

**MACHINE LEARNING PHASES OF MATTER IN
OPEN QUANTUM SYSTEMS**

ZEJIAN LI
(B.Sc.(Hons.), NUS)

**A THESIS SUBMITTED FOR THE DEGREE OF
MASTER OF SCIENCE
DEPARTMENT OF PHYSICS
NATIONAL UNIVERSITY OF SINGAPORE**

2019

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.
This thesis has also not been submitted for any degree in any university previously.



21 August 2019

Acknowledgements

I would like to sincerely thank my supervisor Professor Dimitris Angelakis, who accepted me to his group and guided me throughout the entire project, for the priceless discussions and ideas, which allowed me to overcome the difficulties, for the precious materials he provided and the talks, seminars and group meetings he organised, which deepened my understandings in quantum many-body physics, without which this work would not have been possible.

I would also like to thank all my group-mates for their help and collaboration, with special thanks to Jirawat Tangpanitanon, who helped me solve many problems I encountered, provided guidelines for my progress and helped me correct this thesis; to Supanut Thanasilp and Marc-Antoine Lemonde, for the daily discussions and exchanges of thoughts, and to Wing Chi Yu who helped me on the topic of machine learning.

I thank Professor Cristiano Ciuti, who is one of the authors of relevant works that this thesis is based on, for inviting me to his research group for discussion, without which the progress of this work would have been much more difficult.

Finally, I would like to thank Centre for Quantum Technologies, National University of Singapore and École Polytechnique, who made this project possible.

Contents

Introduction	1
1 Phase transitions in closed and open quantum systems	4
1.1 Closed quantum systems	4
1.1.1 Example: ferromagnetic transverse field Ising model .	5
1.2 Open quantum systems	7
1.2.1 Lindblad master equation	8
1.2.2 Dissipative phase transition	10
1.2.3 Review of recent works on dissipative phase transitions	13
1.2.4 Numerical techniques for probing dissipative phases of matter	15
1.3 Conclusion	16
2 Machine learning quantum phases of matter	17
2.1 Machine learning basics	17
2.2 Recent works applying machine learning to open quantum systems	18
2.3 Artificial neural networks for open quantum systems	19
2.4 Variational Monte Carlo for finding the NESS	24
2.4.1 Variational principle for open quantum systems . . .	25
2.4.2 Evaluating the cost function and its gradient	26
2.4.3 Stochastic sampling	28
2.4.4 Natural gradient adaptation	29
2.4.5 Summary of the VMC training protocol	30
2.5 Conclusion	30
3 Artificial neural network results on driven-dissipative lat- tice models	31
3.1 Dissipative transverse-field Ising model	31
3.2 Driven-dissipative XY model	33
3.3 Driven-dissipative Bose-Hubbard model	37
3.4 Discussion	38
3.4.1 Spectral gap of $\mathcal{L}^\dagger \mathcal{L}$	39

3.5 Conclusion	39
Conclusion and outlook	40
A Log-derivatives of the RBM density matrix ansatz	45

Summary

One of the main challenges in strongly correlated systems is the study of dissipative phases of matter, which, as a quantum many-body problem, is hard in general. This thesis studies a new emerging approach to this problem, which is the application of artificial neural networks combined with machine learning techniques. The latter can be used in finding the many body steady state of an open quantum system with relative high accuracy in one and also two dimensions, where other approximate methods such as tensor networks fail.

This thesis begins with a general introduction to the field, the notion of phases of matter and with a review of existing works on dissipative phases of matter (Chapter 1 and 2).

We then present the method we implement for this study (Chapter 2). The idea consists of first “purifying” the open system by adding ancillary degrees of freedom, whose total wave function can be represented by an artificial neural network (ANN) ansatz called the restricted Boltzmann machine (RBM). The latter is known to be able to well approximate many-body quantum states. By tracing out these ancillary degrees of freedom, we obtain an ansatz for the density matrix, that guarantees to have all the mathematical properties of a valid density matrix. This ANN ansatz can be trained to learn the steady state of the system, whose dynamics are generated by the Lindbladian \mathcal{L} , via a variational Monte Carlo scheme. We also propose a new RBM structure that respects the symmetry of the physical system and implement it in the case of translation invariance.

