together with the profits of the rich dynamics present in quantum systems, point to Quantum Reservoir Computing (QRC) [5, 6, 7] as a topic of great interest. In this context, the transverse-field Ising model has emerged as a candidate system for QRC with discrete variables, with its experimental realisation as a reservoir being already accessible by the current technology. In the work presented here, we will extend the study of this class of systems, both in the more canonical case of a uniform magnetic field and in the case where it exhibits local fluctuations. We will explore these systems more thoroughly from the QRC perspective and characterise some of their purely quantum aspects, so as to lay the groundwork for the identification of possible quantum advantages. For this, we will characterise how the system's ability to predict nonlinear time series is reflected through coherence, a strictly quantum property of the spin network.

This Master thesis aims to be rather self-contained, motivating the investigation in RC and the advantages this model has in the quantum realm first, as well as illustrating its basic inner workings, before moving on to the analysis of the particular models under study. Sections 1.1 and 1.2 set the current scenario in the ML and RC scenes and their translation into a quantum setup. The methodology we follow in order to assess the system's processing capacity is also introduced in this section. Section 1.3 reviews some ideas about quantum coherence and how we will quantify it throughout the study. In order to conclude our theoretical basis, we make some further remarks on the theory or quantum thermalisation in Section 1.4 and move on to the detailed description of the simulation procedure in Section 1.5. At this point we are ready to begin with the numerical simulations for the characterisation of the system's dynamics in Chapter 2. The study of the relation between the system's coherence and its processing capacity is presented in Section 3.1, as well as the assessment of its robustness in the presence

of a noisy environment in 3.2. Finally, we highlight the conclusions and outlook of this work in Chapter 4 and include some complementary notes in the Appendices.

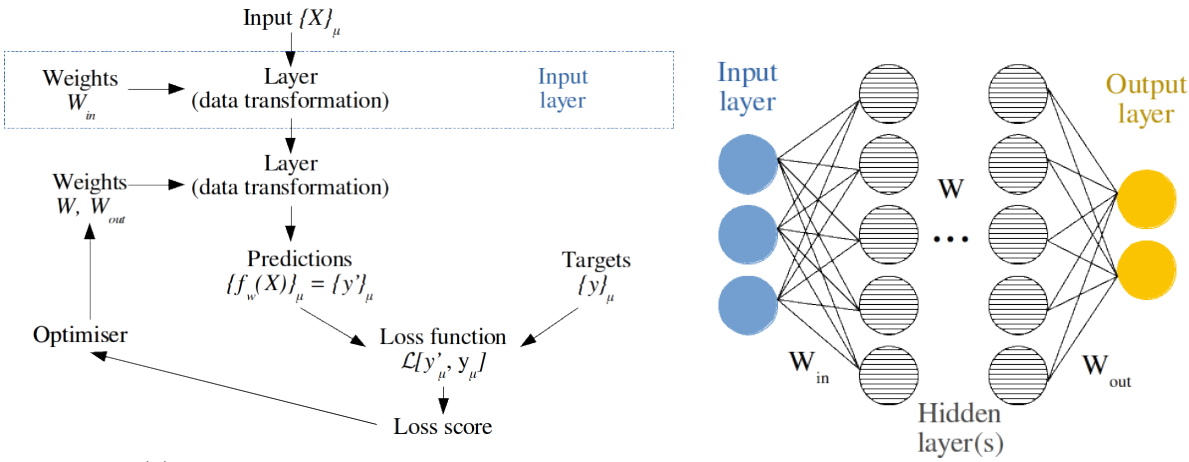
1.1 Relevant Machine Learning concepts

ML technologies are most commonly materialised in Artificial Neural Networks (ANNs), a collection of structures and architectures (either hardware- or software-based) designed to effectively mimic the processing of information that takes place in our brains. These processing tasks may be of two types: static (e.g., classification) or dynamic (e.g., prediction), corresponding to non-temporal and temporal tasks respectively. The first type is addressed with the so-called Feed-Forward Neural Networks (FFNNs), while the second is tackled by Recurrent Neural Networks (RNNs) [1] . The latter manage to retain information about previous entries from the input sequence by introducing closed loops in their inner connections, which allows this past information to influence the current output. In the following, our focus will be centred on RNNs, since these are the ones the concept of Reservoir Computing (RC) originally derives from.

As in every ML instance, the construction of an ANN undergoes two major steps: training and testing. In supervised learning, the framework this work is set in, the network is adjusted during a training phase from a set of training examples in which we provide the system with the desired result. In more strict mathematical terms, we are provided with a training set consisting of n input samples, where each sample may be of dimension p ($X_\mu \in \mathbb{R}^p$, with $\mu = 1, ..., n$) and a series of n outputs, each of dimension d ($y_\mu \in \mathbb{R}^d$), and we are looking for a function f that approximates $f(\mathbf{X})$ to $\mathbf{y}$. For this process to be considered successful, however, the approximation must be good for any other data set the system has not been trained with; only then will we know that the system has properly extracted the universal rule from its particular realisation in the training set. Thus, a testing phase for ensuring that this phenomenon has taken place is indispensable. The most common procedure for finding this function f is to express it in terms of a set of weights $w \in \mathbb{R}^k$, such that the current rule of the network may be expressed as f_w. This f_w is optimised on the basis of a certain loss function $\mathcal{L}[f_w(X_\mu), y_\mu]$ that scores lower the closer the prediction $y' = f_w(X_\mu)$ and the target y are. This minimisation procedure is usually carried out through gradient descent methods [1]. A schematic illustration of this process is presented in Fig. 1.1a.

Fig. 1.1b, on the other hand, shows the different layers present in an ANN. The input layer regulates how the input is leaked into the whole system, a mechanism that remains fixed in the case of RC. In the quantum systems we will focus on, we feed the information through an observable of a single qubit and it spreads out through the system's dynamics, which determines the W_{in}. The output layer and the weights corresponding to the hidden layer (or layers) are thus the ones to undergo the training process. It has long been known that, in order to approximate an arbitrary nonlinear function, the network must have at least one hidden layer [8] (which is where the recurrent connections are accommodated), hence the distinction between the hidden and output layers. This necessary amplification of the parameter space has to do with the impact of dimensionality in the learning process, an effect we will discuss in more detail for the systems studied in this work as well. In short, high dimensionality in

[1] As it usually happens, not everything is black or white. Classification of images may be performed by feeding the pixels to a RNN as a temporal sequence, and a FFNN may perform temporal tasks as well if equipped with a delayed embedding that turns the temporal problem into a spatial one [2].



(a) Generic Neural Network workflow

(b) Typical ANN scheme

Figure 1.1: Outline of the inner workings of ANNs.

prediction tasks is linked to a better performance because it allows the network to read out spatiotemporal dependencies in the input sequence [9]. However, this characteristic cannot be taken to the extreme arbitrarily or one may incur in overfitting, where the system adjusts too much to the training set and fails to predict a different one.

RNNs have achieved great performances in a wide variety of temporal tasks since the 1980s. However, this high accuracy is brought about along with a massive amount of hidden nodes, which makes the training of the hidden layer very demanding computationally. A way around this problem was proposed in the early 2000s with the development of the models of the Echo State Network (ESN) [10] and Liquid State Machine (LSM) [11], the direct precursors of the RC model. The idea they investigated was the requirements for a network in which it was sufficient to train only the readout weights, thus sparing the complex optimisation of the weights in the hidden layer (which remain fixed). In other words, they introduced the concept of reservoir.

1.2 Reservoir Computing

The lesson we learn from LSM and ESN is that, if the network's dynamics is complex enough, it is sufficient to train only the readout layer of the machine. Evidently, the notion of "complex enough" needs to be further determined, which will be the first of the main tasks of this section. The second is presenting a more rigorous definition of the classical RC scheme before we move on to RC based on physical systems instead of *in silico* architectures.

In the RC language, a reservoir machine $\mathcal{M}$ has two essential parts: a reservoir $\mathcal{R}$ and a readout layer ψ.

$$\mathcal{M} = (\mathcal{R}, \psi) \tag{1.1}$$

The reservoir remains fixed over time and it is the one to carry out the processing of information, i.e., the computation in the strict sense. The output layer, on the other hand, is trained

for each particular task and acts merely as a translator from the language of reservoir states to a form of data we can understand. Luckily, since all the nonlinear computation is performed in the reservoir, a linear transformation suffices in order to extract the desired output from the machine. Least squares minimisation is the most common choice for the training of the output layer because of its simplicity, sometimes with some slight regularising additions. This will also be the training methodology of the study presented here.

For a system to successfully perform any information processing task, thus qualifying as a good reservoir, it must satisfy some key properties: separability and the echo state and the fading memory properties [12]. These are, in fact, quite general, so much so that they are often fulfilled by an arbitrary physical system. This realisation brought about the paradigm of physical RC, where the neural network which constitutes the reservoir in conventional RC becomes any suitable physical, dynamical system, and the computation is left to the physical phenomena that naturally take place in it (see Fig. 1.2). Many lines of research are already focusing on this computational scheme because it allows for a most convenient processing of on-line information coming from a bigger physical system, if the reservoir is simply part of the latter. This saves up time and prevents possible losses in the transport of data to the information processing machine, as well as considerably simplifying the implementation of such element in the main system. The studied physical reservoirs are as diverse as the areas that benefit from it, from plain water inside a bucket for pattern recognition [13] to a silicone octopus arm for dynamical data processing [14]. Electronics, biology, brain and cell-related projects, photonics or spintronics are some other areas actively exploring the computational capabilities of the associated physical systems, be it for the purpose of energy-efficient ML hardware, understanding the processing of information on biological systems or simply exploring computation on novel substrates [9].

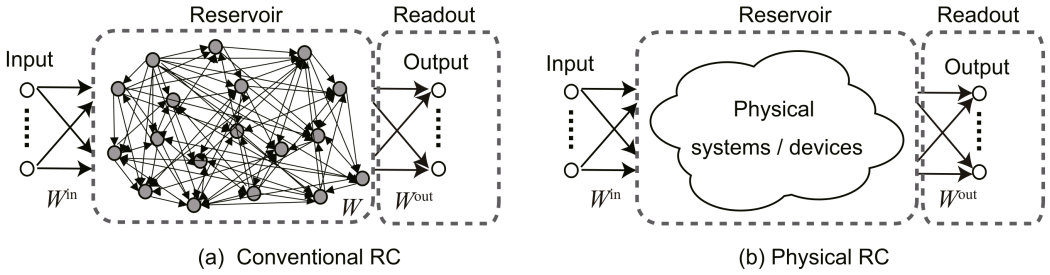


Figure 1.2: The two possible models for RC depending on the nature of the reservoir. Figure taken from [9].

Coming back to the general RC framework, we will now explain in more detail the aforementioned required properties for an arbitrary reservoir, which were separability, the echo state property and the fading memory property. Separability refers to the system's ability to produce different outputs for different inputs, i.e., the map implemented by the reservoir must be injective. The echo state and fading memory properties are strongly interconnected, and in fact it has been proven that a contractivity of the map and a compact input set imply the presence of both of them [15]. In order to be able to perform nonlinear operations, a reservoir must inspect the previously received input upon the arrival of each new instance in order to produce an output. In engineering, such a device is referred to as a filter. For an optimal

computation, we want our system to remember previous data, but only up to a certain point: it is essential that it forgets its initial condition, otherwise we would never be able to get the same output from a given input string. The echo state stresses the filter's reminiscence ability while the fading memory guarantees that the system's starting point will stop being relevant over time, but both characterise the reservoir as a filter. The mathematical definition of the echo state and fading memory properties can be found in Appendix A.

To close this section we proceed to the description of the general mathematical framework of RC, which we will later particularise to the case of our own model and implementations. Consider an arbitrary time series u that maps each time instance $t \in \mathbb{Z}$ to a value of the bounded set $\mathbb{U}$, such that

$$u : \mathbb{Z} \to \mathbb{U} \tag{1.2}$$

$$u_i \in [u_{min}, u_{max}] \qquad u_{min}, u_{max} \in \mathbb{R} \tag{1.3}$$

Since the reservoir is a dynamical system, its current state will inherently depend on the previous one, which amounts to a dependence on the past input history. Thus, given a time series u, the reservoir implements a transformation such that:

$$x_t = T(x_{t-1}, u_t) \tag{1.4}$$

where x is the state of the system and $T(\cdot)$ is the functional encoding the action of the reservoir after the injection of new input. The output layer then translates the state of the reservoir as follows:

$$o_t = h(x_t) \tag{1.5}$$

where $h(\cdot)$ is a functional. This is the most general description we can make of the mapping implemented in our numerical experiments, which will be further illustrated in Section 1.5.

1.2.1 Quantum Reservoir Computing

The benefits of high complexity for ML in general are what makes the quantum realm an interesting playground for all the ideas we have talked about so far. In classical systems, this complexity is achieved by connecting a number of nodes that is between large and massive. However, the high dimensionality of a Hilbert space allows one to obtain such levels of complexity with exponentially less elements to control: note that in a classical reservoir the number of degrees of freedom grows only polynomially with system size ($d.o.f \sim N$), while in a quantum reservoir it does so exponentially ($d.o.f \sim 4^N$ in the case of a spin-1/2 system). This landscape naturally fits into the picture of reservoir computing, since most of the computation takes place by means of nodes we are not accessing through the observables we have chosen to measure (acting as hidden layers) but which may be monitored through merely a few true nodes. We recall that a node corresponds to a degree of freedom of the system, and thus by a true node we are not referring to a physical component of the system (such as a qubit) but to a degree of freedom we directly access, i.e., an independent magnitude that is measured. An illustration of the Quantum Reservoir Computing (QRC) scheme we have just described may be found in Fig. 1.3.

QRC has already proven very successful in the processing of quantum information [6], albeit this work is limited to the task of time series prediction of classical information. We further

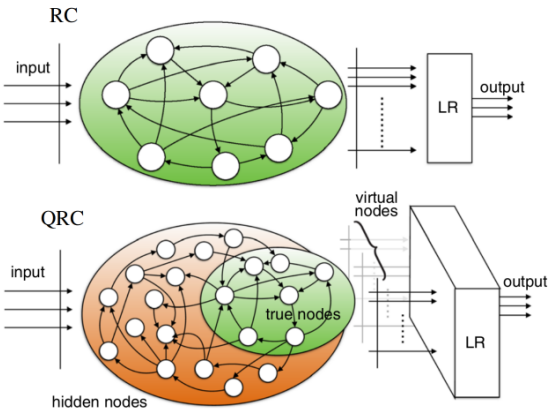


Figure 1.3: Comparison between the classical RC and QRC schemes. The output layer is directly referred to as a linear readout (LR) in this figure, which was taken from [16].

comment that, although the practical implementation of QRC systems remains technologically challenging, today's laboratories are already prepared to build and monitor small versions of these reservoirs with optical systems [17] (an approach particularly promising in terms of scalability), Nuclear Magnetic Resonance (NMR) [18] and cold ions [19, 20]. These last two allow for the implementation of transverse-field Ising reservoirs, a promising QRC candidate proposed in 2017 by [16] that has spurred further research from this perspective, including the work we concern ourselves with here.

1.2.2 Information Processing Capacity

It is hard to define a "better" or "worse" reservoir as an abstract entity beyond the properties already described at the beginning of this section. Once we have good reservoir, i.e., exhibiting separability, the echo state and the fading memory property, we do not expect it to perform optimally at every task. Such reservoir will have finite capacity of retaining past information, for example, which may or may not be enough for providing a good performance in the particular task at hand. This question raises not only on RC, but on the whole ML zoo. To address it, the most common practice is examining how well the system performs on some benchmark tasks, which give an idea of the system's weaknesses and strengths. For example, the timer task is an indicator of how far back the system can reach for previous input. In this task, the system is trained to identify the change of the input string from zeros to ones with a certain delay. If this delay exceeds the system's linear memory capabilities, the system will fail to respond correctly and a reliable threshold for this property is obtained (see Fig. 1.4). When it comes to examining the nonlinear processing capacity of a system, however, the possibilities are much more diverse, and one can only extract a general idea of a system's capacity rather than more concrete information. The NARMA tasks are particularly widespread for this purpose, each of them consisting on approximating a fixed nonlinear polynomial function [16, 18, 21]. Luckily, on 2012 a more general measure of a system's processing capacity was introduced [22]: the IPC (Information Processing Capacity). This measure allows us to scan the system's whole skill-set and see the fraction of it that addresses n-degree nonlinearities,

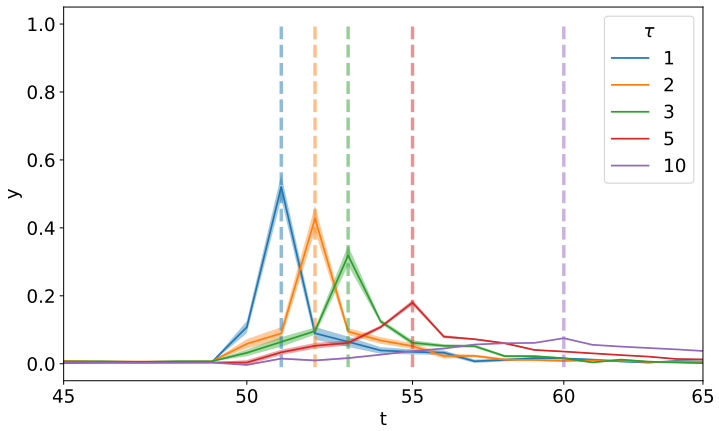


Figure 1.4: Example of the performance of $N = 6$ qubits in the timer task for different delays (τ). y refers to the prediction given by the reservoir, which in this case should approximate to a delta of height 1 for delay (dashed lines).

though it comes with a high computational cost.

