Universitat
de les Illes Balears

DOCTORAL THESIS
2023

RESERVOIR COMPUTING IN QUANTUM SYSTEMS

Rodrigo Martínez Peña

Universitat
de les Illes Balears

DOCTORAL THESIS
2023

DOCTORAL PROGRAMME IN PHYSICS

RESERVOIR COMPUTING IN QUANTUM SYSTEMS

Rodrigo Martínez Peña

THESIS SUPERVISOR: *Dr. Roberta Zambrini*
THESIS SUPERVISOR: *Dr. Miguel Cornelles Soriano*
THESIS TUTOR: *Dr. María Rosa López Gonzalo*

DOCTOR BY THE UNIVERSITAT DE LES ILLES BALEARS

Dr. Roberta Zambrini, Dr. Miguel Cornelles Soriano and Dr. María Rosa López Gonzalo, of the Institute for Cross-Disciplinary Physics and Complex Systems (IFISC)

WE DECLARE:

That the thesis titled "Reservoir Computing in Quantum Systems", presented by Rodrigo Martínez Peña to obtain a doctoral degree, has been completed under our supervision.

For all intents and purposes, we hereby sign this document.

Dr. Roberta Zambrini Thesis supervisor	Dr. Miguel Cornelles Soriano Thesis supervisor
Rodrigo Martínez Peña PhD student	Dr. María Rosa López Gonzalo Thesis tutor

Palma de Mallorca,

Abstract

Abstract of the thesis

Reservoir computing (RC) is a machine learning paradigm that exploits dynamical systems to solve temporal tasks. This technique finds applications in very diverse fields such as weather forecasting, stock market predictions, and communications. Similar to other unconventional computing paradigms inspired by the capabilities of the human brain, RC deals with hardware implementations that aim at overcoming the challenges confronted by digital computation. These challenges include the reduction of the energy budget of digital computation and the speedup of machine learning algorithms. This thesis explores the emerging field of quantum RC. We studied, either by analytical or numerical methods, which are the requirements of complex quantum systems to perform as useful reservoirs, with special attention to quantum spin models. Useful reservoir systems will be defined as those that meet the fundamental requirements that ensure a minimum level of performance from the reservoir. Our main findings are the identification of the dynamical features and input injection favoring quantum reservoirs, and the theoretical conditions for useful reservoirs. In particular, computational capabilities and input response of reservoirs displaying many-body localization are totally degraded by the presence of local integrals of motion while in thermal phases (or at the edge of transition) the operation is optimal. These computational capabilities are characterized by means of the information processing capacity, obtained for the first time for quantum reservoirs, and other benchmark tools such as the short-term memory and nonlinear autoregressive moving average tasks. Characterization of the performance through all these tools allows one to assess the linear and nonlinear contributions of a specific reservoir. Moreover, the input codification mechanism determines the nonlinear response of the reservoir together with the dynamical regime and the election of observable. We demonstrate this relation by showing explicit analytical formulas of the input-output map of the studied reservoir models. Beyond ideal conditions, we explore how all these factors can be affected by the implementation of a quantum reservoir computing experiment with an online protocol, where the extraction of information through measurements is accounted for. Weak measurements are introduced as a possible route to achieve a competitive performance for online temporal processing while keeping a high control over the required experimental resources. Finally, on a theoretical general side, all finite-dimensional quantum reservoir computing models with classical inputs must fulfill the following condition to be, at least, operational: convergent dynamics towards input-dependent fixed points.

Resumen de la tesis

La computación de reservorio (CR) es un paradigma de aprendizaje automático que explota los sistemas dinámicos para resolver tareas temporales. Esta técnica encuentra aplicaciones en campos muy diversos, como la previsión meteorológica, las predicciones bursátiles y las comunicaciones. Al igual que otros paradigmas computacionales no convencionales inspirados en las capacidades del cerebro humano, la CR se ocupa de implementaciones de hardware que pretenden superar los retos a los que se enfrenta la computación digital. Estos retos incluyen la reducción del gasto energético de la computación digital y la aceleración de los algoritmos de aprendizaje automático. Esta tesis explora el emergente campo del CR cuántico. Aquí estudiamos mediante métodos analíticos o numéricos cuáles son los requisitos de los sistemas cuánticos complejos para funcionar como reservorios útiles, con especial atención a los modelos cuánticos de espín. Los sistemas reservorios útiles se definirán como aquellos que cumplen los requisitos fundamentales que aseguran un nivel mínimo de rendimiento del reservorio. Nuestros principales hallazgos son la identificación de las características dinámicas y de la inyección de entrada que favorecen a los reservorios cuánticos, y las condiciones teóricas para definir reservorios útiles. En particular, las capacidades computacionales y la respuesta a la información de entrada de los reservorios que muestran localización están totalmente degradadas por la presencia de integrales locales de movimiento, mientras que en fases térmicas (o al borde de la transición) el funcionamiento es óptimo. Estas capacidades computacionales se pueden caracterizar mediante la medida de la capacidad de procesamiento de la información, obtenida por primera vez para los reservorios cuánticos, y otras herramientas de referencia como las tareas de memoria a corto plazo y de media móvil autorregresiva no lineal. La caracterización del rendimiento mediante todas estas herramientas permite evaluar las contribuciones lineales y no lineales de un reservorio específico. Además, hemos identificado que el mecanismo de codificación de la entrada determina la respuesta no lineal del reservorio junto con el régimen dinámico y la elección del observable. Demostramos esta relación mostrando fórmulas analíticas explícitas del mapa entrada-salida de los modelos de reservorio estudiados. Más allá de las condiciones ideales, exploramos cómo todos estos factores pueden verse afectados por la implementación de un experimento de computación de reservorio cuántico con un protocolo en línea, donde la extracción de información mediante mediciones se tiene en cuenta. Las medidas débiles se presentan como una posible vía para lograr un rendimiento competitivo para el procesamiento temporal en línea, manteniendo al mismo tiempo un alto control sobre los recursos experimentales necesarios. Por último, desde un punto de vista teórico general, todos los modelos de computación de reservorio cuántico de dimensión finita con entradas clásicas deben cumplir la siguiente condición para ser, al menos, operativos: dinámica convergente hacia puntos fijos dependientes de la entrada.

Resum de la tesi

La computació de reservori (CR) és un paradigma d'aprenentatge automàtic que explota els sistemes dinàmics per resoldre tasques temporals. Aquesta tècnica troba aplicacions en camps molt diversos, com la previsió meteorològica, les prediccions borsàries i les comunicacions. Com altres paradigmes computacionals no convencionals inspirats en les capacitats del cervell humà, la CR s'ocupa d'implementacions de hardware que pretenen superar els reptes a què s'enfronta la computació digital. Aquests reptes inclouen la reducció de la despesa energètica de la computació digital i l'acceleració dels algorismes d'aprenentatge automàtic. Aquesta tesi explora el camp emergent del CR quàntic. Aquí estudiem mitjançant mètodes analítics o numèrics quins són els requisits dels sistemes quàntics complexos per funcionar com reservoris útils, amb una atenció especial als models quàntics d'espín. Els sistemes reservoris útils es definiran com aquells que compleixen els requisits fonamentals que asseguren un nivell mínim de rendiment del reservori. Les nostres principals troballes són la identificació de les característiques dinàmiques i de la injecció d'entrada que afavoreixen els reservoris quàntics i de les condicions teòriques per definir reservoris útils. En particular, les capacitats computacionals i la resposta d'entrada dels reservoris que mostren localització estan totalment degradades per la presència d'integrals locals de moviment, mentre que en fases tèrmiques (o vora la transició) el funcionament és òptim. Aquestes capacitats computacionals es caracteritzen mitjançant la capacitat de processament de la informació, obtinguda per primera vegada per als reservoris quàntics, i altres eines de referència com les tasques de memòria a curt termini i de mitjana mòbil autoregressiva no lineal. La caracterització del rendiment mitjançant totes aquestes eines permet avaluar les contribucions lineals i no lineals d'un reservori específic. A més, el mecanisme de codificació de l'entrada determina la resposta no lineal del reservori juntament amb el règim dinàmic i l'elecció de l'observable. Demostrem aquesta relació mostrant fórmules analítiques explícites del mapa entrada-sortida dels models de reservori estudiats. Més enllà de les condicions ideals, explorem com tots aquests factors es poden veure afectats per la implementació d'un experiment de computació de reservori quàntic amb un protocol en línia, on es té en compte l'extracció d'informació mitjançant mesuraments. Les mesures febles es presenten com una possible via per assolir un rendiment competitiu per al processament temporal en línia, mantenint alhora un alt control sobre els recursos experimentals necessaris. Per acabar, des d'un punt de vista teòric general, tots els models de computació de reservori quàntic de dimensió finita amb entrades clàssiques han de complir la condició següent per ser, almenys, operatius: dinàmica convergent cap a punts fixos dependents de l'entrada.

Agradecimientos

Como bien explica Pere Colet, el oficio de científico se transmite de la misma forma que cualquier otro oficio artesanal: de maestros a aprendices. Cuando me uní al gremio del IFISC, fue solo para hacer un máster durante un año, con la grandísima suerte de que concedieron al instituto la (primera) María de Maeztu cuando acabé. Así que aquí estamos, más de cuatro años después, pasando de aprendiz a oficial del gremio.

Comencemos con las formalidades. Esta tesis nunca hubiera sido posible sin la ayuda de toda la gente que me he encontrado a lo largo de estos años. Todos, en mayor o menor medida, han contribuido a construir y consolidar la persona que soy ahora, no solo como científico, sino también como compañero y amigo. La mayor aportación al contenido de esta tesis se debe indudablemente a la inagotable paciencia de mis maestros artesanos, Roberta y Miguel. Su actitud y rigurosidad me han servido como referencia para moldear mi carácter como investigador. Y no solo me han enseñado la parte más técnica del oficio de científico, sino también la parte más humana. De verdad, muchas gracias.

Mis colaboradores más cercanos también han sido una gran influencia: Johannes, Pere Mujal, Antonio y Gian Luca. Muchas gracias a todos por aguantarme, que sé que puedo ser muy pesado con los detalles. Mención también para mi último maestro, Juan Pablo. Contigo no solo descubrí mi faceta más matemática, sino que también me ayudaste a lidiar con los conflictos internos que genera la investigación. Lo último se podría resumir en: "chico, si te quieres dedicar a esto, tendrás que aprender a vivir con la frustación". Porque la investigación científica es una disciplina muy desagradecida. Meses de chocarte constantemente contra una pared intentando resolver un problema... Hasta que un día, de repente, las cosas salen bien. Como también dice Juan Pablo: "ese día paras de trabajar, no vayas a encontrar algún fallo". Y es que la sensación del día que todo sale bien compensa todo lo anterior.

No me quiero olvidar de toda la gente que trabaja (o ha trabajado) en la administración del IFISC y que ha hecho del temible papeleo algo mucho más llevadero, especialmente Marta e Inma. También Maria Quetglas, que aunque de la UIB, nos ha sufrido igualmente. No me olvido tampoco de los mejores técnicos que tantas veces nos han salvado, aún aburridos de nuestras preguntas tontas y estropicios en Nuredduna: Edu, Rubén y Antonia. Todos ellos son los que mantienen a flote este barco. Mención especial para Adrián García, con el cual me adentré en el mundo de la divulgación. Una pena que el covid no nos permitiese continuar.

Continuamos con la sección familia y amigos, con no menos clichés. A mis padres les tengo que dar las gracias básicamente por todo, porque si no fuera por ellos no estaría aquí. De mi hermana Raquel, con la cual he compartido gran parte del viaje del doctorado, he aprendido lo lejos que se puede llegar si uno tiene determinación. Ahora vamos a celebrarlo como se merece. Mil gracias a todos los estudiantes y postdocs del IFISC con los he convivido y compartido estos años, que antes o después, han pasado por las mesas de la S07, el "zulo" o algún despacho del sótano.

Reservo un párrafo aparte para la gente con la que comencé esta aventura en Mallorca y que ahora forman una parte indispensable de mi vida: Alberto, Javi, Irene y María. Con vosotros he compartido la mayor parte de estos años, y os estaré eternamente agradecido por haber estado ahí. Finalmente, quiero dedicar un agradecimiento especial a las personas con las que he compartido , escuchado, tocado, discutido e incluso creado música. Los que me conocen, saben que mi día a día viene acompañado de una banda sonora de trompa. Muchas gracias a Adrià, Pablo, Antonio Lucena, Jesús, Lluís, Alberto, Javi, Catalina, Manu, Irene y María (vosotras en particular os habéis ganado el cielo).

List of publications

This thesis is presented as a compendium of publications. The list of articles detailed below, in chronological order by date of publication, form the main body of the thesis.

Thesis articles

1. 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*, 1–12 (2020).
   Chapter 3.

2. Pere Mujal, Rodrigo Martínez-Peña, Johannes Nokkala, Jorge García-Beni, Gian Luca Giorgi, Miguel C. Soriano and Roberta Zambrini, Opportunities in quantum reservoir computing and extreme learning machines, *Advanced Quantum Technologies*, **4**, 2100027 (2021).
   Chapter 2.

3. Rodrigo Martínez-Peña, Gian Luca Giorgi, Johannes Nokkala, Miguel C. Soriano and Roberta Zambrini, Dynamical phase transitions in quantum reservoir computing, *Physical Review Letters*, **127**, 100502 (1-7) (2021).
   Chapter 4.

4. Pere Mujal, Johannes Nokkala, Rodrigo Martínez-Peña, Gian Luca Giorgi, Miguel C. Soriano and Roberta Zambrini, Analytical evidence of nonlinearity in qubits and continuous-variable quantum reservoir computing, *Journal of Physics: Complexity*, **2**, 045008 (2021).
   Chapter 5.

5. Pere Mujal, Rodrigo Martínez-Peña, Gian Luca Giorgi, Miguel C. Soriano and Roberta Zambrini, Time-series quantum reservoir computing with weak and projective measurements, *npj Quantum Information*, **9**, 16 (2023).
   Chapter 6.

6. Rodrigo Martínez-Peña and Juan-Pablo Ortega, Quantum reservoir computing in finite dimensions, *Physical Review E*, **107**, 035306 (2023).
   Chapter 7.

Other publications

1. Johannes Nokkala, Rodrigo Martínez-Peña, Gian Luca Giorgi, Valentina Parigi, Miguel C. Soriano, and Roberta Zambrini, Gaussian states of continuous-variable quantum systems provide universal and versatile reservoir computing, *Communications Physics*, **4** (1), 53 (2021).

2. Johannes Nokkala, Rodrigo Martínez-Peña, Miguel C. Soriano and Roberta Zambrini, High-performance reservoir computing with fluctuations in linear networks, *IEEE Transactions on Neural Networks and Learning Systems*, **33** (6), 2664-2675 (2021).

3. Pere Mujal, Johannes Nokkala, Rodrigo Martinez-Peña, Jorge García-Beni, Gian Luca Giorgi, Miguel C. Soriano and Roberta Zambrini, Quantum reservoir computing in bosonic networks, *Emerging Topics in Artificial Intelligence (ETAI)*, **11804**, 20-29 (2021).

4. Antonio Sannia, Rodrigo Martínez-Peña, Miguel C. Soriano, Gian Luca Giorgi and Roberta Zambrini, Dissipation as a resource for quantum reservoir computing, *arXiv:2212.12078*, (2022).

Contents

Chapter 1

Introduction

1.1 Context

Over the last decades, different computational visions have emerged as alternatives to the conventional digital paradigm. In particular, quantum and neuromorphic computing are two of the leading approaches that promise to revolutionize this field. On the one hand, quantum computation harnesses unique features such as superposition and entanglement in view of an advantage over classical algorithms [1]. The current development of quantum technologies has achieved noisy intermediate-scale quantum devices [2]. These systems, composed of tens or hundreds of noisy elements, enable us to perform imperfect operations only for a short time, while coherences are still preserved. Even if current systems are noisy, recent advancements like the demonstration of quantum advantage [3–5], the development of quantum simulations [6], and quantum communications [7] demonstrate that quantum technologies are here to stay. Indeed, they already find applications in diverse fields like physics, chemistry, and optimization [2]. Machine learning (ML) is among these possible applications, where the specific features of quantum mechanical systems are exploited to obtain an advantage over their classical counterparts, both with classical and quantum data [8]. On the other hand, neuromorphic computing is a biology-inspired computing paradigm where physical substrates are designed to mimic the brain, calculating faster with low energy consumption. ML and artificial neural networks (ANNs) have demonstrated that they are one of the main pillars of computation nowadays [9, 10]. However, these algorithms usually run on large computing infrastructures, separating processing and memory, slowing them down, and increasing the energy consumption. Neuromorphic computing responds to these demands and ML and ANNs benefit from this progress [11–13].

This thesis is devoted to the analysis and development of a popular neuromorphic computing technique in its quantum version. Quantum reservoir computing (QRC), as an extension of classical reservoir computing (RC), aims to exploit quantum systems as physical substrates to solve ML tasks. The interest in such implementations stems from the possible advantages that QRC platforms could bring over their classical counterparts. Indeed, the number of degrees of freedom that quantum systems possess could boost the information processing performance. Furthermore, available experimental platforms, such as trapped-ion simulators [14, 15], nuclear magnetic resonance (NMR) devices [16], and photonic networks [17] could allow testing this paradigm in the short-medium term.

Finally, QRC would enable the processing of quantum data embedded in fully quantum architectures.

1.1.1 Thesis structure

The aim of this thesis is to tackle questions related to the construction and implementation of **useful quantum reservoir computers**. We will consider useful reservoir systems to be those that fulfill the basic conditions that guarantee a minimum performance from the reservoir. We will test the performance of the quantum reservoir for selected benchmark tasks, which can be seen as a proof of concept validation of the potential of these systems.

This thesis is based on the results of Refs. [18–23]. In Chapter 1 we introduce the fields of classical machine learning and artificial neural networks, to later approach the main topic of this thesis, reservoir computing, and its quantum version. The second part of this chapter introduces the required concepts to understand the context of the publications, such as a brief introduction to many-body localization or quantum measurements. We finish the chapter with a summary of the main contributions of this PhD work. Chapters 2 to 7 reproduce each one of the articles that compose this thesis. Each article is preceded by a short introduction where motivation and personal contributions to the paper in question are presented. Finally, Chapter 8 summarizes the conclusions of the whole thesis, including an outlook on the possibilities that the field of QRC offers in the next few years.

1.2 Machine learning

There has been an unprecedented revolution in the processing of information during the last decades. Originally, data have been a precious treasure, hard to generate, analyze, store, and use. But the development of information and communication technologies has created the opportunity to produce, access, and interpret a vast amount of data, known as **big data** [24]. The technological development that has allowed this expansion has been specifically carried out on digital platforms, where data are easily stored and analyzed. From the hardware point of view, digital computers have experienced an exponential increase of computational power and memory (as indicated by Moore's law) [25], and specialized hardware has been developed for the demands of the big data era, such as supercomputers or the graphical processing units (GPUs) [26]. From the software point of view, new techniques have been developed to generate, analyze and learn from large data sets. **Machine learning** (ML) is halfway between a promoter and a beneficiary of this big data revolution, becoming one of the main research fields of our era on its own and being ubiquitous in practically any other research field. ML can be defined as a set of techniques and algorithms to learn from and make predictions about data, without the need for a precise set of instructions [27, 28]. This rather compact definition does not make justice to the whole spectrum of problems that ML can tackle and currently solve. Just a recent example: Alphafold, a deep learning algorithm developed by the company

Deepmind, is able to predict the 3D structure that proteins will form based only on the amino acid sequence, largely outperforming the previous methods in the field and becoming one of the new cornerstones of bioinformatics [29].

The ML research field represents a broad family of possible implementations, where frontiers between different techniques are becoming more and more blurred by emerging hybrid proposals. Still, three basic types of learning can be distinguished depending on the problem and type of data we are dealing with. **Supervised learning** is the case where we feed inputs with the corresponding correct outputs (or labels) in order to tune the parameters of the algorithm, i.e. to train from examples [30]. In this way, we prepare the machine to make predictions of unseen data. The most common tasks for supervised learning are regression and classification problems. In **unsupervised learning**, instead, the algorithm has to find structure and patterns in the input data without any prior information about the output [31]. The most common example of unsupervised learning is clustering, where inputs are gathered into different groups and data from each group share some properties. Finally, in **reinforcement learning** an agent learns by interacting with its environment while adapting its behavior to maximize a reward [32]. In this thesis, we are interested in supervised training, where reservoir computing has been mostly developed. We introduce next the field of artificial neural networks, which constitutes one of the precedents of RC.

1.2.1 Artificial neural networks

Artificial neural networks (ANNs) are powerful and widely-used supervised models. These brain-inspired techniques are able to solve a large variety of different problems while being only composed of two constituents: **neurons** and **links**. The neuron is the basic unit of the ANN. Inputs are fed into it and a nonlinear function (a.k.a. activation function) processes the information to produce an output. Neurons are connected between them by links whose weights can be adjusted. The network is structured by layers and their connections define the architecture of the ANN. The brain-inspired construction of ANNs allows to process information in an unprecedented way, exploiting the correlations of different neurons at different layers of the network. Then, the architecture of the ANNs is so important that it mainly determines the type of problems that one can solve, becoming its design a craft by itself [33].

A common architecture is represented by **feed-forward neural networks** (FFNNs) [9, 10, 30], which provide good performance for a wide range of problems. FFNNs, as illustrated in Fig. 1.1 (a), can be defined as a series of layers of neurons where information always flows towards the output (or forward direction). The layers can be separated into three categories: **input layer**, **hidden layers**, and **output layer**. The input of the problem (data to be processed) is fed into the input layer and the information is processed at each step by a different hidden layer until it reaches the output layer, where information is read out. The output $y_i^l \in \mathbb{R}$ of a FFNN neuron with index i at a given layer l defines a layer vector $\mathbf{y}^l = (y_1^l, y_2^l, \dots)$. Neurons of different layers are related by:

$$y_i^l = \sigma(\mathbf{w}_i^l \cdot \mathbf{y}^{l-1} + b_i^l), \tag{1.1}$$

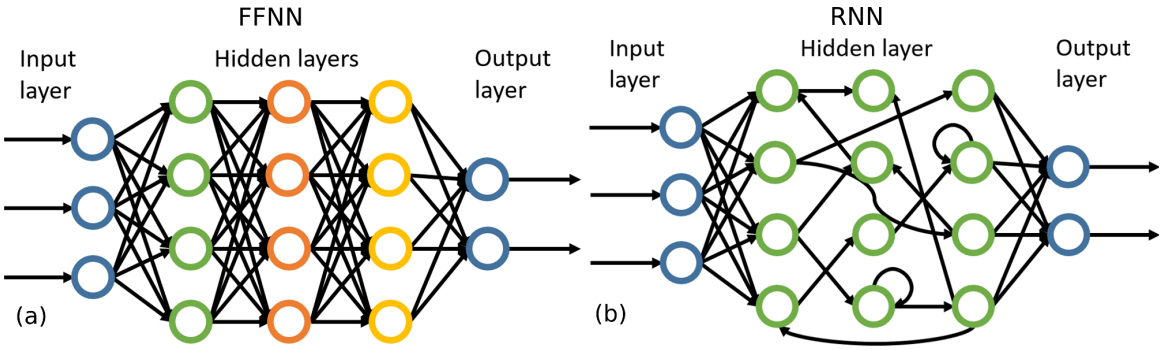

FIGURE 1.1: (a) Scheme of a feed-forward neural network (FFNN). From left to right: input layer, where inputs are codified and might be preprocessed; hidden layers, where neurons process information in parallel at each layer; and output layer, where information is finally extracted. (b) Scheme of a recurrent neural network (RNN). It follows a similar structure to the FFNN, but in its most basic conception we only find a single hidden layer, that allows recurrent connections between the neurons, providing memory to the system.

where σ is the activation function and $\mathbf{y}^{l-1} \in \mathbb{R}^{n_{l-1}}$ is the output vector of the previous layer, that comes as input at layer l. n_{l-1} is the number of neurons of layer $l-1$. The vector $\mathbf{w}_i^l \in \mathbb{R}^{n_{l-1}}$ represents the weighted connections between all neurons of layer $l-1$ and neuron i of layer l, and $b_i^l \in \mathbb{R}$ is a bias term [30]. There are many options for the choice of the activation function, being the sigmoid curve one of the most common ones, and not all neurons need to have the same one. A **sigmoid function** is an increasing, differentiable, and monotonic nonlinear function whose amplitude ranges are constrained. Two of the most representative sigmoid functions are the hyperbolic tangent $\sigma(x) = \tanh(x)$ and the logistic function $\sigma(x) = (1 + e^{-x})^{-1}$, see Fig. 1.2 for an illustration.

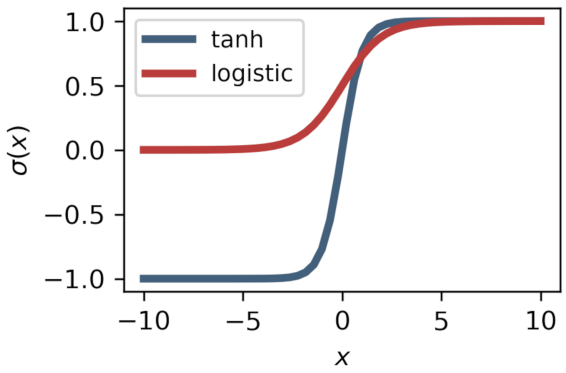

FIGURE 1.2: Two examples of sigmoid functions.

