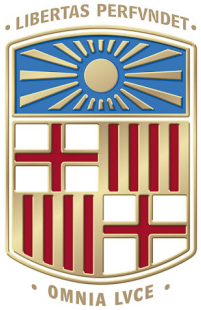


UNIVERSITAT DE BARCELONA

Master in Quantum Science and Technology

Variational Neural Network Methods For Open Quantum Systems

Mateusz Molenda

Supervisor: Mariona Moreno Cardoner, Arnau Rios Huguet

Departament de Física Quàntica i Astrofísica
Universitat de Barcelona

July 07, 2024

Variational Neural Network Methods For Open Quantum Systems

Mateusz Molenda

Supervised by: Mariona Moreno Cardoner, Arnau Rios Huguet

Departament de Física Quàntica i Astrofísica, Universitat de Barcelona (UB), 08028 Barcelona, Spain
07 July 2024

Abstract: This study aims to investigate the application of neural network ansätze to open quantum systems with both local and non-local dissipation. Simulating quantum systems has always been a challenging task due to the rapidly growing dimensions of the Hilbert space with the number of constituents. This is especially severe in the case of open systems, which need to be described by density matrices instead of wavefunctions. Recent developments in machine learning techniques offer a promising alternative to efficiently represent quantum many body systems. In this work, we focus on a specific example of neural networks: Restricted Boltzmann Machines with an additional ancillary layer representing the degrees of freedom of the external bath and ensuring a positive semi-definite density matrix. We implement this ansatz for models of spin-$\frac{1}{2}$ chains in the presence of an external magnetic field, with local and non-local dissipation. For a particular choice of the Hamiltonian parameters, this also corresponds to a realistic model describing a dense chain of two-level atoms that are dipole-dipole interacting and driven by an external field, where collective effects such as superradiance and subradiance emerge. We find that the neural network not only accurately predicts the observables of interest (mean values and correlations in the spin operators), but it also reproduces the exact steady state with high fidelity ($>$0.98). Moreover, the ansatz performs well and yields sensible results even for system sizes too large to be simulated by direct integration of the master equation. Although some error is present, ongoing research suggests that it will be possible to further minimize it.

Mateusz Molenda: mati.00.molenda@gmail.com

Contents

1 Introduction **1**

2 Open quantum systems **2**
2.1 System representation and dynamics 2
2.2 Liouville space representation 3
2.3 Steady state 4

3 Models of interest **4**
3.1 Dissipative Transverse Ising Model 4
3.2 Chain of two-level atoms with light-induced dipole-dipole interactions 5

4 Neural networks for quantum many body physics **6**
4.1 Variational methods 6
4.2 Application to closed quantum systems 6
4.3 Application to open quantum systems 7
4.4 Cost function and optimization 8
4.5 Sampling 10

5 Implementation of RBM based ansatz and results **11**
5.1 Transverse Ising model with local dissipation 11
5.2 Transverse Ising model with non-local dissipation 11
5.3 Hyperparameter tuning 12
5.4 Chain of two-level atoms with light-induced dipole-dipole interactions 14

6 Conclusion and outlook **18**

Bibliography **19**

1 Introduction

In recent decades, machine learning (ML) algorithms have revolutionized numerous fields, offering unprecedented capabilities in data processing, pattern recognition, and predictive modeling [1]. Algorithms that once struggled with simple tasks can now perform complex functions with high accuracy. The impact of ML is evident across various domains. By automating and optimizing processes, ML algorithms enhance efficiency, reduce human error, and open new avenues for scientific exploration [2, 3, 4].

In physics, the advancement of ML algorithms is providing a new way for state representation and simulation of physical systems. Solving quantum many-body problems is notoriously difficult due to the exponential growth of the state space with the number of particles, leading to what is known as the "curse of dimensionality" [5]. Traditional methods, such as perturbation theory and variational approaches, often fall short in capturing the full complexity of many-body systems. This is where machine learning algorithms come into play. By leveraging their ability to identify patterns and make predictions based on large datasets, ML algorithms offer new ways to tackle the computational challenges inherent in many-body physics.

One of such applications are the neural-network quantum states (NQSs) [6]. They are used for the variational representations of a quantum many-body wavefunction. The first application of such networks, using a simple architecture based on Restricted Boltzmann Machines (RBM), was introduced in 2017 by Carleo and Troyer [7], where they successfully used it for calculating the ground state energy. Shortly after, Torlai et al. [8] proposed an ansatz architecture, also based on RBM with an additional ancillary layer representing the environment, that could provide an accurate representation of open quantum systems. As shown in [9, 10, 11], it is capable of predicting the dissipative properties and evolution of spin-$\frac{1}{2}$ chains with local dissipation interacting with an external magnetic field.

Open quantum systems, which interact with their external environment, are essential for understanding a wide range of phenomena in quantum mechanics, from decoherence and dissipation to quantum thermodynamics and quantum information processing [12, 13, 14]. Unlike closed quantum systems, which are isolated from their surroundings, open systems exchange energy and information with their environment, leading to complex dynamics that are often challenging to model and predict. The complexity of open quantum systems arises from the need to account for both the system and its interaction with the environment. Traditional approaches, such as the Lindblad master equation and the stochastic Schrödinger equation, provide frameworks for describing these interactions but can become computationally intensive and intractable for large systems. The NQSs seems to be able to provide a representation for the open systems of the spin-$\frac{1}{2}$ chains. In this work, we aim to reproduce previous results on spin-$\frac{1}{2}$ chains and extend the application to non-local dissipation models, focusing on a dense chain of two-level atoms driven by light. Typically, diluted atomic ensembles treat spontaneous emission as a single-atom process, but we describe a regime where atoms are close enough to couple to the same radiation mode, leading to light-induced dipole-dipole interactions and collective photon emission. We intend to investigate whether the neural network ansatz effectively captures collective radiance and atomic correlations induced by the external bath.

This report will be divided into five more sections. In Sec.2 we explain the theoretical fundamentals of open quantum systems. After that, Sec.3 will provide the description of the physical models that we wish to simulate. Then, in Sec.4 we will explain the methodologies used. In Sec.5 we will present the obtained results. Finally, in Sec.6 we will provide our conclusions and outlook on the possible directions for further development.

2 Open quantum systems

Quantum mechanics traditionally deals with closed systems, where the system is isolated from its environment and evolves deterministically according to the Schrödinger equation. However, most realistic quantum systems interact with their surroundings, leading to the concept of open quantum systems. These interactions introduce non-unitary dynamics that can cause decoherence and dissipation, essential for understanding and further development in fields such as quantum computing, quantum optics and chemical physics [15, 16, 17].

2.1 System representation and dynamics

Open quantum systems are described using the density matrix formalism, as using only state vectors $|\psi\rangle$ defined on a Hilbert space would be an idealized case that cannot fully characterize the statistical mixtures, which we encounter much often in nature. Thus, when using the density matrix ρ, we are able to provide a complete description of the system's statistical state. It has to fulfill a set of properties in order to be valid: hermicity $\rho^\dagger = \rho$, normalization $\mathrm{Tr}\rho = 1$ and positive semi-definiteness $\langle A|\rho|A\rangle \geq 0$ (for any state A). They can be divided into pure state and mixed state density matrices, where the mixed state is a convex sum of pure state density matrices: $\rho_{\mathrm{mix}} = \sum_i p_i \rho_i^{\mathrm{pure}}$ ($p_i \geq 0$, $\sum_i p_i = 1$). The time evolution of the density matrix of a closed system is governed by the Von Neumann's equation:

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho]. \tag{1}$$

However, in an open system the density matrix ρ also describes the interaction with its environment. Such combination lives in the Hilbert space $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$ that is composed of two subsystems, closed system and environment. The generalized Hamiltonian for such systems can be represented as [18]:

$$H = H_S + H_E + H_{SE}, \tag{2}$$

where H_S and H_E are system and environment Hamiltonians respectively, and H_{SE} is the interaction term. This interaction leads to the exchange of energy and information between the system and its surroundings. The Von Neumann's equation, which describes the evolution of only closed systems, is thus insufficient to account for these effects.