The IPC aims to avoid the artificial distinction that is usually made between linear and nonlinear information processing, creating a scale that allows us to compare different dynamical systems on several levels. Each of this levels is associated to a degree d of nonlinearity in the processing of information. This is done by evaluating the system's performance in a class of tasks that consist on approximating a certain polynomial target function of such degree d.

$$\overline{y}_k = \prod_i \mathcal{P}_{d_i}[\tilde{s}_{k-i}] \qquad \sum_i d_i = d \tag{1.6}$$

The $\mathcal{P}_{d_i}$ appearing in Eq. (1.6) refers to the Legendre polynomial of degree d_i, and the $\tilde{s}_{k-i}$ to the input with a delay of i time steps. The performance of the system for each of the tasks described in (1.6) is quantified by a certain capacity C_L, where the L indicates a long-time average over the input string of size L. The sum of all the capacities of the tasks corresponding to a given d (C_d) provides the capacity we associate to that degree. The order in which one should collect all these contributions in order not to miss any, a non-trivial matter, is explained in the supplementary material of [22].

$$C_L = 1 - \frac{\min_{\{W\}} MSE_L(y, \overline{y})}{\langle \overline{y}^2 \rangle_L} \tag{1.7}$$

In Eq. (1.7), $\langle \overline{y}^2 \rangle_L$ is the square average of the target and $\min_{\{W\}} MSE_L(\cdot)$ is the mean squared error between the prediction produced from the weights W (the ones resulting from the minimum least squares fit, analogous to the w we discussed in Section 1.1) and the target function.

$$\langle \overline{y}^2 \rangle_L = \frac{1}{L} \sum_{k=1}^{L} \overline{y}_k^2 \tag{1.8}$$

$$MSE_L(y, \overline{y}) = \frac{1}{L} \sum_{k=1}^{L} (y_k - \overline{y}_k)^2 \tag{1.9}$$

One of the most interesting results in [22] is that the sum of the capacities for all degrees C_d, which we will call $C_{TOT} = C$, is bounded from above by the number of output functions. Furthermore, if the system has fading memory this bound is saturated.

$$0 \leq C = \sum_{d=1}^{\infty} C_d \leq N \tag{1.10}$$

The ∞ in the upper bound of the degree to be computed is not to be feared, since the contributions become less significant as the degree is increased. In practice, for the system sizes concerning us, the summation in (1.10) may be truncated at $d = 6$. We have conveniently named the number of output functions N in Eq. (1.10) because in our system we will obtain our output functions by measuring each of the spins present (a detailed explanation can be found in Section 1.5), so that in our case the number of output functions will coincide with the number of spins in the network. As a result, we will always present the capacity already normalised by this factor in the present work. Some additional notes on the numerical obtention of the IPC are explained in Appendix B.

1.3 The resource theory of coherence

In a way, coherence constitutes one of the conceptual bridges between classical and quantum dynamics. It's an expression of the "quantumness" of a given state, which may be complementary to the presence of correlations. In a closed multipartite system with interactions, coherence usually builds up on its own up to a certain value due to the system's natural dynamics. However, in the presence of interactions with the environment (i.e., dissipation), there is a flow of information leaking out of the system and thus coherence tends to decrease. In the stationary regime ($t \to \infty$) of this process there may be a non-zero coherence, but it will be, in any case, lower than that of the closed system.

The consideration of coherence as a resource has already proven useful in several fields, such as quantum metrology, quantum algorithms, quantum thermodynamics, the witnessing of quantum correlations and even quantum biology and transport phenomena [23]. In this work, we aim to explore how coherence may present itself as a resource for a quantum dynamical system to function as a reservoir. The basic argument is that a proper flow of information within the system (and, with it, a proper performance of the reservoir) can be traced through its coherence, being better the former the higher the latter is. In fact, quantum coherence has already come to the attention of some researchers as a feasible probe for studying a system's different dynamical regimes [24], a line of research this work contributes to.

The coherence measure we are going to use is presented in [23], and it corresponds to the most direct illustration we may have of this magnitude; the sum of the off-diagonal elements of the density matrix:

$$C_{l_1}(\rho) = \sum_{i \neq j} |\rho_{ij}| \tag{1.11}$$

We comment that the somewhat cumbersome notation of C_{l_1} comes from the fact that it a particular case of the general matrix norm C_{l_p}[23]:

$$C_{l_p}(\rho) = \left(\sum_{i \neq j} |\rho_{ij}|^p \right)^{1/p} \tag{1.12}$$

According to definition (1.11), the natural state to take as a reference for how much coherence a system has is the maximally coherent state Ω, for which $\Omega_{ij} = 1/2^N \ \forall \ i, j = 1, ..., 2^N$. It is easy to check that the corresponding maximum coherence for system of N qubits is

$$C_{l_1}(\Omega) = C_{max}(N) = 2^N - 1 \tag{1.13}$$

which corresponds exactly to the number of degrees of freedom in the system in a pure state.

1.4 Notes on the ergodicity of quantum systems

Before we move on to the introduction of our system, it will be interesting to go through some important concepts regarding the ergodicity of quantum systems. Our aim is to build the bridge between the description of a quantum system and that of a reservoir in order to fully dive into QRC. This will consist, on the one hand, on detailing the way in which a quantum system may hold a fading memory, which relates to the concept of thermalisation in quantum mechanics. On the other hand, some aspects of Many-Body Localisation (MBL) will be outlined, since this phenomenon will be present in our system in a number of cases.

The Eigenstate Thermalisation Hypothesis

Describing the quantum analog of what we call ergodicity in the classical world is not a straight-forward task. Under careful inspection, we realise that the common definition is based on the concept of trajectory, and thus cannot be directly translated into the quantum realm. To be more precise, a system is said to be ergodic if all its microstates are accessed with equal probability over a long period of time. This requires that the system completely forgets where it came from, a property that seems impossible to fully achieve in quantum mechanics just by looking at its algebra. Consider a closed system evolving unitarily:

$$|\psi(0)\rangle = \sum_{\alpha} A_\alpha |\alpha\rangle \tag{1.14}$$

$$|\psi(t)\rangle = e^{-i\mathcal{H}t}|\psi(0)\rangle = \sum_{\alpha} A_\alpha e^{-iE_\alpha t}|\alpha\rangle \tag{1.15}$$

The probability of finding the system in a given eigenstate $|\alpha\rangle$ is $p_\alpha = |A_\alpha|^2$, which implies that the initial condition $\{A_\alpha\}$ remains forever encoded in the state of the system even for infinitely long times.

A new approach to the definifion of ergodicity on quantum systems was intoduced by von Neumann in 1929 [25], stating that a closed/open quantum system may be considered thermalised if all its observables reach the expectation values given by the microcanonical/-macrocanonical ensemble. The long-time average of any observable for the system in Eq. (1.15) is described as

$$\langle \mathcal{O} \rangle_\infty = \lim_{T\to\infty} \frac{1}{T} \int_0^T \langle \psi(t)|\mathcal{O}|\psi(t)\rangle dt = \sum_\alpha p_\alpha \langle \alpha|\mathcal{O}|\alpha\rangle \tag{1.16}$$

A way for a system to adjust to von Neumann's definition taking Eq. (1.16) into account is to have each of the system's eigenstates satisfy it individually, i.e.,

$$\langle \alpha|\mathcal{O}|\alpha\rangle = \mathcal{O}_{mc}(E_\alpha) \tag{1.17}$$

Eq. (1.17) summarises the Eigenstate Thermalisation Hypothesis (ETH), which is the theoretical basis for what we currently understand about quantum thermalisation in closed systems.

Many-Body Localisation

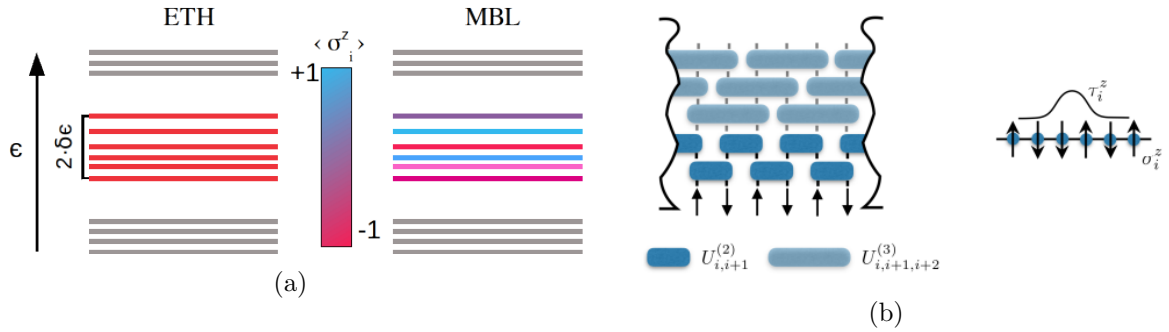
The phenomenon of Many-Body Localisation is fairly rare in natural systems because it implies the lack of thermalisation in the presence of interactions in the infinite time limit. An illustration of this violation of ETH can be seen in Fig. 1.5a, where eigenstates corresponding to the same energy density ϵ present different values of an observable in the infinite-time limit and thus cannot satisfy Eq. (1.17). To be more precise, we are always implicitly considering an interval $[\epsilon - \delta\epsilon, \epsilon + \delta\epsilon]$ due to environmental noise or experimental precision. Thus, E_α in Eq. (1.17) is virtually the same for all eigenstates contained in such band ($E_\alpha \simeq \epsilon$), and $\mathcal{O}_{mc}(E_{\alpha_1}) \simeq \mathcal{O}_{mc}(E_{\alpha_2}) \simeq ... \simeq \mathcal{O}_{mc}(\epsilon)$. The main ingredient in MBL is the presence of disorder [26, 27, 28]. We may think of MBL as a localisation in configuration space resulting from this strong disorder. If this property survives despite the driving of the injection of input, it will amount to the system lacking a fading memory, yet another way to emphasise that a system in this phase cannot constitute a good reservoir. Nonetheless, this phase is still capable of spreading quantum information throughout the system (i.e., the system will reach a stationary state after we feed it input), although it does so at a considerably slower rate than an ergodic system [26, 27].

The properties described above can be explained by the emergent integrability that arises in MBL systems. It so happens that MBL eigenstates are nothing but slightly deformed product states, and thus a quasi-local unitary transformation is sufficient to diagonalise the Hamiltonian in this given product state basis. A unitary operator U is said to be quasi-local if it can be written as the product of n-site operators ($n = 2, 3, 4, ...$)

$$U = \prod_i ... \, U^{(3)}_{i,i+1,i+2} \, U^{(2)}_{i,i+1} \tag{1.18}$$

where the higher n operators induce exponentially smaller rotations. For a system of spins like the one we will concern ourselves with, a fully connected network with a transverse magnetic

field with local fluctuations, in the case of strong disorder (i.e., strong fluctuations) the product state basis will be given by the operators σ_i^z. Thus, the operators $\tau_i^z = U^\dagger \sigma_i^z U$ commute with the Hamiltonian and constitute a complete set of quasi-local integrals of motion, commonly referred to as LIOMs. We remind the reader that, since the action of the high n-site operators is exponentially small, the τ_i^z have a vanishing effect far away from site i, thus giving rise to this localisation. The quasi-local transformation we have just described is represented in Fig. 1.5b.



(a) (b)

Figure 1.5: Schematic illustration of some of the main characteristics of MBL. (a) shows a comparison of the different eigenstates sorted by energy in thermalising (left) and MBL (right) regimes. The different colours indicate the expected value of the z-component of the spin at site i, a local observable, for that eigenstate. Note that the level spacing would not be identical for both cases in reality, but in this schematic representation we have made no such distinction. (b) represents the transformation induced by the quasi-local operator U in the product states so as to produce MBL eigenstates (left figure, illustrated in the tensor network fashion), and in the local operators $\{\sigma_i^z\}$, which produces the set of LIOMs $\{\tau_i^z\}$ (right figure). Figures adapted from [27].

1.5 Methodology

This section is dedicated to the description of the scheme of the numerical simulations. We begin with a system of N spins in an initial state described by a density matrix ρ of size $2^N \times 2^N$. The unitary evolution of the network will be described by a Hamiltonian H, such that at every time step of length Δt the system is transformed as follows:

$$\rho[(k+1)\Delta t] = e^{-iH\Delta t}\rho[k\Delta t]e^{iH\Delta t} \tag{1.19}$$

The task posed to the system is always predicting a time series. All input strings used here are random numbers uniformly distributed between 0 and 1, in both training and testing sets as well as for the warming up set necessary for reaching a stationary state. The data is fed to the system at the beginning of each Δt time step always through the same qubit, which we will identify as the first. This is done by modifying its state according to the following map:

$$\rho_1[s_k] = \begin{pmatrix} 1-s_k & 0 \\ 0 & s_k \end{pmatrix} \tag{1.20}$$

Note here that $\{s\}_k \in [0,1]$, so if the original data does not satisfy this criterion it must be rescaled first. This is our case for the calculation of the IPC, in which the Legendre polynomials are evaluated between -1 and 1, so that $\tilde{s}_k \in [-1,1]$ in Eq. (1.6). In this case, both input strings are related by $\tilde{s}_k = 2s_k - 1$. Then, the information is introduced into the whole system by letting the system evolve naturally from its new state ρ' according to Eq. (1.19), where ρ' is nothing but the tensor product of the new density matrix ρ_1 and the partial trace of the whole system over the first spin.

$$\rho'[k\Delta t] = \rho_1 \otimes \mathrm{Tr}_1 \rho[k\Delta t] \tag{1.21}$$

For the collection of data from the system, we measure the z-component of each of the spins of the network. The choice of this observable is guided by experimental simplicity and following the work of [16]. As a result, our reservoir has N true nodes and a maximum of N independent output functions, which will set the normalisation constant of the total capacity for the system (see Section 1.2.2). Following [16], in our simulations we are always considering a setup where we have a large number of copies of the system, an assumption that allows us to disregard the back-action after measurement. Let it be further noted that when we refer to "a copy" of the system we actually mean a large number of them, since we are directly considering the expected value in our numerical simulations and this physically constitutes the average over several measurements of an identical system. Thus, the data is collected as shown in Eq. (1.22) and the density matrix of the system remains the same afterwards.

$$x_i(k\Delta t) = \langle \sigma_i^z \rangle = \mathrm{Tr}\left[\sigma_i^z \rho(k\Delta t)\right] \tag{1.22}$$

where $i = 1, ..., N$ and σ_i^z is the ordered tensor product of the 2×2 identity matrix for $j \neq i$ and the z-component Pauli matrix for spin i ($\sigma_i^z = I_1 \otimes ... \otimes I_{i-1} \otimes \sigma_i^z \otimes I_{i+1} \otimes ... \otimes I_N$). We also add a constant bias term ($x_{N+1} = 1$) in order to regularise the training algorithm of the output layer.

After an initial number of time steps ζ in which we let the system reach a stationary state, we collect data for L_t time steps and use the least squares method to obtain $N+1$ weights. These are the ones that will be used to make predictions from the testing input string, of length L_{uk}. Usually, we will have $L_t = L_{uk} = L$, although this is not a necessary ratio.