In supervised learning approaches, a part of labeled data are used during training, defining a **target** to approximate. The goal is to tune all the weights $W = \{\mathbf{w}_j^l\}$ and biases $B = \{b_j^l\}$ with input-output examples to obtain the closest approximation to these targets. Let us define a set of n_{sample} training samples

$\{\mathbf{s}_j, \bar{\mathbf{y}}_j\}$, where $\mathbf{s}_j \in \mathbb{R}^m$ and $\bar{\mathbf{y}}_j \in \mathbb{R}^d$ are given as input and target output samples for training respectively, and $\mathbf{y}(\mathbf{s}_j) \in \mathbb{R}^d$ will denote the corresponding final output generated by the ANN for the same input sample. The dimension m corresponds to the dimension of the input and the dimension d denotes both the dimension of the targets and the number of neurons of the last layer. In order to characterize the performance of the setup, a cost function $C(W, B)$ is defined in terms of these variables to measure how good the approximation is. A usual example is the **mean square error** (MSE):

$$C(W, B) = \frac{1}{2n_{\text{sample}}} \sum_{j=1}^{n_{\text{sample}}} ||\bar{\mathbf{y}}_j - \mathbf{y}(\mathbf{s}_j)||^2, \tag{1.2}$$

where $||\cdot||$ denotes the usual euclidean norm. To find the optimal set of parameters W and B one must optimize $C(W, B)$ with respect to the training sample. We note that optimization is generally not trivial and time-consuming due to the presence of local minima or slow convergence. Multivariate optimization is usually performed through gradient descent and its variations, such as **stochastic gradient descent** (SGD) [9, 30]. SGD consists in computing the gradients over small subsets of the total sample set, changing the subset at each iteration of the training process. This allows us to reduce the number of computations per iteration and circumvent local minima. The general approach to compute the gradients is backpropagation [34], which allows us to compute the gradients with respect to all weights and biases in an efficient way since it is entirely based on linear algebra operations [30].

Multilayer FFNNs, commonly known as **deep neural networks**, find a multitude of applications due to their ability to learn representations of data with multiple levels of abstraction. To name just a few, FFNNs have been a major breakthrough in very distinct areas like speech recognition [35], analyzing data from particle physics [36] or content-based recommendations [37]. A particular type of deep FFNN is **convolutional neural networks** (CNNs), which do not display full connectivity between adjacent layers in favor of a better trainability and generalization capability for some specific tasks [38]. CNNs have been also a revolution in areas such as image classification [39], face recognition [40] and natural language processing [41]. Most of the current deep learning approaches mix different learning techniques and architectures to tackle more complex problems. We can find striking examples like AI beating best world players in chess and Go [42], art generation from text [43] or human-like text generation for coding [44].

In contrast to FFNNs, **recurrent neural networks** (RNNs) exhibit an architecture that allows information to remain in the network for several input injections [45, 46]. RNNs were designed to deal with tasks where memory is required, i.e., tasks where the input is fed as a sequence. As shown in the schematic representation of Fig. 1.1 (b), this memory effect is obtained by allowing recurrent connections between the nodes of the hidden layers, that is, information can flow back and forth between them. This creates a single hidden layer where connections represent toward which neuron information is sent at the next time step.

The most basic dynamics of RNNs is usually described by a set of two equations:

$$\begin{aligned} \mathbf{x}_k &= \sigma(W\mathbf{x}_{k-1} + W_{\text{in}}\mathbf{s}_k + \mathbf{b}), \\ \mathbf{y}_k &= o(\mathbf{x}_k), \end{aligned} \tag{1.3}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state of the n-neurons hidden layer at time step k, $\mathbf{s}_k \in \mathbb{R}^m$ is the input vector of dimension m, $W \in \mathbb{R}^{n\times n}$ is the matrix of recurrent connections, $W_{\text{in}} \in \mathbb{R}^{n\times m}$ is the weight matrix between neurons and input layer and $\mathbf{b} \in \mathbb{R}^n$ is the bias vector. The function o reads out the neurons' information and vector $\mathbf{y}_k \in \mathbb{R}^d$ represents its d-dimensional output.

RNNs are very powerful ML techniques but they are also really hard to train. The problem arises during gradient descent training and backpropagation through time [47]. If we unfold an RNN in time, it can be seen as a very deep feedforward network in which all the layers share the same weights. Backpropagated gradients either grow or shrink at each time step, so over long sequences, they usually vanish or explode [48, 49], forbidding optimization. Alternatives have been proposed to circumvent this problem. New training methods [50–52] and different architectures, such as **long short-term memories** (LSTMs) [53] or **gated recurrent units** [54] (GRUs), have allowed these techniques to grow over the last years. Examples of groundbreaking applications are language translation [55, 56], text generation [57], and speech recognition [58]. As previously mentioned, current approaches to deep learning do not employ a single basic architecture to address complex problems. Instead, multilayer FFNNs can be combined with other specific-purpose network architectures, such as RNNs. A representative example of this is the application of LSTM networks for playing real-time video games such as Starcraft II [59], where memory is required to deal with the sequence of observations of the AI player.

Besides the usual applications in the area of speech and text processing, RNNs are often considered the first choice for time-series tasks. Very diverse problems can be successfully solved in this line, such as weather forecasting [60], music generation [61], or even reconstructing the quantum dynamics of superconducting qubits [62]. This last example is one of a long list of classical machine learning applications to physics and in particular quantum problems, see [63–65] and references therein for further examples.

However, conventional RNNs and their variations are not the most efficient options for time-series processing. It has been confirmed by numerical experiments that reservoir computing architectures can obtain similar or better performance with fewer resources [66–68]. To better understand these results, the next section introduces the reservoir computing formalism, where we will explain the nuances of this ML technique.

1.3 Reservoir computing

1.3.1 From digital to analog computing

Most of the ML techniques have been developed for digital computers due to their **accessibility** [69]: they are easy to exploit for all levels of expertise, there

is a unified theory that is taught in all universities in the same way, and digital computation is universal (in the sense it can emulate Turing machines). However, digital computation also has some **associated problems**. In 2015, it was estimated that 10% of the global energy consumption was due to digital computation, and it was (and it is) increasing [70]. In particular, the computational cost of deep learning is becoming more and more economically and environmentally unsustainable [71, 72]. On another side, miniaturization is limited by Moore's Law due to thermal and quantum fluctuations in the development of smaller microprocessors [25]. We also find that the life of electronic hardware components is shortening, speeding up their replacement cycles and producing a vast amount of wasted resources every year [73]. Finally, the von Neumann Bottleneck hinders the performance of digital computers. This refers to the time delay spent in transferring data between processor and memory in platforms with a von Neumann architecture [74]. In principle, there are some ways of alleviating this problem, such as adding hierarchical memory structures inside the CPUs to cache frequently used data. However, if one takes into account energy consumption, these options do not constitute a long-term solution [13].

Different strategies are proposed in the literature to tackle these problems. Here we are concerned with the alternative offered by **neuromorphic computing** [69]. This non-conventional proposal is a brain-inspired computing paradigm where computation and physical substrate go hand by hand. The main characteristics of neuromorphic computing, as a brain-inspired paradigm for both software and hardware, are energy efficiency, parallel computation, and co-location of processing and memory. Currently, the research field of ML is already taking advantage of this progress, and we find reservoir computing as the confluence of both fields [75].

1.3.2 Classical reservoir computing

Reservoir computing (RC) is a broad sub-field of ML where **dynamical systems** are exploited to solve temporal tasks [76]. We will start by introducing classical RC and move to quantum proposals in Sect. 1.4. As commented before, RNNs are paradigmatic examples of ANNs that are used for solving ML tasks where time series are involved. And since an RNN is a dynamical system in itself, one can try to use it for RC.

Indeed, one of the main routes towards the inception of RC came from RNNs [77], simplifying the training of these ANNs but trying to keep a good performance. This idea is condensed in what is known as **echo state networks** (ESNs). An ESN is similarly defined as in Eq. (1.3):

$$\begin{aligned} \mathbf{x}_k &= \tanh(W\mathbf{x}_{k-1} + W_{\text{in}}\mathbf{s}_k + \mathbf{b}), \\ \mathbf{y}_k &= W_{\text{out}}\mathbf{x}_k, \end{aligned} \tag{1.4}$$

where now W and W_{in} are random matrices that are not optimized, and the output function is linear with weights $W_{\text{out}} \in \mathbb{R}^{d\times n}$. Usually, a constant bias is introduced for optimal performance by adding an extra column to matrix W_{out} and adding a constant term to vector $\mathbf{x}_k$. The most common activation function is the

hyperbolic tangent, but different variations have been proposed in the literature over these years, see [78] for a review. While RNNs require a more cumbersome training like backpropagation through time [47] or Atiya-Parlos recurrent learning [50], the election of a linear output layer in the model of Eq. (1.3) is what brings a significantly simplified training, as we will show below.

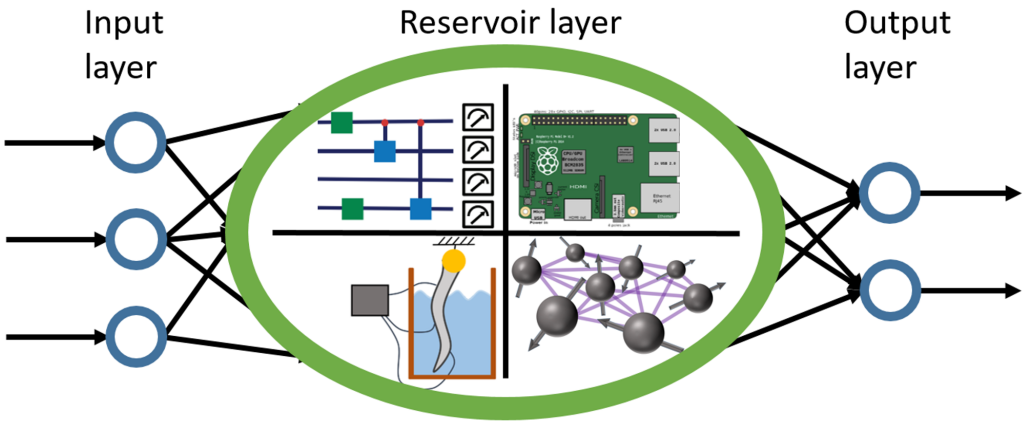

FIGURE 1.3: Scheme of a reservoir computing (RC) system based either on classical or quantum reservoirs. As in the RNN architecture, we find three main components: input layer, reservoir, and output layer. The main difference with RNNs is that the reservoir is a fixed dynamical system that is not tuned, training only the output layer. Different dynamical systems can fulfill the role of the reservoir, such as quantum circuits [79], electronic devices [80], soft robotic arms [81] and quantum spin systems [19].

In a different context, **liquid state machines** (LSMs) were proposed almost at the same time as ESNs to model neural microcircuits [82], introducing a biological perspective. The state equation of an LSM is represented in continuous time as

$$\begin{aligned} \mathbf{x}(t) &= (\mathcal{L}\mathbf{s})(t), \\ \mathbf{y}(t) &= o(\mathbf{x}(t)), \end{aligned} \tag{1.5}$$

where $\mathcal{L}$, the "liquid", transforms input information $\mathbf{s}$ into the liquid state $\mathbf{x}(t)$ and o is again the memoryless readout. Both approaches, ESN and LSM, share the basic idea of avoiding a fine-tuning of the parameters of the dynamical system (known as reservoir). This tuning of parameters (or training) is usually only applied to the weights of an output layer, constructed from the readout information of the reservoir.

As in the RNNs case, we can identify **three** basic **layers** in the RC scheme: an input layer, where information is codified into some of the degrees of freedom of the reservoir; the reservoir layer (or hidden layer), where information is processed via the natural dynamics of the classical or quantum (see Sect. 1.4) system; and output layer, which is usually constructed as a linear combination of some (or all) the degrees of freedom of the reservoir. Figure 1.3 depicts a scheme with the most basic conception.

Let us now introduce the basic mathematical framework for RC. Inputs are defined as infinite discrete-time sequences, given as $\mathbf{s} = (\dots, \mathbf{s}_{-1}, \mathbf{s}_0, \mathbf{s}_1, \dots) \in$

$(\mathbb{R}^m)^{\mathbb{Z}}$ where m is the vector dimension of the data to be processed at each time step. Outputs are also sequences denoted as $\mathbf{y} \in (\mathbb{R}^d)^{\mathbb{Z}}$, being d the vector dimension of the output at each time step. Let $T : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ be the **reservoir map**, where input codification and reservoir dynamics are condensed in T. The update equation at each time step k is written as

$$\mathbf{x}_k = T(\mathbf{x}_{k-1}, \mathbf{s}_k), \tag{1.6}$$

and the output vector is obtained from a readout function of the reservoir states,

$$\mathbf{y}_k = o(\mathbf{x}_k). \tag{1.7}$$

The most common and simplest approach is to take a **linear combination** of some of (or all) the elements of vector $\mathbf{x}_k$:

$$y_k^j = \sum_{i=1}^{M} w_i^j x_k^i + w_0^j, \tag{1.8}$$

with $M \leq n$ the number of vector elements for the output layer, while the set $\{w_i^j\}$ represents the weights of the output element j. Notice that we added the bias term w_0^j mentioned after Eq. (1.4).

Let us consider temporal input sequences of length L. Since we are dealing with a supervised learning scheme, training a linear output reduces to simply perform a linear regression between a target sequence $\bar{\mathbf{y}}$ and the output sequence $\mathbf{y}$ of the reservoir, where a target sequence represents known examples of the solution of the task at hand. The problem to solve is $\bar{\mathbf{y}} = X\mathbf{w}$, where X is the $L \times (M+1)$ reservoir matrix (we added a column of ones for the offsets w_0^j) and $\mathbf{w}$ is the $(M+1) \times d$ weights matrix. A direct solution can be written as

$$\mathbf{w} = X^+ \bar{\mathbf{y}}, \tag{1.9}$$

where X^+ is the Moore-Penrose pseudoinverse of matrix X [76]. While the computation of X^+ shows high numerical stability, it could become memory-expensive for large matrices. The usual approach is to compute the least square solution minimizing the deviation between output $\mathbf{y}$ and target $\bar{\mathbf{y}}$, given by the euclidean norm $||\bar{\mathbf{y}}_k - (X\mathbf{w})_k||$. For example, the LAPACK library uses the singular value decomposition for that purpose [83]. Different strategies to perform the linear regression could bring some numerical instabilities, but regularization techniques can always be adopted, such as **ridge regression** [76]. In this case, the solution is expressed as

$$\mathbf{w} = (X^\top X + \alpha I)^{-1} X^\top \bar{\mathbf{y}}, \tag{1.10}$$

where the ridge parameter $\alpha \geq 0$ is tuned as another hyperparameter of the system.

The simplification of the training brings very important consequences, which are the essence of RC: first, the training is **faster** compared with the case of conventional RNNs, providing a comparable performance with a more efficient computation [66, 67]. Second, it allows us to solve several tasks in **parallel** for a given

input sequence: since we do not tune the reservoir parameters during training, we can construct an output layer for each task that we want to solve and use them in parallel. On the contrary, if we train an RNN for a given specific task, the training of the network for a second task would interfere with the first one, probably forgetting the former. Finally, not tuning the reservoir layer implies that it can be directly **implemented in hardware** [84], such as in electronic circuits [80], photonics [85], optoelectronics [86], with soft robotics [81], with origami [87], or even with a bucket of water [88]. More details about physical implementations will be provided in Sect. 1.3.5.

RC is not the only ML field that exploits dynamical systems to process input information without fine-tuning of the hidden layer. An **extreme learning machine** (ELM) is the approach where the target task is non-temporal, not requiring the memory of the dynamical system [89]. The input and target are not sequences in time but just static collections of data. An ELM is distinct from RC in that the state of the substrate is solely determined by its corresponding input. We can represent this with the following map:

$$\mathbf{x}_l = T(\mathbf{s}_l), \tag{1.11}$$

where T represents here the substrate dynamical map and the index l denotes different instances of the input. Chapter 2 contains a more detailed exposition of RC and ELM approaches and deals with the particularities of their quantum version.

1.3.3 Information processing capacity

In order to compare different unconventional computing approaches, high-level information processing metrics are needed. In particular, dynamical systems, either classical or quantum, can be driven by external temporal signals while they process the input information. Benchmark temporal tasks are usually employed as a test bed to compare the performance of different RC implementations. These tasks can either be reproducing functions of past inputs, such as the short-term memory (STM) [90] and nonlinear autoregressive moving average (NARMA) [50] tasks, or predicting the future values of an input sequence, as for chaotic time series [68] and stochastic processes [91].

Ways in which temporal information is processed can be estimated, and in fact, a task-independent characterization of the linear and nonlinear memory capabilities of an RC system can be carried out. In 2012, Dambre et. al [92] introduced the **information processing capacity** (IPC) with this purpose. The original theory guarantees that the total computational capability of a dynamical system is bounded by the number of linear independent variables that we use for the output layer. Besides, this bound can be only saturated when the system has fading memory.

Let us define the necessary steps to arrive at the mentioned result. The dynamical systems that are considered are in discrete time, such that $\mathbf{x}_k$ represents the degrees of freedom of a dynamical system at time step k. The dynamical system is driven by an external input signal $\mathbf{s}_k \in D_m$, which in principle can have any dimensionality. In general, D_m will be a compact subset of $\mathbb{R}^m$. We consider that

we can access only a finite number M of degrees of freedom from the dynamical system, keeping a resemblance with real experiments.

The dynamical system equation will be of the RC type as in Eq. (1.6) and the input sequence is generated from a uniform distribution. The latter simplifies the mathematical setting but also brings an important point: any measured structure during the characterization will be only due to the dynamical system, the inputs will not have any correlation structure. Then, the measurement protocol is the following. The system is initialized and a number of washout steps τ_{wo} is waited to erase the initial condition information. Then, the input sequence and the M degrees of freedom are recorded for L time steps, and we denote X as the matrix that will contain the M dynamical variables at each time, as defined in the previous section.

To later introduce the IPC, we define the coefficient that characterizes the capabilities of a system to reproduce a target sequence, known as **capacity**:

$$C_L(X, \mathbf{y}) = 1 - \frac{\min_{\mathbf{w}} \mathrm{MSE}_L(\mathbf{y}, \bar{\mathbf{y}})}{\langle \bar{\mathbf{y}}^2 \rangle_L}, \tag{1.12}$$

where $\mathbf{y}$ and $\bar{\mathbf{y}}$ are the prediction and target sequences, and $\mathbf{w}$ is the weight vector of the output layer. The output layer is a linear function, as it is usually done in RC. The MSE of Eq. (1.2) is now redefined as the cost function $\mathrm{MSE}_L(\mathbf{y}, \bar{\mathbf{y}}) = \frac{1}{L}\sum_{k=1}^{L}(y_k - \bar{y}_k)^2$ and the bracket $\langle\rangle_L$ denotes the temporal average for sequences of length L. The correlation coefficient of Eq. (1.12) measures how good is the estimation of our dynamical system to approximate a target. The performance is bounded between 0 and 1 indicating either a null or a perfect approximation (Proposition 3 in [92]).

The main theorem of [92] states that the maximum capacity of a dynamical system, i.e. the sum of contributions for a given set of target functions, is bounded by the number of degrees of freedom:

$$\lim_{L\to\infty} \sum_{s=1}^{S} C_L(X, \bar{\mathbf{y}}) \leq M, \tag{1.13}$$

where we require an infinite-length input sequence and S target functions. The target functions are here defined as a set of orthogonal functions in a Hilbert space of functions of the input. A crucial consequence of the results is that the bound of Eq. (1.13) can be saturated under the conditions of linear independent degrees of freedom and fading memory (plus other mathematical assumptions like infinite initialization time τ_{wo}, the evaluation over a complete set of orthogonal target functions and finite fourth-order moments for both the dynamical system variables and the target functions). For more details about the demonstration, see Theorem 7 in the original reference [92] and the supplementary material.

In practice, the total computational capacity of a system, the IPC, can be computed as follows. First, you choose a set of orthogonal functions for the targets. In our case, we always took Legendre polynomials as we consider unidimensional

uniformly distributed inputs. The targets are constructed as

$$\bar{y}_k = \prod_i \mathcal{P}_{d_i}[s_{k-i}]. \tag{1.14}$$

The functions $\mathcal{P}_{d_i}$ are the **Legendre polynomials** of degree d_i, whose arguments are the inputs with delay i. In this way, we evaluate a set of linear and nonlinear temporal functions of the input. The product of polynomials is further constrained by the condition $\sum_i d_i = d$, i.e. we define all possible targets up to a maximum delay of polynomials equal to d. In general, this maximum delay does not need to be a large number, and in all our numerical experiments the maximum degree ever taken was $d = 9$ [19, 20]. Then, we evaluate all the possible target functions up to delay d fixing the maximum delay $i_{\max}$ that we want to explore. Setting this value is based on overestimating individual capacities: a finite input sequence does not allow obtaining the exact theoretical value of Eq. (1.12). Then we impose a threshold that allows us to cut off the maximum delay $i_{\max}$. More details about the numerical implementation can be found in Chapter 3.

As we said, the appeal of the IPC is that it quantifies the linear and nonlinear memory of a reservoir in a rather task-independent manner. However, it is true that it might not be straightforward to establish a direct connection between the IPC of the reservoir and the performance on specific tasks. An approach to establishing this connection is to compute the IPC of the task itself, identifying the required capabilities. Some examples can be found for the NARMA2, NARMA10, and PAM tasks [93, 94]. Recently, a different method has been proposed to tackle this problem [95]. In this approach, the normalized mean-square error (NMSE) of time-invariant and fading memory target tasks is predicted based on the different contributions to the IPC. The main assumption is that the input for a specific task follows the same probability distribution that is used to evaluate the IPC (as the uniform distribution we use), something that might not hold for most of real tasks. But the motivation of this work is clear: instead of doing a hyperparameter optimization for each specific task, it can be carried out only once for the IPC contributions. This is in fact the main motivation of the work presented in Chapter 3. We computed the IPC for the first time in a QRC system, finding the linear and nonlinear contributions that characterize the performance of a quantum spin model in terms of the different hyperparameters.

More extensions of the IPC formalism have been carried out. In [93], the IPC has been generalized to systems that are not time-invariant. The temporal information processing capacity (TIPC) allows the computation of linear and nonlinear processing contributions of systems that do not forget their initial conditions. This quantity has been already studied in the QRC context with quantum circuits [94]. In an even more recent work [96], the IPC framework is generalized to account for the statistics of a finite number of measurements in the computation of output observables.

1.3.4 Universal approximation property for reservoir computing

In order to establish the quality of a family of reservoirs, it is helpful to introduce the **universal approximation property** (UAP). The UAP implies the existence of a

member of our family of functions that can approximate, with arbitrary precision, the target we intend to reproduce. More formally, UAP means that a proposed family of functions is dense in the space of target functions. This question has resulted in well-known findings for classical machine learning paradigms such as FFNNs, indicating that they can be viewed as universal approximators [97–99]. In the RC context, UAP has been demonstrated for fading memory functionals (a concept that we will introduce later) in different setups with either continuous-time [82, 100–102] or discrete-time [91, 103, 104] dynamics. The requirements that guarantee the UAP are a good road map to understand the mathematical background of many ML techniques, but notice that they do not answer the practical question of *which is the set of hyperparameters that solve a given task*.

In the case of RC, the main conditions for UAP are usually summarized in the echo state [77] and fading memory properties [105]. The **echo state property** (ESP) means that the recurrent relation Eq. (1.6) has a unique solution for each input sequence, and it is only determined by the input history, discarding any possible dependence on the initial condition. To be more precise, given an input sequence space subset $V_m \subset (\mathbb{R}^m)^{\mathbb{Z}}$, a reservoir state-space subset $V_n \subset (\mathbb{R}^n)^{\mathbb{Z}}$ and a reservoir map T, we say that T has the (V_m, V_n)-ESP if for each input sequence $\mathbf{s} \in V_m$ there exists a unique sequence $\mathbf{x} \in V_n$ such that Eq. (1.6) holds. Then, a reservoir map with the (V_m, V_n)-ESP defines an input-output filter $U^T : V_m \to V_n$ that maps each input sequence to its corresponding state-space solution:

$$U^T(\mathbf{s})_k = \mathbf{x}_k. \tag{1.15}$$

From now on we will simply refer to the ESP instead of the (V_m, V_n)-ESP, and the subsets (V_m, V_n) will be clear from the context. Similarly, the reservoir filter can be defined for left-infinite sequences, where $\mathbf{s} = (\ldots, \mathbf{s}_{-1}, \mathbf{s}_0) \in V_m \subset (\mathbb{R}^m)^{\mathbb{Z}_-}$ and $\mathbf{x} = (\ldots, \mathbf{x}_{-1}, \mathbf{x}_0) \in V_n \subset (\mathbb{R}^n)^{\mathbb{Z}_-}$. Reservoir filters with ESP are causal and time-invariant (Proposition 2.1 in [103]), and as such, a bijection with functionals can be established [103, 104]: let $V_m \subset (\mathbb{R}^m)^{\mathbb{Z}}$ and $V_m^- \subset (\mathbb{R}^m)^{\mathbb{Z}_-}$, a functional $H : V_m^- \to \mathbb{R}^n$ defined as $H(\mathbf{s}) := U(\mathbf{s})_0$ fully determines the filter $U : V_m \to (\mathbb{R}^n)^{\mathbb{Z}}$. The functionals will define the elements of the RC families that we use for the study of the UAP.