In order to describe the reduced density matrix of the system, which is defined as $\rho(t) = \mathrm{Tr_E}[\rho_{\mathrm{total}}]$, we will use here the Born-Markov approximation [19]. The Born approximation assumes that the coupling between the system and the environment is weak and it can be treated perturbatively, and also, that the bath correlation functions are not significantly changed by the interaction. On the other hand, the Markov approximation assumes that the bath is memoryless, as its correlation times are much shorter compared to those of the system. These approximations allow to obtain a first order diferential equation for the density matrix, of the Lindblad form [20]:

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] + \sum_{i,j} \frac{\gamma_{ij}}{2} \left(2L_i \rho L_j^\dagger - \{L_j^\dagger L_i, \rho\}\right), \tag{3}$$

where H is an hermitian operator which can be understood as the effective Hamiltonian of the system, γ_{ij} are the dissipation coefficients that can be represented by a positive semi-definite matrix ($\sum_{i,j} x_i^* \gamma_{ij} x_j \geq 0$ for all vectors x) [21], and L_i and L_j are the dissipation operators. The commutator part comes naturally from the density matrix representation of the closed system and accounts for unitary dynamics. The remaining sum terms (Lindblad

terms) represent the coupling to the environment and have non-unitary behaviour, thus the whole evolution of open quantum system is also non-unitary and non-reversible in consequence. This shows that the interaction with the environment modifies the whole system's evolution. The associated operators, called jump operators, are responsible for two important effects: decoherence and dissipation. Decoherence, as for instance, the dephasing mechanism, causes the system to loose its coherent superposition states, leading to classical-like probabilistic mixtures. This loss of coherence is irreversible because the precise information about quantum superpositions is lost to the environment. Dissipation, which can also be considered as a specific form of decoherence, is the energy exchange with the environment that leads to relaxation processes where the system reaches thermal equilibrium with its surroundings. This process is also typically irreversible because the exact energy states of the system before interaction cannot be fully recovered.

2.2 Liouville space representation

Liouville space, also known as the space of operators or superoperator space, is a vector space where operators like the density matrix are treated as vectors. In its core, this is interpreted as a doubled Hiblert space, $\mathcal{H}_L = \mathcal{H} \otimes \mathcal{H}$ [22]. This representation is particularly useful for open quantum systems, where we have to deal with non-unitary dynamics. One of the primary benefits of using Liouville space is the simplification of the Lindblad master equation. In this formulation it becomes a linear differential equation, similar to the Schrödinger equation for pure states. This linearity facilitates analytical and numerical solutions, making it easier to study the system's dynamics. Another advantage is the ability to use standard linear algebra techniques and tools. Operations such as matrix multiplications, eigenvalue decompositions, and solving linear systems become straightforward, leveraging well-established methods from linear algebra. This compatibility enhances computational efficiency and the applicability of various numerical algorithms. In order to see this, let's first start with the representation of the density matrix in its vectorized form, which we will denote as $|\rho\rangle\rangle$,

$$\rho = \sum_{\sigma\tau} \rho_{\sigma\tau} |\sigma\rangle\langle\tau| \Longrightarrow |\rho\rangle\rangle = \frac{1}{C} \sum_{\sigma\tau} \rho_{\sigma\tau} |\sigma, \tau\rangle\rangle, \tag{4}$$

where $|\sigma, \tau\rangle\rangle = |\sigma\rangle \otimes |\tau\rangle$ is the basis configuration that spans $\mathcal{H} \otimes \mathcal{H}$ and $C = \sqrt{\sum_{\sigma\tau} |\rho_{\sigma\tau}|^2}$ is the normalization factor. This way we are able to rewrite Eq.3 as a linear differential equation

$$\partial_t |\rho\rangle\rangle = \mathcal{L} |\rho\rangle\rangle, \tag{5}$$

where $\mathcal{L}$ is the Liouvillian superoperator, and it can be represented as

$$\mathcal{L} = -\frac{i}{\hbar}(H \otimes \mathcal{I} - \mathcal{I} \otimes H^T) + \sum_{ij} \gamma_{ij} \left(L_i \otimes L_j^* - \frac{1}{2} L_i^\dagger L_j \otimes \mathcal{I} - \mathcal{I} \otimes \frac{1}{2} L_j^T L_i^* \right). \tag{6}$$

Here H is the Hamiltonian, γ_{ij} are the dissipation coefficients and L_i, L_j are the jump operators as in Eq.3. We will see later that this formulation proves useful, when considering a choice for the appropriate cost function.

The evolution of the state can then be formally written as :

$$|\rho(t)\rangle\rangle = e^{\mathcal{L}t} |\rho(0)\rangle\rangle, \tag{7}$$

where t is the time and $|\rho(0)\rangle\rangle$ is the initial density matrix of the system at $t = 0$.

2.3 Steady state

A relevant state for many applications in quantum optics and quantum computation is the steady state of the system, which is reached when the time derivative vanishes:

$$\partial_t \rho = 0. \tag{8}$$

The above can also be rewritten in the Liouvillian space as

$$\mathcal{L}|\rho_{ss}\rangle\rangle = 0, \tag{9}$$

which in most cases will correspond to the zero eigenvalue of the Liouvillian superoperator. It is also essential to analyze the stability of the steady state. Perturbations or deviations from the steady state should decay over time if the steady state is stable. It can be interpreted as the stabilization of the interactions with the environment. In experiments and practical applications, it is often desirable to study the behavior of systems after reaching the steady state. This is because many measurements and observations are more meaningful when the system has settled into a stable state, allowing to study its properties without the complications of transient dynamics and with observables remaining constant over time.

It has been proven for systems with finite Hilbert space dimension, that the significant majority has a unique steady state [23, 24]. Nevertheless, it is possible to construct the system in such way that it will have more than one unique steady state [25]. A simple example of this situation is a single qubit which undergoes dephasing on two of the three axes of the Bloch sphere [23, 24].

3 Models of interest

In this section we will present the different models that will be studied using the neural network approach. The first model is the paradigmatic transverse Ising model of a chain of spin-$\frac{1}{2}$ particles with both local and non-local dissipation. The second model corresponds to a dense one-dimensional chain of two-level atoms driven by light and dipole-dipole interacting, which can collectively radiate. Our goal is to determine whether the steady state of these models can be efficiently described using a neural network variational ansatz adapted to open quantum systems.

3.1 Dissipative Transverse Ising Model

Spin chains are a fundamental topic in condensed matter physics and quantum mechanics. Moreover, spin chains have practical implications in the development of quantum computing and information storage technologies. One of the most studied models in this field is the transverse Ising model, which provides a simple yet powerful framework to understand phase transitions and critical phenomena. The Ising model, with its clear mathematical formulation, allows for the exploration of key concepts such as spontaneous magnetization and symmetry breaking. Studying spin chains through the Ising model also offers insights into more complex systems, helping to develop computational techniques and theoretical tools that are widely applicable in physics and material science. Its Hamiltonian can be written as:

$$H = \frac{V}{4} \sum_{i=0}^{L-1} \sigma_i^z \sigma_{i+1}^z + \frac{g}{2} \sum_{i=0}^{L-1} \sigma_i^x. \tag{10}$$

The first terms describes spin-spin interactions, which depend only on the nearest-neighbor alignment in the z-direction, with V being the coupling strength. The second term is the uniform transverse magnetic field acting in the x-direction, with g being the field strength.

The model will be solved for two types of dissipation: local and non-local. For the local case we will consider 1-body dissipation operators, which are given by $L_i = \sigma_i^- = \frac{1}{2}(\sigma_i^x - i\sigma_i^y)$. This physically represents the spin-flip processes, where the jump operators act destructively locally on each site. The non-local dissipation will be modeled here as 2-body nearest-neighbor operators (for a more precise definition, see Eq.33).

The local dissipation case will be used as a benchmark of our method, as the results are already given in [9]. The non-local case, instead, is still unexplored in relation to neural networks, and we wish to see if the simulation is able to provide an accurate representation of the steady state of the system.

3.2 Chain of two-level atoms with light-induced dipole-dipole interactions

The interaction between neutral atoms and light is a fundamental aspect of quantum physics, underpinning a wide range of phenomena and technologies [26]. From the precision measurements in atomic clocks [27] to the manipulation of quantum states in quantum information science [28], understanding these interactions is crucial for both theoretical and practical advancements. At the heart of these interactions is the process of absorption and emission of photons by atoms. When a neutral atom encounters a photon, it can absorb the photon's energy, elevating an electron to a higher energy level. Here we will explain the model describing this process which will be captured by the neural network.

We want to start by considering a single neutral atom in a quantized field, in order to understand how it behaves. It is also necessary to introduce a set of assumptions. We will be working with two-level atoms interacting with monochromatic light, meaning that we are considering only near-resonant interactions, since transitions to off-resonant states have weaker coupling to the laser field. The atomic states are denoted as $|g\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (ground state) and $|e\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (excited state), with transition frequency ω_0. Due to interactions with the environment, which in our case is the quantized electromagnetic field, an atom in the excited state can spontaneously emit a photon, decaying to the ground state. This is possible because of the field's vacuum fluctuations. The evolution of this system is governed by a Lindblad equation, which for the case of one atom is given by:

$$\frac{d\rho}{dt} = -i[H, \rho] + \Gamma_0 \left(\sigma^- \rho \sigma^+ - \frac{1}{2}\{\sigma^+ \sigma^-, \rho\} \right), \tag{11}$$

where Γ_0 is the spontaneous emission rate, and lowering and raising atomic operators are respectively $\sigma^- = |g\rangle\langle e|$, $\sigma^+ = |e\rangle\langle g|$. The total Hamiltonian H contains the free atom contribution and when it is present, the interaction with an additional external driving field (in our case, classical). The final form will be similar to the one in Eq.2:

$$H = H_{\mathrm{A}} + H_{\mathrm{AF}} = \hbar\omega_0 |\mathrm{e}\rangle\langle \mathrm{e}| + \hbar g(\sigma^- + \sigma^+), \tag{12}$$

where H_{A} is the atomic part, and H_{AF} is the interaction part with the classical field, with g being the corresponding Rabi frequency.