Finally (Chapter 3), we apply our method to the 1D and 2D dissipative transverse field Ising model, the 1D driven-dissipative XY model and the 1D driven-dissipative Bose-Hubbard model, obtaining overall satisfying results, which are also novel by the completion of this work, to the best of our knowledge. This study successfully extends the ANN approach to new physical models, showing great potential in this field of research.

List of Figures

1.1	Two types of non-analyticity in a quantum phase transition	5
1.2	Schematic diagram of a driven-dissipative QED cavity array	12
1.3	Bunching-antibunching transition in the XY model	14
1.4	A 72-site circuit QED lattice	16
2.1	Graphical representation of the RBM wave function ansatz .	21
3.1	Convergence of the cost function and the observables in the TFI model	32
3.2	1D TFI NESS observables, cost function and spectral gap of $\mathcal{L}^\dagger \mathcal{L}$	34
3.3	2D TFI NESS observables and the cost function	35
3.4	1D XY NESS observables and the cost function	36
3.5	Comparison between the cost function values of the standard and the symmetrised RBMs	37
3.6	1D BH NESS observable and the cost function	38

List of Tables

1.1	Comparison between different types of phase transitions . . .	11
2.1	Summary of recent works applying ANN to open quantum systems	20

Introduction

One of the fascinating aspects of our universe is that given its immensity and the number of particles it contains, our world is not (yet) in a complete disorder. Indeed, any material at a real-life scale consists of a large number of particles, and should be considered as a *many-body system*. For example, a few drops of water can contain as many as 10^{23} electrons. The particles in a many-body system can interact and collaborate with each other, forming different *phases* of matter. A phase is a regime of matter having essentially uniform physical properties throughout the material that is distinct from other phases. Under the change of external conditions, the material can undergo an abrupt transition from one phase to another, which is referred to as a *phase transition*.

In classical physics, thermal phase transitions happen at finite temperatures and are driven by thermal fluctuations. In closed quantum systems, phase transitions happen at absolute zero temperature, where the system is in its ground state, and are due to quantum fluctuations, which are thus referred to as quantum phase transitions. However, realistically speaking, all quantum systems (unless we talk about the entire universe as the system) are open systems, as they are constantly interacting with the environment. For example, in photonic systems, photons are continuously dissipating out of the system and must be pumped into the system from an external source. In this case, the system will reach a non-equilibrium steady-state, where the pump and loss compensate each other. Phase transitions in an open quantum system will then happen on the steady-states instead, and are therefore referred to as dissipative phase transitions.

Identifying phases of matter has been at the very heart of material science, quantum chemistry and condensed matter physics [1], not only because of the richness in physics but also due to the potential of valuable applications. One spectacular manifestation of quantum phase transition is superconductivity. In certain materials, when the temperature drops below a critical value, electrons pair up and condense into a state of “quantum fluid”, allowing electric current to pass through without any dissipation.

In general, studying quantum phases of matter and the transitions between them involves solving the Schrödinger equation, a linear partial dif-

ferential equation that describes the dynamics of the state vector $|\Psi\rangle$, that lives in a complex vector space called the Hilbert space and contains all the information one can know about the system. In the open system regime, as we will derive in Chapter 1, the state is described by a density matrix $\hat{\rho}$ whose dynamics are governed by the master equation under certain approximations. However, as the dimension of the Hilbert space grows exponentially with the size of the system, it is extremely hard to exactly solve a many-body system at any real-life scale. In fact, the exact computer simulation of a system containing only 50 spin-1/2 particles, such as electrons, would require several petabytes of RAM assuming each complex coefficient of the vector $|\Psi\rangle$ takes 1 byte to store, which is on the order of magnitude of the memory capacity of best supercomputers today [2]. Therefore, the exact solution for a quantum many-body problem is usually not feasible given today's computer technology.

To tackle this difficulty, many sophisticated numerical techniques have been developed aiming to approximately represent states of an open many-body system with much less information, including tensor networks, Monte Carlo methods and others [3, 4, 5]. In recent years, artificial neural networks, which are inspired from the immense network of neurons in our own brains, have been proposed to solve quantum many-body problems and ever since, neural network quantum states combined with machine learning techniques are emerging as powerful tools for solving for ground states in closed quantum systems as well as their dynamics in the Hilbert space [6]. However, the generalisation of neural network states to many-body problems in open quantum systems is still in its very infancy. This thesis will be focusing on the application of artificial neural networks in finding the steady states of various open quantum systems, where non-trivial phases of matter may emerge.