The **fading memory property** (FMP) is closely connected to the ESP. It is present when two input sequences that are close in the recent past produce outputs that are close in the present. This is a stronger condition than the usual definition of continuity and, as we will see below, it also implies the erasure of initial conditions from the system. Let us recall the definition of continuity in metric spaces. A function between metric spaces $f : \mathcal{X} \to \mathcal{Y}$ is continuous if for any $x \in \mathcal{X}$ and any $\epsilon > 0$ there exists a $\delta(x, \epsilon) > 0$ such that for any $x' \in \mathcal{X}$ it holds that

$$d(x, x') < \delta(x, \epsilon) \Rightarrow d(f(x), f(x')) < \epsilon, \tag{1.16}$$

where $d(\cdot, \cdot)$ is the metric. The FMP extends this notion of continuity to a metric sequence space with a weighted norm. A weighted norm is a decreasing sequence $w : \mathbb{N} \to (0, 1]$ with $w_0 = 1$ and $\lim_{k\to\infty} w_k = 0$. The weighted norm

$||\cdot||_w$ on $(\mathbb{R}^m)^{\mathbb{Z}_-}$ is defined as

$$||\mathbf{s}||_w := \sup_{k\in\mathbb{Z}_-}\{w_{-k}||\mathbf{s}_k||\}. \tag{1.17}$$

We define the spaces $l_-^w(\mathbb{R}^m)$ and $l_-^w(\mathbb{R}^n)$ as the Banach spaces formed by the elements of $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively with finite weighted norm $||\cdot||_w$ (see Appendix A.2 in [103]) . Now we can apply the continuity condition. Let w being a weighting sequence, $V_m \subset l_-^w(\mathbb{R}^m)$ and $V_n \subset l_-^w(\mathbb{R}^n)$, a filter U has the FMP with respect to w if the map $U : (V_m, ||\cdot||_w) \to (V_m, ||\cdot||_w)$ is continuous. That is, for any $\mathbf{s} \in V_m$ and any $\epsilon > 0$ there exists a $\delta(\mathbf{s}, \epsilon) > 0$ such that for any $\mathbf{s}' \in V_m$ it holds that

$$||\mathbf{s} - \mathbf{s}'||_w < \delta(\mathbf{s}, \epsilon) \Rightarrow ||U(\mathbf{s}) - U(\mathbf{s}')||_w < \epsilon. \tag{1.18}$$

Equations (1.15) and (1.18) define in a rigorous way the key features of a RC system. Furthermore, the ESP and FMP definitions can be extended to the entire reservoir system with an output layer. Let us define $\mathbf{y}_k = o(\mathbf{x}_k) \in \mathbb{R}^d$ as a continuous function of the reservoir variables. We call $U_o^T(\mathbf{s})_k := o(U^T((\mathbf{s})_k) = \mathbf{y}_k$ the entire reservoir filter, and since it is causal and time-invariant by construction, it has an associated functional $H_o^T := o(U^T((\mathbf{s})_0) = \mathbf{y}_0$.

A convenient sufficient condition that simultaneously provides the ESP and the FMP can be expressed in terms of contractivity of the RC map, and it holds for any weighting sequence w when considering compact subsets $D_n \in \mathbb{R}^n$ and $D_m \in \mathbb{R}^m$ as state and input spaces respectively (Theorem 3.1 and Proposition 2.11 in [103]). Given subsets $D_n \subset \mathbb{R}^n$ and $D_m \subset \mathbb{R}^m$, a reservoir map T is strictly contractive if there exists $0 < r < 1$ such that

$$||T(\mathbf{x}, \mathbf{s}) - T(\mathbf{x}', \mathbf{s})|| \leq r||\mathbf{x} - \mathbf{x}'|| \tag{1.19}$$

for all $\mathbf{x}, \mathbf{x}' \in D_n, , \mathbf{s} \in D_m$ and a given norm $||\cdot||$.

Further requirements generally depend on the universality statement one intends to prove, such as the **separability condition** and **polynomial algebra condition** [104]. Separability refers to the ability of RC systems to discriminate inputs and therefore requires the existence of elements of a given RC family that enable to differentiate any pair of inputs. Let us define a compact metric space E and $C(E)$ the set of real-valued continuous functions on E. We say that a subset $A \subset C(E)$ has the separation property if for any distinct pair of points $a, b \in E$, there exists a function $f \in A$ that fulfills $f(a) \neq f(b)$. In the RC case, separability is satisfied by an RC family if for any pair input sequences which differed in the past, $\mathbf{s}_1 \neq \mathbf{s}_2$, there exists a reservoir functional such that $H_o^T(\mathbf{s}_1) \neq H_o^T(\mathbf{s}_2)$. An RC family forms a polynomial algebra if it is closed under addition and product of its elements. Let $o_i \in C(D_{n_i})$, $T_i : D_{n_i} \times D_m \to D_{n_i}$, $i \in \{1, 2\}$ and $\lambda \in \mathbb{R}$. Then, product and linear combination are defined as [104]:

$$\begin{aligned} &H_{o_1}^{T_1} \cdot H_{o_2}^{T_2} = H_o^T, \text{ with } o := o_1 \cdot o_2 \in C(D_{n_1} \times D_{n_2}). \\ &H_{o_1}^{T_1} + \lambda H_{o_2}^{T_2} = H_{o'}^T, \text{ with } o' := o_1 + \lambda o_2 \in C(D_{n_1} \times D_{n_2}), \end{aligned} \tag{1.20}$$

where $T : (D_{n_1} \times D_{n_2}) \times D_m \to (D_{n_1} \times D_{n_2})$ is given by

$$T((\mathbf{x}_{1k}, \mathbf{x}_{2k}), \mathbf{s}_k) := (T_1(\mathbf{x}_{1k}, \mathbf{s}_k), T_2(\mathbf{x}_{2k}, \mathbf{s}_k)), \tag{1.21}$$

for any $(\mathbf{x}_{1k}, \mathbf{x}_{2k}) \in D_{n_1} \times D_{n_2}$, $\mathbf{s}_k \in D_m$ and $k \in \mathbb{Z}_-$. All these conditions are the ingredients of the Stone-Weierstrass theorem, which reads as follows [104, 106]: Let E be a compact metric space. If $A \subset C(E)$ is a subalgebra that contains the constant functions and separate points of E, then A is dense in $C(E)$. Of course, this is only a sufficient route to prove UAP of RC systems. Different strategies might be required and different conditions may arise when one is faced with a different setting, such as when one cannot make use of the polynomial algebra condition as with ESNs [103], or when using the L^p norms criteria [91].

1.3.5 Physical reservoir computing

RC is an amenable technique for hardware implementations, generalizing initially introduced algorithms (see Fig. 1.4 and Ref. [107] for several examples). In fact, a specific term was coined to differentiate software applications from a variety of experimental implementations known as **physical reservoir computing** (PRC) [107, 108].

In this context, it is crucial the dynamical system choice. It is usually claimed that if the dynamical system is complex enough, it might be a good candidate for the reservoir layer. However, some prerequisites should be fulfilled to ensure a successful computation [107, 109]. First, we need **reproducibility** of the input-output relation, i.e., the same input sequence should always produce the same output sequence. This, although may look naive, it is a rather strong condition. For instance, it requires the ESP to avoid any dependence on the initial condition of the system. The reproducibility relation also implies some degree of stability against experimental noise, especially in our context of analog computation. But although noise may seem a problem at first, it can be shown that in some situations we can control its effect during the training [80, 110]. Second, it is usually required a **high dimensional reservoir space** [111]. High dimensionality facilitates the separation of inputs for classification tasks and improves the finding of spatiotemporal dependencies of inputs in prediction tasks. Third, if we want our reservoir to be as general purpose as possible, it must exhibit **rich nonlinear dynamics**.

Many types of dynamical systems have been proposed for PRC. Some representative examples may be found in photonics [84], spintronics [118], mechanical systems [81, 119], nanomaterials [120] and quantum systems [113]. See [107, 109] for a comprehensive list of implementations. The goal of exploring such a variety is, on the one hand, trying to find physical systems that are able to perform complex enough computations, and on the other hand, pushing them to the limit to extract some advantage from their particular physical properties. In fact, speed and energy efficiency have been two of the main goals in the latest years. Photonic and electronic implementations are becoming the best exponents of this trend [84, 109], but new physical systems could enter into the race.

Indeed, quantum systems are becoming an interesting playground for these machine learning techniques [18]. The interest in quantum systems in RC is at

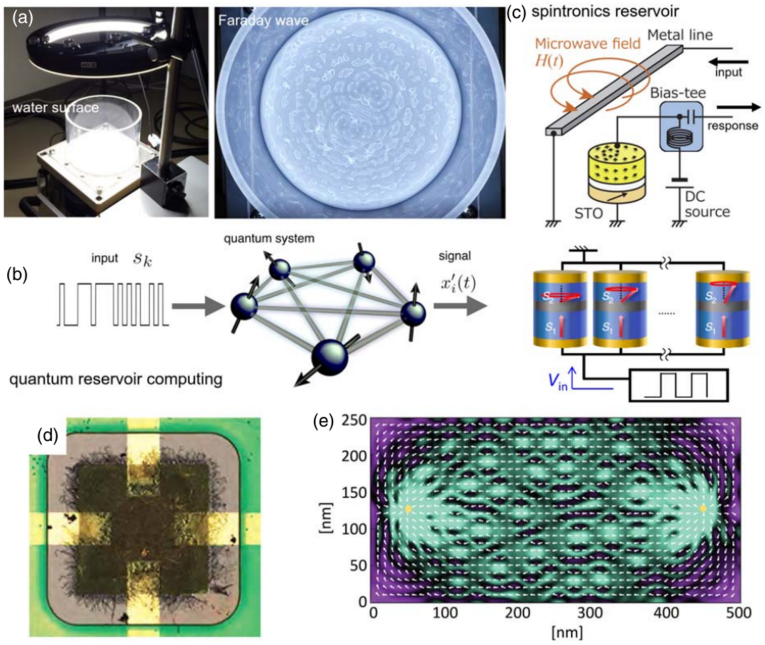

FIGURE 1.4: Different types of physical reservoir computers. (a) A physical liquid state machine that exploits Faraday waves [112]. (b) Quantum reservoir composed of quantum spins [113]. (c) Different spintronics reservoir. The upper scheme shows a reservoir that exploits vortex-type spintronics [114] while the lower one exploits spatially multiplexed magnetic tunnel junctions [115]. (d) Turing B-type atomic switch networks [116]. (e) Skyrmion network embedded in frustrated magnetic films [117]. Figure reprinted from [107].

least four-fold: i) exploit the **large number of degrees of freedom** that Hilbert space can exhibit for few-particle systems; ii) the possibility to implement them in **current experimental devices**; iii) the **possible quantum advantages** that could be found respect to other physical models; and iv) extend the RC techniques for **quantum input processing and tasks** in fully quantum architectures. In the next section, we will introduce quantum systems as reservoir computers, exploring their implications and exposing the state-of-the-art in the field.

1.4 Quantum reservoir computing

Quantum reservoir computing (QRC) extends the RC framework to the quantum regime. In its most basic conception, QRC represents the case where the reservoir is a quantum dynamical system. Then, a QRC state-space transformation is defined by:

$$\begin{cases} \rho_k = T(\rho_{k-1}, \mathbf{s}_k), \\ \mathbf{y}_k = o(\rho_k). \end{cases} \tag{1.22}$$

Here input sequences $\mathbf{s}$ are real-valued, i.e. classical, and the output layer o maps quantum states to real values. Both assumptions could be relaxed considering quantum input and output [18]. Quantum states are represented as density matrices $\rho_k \in \mathcal{S}(\mathcal{H})$, with $\mathcal{S}(\mathcal{H})$ being the space of Hermitian, positive semi-definite and trace-one operators. The symbol $\mathcal{H}$ represents the Hilbert space of the quantum substrate. The input codification and dynamics of the reservoir are gathered in a single quantum channel T. Quantum channels are linear maps $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ that are **completely positive and trace preserving** (CPTP), where $\mathcal{B}(\mathcal{H})$ is the space of bounded operators that act over the Hilbert space $\mathcal{H}$. We say that T is positive when it maps positive semi-definite operators to positive semi-definite operators. Completely positive maps are those positive maps which, when extended to a larger space using the tensor product $T \otimes I_l$, with I_l the identity map in dimension l, also yield a positive map for any $l \in \mathbb{N}$. Finally, trace-preserving maps keep the same trace of operators after their application.

Linear maps are CPTP if and only if they can be written with a **Kraus decomposition** [121], i.e., there exists a set of operators $\{K_i\}_{i\in X}$ such that for all $A \in B(\mathcal{H})$:

$$T(A) = \sum_{i\in X} K_i A K_i^\dagger, \tag{1.23}$$

where $\sum_{i\in X} K_i^\dagger K_i = I$ and X is an index set of cardinality at most d^2, with $d = \dim(\mathcal{H})$. A natural way to insert the input dependence of the reservoir map equation (1.22) is using the same Kraus decomposition:

$$T(\rho_{k-1}, \mathbf{s}_k) = \sum_{i\in X} K_i(\mathbf{s}_k)\rho_{k-1}K_i^\dagger(\mathbf{s}_k). \tag{1.24}$$

Choosing a convenient quantum channel T brings us back to the physical reservoir choice problem because quantum reservoir dynamics should be complex enough to guarantee successful computations. First, few-particle quantum systems can easily provide high dimensionality, which makes them amenable for QRC. Besides, linear independence of observables, which allows exploiting the high dimensionality, can be obtained with the right choice of the quantum model, for example avoiding symmetries. These points are met for several QRC implementations, and in particular, will be shown to be true for the QRC spin model studied in this thesis in Chapters 3, 4, and 6.

Second, nonlinearity could be a more delicate matter since quantum dynamics is linear per se. Nevertheless, input codification, together with the reservoir dynamics and observable election, can determine the nonlinear input-output transformation [122]. Chapter 5 explicitly tackles this question for our QRC spin model and the harmonic oscillator model proposed in [123]. Third, we also require the reproducibility of the input-output relation. This implies that stochastic reservoir dynamics could not be desired. How do we conciliate this with quantum mechanics? We will consider average observables in the output layer such that the effects of quantum measurements are mitigated and the input-output relation is guaranteed. However, we do not discard the possibility of employing stochastic reservoir dynamics as an information-processing tool in future research. Furthermore,

as seen in the previous section, RC requires ESP and FMP as fundamental ingredients for a reservoir to be useful. Then, these conditions imply reservoir dynamics that becomes independent of the initial conditions. Some degree of dissipation must be introduced into the system in order to observe input-dependent dynamics, meaning that unitary dynamics alone is not enough for QRC. We will approach this question in Chapter 7 in a very general QRC framework.

We move now to the description of the three RC layers as depicted in Fig. 1.3: reservoir, input, and output.

1.4.1 Quantum reservoir

In most cases, a quantum reservoir will represent the mathematical model that processes the input information. But as discussed in Sect. 1.3.5, hardware approaches are also considered in RC, and quantum systems are not an exception. Then, depending on the context, the quantum reservoir will refer to either the **quantum physical substrate** or the **mathematical model**. This dichotomy becomes relevant if a given quantum model may be simulated in different physical platforms, such as with qubits, distinguishing between "physical" reservoirs and "model" reservoirs.

With respect to physical substrates, a multitude of platforms have been envisioned to work as reservoirs since the beginning of the field. Section 2.4 in Chapter 2 contains a summary of the main ones up to 2021, including NMR in molecules [16], trapped ions [15, 124], quantum circuits [125, 126] and photonic platforms [127, 128]. Along 2022, new platforms have been proposed for QRC, such as Rydberg atoms for quantum spins [129], Josephson mixers as two coupled oscillators [130], and closed loops of optical pulses working as an ensemble [131]. With respect to QRC models, the exploration has grown as well, with single qudits [132] or quantum master equation of quantum spins with tuned losses [133] as examples. The first proposal was based on the transverse-field Ising model [113], and it will be described in detail in Sect. 1.4.4.

Experimental implementations of QRC are already in their infancy. Few published works that use quantum reservoirs to solve temporal tasks have been reported on quantum computers, either using a reset rate as the source of dissipation [126, 134, 135] or harnessing the natural noise of quantum devices [79, 94, 125]. Yet, as commented before, a multitude of other platforms are good candidates as quantum reservoirs and experimental realizations are expected in the near term [17].

Before moving to the input and output layers, we notice that in some cases reservoir and input layers are intertwined and there is not a clear distinction between them in the mathematical model. As discussed in Sect. 1.4.4, one can generally proceed by introducing a specific CPTP map for the input codification (input layer) and a specific CPTP map for information processing (reservoir layer). However, we may find some examples where there is not a boundary line, as in some implementations with quantum circuits [79, 94, 125]. We might also find quantum channels where input and dynamics are merged into one single map. This can happen when the input is fed as part of a model hyperparameter [122],

as can be seen for instance in [133]. Under these circumstances, we will still denote the quantum model as the reservoir, bearing in mind that it includes the input codification.

1.4.2 Input and output layers

A quantum reservoir has to be accompanied by a proper input encoding and output extraction because both input and output are restricted by the physics of the reservoir. In the seminal work of Fujii and Nakajima [113], classical inputs are codified in the state of one qubit, and the exact expected values of observables of the system are used for the construction of the output layer. We will apply the same scheme as the basis of our work in Chapters 3 to 6. The ancillary codification has been continued in several articles, either with quantum spins models [19, 20, 136–142], quantum circuits [126, 134] or continuous-variable systems [110, 123, 131].

Other input codifications have been proposed in the literature, either introducing classical information through the variation of reservoir parameters [122, 129, 130, 132, 133, 143, 144], quantum gates [79, 94, 125, 135, 145] or directly using quantum states as inputs [146–150].

Classical output information is required when the target is classical. Then, although there might be several options for the input codification, an output layer for classical tasks is enforced to be constructed by measurements over the reservoir. Generally speaking, quantum measurements modify the state of the observed system, while the measurement outcome (and therefore the state after a measurement) is given with a certain probability. The basic formalism about quantum measurements is introduced in Sect. 1.5. Under these circumstances, the number of measurements and even the type of measurement become fundamental ingredients at the moment of designing a QRC system (see Chapter 6). However, QRC is not limited to deal with classical outputs, as shown by some recent works [148–150].

One of the main motivations for using quantum substrates for RC is the presence of a large number of degrees of freedom in few-elements quantum systems. A large number of degrees of freedom increases the chances to improve the performance. In this context, multiplexing techniques have been proposed in the literature to try to maximize the information extracted from the available degrees of freedom. They consist on increasing the number of output variables either by sampling the RC system more frequently or adding more reservoir layers. On the one hand, **temporal multiplexing** harnesses the complex dynamics of quantum many-body systems to multiply the number of output observables by a factor equal to the number of snapshots that we take from the dynamics [113]. Remember that the RC output layer was defined in Eq. (1.8) for a number of output observables M that are sampled only once after each input injection. In temporal multiplexing, these M output observables are sampled V times between two input injections and this is repeated for all inputs, obtaining a total of MV observable nodes for the output layer. Given an observable $\langle O \rangle$, the snapshots at input step k are evenly divided as $\langle O \rangle (k\Delta t + v\Delta t/V)$, with $1 \leq v \leq V$. Temporal multiplexing can significantly increase the performance of QRC models, although

it could require a large number of experimental resources. Besides, there is a limit in the size of V given by the moment at which two consecutive snapshots become linear dependent. In that situation, increasing V would not introduce new information into the output layer. An implementation of this technique and further details will be found in Chapter 3. On the other hand, **spatial multiplexing** focuses on exploiting several reservoir layers at the same time, parallelizing the processing of information [136]. The latter, although may seem more experimental-friendly than temporal multiplexing in terms of resources (using many copies of the system), requires non-identical reservoir layers. Otherwise, observables from each reservoir would be linearly dependent and there would not be any extra contribution to the full (combined) output layer.

1.4.3 Universal approximation property considerations for quantum reservoir computing

Another important aspect of the QRC research is about theoretical results of the UAP. In the case of qubits, sufficient conditions for the ESP and UAP were provided constraining the norm of the quantum dynamical map [139]. These results were limited to the input codification proposed by [113], but a variation in terms of a family of quantum circuits was guaranteed to have the UAP as well [126], further extended in [151]. Later, a similar definition of ESP is introduced in [138], where the focus moves to the eigenvalues of the dynamical map. With respect to other physical platforms, the UAP has also been proven for networks of linear oscillators with Gaussian states [123]. In this work, we showed that by encoding the input in different families of single-mode Gaussian states, like thermal and squeezed, a polynomial readout ensures all nonlinear input combinations and separability can be proven.

Chapter 7 has been our main contribution to the theory of QRC with finite-dimensional quantum systems and classical inputs. This work provides necessary and sufficient conditions for obtaining operational reservoirs. We consider operational QRC systems to be those that fulfill the ESP and FMP and that guarantee an input-dependent mapping for an arbitrarily long input sequence. Conclusions about the design of QRC systems can be drawn from these findings, and we provide several examples to support our conclusions. We remark that operational reservoirs are not necessarily useful, as will be evident for the model we introduce here below in terms of its hyperparameters.

1.4.4 Fujii and Nakajima reservoir model

From chapters 3 to 6 we will focus on reservoir layers based on qubit networks evolving unitarily under the transverse-field Ising model. The **natural dynamics** of the closed quantum systems in the Schrödinger picture is governed by

$$\rho(t) = e^{-iHt}\rho(0)e^{iHt}, \tag{1.25}$$

where H is the Hamiltonian, ρ is the reservoir state, and we are setting $\hbar = 1$.

As our work is mainly developed with qubits, it is convenient to provide some basic definitions to also set the notation. A **qubit** is a two-level quantum system, which lives in a 2-dimensional complex Hilbert space $\mathbf{C}^2$. Its state can be represented as a linear combination of linear independent bounded operators acting on $\mathbf{C}^2$. These operators could be chosen to be the identity matrix I, and the three Pauli matrices σ^z, σ^x and σ^y:

$$\rho = \frac{1}{2}(I + a_z\sigma^z + a_x\sigma^x + a_y\sigma^y), \tag{1.26}$$

where the a_i's are just the projections of the state ρ in the directions of the different Pauli matrices, i.e., the inner product in the Hilbert space known as Hilbert-Schmidt product: $a_i = \langle\sigma^i\rangle = \mathrm{tr}(\sigma^i\rho)/2$. In fact, the state of a qubit can be represented in a 3-dimensional space known as the Bloch sphere, where each component a_i represents each one of the space coordinates. A convenient (and conventional) basis set to describe the matrix elements of the state ρ are the eigenstates of σ^z, such that we can write the density matrix as

$$\rho = \rho_{00}\left|0\right\rangle\left\langle 0\right| + \rho_{01}\left|0\right\rangle\left\langle 1\right| + \rho_{01}^*\left|1\right\rangle\left\langle 0\right| + (1-\rho_{00})\left|1\right\rangle\left\langle 1\right|, \tag{1.27}$$

where we already applied the normalization of the state ($\mathrm{tr}(\rho) = 1$) and the hermiticity of the matrix.

The matrix formulation can be extended to an arbitrary number of qubits, enlarging the dimension of the Hilbert space through the tensor product. In the case of N qubits, the multi-qubit state is represented by

$$\rho = \frac{1}{2^N}\sum_{i=1}^{4^N} a_i B_i, \tag{1.28}$$

where $\{B_i\}$ forms an orthogonal basis in the space of bounded operators that act over our new Hilbert space $\mathbf{C}^{2N}$. The 4^N operators that we usually take are constructed as all the possible combinations of tensor products between the basic operators $\{I, \sigma^z, \sigma^x, \sigma^y\}$ of each qubit. An example could be $B_i = \sigma_1^x \otimes \sigma_2^y \otimes I_3 \otimes \cdots \otimes \sigma_N^z$, and a_i would be computed as $a_i = \langle B_i\rangle = \mathrm{tr}(B_i\rho)$. This produces $4^N - 1$ linear independent operators (disregarding the identity) that are observables. It means that, in principle, a quantum system of N qubits can offer $4^N - 1$ degrees of freedom to perform computations in our RC framework. Finally, we can also describe matrix elements of the multi-qubit state ρ with the 2^N simultaneous eigenstates of all of the $\{\sigma_i^z\}$.

Having defined the elementary constituents of our RC architecture based on qubits, we need to introduce the dynamical model that will allow the processing of information. Part of the work of this thesis gravitates around a very specific Hamiltonian: **the transverse-field Ising model**. Introduced in the context of QRC by Fujii and Nakajima [113], this model has been a constant link between the first works in the field and our contributions. The former proposal reads as

$$H = \sum_{i>j=1}^{N} J_{ij}\sigma_i^x\sigma_j^x + h\sum_{i=1}^{N}\sigma_i^z, \tag{1.29}$$

where N is the number of qubits, h is an homogeneous external magnetic field and J_{ij} is the coupling strength between qubits i and j. These couplings are usually given by a random uniform distribution over an interval $[-J_s/2, J_s/2]$, where J_s denotes the maximum absolute coupling strength.