When a chain of atoms couple to the same radiation modes, the atoms experience dipole-dipole interactions mediated by the exchange of a virtual photon, together with

collective dissipation. The master equation then generalizes into the form:

$$\frac{d\rho}{dt} = -i[H,\rho] + \sum_{i\neq j} \frac{\Gamma_{ij}}{2}\left(2\sigma_j^- \rho \sigma_i^+ - \{\sigma_i^+\sigma_j^-,\rho\}\right), \tag{13}$$

where $H = H_{\mathrm{A}} + H_{\mathrm{AF}}$, with $H_{\mathrm{A}} = \sum_{ij} J_{ij}\sigma_i^+\sigma_j^-$ corresponding to the light-mediated dipole-dipole atomic interactions, and $H_{\mathrm{AF}} = \sum_i g_i(\sigma_i^- + \sigma_i^+)$ to the interaction with the external field (with local Rabi frequency g_i). The parameters J_{ij} and Γ_{ij} are, respectively, the dispersive and dissipative light-mediated couplings, which depend on the interparticle distance as we will see later in more detail.

4 Neural networks for quantum many body physics

In recent years, neural networks have emerged as a powerful tool in the study of many-body physics, revolutionizing our understanding and approach to complex quantum systems [7, 8, 9]. One of the main problems in quantum many-body physics is finding the computationally efficient manner of expressing the wave function. If we would consider a N-spin many-body quantum system, the number of parameters that we need for the most general wavefunction grows exponentially as 2^N. The general wavefunction equation for this case can be represented as a superposition of spin states:

$$|\Psi\rangle = \sum_{\sigma_1,\sigma_2,\ldots,\sigma_N} C_{\sigma_1,\sigma_2,\ldots,\sigma_N}|\sigma_1\rangle \otimes |\sigma_2\rangle \otimes \ldots \otimes |\sigma_N\rangle, \tag{14}$$

where $|s\rangle = |\sigma_1\rangle \otimes |\sigma_2\rangle \otimes \ldots \otimes |\sigma_N\rangle$ are the basis vectors spanning the Hilbert space of interest and $C_{\sigma_1,\sigma_2,\ldots,\sigma_N}$ are the coefficients of each of those vectors.

4.1 Variational methods

In order to tackle the problem of exponential cost, variational methods have been developed. The variational principle, which underpins these methods, offers a systematic way to approximate the ground state and low-lying excited states of many-body systems. The usefulness of the variational ansatz lies in its ability to balance accuracy and computational feasibility. By constructing trial wave functions with adjustable parameters, the variational method allows to optimize these parameters. This approach not only provides insights into the ground state properties but also helps in understanding the correlations and dynamics of the system [7, 10, 11].

We can distinguish between two main types of useful variational ansätze. The first type allows us to compute exact expectation values with a polynomial computational cost. The second type enables approximate expectation values, where the accuracy can be systematically improved with a polynomial increase in computational cost relative to the system size [29]. Here our interest will be the second type, which can be called Computationally Tractable States. They have to obey the following conditions: (a) it is possible to sample in poly(n) time with classical means from the probability distribution, $P(s) = \frac{|\langle s|\Psi\rangle|^2}{\langle\Psi|\Psi\rangle}$, (b) amplitudes for any basis element, $\Psi(s) = \langle s|\Psi\rangle$, can be computed in poly(n) time on a classical computer [30].

4.2 Application to closed quantum systems

Here we will start introducing and explaining the ansatz architecture and application to the closed systems, where the objective will be mainly to find the ground state of the

given system. It is beneficial to first understand this case in order to fully comprehend the necessary differences when transitioning to open quantum systems. There are many models that are being tested, however we want to focus on the Restricted Boltzmann Machine (RBM) proposed by Geoffrey Hinton [31]. This is an energy based model, where the general energy function is given by

$$E(v,h) = -a^T v - b^T h - v^T W h = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i W_{ij} h_j, \tag{15}$$

where $v = (v_1, v_2, ..., v_N)$ and $h = (h_1, h_2, ..., h_M)$ are neuron variables of the input and hidden layers, $a = (a_1, a_2, ..., a_N)$ and $b = (b_1, b_2, ..., b_M)$ are local biases, and W are the weights between them. Both input and hidden layers have binary values. The term "restricted" comes from the fact that connections between layers of the same type is not allowed (see Fig.1).

We can make use of this network in order to describe a many-body quantum system that will have two local degrees of freedom (chains of spin-$\frac{1}{2}$ particles, two level neutral atoms, etc.). So, if we assume that the Hilbert space of our system will be spanned by the computational basis $|\sigma\rangle$, where $\sigma = (\sigma_1, \sigma_2, ..., \sigma_N)$, where each σ_i has two local degrees of freedom, the variational wavefunction can be represented as

$$\Psi(\sigma) = \sum_{h_j} e^{E(\sigma,h)} = \sum_{h_j} e^{-\sum_i a_i\sigma_i - \sum_j b_j h_j - \sum_{i,j} \sigma_i W_{ij} h_j}, \tag{16}$$

where $h_j = \{-1, 1\}$ and the summation runs over the M hidden spin variables. Now comes the most special part about this type of ansatz, which is that we are able to trace out analytically the hidden layer variables, so the wavefunction will read

$$\Psi(\sigma) = e^{-\sum_i a_i\sigma_i} \prod_i^M 2\cosh\left(b_i - \sum_j \sigma_j W_{ij}\right). \tag{17}$$

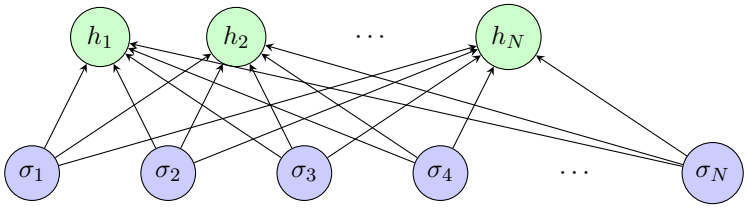
The weights are allowed to take complex values, in order to provide the description of wavefuntion parameters. This type of ansatz has been successfully applied in [7], for calculation of the ground state energy and simulation of unitary dynamics of a closed systems of spin-$\frac{1}{2}$ atomic chains, governed by the transverse Ising or the Heisenberg models.

4.3 Application to open quantum systems

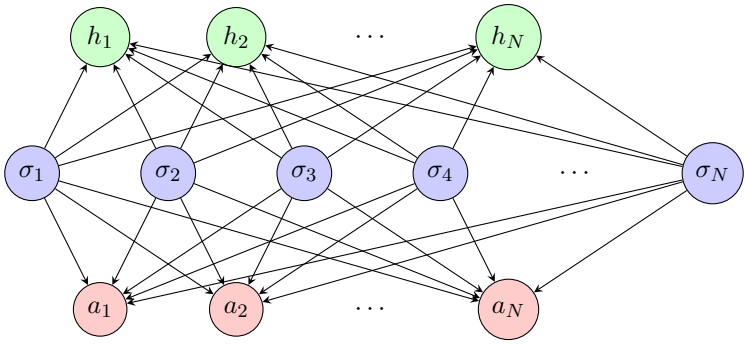
We will now move on to the case of open quantum systems, which poses additional challenges due to the requirements of the density matrix described in Sec.2. The main difficulty here is due to the fact that we have to describe our state using a density matrix, which has to be hermitian, positive semi-definite and has trace one. To overcome this challenge we make use of the ansatz proposed in [8], which is called Neural Density Matrix ansatz (NDM). Here to ensure the validity of the density matrix, it has been suggested to include an additional layer called ancillary layer, that is responsible for purification of the density matrix and representation of the degrees of freedom of the environment. The basis vector set here can be described as $\{|\sigma, a\rangle\}$, where $a = (a_1, a_2, ..., a_N)$. In this formulation the reduced density matrix of a system can be written as

$$\rho(\sigma, \sigma') = \sum_a \psi(\sigma, a)\psi^*(\sigma', a), \tag{18}$$

where we want to represent $\psi(\sigma, a)$ with the neural network. We choose the functions to be of the form proposed in [8]:



(a) RBM with Hidden Layer Only



(b) RBM with Hidden and Ancillary Layer

Figure 1: Graphical representation of a Restricted Boltzmann machine for (a) closed system and (b) open systems. Each neuron (node) of the input (purple), hidden (green) or ancillary (pink) layer can take binary values. There are no interalayer connections, only between input and hidden or ancillary layers. The weights between different layers can take complex values, except for the connections between input and ancillary layer, which are real due to restrictions imposed by the density matrix. The input is the basis state configuration with two local degrees of freedom, $|s\rangle = |\sigma_1\rangle \otimes |\sigma_2\rangle \otimes ... \otimes |\sigma_N\rangle$.

$$\psi(\sigma, a) = \sqrt{\mathcal{P}_{v_A}(\sigma, a)} e^{-\frac{1}{2}\log\left(\mathcal{P}_{v_\theta}(\sigma,a)\right)}, \tag{19}$$

where $\mathcal{P}_{v_A}(\sigma, a)$ and $\mathcal{P}_{v_\theta}(\sigma, a)$ are amplitude and phase-related functions respectively. They can be calculated from the general energy-like function, similar to Eq.15, but including the purification. So, we get $\mathcal{P}_v(\sigma, a) = \sum_h e^{-\mathrm{E}_v(\sigma,a,h)}$ $(v \in \{v_A, v_\theta\})$, where the energy function is

$$E_v(\sigma, a, h) = \sum_i \sigma_i b_i^{(\sigma)} + \sum_j a_j b_j^{(a)} + \sum_k h_k b_k^{(h)} + \sum_{ik} \sigma_i W_{ik} h_k + \sum_{ij} \sigma_i U_{ij} a_j, \tag{20}$$

where σ_i, a_j, h_k represent the neuron values of the input, ancillary and hidden layers, respectively, and W_{ik}, U_{ij} are the weights between input-hidden, and input-ancillary layers. The graphical representation of both architectures is presented in Fig.1.