Error estimation

Throughout this work we are dealing with randomly generated systems from a given set of parameters, and thus it is frequently necessary to average over a number of systems N_{sys} in order to obtain an accurate picture of the general behaviour under study. When measuring a certain magnitude A in the stationary regime (the coherence, for example), we proceed as follows: after letting the system n reach such stationary regime with ζ time steps, we measure it in the following L time steps and obtain a mean a_n and a variance Δa_n from that data string. Once we have these values for each different system, we consider the final measure of the desired magnitude to be the mean value of the a_n, and its associated deviation ΔA is

calculated by simple error propagation.

$$\overline{A} = \frac{1}{N_{sys}} \sum_{n=1}^{N_{sys}} a_n \tag{1.23}$$

$$\Delta A = \frac{1}{N_{sys}} \sqrt{\sum_{n=1}^{N_{sys}} (\Delta a_n)^2} \tag{1.24}$$

In any case, the biggest contribution to the deviation from the mean $\overline{A}$ comes from the different system realisations rather than the fluctuations relative to a single system.

Chapter 2

The system

One of the interesting candidates for QRC are networks of spins, the paradigmatic example of QRC with discrete variables. The proposed models of this type can be brought to life in a lab with trapped ions or NMR (Nuclear Magnetic Resonance) settings. The work presented here contributes to the research into the performance of transverse-field Ising systems as reservoirs, opened by [16] and continued in [29, 21, 30].

We consider an ensemble of N spins with random all-to-all interactions drawn from a uniform distribution, and subject to a magnetic field in the transverse direction. It is akin to the Sherrington-Kirpatrick model [31], one of the simplest [1] devised to explain the spin glass (SG) phenomenology. In this model the magnetic field is the same at every site, so that the system is described by the following Hamiltonian:

$$\mathcal{H} = \sum_{i>j} J_{ij}\sigma_i^x\sigma_j^x + h\sum_i \sigma_i^z \tag{2.1}$$

where σ_i^α is the α-component of the spin acting on qubit i and the $\{J_{ij}\}$ are uniformly distributed over the interval $[-J_s/2, +J_s/2]$. We further comment that throughout this work we have taken $J_s = 1$ in order to set our energy scale. Systems like the one described in Eq. (2.1) already are experimentally feasible with the technology of trapped ions, as demonstrated in [20], although here the coupling constants $\{J_{ij}\}$ were not drawn from a uniform distribution but decrease algebraically with the distance between spins. As it will be confirmed in our analysis, this model harbours a quantum phase transition between an ergodic and a non-ergodic phase, the latter materialising in a SG phase.

Afterwards, we will consider a more general version of this model in which we allow the magnetic field to have random local fluctuations. These will be bounded by a parameter W and uniformly distributed, such that

$$\mathcal{H} = \sum_{i>j} J_{ij}\sigma_i^x\sigma_j^x + \sum_i h_i\sigma_i^z \tag{2.2}$$

[1] In the sense that it is exactly solvable, but the mathematical machinery required to do so is, in fact, quite sophisticated. In Sherrington-Kirpatrick, the couplings are drawn from a gaussian distribution instead of a uniform one.

where $h_i = h_{mean} + w_i$ for $w_i \in [-W, +W]$. This introduction of further disorder allows for the appearance of Many-Body Localisation (MBL) for high W. A more exhaustive discussion of the MBL phase is presented in Section 1.4, so here we simply recall that it is non-ergodic, making this regime unsuitable for computation. MBL has been identified in similar spin systems as well, as the Sherrington-Kirpatrick model [32]. Although the model in Eq. 2.2 has been less studied in the literature (at the time of the redaction of this work), the phase diagram it produces has already been characterised in other works, both theoretical [21] and experimental [19].

Determination of an appropriate time scale

Before the beginning of the study, a suitable time step between inputs Δt for capturing the dynamics of the system must be set, long enough for the information fed to be properly processed. There is a threshold Δt setting this distinction for a given set of system parameters, which corresponds to the minimum time the system needs to relax after the quench produced by the injection of input. Here, we have numerically determined the threshold Δt (as shown in Fig. 2.1) for a given set of parameters and chosen to fix a value sufficiently far from this threshold, $\Delta t = 10$, so that it is still suitable for other system sizes we will be considering. We will mainly focus on the case $N = 5$ because it provides a good compromise between the visibility of interesting effects and computational time, but we will occasionally consider $N = 6$ and $N = 10$ as well.

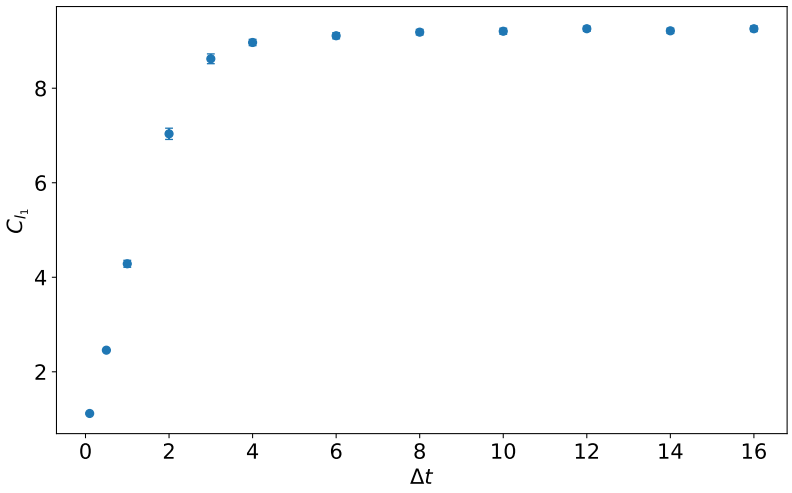
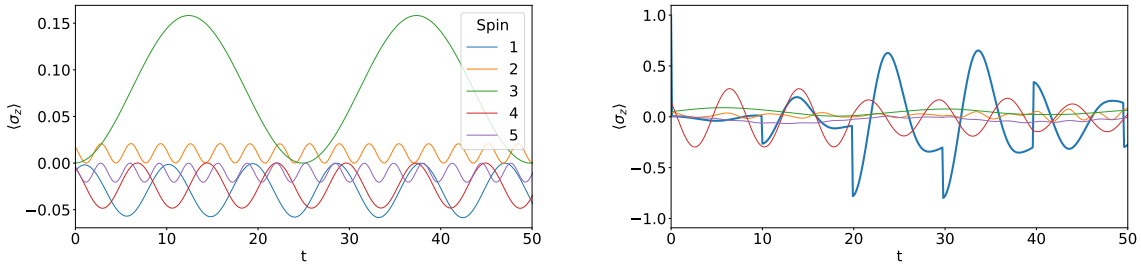


Figure 2.1: Stationary value of the coherence starting from an incoherent initial condition and averaging over 100 systems after $\zeta = 10^3$ steps for $N = 6$, $h = 1$ in the uniform field Hamiltonian. The point corresponding to $\Delta t = 0.1$, however, was calculated with $\zeta = 10^4$, since for such a small interval between inputs the system needs much more time to reach stationary dynamics.

2.1 Uniform magnetic field

The model depicted in Eq. (2.1) interpolates between the dynamics of independent spins under the action of a magnetic field h in the limit $h \gg J_s$ (which we will refer to as the paramagnetic regime) and the dynamics of a SG for $h \ll J_s$. From this very description we can already presume that this system will not constitute an appropriate reservoir at least in the latter case, since SGs have long been known to retain memory of their initial condition. For $h \gg J_s$, however, the fading memory condition will be satisfied (in the presence of input) because of the strong influence of the magnetic field, but the flow of information back and forth throughout the system will also be hindered by its rigidity, which should somehow limit its processing capacity. Since the network won't be able to store and mix previous input as effectively, we expect these systems to have a lower non-linear processing capacity. The intermediate regime, on the other hand, should hold the most interesting properties of both extremes from the point of view of reservoir computing: an external magnetic field strong enough to dilute the initial condition but with important interactions between nodes, allowing for the proper flow of past inputs in the network.

We can already catch a glimpse of these differences by looking at the unitary evolution of the z-component of the spins. We start by looking at the dynamics in the SG regime, both unitary (see (2.1)) and with the insertion of input (adding the mapping (1.21)).



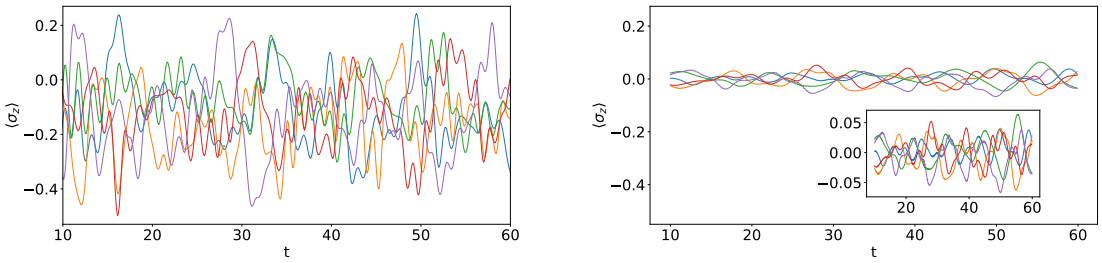
(a) Unitary evolution

(b) Evolution with input

Figure 2.2: Typical evolution of the system described by the Hamiltonian in Eq. (2.1) for $h = 0.01$ (a) and with the addition of the reservoir computing map (b). The thicker line in (b) corresponds to the spin through which random input is being fed to the system (spin 1), visible in the "kicks" it experiences. The point $t = 0$ corresponds to when we begin the measurement of the system, after it has reached the stationary state starting from a maximally coherent initial condition.

Fig.2.2a already shows how, in this regime, spins establish themselves in stationary states with their own periodic orbits, without interfering with each other any further. This independence persists when we introduce our data into the system, as we can see in Fig.2.2b, where the rest of spins remain essentially idle to the activity of the first one. Since information does not flow properly through the system, our previous assertion of it constituting an inadequate reservoir in this parameter regime is confirmed.

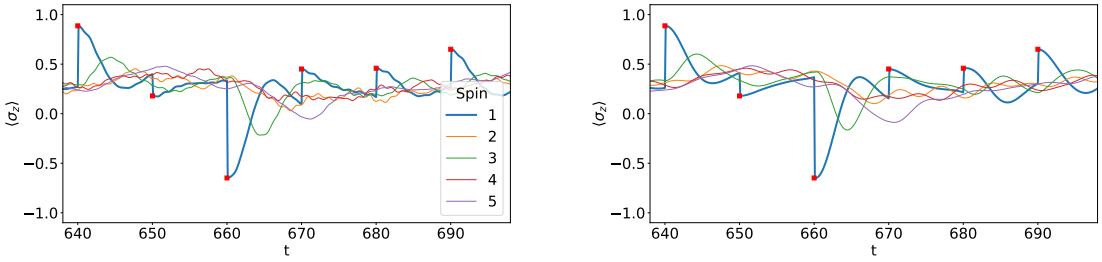
We may now look at the more promising cases for the purpose of RC: the intermediate and paramagnetic regimes. Figs. 2.3 and 2.4 present the evolution of a typical system in both scenarios for the non-dissipative and the dissipative case, respectively.



(a) Intermediate regime (b) Paramagnetic regime

Figure 2.3: Unitary evolution of all spins for the intermediate and paramagnetic cases for the same system with (a) $h = 1$ and (b) $h = 100$ in the stationary regime after starting from a maximally coherent initial condition. The inset in (b) presents a zoomed-in version of the trajectories in this regime, which have a considerably smaller amplitude than the intermediate case. The colour coding of the figure is explained in Fig. 2.2.

The difference between regimes that appears in the unitary dynamics of the system is in good agreement with our previous argument of the magnetic field limiting the flow of information through the network. Fig. 2.3 shows that the amplitude of oscillation of the spins is about an order of magnitude lower for $h = 100$, which points to a weaker response to the changes suffered by the rest of the network overall.



(a) Intermediate regime (b) Paramagnetic regime

Figure 2.4: Spin trajectories for the intermediate and paramagnetic cases, for a system with identical coupling strengths and (a) $h = 1$ and (b) $h = 100$, when the same input sequence is fed to both of them (red squares). Results correspond to the stationary regime after starting from a maximally coherent initial condition.

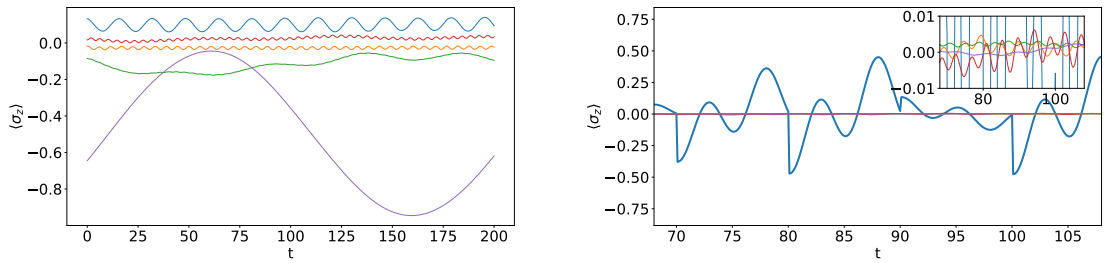
Despite the considerable difference observed between both regimes during the unitary evolution of the system in Fig. 2.3, Fig. 2.4 shows that the driving of the injection of input is strong enough to induce a very similar response in both cases. Note that in the case where we are introducing the input, the rest of spins follow the dynamics of the first. This is clearly seen, for example, around $t = 660$ in Fig. 2.4, where the peak of spin 1 produced by the injection of input is translated to spins 3 and 5 at $t = 665$ and $t = 671$, respectively.

In conclusion, this preliminary study of the system by looking at the time evolution of

its observables has revealed the following: in the low-h regime of the system described in (2.1), we expect computation to be unfeasible, since it remains in a spin-glassy phase and thus cannot forget its initial condition. The other two regimes, however, seem much more promising, since a noticeable magnetic field allows for the dilution of the initial condition and for the precession of spins in planes more parallel to each other. This last feature facilitates the appearance of resonances, which allow information to flow through the whole system.

2.2 Disordered magnetic field

We will now perform a similar study on the model from Eq. (2.2). This Hamiltonian has also been studied from the perspective of QRC with a disorder distribution different from the one presented here in [30] and with our same distribution in [21]. In particular, we are interested in the effect of this further disorder, characterised by W and introduced via spatial fluctuations in the transverse magnetic field. The theory presented in Section 1.4 predicts the appearance of an MBL phase at high disorder ($W \gg J_s$), in which computation will not be possible due to the violation of ETH (and, equivalently, the fading memory condition). Here we will merely present the evolution of the observables relative to $h_{mean} = 0.01$ since if $W \ll h_{mean}$ results will be qualitatively similar to those of a uniform magnetic field. Thus, we are going to set ourselves in what we previously referred to as the SG phase and see to which other dynamical regimes disorder takes us.

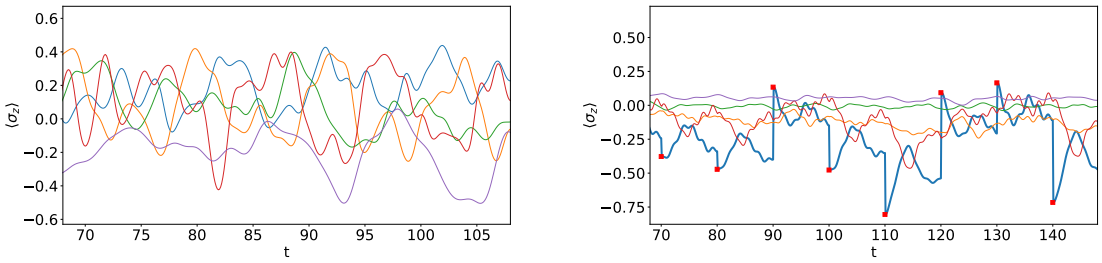


(a) Unitary evolution

(b) Evolution with input

Figure 2.5: Spin trajectories for the cases with unitary evolution (a) and input injection (b), for systems with identical coupling strengths $\{J_{ij}\}$, $W = 0.01$ and $h_{mean} = 0.01$. The colour code for the different spins is the same as in the previous section. Results correspond to the stationary regime after starting from a maximally coherent initial condition, and the colour coding of the figure is explained in Fig. 2.2.

We see that Fig. 2.5 shares its main features with Fig. 2.2: in the unitary case, each spin follows its own independent orbit, although somewhat distorted by the fluctuations of the magnetic field in the former case. The input map does bring about more differences, like the considerable reduction in the amplitude of oscillation of the rest of spins (whose origin remains unclear), but in any case none of them are able to follow the input signal. Therefore, this region remains unsuitable for RC, as the disorder is not enough to spur the transport of information.



(a) Unitary evolution

(b) Evolution with input

Figure 2.6: Spin trajectories for the cases with unitary evolution (a) and input injection (b) for the same system with $W = 1$ and $h_{mean} = 0.01$. Results correspond to the stationary regime after starting from a maximally coherent initial condition, and the colour coding of the figure is explained in Fig. 2.2.