Disorder in the couplings helps to avoid any redundancy in the dynamics of the observables. For example, if $J_{ij} = J_s > 0$ i.e. all couplings are identical, the dynamics of each qubit becomes identical, hindering the processing capability. Since our goal is to exploit as many observables as possible, we consider instead a random distribution of the couplings, following the common convention of disordered connectivity matrix of traditional RC. However, one could try to tune different coupling configurations to see which one improves results with respect to specific problems. These new configurations could be artificial, like star or ring networks, while other examples could be naturally found in experimental setups, like in long-range interactions in ion-trap experiments [140]. With respect

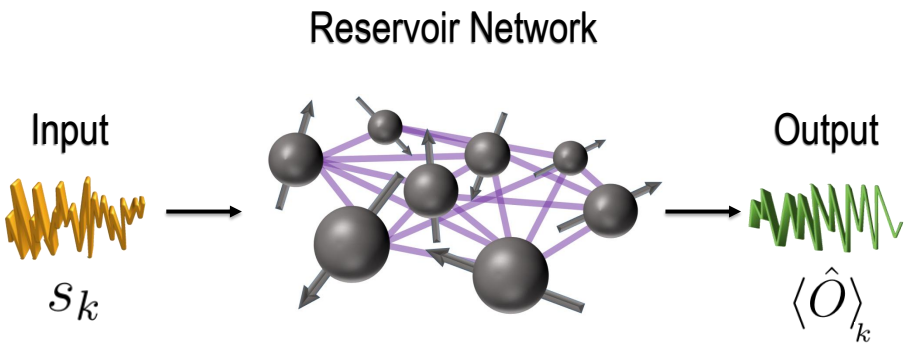

FIGURE 1.5: Scheme of a quantum spin model reservoir. The input layer consists in feeding the input to one of the spins, rewriting its state. Then, the unitary dynamics of the model processes the input information, extracting the output through some of the observables of the system. Figure reprinted from [21] and Chapter 5.

to the physics of Eq. (1.29), this Hamiltonian possesses a **parity $\mathcal{Z}_2$ symmetry**. Intuitively, this symmetry means that the Hamiltonian is invariant under a global spin flip in the x-direction. Mathematically, the parity operator $P = \prod_i^N \sigma_i^z$ commutes with the Hamiltonian ($[P, H] = 0$), and the shape of the eigenstates will depend on the ratio h/J_s. If $h = 0$, the eigenstates are just all the possible configurations of spins aligned in the x axis ($|n\rangle = |\leftarrow\rightarrow\rightarrow \ldots\rangle$). When $h/J_s << 1$, the eigenstates look like cat states as the superposition of states aligned in the x axis: $|n_\pm\rangle \simeq \frac{1}{\sqrt{2}}(|\leftarrow\rightarrow\rightarrow \ldots\rangle \pm |\rightarrow\leftarrow\leftarrow \ldots\rangle)$, where $\pm$ indicates the parity of the eigenstate. In the other extreme, when $h/J_s >> 1$ the Hamiltonian becomes effectively the term $h\sum_{i=1}^N \sigma_i^z$, being the total spin in the z direction a nearly conserved quantity. Then, the eigenstates are very close to spins aligned in the z axis ($|n\rangle \simeq |\uparrow\uparrow\downarrow\uparrow \ldots\rangle$).

Chapters 3 and 6 work with Eq. (1.29) as the Hamiltonian of our spin reservoirs. In chapter 4 we proposed an extension to explore more dynamical regimes,

later exploited in Chapter 5. The extended Hamiltonian is inspired by the ion-trap experiments of [14] and relaxes the assumption of an homogeneous external magnetic field:

$$H = \sum_{i>j=1}^{N} J_{ij}\sigma_i^x\sigma_j^x + \frac{1}{2}\sum_{i=1}^{N}(h + D_i)\sigma_i^z, \tag{1.30}$$

where D_i is a local magnetic field drawn from a random uniform distribution in the interval $[-W, W]$, being W the disorder strength. The introduction of local disorder is one of the routes to the phenomenon of many-body localization, which will be explained in Sect. 1.6.2.

Let us now introduce the reservoir and input injection maps. Here we adopt the Fujii and Nakajima framework following an **amplitude encoding scheme**. As illustrated in Fig. 1.5, a unidimensional input s_k is fed to a single qubit, which we will call qubit one. The state of this qubit is reinitialized every time step as

$$|\psi_{s_k}\rangle = \sqrt{1 - s_k}\,|0\rangle + \sqrt{s_k}\,|1\rangle\,, \quad s_k \in [0, 1], \tag{1.31}$$

which is represented with the density matrix $\rho_1 = |\psi_{s_k}\rangle\langle\psi_{s_k}|$. This is a convenient encoding and variations are proposed in Chapter 5, finding different non-linear responses in terms of the codification. Input injection of ρ_1 corresponds to "rewriting" this density matrix, leading to a full reservoir state updated with the input information:

$$\rho_k' = \rho_1 \otimes \mathrm{tr}_1(\rho_{k-1}), \tag{1.32}$$

where ρ_k' is the updated reservoir state and tr_1 is the partial trace over the first qubit. There are some strategies to implement this input protocol:

- Equation (1.32) can be approximated by a strong local dissipation over the first qubit followed by a quantum rotation gate [133]:

$$\rho_1 \otimes \mathrm{tr}_1(\rho_{k-1}) \approx R_{y1}(s_k)e^{\mathcal{D}_1\Delta t_d}\rho_{k-1}, \tag{1.33}$$

  where $e^{\mathcal{D}_1\Delta t_d}$ is a dissipative Markovian dynamical map that sends the first qubit to the ground state for $\Delta t_d \to \infty$, and $R_{y1}(s_k)$ is a rotation gate in the y direction of the first qubit that introduces the input.

- By performing an unselective projective measurement over the first qubit (see Sect. 1.5), we are actually implementing the partial trace operation $\mathrm{tr}_1(\rho_{k-1})$. Since unselective measurements would discard the outcome of the first qubit measurement, we could not know its state to prepare a rotation gate. However, we can prepare an auxiliary qubit with state ρ_1 and perform a SWAP operation between this ancilla and the first qubit to exchange their states. This method requires the preparation of a new ancilla for each input codification.

- Another possibility is to perform the measurement over the first qubit and register the outcome. Depending on this outcome, a rotation that encodes the input is prepared. In this situation, expected values of observables would be computed by averaging over several measurements at each time step, following a similar scheme to Ref. [126].

Finally, we let the system, i.e. the full quantum network, evolve under its natural dynamics during a Δt time to process the information:

$$\rho_k = e^{-iH\Delta t}\rho_1 \otimes \mathrm{tr}_1(\rho_{k-1})e^{iH\Delta t}. \tag{1.34}$$

Equation 1.34 defines the QRC map (whose CPTP properties can be easily demonstrated). As we will see in our numerical experiments, Eq. (1.34) provides the ESP, necessary for the optimal performance of the model. We note that unitary dynamics by itself cannot produce a dissipation that erases the initial conditions. It is the interplay between the partial trace and unitary dynamics of the previous time step that makes it compatible with the RC requirements. In particular, $e^{-iH\Delta t}$ creates entanglement between the ancilla qubit (qubit one) and the rest of the reservoir, and the operation of taking partial trace becomes an irreversible process when there are correlations between the two subsystems, erasing information.

As we will be dealing with classical targets, we will construct the **output layers with the expected value of observables**, using a subset of the degrees of freedom of our quantum system. In the limit of infinite measurements (no statistical error) and no measurement backaction, observables are computed as

$$\langle O \rangle (t) = \mathrm{tr}\left(O\rho(t)\right). \tag{1.35}$$

This choice can be justified if the experimental protocol consists of a large number of measurements at each time step, with the corresponding repetition of the dynamics up to the measurement instant (see the restarting and rewinding protocols in Chapter 6). Another possibility is to perform ensemble measurements over many copies of the system such that backaction is negligible, as it is done in NMR experiments [16].

The output layer is constructed then as a linear combination of observables:

$$y_k = \sum_{i=1}^{M} w_i \langle O_i \rangle (k\Delta t) + w_0, \tag{1.36}$$

where M indicates the size of the subset of observables. The most common choice will be the local spin projections $\langle \sigma_i^x \rangle$, $\langle \sigma_i^y \rangle$ and $\langle \sigma_i^z \rangle$ for $1 \leq i \leq N$, but we will also consider two spin correlations $\langle \sigma_i^a \sigma_j^b \rangle$ in some chapters. The explicit input dependence of Eq. (1.36) is shown in Chapter 5, where an analysis of the nonlinear input-output relation is carried out for several encoding maps. In particular, we demonstrate that the nonlinear response of the output layer depends on the chosen observables, the number of qubits used for the input injection, and the dynamical regime of the reservoir model.

In this section, we have dealt with closed quantum systems, that is, those that are not connected to an external environment or any measurement apparatus. We remark that none of our works will include an environment, but Chapter 6 deals specifically with the problem of applying a measurement process. The motivation to describe the measurement process in QRC is to depart from the ideal situation of infinite projective measurements, where a significant amount of resources is usually required to estimate the output layer if no ensemble measurement is

available. We propose a more realistic scenario where weak measurements are used to balance the effect of quantum measurements (backaction and finite statistics) with respect to the information we extract from the system, keeping a good performance with a reasonable amount of resources.

1.5 Quantum measurements

The previous section introduces the reservoir model that we employed in Chapters 3 to 6. However, Chapter 6 also presents QRC experimental protocols for which quantum measurements are a basic ingredient. For this reason, we introduce here the quantum measurement formalism, which is fundamental for the statistical interpretation of quantum mechanics. The measurement process introduces two aspects to quantum mechanics experiments: the change in the state of measured systems, known as **backaction**, and the **finite statistics of observations**. Here, we will focus on quantum measurement schemes that allow the monitoring of a quantum reservoir driven by a sequential input.

1.5.1 Projective measurements

The **projective** (sometimes denoted as ideal) **measurement** scheme of most quantum mechanics textbooks relies on the treatment of Born and Von Neumann [152] and the later extension by Lüders [153]. Let us start by supposing we want to measure a quantum system $\mathcal{S}$, which is represented by a finite-dimensional Hilbert space $\mathcal{H}_\mathcal{S}$, and a state ρ. We are interested in measuring a particular observable of the system named O, which we consider discrete. Observables are described by Hermitian operators ($O = O^\dagger$) and have a spectral decomposition,

$$O = \sum_j o_j \Pi_j, \tag{1.37}$$

where Π_j are called the projectors on the eigenbasis of O with eigenvalues o_j. If the eigenvalues are non-degenerate, the projector is a rank-1 operator defined as $\Pi_j = |o_j\rangle \langle o_j|$, where the orthonormal eigenvectors $|o_j\rangle$ form a basis. Otherwise, degeneracy has to be taken into account and the projector becomes $\Pi_j = \sum_{l=1}^{d_j} |o_j, l\rangle \langle o_j, l|$, where d_j accounts for the degeneracy of eigenvalue o_j. The sum of all degeneracies must add up to the dimension of the Hilbert space, $\sum_j d_j = d_\mathcal{S}$, with $d_\mathcal{S} = \dim(\mathcal{H}_\mathcal{S})$. Projectors fulfill the orthonormal relation $\Pi_i \Pi_j = \delta_{ij} \Pi_i$ and the completeness relation

$$\sum_j \Pi_j = I. \tag{1.38}$$

Given an initial state ρ for system $\mathcal{S}$, the projective postulate states that the only possible result of measuring the observable O on system $\mathcal{S}$ is one of the

eigenvalues o_j, with probability $P_j = \mathrm{tr}(\rho\Pi_j)$, leaving the system after measurement in the state

$$\rho'_j = \frac{\Pi_j \rho \Pi_j}{P_j}, \tag{1.39}$$

where the prime refers to the perturbed state by the measurement and the probabilities are normalized to 1, $\sum_j P_j = 1$, given the completeness relation of Eq. (1.38). The updated state of Eq. (1.39), known as the **conditional state**, is the projection of ρ to one of the eigenstates of O, and the measurement is known as **selective**. If a measurement is performed but the result is ignored, named **non-selective measurement**, then the final **unconditional state** is an average over all possible outcomes:

$$\rho' = \sum_j \rho'_j P_j = \sum_j \Pi_j \rho \Pi_j. \tag{1.40}$$

The expected value of observable O can be easily derived from the previous definitions:

$$\langle O \rangle = \sum_j o_j P_j = \mathrm{tr}(\rho \sum_j o_j \Pi_j) = \mathrm{tr}(\rho O), \tag{1.41}$$

Equation (1.41) is the building block of Eq. (1.35) when considering projective measurements. But the expected value of observables can be obtained from more general measurements.

1.5.2 Generalized measurements

Projective measurements are the simplest case of quantum measurements, but this formalism has many limitations. First, the number of possible outcomes is limited by the dimension of the Hilbert space. It could be often desired to have more outcomes than the dimension of the Hilbert space while keeping positivity and normalization of probabilities distributions. Second, we may find that the final state of our measured system is not an eigenstate of the observable we are measuring, difficulting the well-known repeatability of projective measurements. Consider as an example the position detection of a photon in a silver screen, which destroys the photon in the process and makes it impossible to repeat the measurement. Third, one never measures directly the system. There is always a chain of interactions between the system of interest and other systems, such as the measurement apparatus or the environment. However, even considering the interaction between different agents during the measurement, we need to introduce a cut in the chain at a certain point to introduce a projective measurement and be consistent with the quantum mechanics postulates [154].

All these issues can be formally addressed by relaxing the assumptions on the operators that describe the measurement process. Let us consider a discrete set $\mathcal{M}$ of possible outcomes $m \in \mathcal{M}$. The extension to a set of continuous values is straightforward. The space $\mathcal{M}$ can be considered as a sample space where the elements m are the events. Let us define a positive operator E_m, called **effect**, which satisfies the completeness condition:

$$\sum_{m \in \mathcal{M}} E_m = I. \tag{1.42}$$

The measurement outcome m will be a random number given by the probability distribution $P_m = \mathrm{tr}(\rho E_m)$, such that, together with the completeness condition, the total probability is normalized, $\sum_{m\in\mathcal{M}} P_m = 1$.

For the case of a selective measurement over a state ρ, the conditional state is given by

$$\rho'_m = \frac{\Phi_m(\rho)}{P_m}, \tag{1.43}$$

where the **operation** $\Phi_m(\rho)$ depends on the effect E_m. $\Phi_m(\rho)$ is assumed to obey the consistency condition to ensure the normalization of density matrices:

$$\mathrm{tr}(\Phi_m(\rho)) = \mathrm{tr}(\rho E_m), \tag{1.44}$$

in such a way that $\Phi_m(\rho)/P_m$ is a CPTP map. For non-selective measurements, the unconditional state is

$$\rho' = \sum_{m\in\mathcal{M}} \rho'_m P_m = \sum_{m\in\mathcal{M}} \Phi_m(\rho). \tag{1.45}$$

One can check the von Neumann-Lüders scheme of projective measurements is a special case of the above setting. Indeed, the effect would correspond to a projection, i.e. $E_m = \Pi_m$, where the operation over the measured state is given by $\Phi_m(\rho) = \Pi_m \rho \Pi_m$.

A natural generalization of the projection example is given by an operation map like

$$\Phi_m(\rho) = \Omega_m \rho \Omega_m^\dagger, \tag{1.46}$$

with the corresponding effect $E_m = \Omega_m^\dagger \Omega_m$. The linear operator Ω_m still satisfies the completeness condition as $\sum_{m\in\mathcal{M}} \Omega_m^\dagger \Omega_m = I$.

This concept of generalized measurement leads to the more general idea of **positive operator-valued measure** (POVM), which associates effects E_m with measurement outcomes m. The most general effect that one can construct is the following:

$$E_m = \sum_l \Omega_{ml}^\dagger \Omega_{ml} \leq I, \tag{1.47}$$

where the operation over the measured state is

$$\Phi_m(\rho) = \sum_l \Omega_{ml} \rho \Omega_{ml}^\dagger. \tag{1.48}$$

The construction of this general case is based on the definition of CPTP maps introduced in Sect. 1.4 and its Kraus decomposition (see Eq. (1.23)), where l accounts for the number of Kraus operators. Further details about generalized measurements can be found in Refs. [154, 155].

1.5.3 Indirect measurements

So far we have not given any hint about the nature of the measurement itself. We have not commented on how the measurement is performed, what kind of apparatus is used, what distinguishes measurement from other types of interactions,

etc. The **indirect** or ancilla **scheme** goes in this direction. The **ancilla** system, also called meter, probe, or pointer that we will denote by $\mathcal{A}$, is a detection device that will be connected with the system we want to measure $\mathcal{S}$. We will only consider a measurement chain with one extra element because the difference with adding further stages is negligible when the ancilla undergoes a rapid decoherence process [154]. Our indirect scheme is hence composed of three elements: the system under measurement $\mathcal{S}$ with Hilbert space $\mathcal{H}_\mathcal{S}$, the ancilla system $\mathcal{A}$ connected to $\mathcal{S}$ which lives in a Hilbert space $\mathcal{H}_\mathcal{A}$, and a classical apparatus by which a projective measurement is performed over the ancilla.

The aim of the indirect scheme is to obtain information about $\mathcal{S}$ from the correlations built up between system $\mathcal{S}$ and ancilla $\mathcal{A}$ during their interaction by measuring the ancilla. The preparation of the measurement follows three steps. First, before the interaction between the system and the ancilla starts at $t = 0$, the ancilla is prepared in a well-defined state $\rho_\mathcal{A}$, while the system is in a state $\rho_\mathcal{S}$. Second, the interaction between the system and the ancilla starts at $t = 0$ and finishes at some $t = \tau > 0$, before the classical apparatus performs the projective measurement over the ancilla. Finally, the classical apparatus reads out the ancilla, and the von Neumann-Lüders measurement postulate is applied.

In the following, we will derive the expressions of effects and post-measurement states of the indirect measurement scheme. We start by assuming that system and ancilla are completely independent of each other at $t = 0$ (before interacting), such that the initial global state of system and ancilla is $\rho_{\mathcal{SA}}(0) = \rho_\mathcal{S} \otimes \rho_\mathcal{A}$. The Hamiltonian of the total system is

$$H(t) = H_\mathcal{S} + H_\mathcal{A} + H_{\text{int}}(t), \tag{1.49}$$

where $H_\mathcal{S}$ and $H_\mathcal{A}$ describe the free evolution of system and ancilla respectively, and $H_{\text{int}}(t)$ is the interaction Hamiltonian. The latter is assumed to vanish when outside of the time interval $[0, \tau]$. The unitary operator of the time evolution for this interval is given by

$$\mathcal{U} \equiv \mathcal{U}(\tau, 0) = \mathcal{T} \exp\left(-i \int_0^\tau dt H(t)\right), \tag{1.50}$$

where $\mathcal{T}$ denotes here the chronological time ordering. After applying the unitary time evolution, the composite system state is

$$\rho_{\mathcal{SA}}(\tau) = \mathcal{U} \rho_\mathcal{S} \otimes \rho_\mathcal{A} \mathcal{U}^\dagger. \tag{1.51}$$

An important consequence of this evolution is that system and ancilla become entangled and cannot be written as product states anymore. Once the unitary operator has acted, an ancilla measurement produces an outcome. The classical apparatus performs a projective measurement over the ancilla by probing an observable R (which acts in the Hilbert space $\mathcal{H}_\mathcal{A}$). We here assume that R has a non-degenerate and discrete spectrum, writing its decomposition as $R = \sum_m r_m |r_m\rangle \langle r_m|$, where eigenstates $|r_m\rangle$ form an orthonormal basis. Notice that although the measurement is performed over the ancilla, the system will be influenced too because of the entanglement. We also assume that the evolution

of the system and the ancilla is negligible during the measurement of the classical apparatus, considering the measurement as instantaneous. The probability of obtaining an outcome r_m is then

$$P_m = \mathrm{tr}((I_{\mathcal{S}} \otimes |r_m\rangle \langle r_m|)\rho_{\mathcal{SA}}(\tau)) = \mathrm{tr}(\mathcal{U}^\dagger (I_{\mathcal{S}} \otimes |r_m\rangle \langle r_m|)\mathcal{U}(\rho_{\mathcal{S}} \otimes \rho_{\mathcal{A}})), \quad (1.52)$$

where the trace is taken over the whole Hilbert space $\mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{A}}$. Denoting the partial traces over $\mathcal{H}_{\mathcal{S}}$ and $\mathcal{H}_{\mathcal{A}}$ as $\mathrm{tr}_{\mathcal{S}}$ and $\mathrm{tr}_{\mathcal{A}}$ respectively, the last expression can be rewritten as

$$P_m = \mathrm{tr}_{\mathcal{S}}(E_m \rho_{\mathcal{S}}), \quad (1.53)$$

where

$$E_m \rho_{\mathcal{S}} = \mathrm{tr}_{\mathcal{A}}(\mathcal{U}^\dagger (I_{\mathcal{S}} \otimes |r_m\rangle \langle r_m|)\mathcal{U}(\rho_{\mathcal{S}} \otimes \rho_{\mathcal{A}})). \quad (1.54)$$

The projection postulate, given a selective indirect measurement, produces the following global conditional state:

$$\rho'_{\mathcal{SA}m} = \frac{(I_{\mathcal{S}} \otimes |r_m\rangle \langle r_m|)\rho_{\mathcal{SA}}(\tau)(I_{\mathcal{S}} \otimes |r_m\rangle \langle r_m|)}{P_m}. \quad (1.55)$$

If we only focus on the system of interest $\mathcal{S}$, the final state is

$$\rho'_{\mathcal{S}m} = \mathrm{tr}_{\mathcal{A}}(\rho'_{\mathcal{SA}m}) = \frac{\langle r_m|\mathcal{U}\rho_{\mathcal{S}} \otimes \rho_{\mathcal{A}}\mathcal{U}^\dagger|r_m\rangle}{P_m}, \quad (1.56)$$

where we notice that the unconditional state should read $\rho'_{\mathcal{S}} = \sum_m \rho'_{\mathcal{S}m} P_m$.

Introducing the spectral decomposition of the initial density matrix of the ancilla,

$$\rho_{\mathcal{A}} = \sum_l p_l |\psi_l\rangle \langle \psi_l|, \quad (1.57)$$

the operation map that acts over the system can be written as

$$\Phi_m(\rho_{\mathcal{S}}) = \sum_l \Omega_{ml} \rho_{\mathcal{S}} \Omega_{ml}^\dagger, \quad (1.58)$$

where we define the operators $\Omega_{ml} := \sqrt{p_l} \langle r_m|\mathcal{U}|\psi_l\rangle$ and the effects are given by Eq. (1.47), i.e. $E_m = \sum_l \Omega_{ml}^\dagger \Omega_{ml}$. In the particular case of an initial pure state for the ancilla, $\rho_{\mathcal{A}} = |\psi\rangle \langle \psi|$, the operators and effects become $\Omega_m = \langle r_m|\mathcal{U}|\psi\rangle$ and $E_m = \Omega_m^\dagger \Omega_m$ respectively. We further notice that the ancilla states $|\psi_l\rangle$ are not, in general, eigenstates of the observable R, but superpositions of such eigenstates.

1.5.4 Continuous-variable indirect measurement

In the previous section, we introduced the indirect measurement scheme for the simple case of an observable R of the ancilla with a discrete spectrum. We now generalize to the case of an **observable R with a continuous spectrum**, $R := \int r dr |r\rangle \langle r|$, introducing a more realistic implementation closely connected with the measurement model described in Chapter 6. The observable R is going to be now the momentum operator in a specific direction. Therefore, the classical measurement apparatus will measure a certain momentum component of the ancilla

particle. Our goal is to exploit the indirect measurement scheme to measure an observable O of the system $\mathcal{S}$. We take O as a discrete, non-degenerate observable $O = \sum_j o_j \left|o_j\right\rangle \left\langle o_j\right|$. To this end, we will consider an interaction Hamiltonian that couples the observable O to the canonically conjugate position operator of R, which we will denote as Q and that fulfills the commutation relation $[Q, R] = i$ (for $\hbar = 1$). The interaction Hamiltonian reads as

$$H_{\text{int}}(t) = G(t)OQ, \tag{1.59}$$

where we assume that the time-dependent coupling $G(t)$ vanishes outside the interval $[0, \tau]$. We make a further assumption by considering that the interaction time τ is so short that the free evolution $H_{\mathcal{S}} + H_{\mathcal{A}}$ can be neglected over this window of time. The unitary evolution becomes then

$$\mathcal{U} = \exp\left(-i \int_0^{\tau} dt G(t) OQ\right) = e^{-igOQ}, \tag{1.60}$$

where g is the integrated coupling strength. This parameter indicates how strong will be the correlation between the system and the ancilla before the readout. For simplicity, we assume that the initial state of the ancilla is a pure state $\rho_{\mathcal{A}} = \left|\psi\right\rangle \left\langle\psi\right|$. We can comment now that the choice of this H_{int} is based on the fact that the operator $e^{i\lambda Q}$, with $\lambda \in \mathbb{R}$, acts as a translation operator for wave functions like $\phi(p) \equiv \langle p|\psi\rangle$ of the ancilla. Therefore, the operators Ω_p, where p represents a momentum outcome detected by the classical apparatus, are given by

$$\begin{aligned} \Omega_p &= \langle p|\mathcal{U}|\psi\rangle = \langle p|e^{-igOQ}|\psi\rangle = \sum_j \left|o_j\right\rangle \langle p|e^{-igo_jQ}|\psi\rangle \left\langle o_j\right| \\ &= \sum_j \left|o_j\right\rangle \langle p|e^{-igo_jQ}|\psi\rangle \left\langle o_j\right| = \sum_j \phi(p + go_j) \left|o_j\right\rangle \left\langle o_j\right|. \end{aligned} \tag{1.61}$$

It follows that the probability density of obtaining a momentum outcome p is

$$P(p) = \text{tr}_{\mathcal{S}}(E_p \rho_{\mathcal{S}}) = \sum_j \left|\phi(p + go_j)\right|^2 \left\langle o_j\right|\rho_{\mathcal{S}}\left|o_j\right\rangle. \tag{1.62}$$

This measurement scheme illustrates how the pre-measurement interaction shifts the ancilla state by an amount go_j. Reading out the momentum constitutes then an approximate measurement of the observable O by the proximity of p to one of the values go_j (assuming that the initial momentum of the ancilla is zero, $\langle R\rangle(0) = \langle\psi|R|\psi\rangle = 0$). All relevant quantities can be now related to the shifted states. In fact, further details can be obtained if one considers a specific initial wave function, such as a Gaussian probability density:

$$\phi(p) = \langle p|\psi\rangle = \frac{1}{(2\pi\sigma^2)^{1/4}} e^{\frac{-p^2}{2\sigma^2}}, \tag{1.63}$$

where σ is the Gaussian width. This is the starting point of the measurement model of Chapter 6, which is described in more detail in its supplementary information Section II.B.