4.4 Cost function and optimization

In variational methods, selecting the appropriate cost function is crucial for the effectiveness and accuracy of the method. The cost function must meet specific requirements to ensure reliable results. Primarily, it needs to be real-valued. These properties guarantee that the function can be minimized effectively, leading to meaningful physical interpretations and solutions.

For closed systems, and to find the ground state of the system, the energy or the expectation value of the Hamiltonian is the most natural choice of the cost function. In such case the minimization is done iteratively over:

$$E_T = \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}, \tag{21}$$

where E_T is the expectation value of Hamiltonian, and $|\Psi_T\rangle$ represents the trial state vector of the closed system. Here E_T is the upper bound on the ground state energy $E_{\min}$, $E_T \geq E_{\min}$.

However, when dealing with open systems, the selection of a suitable cost function becomes slightly more intricate, due to interaction with the environment, leading to more complex dynamics. In order to find the steady state we need to minimize the derivative of the density matrix. In this context, the Liouvillian operator $\mathcal{L}$ comes into play. The Liouvillian operator is generally non-Hermitian, which means its eigenvalues are complex. This complexity arises from the open nature of the system, where energy and particles can exchange with the surroundings, leading to dissipative processes.

Despite this complexity, there is a way to construct a suitable cost function for open systems. By considering the product of the Liouvillian operator and its Hermitian conjugate, $\mathcal{L}^\dagger\mathcal{L}$, we obtain an operator with a real-valued, non-negative spectrum. This transformation is significant because it aligns with the requirements for the cost function in variational methods. The eigenvalues of $\mathcal{L}^\dagger\mathcal{L}$ are non-negative, and the lowest eigenvalue is zero. This zero eigenvalue corresponds to the steady state of the system, which is a critical aspect of open systems as it represents the long-term behavior of the system when it reaches equilibrium.

Thus, $\mathcal{L}^\dagger\mathcal{L}$ emerges as the optimal choice for the cost function in variational methods for open systems. It not only meets the necessary criteria of being real-valued but also provides valuable insights into the steady-state properties of the system. The cost function can be described as:

$$\mathcal{C} = \langle\langle \mathcal{L}^\dagger\mathcal{L} \rangle\rangle = \frac{\langle\langle \rho|\mathcal{L}^\dagger\mathcal{L}|\rho\rangle\rangle}{\langle\langle \rho|\rho\rangle\rangle}, \tag{22}$$

where $|\rho\rangle\rangle$ is density matrix expressed as a vector in Liouville space.

So our task now essentially becomes the optimization problem of finding the global minimum of $\langle\langle \mathcal{L}^\dagger\mathcal{L}\rangle\rangle$, which we know happens when $\mathcal{L}|\rho\rangle\rangle = 0$. In order to do that, we will use the stochastic gradient descent method combined with stochastic reconfiguration. Stochastic gradient descent is a well-known algorithm for cost function minimization [32]. It requires calculating the gradient of the cost function at each step of the minimization process and proposes that the state vector is updated by moving in the direction opposite to the gradient, which corresponds to the steepest descent direction:

$$|\rho_{\text{new}}\rangle\rangle = |\rho_{\text{old}}\rangle\rangle - \mu\nabla\mathcal{C}(\rho_{\text{old}}), \tag{23}$$

where μ is the learning rate, a scalar number, that allows us to control the speed of the convergence.

Stochastic gradient descent often alone might no be enough, thus it can be improved by including stochastic reconfiguration, a method that considers the geometry of the parameter space, which leads to more efficient optimization [33]. We do that by calculating the metric tensor S, which represents the overlap between small changes in the state vector. The metric tensor is given by the expectation value of the Fisher information matrix:

$$S_{ik} = \langle O_i^\dagger O_k\rangle - \langle O_i^\dagger\rangle\langle O_k\rangle, \tag{24}$$

where $i, k = 1, 2, ..., N_p$ (N_p is the number variational parameters), and O_i, O_k are the log derivatives of the density matrix

$$\langle\sigma|O_k|\sigma'\rangle = \partial_k \log \rho_\chi(\sigma, \sigma'), \tag{25}$$

where $\chi = \{\chi_1, \chi_2, ..., \chi_{N_p}\}$ is the set of variational parameters. Now, instead of directly using the gradient for the update, we need to solve the linear system:

$$\delta\rho = S^{-1}\nabla\mathcal{C}(\rho), \tag{26}$$

where $\delta\rho$ represents the update direction for the state vector and at each iteration we perform the optimization according to this rule:

$$|\rho_{\text{new}}\rangle\rangle = |\rho_{\text{old}}\rangle\rangle - \mu\delta\rho. \tag{27}$$

4.5 Sampling

The evaluation of the expectation values in Eq.24 at each iteration step, as well as the observables of interest for the converged solution, cannot be performed exactly using the full Hilbert space due to computational limitations. Instead, we need to sample the value of the operators over a large enough set of random spin configurations in an unbiased manner, as we will discuss in the following.

We will start with a randomly generated initial configuration, which will be changing locally, based on a predefined condition. This way we ensure exploring the biggest possible part of the Hilbert space. For that, we want to use the Metropolis-Hastings algorithm [34]. The algorithm allows us to randomly sample from a distribution, due to combining Monte Carlo sampling with the dynamics of Markov chains (MCMC). The advantage of MCMC is that after a certain number of iterations all states will be sampled in a statistically representative way, regardless of the staring point [35]. Thus it is beneficial to discard the initial points and choose samples after a certain number of iterations. Metropolis-Hastings specifies the transition kernel $T(\sigma \to \sigma')$, which is the probability that state σ transitions to σ'. Markov process is defined in a way, that it satisfied detailed balance condition

$$P(\sigma)T(\sigma \to \sigma') = P(\sigma')T(\sigma' \to \sigma). \tag{28}$$

The transition probability can be split into two quantities

$$T(\sigma \to \sigma') = Q(\sigma \to \sigma')A(\sigma \to \sigma'). \tag{29}$$

where $Q(\sigma \to \sigma')$ is the proposed transition distribution and $A(\sigma \to \sigma')$ is the acceptance probability to go from current configuration to the next, defined such that

$$A\left(\sigma \to \sigma'\right) = \min\left(1, \frac{P\left(\sigma'\right)Q\left(\sigma' \to \sigma\right)}{P(\sigma)Q\left(\sigma \to \sigma'\right)}\right). \tag{30}$$

Because we consider the ratio $P(\sigma')/P(\sigma)$ the probability distribution does not need to be normalized, thus we also do not need to normalize the variational ansatz, due to definition of Born distribution. In order to determine whether the sample is accepted, we draw a number from the uniform distribution $U \in (0, 1)$. If $A\left(\sigma \to \sigma'\right) \geq U$ then the step is accepted and analogously if $A\left(\sigma \to \sigma'\right) < U$ then it is rejected, and the state does not change. As mentioned before our state will correspond to the chain of basis states with two local degrees of freedom $\sigma = (\sigma_1, \sigma_2, ..., \sigma_i, ..., \sigma_N)$, where they are represented by binary values. The transition will correspond here to randomly flipping one of the states $\sigma = (\sigma_1, \sigma_2, ..., \sigma_i', ..., \sigma_N)$ and if we accept the new configuration. It is better to choose kernel that acts locally because it is harder to find normalized global kernel, also global changes can often lead to unexpected errors if we consider bigger systems.

5 Implementation of RBM based ansatz and results

In this section we present the implementation of the RBM to the previously described models and the results using this approach. The neural network ansatz has been previously successfully applied to spin-$\frac{1}{2}$ atomic chains with local dissipation, to find the steady state and the dynamics of the system [9, 10, 11]. These results have been highly promising, yet the focus has been predominantly to cases involving local dissipation. Here we want to explore the possibility of including non-local dissipation operators and check if the neural network ansatz is also capable of accurate representation of such systems. The simulations were performed with the help of NetKet, an open-source project with optimized methods for the study of quantum many-body physics using artificial neural networks. [36].