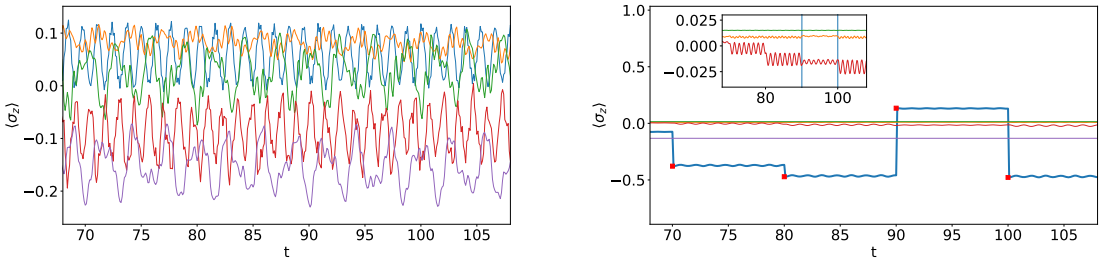
The picture presented in Fig. 2.6 as we increase W is now more reminiscent of what we had observed in the ergodic regime for the case of a uniform magnetic field, in particular to the intermediate regime due to the amplitude of the unitary trajectories and a certain lack of smoothness in the case with the input map. In the multiple-valley picture of a spin glass [2] , the disorder has flattened out the landscape and the system is now able to diffuse and process information efficiently.

However, when disorder is too strong we return to a situation in which the landscape's local minima are shaped by the disorder on the magnetic field, and the system gets stuck in them over again. This corresponds to the MBL phase, represented in Fig. 2.7. Despite the feasible analogy, one must not equate the phases that appear in both extremes (SG and MBL), as they are qualitatively different states born through different mechanisms. In particular, MBL is a strictly quantum phenomena while a spin glass may exist without the need to introduce the concept of superposition [26].
Fig. 2.7b shows best the many-body character of the states accessible to the system, as the input kicks generate a minimal little effect over the rest of spins and the first spin remains essentially static around the position it was set in. This picture proves consistent with the phenomenon of localisation and evidences that this system cannot constitute a good reservoir, as we had already predicted from the theory.

Previous studies [21] report a full characterisation depending on magnetic field and disorder parameters of the regimes we have inspected in this chapter. In particular, ergodicity is identified by looking at more specific features, like the average level spacing or the convergence of two final states starting from different initial conditions. We present some of these results in Fig. 2.8, as the ergodicity and non-ergodicity of the different phases will prove to be of great importance for the system's IPC. This characterisation has been made by looking at the spacing between the system's energy levels, more precisely the ratio between adjacent gaps r:

$$r = \frac{\min\left(\delta_n, \delta_{n+1}\right)}{\max\left(\delta_n, \delta_{n+1}\right)} \tag{2.3}$$

[2] Spin glasses are well illustrated by their free energy landscape in state space, which presents many local minima. The system tends to get stuck in these valleys, thus failing to explore the rest of the phase space.



(a) Unitary evolution

(b) Evolution with input

Figure 2.7: Spin trajectories for the cases with unitary evolution (a) and input injection (b) for the same system with $W = 10$ and $h_{mean} = 0.01$. The inset in (b) shows a close-up over some of the spins, so that the abrupt change of their state with the injection of new input may be noticed. Results correspond to the stationary regime after starting from a maximally coherent initial condition, and the colour coding of the figure is explained in Fig. 2.2.

where $\delta_n = E_n - E_{n-1}$. It is an established fact that, for a thermalising system, the Hamiltonian's eigenvalues are distributed according to Wigner-Dyson statistics, while localisation generates a Poisson distribution [26]. According to [33], a Hamiltonian satisfying ETH must have $\langle r \rangle \simeq 0.535$, while one that presents localisation will have $\langle r \rangle \simeq 0.386$. Thus, the lighter regions in Fig. 2.8 indicate that the system is ergodic for that set of parameters, while the dark regions indicate the contrary. Looking past the slight deformation in shape due to the difference in system size, in the next chapter we will be able to properly confirm that the intermediate and paramagnetic regimes we have previously referred to belong to the ergodic area, while the SG and MBL regimes stand in the dark region. As a last reminder before we move on to the study of the system's coherence and IPC, we once again stress the fact that ergodicity is a necessary requirement for a system to be able to act as a reservoir, since it guarantees that information spreads through the reservoir and that the system tends to a state independent of the initial condition.

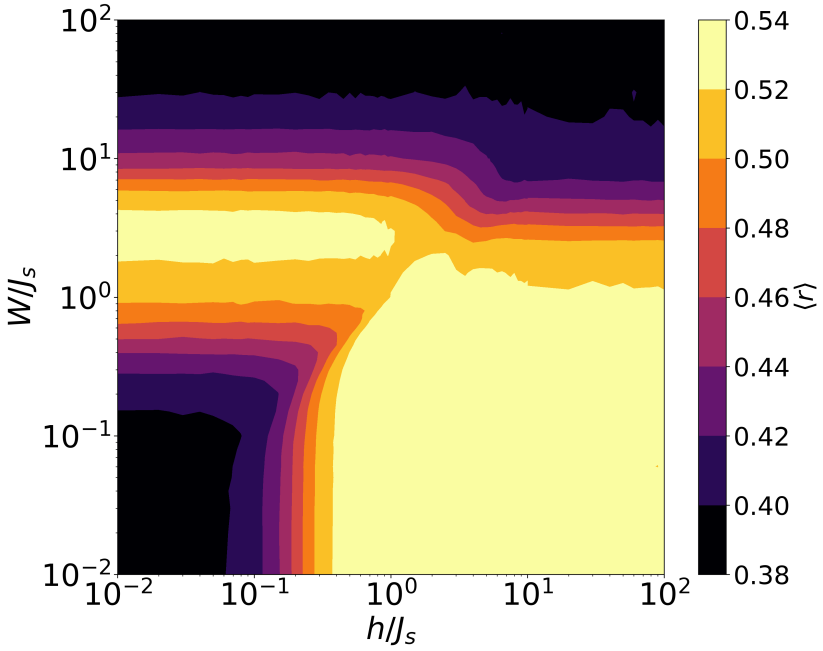


Figure 2.8: Level spacing ratio $\langle r \rangle$ as a function of the magnetic field and disorder for a system of $N = 10$ spins. Results are averaged over 1200 realisations. This plot was extracted from [21] with the authors' permission.

Chapter 3

Results

We will now move on to the presentation of two different sets of results. First, we will identify the different regimes we examined in the previous section through the system's coherence, associating it to its information processing capacity in that region. Afterwards, we will analyse the system's response when it is submitted to an undesired dissipation process. Our goal here is to explore purely quantum aspects of such system, which constitute a fundamental tool towards the exploration of possible quantum advantages. We will also probe the robustness of this model in the presence of decoherence, providing a more realistic picture of the performance of an experimental realisation of our reservoir.

3.1 Coherence vs. IPC

Now that we have familiarised ourselves with the system, both by itself and as a reservoir, we may focus on the analysis of its coherence and the information about the processing capacity we may extract from it. We begin by presenting a heatmap of the stationary coherence of the system as a function of W and h_{mean} (i.e., directly considering the more general case of Hamiltonian (2.2)) in order to relate the different phases we identified in the previous section with our coherence measure. For the realisation of this set of simulations we have always considered a completely incoherent initial state with all spins pointing upwards. This is relevant in the case of the non-ergodic phases, since they will reach a higher stationary coherence if their initial condition is the maximally coherent state than if they start start from a maximally incoherent one. The obtained results are presented in Figs. 3.1 and 3.2.

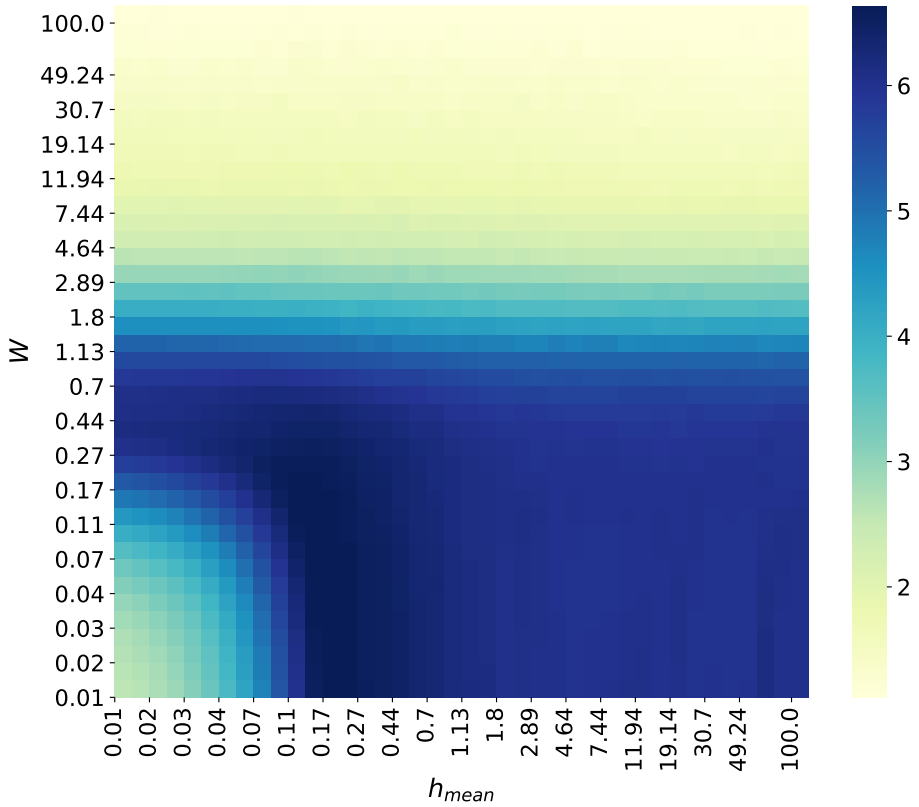


Figure 3.1: Stationary coherence $\langle C_{l_1} \rangle$ for a system of $N = 5$ spins. The resulting value of each point has been averaged over 100 systems, within which we have also averaged over $L = 1000$ measurements taken after $\zeta = 1000$ steps.

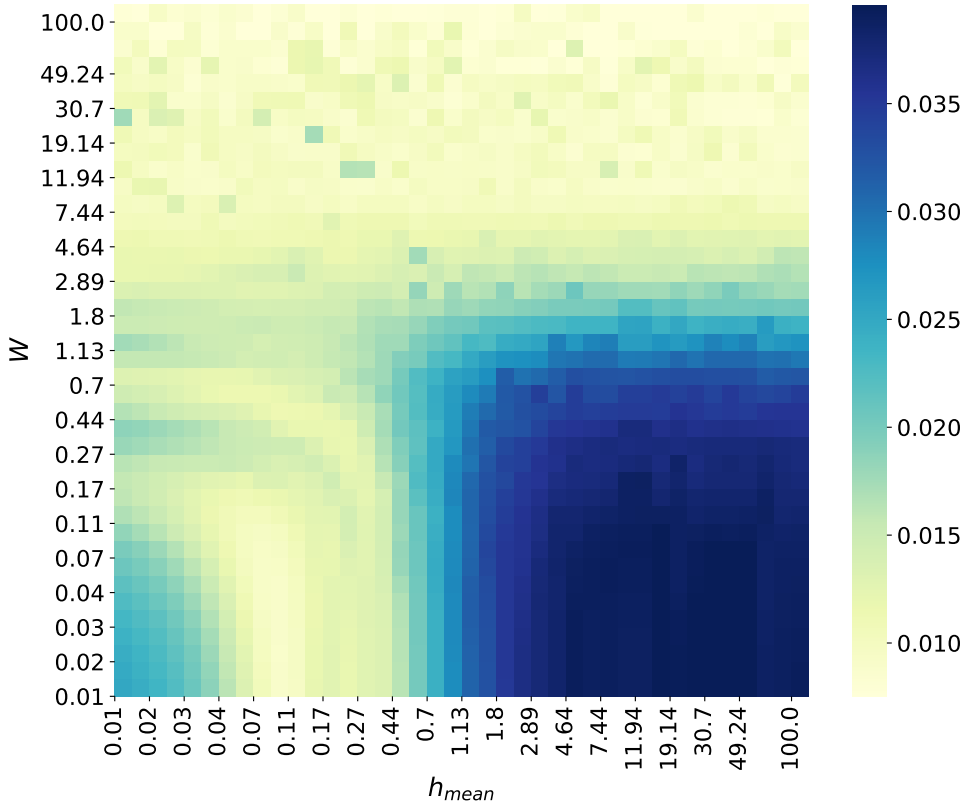


Figure 3.2: Associated deviation from the average stationary coherence ΔC_{l_1}. Simulation parameters are the ones stated in the caption of Fig. 3.1.

Despite the deformation caused by the different sizes of the network, Fig. 3.1, where $N = 5$, presents the same structure as Fig. 2.8, where $N = 10$, establishing a link between ergodicity and a high coherence. Indeed, this measure is also able to indicate, e.g., when a moderate amount of disorder ($0.15 < W < 1.5$) brings the system into the ergodic regime at low h_{mean}, pulling it away from the SG phase. Coherence seems to indicate a further distinction within the ergodic area that is not visible in Fig. 2.8: the area around the SG phase presents a somewhat higher average coherence than the one with high h_{mean}, which presents a very stable average value ($h_{mean} \gtrsim 0.1$ and $W \lesssim 2$). This plateau is more visible in Fig. 3.2, where it appears as the square where the error is highest. We will refer to this area as the "deep ergodic" region, and it corresponds to what we had referred to as the paramagnetic regime in the previous chapter. The rest of the ergodic area, corresponding to the one where coherence is highest, we will identify as the intermediate region. We will further invertigate the differences between both regimes from the reservoir perspective throughout this chapter. For the sake of a more direct comparison between the information provided by Fig. 2.8 and the system's coherence, we present in Fig. 3.3 another similar heatmap. Despite having considerably less resolution, we see that the quantitative value of the parameters at the boundaries of the high-coherence area roughly coincide with those of the ergodic region in Fig. 2.8.

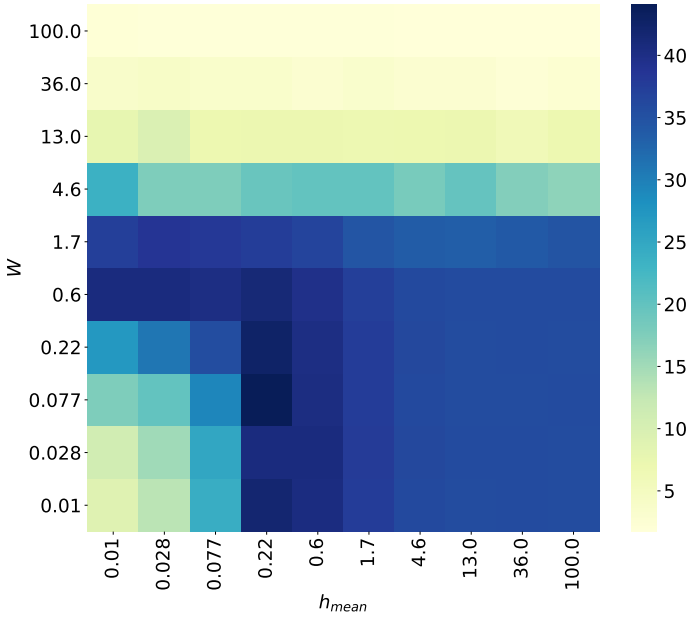
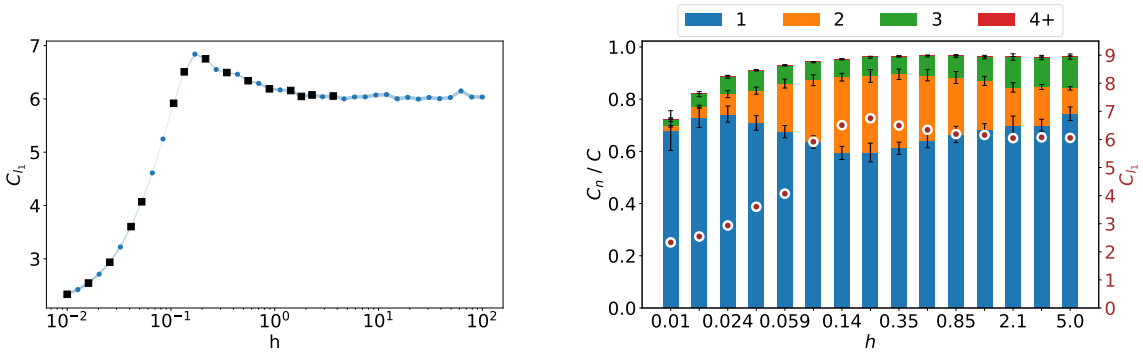


Figure 3.3: Average stationary coherence $\langle C_{l_1} \rangle$ for a system of $N = 10$. Despite the reduced set of points calculated here, the structure present in Fig. 3.1 is still observed.