1.6 Dynamical phase transitions in closed quantum systems

This thesis is devoted to a large extent to the use of complex closed quantum systems as reservoir computers. Interestingly, in the last few years, the statistical mechanics of closed quantum systems has been a very active research field. This is in part due to the interest in studying fundamental properties of many-body atomic systems (like ultracold atoms [156] and trapped ions [157]), motivated by the increasing access to experiments of nearly isolated quantum systems, but also because of the interest to control these fundamental properties, with the development of new quantum technologies in mind as computation or simulation.

In this section, we will introduce two dynamical regimes and the main mechanisms that describe them: the thermal and MBL dynamical phases. We start by describing the thermal regime and how it is understood in closed quantum systems.

1.6.1 Eigenstate thermalization hypothesis

Thermalization in closed quantum systems has been one of the hottest topics in many-body physics during the last years, with pioneering publications even back to the first days of quantum mechanics [158] (original in [159]), but with more recent contributions that settled the grounds of the field [160–162]. However, these latter works were almost forgotten until the work by Rigol et al. [163], in part due to experimental limitations.

Let us now proceed by introducing the concept of thermalization. In thermodynamics, macroscopic physical systems are in thermal equilibrium with an environment when they reach the same temperature after exchanging energy and/or particles, staying in a stationary state. At thermal equilibrium, these physical systems become fully characterized by just a few extensive conserved quantities, such as temperature, chemical potential, etc, suggesting that this process of thermalization is associated with the erasure of information of the system's initial state [164].

The thermodynamic description of macroscopic classical objects lacks a connection with microscopic dynamics until statistical mechanics comes into play. Under this formalism, the macroscopic deterministic dynamics is replaced by a probabilistic description of the microscopic system, which is usually complemented with the **ergodic hypothesis**: at long times, the system visits all accessible microstates with equal probability. By accessible we mean that the dynamics might be constrained by, for example, energy conservation, exploring only a subregion of the total phase space. The formal formulation of this hypothesis is presented as

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t o(t')dt' = \frac{1}{V} \int_V o(x)dx, \tag{1.64}$$

where o is a classical macroscopic observable, t is time and V is the volume of phase space that the system can explore. The left-hand side is equal to the time-independent right-hand side, meaning that the system reaches a stationary

value of the observable after the deterministic macroscopic dynamics. Then, the right-hand side integrates over the volume V of phase space, averaging over the equally weighted accessible microstates. Notice that given that the right-hand side does not depend on the initial conditions, the stationary value neither does, becoming one of the main characteristics of ergodicity. We further remark that this concept introduces thermalization in a weak sense: ergodicity deals with long-time averages of observables, while we are interested in the value of observables at long times.

In the language of quantum mechanics, we will work with thermalization in the strong sense, meaning that the values of observables will approach, at a given time, their equilibrium value predicted by statistical mechanics, and remain close to it at almost all subsequent times. Indeed, closed quantum systems exhibit situations where the expectation value of local observables arrives at a stationary state that agrees with microcanonical predictions, "forgetting" their initial conditions. See Fig. 1.6 for a numerical example. Then, there is an apparent paradox

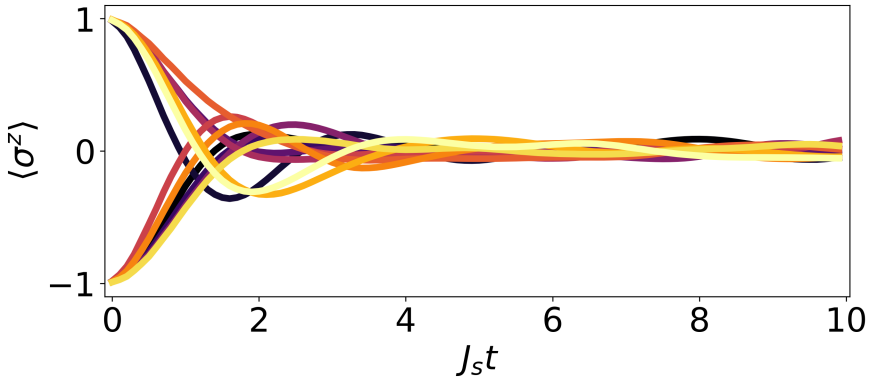

FIGURE 1.6: Dynamics of spin projections in the z direction for the transverse-field Ising model of Eq. (1.30). Parameters are $h = 10$, $J_s = 1$, $W = 0$ and $N = 12$.

between thermalization and unitary dynamics, which cannot erase any information. The explanation of why an isolated quantum system can exhibit thermalization is that smaller subsystems use the rest of the system as if it were a bath. Hence, the information about the initial conditions of the system is hidden in the global operators. In fact, in the thermodynamic limit, this information would be hidden in those degrees of freedom that are not observables, i.e. infinite dimensional Hermitian operators. Therefore, quantum thermalization in closed systems is not about forgetting but hiding the information about the initial condition.

To understand the main ingredients of thermalization in closed quantum systems, we are going to consider a quantum system initially prepared in a pure state $|\psi(0)\rangle$ and evolve it under a time-independent Hamiltonian H. We remark that the following analysis can be straightforwardly generalized to density matrices. The eigenvalues and eigenvectors of H are given by $H|m\rangle = E_m|m\rangle$. The initial state can be decomposed in the eigenbasis of H as $|\psi(0)\rangle = \sum C_m|m\rangle$, where $p_m = |C_m|^2$ is the probability of finding the initial state at the eigenstate $|m\rangle$. The

time-evolved state would be given by

$$|\psi(t)\rangle = e^{-iHt}\,|\psi(0)\rangle = \sum_m C_m e^{-iE_m t}\,|m\rangle\,. \tag{1.65}$$

The time evolution of some observable is then

$$\langle O\rangle\,(t) = \langle\psi(t)|O|\psi(t)\rangle = \sum_m p_m\,\langle m|O|m\rangle + \sum_{m,n\neq m} C_m^* C_n e^{i(E_m-E_n)t}\,\langle m|O|n\rangle\,. \tag{1.66}$$

As commented before, it is considered that an observable thermalizes if, after some transient time, the expectation value of this observable agrees with microcanonical predictions, but we also further require that the temporal fluctuations of the expectation value around the prediction are small at most later times. These two requirements imply that the long-time average accurately describes the value of the expected value $\langle O\rangle\,(t)$ after relaxation time, with a value close to the microcanonical prediction.

Now we face the problem of matching Eq. (1.66) with the previous requirements. The infinite-time average $\langle O\rangle_\infty$ is given by

$$\langle O\rangle_\infty = \lim_{t\to\infty}\frac{1}{t}\int_0^t \langle O\rangle\,(t')dt' = \sum_m p_m\,\langle m|O|m\rangle\,, \tag{1.67}$$

where the second term of Eq. (1.66) averages to zero (considering that there are no degeneracies). How can Eq. (1.67) agree with the microcanonical ensemble if the probabilities p_m are constant in time? Besides, the eigenenergies of many-body systems get exponentially close when we increase their size, and one could need to wait an exponentially long time to observe the second term of Eq. (1.66) disappear, in contradiction with the requirement of thermalization at a finite time (and with experience).

A first solution could be given by the **random matrix theory** (RMT) [165, 166]. Originally developed to understand the spectra of heavy nuclei [167–170], RMT lies at the heart of the so-called field of **quantum chaos** [171]. Under the RMT hood, if a Hamiltonian H were a random matrix generated by one of the ensembles of this theory, diagonal elements $\langle m|O|m\rangle$ would be independent of the eigenstate $|m\rangle$ and the off-diagonal terms $\langle m|O|n\rangle$ would be exponentially small in system size. The eigenbasis $\{|m\rangle\}$ is given by random orthonormal vectors, where its elements (in the computational basis) are well approximated by Gaussian random numbers with unit variance. Let us consider an observable with a discrete non-degenerated decomposition, i.e. $O = \sum_i o_i\,|o_i\rangle\,\langle o_i|$. The RMT ansatz for its matrix elements in the basis $\{|m\rangle\}$ is

$$\langle m|O|n\rangle \approx \overline{O}\delta_{mn} + \sqrt{\frac{\overline{O^2}}{d}}R_{mn} \tag{1.68}$$

where d is the dimension of the Hilbert space, R_{mn} is a random variable with zero mean and unit variance and the upper bar over operators accounts for an average

over random eigenvectors $|m\rangle$, being:

$$\begin{aligned} \overline{O} &:= \overline{\langle m|O|m\rangle} = \frac{1}{d}\sum_i o_i, \\ \overline{O^2} &:= \overline{\langle m|O^2|m\rangle} = \frac{1}{d}\sum_i o_i^2, \end{aligned} \tag{1.69}$$

see [172] for more details. Under these conditions, we would find that observables thermalize in the sense specified above: the long-time average becomes independent of the initial conditions by

$$\langle O\rangle_\infty = \sum_m p_m \langle m|O|m\rangle \approx \overline{O}\sum_m p_m = \overline{O}, \tag{1.70}$$

However, RMT falls short because physical Hamiltonians are not truly random. In particular, the microcanonical ensemble prediction is energy-dependent while $\overline{O}$ is not, and relaxation times are observable-dependent. Hence, RMT does not provide all the detailed information that diagonal and off-diagonal matrix elements of observables contain in real systems.

The most promising explanation to solve this conundrum was provided by the groundbreaking works of Srednicki and Deutsch [160–162]. Srednicki's ansatz is known as the **eigenstate thermalization hypothesis** (ETH), and it can be formulated as the assumption:

$$\langle m|O|n\rangle = O(\overline{E})\delta_{mn} + e^{-S_{\text{th}}(\overline{E})/2} f_O(\overline{E},\omega) R_{mn}, \tag{1.71}$$

where $\overline{E} = (E_m + E_n)/2$, $\omega = E_n - E_m$, and $e^{S_{\text{th}}(\overline{E})}$ is the density of states at energy $\overline{E}$, therefore relating $S_{\text{th}}(\overline{E})$ to the thermodynamic entropy. The terms $O(\overline{E})$ and $f_O(\overline{E},\omega)$ are smooth functions, and R_{mn} is again a random real (or complex) variable with zero mean and unit variance. The ETH ansatz (1.71) differs from the RMT ansatz (1.68) in: i) the energy dependence of the diagonal elements and ii) the envelope function $f_O(\overline{E},\omega)$ that goes with the off-diagonal terms. Then, the former ansatz reduces to the latter when one focuses on a very narrow energy window where $f_O(\overline{E},\omega)$ is constant (realizing that $e^{-S_{\text{th}}(\overline{E})/2} \propto 1/\sqrt{d}$). As we will see, the ETH ansatz is a sufficient condition for thermalization. This, together with the fact that RMT satisfies the ETH ansatz, means that Hamiltonians which follow the RMT are expected to thermalize. Later we will see how to exploit RMT indicators to detect thermalizing Hamiltonians.

Let us now see how ETH implies thermalization and ergodicity. We assume that the expected value of the total energy

$$\langle E\rangle := \langle\psi(0)|H|\psi(0)\rangle = \sum_m |C_m|^2 E_m \tag{1.72}$$

has a quantum uncertainty

$$\delta E^2 := \langle\psi(0)|H^2|\psi(0)\rangle - \langle\psi(0)|H|\psi(0)\rangle^2 = \sum_m |C_m|^2 (E_m - \langle E\rangle)^2 \tag{1.73}$$

that is sufficiently small. This is a natural assumption since for d large, initial states of physical interest fulfill $\delta E \sim E/\sqrt{d}$. Therefore, eigenenergies are found in a small window such that we can expand $O(E_m)$ in powers of $E_m - \langle E \rangle$ as $O(E_m) \approx O(\langle E \rangle) + (E_m - \langle E \rangle)O'(\langle E \rangle) + \frac{1}{2}(E_m - \langle E \rangle)^2 O''(\langle E \rangle)$. Inserting this approximation in Eq. (1.67), we obtain

$$\langle O \rangle_\infty \approx O(\langle E \rangle) + \frac{1}{2}\delta E^2 O''(\langle E \rangle). \tag{1.74}$$

We can also observe that assuming ETH matrix elements, the microcanonical prediction equals $O(\langle E \rangle)$:

$$O_{MC} = \frac{1}{\Omega}\sum{}' O_{mm} \approx \frac{1}{\Omega}\sum{}' O(E_m) \approx \frac{O(\langle E \rangle)}{\Omega}\sum{}' 1 = O(\langle E \rangle), \tag{1.75}$$

where $\sum' = \sum_{\langle E \rangle < E_m < \langle E \rangle + \delta E}^{\Omega}$ and Ω is the number of microstates whose energy lies in the shell $\langle E \rangle < E_m < \langle E \rangle + \delta E$. Then, ergodicity is displayed in the quantum system by the relation $\langle O \rangle_\infty \approx O_{MC}$. On the other hand, the long-time average of temporal fluctuations of $\langle O \rangle$ is

$$\sigma_O^2 = \lim_{t\to\infty} \frac{1}{t}\int_0^t \left(\left(\langle O \rangle (t') \right)^2 - \langle O \rangle_\infty^2 \right) dt' \leq \max_{m \neq n} |\langle m|O|n \rangle|^2 \propto e^{-S_{\text{th}}(\overline{E})}. \tag{1.76}$$

The time fluctuations of $\langle O \rangle$ are exponentially small in system size since $S_{\text{th}}(\overline{E}) \propto N$ is an extensive quantity, where N is the number of particles. This equation tells us that at almost any time the expectation value $\langle O \rangle$ is close to the result of Eq. (1.67). Notice that these fluctuations are the deviation of the expected value, not the experimental measurement fluctuations of the observable. It can be shown that the latter, when taking the long-time average, are approximated by the fluctuations predicted by the microcanonical ensemble [172]. Finally, we can comment on the relaxation time scale. In general, it depends on the observable, the initial state, and the Hamiltonian that drives the dynamics. But as numerical and real experiments show, the relaxation time scales of observables are not exponentially large [172].

While it is not clear yet if the ETH is a necessary condition for thermalization, neither it is known which observables satisfy the ansatz, it is believed that it holds for very general many-body quantum systems with strong interactions and for all their physical observables (i.e. finite-dimensional) [172]. This ansatz has implications for the structure of eigenstates and their entanglement properties. For an eigenstate $|m\rangle$ obeying the ETH, all the observables with support over a small subsystem A will have thermal expectation values, implying that the reduced density matrix $\rho_A = \text{tr}_B(|m\rangle\langle m|)$ (B is the complement of A) is a thermal state.

1.6.2 Many-body localization

On the extreme side, we can find isolated many-body quantum systems that break the ergodicity of the ETH. There are some known routes to escape from it,

such as integrability, characterized by an extensive number of conserved quantities [173] or quantum many-body scars, where a specific type of initial conditions never reach thermal equilibrium because quantum systems can have an extensive number of nonergodic eigenstates [174]. Another possible route is given by **localization**. Phillip Anderson was the first one to notice that wave function localization could be present in noninteracting systems [175], but recently, localization in interacting systems has gained a lot of attention since the publication of the work by Basko et al. [176]. This localization with interactions defines the **many-body localization** (MBL) regime, where the dynamics of local observables retain a memory of their initial conditions, see Fig. 1.7. We notice that MBL is the only known example where ETH can be violated by all the eigenstates in a nonintegrable system. This means that MBL is a phase of matter where localization happens at finite energy density, i.e. far away from the ground state. However, it could happen that transitioning from ergodic to MBL, not all eigenstates localize. The presence of this phenomenon, named mobility edge [176], implies that there is a definite energy density in the spectrum at which the properties of the Hamiltonian eigenstates change, separating localized and ergodic eigenstates. This means that the initial condition determines whether the system thermalizes or remains localized, depending on its energy.

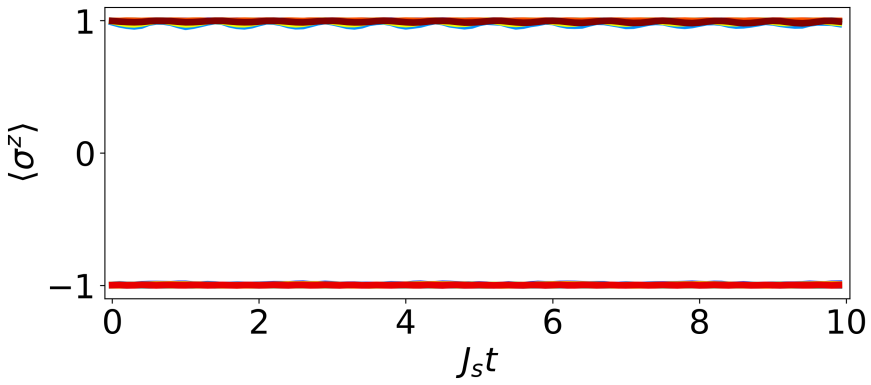

FIGURE 1.7: Dynamics of spin projections in the z direction for the transverse-field Ising model of Eq. (1.30). Parameters are $h = 1$, $J_s = 1$, $W = 100$ and $N = 12$.

The existence of the MBL regime has not only been shown in numerical experiments (see the reviews [177–179] for early contributions), but it has been also theoretically demonstrated for one-dimensional spin models [180, 181] and experimentally shown for a multitude of platforms, such as trapped ions [14], superconducting qubits [182, 183], bosons [184] and fermions [185] in optical lattices, and nuclear spins [186]. However, for $d > 1$ dimensions and long-range interactions, it is not clear if the MBL phase is stable. Indeed, localization would be harder to sustain, not knowing which are the conditions that guarantee the MBL phase stability in the thermodynamic limit. Since our work deals with finite-dimensional quantum systems (used as complex reservoirs), we will be only concerned with those finite systems that exhibit MBL features and get localized up to experimentally relevant time scales.

The usual mechanism to escape ergodicity in the MBL context is through **local disorder**, which introduces a new kind of integrability. And this emergence of integrability is compatible with the observation of memory of initial conditions that we commented on before. The effective theory that we will describe below provides a complete description of the most important properties of the MBL regime, such as the absence of transport, logarithmic dynamics of entanglement, dynamics of local observables, and the breakdown of ETH. Let us introduce it through a simple example, which will help us to gain some intuition. A trivial limit of MBL would be the case of an integrable system, as a chain of quantum spins with Hamiltonian

$$H = H_0 = J \sum_{i=1}^{N-1} \sigma_i^z \sigma_{i+1}^z + \sum_{i=1}^{N} \frac{h_i}{2} \sigma_i^z, \tag{1.77}$$

where N is the number of spins, σ_i^z is the Pauli matrix in the z axis at site i, J is the coupling strength and h_i is a local magnetic field different for each spin, i.e. $|h_i| \neq |h_j|$ for $i \neq j$ (avoiding degeneracy in the spectrum). For simplicity, we assume that h_i are randomly distributed in an interval $h_i \in [-W, W]$, where W is the disorder strength. The Hamiltonian of Eq. (1.77) commutes with all σ_i^z and any product of them. Hence, in this integrable model, there is no thermalization because of the extensive number of conserved quantities that it presents. And independently of the magnitude of disorder W, the eigenstates of Eq. (1.77) are spins aligned in either the parallel or antiparallel direction of the field.

Now, we add a weak nearest-neighbor interaction to the Hamiltonian in the other directions with strength g:

$$H = H_0 + g \sum_{i=1}^{N-1} (\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y). \tag{1.78}$$

This Hamiltonian is known as the Heisenberg model, which is non-integrable and the paradigmatic example in (the numerical analysis of) MBL [187–189]. Equation (1.78) is no longer diagonal in the z basis. Keeping a low value of g, such that $\mathcal{O}(g/W) \ll 1$ with $\mathcal{O}(J/W) \sim 1$, we can construct a set of "dressed" operators τ_i^z such that $[H, \tau_i^z] = [\tau_i^z, \tau_j^z] = 0$. These τ_i^z are called **local integrals of motion** (LIOMs) [190, 191]. They are quasi-local operators whose support decays exponentially with the distance to the corresponding spin i. One can connect each dressed operator τ_i^z with its original integral of motion σ_i^z through a quasilocal unitary transformation: $\tau_i^z = U \sigma_i^z U^\dagger$. A unitary transformation is defined as quasilocal if it can be factorized into a sequence of local unitary operators as $U = \prod_i \cdots U_{i,i+1,i+2}^{(3)} U_{i,i+1}^{(2)}$, such that the long-range unitary operators fulfill $||I - U_{i,\ldots,i+n}^{(n)}||^2 < e^{-n/\xi}$ over the Frobenius norm, where ξ is the localization length scale. This localization scale, which controls the locality of τ_i^z, is what distinguishes the MBL regime from the ergodic one. In principle, one could also define operators τ_i^z for thermal regimes, but they would be highly nonlocal and therefore useless [179].

We can form a complete basis of operators with the new quasilocal operators,

such as the Pauli matrix basis that we employed in Sect. 1.4.4. For that, we need to introduce the quasilocal terms $\tau_i^{x,y} = U\sigma_i^{x,y}U^\dagger$, and now every operator can be decomposed in terms of the operators $\tau_i^{x,y,z}$. Hence, we can exploit this representation to show the general form of an MBL Hamiltonian. Since all τ_i^z are integrals of motion, i.e. $[H, \tau_i^z] = 0$, the Hamiltonian cannot contain the operators $\tau_i^{x,y}$ because they do not commute, i.e. $[\tau_i^{x,y}, \tau_i^z] = U[\sigma_i^{x,y}, \sigma_i^z]U^\dagger \neq 0$. Then, the general expression of an MBL Hamiltonian can be written as:

$$H_{\text{MBL}} = \sum_i \tilde{h}_i \tau_i^z + \sum_{i>j} J_{ij}\tau_i^z\tau_j^z + \sum_{i>j>k} J_{ijk}\tau_i^z\tau_j^z\tau_k^z + \cdots . \tag{1.79}$$

The parameters $\tilde{h}_i$ denote the effective external fields acting over each l-bit (the localized bits, sometimes used in the literature referring to the operators τ_i^z), while the couplings

$$J_{ij} \propto e^{-|i-j|/\kappa}, \; J_{ijk} \propto e^{-|i-k|/\kappa}, \ldots \tag{1.80}$$

decay exponentially with the separation between the l-bits. It can be shown that the interaction length scale κ satisfies the inequality $\kappa^{-1} \geq (\xi^{-1} + \ln 2)/2$ [179].

The emergent integrability of the MBL phase explains the breakdown of ergodicity since local observables retain memory of the initial states, encoded in the initial values of the LIOMs. It also explains many other features of the MBL regime, that in fact, are quantified as deviations from ETH, such as area law entanglement entropy [190, 192], Poisson level statistics of eigenenergies [188, 193], logarithmic growth of entanglement [194, 195] and dynamics of local observables after a quench [187, 194]. We now formally define these concepts, which can be exploited to detect a phase transition between both regimes. The main figure of merit we are going to introduce is Eq. (1.86), used in Chapter 4.

1.6.3 Dynamical phase transition

The differences between the ergodic and the MBL phase can be seen in many indicators. One of the most characteristic features of the ETH ansatz and the MBL phase is the entanglement properties of eigenstates. Let us take an eigenstate $|m\rangle$ and define two partitions A and B of its Hilbert space, such that we can define the reduced state $\rho_A = \text{tr}_B(|m\rangle\langle m|)$, being B the complement of A. If the state $|m\rangle$ could be written as a pure state product, $|m\rangle = |m\rangle_A \otimes |m\rangle_B$, then ρ_A would be pure and uncorrelated from subsystem B. Otherwise, it is a mixed state. A way of quantifying the entanglement between the two partitions is through the **entanglement entropy**:

$$S_{\text{ent}} = -\text{tr}(\rho_A \ln \rho_A), \tag{1.81}$$

which is the von Neumann entropy of subsystem ρ_A. Then, for a pure state ρ_A we find $S_{\text{ent}} = 0$. If mixed, $S_{\text{ent}} \geq 0$.

A random eigenstate $|m\rangle$ that fulfills the ETH will exhibit a thermal density matrix when it is reduced to a sufficiently small subsystem A. Then the entanglement entropy of A equals the thermodynamic entropy (as introduced in Eq. (1.71)) in the thermodynamic limit: $S_{\text{ent}} = S_{\text{th}}$. And since the thermodynamic entropy is an extensive quantity, the entanglement entropy follows a **volume law**

as $S_{\text{ent}} \propto \text{vol}(A)$. It can be also shown that in the case of random vectors in Hilbert space, the average (over random pure states) entanglement entropy is computed as $\overline{S_{\text{ent}}} \approx \ln(d_A) - \frac{d_A}{2d_B}$ with $1 \ll d_A \leq d_B$, where d_A and d_B are the Hilbert space dimensions of partitions A and B respectively [196]. The connection between the volume law and the previous formula for random vectors can be made explicit for a concrete example, as in a unidimensional chain of qubits. Splitting a chain of N qubits into two partitions of equal size, the partition dimensions are $d_A = d_B = 2^{N/2}$. The average entanglement entropy would read

$$\overline{S_{\text{ent}}} \approx \frac{1}{2}(N \ln 2 - 1), \tag{1.82}$$

which is proportional to the chain's length, following the volume law. On the contrary, it has been shown that high energy eigenstates of MBL systems exhibit an **area law** entanglement entropy [190, 192]. This means that the entanglement entropy of A scales proportional to the volume of the boundary ∂A of A. The usual argument to understand this is as if the MBL eigenstates were obtained from a local perturbation that entangles particles between A and B only over a distance ξ away from the boundary ∂A [179, 190]. This local perturbation is supposed to decay exponentially in distance from the boundary.