The first chapter of this manuscript is aimed at introducing the notion of phase transitions in both closed quantum systems and open quantum systems. For the latter, we will derive the Lindblad master equation, and review some of the works in the domain of dissipative phase transition.

In Chapter 2, we will introduce the machine-learning algorithms for open quantum many-body systems. The problem of finding the steady state of an open quantum system will be formulated with a variational approach, and we will use a variational ansatz based on an artificial neural network called the restricted Boltzmann machine to represent the density matrix of the system. In particular, we implement a new neural network structure specially designed for translational-invariant lattice models. The neural network ansatz will then be trained with a reinforcement learning scheme, giving the solution for the steady state of the system.

In Chapter 3, we present our novel results on various driven-dissipative lattice models, including the Bose-Hubbard and Ising spin models, which demonstrate the universality and strength of the neural network approach.

Chapter 1

Phase transitions in closed and open quantum systems

In this chapter, we will briefly review the theory of phase transitions of matter in a quantum-mechanical context, for both closed systems and open systems.

We will begin with an introduction to the notion of quantum phase transitions in closed systems, with the example of the quantum Ising model. We then extend our discussion to the driven-dissipative regime in open systems, where the dynamics are governed by the master equation and that the steady state can exhibit dissipative phase transitions.

1.1 Closed quantum systems

Unlike classical phase transitions driven by thermal fluctuations at a finite temperature, quantum phase transitions happen at absolute zero temperature. In a closed system, this corresponds to the ground state of the system, that can show abrupt changes under the change of external parameters. To make this idea concrete, we can consider a Hamiltonian for some lattice model of the form

$$\hat{H}(g) = \hat{H}_1 + g\hat{H}_2, \quad (1.1)$$

where g is a dimensionless parameter. Denoting the ground state of the Hamiltonian as $|G\rangle$, a quantum phase transition happens when the dependence of the ground state energy on the parameter g ,

$$E(g) = \langle G | \hat{H}(g) | G \rangle, \quad (1.2)$$

is non-analytic, as illustrated in Figure 1.1. For a finite lattice, $E(g)$ is usually smooth and analytic except when \hat{H}_1 is a conserved quantity of the system, i.e. $[\hat{H}_1, \hat{H}_2] = 0$, creating a level crossing for certain value of g

and thus a non-analyticity. Another possible source of non-analyticity is when the lattice is infinite, i.e. in the thermodynamic limit, where the gap between the first excited state and the ground state may close. The latter has richer possibilities and is more closely related to the classical phase transition.

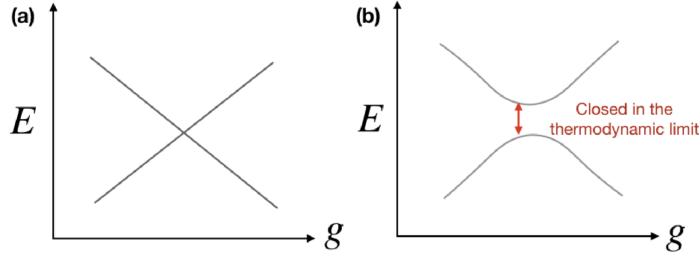


Figure 1.1: Two types of non-analyticity in the ground state energy leading to a quantum phase transition. Spectrum (only the two lowest energies are shown) of the Hamiltonian $\hat{H}(g) = \hat{H}_1 + g\hat{H}_2$ for (a) $[\hat{H}_1, \hat{H}_2] = 0$ and (b) $[\hat{H}_1, \hat{H}_2] \neq 0$.