We now move on to establishing the relation between coherence and information processing capacity of the system. For this purpose, we will examine some sections of the heatmap in Fig. 3.1 and compare them with the evolution of the IPC along those lines.
We begin by looking at the scenario presented by the Hamiltonian (2.1), where $W = 0$, in Fig. 3.4. A couple of features are to be highlighted here. First, we observe that the maximum capacity ($C/N = 1$) is reached just as we enter the high coherence area (which we have already linked to the ergodic regime), where $C_{l_1} > 6$, and corresponding to $h \simeq 0.1$. Thus, we are able to equate a number of features that characterise the system from different perspectives: high coherence, ergodicity and a saturated Information Processing Capacity. We will see that this correspondence remains unchanged throughout the analysis carried out in this section. We also observe that the maximum of the coherence coincides with the minimum of linear capacity and thus with a maximum in the nonlinear processing capacity, as we had predicted at the beginning of Chapter 2. However, we notice that the difference of the linear capacity with respect to the deep ergodic phase (a representative of which is the last bar in Fig. 3.4b) remains at a 10% over the total capacity. That is, we go from a linear capacity of 0.6 at the minimum to 0.7 in the deep ergodic phase. We will always refer to percentages in this sense in this discussion. On the other hand, the third-degree capacity is also incremented by an approximate 10%, so the most striking difference in the second-degree capacity, which loses a 20% over the total capacity. A study performed in larger sizes confirmed this trend for these cases as well.

We will now look at a fixed h_{mean} in order to observe the variation in the processing capacity induced by disorder. We choose $h_{mean} = 0.002$, which lies outside the charted heatmap, in order to probe the deep SG phase as well. We will also consider a larger system size in this case, $N = 6$, with the aim of attaining a better resolution, since in the low-h_{mean} regime the



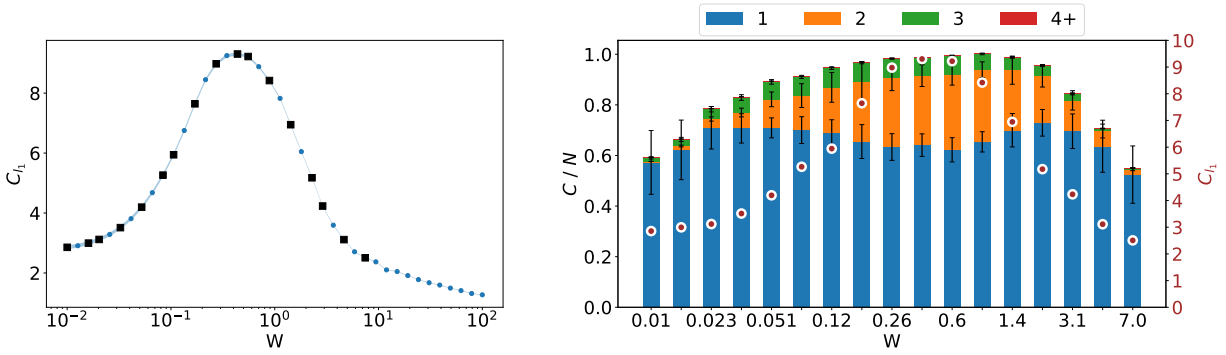
(a) (b)

Figure 3.4: Measured coherence (a) and IPC (b) for a system of $N = 5$ and $W = 0$. Results in (a) are averaged over 100 systems and systems, and over $L = 1000$ measures taken after $\zeta = 1000$ steps. The black squares correspond to the approximate location of the points where the IPC in (b) was evaluated, and the shaded area represents the associated uncertainty (negligible in this case, appearing just as a slightly thicker line). The maximum coherence is reached at $h = 0.22$, which corresponds to the minimum of the linear capacity in (b). In (b), results were averaged over 100 different systems. The full capacity is reached at around $h = 0.1$, that is, shortly before the coherence maximum, and the brown dots illustrate the corresponding coherence at that point, quantified in the right axis.

system exhibits more fluctuations and larger sizes always provide a lower error. The results, presented in Fig. 3.5, show the correspondence between a saturated IPC and high coherence as well, since the region of full processing capacity ($0.18 \lesssim W \lesssim 1.4$) coincides with that of $C_{l_1} \gtrsim 6$. We note that this reference value is not absolute, it only coincides with what we observed in the case $W = 0$, $N = 5$ because these system sizes are close enough. The higher value of the maximum coherence in Fig. 3.5a with respect to that in Fig. 3.4a is only due to the different system sizes: it has been checked that otherwise they would be approximately the same. This is consistent with the fact that they are both situated in the intermediate regime.

As a side note, we comment that the maximum coherence of each heatmap section, as well as the coherence of the deep ergodic region, are not simply proportional to the maximum coherence achievable at that system size. In fact, the fraction of coherence relative to $C_{max}(N)$ (see Eq. (1.13)) is exponentially decreased as N increases, as it is shown in Fig. 3.6. This already points to the fact that the saturation of the IPC of the system may be sustained even with a lower coherence than what the system actually exhibits, an issue we will get back to when we analyse the robustness of the reservoir in the presence of noise in Section 3.2.

Finally, we will further investigate the differences between the intermediate and the paramagnetic regimes with disorder by analysing the transverse cut at $W = 0.5$, again in a system with $N = 5$ spins. The results, presented in Fig. 3.7, highlight the fact that this distinction has no significant consequence on the IPC in the presence of disorder. This is consistent with what was observed in Fig. 2.6, which evidences how the input map drives the two regimes into virtually the same dynamics. This similarity is also reflected in the little variation of the system's coherence for the different h_{mean}, which remains between 5.7 and 6.3.
We also highlight the difference between the IPC for the higher values of the magnetic field



(a) (b)

Figure 3.5: Measured coherence (a) and IPC (b) for a system of $N = 6$ and $h = 0.002$. Results in (a) are averaged over 100 systems, and over $L = 1000$ measures taken after $\zeta = 1000$ steps (and $\zeta = 10^4$for the first 10 points, in order to overcome the slow dynamics of the SG regime). The black squares correspond to the approximate location of the points where the IPC in (b) was evaluated, and the shaded area represents the associated error (negligible in this case, appearing just as a slightly thicker line). The maximum coherence is reached at $W = 0.43$, which corresponds to the minimum of the linear capacity in (b) within error bars. In (b), results were averaged over 100 different systems, and the brown dots illustrate the corresponding coherence at that point, quantified in the right axis. The full capacity is reached at around $0.18 \lesssim W \lesssim 1.4$.

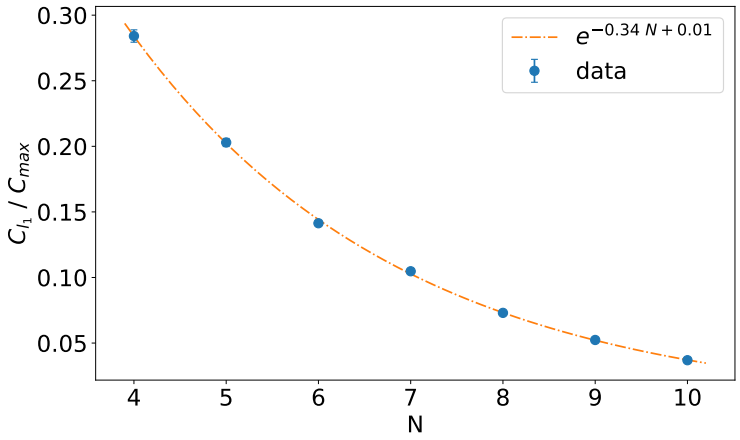
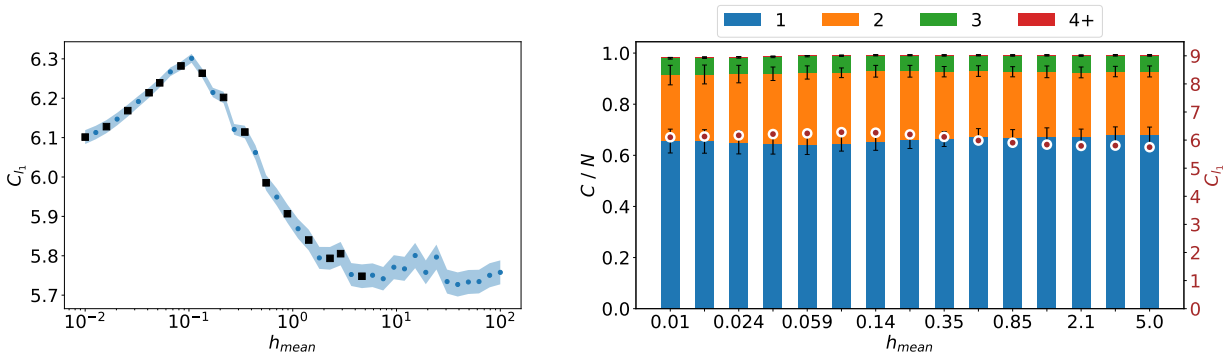


Figure 3.6: Fraction of coherence over the maximally coherent state at the point $h = 1$, $W = 0$. Results are averaged over 100 stationary measurements after $\zeta = 1000$ steps and over 10 different systems.

shown in Fig. 3.7b and those for the case $W = 0$, in Fig. 3.4b. We see that in the absence of disorder the contribution to the IPC coming from higher-order degrees ($n > 2$) is more important than in the $W = 0.5$ case. This may be explained in the semiclassical picture, where a uniform magnetic field concentrates the dynamics of the spins in planes parallel to each other, thus easing the flow of information between them and producing higher degrees of nonlinearity. However, disorder disrupts this effect. In fact, in the disordered case the system's processing capacity seems more similar to that of the intermediate regime for $W = 0$, where the second-degree nonlinearities are enhanced as a whole.



(a) (b)

Figure 3.7: Measured coherence (a) and IPC (b) for a system of $N = 5$ and $W = 0.5$, averaging over 100 systems in both cases. Results in (a) are also averaged over $L = 1000$ measures taken after $\zeta = 1000$ steps. The black squares correspond to the approximate location of the points where the IPC in (b) was evaluated, and the shaded area represents the associated error. The maximum coherence is reached at $h_{mean} = 0.11$. In (b), results were averaged over 100 systems, and the brown dots illustrate the corresponding coherence at that point, quantified in the right axis.

3.2 Robustness to noise

In practice, any system is inevitably subject to some degree of noise, so it is of great interest to study how well the system will perform under non-ideal conditions. To this end, in this section we analyse the robustness of our reservoir through the implementation of a dissipative process between measurements. The particular transformations we have chosen to implement are the phase flip and bit flip channels [34, 35], which are standard ways of simulating decoherence. Such noise acting on a single spin amounts to performing the following transformation:

$$\text{Phase flip: } \mathcal{E}_z[\rho] = (1-p)\rho + p\ \sigma_z\rho\sigma_z \tag{3.1}$$

$$\text{Bit flip: } \mathcal{E}_x[\rho] = (1-p)\rho + p\ \sigma_x\rho\sigma_x \tag{3.2}$$

where p is the error probability. This p is given by a certain noise strength γ and the duration of the time interval during which the process may take place, δt, as follows:

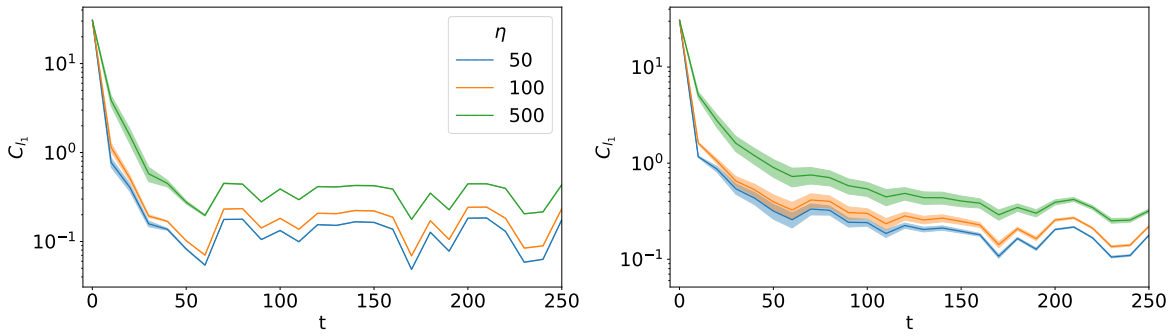
$$p = p_{err} = \frac{1 - e^{-2\gamma\delta t}}{2} \tag{3.3}$$

The parametrisation defined in Eq. (3.3) is a standard choice, since after the phase flip noise acts for a time δt the off-diagonal elements of the density matrix are reduced by a factor $e^{-\gamma\delta t}$ (the diagonal remains intact). A deeper characterisation of the action of both channels can be found in Appendix C, but we can already expect for their effects to be different: the magnetic field concentrates the dynamics around the z-direction, so we expect the bit flip noise, which projects the system onto the x-axis, to erase more information than the projection over the z-axis caused by phase flip.
The generalisation of this process to a system with more than one spin is not straightforward, since many different possibilities arise, but we have chosen to consider that dissipation happens independently in all qubits, following the scarce literature on the subject [35]. In this protocol, we implement η dephasing steps between inputs, where on each dephasing step we first let the system evolve unitarily for a time $\delta t = \Delta t/\eta$ and then apply the dissipative map:

$$\begin{aligned}\mathcal{E}_\alpha[\rho] &= (1-p)^N\rho + (1-p)^{N-1}p\sum_i \sigma_\alpha^i\rho\sigma_\alpha^i \\ &+ (1-p)^{N-2}p^2\sum_{i\neq j}\sigma_\alpha^i\sigma_\alpha^j\rho\sigma_\alpha^i\sigma_\alpha^j + ... + p^N\sigma_\alpha^1...\sigma_\alpha^N\rho\sigma_\alpha^1...\sigma_\alpha^N\end{aligned} \tag{3.4}$$

where $\alpha = \{z, x\}$ and $i, j = 1, ..., N$, so that σ_α^i corresponds to the Pauli matrix of coordinate α acting on spin i. It is only after the series of dephasing steps have taken place, that is, right before feeding new input, that we perform the measurement.

To begin the study of the introduction of such noise, we first observe the effect the number of dephasing steps we introduce between measurements, presented on Fig. 3.8. The results from 3.8 are explained by the fact that applying the decohering operation for smaller time intervals results in a lower error probability, and thus an effectively weaker effect. Notice the dependence of the time it takes the system to reach a stationary regime on the effective strength of the noise (higher p_{err} reach the stationary state faster) and on the dynamical



(a) $W = 0$ (b) $W = 3$

Figure 3.8: Measured coherence for a system of parameters $h_{mean} = 1$ and the specified disorder W subjected to phase flip noise of strength $\gamma = 0.9$ and different number of dephasing steps, always starting from the maximally coherent state. Results are averaged over 10 different systems subjected to the same input sequence for each η. We highlight the fact that δt, the time between decohering events, is changed for the different values of η as $\delta t = \Delta t/\eta = 10/\eta$, which also implies a lower p_{err}. In particular, $p_{err} = 0.15,\ 0.08,\ 0.02$ for $\eta = 50,\ 100,\ 500$.

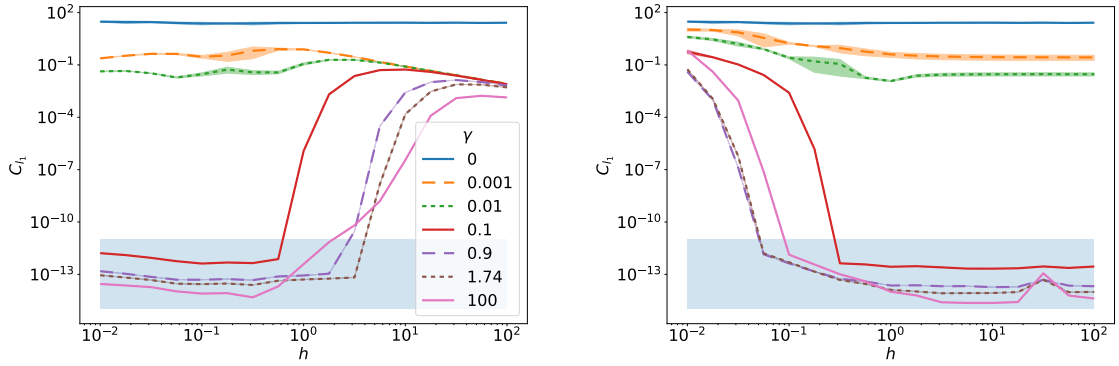
regime of the system. This last feature stands out comparing Figs. 3.8a and 3.8b, where we see that the point situated in the ergodic regime ($W = 0$, $h_{mean} = 1$) reaches the stationary regime considerably faster than its MBL counterpart ($W = 3$, $h_{mean} = 1$), consistently with the dynamical properties of this phase. Throughout the rest of this analysis we will fix the number of dephasing steps as $\eta = 50$, which implies $\delta t = 0.2$. Moreover, we will always start from a maximally coherent initial condition, in order to better appreciate the system's tendency to lose it as an effect of noise.
Before proceeding to the rest of the study we comment that, for the parameters we have already fixed ($\Delta t = 10$, $\eta = 50$) and the number of measurements we usually average over ($L = 100 - 1000$), we will only see some effect of the noise for $\gamma \gtrsim 10^{-3}$, so we will focus on such noise strengths. The particular set of γ to be studied and their corresponding error probabilities are presented in Table 3.1.