The entanglement entropy is not the only quantity that can detect differences in the spectrum between phases. The ETH implies a high sensitivity of eigenstates to small local perturbations of the Hamiltonian, and the sensitivity of thermal eigenstates implies **level repulsion** [172]. Level repulsion is in fact one of the main features of the **Wigner-Dyson distribution**, used to describe the level statistics of quantum chaotic systems, and as commented in Sect. 1.6.1, quantum chaos is an indicator of thermalization. The Wigner-Dyson distribution of a 2×2 Hamiltonian has the general form

$$P_{\text{WD}}(w) = A_\beta w^\beta e^{-B_\beta w^2}, \tag{1.83}$$

where $w := E_2 - E_1$ is the energy gap, the coefficients A_β and B_β are given by normalization and β depends on the symmetry of the Hamiltonian. Time-reversal symmetry makes the Hamiltonian a real symmetric matrix, having $\beta = 1$, while $\beta = 2$ means the absence of this symmetry (not shown here, see [172]). Observe that the probability of a vanishing energy gap goes to zero, $\lim_{w\to 0} P_{\text{WD}}(w) = 0$, defining the so-called level repulsion. Generalizing the Wigner-Dyson distribution, one can define an ensemble of matrices drawn from a Gaussian distribution with probability [172, 197]:

$$P_{\text{WD}}(H) \propto e^{-\frac{\beta}{2a^2}\text{tr}(H^2)}, \tag{1.84}$$

where a sets the energy scale. As before, $\beta = 1$ refers to real symmetric matrices, denoted as the Gaussian orthogonal ensemble (GOE), while $\beta = 2$ corresponds to Hermitian matrices, denoted as the Gaussian unitary ensemble (GUE). The nearest-neighbor level-spacing distribution $P_{\text{WD}}(w)$ that one would derive from Eq.(1.84) does not have a closed analytical form. However, it is qualitatively and quantitatively close to the simple case of two levels in Eq. (1.83) [197].

In contrast, for strong disorder in MBL systems the energy level repulsion disappears and the level statistics approaches a **Poisson distribution** [188, 193]. The absence of level repulsion means that it is likely to find energy levels arbitrarily close to each other within an energy interval. This is a consequence of emergent integrability and can be explained as follows. For an extensive number of integrals of motion (as in integrable models) in many-particle systems, the eigenstates behave as independent random variables. Therefore, the distribution of energy level spacings corresponds to an uncorrelated random spectrum, being described by a Poisson distribution [172]. For a mean level spacing set to one, the distribution of energy gaps is

$$P_{\text{Poisson}}(w) = e^{-w}, \tag{1.85}$$

which fulfills the absence of level repulsion $P_{\text{Poisson}}(0) = 1$. The statement that energy eigenvalues behave as a sequence of independent random variables in integrable quantum systems whose corresponding classical counterpart is integrable is known as the Berry–Tabor conjecture [198]. This conjecture describes what is actually seen in many quantum systems, even without a classical counterpart, but one can find examples for which it fails, such as a single particle in a harmonic potential [199].

The current analysis of spectral level spacing for distinguishing ETH from MBL is usually carried out through the so-called **gap ratio** [193]:

$$r_n = \frac{\min(\delta_n, \delta_{n+1})}{\max(\delta_n, \delta_{n+1})}, \tag{1.86}$$

where $\delta_n = E_{n+1} - E_n$ is the energy gap between consecutive levels in the spectrum. It has been shown that the average gap ratio (averaged over random realizations of the Hamiltonian) for a spectrum of Poissonian statistics is $r_{\text{Poisson}} = 2\ln 2 - 1 \approx 0.38$, while in a GOE ensemble of random matrices it was found $r_{\text{GOE}} \approx 0.53$ [200]. Since the transverse-field Hamiltonian of Eq. (1.30) has time-reversal symmetry (described by the GOE ensemble when thermalizing), we will make use of the previous values to classify its dynamical regimes in Chapter 4.

So far we have commented on order parameters that depend on the eigenvalues and their statistics. But we are dealing with dynamical phases: nonequilibrium regimes where physical quantities can exhibit different behavior in time. The most obvious order parameter with time dependence would be the **long-time value of local observables**. Compare for example Figs. 1.6 and 1.7. This feature has been actually employed to experimentally demonstrate dynamical phase transitions, for instance in ion-trap [14, 15] and cold atoms experiments [201, 202]. But other physical quantities can be evaluated in time to detect the ergodicity breaking in MBL [187, 203, 204]. One of the most common ones is the **entanglement entropy dynamics**. The idea is very similar to the eigenstate entanglement entropy detection: we divide the system into two parts, A and B, where B is the complement of A. Generally, one takes half of the system for each partition. Then, the system is usually initialized as a product state, $|\psi(0)\rangle = |\psi(0)\rangle_A \otimes |\psi(0)\rangle_B$. The system evolves under the unitary dynamics and the entanglement entropy is

evaluated at each time t as

$$S_{\text{ent}}(t) = -\text{tr}(\rho_A(t) \ln \rho_A(t)). \tag{1.87}$$

Starting from a product state, $S_{\text{ent}}(0) = 0$ and $S_{\text{ent}}(t) > 0$ for any finite t. The entanglement entropy grows until the system becomes entangled and it saturates. The key point is the temporal dependence of this growth. In the thermal regime, the growth is ballistic, $S_{\text{ent}}(t) \sim t$, while in the localization regime under strong disorder, the growth is logarithmic, $S_{\text{ent}}(t) \sim \ln t$.

Of course, all these differences in entanglement and eigenvalue statistical properties must have an impact on how closed quantum systems process information. An example of this is that systems exhibiting MBL can furnish quantum memories at finite temperature [205], enable quantum batteries [206], or prevent overheating in Floquet systems [207, 208]. Localization of quantum spin models can also enhance the trainability in quantum machine learning [209]. On the other hand, localization has been reported as potentially negative for quantum random walk algorithms [210, 211] or quantum annealing [212, 213], although adiabatic optimization algorithms can be successfully defined around the transition point [214]. The central goal of Chapter 4 is to understand the role of thermalization and localization in the specific scenario of QRC, demonstrating that the optimal information processing abilities in QRC are not only preferred in the ergodic phase but also have significant benefits during the onset of this regime.

1.7 Main original contributions of this thesis

We finish the introduction with a summary of the results of this thesis, presenting the problems that we tackled from a chronological perspective. For this, we will proceed by briefly reviewing the QRC literature, summarizing the outcome of each chapter within the context of publications in the QRC field. We would like to warn the reader that some of the articles discussed here may have a publication date distant from their first appearance in the *arXiv*. Besides, the QELM field has been developed almost in parallel in time with respect to the QRC field, but we will focus on the QRC field in this discussion for the sake of conciseness.

The field of QRC started with the publication of the quantum reservoir spin model of Fujii and Nakajima in 2017 [113], presented in Sect. 1.4.4. This work was a natural continuation of the exploration of classical physical implementations for RC and neuromorphic platforms that had started some years before [75, 109]. The proposal was based on exploiting a transverse-field Ising model as the reservoir, with amplitude encoding as the input mechanism and the expected value of observables for the readout. Then, it immediately followed a series of articles that contributed to extend the understanding of the same model. These contributions include further explorations of the spin network topology [137], the introduction of a deep reservoir scheme through spatial multiplexing [136], and a proof-of-principle QELM experiment with NMR [16]. Soon after these publications, Chen and Nurdin analyzed sufficient conditions for the ESP and FMP in QRC [139], accompanied by numerical experiments of a variation of the model of Sect. 1.4.4.

Up to this point, we realized that all the previous studies only used local spin projections as the observables for the output layer, that is, nobody had explored the possibility of using the rest of the degrees of freedom for computation in QRC. Besides, although some benchmark tasks had been analyzed [113, 137, 139], it was not clear the computational capabilities that these observables could offer in the spin model. Therefore, our first work focused on exploring these capabilities for both local and second-order observables [19] (Chapter 3). For doing this, we took a novel approach by calculating the IPC within the QRC framework, assessing the linear and nonlinear memory capabilities of the system. By analyzing the convergence speed with respect to the input injection time, we also identified the optimal conditions for input driving that would enhance computational speed and performance of the studied model. This investigation was crucial in revealing the intricate nature of the Ising model as a reservoir, demonstrating how its hyperparameters strongly influence its performance and highlighting the benefits offered by the large dimension of the Hilbert space.

Meanwhile, in 2019 Chen et al. had been working on the first IBM experimental proposal for QRC [126, 215], studying the ESP, FMP, and UAP in a quantum circuit platform with a reset rate as the dissipation source. A second article on an IBM implementation with a different quantum circuit design followed soon after [125], characterizing its linear memory capacity. Around the same time, with Nokkala et al. [123] we started to work on the first proposal of a continuous-variable QRC model with Gaussian states, and studied the conditions that provide universality in this platform. Also in 2020, a perspective article appeared including an overview of quantum neuromorphic computing with a mention to QRC [216]. Later that year, Tran and Nakajima continued exploring new ways of improving the performance and scalability of QRC models [138], and two book chapters were released [108]: an introduction to both the QRC field [217] and the NMR experimental platform they previously employed for QELM [218].

All the research mentioned above focused either on heuristically exploring new reservoir architectures or improving the performance of the already existing ones. This is in fact a common trend in the ML research: new algorithms are proposed and optimized in the search for the best results, sometimes postponing the detailed understanding of models. After publishing our first work [19], we decided to fill this gap in the quantum spin model. Reference [20] (Chapter 4) aimed to gain a comprehensive understanding of the relationship between the hyperparameters of the quantum spin model and its performance. Our investigation revealed the relation between these factors and the dynamical regimes of the unitary dynamics. Specifically, we found that thermalization within the unitary dynamics establishes favorable conditions for RC, while localization has a detrimental effect. Furthermore, we observed that at the transition point between the thermalization and localization regimes, there might exist a peak in performance that depends on the specific task being addressed. While we identified the roles of the dynamical regimes for QRC and characterized them in all detail, it is worth mentioning that we were not the only ones thinking about this problem. The first publication of Fujii and Nakajima gives some hints about the dynamical regime dependence of the quantum spin model [113], and Xia et al. released a related work almost at the same time [140].

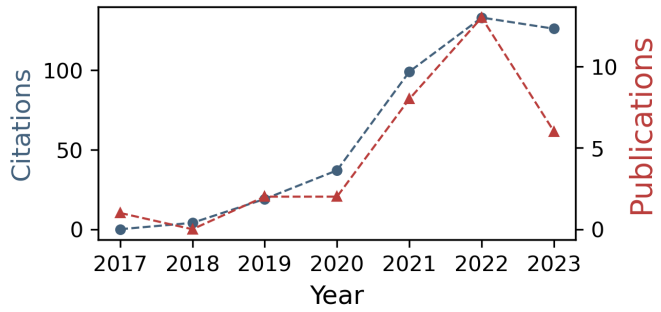

FIGURE 1.8: Number of citations (left axis) and number of publications (right axis) in the field of QRC between the years 2017 and 2023. Data has been obtained by searching "reservoir computing" and "quantum" in Scopus [219]. Preprints are not included and QELM and other unrelated references have been manually removed.

We find a few more publications during 2021, see Fig. 1.8. In Ref. [110], Nokkala et al. continued evaluating the capabilities of continuous variables quantum reservoirs, further extending the applicability of these models to process quantum time series [148]; a few months before, the application of quantum spin reservoirs to learn the tomography of temporal quantum maps had already been proposed [149]; Kalfus et al. explored the impact of the Hilbert space dimension of a single qudit reservoir [132]; Suzuki et al. proposed a new quantum circuit implementation where the natural hardware noise is exploited for QRC [79]. That summer, two review articles were published, Refs. [18] (Chapter 2) and [220]. It is not a coincidence that both publications appeared almost at the same time: the QRC and QELM fields were burgeoning enough to be reviewed, foreseeing some perspective on the future directions that the fields could take. In fact, some review articles with a broader scope had already devoted paragraphs to them [2, 109, 216].

Again, some fundamental questions were not being addressed up to this point. In particular, the origin of nonlinearity in QRC was not completely clear. In a prior study [123], we successfully demonstrated that nonlinearity of a continuous-variable implementation using Gaussian states can be entirely attributed to the input codification. A more general description was provided by Govia et al., in which again the input encoding becomes the determinant factor to provide nonlinearity [122]. But input encoding is not the only ingredient. In the study presented in Ref. [21] (Chapter 5), we derived analytical expressions for the expected values of observables of the Fujii and Nakajima model based on the chosen input codification, revealing their dependence on the observable itself, input preprocessing, and the number of qubits involved in the encoding map. We presented the QRC spin model alongside the harmonic oscillator model from Ref. [123] for the sake of comparison.

While we were working on this, experimental implementations on the IBM quantum computers have been already carried on, but they relied on restarting the protocol at each time step to obtain the expected value of the output observables. Reinjecting the input sequence from the first codification at each time step

yields a time complexity that increases quadratically with the length of the sequence. Since the beginning of 2021, we had been wondering how to implement a more efficient protocol to process time series. We focused on finding an alternative strategy where an external memory of inputs would not be required to be processed once the quantum reservoir is trained. That is, we wanted to propose a scheme where the quantum reservoir could make predictions from a directly injected input sequence, without any storage of the input information. Our proposal was to implement weak measurements over an ensemble of copies of QRC systems [22] (Chapter 6). This approach allows us to keep a balance between performance and the effect of backaction, the retrieved information after each measurement, and the statistical noise of finite sampling.

After a few years of research, the QRC field was becoming attractive to more researchers. Vintskevich and Grigoriev proposed to enhance the performance of quantum reservoirs with spatial multiplexing by connecting them with nonselective measurements [221]; in Ref. [147], Khan et al. analyze the transition between the classical and quantum regime of a Kerr oscillator reservoir, with quantum states as inputs and output measurements determined by the experimental platform; Araiza Bravo et al. studied the construction of a quantum reservoir from an array of Rydberg atoms [129]; Dudas et al. numerically studied coupled quantum oscillators [222] and its implementation in Josephson junctions [130]; Llodrà et al. compared the memory capacity of bosons and fermions with a model inspired by the one presented in Sect. 1.4.4 [223]; García-Beni et al. proposed a photonic platform for online QRC based on an ensemble of reservoirs formed by identical optical pulses recirculating through a closed loop [131]; Sannia et al. explored the memory and forecasting capabilities of quantum dissipative systems with tuneable local losses, given that this family of models possesses the universality property [133]. Also, new studies continued to explore experimental implementations. A hybrid scheme between classical RC and quantum computation scheme for predicting convection flow was experimentally tested [224]; Kubota et al. numerically evaluated the TIPC for quantum circuit reservoir under different hardware noise channels and [94]; Molteni et al. explored the optimization of quantum circuits with reset rate [134]; lastly, Spagnolo et al. demonstrated the viability of optical quantum memristors for performing QRC [17].

Finally, more contributions have appeared in 2023. Götting et al. studied the relation between entanglement, STM capacity, and the effective Hilbert space dimension of the spin model of Sect. 1.4.4 [142]; Motamedi et al. evaluated the performance of a single nonlinear oscillator in a time-series prediction task with respect to a measure of quantumness [144]; Fry et al. extended the previous works on quantum circuits with hardware noise [79, 94] optimizing the very same noise channels [135]; Xia et al. optimized hyperparameters of a quantum reservoir to efficiently solve multiple tasks with the same system [141]; Mlika et al. tackled the problem of predicting the trajectory of mobile users with a quantum circuit implementation [225]. In this context of new research groups and very diverse publications, we published our last work [23] (Chapter 7). A detailed analysis of the requirements that make quantum reservoirs suitable for time-series processing was missing. For finite-dimensional quantum systems with classical inputs

and an infinite number of measurements, we determined that the crucial criterion for a reservoir to be valuable is the presence of strictly contractive dynamics leading to input-dependent fixed points.

Chapter 2

Opportunities in Quantum Reservoir Computing and Extreme Learning Machines

2.1 Motivation and contribution

By the time this work was published, the field of QRC was burgeoning. We decided to make the effort of gathering all the available information up to that moment, analyzing the status of the field, and giving some hints about future lines that could be explored providing a perspective on QRC. Beyond QRC, we also described the state-of-the-art in quantum extreme learning machines (QELM), not devoted to temporal tasks, since both concepts are closely related. In QELM and QRC, quantum dynamical systems are exploited as physical substrates for supervised machine learning tasks, with the common characteristic of no fine-tuning of the quantum systems during training. The main differences between RC and ELM (and their quantum versions) are highlighted in Sect. 2 of the published paper, with emphasis on the presence (RC) or absence (ELM) of memory in the hidden layer. Although they differ in the nature of the target tasks, actually most of the physical substrates proposed for QELM could be used in QRC, and vice versa. Since both QRC and QELM were developed almost at the same time and often not clearly identified in the literature, it made sense to review them together. As a complement to this chapter, Sect. 1.4 updates in some aspects the state-of-the-art of the QRC literature.

Due to the nature of this work and although all the authors contributed to the manuscript, it was necessary to initially divide the workload by different topics. In particular, my contributions to the text are especially reflected in all the sections that are related to spin models. The outcome of this work is the publication of an accessible introduction for newcomers, but also a good reference to consult for experts in the field.

2.2 Published paper

Pere Mujal, Rodrigo Martínez-Peña, Johannes Nokkala, Jorge García-Beni, Gian Luca Giorgi, Miguel C. Soriano and Roberta Zambrini, Opportunities in quantum reservoir computing and extreme learning machines, *Advanced Quantum Technologies*, **4**, 2100027 (2021).

https://doi.org/10.1002/qute.202100027

Chapter 3

Information Processing Capacity of Spin-Based Quantum Reservoir Computing Systems

3.1 Motivation and contribution

The characterization of RC systems is one of the main problems that the RC community has to address when a new physical substrate is proposed. By characterization we mean to understand what are the capabilities of an RC system, being the performance in different tasks a crucial issue in view of applications. This characterization could be related to the more general capacity problem of neural networks, where the committed errors during task executions are estimated in terms of the hyperparameters of the system.

In principle, there is no way of completely characterizing RC systems such that the performance for any arbitrary target task can be anticipated. A usual approach is to examine some benchmark tasks that the community considers appropriate proxies. For instance, dynamical systems with chaotic regimes, such as the Lorenz or Mackey-Glass models are used as indicators of the ability to predict chaotic time series [226, 227], while the short-term memory task is a good indicator of the capacity of a system to remember linear functions of past inputs [90].

In our case, we characterized the memory of the recently proposed QRC model of spins [113], since due to its novelty it was not clear which kind of tasks it could solve, although some benchmark tasks were already analyzed [113, 137, 139]. In fact, we were mainly concerned with two points: 1) which nonlinear tasks it could solve since the form and origin of nonlinearity in QRC were still unclear at that moment; and 2) what were the capabilities of a QRC system beyond the output layer made of single spin observables, as had been constructed in previous works. We decided to characterize the memory of the RC system with the IPC, which quantifies in a fairly objective way the different linear and nonlinear contributions to the memory of an RC system. We computed the IPC in terms of different hyperparameters, such as the input injection rate Δt, the number of spins N, and the external magnetic field h. This, with a previous analysis of the convergence speed in terms of Δt, allowed us to find optimal conditions for the input driving in terms of speed and performance. Finally, we provided different alternatives for the output variables after evaluating the effect of temporal

multiplexing and correlation observables in the output layer. We further notice that this paper pioneered the use of the IPC to characterize the capabilities of a quantum reservoir.

To follow this chapter, Sects. 1.4.4 and 1.3.3 provide an introduction of the basic concepts. Section 1.4.4 contains a detailed description of the model we employed while Sect. 1.3.3 defines the IPC. As the first author, I led this project, performing all the numerical simulations and (with all the authors) contributing to analyze and interpret the results.

3.2 Published paper

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*, 1–12 (2020).
https://doi.org/10.1007/s12559-020-09772-y

Chapter 4

Dynamical phase transitions in quantum reservoir computing

4.1 Motivation and contribution

While preparing the previous chapter's article, we realized that something important was affecting the RC performance when varying hyperparameters like h, J_s, and Δt. For example, the decrease of the IPC in Fig. 8 (a) for a small magnetic field h was unintelligible at that moment. We had the intuition that the saturation of the IPC for large values of Δt was connected with some thermalization time scale. After some tests and revision of the literature, we realized that the value of these hyperparameters determines the dynamical regime of our closed quantum system. Then, the motivation behind this chapter was to contribute to the identification of "good" reservoirs, answering the particular question: what is the role of dynamical phases in QRC?

We studied the case of a transverse-field Ising model, generalizing the Hamiltonian proposed in [113] and exploring its different dynamical regimes. We actually provided a complete phase diagram of this model establishing the presence of ETH, MBL, and spin glass phases when changing the relative strength of the homogeneous magnetic field and the heterogeneous disordered magnetic field. Our conclusions can be extrapolated to any QRC model where unitary dynamics processes the input information. Indeed, the physical interpretation that we provide in terms of dynamical regimes like thermalization or MBL intuitively assures that different dynamical systems within the same phase should provide similar RC properties.

The outcome of this work is understanding the role of dynamical regimes in QRC systems composed of an input protocol map and an entangler unitary dynamics. The thermal regimes of the unitary part provide the appropriate conditions for temporal information processing, while localization regimes like MBL are detrimental. Interestingly, the transition between these regimes can offer an improvement in performance, as shown in Fig. 4. However, as Fig. S5 shows for nonlinear targets, this is not necessarily true for any task.

Section 1.6 offers an introduction to the thermalization of closed quantum systems, MBL, and dynamical phase transitions. This, with the content of Sect. 1.4.4, contributes to comprehending the present chapter. As in the previous chapter, I led this research project by refining the research idea toward dynamical phase

transitions, performing all the simulations, and providing a physical interpretation.

4.2 Published paper

Rodrigo Martínez-Peña, Gian Luca Giorgi, Johannes Nokkala, Miguel C. Soriano and Roberta Zambrini, Dynamical phase transitions in quantum reservoir computing, *Physical Review Letters*, **127**, 100502 (1-7) (2021).
https://doi.org/10.1103/PhysRevLett.127.100502

Chapter 5

Analytical evidence of nonlinearity in qubits and continuous-variable quantum reservoir computing

5.1 Motivation and contributions

The first explanations about the origin of nonlinearity in QRC were not satisfactory. In particular, explicit expressions of the input dependence on the output were missing in QRC, including the Fujii and Nakajima model. With such analytical expressions, we could have a better intuition of which kind of nonlinear functions the models could solve. Indeed, in a previous work, we proved that the nonlinearity of a continuous-variable implementation with Gaussian states was completely determined by the input codification [123]. This allowed us to explain why the proposed Gaussian state implementation could not solve tasks like the parity check with a linear output layer made of first and second moments, since reservoir equations were missing the products of inputs at different time steps. To solve the parity check task, a careful construction of the output layer was required, such that it contained the desired input products. Besides, we observed that different input encodings were responsible for shifting the nonlinear capabilities of the continuous-variable model from completely linear to highly nonlinear.

The input encoding that we used in Ref. [123] is the same ancillary input codification scheme that Fujii and Nakajima proposed in [113]. Therefore, after the experience we gained with the continuous-variable work, we expected to obtain similar conclusions with the qubit model, i.e., that the nonlinearity of the Fujii and Nakajima model should be determined by the input codification (for a linear output layer). However, the input mechanism does not necessarily uniquely determine the nonlinear response of a general QRC system. In Ref. [21] we obtained analytical expressions of the expected values of observables in terms of the input, showing a dependence on the chosen observable, the input preprocessing, and the number of qubits that participate in the encoding map. The QRC spin is presented together with the harmonic oscillator model of Ref. [123] for the sake of comparison. These results are complemented in Chapter 6 clarifying the role of measurement: no further nonlinearities are introduced into the QRC spin model by applying any of the experimental protocols we explored. Furthermore, we show that the dynamical regime of the quantum spin model can play a determinant role, finding that the MBL regime completely neutralizes the input response,

an aspect we have fully characterized in Chapter 4. This is also found in other implementations such as quantum master equations or parameterized unitaries, showing that input codification and reservoir dynamics together determine the nonlinearity [122].

The most relevant introductory section for this chapter is Sect. 1.4, as it contains a general description of quantum reservoirs and input codification. The original idea of this chapter was developed by Dr. Pere Mujal, where explicit analytical expressions of the quantum spin reservoir equations were derived in terms of the input. In this case, my contribution is devoted to the numerical aspects of the spin model section, including the discussion of the analytical results. The learned conclusions about the nonlinear input dependence of the continuous-variable case were also included for the sake of completeness.

5.2 Published paper

Pere Mujal, Johannes Nokkala, Rodrigo Martínez-Peña, Gian Luca Giorgi, Miguel C. Soriano and Roberta Zambrini, Analytical evidence of nonlinearity in qubits and continuous-variable quantum reservoir computing, *Journal of Physics: Complexity*, **2**, 045008 (2021).

https://doi.org/10.1088/2632-072X/ac340e

Chapter 6

Time-Series Quantum Reservoir Computing with Weak and Projective Measurements

6.1 Motivation and contribution

This chapter is the beginning of a new exploration line, that is, the study of experimental protocols for QRC. This work was actually motivated by a crucial question in view of (QRC) implementations of time series processing: how to efficiently monitor and extract meaningful information from a quantum reservoir?

Let us introduce some context to understand the previous question. The classical RC technique is implemented as an online protocol, i.e., after the supervised training, one expects to continuously extract outputs at each time step with no halts in the algorithm. However, online monitoring in quantum systems leads to the backaction of quantum measurements, which introduces two effects. The first one is the memory dissipation that measurements introduce into the reservoir dynamics. In the case of projective measurements, the collapse of an eigenvector of the measured observable will erase most of the information about the input. If indirect measurements are applied, the collapse will be more or less evident in terms of the measurement strength (interaction strength between measurement apparatus and system). Second, the outcome of a measurement is stochastic, so we need to repeat the measurement process several times at a given time step to obtain some statistics.