1.1.1 Example: ferromagnetic transverse field Ising model

The ferromagnetic transverse field Ising (TFI) model is a model for quantum magnetism and is one of the simplest interacting models to show a phase transition. It describes a lattice of two level systems, such as spin-1/2 particles, with a fixed external magnetic field along the x direction and with nearest-neighbour coupling via $z - z$ interaction, defined by the following Hamiltonian:

$$\hat{H}_{\text{ferro}}^{\text{TFI}} = -J \sum_{\langle i,j \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z - gJ \sum_i \hat{\sigma}_i^x, \quad (1.3)$$

where $\langle i, j \rangle$ denotes nearest neighbour pairs of lattice indices, J is the interaction strength and sets the energy scale, g is the dimensionless parameter characterising the strength of the transverse magnetic field and $\hat{\sigma}_i^z$ and $\hat{\sigma}_i^x$ are the Pauli matrices acting on site i . We stay in the regime $J > 0$, where the interaction is called ferromagnetic.

The phases of the TFI model can be understood qualitatively by investigating the limiting cases for the parameter g .

Weak magnetic field limit ($g \rightarrow 0$)

In this limit, the $z - z$ interaction dominates and all spins will be aligned in the same direction, giving the ground state

$$|G\rangle = \prod_i |\uparrow\rangle_{z,i} \text{ or } |G\rangle = \prod_i |\downarrow\rangle_{z,i}, \quad (1.4)$$

which is degenerate. A thermodynamic system will pick one of the two possible ground states, commonly referred to as “spontaneous symmetry breaking”. In this phase, the system is magnetically ordered and has a non-zero total magnetisation along z direction.

Strong magnetic field limit ($g \rightarrow \infty$)

In this limit, the external field dominates and forces all the spins to align with it. As a result, the ground state is

$$|G\rangle = \prod_i |\uparrow\rangle_{x,i}, \quad (1.5)$$

and the z component of the spins are totally not correlated. In this phase, the system is paramagnetic and has zero total magnetisation along z .

As the ground states of these two cases are qualitatively different, we can expect a quantum phase transition to happen at some critical value $g = g_c$. In the following, we will show that the TFI model can be exactly solved and that the energy levels will cross at this critical value.

Exact solution

In the one-dimensional case, following the treatment in [7], we can show that the Hamiltonian can be first written in terms of spinless fermionic operators by performing a Jordan-Wigner transformation after rotating the system by 90° about the y -axis:

$$\hat{\sigma}_i^z \mapsto \hat{\sigma}_i^x \text{ and } \hat{\sigma}_i^x \mapsto -\hat{\sigma}_i^z, \quad (1.6)$$

$$\hat{H}_{\text{ferro}}^{\text{TFI}} = -J \sum_i (\hat{c}_i^\dagger \hat{c}_{i+1} + \hat{c}_{i+1}^\dagger \hat{c}_i + \hat{c}_{i+1}^\dagger \hat{c}_i^\dagger + \hat{c}_{i+1} \hat{c}_i - 2g\hat{c}_i^\dagger \hat{c}_i - g), \quad (1.7)$$

where

$$\hat{c}_i = \left(\prod_{j < i} \hat{\sigma}_j^z \right) \hat{\sigma}_i^+, \quad \hat{c}_i^\dagger = \left(\prod_{j < i} \hat{\sigma}_j^z \right) \hat{\sigma}_i^-, \quad (1.8)$$

and

$$\hat{\sigma}_i^\pm = \frac{1}{2}(\hat{\sigma}_i^x \pm i\hat{\sigma}_i^y) \quad (1.9)$$

are the Pauli transition matrices. Then, the Hamiltonian can be diagonalised by the Fourier transform

$$\hat{c}_k = \frac{1}{\sqrt{N}} \sum_j \hat{c}_j e^{ikja}, \quad (1.10)$$

where

$$k = \frac{2\pi}{Na} m, \quad m \in \{0, 1, \dots, N-1\}, \quad (1.11)$$

N is the number of sites and a is the inter-site spacing. The Hamiltonian becomes

$$\hat{H}_{\text{ferro}}^{\text{TFI}} = J \sum_k [2(g - \cos(ka) \hat{c}_k^\dagger \hat{c}_k - i \sin(ka))(\hat{c}_{-k}^\dagger \hat{c}_k^\dagger + \hat{c}_{-k} \hat{c}_k) - g]. \quad (1.12)$$