γ	$1{\cdot}10^{-3}$	$2{\cdot}10^{-3}$	0.01	0.1	0.9	1.74	100
p_{err}	$2{\cdot}10^{-4}$	$4{\cdot}10^{-4}$	$2{\cdot}10^{-3}$	0.02	0.15	0.25	0.5

Table 3.1: Studied noise intensities γ and their corresponding error probabilities p_{err} according to (3.3).

It will be useful to study the effect the different channels have on the system in the absence of the input map first (Fig. 3.9). This way, we may take note of the action of both kinds of noise by themselves, without the dissipation caused by the repeated injection of input and the partial trace. We will do so through the preferred measure of this work, C_{l_1}. In order to observe an overall change in the system through the probing of coherence, we will scan the whole spectrum of h we have been considering with a fixed $W = 0$. We have focused the gross part of this analysis on this simplest case scenario of (2.1) in order to better understand the effects of dissipation. However, some interesting remarks will be made about a case with

disorder at the end of this section.



(a) Phase flip

(b) Bit flip

Figure 3.9: Coherence for systems with $W = 0$ without the injection of input at different noise intensities. Measurements were taken after $\zeta = 1000$ steps and averaged over 8 different systems, which were initiated at the maximally coherent state. The shaded area on the bottom of the plots indicates the region where the obtained results must be considered with care, as they may be affected by numerical precision.

Fig. 3.9 evidences some important differences between decoherence channels in the regimes present for $W = 0$: the SG phase for low h and the ergodic phase for $h \gtrsim 0.1$, within which we had the intermediate region ($0.1 \lesssim h \lesssim 0.3$) and the deep ergodic or paramagnetic region ($h \gtrsim 0.3$). Nonetheless, we will merely focus on the latter, as it is the broader region of interest for QRC. We notice that both channels begin having a strong effect starting from $\gamma = 0.1$, but the bit flip channel shows a higher disruptive power in the ergodic regime since coherence remains essentially static at a low value in this region. For the phase flip case, however, coherence builds up for increasing h, even making the effect of noise negligible for high enough h. It is also interesting to note the displacement of the coherence curves produced by the various phase noise intensities in 3.9, which suggests that noise is distorting the threshold of the ergodic phase. This effect is particularly important in the phase flip case (see 3.9a) where, for increasing γ, it takes the system an increasingly higher magnetic field to reach the asymptotic stationary coherence set by lower noise intensities.

Once we have characterised some main features of the noise channels under consideration, we may move on to the central subject of interest: the robustness of the reservoir to the action of such noise. We will first examine the coherences present in the system, in an analog picture to the one in Fig. 3.4a for the set of noise strengths under consideration.
The results of Fig. 3.10 suggest that bit flip noise noise starts being truly disruptive at around $\gamma \sim 0.1$, where we find evidence that the deep ergodic phase may be close to its erradication in the fact that the coherence is almost the same as it was in the SG phase, where the system was not ergodic. Actually, for the highest noise intensities we see that it is even lower, nearly dropping to zero, suggesting a breakdown of ergodicity.
The consequences of the bit flip channel, however, seem to be less dramatic even up to $p_{err} = 0.5$, as the plateau we associated to the deep ergodic phase in Section 3.1 appears in any case. In fact, it even keeps the structure of a peak at the transition between phases we

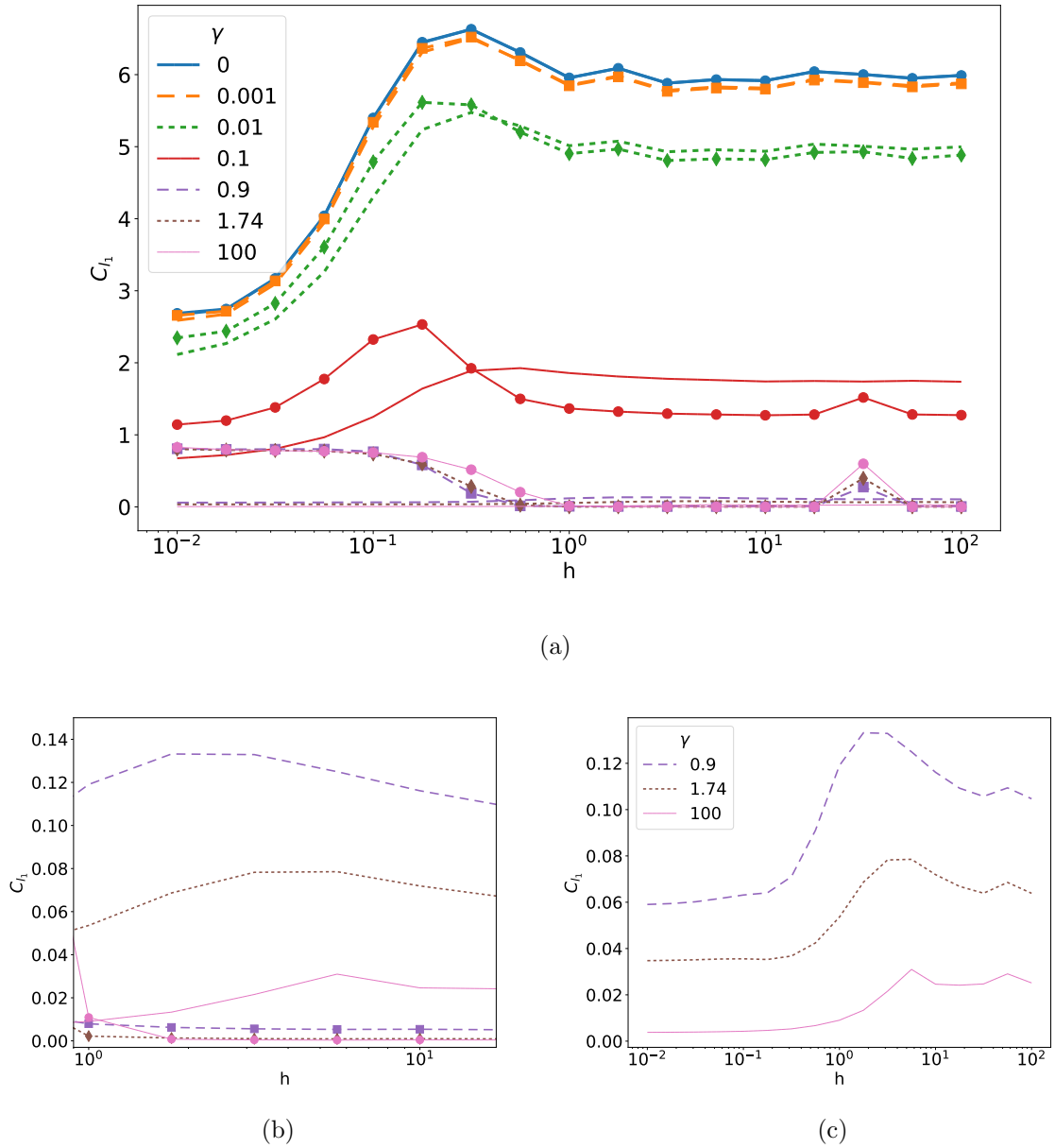
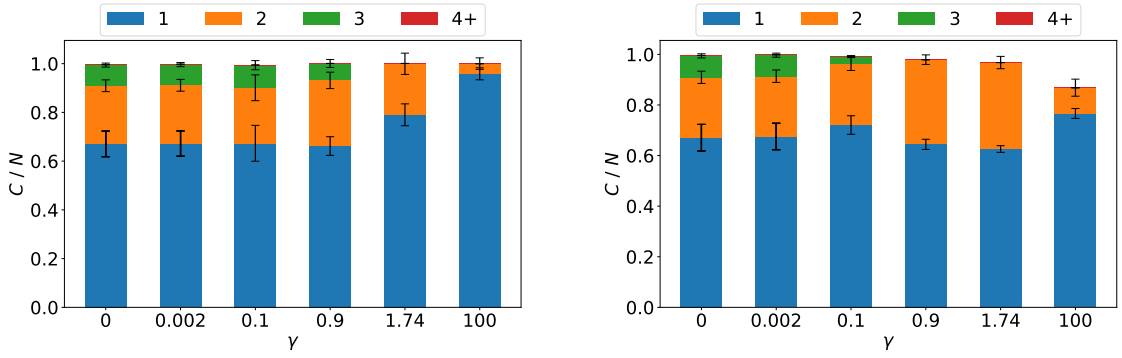


Figure 3.10: Stationary coherence of the reservoir computer as a function of h with $W = 0$ for different noise intensities. The lines without a marker correspond to phase flip noise and the ones with markers to bit flip (both cases are indistinguishable for $\gamma = 0, 0.001$). Results are averaged over 1000 measurements taken after $\zeta = 1000$ steps and over 8 different systems. (a) presents all the obtained curves, while (b) and (c) show close-ups of the beginning of the ergodic phase and the whole phase flip curve for high noise intensities, respectively. Error bars have been left out of these plots for the sake of clarity, but we comment that they are rather small: the relative error ranges between 3% for low γ and 0.3% for high γ.

had previously linked to the intermediate phase, visible in Fig. 3.10c.
Nonetheless, some differences do appear as well for $\gamma \gtrsim 0.1$: fluctuations are considerably flattened out (we remind the reader that the relative error for this noise intensity is around 0.3%, an order of magnitude lower than it was for lower γ) and the plateau develops at increasing h as we turn up the noise, an effect we had already seen for the unitary case in Fig. 3.9a. To know more about how these features affect the nature of our reservoir, however, we must directly look into the system's IPC. The high computational cost of simulating a system with decoherence and the complexity of the different dynamics of each possible regime force us to focus our understanding efforts in selected points. We have mainly focused on the ergodic regime, where the system may act as a reservoir, but we will also present an analysis of a point in the MBL region.

3.2.1 Ergodic region

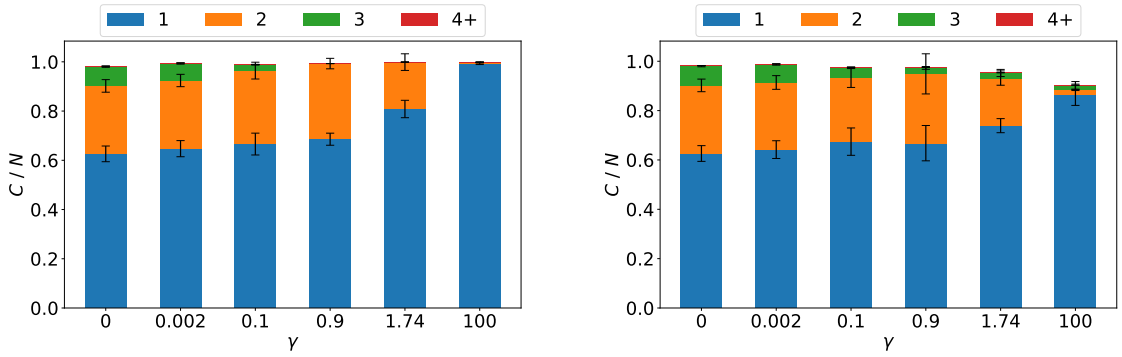
As representatives of the computationally adept regime, we have chosen to study the points $h = 1$, $W = 0$, located in the deep ergodic phase but not too far away from the intermediate region, and $h = 0.02$, $W = 0.3$, located in the intermediate region. However, since the coherence analysis we have presented corresponds to the case without disorder, we will only comment on the first point when extracting conclusions in relation to this magnitude.



(a) Phase flip

(b) Bit flip

Figure 3.11: IPC at $W = 0$, $h = 1$ for a system subjected to increasing noise strengths. Results are averaged over 10 different systems. In (b), the bars relative to $\gamma = 0.9,\ 1.74$ are slightly underfilled, with $C/N = 0.98,\ 0.95$ when a filled bar is considered for $C/N > 0.99$.



(a) Phase flip

(b) Bit flip

Figure 3.12: IPC at $W = 0.3$, $h = 0.02$ for a system subjected to increasing noise strengths. Results are averaged over 10 different systems. In (b), the first bar that fails to saturate corresponds to $\gamma = 1.74$, although the high error of $\gamma = 0.9$ must be taken into account. This elevated error indicates that this point is most likely close to the transition to an unsaturated capacity of the reservoir.

What we observed in Figs. 3.10a, 3.10c together with the results of Fig. 3.11a (and Fig. 3.12a) help us get a clearer picture of the robustness of a good performance of reservoir against the phase flip channel. This type of noise seems to attack the correlations that make up the non-linear processing capacity of the system, leaving the most essential pathways for the flow of information along the system untouched. The story told by Figs. 3.10a and 3.11b, however, is rather different: for $\gamma \geq 0.9$, the IPC fails to saturate and the reservoir ceases to be operational. In order to dive deeper into this matter, we will make a further comparison between the state of the system and its factorised analog ϱ, defined as

$$\varrho = \mathrm{Tr}_{\{i\}_{i\neq 1}}\rho \otimes ... \otimes \mathrm{Tr}_{\{i\}_{i\neq N}}\rho = \bigotimes_{i=1}^{N} \rho_i \tag{3.5}$$

This will allow us to observe how far the system is from being completely factorised in each case.
For this comparison we will make use of the mutual information measure $\mathcal{I}$, which measures the distance between the true state of the system ρ and its factorised counterpart ϱ.

$$\mathcal{I}(\rho) = S(\varrho) - S(\rho) = \sum_{i=1}^{N} S(\rho_i) - S(\rho) \tag{3.6}$$

where $S(\rho)$ refers to the von Neumann entropy of state ρ. Notice that $\mathcal{I}(\rho) = 0$ when $\rho = \varrho$, in which case noise has completely cut off the 'communication' between spins. We have chosen this measure instead of a coherence-based one, as could be the normalised coherence difference between both states $[C_{l_1}(\rho) - C_{l_1}(\varrho)]/C_{max}$ because the mutual information is built from the truly fundamental structure of the state, thus making it much more general.
Fig. 3.13 confirms that the bit flip cases $\gamma = 0.9, 1.74$ and 100, for which the reservoir's proper operation is lost according to Fig. 3.11b, are completely factorised. It also shows that a very few correlations between spins need to be preserved for the reservoir to retain its

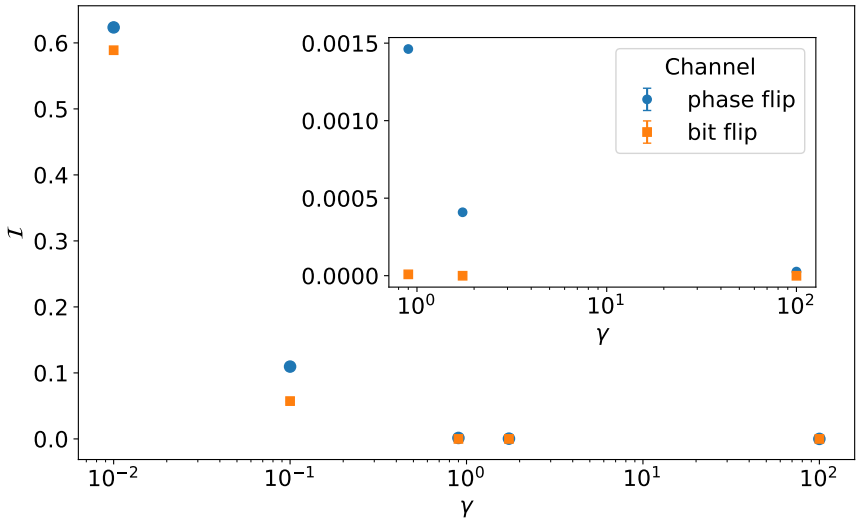


Figure 3.13: Mutual information for $W = 0$, $h = 1$ resulting from the average of $L = 100$ measurements taken after $\zeta = 1000$ steps and averaged over 10 different systems. The inset shows a close-up of the last three points, for which the bit flip channel scores an absolute 0 within numerical precision.

fading memory. For example, the phase flip point $\gamma = 100$ presents a mutual information of the order of 10^{-5}. In broader terms, Fig. 3.13 manifests that both channels have a tendency to factorise the state of the system. This stems from the fact that both types of noise are local and affect spins independently, which takes its toll on the non-linear capacity of the system, as shown in Figs. 3.11 and 3.12 as a whole. The case of the second-degree capacity at $\gamma = 0.9$ and 1.74 in 3.11b seems to escape this tendency, an issue whose explanation requires further research, but the capacity relative to higher orders does abide by this tendency here as well.