The combination of these two effects, represented by the measurement strength and the number of measurements respectively, is introduced in the proposed online protocol, and both effects determine the answer to the question formulated above. On the one hand, the measurement strength controls the trade-off between how much information is extracted in a single measurement and the dissipative effect, and we proposed to use weak measurements to maximize this trade-off. On the other hand, the number of measurements is directly related to the resources of the experimental protocol. The number of measurements in this protocol refers for instance to the number of copies of the system, feeding the input information in all copies at the same time. The performance of the online protocol is tested against two other protocols with halt and repetition, the restarting and rewinding protocols, in order to make a fair comparison between performance and resources.

Again, Sect. 1.4.4 of the introduction is relevant to follow this chapter because of the model description it contains, while Sect. 1.5 introduces the quantum measurement formalism, providing some of the basic ingredients that led us to the present publication. As in the previous chapter, this work was led by Dr. Pere Mujal, establishing a thriving collaboration in both numerical and analytical aspects of the QRC experimental protocols.

6.2 Published paper

Pere Mujal, Rodrigo Martínez-Peña, Gian Luca Giorgi, Miguel C. Soriano and Roberta Zambrini, Time-series quantum reservoir computing with weak and projective measurements, *npj Quantum Information*, **9**, 16 (2023).
https://doi.org/10.1038/s41534-023-00682-z

Chapter 7

Quantum reservoir computing in finite dimensions

7.1 Motivation and contribution

Approaching the end of this thesis, we looked back over the last years and we realized that there were only a few theoretical results in the QRC field. As commented in the introduction, sufficient conditions for the UAP have been obtained for qubit platforms [126, 139, 151], finding a few other publications where the quantum ESP is discussed as well [138, 149]. We also contributed to the theory of QRC by finding a necessary and sufficient condition for the ESP and FMP and studying the UAP of continuous-variable systems with Gaussian states [123].

However, we found this material insufficient. Although RC properties are usually hard to analyze in nonlinear RC systems, quantum reservoir dynamics is intrinsically linear. And indeed, linear dynamics considerably simplifies the RC theory [91, 104, 123, 228–232]. Another student of the team was actually able to show that a different QRC approach based on the master equation description displays also UAP [133]. Then, we tackled the question of minimal ingredients for operational QRC systems. We had the intuition that some general statements could be derived for QRC systems under very few assumptions. We started by exploring a simple setup: a finite-dimensional system whose reservoir dynamics is driven by a CPTP map with classical inputs. To do so, we realized that a different approach to previous theoretical works was required. We based our study on using a different representation of quantum systems, in such a way that the RC properties analysis could be simplified. We found the state-affine system (SAS) representation to be the desired one [104].

The findings of the chapter can be summarized in a single sentence: the necessary and sufficient condition that makes a quantum reservoir valuable is strictly contractive dynamics towards input-dependent fixed points. The intuition behind this formal result is that dissipation and contractive dynamics make the system converge to an attractor where the system will eventually "die" if the attractor does not respond to the input injections. But if this attractor depends on the input, the system will follow its lead. This work connects with two important elements of the theory of QRC systems: the UAP and the proper design of a quantum reservoir. Future work will be devoted to tackling both problems, completely characterizing the UAP in quantum finite dimensional systems and understanding the physical conditions that optimize the performance in QRC systems.

This paper is intended to be self-contained in many aspects. Section 1.3 and in particular Sect. 1.3.4 can help to gain some intuition about the mathematical properties of reservoir computers. This work was produced during my research stay at the Nanyang Technological University of Singapore in collaboration with Professor Juan-Pablo Ortega. I led the project with the proposal, demonstration, and analysis of the results.

7.2 Published paper

Rodrigo Martínez-Peña and Juan-Pablo Ortega, Quantum reservoir computing in finite dimensions, *Physical Review E*, **107**, 035306 (2023).
https://doi.org/10.1103/PhysRevE.107.035306

Chapter 8

Conclusions and outlook

This thesis has contributed to the development of the QRC field in different aspects. From the beginning, special attention has been devoted to complex network platforms that provide rich enough dynamics to process temporal series. Such an interest in quantum complex systems as physical substrates for RC aroused because of the following main reasons: the possibility to disclose the relation between complex quantum dynamics and information processing; the large number of degrees of freedom that quantum systems can exhibit with a few particles; the current availability of these experimental platforms; the possible experimental advantages that we could find with respect to classical models; and the potential extension of RC techniques to deal with quantum data. Most of our work has been devoted to numerically and analytically exploring quantum spin models as reservoirs, with special attention to ensemble quantum systems. NMR experiments are suited for this approach [16], where one can measure the average value of an observable with a single shot and we can find a rich dynamical landscape [186, 233, 234], but their use for temporal tasks remains unclear. Other platforms like trapped-ion quantum simulators [14, 15], superconducting circuits [183, 235] or integrated photonics [17] are already well established to serve as QRC systems as well.

Most of our results (Chapters 3 to 6) address different fundamental and implementation issues exploring in depth the Fujii and Nakajima model [113]. The specific setting that we have analyzed during these years is a QRC system formed by the composition of an input CPTP map and a unitary reservoir map. The input map consists on the reinitialization of one qubit with a superposition in terms of the input. The unitary dynamics processes the input information and it is given by the transverse-field Ising model. The simplest version we used contained a homogeneous magnetic field in the z axis, while the posterior version we adopted included a random heterogeneous magnetic field, increasing the richness of the dynamical regimes of the model. The joint action of the unitary map with the reinitialization of a single qubit is determinant to obtain the dissipation that guarantees the ESP and FMP. The final element of the QRC system under study is an output layer formed by the expected value of the spin observables.

The first question, addressed in Chapter 3, was devoted to understanding the computational capabilities of this QRC system of quantum spins. While different tasks were reported in [113], we computed for the first time in QRC the IPC, which measures the linear and nonlinear memory of the system. Assisted by an analysis of the convergence speed in terms of Δt, we found optimal conditions for the input driving in terms of computational speed and performance,

also showing different alternatives for the output layer in terms of temporal multiplexing and correlation observables. Actually, this study was crucial to unveil the complexity of the Ising model as a reservoir with a performance strongly influenced by its hyperparameters. This approach displayed the advantage offered by the dimension of the Hilbert space. Chapter 4 pursued a firm understanding of this relation between the hyperparameters of the quantum spin model and its performance. We discovered that a direct connection between these elements is established through the dynamical regimes of the unitary dynamics. In fact, we concluded that thermalization in the unitary dynamics provides the appropriate conditions for RC while localization is detrimental. At the transition between thermalization and localization regimes, it is possible to observe a peak in performance depending on the task that we address. Chapters 3 and 4 constituted a fundamental block in the development of this thesis. Intensive but feasible numerical calculations of tools like the IPC and other benchmark tasks allowed us to pave a path that our group still relies on to understand the capabilities of quantum reservoirs [131, 223].

Chapter 5 deals with the specific problem of understanding the origin of nonlinearity in the models that we had been employing. This question is not only relevant in QRC, but also in the whole quantum ML field since nonlinearity is a fundamental ingredient for the expressivity of ML models while quantum dynamics is linear. Of course, linearity in the state does not preclude nonlinearity in the input-output relation. Indeed, explicit analytical expressions were provided for QRC models of spins and also for the case of linear boson dynamics: the input codification, with its preprocessing, determines the nonlinear response of the QRC system. However, the natural dynamics of the reservoir, determined by the hyperparameters of the model, can affect this response too, as we showed for the dynamical regimes of the transverse-field Ising model in Chapter 4. We emphasize that different input codification strategies for a given dynamical model, or otherwise, different dynamical models for the same input codification, can completely change the nonlinear response. This was vividly illustrated in the quantum master equation examples of Chapter 7. The input dependence of these models can be intricate, exhibiting trigonometric and hyperbolic functions in an organic way. But what is more striking is that just a change in the direction of the external magnetic field, as in the models of Example 3 and Sect. IV A of Chapter 7, which have the same dissipation, can totally modify the nonlinear response (and the viability of the reservoir).

We concluded the study of QRC based on quantum spin models with the conception of an online experimental protocol in Chapter 6. Indeed, a crucial aspect to deal with temporal data analysis is the possibility to process information sequentially. Our goal was to prove that it is possible to find a compromise between performance and resource efficiency for a protocol where there is no need to repeat input encoding steps nor need to store any input externally, and still properly consider the measurement backaction in the reservoir dynamics. The key of our proposal relies on indirect weak measurements over an ensemble. We find that utilizing measurements of the proper strength with the OLP could potentially outperform the RWP, even when dealing with higher-order moments of observables. This can ultimately lead to an efficient exploitation of the Hilbert

space. In this line, an independent approach for a photonic implementation was also proposed in Ref. [131].

Chapter 7, despite being our last article, is the beginning of a thriving collaboration where we take a step back from the study of specific models. Instead, we adopted a general viewpoint of the QRC framework, sacrificing some experimental details in exchange for a deeper insight into the physical aspects that make QRC systems operational. In the precise case of finite dimensional quantum systems with classical inputs and an infinite number of measurements, we find that the necessary and sufficient condition that makes a reservoir valuable is strictly contractive dynamics towards input-dependent fixed points. In this chapter, we also establish a connection between experimental protocols and the viability of the QRC model. The weak measurement scheme we proposed in Chapter 6 can be represented as a dephasing channel in the measurement direction when averaging in the limit of an infinitely large ensemble. Then, as discussed in Sect. IV of Chapter 7, not all CPTP maps might be employed with this measurement scheme since the fixed point of the total quantum channel could become input-independent. For instance, unitary maps would not fit because the composition map would be unital, becoming useless for long input sequences as described in Theorem 1. Hence, we see that our general picture can relate measurement schemes and reservoir dynamics in a formal way, being able to design operational QRC systems under backaction effects as well. Indeed, our framework can cover online protocols with backaction as well as offline protocols with repetition, concluding that our results are valuable for the design of experimental platforms.

Chapter 2 precedes the research chapters of this thesis because of the global perspective that it offers. The time elapsed between the first publication of Fujii and Nakajima [113] and our review [18] allowed us to understand the large number of physical platforms that are good candidates for QRC (and QELM as well) and identify the main challenges of the field. Since its publication, new articles have confirmed some of our research avenue predictions (exposed in Sect. 3 of Chapter 2). For example, the quantum circuits experimental line for QRC has continued, where noise can be incorporated as part of the RC dynamics [79, 94, 134]. As commented at the end of Chapters 2 and 7, caution is advised since the size of temporal sequences that noisy circuit platforms (and any noisy quantum device) can process is limited by the decoherence time scale. Another challenging aspect that we noticed and that we have started to explore is the design of online experimental protocols for time series processing. Chapter 6 was the response to this initially identified challenge, where we proposed a middle-point solution between the theoretical idea of ensemble measurements without backaction and strong projective measurements: weak measurements over several samples at each time step. However, up to this moment, there is no experimental implementation of online protocols with quantum hardware. We further notice that some advancements were also made in the QRC theory of quantum temporal task [148–150].

However, due to the early age of the QRC field, the list of open problems is large and wide open to increase with each breakthrough. To start with, several directions can be taken from Chapter 6. Apart from the obvious exploration of

different types of quantum measurements and reservoir models, a very important question is how to experimentally and efficiently exploit the large number of degrees of freedom that complex quantum systems offer. A possible path could be the randomization of the measurement process [236], as it is done in shadow tomography, which still requires averaging over samples but can be proven to be efficient to estimate a set of local observables [237]. Besides, state-of-the-art quantum platforms would already allow to implement the devised experimental protocols. Several experimental groups with the means to perform QRC have shown their interest (e.g. [17]). Another aspect from Chapter 6 that is quite open to debate is the exploration of online protocols with stochastic reservoir dynamics. The idea would be to try to reduce the number of measurements to the minimum, with even single-shot measurements at each time step. Despite finding that a small number of samples give a poor performance in qubit systems in Chapter 6, there could be situations where a quantum stochastic reservoir is still operational. The intuition behind this statement is that a small amount of noise in the reservoir (either classical or quantum) dynamics should be tolerable, with a transition between operational and useless in terms of noise strength. This transition might only happen at the quantum-classical limit, but the exploration could be worth it.

The fact that Chapter 7 is just our first dive into a QRC theory leaves the door open to many possibilities. Studying particular CPTP maps like Markovian master equations, extending the theory to non-ideal situations such as a finite number of measurements, or the inclusion of arbitrary POVMs into the reservoir dynamics are just a few examples. More ambitious directions include the study of infinite-dimensional quantum systems or trying to find the most general as possible statement about the UAP with QRC systems. Different encoding strategies and dynamical maps can be explored as well. We recently studied the implementation of Markovian quantum dynamics in spin models, where the input is codified as a Hamiltonian parameter such as an external magnetic field [133]. This implementation, when compared with Fujii and Nakajima model with the same number of qubits in the reservoir and the same output observables, outperformed the latter for such dispair targets such as the NARMA task and the Mackey-Glass time-series prediction.

Finally, one of the driving forces of studying QRC systems is the possibility to find advantages over classical systems. These advantages could be of any type: performance, time, or even energy. At this moment, there are a few works suggesting an advantage in terms of physical resources such as reservoir size [94, 113, 138, 139]. However, we consider that a systematic exploration could be worth it. In fact, we suspect that any possible advantage of quantum versus classical might come from the study of quantum temporal tasks. Quantum systems are native platforms for processing quantum information and it makes sense that any advantage emerges in this particular setup.

Acronyms

ML	**M**achine **L**earning
QML	**Q**uantum **M**achine **L**earning
NMR	**N**uclear **M**agnetic **R**esonance
GPU	**G**raphical **P**rocessing **U**nit
ANN	**A**rtificial **N**eural **N**etwork
FFNN	**F**eed-**F**orward **N**eural **N**etwork
CNN	**C**onvolutional **N**eural **N**etwork
RNN	**R**ecurrent **N**eural **N**etwork
LSTM	**L**ong **S**hort-**T**erm **M**emory
GRU	**G**ated **R**ecurrent **U**nit
MSE	**M**ean **S**quare **E**rror
NMSE	**N**ormalized **M**ean **S**quare **E**rror
SGD	**S**tochastic **G**radient **D**escent
RC	**R**eservoir **C**omputing
PRC	**P**hysical **R**eservoir **C**omputing
QRC	**Q**uantum **R**eservoir **C**omputing
ESN	**E**cho **S**tate **N**etwork
LSM	**L**iquid **S**tate **M**achine
UAP	**U**universal **A**pproximation **P**roperty
ESP	**E**cho **S**tate **P**roperty
FMP	**F**ading **M**emory **P**roperty
CPTP	**C**ompletely **P**ositive and **T**race **P**reserving
ELM	**E**xtreme **L**earning **M**achine
QELM	**Q**uantum **E**xtreme **L**earning **M**achine
POVM	**P**ositive **O**perator-**V**alued **M**easurement
RMT	**R**andom **M**atrix **T**heory
GOE	**G**aussian **O**rthogonal **E**nsemble
GUE	**G**aussian **U**nitary **E**nsemble
ETH	**E**igenstate **T**hermalization **H**ypothesis
MBL	**M**any-**B**ody **L**ocalization
LIOM	**L**ocal **I**ntegral **O**f **M**otion
IPC	**I**nformation **P**rocessing **C**apacity
TIPC	**T**emporal **I**nformation **P**rocessing **C**apacity
NARMA	**N**onlinear **A**utoregressive **M**oving **A**verage
STM	**S**hort-**T**erm **M**emory
RSP	**R**e**S**tarting **P**rotocol
RWP	**R**e**W**inding **P**rotocol
OLP	**O**n**L**ine **P**rotocol
SAS	**S**tate-**A**ffine **S**ystem

Bibliography

1. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2,** 79 (2018).
2. Bharti, K. *et al.* Noisy intermediate-scale quantum algorithms. *Reviews of Modern Physics* **94,** 015004 (2022).
3. Arute, F. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574,** 505–510 (2019).
4. Zhong, H.-S. *et al.* Quantum computational advantage using photons. *Science* **370,** 1460–1463 (2020).
5. Huang, H.-Y. *et al.* Quantum advantage in learning from experiments. *Science* **376,** 1182–1186 (2022).
6. Daley, A. J. *et al.* Practical quantum advantage in quantum simulation. *Nature* **607,** 667–676 (2022).
7. Chen, Y.-A. *et al.* An integrated space-to-ground quantum communication network over 4,600 kilometres. *Nature* **589,** 214–219 (2021).
8. Biamonte, J. *et al.* Quantum machine learning. *Nature* **549,** 195–202 (2017).
9. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *nature* **521,** 436–444 (2015).
10. Goodfellow, I., Bengio, Y. & Courville, A. *Deep learning* (MIT press, 2016).
11. Benjamin, B. V. *et al.* Neurogrid: A mixed-analog-digital multichip system for large-scale neural simulations. *Proceedings of the IEEE* **102,** 699–716 (2014).
12. Furber, S. B., Galluppi, F., Temple, S. & Plana, L. A. The spinnaker project. *Proceedings of the IEEE* **102,** 652–665 (2014).
13. Indiveri, G. & Liu, S.-C. Memory and information processing in neuromorphic systems. *Proceedings of the IEEE* **103,** 1379–1397 (2015).
14. Smith, J. *et al.* Many-body localization in a quantum simulator with programmable random disorder. *Nature Physics* **12,** 907–911 (2016).
15. Zhang, J. *et al.* Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator. *Nature* **551,** 601–604 (2017).
16. Negoro, M., Mitarai, K., Fujii, K., Nakajima, K. & Kitagawa, M. Machine learning with controllable quantum dynamics of a nuclear spin ensemble in a solid. *arXiv preprint arXiv:1806.10910* (2018).
17. Spagnolo, M. *et al.* Experimental photonic quantum memristor. *Nature Photonics* **16,** 318–323 (2022).
18. Mujal, P. *et al.* Opportunities in quantum reservoir computing and extreme learning machines. *Advanced Quantum Technologies* **4,** 2100027 (2021).

19. Martínez-Peña, R., Nokkala, J., Giorgi, G. L., Zambrini, R. & Soriano, M. C. Information processing capacity of spin-based quantum reservoir computing systems. *Cognitive Computation*, 1–12 (2020).
20. Martínez-Peña, R., Giorgi, G. L., Nokkala, J., Soriano, M. C. & Zambrini, R. Dynamical phase transitions in quantum reservoir computing. *Physical Review Letters* **127,** 100502 (2021).
21. Mujal, P. *et al.* Analytical evidence of nonlinearity in qubits and continuous-variable quantum reservoir computing. *Journal of Physics: Complexity* **2,** 045008 (2021).
22. Mujal, P., Martínez-Peña, R., Giorgi, G. L., Soriano, M. C. & Zambrini, R. Time-series quantum reservoir computing with weak and projective measurements. *npj Quantum Information* **9,** 16 (2023).
23. Martínez-Peña, R. & Ortega, J.-P. Quantum reservoir computing in finite dimensions. *Phys. Rev. E* **107,** 035306 (2023).
24. Kitchin, R. *The data revolution: Big data, open data, data infrastructures and their consequences* (Sage, 2014).
25. Waldrop, M. M. The chips are down for Moore's law. *Nature News* **530,** 144 (2016).
26. Singh, D. & Reddy, C. K. A survey on platforms for big data analytics. *Journal of big data* **2,** 1–20 (2015).
27. Michalski, R. S., Carbonell, J. G. & Mitchell, T. M. *Machine learning: An artificial intelligence approach* (Springer Science & Business Media, 2013).
28. Russell, S. J. *Artificial intelligence a modern approach* (Pearson Education, Inc., 2010).
29. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596,** 583–589 (2021).
30. Nielsen, M. A. *Neural networks and deep learning* (Determination press San Francisco, CA, USA, 2015).
31. Duda, R. O., Hart, P. E., *et al. Pattern classification* (John Wiley & Sons, 2006).
32. Sutton, R. S. & Barto, A. G. *Reinforcement learning: An introduction* (MIT press, 2018).
33. Basheer, I. A. & Hajmeer, M. Artificial neural networks: fundamentals, computing, design, and application. *Journal of microbiological methods* **43,** 3–31 (2000).
34. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *nature* **323,** 533–536 (1986).
35. Hinton, G. *et al.* Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine* **29,** 82–97 (2012).
36. Baldi, P., Sadowski, P. & Whiteson, D. Searching for exotic particles in high-energy physics with deep learning. *Nature communications* **5,** 1–9 (2014).

37. Covington, P., Adams, J. & Sargin, E. *Deep neural networks for youtube recommendations* in *Proceedings of the 10th ACM conference on recommender systems* (2016), 191–198.
38. LeCun, Y. *et al.* Handwritten digit recognition with a back-propagation network. *Advances in neural information processing systems* **2** (1989).
39. Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Communications of the ACM* **60,** 84–90 (2017).
40. Taigman, Y., Yang, M., Ranzato, M. & Wolf, L. *Deepface: Closing the gap to human-level performance in face verification* in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2014), 1701–1708.
41. Collobert, R. *et al.* Natural language processing (almost) from scratch. *Journal of machine learning research* **12,** 2493–2537 (2011).
42. Silver, D. *et al.* A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* **362,** 1140–1144 (2018).
43. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C. & Chen, M. Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125* (2022).
44. Chen, M. *et al.* Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374* (2021).
45. Lipton, Z. C., Berkowitz, J. & Elkan, C. A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019* (2015).
46. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena* **404,** 132306 (2020).
47. PJ, W. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE* **78,** 1550–1560 (1990).
48. Bengio, Y., Simard, P. & Frasconi, P. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks* **5,** 157–166 (1994).
49. Pascanu, R., Mikolov, T. & Bengio, Y. Understanding the exploding gradient problem. *CoRR, abs/1211.5063* **2,** 1 (2012).
50. Atiya, A. F. & Parlos, A. G. New results on recurrent network training: unifying the algorithms and accelerating convergence. *IEEE transactions on neural networks* **11,** 697–709 (2000).
51. Sutskever, I. *Training recurrent neural networks* (University of Toronto Toronto, ON, Canada, 2013).
52. Pascanu, R., Mikolov, T. & Bengio, Y. *On the difficulty of training recurrent neural networks* in *International conference on machine learning* (2013), 1310–1318.
53. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural computation* **9,** 1735–1780 (1997).

54. Cho, K. *et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078* (2014).
55. Mikolov, T., Karafiát, M., Burget, L., Cernockỳ, J. & Khudanpur, S. *Recurrent neural network based language model.* in *Interspeech* **2** (2010), 1045–1048.
56. Sutskever, I., Vinyals, O. & Le, Q. V. Sequence to sequence learning with neural networks. *Advances in neural information processing systems* **27** (2014).
57. Sutskever, I., Martens, J. & Hinton, G. E. *Generating text with recurrent neural networks* in *ICML* (2011).
58. Graves, A., Mohamed, A.-r. & Hinton, G. *Speech recognition with deep recurrent neural networks* in *2013 IEEE international conference on acoustics, speech and signal processing* (2013), 6645–6649.
59. Vinyals, O. *et al.* Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* **575,** 350–354 (2019).
60. Qing, X. & Niu, Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy* **148,** 461–468 (2018).
61. Nayebi, A. & Vitelli, M. Gruv: Algorithmic music generation using recurrent neural networks. *Course CS224D: Deep Learning for Natural Language Processing (Stanford),* 52 (2015).
62. Flurin, E., Martin, L. S., Hacohen-Gourgy, S. & Siddiqi, I. Using a recurrent neural network to reconstruct quantum dynamics of a superconducting qubit from physical observations. *Physical Review X* **10,** 011006 (2020).
63. Dunjko, V. & Briegel, H. J. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Reports on Progress in Physics* **81,** 074001 (2018).
64. Carleo, G. *et al.* Machine learning and the physical sciences. *Reviews of Modern Physics* **91,** 045002 (2019).
65. Carrasquilla, J. Machine learning for quantum matter. *Advances in Physics: X* **5,** 1797528 (2020).
66. Gallicchio, C., Micheli, A. & Pedrelli, L. Comparison between DeepESNs and gated RNNs on multivariate time-series prediction. *arXiv preprint arXiv:1812.11527* (2018).
67. Bacciu, D., Di Sarli, D., Gallicchio, C., Micheli, A. & Puccinelli, N. *Benchmarking Reservoir and Recurrent Neural Networks for Human State and Activity Recognition* in *International Work-Conference on Artificial Neural Networks* (2021), 168–179.
68. Shahi, S., Fenton, F. H. & Cherry, E. M. Prediction of chaotic time series using recurrent neural networks and reservoir computing techniques: A comparative study. *Machine learning with applications* **8,** 100300 (2022).
69. Jaeger, H. Towards a generalized theory comprising digital, neuromorphic and unconventional computing. *Neuromorphic Computing and Engineering* **1,** 012002 (2021).
70. Andrae, A. S. & Edler, T. On global electricity usage of communication technology: trends to 2030. *Challenges* **6,** 117–157 (2015).