Finally, with a Bogoliubov transformation defining the new operators:

$$\hat{\gamma}_k = \cos \frac{\theta_k}{2} \hat{c}_k - i \sin \frac{\theta_k}{2} \hat{c}_{-k}^\dagger, \quad \text{with } \tan \theta_k = \frac{\sin(ka)}{\cos(ka) - g}, \quad (1.13)$$

the diagonalised Hamiltonian becomes

$$\hat{H}_{\text{ferro}}^{\text{TFI}} = \sum_k \epsilon_k (\hat{\gamma}_k^\dagger \hat{\gamma}_k - \frac{1}{2}), \quad (1.14)$$

where

$$\epsilon_k = 2J \sqrt{2 + g^2 - 2g \cos k} \quad (1.15)$$

gives the energy gap between the ground state and the first excited state for momentum number k . The smallest gap is at $k = 0$, which vanishes at $g = g_c = 1$. This level crossing signifies the transition between the ordered phase and the paramagnetic phase.

1.2 Open quantum systems

In real-life experiments and applications, a finite coupling between the quantum system and the environment cannot be avoided. In the presence of external drive and dissipation, the system will evolve towards the non-equilibrium steady state (NESS), which is, in general, dramatically different from the ground state. As we will now derive, assuming that the environment is memoryless and weakly coupled to the system, the dynamics of an open quantum system can be described by the Lindblad master equation.

1.2.1 Lindblad master equation

To start our derivation, let us consider the time evolution of an open quantum system as a part of a larger closed system formed by the system and the environment, that undergoes a unitary time evolution generated by the total Hamiltonian. We assume that the system only starts interacting with the environment from time $t = 0$, which means that the initial state of the system and the environment can be written as a separable density matrix

$$\hat{\rho}(0) = \hat{\rho}_S(0) \otimes |\psi\rangle_E \langle\psi|_E, \quad (1.16)$$

where $\hat{\rho}_S$ is the density matrix for the system and $|\psi\rangle_E$ is the initial state of the environment, assumed to be a pure state. Denoting the time evolution (unitary) operator of the system and the environment from time 0 to t with \hat{U}_{SE} , we can write the evolution as

$$\hat{\rho}(t) = \hat{U}_{SE}(\hat{\rho}_S(0) \otimes |\psi\rangle_E \langle\psi|_E) \hat{U}_{SE}^\dagger. \quad (1.17)$$

This is equivalent to solving the Schrödinger equation for the system and the environment. However, it is usually unfeasible due to the fact that the environment may contain infinitely many degrees of freedom. Instead, if we are only interested in the evolution of the system, we can focus on the density matrix describing the system by performing a partial trace on the environment degrees of freedom:

$$\hat{\rho}_S(t) = \text{Tr}_E[\hat{\rho}(t)] \quad (1.18)$$

$$= \sum_{k=1}^{N_E} \langle k|_E \hat{U}_{SE}(\hat{\rho}_S(0) \otimes |\psi\rangle_E \langle\psi|_E) \hat{U}_{SE}^\dagger |k\rangle_E \quad (1.19)$$

$$= \sum_{k=1}^{N_E} \langle k|_E \hat{U}_{SE} |\psi\rangle_E \hat{\rho}_S(0) \langle\psi|_E \hat{U}_{SE}^\dagger |k\rangle_E, \quad (1.20)$$

where $\{|k\rangle_E, k = 1, 2, \dots, N_E\}$ denotes an orthonormal basis for the environment. Defining the operators that act on the system's Hilbert space

$$\hat{M}_k = \langle k|_E \hat{U}_{SE} |\psi\rangle_E, \quad (1.21)$$

The evolution for the system's density matrix can now be written as

$$\hat{\rho}_S(t) = \sum_k \hat{M}_k \hat{\rho}_S(0) \hat{M}_k^\dagger. \quad (1.22)$$