In order to give a consistent interpretation of all the results we have revised in this section, we must recall the physical effect of the decoherence channels under study. The bit flip channel corresponds to the most familiar form of dissipation, since the bit flip error exists in classical systems as well. As we anticipated earlier, this channel was bound to erase information more effectively since the magnetic field concentrates the dynamics along the z-direction, while the bit flip operation projects spins into a perpendicular direction. On the other hand, the phase flip channel is exclusive to qubits since it erases the relative phase between our basis states, an inherently quantum magnitude. It does so by means of a projection onto the z-axis, so in this case most of the information encoded in local populations will remain intact. Indeed, from the results shown up to this point it remains clear that bit flip noise is much more destructive for the reservoir than phase flip. Some further, more specific evidence for this may be provided in the case of quantum correlations through an analysis of the concurrence, a measure of entanglement that, consequently, only picks up on quantum correlations [36]. The definition of the concurrence is presented below:

$$\mathcal{C}(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} \tag{3.7}$$

where the λ_i are the eigenvalues of the operator $R = \sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}}$ in decreasing order, ρ is the density matrix of a two-qubit state and $\tilde{\rho}$ is its bit-flipped counterpart, $\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y)$. This measure is bounded between 0 and 1, with the former corresponding to a product state and the latter to a maximally entangled one (i.e., a Bell state).
Since concurrence is defined for a two-qubit state, each measurement is the arithmetic average of the concurrences of all possible combinations of spins, where the two-spin matrices are obtained tracing over the rest of spins. Unfortunately, concurrence provides a very small value for low entanglement, so much so that high noise intensities become unreachable because they go beyond our numerical precision. For this reason, we merely present the obtained values for the highest γ within the ones relevant to the robustness analysis that we were able to access through this measure ($\gamma = 0.1$) for both channels, as well as the concurrence for the noiseless case. These results are provided in Table 3.2, where we see that this noise intensity has already eliminated the quantum correlations we could have observed in the case of bit flip, leading to the conclusion that this channel erases quantum correlations more efficiently than the dephasing operation. This can be related to the faster disappearance of the n-degree capacity for $n > 2$ in Fig. 3.11b than in 3.11a, suggesting a connection between these higher nonlinearities and quantum correlations. There is some further evidence supporting that last statement in the study of the IPC of the previous section (see Fig. 3.4b in particular), where the third-order nonlinear capacity was highest in the paramagnetic regime. As the presence of the transverse magnetic field is what brings out the quantum character of the Ising model, quantum correlations are expected to be stronger in the high-h regime, where they induce the resonances we briefly mentioned in the previous section. Fig. 3.14 illustrates this by presenting the concurrence in the absence of decoherence for the case $W = 0$.

γ , noise channel	0 , -	0.1, phase flip	0.1 , bit flip
$\mathcal{C}$	$(2.57 \pm 0.02) \cdot 10^{-4}$	$(5.85 \pm 0.02) \cdot 10^{-5}$	$(1.450 \pm 0.004) \cdot 10^{-5}$

Table 3.2: Resulting coherence for $W = 0$, $h = 1$ after an average over $L = 100$ measurements taken after $\zeta = 1000$ steps and averaged over 10 different systems.

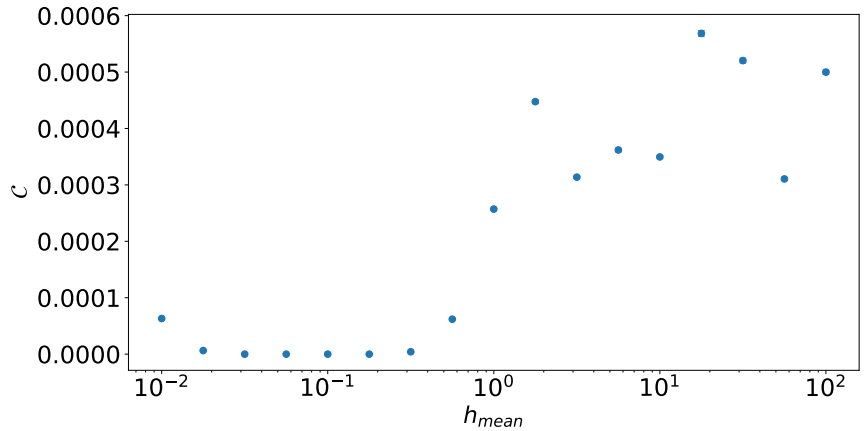


Figure 3.14: Concurrence for $W = 0$ without any kind of noise, resulting from the average of $L = 100$ measurements taken after $\zeta = 1000$ steps and further averaged over 10 different systems. Error bars are too small to be visible in this plot.

3.2.2 MBL regime

As a last remark, we highlight the interest a certain amount of noise may have in order to push the reservoir into its operational regime rather than out of it. A system originally standing in the MBL phase is a suitable candidate for the process we have just mentioned: we recall that the lack of ergodicity in this dynamical phase comes from the strong quantum correlations locking spins into a robust, many-body state that is unable to forget its initial condition and to process each input non-locally in the reservoir. Fig. 3.15 shows that the MBL phase arising for large W does, indeed, present considerably more correlations of quantum origin, even though entanglement appears to diminish for increasing W deeper in the MBL phase. The effect of noise is displayed in Fig. 3.16, where we examine the robustness of the IPC for a system with $W = 3$ and $h_{mean} = 1$, which is located in the MBL phase but not too far from the ergodic frontier.

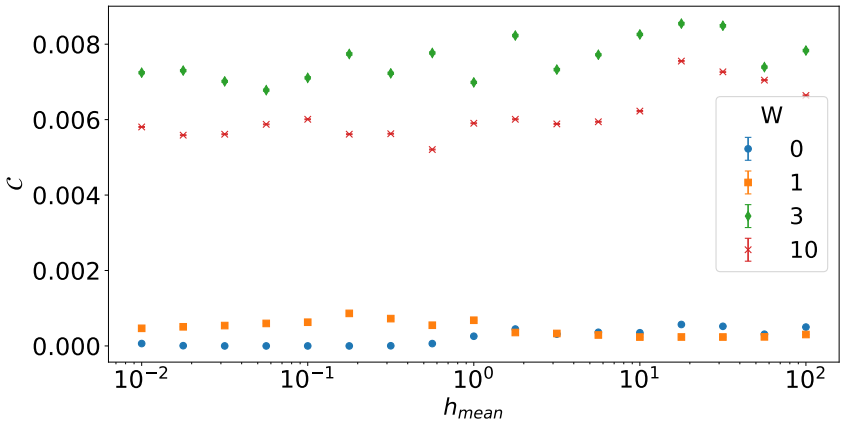
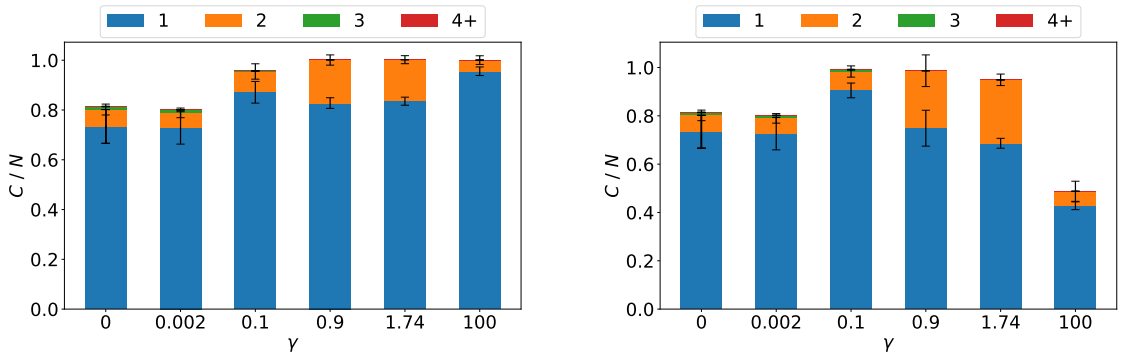


Figure 3.15: Concurrence without any kind of noise, resulting from the average of $L = 100$ measurements taken after $\zeta = 1000$ steps and further averaged over 10 different systems. Error bars are too small to be visible in this plot.

The results of Figs. 3.16 and 3.17 are consistent with a mechanism where the destruction of correlations appears together with an improvement of the IPC in the case of phase flip, up to the point where it is fully restored for high noise intensities. This happens for the bit flip channel as well, but only for a small window before the system suffers too much damage. We also see that the irruption into the operational phase happens sooner in the bit flip case ($\gamma \approx 0.1$) than for phase flip ($\gamma \approx 0.9$) because of the stronger disruptive effect of the former. A way to preserve some degree of nonlinear processing for these reservoirs (greater than second-degree) remains a subject for future research, as well as a deeper analysis of the role of quantum correlations in this phase.



(a) Phase flip

(b) Bit flip

Figure 3.16: IPC at $W = 3$, $h = 1$ for a system subjected to increasing noise strengths. Results are averaged over 10 different systems. In (a), the first bar to be completely filled is that with $\gamma = 0.9$, while in (b) the only full bar corresponds to $\gamma = 0.1$ (although $\gamma = 0.9$ is filled within error).

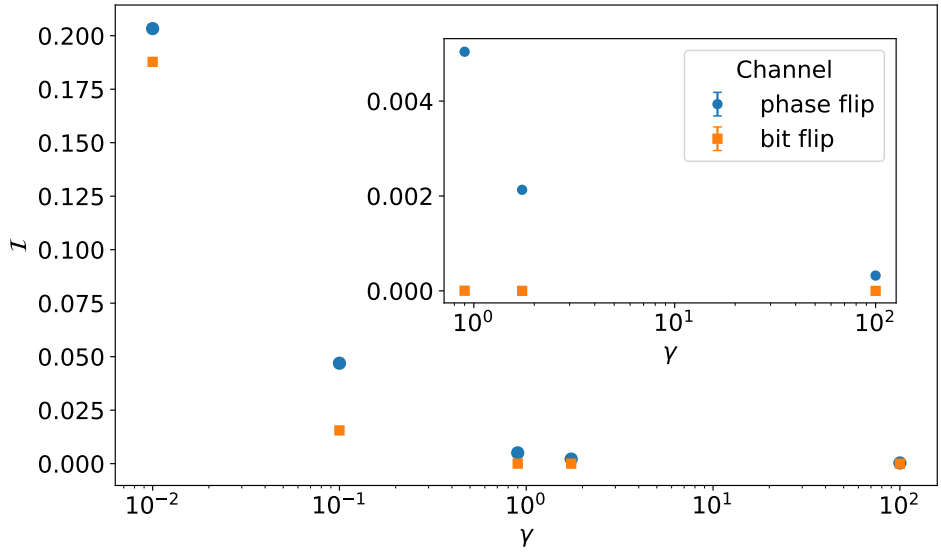


Figure 3.17: Mutual information for $W = 3$, $h = 1$ resulting from the average of $L = 100$ measurements taken after $\zeta = 1000$ steps and averaged over 10 different systems. The inset shows a close-up of the last three points, for which the bit flip channel scores an absolute 0 within numerical precision.

Chapter 4

Conclusions and outlook

In this work we have elaborated the path set out by [16], where the transverse-field Ising model was proposed as a reservoir, and continued in [29], where its performance was examined from the point of view of the IPC. We have extended this analysis to the case of a magnetic field with disorder and explored the reservoir's capacity in the different dynamical regimes that arise, which were already discussed in [21] in the context of dynamical phase transitions. Original results are presented in the context of coherence, correlations and the effects of noisy channels.

Our study has revealed a strong correspondence between the ability of a system to build up coherence and its respective processing capacity. The obtained results in this respect show that, starting from an initially incoherent state, the systems that reach a high stationary coherence are those that exhibit a proper performance as reservoirs. Furthermore, these have also been linked to an ergodic regime of the system. We stress that this correspondence does not have to do with the absolute value of the coherence, but rather with the absolute value of the change in coherence from the initial state up to the stationary one. A question that remains open for future research is whether we would obtain a similar phase diagram if we looked at the coherence along the x-axis, i.e., whether we would be able to identify the ergodic regions through the change in coherence as well by looking at the system from this direction. In any case, we also report that no identifiable threshold over which the reservoir holds the fading memory property has been found, so the notion of "large" or "small" change depends on the dynamical regime and the presence of noise. It has been shown that the reservoir may function properly with very little coherence, as it is the case for strong phase flip noise, but the results relative to bit flip link the complete factorisation of the state to the loss of a fading memory. Our results also show that bit flip noise is more disruptive than phase flip for the system's ability to process information. We mainly attribute this asymmetry to the presence of the magnetic field, which favours the z-direction in space, but it may be interesting to see how much the system's robustness is improved (if at all) for the bit flip channel when we introduce the input through the first spin's coherences rather than through a diagonal state (e.g., though a mixed input state as the one we have used in this work of the eigenstates of σ^x rather than those of σ^z). It would also be interesting to address the effect of noise in the IPC when our measured observable is non-local. We expect the decoherence process analysed here to be more disruptive in this case, and it may provide some more insight into how information stops flowing through the network.

As for our particular reservoir under study, the transverse-field Ising model has been found

to exhibit a good robustness against noise. In fact, noise may even be an instrument to turn our spin network into an operational reservoir (when it is set in the MBL regime), turning its disruptive power into an asset. A way to make a more exhaustive analysis of the constructive effect of noise would be to obtain another heatmap showing the system's different dynamical regimes under the presence of relatively high noise intensities. It would be very expensive computationally to perform this study through coherence, as we have done throughout this work, but a good alternative would be to analyse the scaling of entanglement entropy with system size. Since in the MBL and ergodic phases this property scales according to area and volume, respectively [26, 27, 28, 37], this would allow us to detect a possible displacement of this frontier with respect to the noiseless case. The issue of restoring the system's nonlinear processing capacity, which is undermined by decoherence, is a subject for future research as well. However, the methodologies of space and time multiplexing ([18] and [16, 29] respectively) present promising approaches to this problem. Overall, our analysis reinforces the transverse-field Ising model's position as a promising candidate for QRC.

Ultimately, our results contribute to the hopeful perspectives QRC has to offer. A similar study to the one presented here considering quantum input instead of a classical one could pave the way towards quantum RC processors which would be truly useful for real experiments. For example, a very appealing task for these auxiliary quantum processors would be the simulation of other quantum systems that are highly expensive or complex to realise experimentally inside a larger, easily controllable reservoir. Moreover, thanks to the structure of RC it would suffice to have a single such reservoir and a collection of trained readouts for each of the systems one wishes to simulate, a solution that saves up space, resources and the time it would take to build different setups. Combining this line of thought with what has already been achieved in [6], we could even obtain several magnitudes of the simulated system at once, including those we cannot access experimentally but which carry essential information, such as entropies. This possibility alone already alters notably the aspect of a quantum laboratory as well as the scope of the knowledge it can uncover, boosting a faster development of new science and technologies.

Appendix A

Definition of the echo state and fading memory properties

The echo state property

Let u be a fixed, but otherwise arbitrary left infinite input sequence up to a time t, $u[-\inf : t] = (..., u_{t-2}, u_{t-1}, u_t)$, with $t \in \mathbb{Z}$. Let x_t be the state of the dynamical system at time t, which may be described as

$$x_t = H(x_{t-1}, u_t) \tag{A.1}$$

The dynamical system implementing the transformation H is said to have the echo state property if, for any two trajectories $x[-\infty : t]$, $x'[-\infty : t]$ that were initially close to each other we have $x_t = x'_t$. In that case, there exists an input echo function $\mathcal{E}$ such that the current state of the system may be written as

$$x_t = \mathcal{E}(..., u_{t-2}, u_{t-1}, u_t) \tag{A.2}$$

The fading memory property

Consider an input string u and its corresponding output state x according to the dynamical map

$$x_t = \mathcal{R}(u)(t) \tag{A.3}$$

For every $\epsilon > 0$ and input u we may define the ball centred on u of radius ϵ on input space, which is mapped into a ball centred around x of radius $\delta(u, \epsilon)$ by virtue of (A.3). Such dynamical map is said to have the fading memory property if, for two ϵ-close inputs u, v on the interval $t \in [t_0 - T, t_0]$ ($|u(t) - v(t)| < \epsilon$), we have that

$$|\mathcal{R}(u)(t_0) - \mathcal{R}(v)(t_0)| < \delta \tag{A.4}$$

Appendix B

Numerical aspects of IPC

We hereby explain the parameters that must be adjusted for the obtention of the IPC, as well as the ones we have deemed suitable for the presented simulations.