5.1 Transverse Ising model with local dissipation

The first model that will be studied is the 1D transverse field Ising model with local dissipation. This will be used as a benchmark, as the results have been presented in [9]. The Lindblad master equation is

$$\frac{d\rho}{dt} = -i[H,\rho] + \frac{\gamma}{2}\sum_{i=0}^{L-1}\left(2\sigma_i^-\rho\sigma_i^+ - \left\{\sigma_i^+\sigma_i^-,\rho\right\}\right), \tag{31}$$

where the Hamiltonian is the same as in Eq.10, the dissipation operators σ_i^-, σ_i^+ act locally on each side and γ is the dissipation rate, with the same value for all terms. The system has periodic boundary condition such that $\sigma_L^a = \sigma_0^a$ (where $a \in \{x,y,z\}$). The results for the mean magnetization value $\langle S_a\rangle$ (a = x,y,z) of the steady state, as a function of the ratio g/γ are presented in Fig.2 (orange dots). In the same figure, we also plot the results obtained by exact diagonalization of the Liouvillian superoperator (blue solid line). We can see that the simulations accurately predict magnetizations in y- and z-direction when the system is in the steady state. However the magnetization in x-direction shows a visible discrepancy for values $g/\gamma \sim 1$. In order to fix it we tried different configurations of network parameters (α, β), and also different learning rate values for the stochastic gradient descent, however without significant improvements. The similar problem was encountered in [9], which may suggest that this error could come from the ansatz architecture itself. Nevertheless, the fidelity between the variational ansatz and the steady state obtained by exact diagonalization of the Liouvillian, defined as

$$\mathcal{F}(\rho,\sigma) = \left(\mathrm{Tr}\left[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right]\right)^2, \tag{32}$$

is larger than 0.997 for all the studied points.

5.2 Transverse Ising model with non-local dissipation

We add now non-local dissipation to the previous 1D Ising model, set as two-body nearest-neighbour operators in the Lindblad term:

$$D[\sigma] = \frac{\gamma}{2}\sum_{i=0}^{L-1}\left(2\sigma_i^-\rho\sigma_{i+1}^+ - \left\{\sigma_{i+1}^+\sigma_i^-,\rho\right\}\right), \tag{33}$$

where V is the strength of the nearest neighbor interaction, g is the amplitude of the magnetic field along the x-axis and γ is the dissipation magnitude, set to the same values for all terms in the sum. We also assume periodic boundary conditions such that $\sigma_L^a = \sigma_0^a$ (where

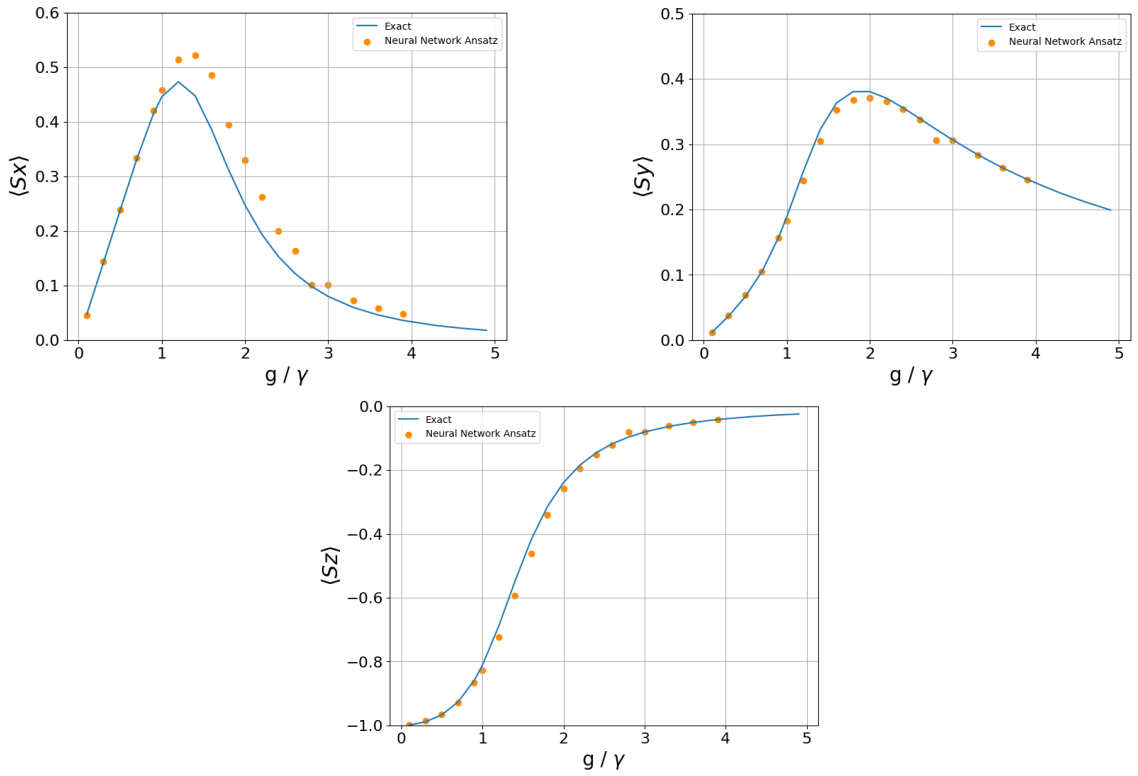


Figure 2: Transverse Ising model with local dissipation: Mean magnetization as a function of g/γ and for fixed $V/\gamma = 2$. The neural network ansatz is represented by the orange dots while the exact solution as the blue line, which was calculated by diagonalization of the Liouvillian superoperator. The number of particles used was $L = 5$. The learning rate of stochastic gradient descent was 0.01. The minimization was done for 5000 steps and the number of samples used at each iteration step was 4000. The fidelity between exactly calculated density matrix and the variational state was ~ 0.997 and the cost function was of the order 10^{-3}. The network parameters were set to $\alpha = 1, \beta = 1$ for $g/\gamma < 2$ and $\alpha = 1, \beta = 2$ for $g/\gamma \geq 2$.

$a \in \{x, y, z\}$). The results are presented in Fig.3. What we observe is that the variational ansatz does a good job finding the steady state and the corresponding observables. For both magnetizations in y- and z-directions the predictions match perfectly the exact diagonalization calculation. Unfortunately, for the magnetization in x-direction we see again that for values $g/\gamma \geq 1.5$ there is a little discrepancy between exact calculation and simulation, yet this time of the smaller value. Nevertheless, the overall behaviour obtained by the neural network closely approaches that of the exact diagonalization solution. Moreover, we have checked that the fidelity between the variational and the exact density matrices is larger than 0.992. This shows that the neural network provides a good representation of the state even when the dissipation is non-local.

5.3 Hyperparameter tuning

In theory the representation power of our ansatz can be improved by increasing the initial hyperparameters. Two most important ones here are α and β, which are the density of the hidden layer, N_h/N, and the density of the ancillary layer, N_a/N, respectively. The possible offset here is the computational time. It is then worth checking if there exists an optimal set of parameters, that gives us the most accurate results while not consuming