Note that the unitarity of the time evolution operator \hat{U}_{SE} leads directly to the “normalisation”:

$$\sum_k \hat{M}_k^\dagger \hat{M}_k = \mathbb{1}. \quad (1.23)$$

In the following, we drop the subscript “ S ” of the density matrix of the system since there is no more ambiguity. In fact, this representation of the evolution of an open quantum system is general, as guaranteed by the *Kraus representation theorem*, which states that any linear and completely positive operator $\mathcal{M} : \hat{\rho} \mapsto \mathcal{M}(\hat{\rho})$ in a space of dimension N_S^2 (N_S being the dimension of the Hilbert space of the system whose density matrix is $\hat{\rho}$) that preserves the trace and Hermiticity of $\hat{\rho}$ can be written in the form:

$$\mathcal{M}(\hat{\rho}) = \sum_{k=1}^K \hat{M}_k \hat{\rho} \hat{M}_k^\dagger, \quad (1.24)$$

where $K < N_S^2$ is the *Kraus number* and the \hat{M}_k ’s are called the *Kraus operators*, with the normalisation

$$\sum_{k=1}^K \hat{M}_k^\dagger \hat{M}_k = \mathbb{1}. \quad (1.25)$$

To be able to describe the evolution of the system with a differential equation, we have to further assume that the evolution is Markovian, which means that the environment has no memory, such that $\hat{\rho}(t + dt)$ only depends on $\hat{\rho}(t)$ and not on the state of the system at earlier times, which is a good approximation if the environmental correlations decay on a time scale that is short enough compared to those of the system. Let us consider an infinitesimal time evolution of $\hat{\rho}(t)$ from Equation (1.24):

$$\hat{\rho}(t + \delta t) = \sum_k \hat{M}_k(\delta t) \hat{\rho}(t) \hat{M}_k^\dagger(\delta t). \quad (1.26)$$

We wish to only keep the first order terms in δt , which implies that the Kraus operators should be expanded as

$$\hat{M}_k = \hat{M}_k^{(0)} + \sqrt{\delta t} \hat{M}_k^{(1)} + \delta t \hat{M}_k^{(2)} + \dots. \quad (1.27)$$

To ensure the form $\hat{\rho}(t + \delta t) = \hat{\rho}(t) + \delta t \dot{\rho}$, we can rewrite the expansion for the Kraus operators in the form

$$\begin{aligned} \hat{M}_0 &= \mathbb{1} + \delta t(-i\hat{H} + \hat{K}) + \mathcal{O}(\delta t^2), \\ \hat{M}_k &= \sqrt{\delta t} \hat{L}_k + \mathcal{O}(\delta t), \quad \text{for } k > 0, \end{aligned} \quad (1.28)$$

Where \hat{H} and \hat{K} are Hermitian. Equation (1.25) requires that

$$\hat{K} = -\frac{1}{2} \sum_{k>0} \hat{L}_k^\dagger \hat{L}_k. \quad (1.29)$$

Substituting equations (1.28) and (1.29) back into Equation (1.26), and taking the limit $\delta t \rightarrow 0$, we arrive at the *Lindblad master equation*:

$$\dot{\hat{\rho}}(t) = \mathcal{L}[\hat{\rho}(t)] = -i[\hat{H}, \hat{\rho}(t)] + \sum_k \left[\hat{L}_k \hat{\rho}(t) \hat{L}_k^\dagger - \frac{1}{2} \hat{L}_k^\dagger \hat{L}_k \hat{\rho}(t) - \frac{1}{2} \hat{\rho}(t) \hat{L}_k^\dagger \hat{L}_k \right], \quad (1.30)$$

Where $[\hat{H}, \hat{\rho}]$ is the commutator defined as $[\hat{H}, \hat{\rho}] = \hat{H}\hat{\rho} - \hat{\rho}\hat{H}$. This term, that is also present in the Liouville–von Neumann equation, is responsible for the unitary part of the evolution in the absence of coupling with the environment, which allows us to identify \hat{H} as the Hamiltonian (we set $\hbar = 1$ throughout this thesis). The super-operator \mathcal{L} is called the *Lindbladian* and the operators \hat{L}_k are called the *Lindblad operators* or *quantum jump operators*, representing the possible quantum jumps of the system due to coupling with the environment.

Note that the form of the master equation that we have derived is general for open quantum systems coupled to Markovian environments. To figure out the expression for the Lindblad operators \hat{L}_k , we need to consider specific models with more assumptions. As an example, for a bosonic lattice system coupled to a Markovian bath (environment) of harmonic oscillators and under the following assumptions,

- Born approximation: the system and the bath remain separable throughout the evolution and the state of the environment remains unchanged;
- Rotating wave approximation: all fast rotating terms in the interaction picture can be neglected;

the Lindblad operators can be identified to be $\hat{L}_i = \sqrt{\gamma_i} \hat{b}_i$, where \hat{b}_i is the bosonic annihilation operator for site i and γ_i is called the *decay rate*. In the case of a lattice of two level systems (for example, spin-1/2 particles) this operator becomes $\sqrt{\gamma_i} \hat{\sigma}_i^-$.