71. Thompson, N. C., Greenewald, K., Lee, K. & Manso, G. F. The computational limits of deep learning. *arXiv preprint arXiv:2007.05558* (2020).
72. Patterson, D. *et al.* Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350* (2021).
73. Coalition, E.-w. *A new circular vision for electronics: Time for a global reboot* in (2019).
74. Backus, J. Can programming be liberated from the von Neumann style? A functional style and its algebra of programs. *Communications of the ACM* **21,** 613–641 (1978).
75. Marković, D., Mizrahi, A., Querlioz, D. & Grollier, J. Physics for neuromorphic computing. *Nature Reviews Physics* **2,** 499–510 (2020).
76. Lukoševičius, M. & Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Computer Science Review* **3,** 127–149 (2009).
77. Jaeger, H. 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,** 13 (2001).
78. Sun, C., Song, M., Hong, S. & Li, H. A review of designs and applications of echo state networks. *arXiv preprint arXiv:2012.02974* (2020).
79. Suzuki, Y., Gao, Q., Pradel, K. C., Yasuoka, K. & Yamamoto, N. Natural quantum reservoir computing for temporal information processing. *Scientific reports* **12,** 1–15 (2022).
80. Estébanez, I., Fischer, I. & Soriano, M. C. Constructive role of noise for high-quality replication of chaotic attractor dynamics using a hardware-based reservoir computer. *Physical Review Applied* **12,** 034058 (2019).
81. Nakajima, K., Hauser, H., Li, T. & Pfeifer, R. Information processing via physical soft body. *Scientific reports* **5,** 1–11 (2015).
82. Maass, W., Natschläger, T. & Markram, H. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural computation* **14,** 2531–2560 (2002).
83. Anderson, E. *et al. LAPACK Users' Guide* Third. ISBN: 0-89871-447-8 (paperback) (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999).
84. Van der Sande, G., Brunner, D. & Soriano, M. C. Advances in photonic reservoir computing. *Nanophotonics* **6,** 561–576 (2017).
85. Brunner, D., Soriano, M. C. & Van der Sande, G. *Photonic reservoir computing: optical recurrent neural networks* (Walter de Gruyter GmbH & Co KG, 2019).
86. Paquot, Y. *et al.* Optoelectronic reservoir computing. *Scientific reports* **2,** 287 (2012).
87. Bhovad, P. & Li, S. Physical reservoir computing with origami and its application to robotic crawling. *Scientific Reports* **11,** 1–18 (2021).
88. Fernando, C. & Sojakka, S. *Pattern recognition in a bucket* in *European conference on artificial life* (2003), 588–597.

89. Huang, G.-B., Wang, D. H. & Lan, Y. Extreme learning machines: a survey. *International journal of machine learning and cybernetics* **2,** 107–122 (2011).

90. Jaeger, H. *Short term memory in echo state networks. gmd-report 152* in *GMD-German National Research Institute for Computer Science (2002), http://www. faculty. jacobs-university. de/hjaeger/pubs/STMEchoStatesTechRep. pdf* (2002).

91. Gonon, L. & Ortega, J.-P. Reservoir computing universality with stochastic inputs. *IEEE transactions on neural networks and learning systems* **31,** 100–112 (2019).

92. Dambre, J., Verstraeten, D., Schrauwen, B. & Massar, S. Information processing capacity of dynamical systems. *Scientific reports* **2,** 1–7 (2012).

93. Kubota, T., Takahashi, H. & Nakajima, K. Unifying framework for information processing in stochastically driven dynamical systems. *Physical Review Research* **3,** 043135 (2021).

94. Kubota, T. *et al.* Quantum Noise-Induced Reservoir Computing. *arXiv preprint arXiv:2207.07924* (2022).

95. Hülser, T., Köster, F., Lüdge, K. & Jaurigue, L. Deriving task specific performance from the information processing capacity of a reservoir computer. *Nanophotonics* (2022).

96. Hu, F. *et al.* Fundamental Limits to Expressive Capacity of Finitely Sampled Qubit-Based Systems. *arXiv preprint arXiv:2301.00042* (2022).

97. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems* **2,** 303–314 (1989).

98. Hornik, K., Stinchcombe, M. & White, H. Multilayer feedforward networks are universal approximators. *Neural networks* **2,** 359–366 (1989).

99. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural networks* **4,** 251–257 (1991).

100. Maass, W. & Sontag, E. D. Neural systems as nonlinear filters. *Neural computation* **12,** 1743–1772 (2000).

101. Maass, W., Natschläger, T. & Markram, H. Fading memory and kernel properties of generic cortical microcircuit models. *Journal of Physiology-Paris* **98,** 315–330 (2004).

102. Maass, W., Joshi, P. & Sontag, E. D. Computational aspects of feedback in neural circuits. *PLoS computational biology* **3,** e165 (2007).

103. Grigoryeva, L. & Ortega, J.-P. Echo state networks are universal. *Neural Networks* **108,** 495–508 (2018).

104. Grigoryeva, L. & Ortega, J.-P. Universal discrete-time reservoir computers with stochastic inputs and linear readouts using non-homogeneous state-affine systems. *Journal of Machine Learning Research* **19,** 1–40 (2018).

105. Boyd, S. & Chua, L. Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Transactions on circuits and systems* **32,** 1150–1161 (1985).

106. Dieudonné, J. *Foundations of modern analysis* (Read Books Ltd, 2011).

107. Nakajima, K. Physical reservoir computing—an introductory perspective. *Japanese Journal of Applied Physics* **59,** 060501 (2020).

108. Nakajima, K. & Fischer, I. *Reservoir Computing* (Springer, 2021).

109. Tanaka, G. *et al.* Recent advances in physical reservoir computing: A review. *Neural Networks* **115,** 100–123 (2019).

110. Nokkala, J., Martínez-Peña, R., Zambrini, R. & Soriano, M. C. High-Performance Reservoir Computing with Fluctuations in Linear Networks. *IEEE Transactions on Neural Networks and Learning Systems* (2021).

111. Appeltant, L. *et al.* Information processing using a single dynamical node as complex system. *Nature communications* **2,** 1–6 (2011).

112. Nakajima, K. & Aoyagi, T. The memory capacity of a physical liquid state machine. *IEICE Technical Report; IEICE Tech. Rep.* **115,** 109–112 (2015).

113. Fujii, K. & Nakajima, K. Harnessing disordered-ensemble quantum dynamics for machine learning. *Physical Review Applied* **8,** 024030 (2017).

114. Tsunegi, S. *et al.* Physical reservoir computing based on spin torque oscillator with forced synchronization. *Applied Physics Letters* **114,** 164101 (2019).

115. Furuta, T. *et al.* Macromagnetic simulation for reservoir computing utilizing spin dynamics in magnetic tunnel junctions. *Physical Review Applied* **10,** 034063 (2018).

116. Stieg, A. Z. *et al. Emergent criticality in complex turing B-type atomic switch networks* 2012.

117. Prychynenko, D. *et al.* Magnetic skyrmion as a nonlinear resistive element: a potential building block for reservoir computing. *Physical Review Applied* **9,** 014034 (2018).

118. Torrejon, J. *et al.* Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547,** 428 (2017).

119. Caluwaerts, K. *et al.* Design and control of compliant tensegrity robots through simulation and hardware validation. *Journal of the royal society interface* **11,** 20140520 (2014).

120. Moon, J. *et al.* Temporal data classification and forecasting using a memristor-based reservoir computing system. *Nature Electronics* **2,** 480–487 (2019).

121. Kraus, K., Böhm, A., Dollard, J. D. & Wootters, W. *States, Effects, and Operations Fundamental Notions of Quantum Theory: Lectures in Mathematical Physics at the University of Texas at Austin* (Springer, 1983).

122. Govia, L., Ribeill, G., Rowlands, G. & Ohki, T. Nonlinear input transformations are ubiquitous in quantum reservoir computing. *Neuromorphic Computing and Engineering* **2,** 014008 (2022).

123. Nokkala, J. *et al.* Gaussian states of continuous-variable quantum systems provide universal and versatile reservoir computing. *Communications Physics* **4,** 1–11 (2021).

124. Pino, J. M. *et al.* Demonstration of the trapped-ion quantum CCD computer architecture. *Nature* **592,** 209–213 (2021).

125. Dasgupta, S., Hamilton, K. E. & Banerjee, A. *Characterizing the memory capacity of transmon qubit reservoirs* in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)* (2022), 162–166.

126. Chen, J., Nurdin, H. I. & Yamamoto, N. Temporal information processing on noisy quantum computers. *Physical Review Applied* **14,** 024065 (2020).

127. Nokkala, J. *et al.* Reconfigurable optical implementation of quantum complex networks. *New Journal of Physics* **20,** 053024 (2018).

128. Cai, Y. *et al.* Multimode entanglement in reconfigurable graph states using optical frequency combs. *Nature communications* **8,** 15645 (2017).

129. Bravo, R. A., Najafi, K., Gao, X. & Yelin, S. F. Quantum reservoir computing using arrays of Rydberg atoms. *PRX Quantum* **3,** 030325 (2022).

130. Dudas, J., Plouet, E., Mizrahi, A., Grollier, J. & Marković, D. Quantum reservoir neural network implementation on a Josephson mixer. *arXiv preprint arXiv:2209.03221* (2022).

131. García-Beni, J., Giorgi, G. L., Soriano, M. C. & Zambrini, R. Scalable photonic platform for real-time quantum reservoir computing. *arXiv preprint arXiv:2207.14031* (2022).

132. Kalfus, W. *et al.* Hilbert space as a computational resource in reservoir computing. *Physical Review Research* **4,** 033007 (2022).

133. Sannia, A., Martínez-Peña, R., Soriano, M. C., Giorgi, G. L. & Zambrini, R. Dissipation as a resource for Quantum Reservoir Computing. *arXiv preprint arXiv:2212.12078* (2022).

134. Molteni, R., Destri, C. & Prati, E. Optimization of the memory reset rate of a quantum echo-state network for time sequential tasks. *Physics Letters A,* 128713 (2023).

135. Fry, D., Deshmukh, A., Chen, S. Y.-C., Rastunkov, V. & Markov, V. Optimizing quantum noise-induced reservoir computing for nonlinear and chaotic time series prediction. *arXiv preprint arXiv:2303.05488* (2023).

136. Nakajima, K., Fujii, K., Negoro, M., Mitarai, K. & Kitagawa, M. Boosting computational power through spatial multiplexing in quantum reservoir computing. *Physical Review Applied* **11,** 034021 (2019).

137. Kutvonen, A., Fujii, K. & Sagawa, T. Optimizing a quantum reservoir computer for time series prediction. *Scientific reports* **10,** 1–7 (2020).

138. Tran, Q. H. & Nakajima, K. Higher-order quantum reservoir computing. *arXiv preprint arXiv:2006.08999* (2020).

139. Chen, J. & Nurdin, H. I. Learning nonlinear input–output maps with dissipative quantum systems. *Quantum Information Processing* **18,** 1–36 (2019).

140. Xia, W., Zou, J., Qiu, X. & Li, X. The reservoir learning power across quantum many-body localization transition. *Frontiers of Physics* **17,** 1–9 (2022).

141. Xia, W. *et al.* Configured Quantum Reservoir Computing for Multi-Task Machine Learning. *arXiv preprint arXiv:2303.17629* (2023).

142. Götting, N., Lohof, F. & Gies, C. Exploring quantum mechanical advantage for reservoir computing. *arXiv preprint arXiv:2302.03595* (2023).

143. Govia, L., Ribeill, G., Rowlands, G., Krovi, H. & Ohki, T. Quantum reservoir computing with a single nonlinear oscillator. *Physical Review Research* **3,** 013077 (2021).

144. Motamedi, A., Zadeh-Haghighi, H. & Simon, C. Correlations Between Quantumness and Learning Performance in Reservoir Computing with a Single Oscillator. *arXiv preprint arXiv:2304.03462* (2023).

145. Sakurai, A., Estarellas, M. P., Munro, W. J. & Nemoto, K. Quantum Extreme Reservoir Computation Utilizing Scale-Free Networks. *Physical Review Applied* **17,** 064044 (2022).

146. Ghosh, S., Opala, A., Matuszewski, M., Paterek, T. & Liew, T. C. Quantum reservoir processing. *npj Quantum Information* **5,** 1–6 (2019).

147. Khan, S. A., Hu, F., Angelatos, G. & Türeci, H. E. Physical reservoir computing using finitely-sampled quantum systems. *arXiv preprint arXiv:2110.13849* (2021).

148. Nokkala, J. Online quantum time series processing with random oscillator networks. *Scientific Reports* **13,** 7694 (2023).

149. Tran, Q. H. & Nakajima, K. Learning temporal quantum tomography. *Physical Review Letters* **127,** 260401 (2021).

150. Tran, Q. H., Ghosh, S. & Nakajima, K. Quantum-Classical Hybrid Information Processing via a Single Quantum System. *arXiv preprint arXiv:2209.00497* (2022).

151. Chen, J. *Nonlinear Convergent Dynamics for Temporal Information Processing on Novel Quantum and Classical Devices* PhD thesis (UNSW Sydney, 2022).

152. Von Neumann, J. *Mathematische grundlagen der quantenmechanik* (Springer-Verlag, 2013).

153. Lüders, G. Über die Zustandsänderung durch den Meßprozeß. *Annalen der Physik* **443,** 322–328 (1950).

154. Wiseman, H. M. & Milburn, G. J. *Quantum measurement and control* (Cambridge university press, 2009).

155. Breuer, H.-P., Petruccione, F., *et al. The theory of open quantum systems* (Oxford University Press on Demand, 2002).

156. Bloch, I., Dalibard, J. & Zwerger, W. Many-body physics with ultracold gases. *Reviews of modern physics* **80,** 885 (2008).

157. Blatt, R. & Roos, C. F. Quantum simulations with trapped ions. *Nature Physics* **8,** 277–284 (2012).

158. Von Neumann, J. Proof of the ergodic theorem and the H-theorem in quantum mechanics. *The European Physical Journal H* **35,** 201–237 (2010).

159. Von Neumann, J. Proof of the Ergodic theorem and the H-Theorem in the new Mechanics. *Z. Phys* **57,** 30–70 (1929).

160. Deutsch, J. M. Quantum statistical mechanics in a closed system. *Physical review a* **43,** 2046 (1991).

161. Srednicki, M. Chaos and quantum thermalization. *Physical review e* **50,** 888 (1994).

162. Srednicki, M. The approach to thermal equilibrium in quantized chaotic systems. *Journal of Physics A: Mathematical and General* **32,** 1163 (1999).

163. Rigol, M., Dunjko, V. & Olshanii, M. Thermalization and its mechanism for generic isolated quantum systems. *Nature* **452,** 854–858 (2008).

164. Pathria, R. K. *Statistical mechanics* (Elsevier, 2016).

165. Edelman, A. & Rao, N. R. Random matrix theory. *Acta numerica* **14,** 233–297 (2005).

166. Paul, D. & Aue, A. Random matrix theory in statistics: A review. *Journal of Statistical Planning and Inference* **150,** 1–29 (2014).

167. Wigner, E. P. Characteristic vectors of bordered matrices With infinite dimensions. *Annals of Mathematics* **62,** 548–564 (1955).

168. Wigner, E. P. Characteristic vectors of bordered matrices With infinite dimensions II. *Annals of Mathematics* **65,** 203–207 (1957).

169. Wigner, E. P. On the Distribution of the Roots of Certain Symmetric Matrices. *Annals of Mathematics* **67,** 325–327 (1958).

170. Dyson, F. J. Statistical theory of the energy levels of complex systems. I. *Journal of Mathematical Physics* **3,** 140–156 (1962).

171. Stöckmann, H.-J. *Quantum chaos: an introduction* (American Association of Physics Teachers, 2000).

172. D'Alessio, L., Kafri, Y., Polkovnikov, A. & Rigol, M. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. *Advances in Physics* **65,** 239–362 (2016).

173. Sutherland, B. *Beautiful models: 70 years of exactly solved quantum many-body problems* (World Scientific, 2004).

174. Serbyn, M., Abanin, D. A. & Papić, Z. Quantum many-body scars and weak breaking of ergodicity. *Nature Physics* **17,** 675–685 (2021).

175. Anderson, P. W. Absence of diffusion in certain random lattices. *Physical review* **109,** 1492 (1958).

176. Basko, D. M., Aleiner, I. L. & Altshuler, B. L. Metal–insulator transition in a weakly interacting many-electron system with localized single-particle states. *Annals of physics* **321,** 1126–1205 (2006).

177. Nandkishore, R. & Huse, D. A. Many-body localization and thermalization in quantum statistical mechanics. *Annu. Rev. Condens. Matter Phys.* **6,** 15–38 (2015).

178. Alet, F. & Laflorencie, N. Many-body localization: An introduction and selected topics. *Comptes Rendus Physique* **19,** 498–525 (2018).

179. Abanin, D. A., Altman, E., Bloch, I. & Serbyn, M. Colloquium: Many-body localization, thermalization, and entanglement. *Reviews of Modern Physics* **91,** 021001 (2019).

180. Imbrie, J. Z. Diagonalization and many-body localization for a disordered quantum spin chain. *Physical review letters* **117,** 027201 (2016).

181. Imbrie, J. Z. On many-body localization for quantum spin chains. *Journal of Statistical Physics* **163,** 998–1048 (2016).

182. Roushan, P. *et al.* Spectroscopic signatures of localization with interacting photons in superconducting qubits. *Science* **358,** 1175–1179 (2017).

183. Xu, K. *et al.* Emulating many-body localization with a superconducting quantum processor. *Physical review letters* **120,** 050507 (2018).

184. Choi, J.-y. *et al.* Exploring the many-body localization transition in two dimensions. *Science* **352,** 1547–1552 (2016).

185. Schreiber, M. *et al.* Observation of many-body localization of interacting fermions in a quasirandom optical lattice. *Science* **349,** 842–845 (2015).

186. Wei, K. X., Ramanathan, C. & Cappellaro, P. Exploring localization in nuclear spin chains. *Physical review letters* **120,** 070501 (2018).

187. Žnidarič, M., Prosen, T. & Prelovšek, P. Many-body localization in the heisenberg x x z magnet in a random field. *Physical Review B* **77,** 064426 (2008).

188. Pal, A. & Huse, D. A. Many-body localization phase transition. *Physical review b* **82,** 174411 (2010).

189. Luitz, D. J., Laflorencie, N. & Alet, F. Many-body localization edge in the random-field Heisenberg chain. *Physical Review B* **91,** 081103 (2015).

190. Serbyn, M., Papić, Z. & Abanin, D. A. Local conservation laws and the structure of the many-body localized states. *Physical review letters* **111,** 127201 (2013).

191. Huse, D. A., Nandkishore, R. & Oganesyan, V. Phenomenology of fully many-body-localized systems. *Physical Review B* **90,** 174202 (2014).

192. Bauer, B. & Nayak, C. Area laws in a many-body localized state and its implications for topological order. *Journal of Statistical Mechanics: Theory and Experiment* **2013,** P09005 (2013).

193. Oganesyan, V. & Huse, D. A. Localization of interacting fermions at high temperature. *Physical review b* **75,** 155111 (2007).

194. Bardarson, J. H., Pollmann, F. & Moore, J. E. Unbounded growth of entanglement in models of many-body localization. *Physical review letters* **109,** 017202 (2012).

195. Serbyn, M., Papić, Z. & Abanin, D. A. Universal slow growth of entanglement in interacting strongly disordered systems. *Physical review letters* **110,** 260601 (2013).

196. Page, D. N. Average entropy of a subsystem. *Physical review letters* **71,** 1291 (1993).

197. Alhassid, Y. The statistical theory of quantum dots. *Reviews of Modern Physics* **72,** 895 (2000).

198. Berry, M.-V. Tabor, M1977RSPSA. 356. 375B: Level clustering in the regular spectrum. vol. 356. *Proc. R. Soc. Lond. A,* 375–394 (1977).

199. Pandey, A. & Ramaswamy, R. Level spacings for harmonic-oscillator systems. *Physical Review A* **43,** 4237 (1991).

200. Atas, Y., Bogomolny, E., Giraud, O & Roux, G. Distribution of the ratio of consecutive level spacings in random matrix ensembles. *Physical review letters* **110,** 084101 (2013).

201. Muniz, J. A. *et al.* Exploring dynamical phase transitions with cold atoms in an optical cavity. *Nature* **580,** 602–607 (2020).

202. Smale, S. *et al.* Observation of a transition between dynamical phases in a quantum degenerate Fermi gas. *Science Advances* **5,** eaax1568 (2019).

203. Luitz, D. J., Laflorencie, N. & Alet, F. Extended slow dynamical regime close to the many-body localization transition. *Physical Review B* **93,** 060201 (2016).

204. Doggen, E. V. *et al.* Many-body localization and delocalization in large quantum chains. *Physical Review B* **98,** 174202 (2018).

205. Huse, D. A., Nandkishore, R., Oganesyan, V., Pal, A. & Sondhi, S. L. Localization-protected quantum order. *Physical Review B* **88,** 014206 (2013).

206. Rossini, D., Andolina, G. M. & Polini, M. Many-body localized quantum batteries. *Physical Review B* **100,** 115142 (2019).

207. Ponte, P., Chandran, A., Papić, Z & Abanin, D. A. Periodically driven ergodic and many-body localized quantum systems. *Annals of Physics* **353,** 196–204 (2015).

208. Lazarides, A., Das, A. & Moessner, R. Fate of many-body localization under periodic driving. *Physical review letters* **115,** 030402 (2015).

209. Tangpanitanon, J., Thanasilp, S., Dangniam, N., Lemonde, M.-A. & Angelakis, D. G. Expressibility and trainability of parametrized analog quantum systems for machine learning applications. *Physical Review Research* **2,** 043364 (2020).

210. Keating, J. P., Linden, N., Matthews, J. C. & Winter, A. Localization and its consequences for quantum walk algorithms and quantum communication. *Physical Review A* **76,** 012315 (2007).

211. Schreiber, A *et al.* Decoherence and disorder in quantum walks: from ballistic spread to localization. *Physical review letters* **106,** 180403 (2011).

212. Altshuler, B., Krovi, H. & Roland, J. Anderson localization makes adiabatic quantum optimization fail. *Proceedings of the National Academy of Sciences* **107,** 12446–12450 (2010).

213. Laumann, C. R., Moessner, R., Scardicchio, A. & Sondhi, S. L. Quantum annealing: The fastest route to quantum computation? *The European Physical Journal Special Topics* **224,** 75–88 (2015).

214. Wang, H., Yeh, H.-C. & Kamenev, A. Many-body localization enables iterative quantum optimization. *Nature communications* **13,** 5503 (2022).

215. Chen, J., Nurdin, H. I. & Yamamoto, N. *Towards single-input single-output nonlinear system identification and signal processing on near-term quantum computers* in *2019 IEEE 58th Conference on Decision and Control (CDC)* (2019), 401–406.

216. Marković, D. & Grollier, J. Quantum neuromorphic computing. *Applied physics letters* **117,** 150501 (2020).

217. Fujii, K. & Nakajima, K. Quantum reservoir computing: a reservoir approach toward quantum machine learning on near-term quantum devices. *Reservoir Computing: Theory, Physical Implementations, and Applications,* 423–450 (2021).

218. Negoro, M., Mitarai, K., Nakajima, K. & Fujii, K. Toward NMR quantum reservoir computing. *Reservoir Computing: Theory, Physical Implementations, and Applications,* 451–458 (2021).

219. *Scopus*. https://www.scopus.com. Accessed: 2023-06-27.

220. Ghosh, S., Nakajima, K., Krisnanda, T., Fujii, K. & Liew, T. C. Quantum neuromorphic computing with reservoir computing networks. *Advanced Quantum Technologies* **4,** 2100053 (2021).

221. Vintskevich, S. & Grigoriev, D. *Computing with two quantum reservoirs connected via optimized two-qubit nonselective measurements* in *Quantum 2.0* (2022), QW2A–48.

222. Dudas, J., Grollier, J. & Marković, D. *Coherently coupled quantum oscillators for quantum reservoir computing* in *2022 IEEE 22nd International Conference on Nanotechnology (NANO)* (2022), 397–400.

223. Llodrà, G., Charalambous, C., Giorgi, G. L. & Zambrini, R. Benchmarking the Role of Particle Statistics in Quantum Reservoir Computing. *Advanced Quantum Technologies,* 2200100 (2022).

224. Pfeffer, P., Heyder, F. & Schumacher, J. Hybrid quantum-classical reservoir computing of thermal convection flow. *Physical Review Research* **4,** 033176 (2022).

225. Mlika, Z., Cherkaoui, S., Laprade, J. F. & Corbeil-Letourneau, S. User Trajectory Prediction in Mobile Wireless Networks Using Quantum Reservoir Computing. *arXiv preprint arXiv:2301.08796* (2023).

226. Jaeger, H. & Haas, H. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *science* **304,** 78–80 (2004).

227. Lu, Z. *et al.* Reservoir observers: Model-free inference of unmeasured variables in chaotic systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **27,** 041102 (2017).

228. Ganguli, S., Huh, D. & Sompolinsky, H. Memory traces in dynamical systems. *Proceedings of the national academy of sciences* **105,** 18970–18975 (2008).

229. Couillet, R., Wainrib, G., Sevi, H. & Ali, H. T. The asymptotic performance of linear echo state neural networks. *The Journal of Machine Learning Research* **17,** 6171–6205 (2016).

230. Tiňo, P. Asymptotic Fisher memory of randomized linear symmetric Echo State Networks. *Neurocomputing* **298,** 4–8 (2018).

231. Gonon, L., Grigoryeva, L. & Ortega, J.-P. Memory and forecasting capacities of nonlinear recurrent networks. *Physica D: Nonlinear Phenomena* **414,** 132721 (2020).

232. Tino, P. Dynamical systems as temporal feature spaces. *The Journal of Machine Learning Research* **21,** 1649–1690 (2020).

233. Zobov, V. E. & Lundin, A. A. On the effect of an inhomogeneous magnetic field and many-body localization on the increase in the second moment of multiple-quantum NMR with time. *JETP Letters* **105,** 514–518 (2017).

234. Peng, P. *NMR studies of quantum thermalization* PhD thesis (Massachusetts Institute of Technology, 2022).

235. Houck, A. A., Türeci, H. E. & Koch, J. On-chip quantum simulation with superconducting circuits. *Nature Physics* **8,** 292–299 (2012).

236. Elben, A. *et al.* The randomized measurement toolbox. *Nature Reviews Physics,* 1–16 (2022).

237. Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nature Physics* **16,** 1050–1057 (2020).