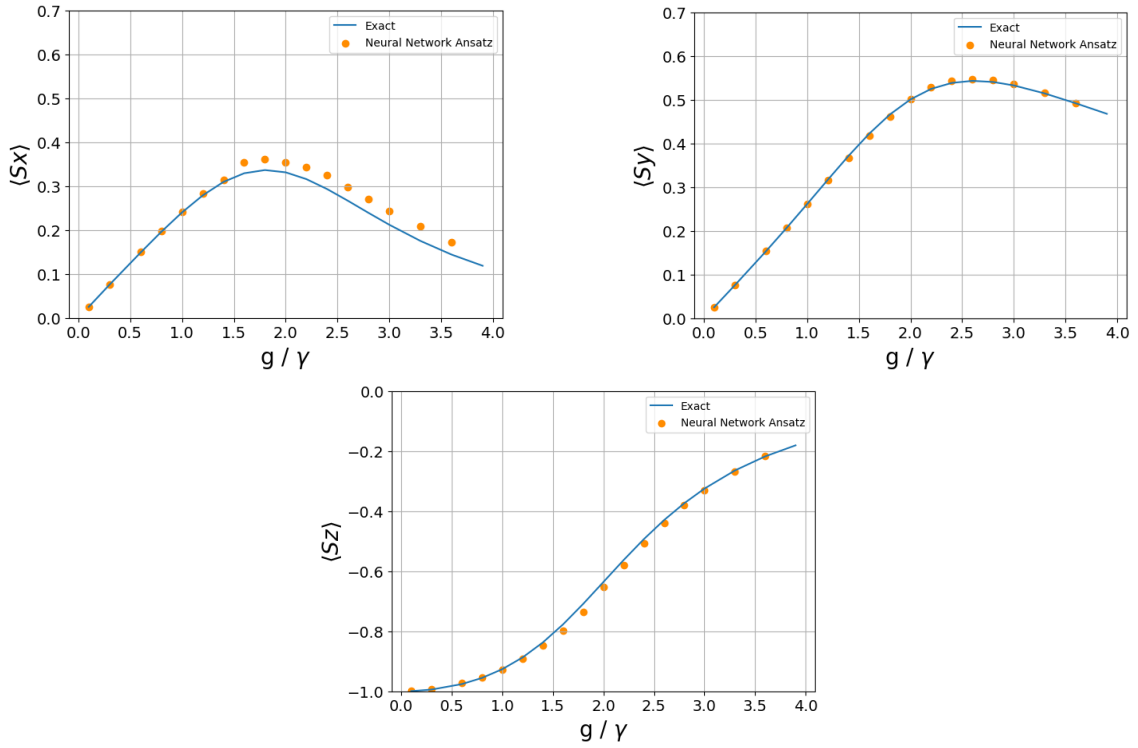
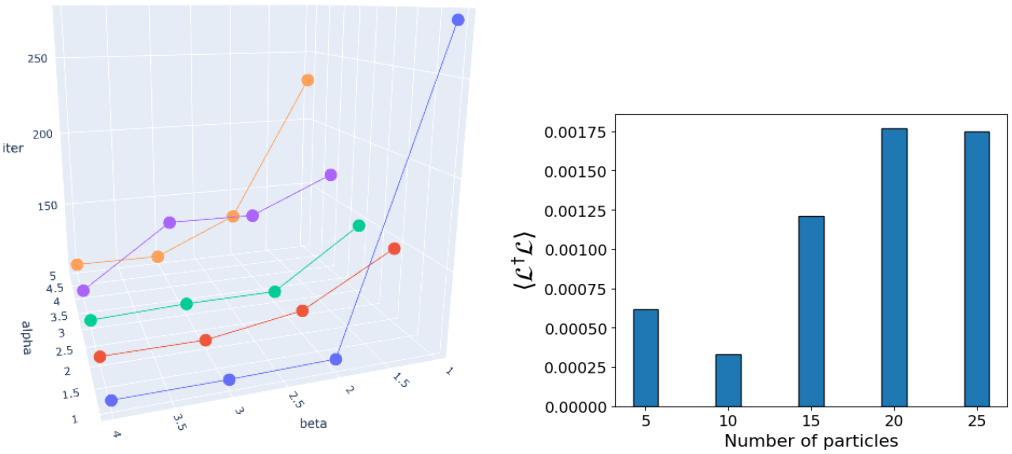


Figure 3: Transverse Ising model with **non-local** dissipation (set as two-body nearest-neighbours dissipators): Mean magnetization as a function of g/γ and for fixed $V/\gamma = 2$. The neural network ansatz is represented by the orange dots while the exact solution as the blue line, which was calculated by diagonalization of the Liouvillian superoperator. The number of particles used was $L = 5$. The learning rate of stochastic gradient descent was 0.01. The minimization was done for 5000 steps and the number of samples used at each iteration step was 4000. The fidelity between exactly calculated density matrix and the variational state was ~ 0.992 and the cost function was of the order 10^{-3}. The network parameters were set to $\alpha = 1, \beta = 1$ for $g/\gamma < 2$ and $\alpha = 1, \beta = 2$ for $g/\gamma \geq 2$.

too much time. We performed such analysis for the 1D Ising model, and the results are presented in Fig.4. In the Fig.4a we have the analysis of the speed of convergence for different values of α and β. The limit value for $\langle\langle \mathcal{L}^\dagger \mathcal{L}\rangle\rangle$ was set to 0.01 and we were calculating the minimum number of iterations needed to go past this value, in order to establish possibly the best and most efficient set of parameters. What can be observed is that the most promising set of parameters is $\alpha = 1, \beta = 2$. It shows highest accuracy-computational time reward compared to other configurations. It is worth noting that while increasing the parameters the cost function continues to converge in less number of iterations, which in theory is an improvement, however it is of such insignificant order (1 or 2 iterations difference), that it does not compensate for longer computational time, which is significantly larger. In some cases we can also observe a situation that increasing the parameters might worsen the convergence, as it is the case or $\alpha = 4, \beta = 2$ and $\alpha = 4, \beta = 3$.

Having found the optimal set of hyperparameters, we then wanted to see how the convergence depends on the system size. This was also done for 1D Ising model in non-local dissipation. We decided to set the target number of iterations to 1000 and save the final values for each of the configurations. The results are presented in Fig.4b. As expected, we can see that with larger size of the system the neural network takes longer to converge. However the final value of the cost function for all simulations is of the same order, the

largest one being for $N = 25$ with cost function value equal to ~ 0.0017.



(a) Convergence speed for different α and β configurations.

(b) Steady state convergence for different system sizes.

Figure 4: Analysis of the convergence accuracy for network parameters for the 1D Ising model with non-local dissipation ($g/\gamma = 0.7$, $V/\gamma = 2$). (a) Number of iterations as a function of network hyperparameters $\alpha = N_h/N$ and $\beta = N_a/N$ to achieve the cost function value 0.01, for the system size $L = 5$. The best set of parameters in terms of accuracy-computational time seems to be $\alpha = 1$ and $\beta = 2$. (b) Cost function for different system sizes after 1000 iterations. In all cases we achieve a cost function value below 0.002.

5.4 Chain of two-level atoms with light-induced dipole-dipole interactions

The main model of interest is a chain of neutral two-level atoms at fixed positions, interacting with light. Due to the coupling with common electromagnetic modes, the atoms experience dipole-dipole interactions and collective scattering of photons, or equivalently, non-local dissipation. As we have seen before, neural network does descent job when used to represent spin-$\frac{1}{2}$ chains. We can understand the previous part of the work as tuning the neural network parameters (number of neurons of the hidden and ancillary layer, and learning rate) on a simpler model. We will now apply these values to our main model of interest, which consists of two-level atoms (with ground and excited states) in an environment of photonic modes. The Hamiltonian of our system will be

$$\begin{aligned} H &= H_{dd} + \Omega \sum_i \sigma_i^x, \\ H_{dd} &= \sum_{i \neq j} \frac{J_{ij}}{2} (\sigma_i^+ \sigma_j^- + \sigma_i^- \sigma_j^+), \end{aligned} \tag{34}$$

where J_{ij} is the coherent interaction strength between the particles, H_{dd} is the dipole-dipole Hamiltonian part and $\Omega \sum_i \sigma_i^x$ is the driving field acting as a longitudinal magnetic field, with Ω being the Rabi frequency. Here the Lindblad dissipation part will take the form:

$$\mathcal{D}[\rho] = \sum_{i,j} \frac{\Gamma_{ij}}{2} \left(2\sigma_j^- \rho \sigma_i^+ - \{\sigma_i^+ \sigma_j^-, \rho\} \right), \tag{35}$$

where σ^+, σ^- are the atomic creation and annihilation operators, and Γ_{ij} is the correlated dissipation. Here both J_{ij} and Γ_{ij} are dependent on the distance between the atoms, and

can be calculated from the Green´s tensor. This tensor inherently captures how electromagnetic fields propagate through a medium, including phase shifts, attenuation, and the influence of boundaries and interfaces. It is essential for calculating coherent interaction and correlated dissipation because it encapsulates the full electromagnetic response of the medium between the emitters. The real part of the Green's tensor determines the coherent exchange of energy (coherent interaction), while the imaginary part accounts for the dissipation and loss mechanisms (correlated dissipation). In this case its form will be:

$$\mathcal{G} = -\frac{3}{4}\Gamma_0 \frac{e^{ikz_{ij}}}{(kz_{ij}^3)}(k^2 z_{ij}^2 + ikz_{ij} - 1) \tag{36}$$

where z_{ij} is the relative distance between the atom's positions $|z_i - z_j|$ and $k = \frac{2\pi}{\lambda}$ is the wavevector, Γ_0 is the single atom spontaneous emission rate, J_{ij} and Γ_{ij} are the real and imaginary parts of the Green's tensor, respectively

$$\begin{aligned} J_{ij} &= -\mathrm{Re}[\mathcal{G}_{ij}] \\ \Gamma_{ij} &= 2\mathrm{Im}[\mathcal{G}_{ij}]. \end{aligned} \tag{37}$$

Additionally, we do not include periodic boundary conditions, as this would require to model the spacing between atoms as two variable function rather than one variable. The system is visualized in Fig.5.