1.2.2 Dissipative phase transition

For an open quantum system whose dynamics are governed by the Lindblad master equation, it will finally reach the non-equilibrium steady state (NESS), characterised by

$$\frac{\partial}{\partial t} \hat{\rho}_{\text{NESS}} = \mathcal{L}[\hat{\rho}_{\text{NESS}}] = 0. \quad (1.31)$$

Dissipative phases of matter then refer to the different phases exhibited by the NESS of the open system. A dissipative phase transition happens when there is an abrupt change (non-analyticity) in the NESS of the system under the change of some external parameter. An analogy can be drawn

	Thermal phase transition (finite temperature)	Quantum phase transition ($T = 0$)	Dissipative phase transition (non-equilibrium)
System operator	Hamiltonian \hat{H}	Hamiltonian \hat{H}	Lindbladian \mathcal{L}
Relevant quantity	Free energy $F(\hat{\rho}) = \langle \hat{H} \rangle_{\hat{\rho}} - T \langle \hat{S} \rangle_{\hat{\rho}}$	Eigenvalues of \hat{H} : $E : \hat{H} \psi\rangle = E \psi\rangle$	Eigenvalues of \mathcal{L} : $\lambda : \mathcal{L}[\hat{\rho}] = \lambda \hat{\rho}$
Relevant state	Gibbs state, minimising F	Ground state, minimising E	Steady state, $\mathcal{L}[\hat{\rho}] = 0$
Phase transition	Non-analyticity in the Gibbs state	Non-analyticity in the ground state	Non-analyticity in the steady state

Table 1.1: Comparison between the thermal phase transition, the quantum phase transition and the dissipative phase transition.

between the dissipative phase transition and the quantum phase transition, as we present in Table 1.1 [8]. (For completeness, we also include the case of thermal phase transition therein.)

The evaluation of an observable \hat{O} , such as an order parameter, at NESS is carried out by

$$\langle \hat{O} \rangle_{\text{NESS}} = \text{Tr} [\hat{O} \hat{\rho}_{\text{NESS}}]. \quad (1.32)$$

To have a concrete example illustrating the concept of dissipative phase transition, let us look at the driven-dissipative Bose-Hubbard model as well as its hard core boson limit, the driven-dissipative XY model.

Driven-dissipative photonic Bose-Hubbard model

The driven-dissipative Bose-Hubbard (BH) model [9, 10] describes an array of coupled cavities with Kerr nonlinearity pumped by a coherent external laser source, as illustrated in Figure 1.2. It is defined by the Hamiltonian:

$$\hat{H}_{\text{BH}} = -J \sum_{\langle i,j \rangle} \hat{b}_i^\dagger \hat{b}_j + \sum_i \omega_c \hat{b}_i^\dagger \hat{b}_i + \sum_i \frac{U}{2} \hat{b}_i^\dagger \hat{b}_i^\dagger \hat{b}_i \hat{b}_i + \sum_i (F e^{-i\omega_p t} \hat{b}_i^\dagger + F^* e^{i\omega_p t} \hat{b}_i), \quad (1.33)$$

where J is the nearest neighbour hopping amplitude, \hat{b}_i^\dagger and \hat{b}_i are the bosonic creation and annihilation operators for site i , $\langle i,j \rangle$ denotes all pairs of nearest neighbour lattice indices, ω_c is the cavity frequency, U is the strength of the Kerr nonlinearity, F is the amplitude of the coherent