We first comment that, in order to get reasonable approximations to the full capacity, we have always used training and testing strings of $L = 10^5$ elements. Once that parameter is fixed, the most important bound to determine is the threshold to consider a contribution to the IPC valid, since there is a noisy background resulting from numerical precision. In order to identify it, one must first run a simulation attempting to pick up as many contributions as possible, which will result in a picture similar to Fig. B.1.

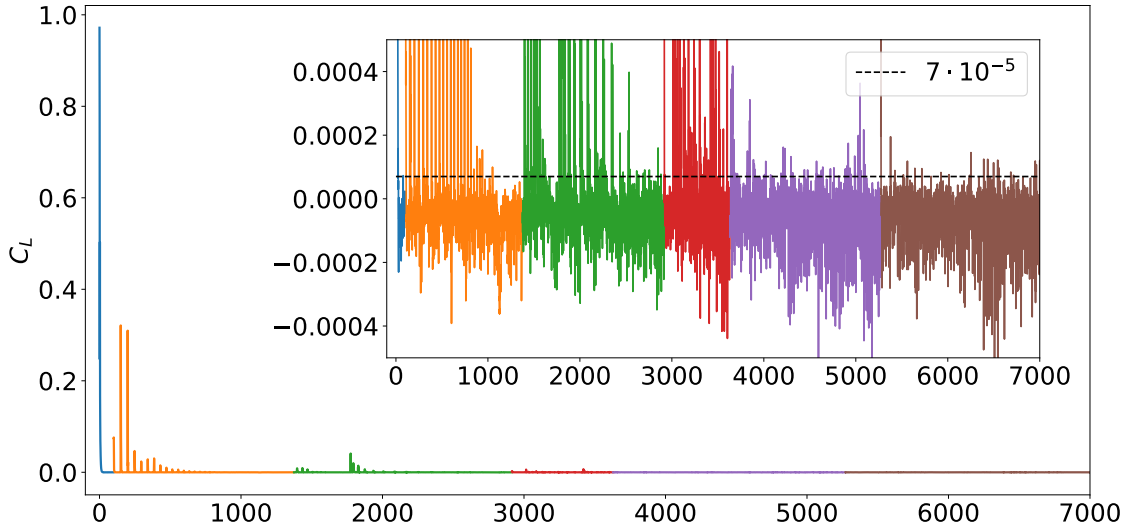


Figure B.1: Example of all the registered contributions to the IPC for a typical system in the ergodic region. The different colours refer to different degrees of the target function. The noisy contributions are visible in the inset of the figure with a suitable threshold.

In our case, the optimal thresholds found vary between $5 \cdot 10^{-5}$ and $5 \cdot 10^{-4}$ depending on the simulated process. The noise depends on the number of observables measured (N)

and the length of the input string ($L = 10^5$), such that it is of the order of $\mathcal{O}(N/L)$ [22]. Moreover, one must take into account that longer simulations such as the ones of Section 3.2 produce a higher accumulated error, which slightly increases this threshold. In any case, it has been observed that being sufficiently close to the optimal threshold suffices in order to have a realistic picture of the IPC. For example, the full normalised capacity may appear as 0.97 instead of 0.997, as it would with the optimal threshold. However, the deviation introduced by the different realisations of the system is still more significant in this case.

We must also have an idea of the maximum delay of the input string we want to consider for each degree, at least in the cases where we have no clue of how far back it may reach, in order to perform the threshold analysis. A good strategy is to look just at the linear capacity for delays that go really far back, and the maximum delay to provide a contribution in higher degrees will always be lower. The meaning of "really far back" depends entirely on the system at hand: it may be 100 or 1000 just as easily, but either way such maximum delays are not a big computational hindrance for the linear capacity.

Following the order in the recollection of contributions described in the Supplementary Material of [22] and taking into account the bounds we have just described, we are guaranteed to obtain the best approximation to the system's IPC allowed by our numerical precision.

Appendix C

Decoherence channels

We hereby give an illustration of the action of the bit flip and phase flip channels, since their effect on the Bloch sphere is useful for the interpretation of some of the results in Section 3.2. For this, we present the effect of this channels on a single qubit described by a density matrix ϱ:

$$\varrho = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix}$$

The phase flip channel

According to Eq. (3.1), the phase flip channel induces a transformation that results in an exponential decrease of the coherence. It can be easily checked that a single application of the dephasing operation to ϱ results in

$$\mathcal{E}[\varrho] = \begin{pmatrix} a_{00} & (1-2p)a_{01} \\ (1-2p)a_{10} & a_{11} \end{pmatrix} \mathcal{E}^N[\varrho] = \begin{pmatrix} a_{00} & (1-2p)^N a_{01} \\ (1-2p)^N a_{10} & a_{11} \end{pmatrix} \tag{C.1}$$

where p is the error probability. As a consequence, if we were to measure the coherence according to definition (1.11), we would find that

$$C_{l_1}(\mathcal{E}[\varrho]) = |1-2p|^N C_{l_1}(\varrho) \tag{C.2}$$

Thus, this noise tends to project the spin towards the z axis, turning an initial Bloch sphere of pure states into an ellipsoid with its longer axis located along coordinate z.

The bit flip channel

The mapping for this channel, defined in Eq. (3.2), yields the following result on ϱ:

$$\mathcal{E}[\varrho] = \begin{pmatrix} (1-p)a_{00} + pa_{11} & (1-p)a_{01} + pa_{10} \\ pa_{01} + (1-p)a_{10} & pa_{00} + (1-p)a_{11} \end{pmatrix} \tag{C.3}$$

$$\mathcal{E}^2[\varrho] = \begin{pmatrix} ... & (1-p)^2 a_{01} + 2p(1-p)a_{10} + p^2 a_{01} \\ (1-p)^2 a_{10} + 2p(1-p)a_{01} + p^2 a_{10} & ... \end{pmatrix} \tag{C.4}$$

One soon realises that the new off-diagonal elements are almost like Newton's binomial, but with a_{01} and a_{10} alternating between the coefficients. The individual coherences are thus hard

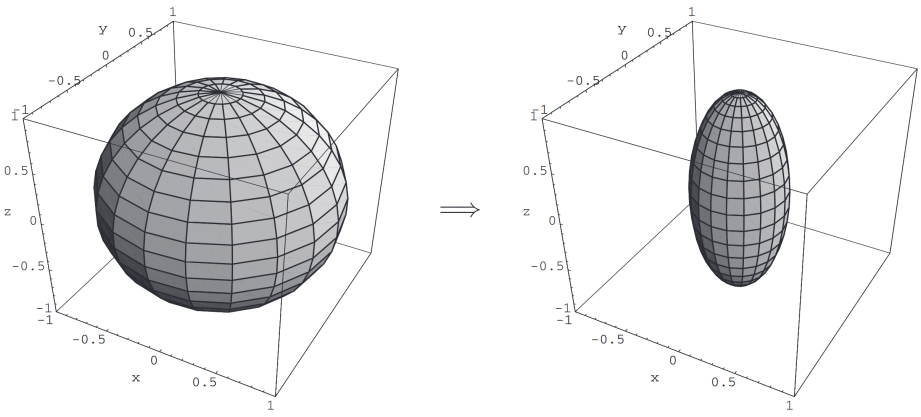


Figure C.1: Phase flip transformation on the Bloch sphere. Figure taken from [34].

to describe analytically, but the associated coherence of ϱ takes a much simpler form:

$$C_{l_1}(\mathcal{E}[\varrho]) = C_{l_1}(\mathcal{E}^2[\varrho]) = C_{l_1}(\mathcal{E}^N[\varrho]) = C_{l_1}(\varrho) \tag{C.5}$$

We thus see that this type of noise does not harm the individual qubit's coherence at all. This observation and the physical interpretation of the way we have extended the action of this noise to the case of a multipartite system (namely, as a Markov process where each qubit may be affected by decoherence independently) leads us to expect the phase flip noise to be more disruptive for the coherence than bit flip in the situations where the density matrix is closer to that of factorised qubit states, i.e., when the flow of information through the system is weak.
On a final note, the visual picture of the bit flip operation is presented below:

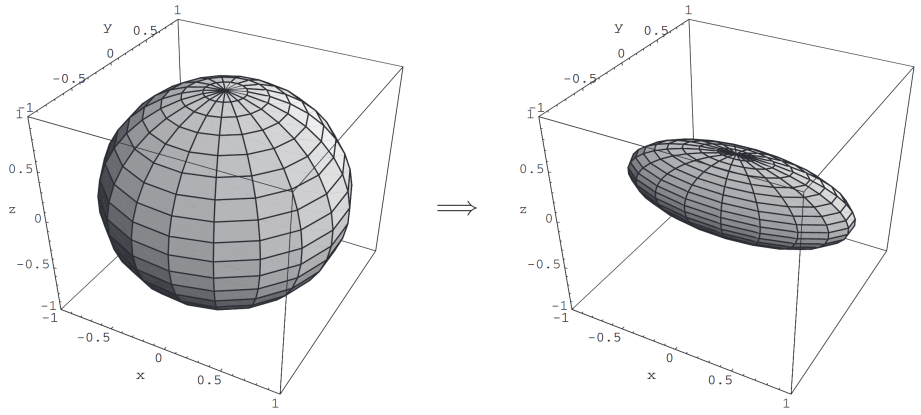


Figure C.2: Bit flip transformation on the Bloch sphere. Figure taken from [34].

Bibliography

[1] Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie Vogt-Maranto, and Lenka Zdeborová. Machine learning and the physical sciences. *Rev. Mod. Phys.*, 91:045002, 2019.

[2] Benjamin Schrauwen, David Verstraeten, and Jan Campenhout. An overview of reservoir computing: Theory, applications and implementations. *Proceedings of the 15th European Sympsosium on Artificial Neural Networks*, pages 471–482, 2007.

[3] Susan Stepney, Steen Rasmussen, and Martyn Amos. *Computational Matter*. Springer Publishing Company, Incorporated, 1st edition, 2018.

[4] Daniel Brunner, Miguel C Soriano, and Guy Van der Sande. *Photonic Reservoir Computing*. De Gruyter, 2019.

[5] Aki Kutvonen, Keisuke Fujii, and Takahiro Sagawa. Optimizing a quantum reservoir computer for time series prediction. *Scientific Reports*, 10(1):14687, 2020.

[6] Sanjib Ghosh, Andrzej Opala, Michał Matuszewski, Tomasz Paterek, and Timothy C. H. Liew. Quantum reservoir processing. *npj Quantum Information*, 5(1):35, 2019.

[7] Danijela Markovic and Julie Grollier. Quantum neuromorphic computing. *arXiv:arXiv:2006.15111*, 2020.

[8] Marvin Minsky and Seymour Papert. *Perceptrons: An Introduction to Computational Geometry*. MIT Press, Cambridge, MA, USA, 1969.

[9] Gouhei Tanaka, Toshiyuki Yamane, Jean Benoit Héroux, Ryosho Nakane, Naoki Kanazawa, Seiji Takeda, Hidetoshi Numata, Daiju Nakano, and Akira Hirose. Recent advances in physical reservoir computing: A review. *Neural Networks*, 115:100 – 123, 2019.

[10] Herbert Jaeger. The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, 148, 2001.

[11] Wolfgang Maass, Thomas Natschläger, and Henry Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural computation*, 14:2531–60, 2002.

[12] Zoran Konkoli. *On Reservoir Computing: From Mathematical Foundations to Unconventional Applications*, volume 22. Springer International Publishing, 2017.

[13] Chrisantha Fernando and Sampsa Sojakka. Pattern recognition in a bucket. *Proc. of ECAL*, 2801:588–597, 2003.

[14] Kohei Nakajima, Helmut Hauser, Tao Li, and Rolf Pfeifer. Information processing via physical soft body. *Scientific reports*, 5:10487, 2015.

[15] Lyudmila Grigoryeva and Juan-Pablo Ortega. Echo state networks are universal. *Neural Networks*, 108:495 – 508, 2018.

[16] Keisuke Fujii and Kohei Nakajima. Harnessing disordered-ensemble quantum dynamics for machine learning. *Phys. Rev. Applied*, 8:024030, 2017.

[17] Johannes Nokkala, Rodrigo Martínez-Peña, Gian Luca Giorgi, Valentina Parigi, Miguel Soriano, and Roberta Zambrini. Gaussian states provide universal and versatile quantum reservoir computing. *arXiv:2006.04821*, 2020.

[18] Kohei Nakajima, Keisuke Fujii, Makoto Negoro, Kosuke Mitarai, and Masahiro Kitagawa. Boosting computational power through spatial multiplexing in quantum reservoir computing. *Phys. Rev. Applied*, 11:034021, 2019.

[19] Jacob Smith, Aaron Lee, Philip Richerme, Brian Neyenhuis, Paul W Hess, Philipp Hauke, Markus Heyl, David A Huse, and Christopher Monroe. Many-body localization in a quantum simulator with programmable random disorder. *Nature Physics*, 12(10):907–911, 2016.

[20] J Zhang, G Pagano, P. W. Hess, A Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z-X Gong, and Christopher Monroe. Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator. *Nature*, 551, 2017.

[21] Rodrigo Martínez Peña, Gian Luca Giorgi, Johannes Nokkala, Miguel C. Soriano, and Roberta Zambrini. Dynamical phase transitions in quantum reservoir computing. *arXiv*, to be published.

[22] Joni Dambre, David Verstraeten, Benjamin Schrauwen, and Serge Massar. Information processing capacity of dynamical systems. *Scientific Reports*, 2(1):514, 2012.

[23] Alexander Streltsov, Gerardo Adesso, and Martin B. Plenio. Colloquium: Quantum coherence as a resource. *Rev. Mod. Phys.*, 89:041003, 2017.

[24] Sayandip Dhara, Alioscia Hamma, and Eduardo R. Mucciolo. Quantum coherence in ergodic and many-body localized systems. *Phys. Rev. B*, 102:045140, 2020.

[25] Luca D'Alessio, Yariv Kafri, Anatoli Polkovnikov, and Marcos Rigol. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. *Advances in Physics*, 65(3):239–362, 2016.

[26] Dmitry A. Abanin, Ehud Altman, Immanuel Bloch, and Maksym Serbyn. Colloquium: Many-body localization, thermalization, and entanglement. *Rev. Mod. Phys.*, 91:021001, 2019.

[27] Fabien Alet and Nicolas Laflorencie. Many-body localization: An introduction and selected topics. *Comptes Rendus Physique*, 19(6):498 – 525, 2018.

[28] Antonello Scardicchio and Thimothée Thiery. Perturbation theory approaches to anderson and many-body localization: some lecture notes. *arXiv:1710.01234v1*, 2017.

[29] Rodrigo Martínez Peña, Johannes Nokkala, Gian Luca Giorgi, Roberta Zambrini, and Miguel C. Soriano. Information processing capacity of spin-based quantum reservoir computing systems. *Cognitive computation*, accepted.

[30] Aki Kutvonen, Keisuke Fujii, and Takahiro Sagawa. Optimizing a quantum reservoir computer for time series prediction. *Scientific Reports*, 10(1):14687, 2020.

[31] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Phys. Rev. Lett.*, 35:1792–1796, 1975.

[32] Sudip Mukherjee, Sabyasachi Nag, and Arti Garg. Many body localization-delocalization transition in quantum sherrington-kirkpatrick model. *Physical Review B*, 97:144202, 2018.

[33] Yasar Yilmaz, Eugene Bogomolny, O Giraud, and G Roux. Distribution of the ratio of consecutive level spacings in random matrix ensembles. *Physical review letters*, 110:084101, 2013.

[34] Michael Nielsen and Isaac Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[35] Jiayin Chen and Hendra I. Nurdin. Learning nonlinear input-output maps with dissipative quantum systems. *Quantum Information Processing*, 18(7):198, 2019.

[36] William K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245–2248, 1998.

[37] Emanuele Levi, Markus Heyl, Igor Lesanovsky, and Juan P. Garrahan. Robustness of many-body localization in the presence of dissipation. *Phys. Rev. Lett.*, 116:237203, 2016.