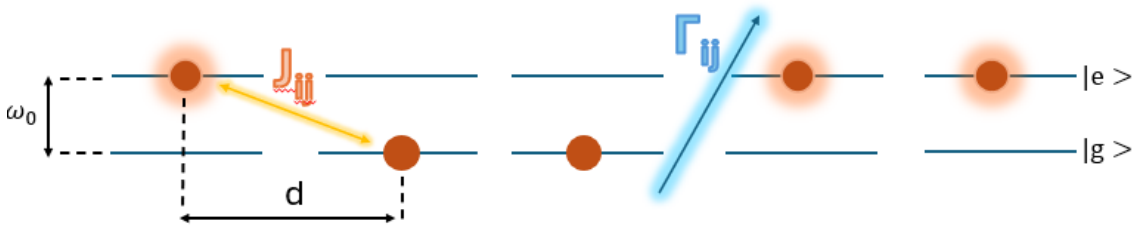


Figure 5: Representation of the two-level neutral atomic chain, for 5 atoms, with electromagnetic field acting as en environment. The atoms interact will all the others via coherent couplings (J_{ij}), representing the exchange of an excitation between them, and dissipative couplings (Γ_{ij}) leading to collective spontaneous emission of photons. Both long and short range interactions can be achieved depending on the spacing between them.

We will now describe what are the observables that we want to measure and the expected results. We start by exploring the excitation density in the stationary state, $n_s = \sum_i < \sigma_i^+ \sigma_i^- > /N$ as a function of Ω/Γ_0, that is, the Rabi frequency in units of the single atom decay rate. When $\Omega \ll \Gamma_0$ the dissipation is dominating the dynamics, so in the steady state there will be few excited atoms, $n_s \approx 0$. The opposite regime corresponds to $\Omega \gg \Gamma_0$. Here, as shown in [37], the steady state is a completely mixed state, thus the value of the excitation density should be $n_s \approx 1/2$. The results for this observable are presented in Fig.6. We can observe that the ansatz gives sensible results. For $N = 8$ particles the neural network shows some discrepancy for values $0.5 < \Omega/\Gamma_0 < 1.5$, however the correct behaviour is kept and the values are not too far off. The similar problems were observed for the magnetization in x-direction for the spin chains, which might suggest that the source of the error could lie not in the simulation itself, but rather in the computational limits of our ansatz. Apart from that we see that the simulation correctly converges to values $n_s \approx 1/2$ when $\Omega \gg \Gamma_0$. For N = 15 atoms the exact diagonalization cannot be performed due to the size of the Liouvillian superoperator. This demonstrates the advantage of potential usage of neural networks, as it is able to perform those calculations. We

can verify that the results are sensible, because for the exact calculation we see that the system converges slower to the $n_s \approx 1/2$ with increasing Ω/Γ_0. This is correct with the theoretical assumption, as we would expect that the longer the atomic chain is, a larger number of photons, and thus, a larger value of the driving Rabi frequency Ω is needed to reach the saturation.

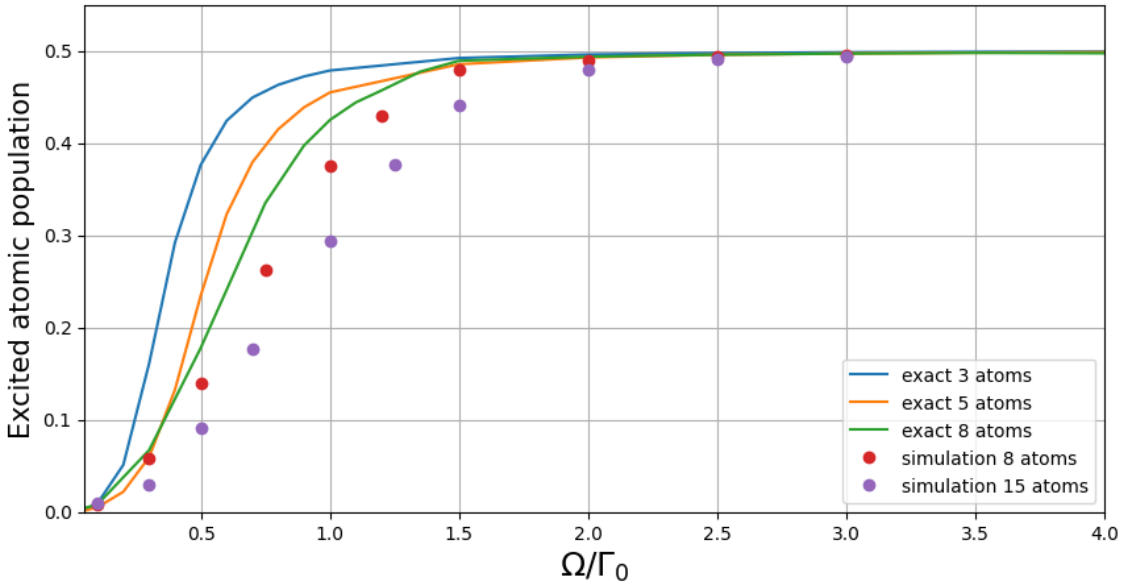


Figure 6: Chain of two-level atoms: Excited state population (per site) as a function of the Rabi frequency Ω in units of single atom decay rate Γ_0, for different chain lengths. The cases $N = 3, 5, 8$ (solid lines) are calculated using exact diagonalization of the Liouvillian. The cases $N = 8, 15$ (circles) are obtained with the variational neural network ansatz. In this case the cost function is of the order $10^{-2} - 10^{-3}$ for all points. The minimization was done for 5000 steps. For $L = 8$ the number of samples used was 6000, while for $L = 15$ was number of samples was 20000. The network parameters were set to $\alpha = 1, \beta = 2$.The learning rate of stochastic gradient descent was 0.01. The fidelity for $L = 8$ was ~ 0.97.

Another interesting quantity of this system to look into is the average correlated photon emission, which arises as a consequence of the non-local character of the dissipation operators [38]. The average emission rate, k_s, can be calculated as a sum over all possible decay channels and when the system is in the steady state ρ_{ss}, its value corresponds to:

$$k_s = \sum_{ij} \Gamma_{ij} < \sigma_i^+ \sigma_j^- >= \sum_{ij} \Gamma_{ij} \mathrm{Tr}[\sigma_i^+ \sigma_j^- \rho_{ss}], \tag{38}$$

where Γ_{ij} is the correlated dissipation. From this quantity we need to subtract the independent atom photon emission $N\Gamma_0 n_s$. Thus the average correlated photon emission is given by:

$$k_s - N\Gamma_0 n_s = \sum_{ij} \Gamma_{ij} < \sigma_i^+ \sigma_j^- > -N\Gamma_0 n_s = \sum_{ij} \Gamma_{ij} \mathrm{Tr}[\sigma_i^+ \sigma_j^- \rho_{ss}] - N\Gamma_0 n_s. \tag{39}$$

The expected behavior is as follows. For low values of the driving Rabi frequency we expect a small photon emission rate as the excited state population is very small. On the other hand, for very large values of the field intensity or Rabi frequency, the atoms are saturated and the state corresponds to a product state, and they behave as if they were independent. Thus, in this limit, although the total emission is large, the correlated part

also vanishes. We thus expect a non-trivial maximum in the correlated emission, which strongly depends on the inter-particle distances. As the particles become further apart, the dipole-dipole couplings become smaller, thus leading to a flatter behavior closer to zero.

The results of the correlated emission are presented in Fig.7. We can see that for most of the explored cases the neural network successfully catches the correlated emission behaviour. For $d/\lambda = 0.16$ it shows a slower convergence near the peak, and leads to some discrepancy with respect to the exact calculation results. However for the rest of the points both curves are close to each other. For $d/\lambda = 0.24$, we see the similar problem, near the peak, some of the values are off, yet later they are matching the exact solution. For $d/\lambda = 0.8$, we expect to see the atoms already behaving as independent, due to large spacing between them, so the correlated emission should be zero throughout. Excluding the first point, which shows the discrepancy, the rest of the calculated values seems to catch that behaviour correctly.

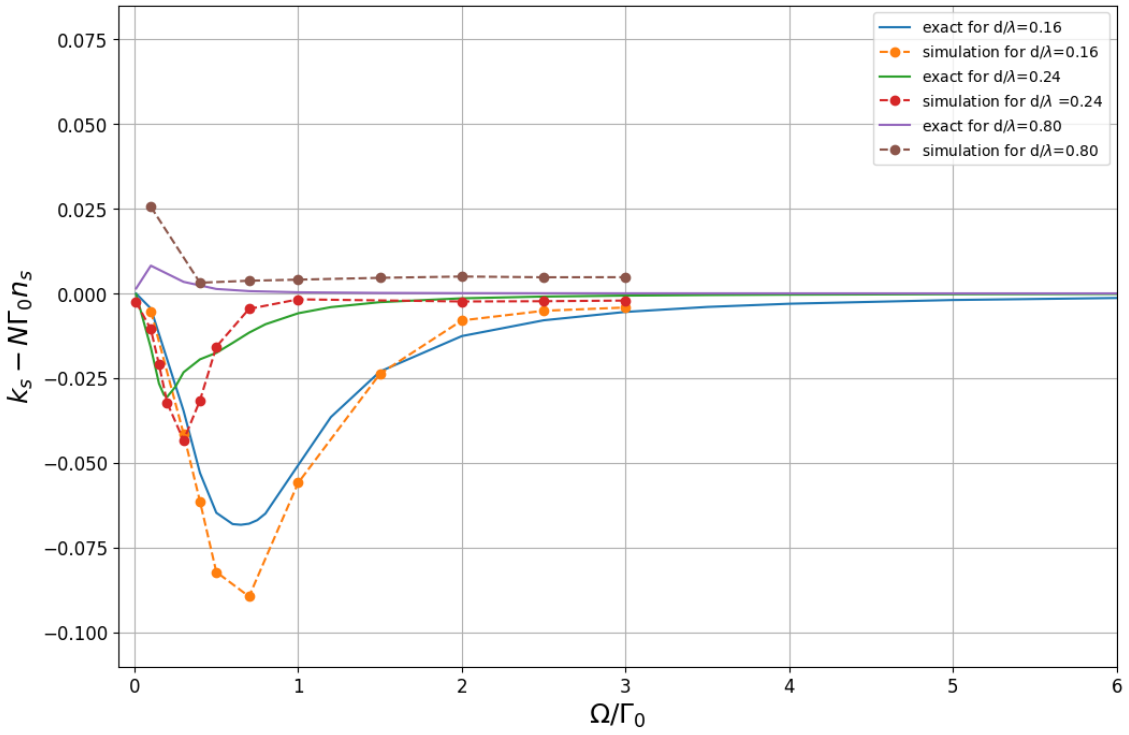


Figure 7: Chain of two-level atoms: Correlated photon emission as a function of Ω/Γ_0, for different values of the lattice constant ($d/\lambda = 0.16, 0.24, 0.8$). Circles denote the calculation with the variational neural network ansatz, while the solid line is the result when exactly diagonalizing the Liouvillian. The number of atoms is $N = 6$.

The source of the error here could be due to many factors. For all the points the fidelity was high, ~ 0.98, which means that both density matrices were close to each other. Here the parameter tuning played an important role, because even if the cost function was low, fidelity happened to be well off. In some cases the solution was to change the learning rate, in others manipulation of the density of neurons in the hidden and ancillary layers. Here a clear pattern was not observed, which means that for each point, the tuning had to be done individually. The discrepancy that we get, in spite of the high fidelity, could mean that this final error is not only the optimization problem. It can then be argued that this error could be reduced by modifying the ansatz. For instance, by adding more hidden or ancillary layers, or by construction of other type of suitable architecture for this task.

6 Conclusion and outlook

In this thesis, we explored the application of neural network ansatz, to the study of open quantum systems. We demonstrated the viability and efficiency of using neural network-based variational methods for modeling the steady states and dynamical properties of such systems, focusing particularly on spin chains and neutral atomic chains interacting with light.

Our investigation confirmed that the neural network ansatz is capable of approximating the steady states of open quantum systems with high fidelity, often achieving values around 0.98. This high fidelity indicates that the density matrices generated by the neural network are very close to those obtained through exact methods. The RBM-based approach successfully captures the essential physics of the models considered, including both local and non-local dissipation scenarios.

Despite these successes, certain challenges were encountered, like parameter tuning and convergence near peak values of correlated photon emission or discrepancy in values when measuring x-direction magnetization. These difficulties suggest that while the current RBM ansatz is effective, there is room for improvement. Adjustments to the ansatz architecture, such as incorporating additional hidden or ancillary layers, could potentially enhance performance and reduce discrepancies.

One of ideas that for now offers a potential solution could be Liouvillie Denisty Machine (LDM), proposed by Simon Kothe and Peter Kirton [39]. It shows a possibility of representing the density matrix, as a NQSs, directly in Liouvillie space. So far, the results suggest that this type of ansatz works better for systems with high number of correlations. If proven successful, this type of encoding could then be extended to deep RBMs [40] or Recurrent Neural Networks [41], which also have shown accurate results for closed systems.

In summary, the neural network ansatz represents a promising method for addressing the challenges posed by open quantum systems. By leveraging the strengths of machine learning, this approach opens new pathways for research and application in quantum mechanics, offering a scalable and flexible framework for exploring the dynamics and properties of open quantum systems.

Bibliography

[1] S. Lapuschkin, S. Wäldchen, A. Binder, G. Montavon, W. Samek, and K.-R. Müller, Nature Communications **10**, 1096 (2019).

[2] J. Frankle and M. Carbin, "The lottery ticket hypothesis: Finding sparse, trainable neural networks," (2019), arXiv:1803.03635 .

[3] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T. S. Chua, "Neural collaborative filtering," (2017), arXiv:1708.05031 .

[4] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, in *2014 IEEE Conference on Computer Vision and Pattern Recognition* (2014) pp. 1701–1708.

[5] P. A. M. Dirac, Proc. R. Soc. Lond. A **123**, 714 (1929).

[6] H. Lange, A. Van de Walle, A. Abedinnia, and A. Bohrdt, "From architectures to applications: A review of neural quantum states," (2024), arXiv:2402.09402 .

[7] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[8] G. Torlai and R. G. Melko, Phys. Rev. Lett. **120**, 240503 (2018).

[9] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, Phys. Rev. Lett. **122**, 250503 (2019).

[10] A. Nagy and V. Savona, Phys. Rev. Lett. **122**, 250501 (2019).

[11] M. J. Hartmann and G. Carleo, Phys. Rev. Lett. **122**, 250502 (2019).

[12] H. P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, 2006).

[13] W. H. Zurek, Rev. Mod. Phys. **75**, 715 (2003).

[14] J. Goold, M. Huber, A. Riera, L. d. Rio, and P. Skrzypczyk, J. Phys. A: Math. Theor. **49**, 143001 (2016).

[15] P. W. Shor, Phys. Rev. A **52**, R2493 (1995).

[16] B. J. Berne and J. E. Straub, Curr. Opin. Struct. Biol. **7**, 181 (1997).

[17] S. Chu, Science **242**, 365 (1988).

[18] H. P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, 2002).

[19] C. Gardiner and P. Zoller, *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*, Springer Series in Synergetics (Springer, 2004).

[20] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, J. Math. Phys. **17**, 821 (1976).

[21] G. Schaller, *Open Quantum Systems Far from Equilibrium*, Lecture Notes in Physics, Vol. 881 (Springer, 2014).

[22] N. Yoshioka and R. Hamazaki, Phys. Rev. B **99**, 214306 (2019).

[23] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, Phys. Rev. A **57**, 3049 (1998).

[24] D. Nigro, J. Stat. Mech. **2019**, 043202 (2019).

[25] V. V. Albert and L. Jiang, Phys. Rev. A **89**, 022118 (2014).

[26] P. Meystre and M. Sargent, *Elements of Quantum Optics*, SpringerLink: Springer e-Books (Springer Berlin Heidelberg, 2007).

[27] A. D. Ludlow, M. M. Boyd, J. Ye, E. Peik, and P. O. Schmidt, Rev. Mod. Phys. **87**, 637 (2015).

[28] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).

[29] A. Dawid, J. Arnold, B. Requena, A. Gresch, M. Płodzień, K. Donatella, K. A. Nicoli, P. Stornati, R. Koch, M. Büttner, R. Okuła, G. Muñoz-Gil, R. A. Vargas-Hernández, A. Cervera-Lierta, J. Carrasquilla, V. Dunjko, M. Gabrié, P. Huembeli, E. van Nieuwenburg, F. Vicentini, L. Wang, S. J. Wetzel, G. Carleo, E. Greplová,

R. Krems, F. Marquardt, M. Tomza, M. Lewenstein, and A. Dauphin, "Modern applications of machine learning in quantum sciences," (2023), arXiv:2204.04198 [quant-ph] .
[30] M. Van den Nest, "Simulating quantum computers with probabilistic methods," (2010), arXiv:0911.1624 [quant-ph] .
[31] G. E. Hinton, Trends Cogn. Sci. **11**, 428 (2007).
[32] H. Robbins and M. Monro, Ann. Math. Statist. **22**, 400 (1951).
[33] S. Sorella and L. Capriotti, Phys. Rev. B **61**, 2599 (2000).
[34] W. K. Hastings, Biometrika **57**, 97 (1970).
[35] S. Brooks, J. R. Stat. Soc. Ser. D **47**, 69 (1998).
[36] F. Vicentini, D. Hofmann, A. Szabó, D. Wu, C. Roth, C. Giuliani, G. Pescia, J. Nys, V. Vargas-Calderón, N. Astrakhantsev, and G. Carleo, SciPost Physics Codebases (2022), 10.21468/scipostphyscodeb.7.
[37] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller, Phys. Rev. A **78**, 042307 (2008).
[38] B. Olmos, D. Yu, and I. Lesanovsky, Phys. Rev. A **89**, 023616 (2014).
[39] S. Kothe and P. Kirton, Phys. Rev. A. **109** (2024).
[40] H. Hu, L. Gao, and Q. Ma, "Deep restricted boltzmann networks," (2016), arXiv:1611.07917 .
[41] A. Sherstinsky, Physica D **404**, 132306 (2020).