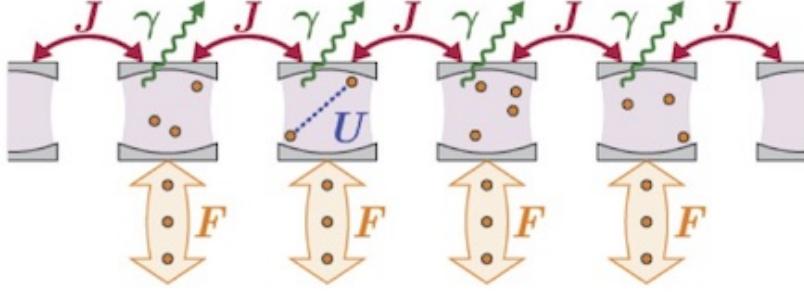


Figure 1.2: Schematic diagram of a driven-dissipative QED cavity array, where F is the pumping amplitude, U is the strength of the Kerr nonlinearity, J is the nearest neighbour hopping amplitude and γ is the decay rate [10].

pumping and ω_p is the pumping frequency. Note that without loss of generality, we can assume F to be real. We admit the Lindblad operator for each site to be $\sqrt{\gamma} \hat{b}_i$, as discussed in Section 1.2.1, which allows us to write the master equation for the system:

$$\frac{\partial}{\partial t} \hat{\rho} = -i[\hat{H}_{\text{BH}}, \hat{\rho}] + \sum_i \frac{\gamma}{2} (2\hat{b}_i \hat{\rho} \hat{b}_i^\dagger - \hat{b}_i^\dagger \hat{b}_i \hat{\rho} - \hat{\rho} \hat{b}_i^\dagger \hat{b}_i). \quad (1.34)$$

Note that the Lindbladian is time-dependent due to the pumping term in the Hamiltonian. To facilitate our discussion, we can transform into the frame rotating at the pump frequency via the unitary transformation:

$$\rho \rightarrow \hat{U} \hat{\rho} \hat{U}^\dagger, \quad (1.35)$$

where $\hat{U} = e^{i\omega_p t \sum_j \hat{b}_j^\dagger \hat{b}_j}$. This transformation will leave the master equation in the same form except that the Hamiltonian is now changed to the time-independent rotating frame Hamiltonian:

$$\hat{H}_{\text{BH}}^{\text{rf}} = -J \sum_{\langle i,j \rangle} \hat{b}_i^\dagger \hat{b}_j + \sum_i \Delta \hat{b}_i^\dagger \hat{b}_i + \sum_i \frac{U}{2} \hat{b}_i^\dagger \hat{b}_i^\dagger \hat{b}_i \hat{b}_i + \sum_i F(\hat{b}_i^\dagger + \hat{b}_i), \quad (1.36)$$

where $\Delta = \omega_c - \omega_p$ is the detuning between the cavity and the pumping frequency. From now on, we will work in the rotating frame. The steady state of the system will be a compromise between the pumping, which increases the number of bosons in the system, and the dissipation, which leaks the bosons to the environment.

For a simple discussion of the dissipative phase transition, we now take the hard core boson limit by setting the nonlinearity $U/J \rightarrow \infty$. As a result, each cavity can have at most 1 boson due to the strong repulsion quantified by U , becoming effectively a two level system. As the Hamiltonian can now be described using Pauli matrices and that the hopping

terms can be written in the form of XY coupling, it is also known as the XY Hamiltonian.

Dissipative phase transition in the XY model

The driven-dissipative XY Hamiltonian can be derived directly from that for the BH model (Equation (1.36)), by replacing the operators with their counterparts (the Pauli transition matrices) for two level systems:

$$\hat{H}_{XY}^{\text{rf}} = -J \sum_{\langle i,j \rangle} \hat{\sigma}_i^+ \hat{\sigma}_j^- + \sum_i \Delta \hat{\sigma}_i^+ \hat{\sigma}_i^- + \sum_i F(\hat{\sigma}_i^+ + \hat{\sigma}_i^-). \quad (1.37)$$

The corresponding master equation is

$$\frac{\partial}{\partial t} \hat{\rho} = -i[\hat{H}_{XY}^{\text{rf}}, \hat{\rho}] + \sum_i \frac{\gamma}{2} (2\hat{\sigma}_i^- \hat{\rho} \hat{\sigma}_i^+ - \hat{\sigma}_i^+ \hat{\sigma}_i^- \hat{\rho} - \hat{\rho} \hat{\sigma}_i^+ \hat{\sigma}_i^-). \quad (1.38)$$

Note that the equations are written in the rotating frame. To depict the phase transition, we consider the normalised two-site correlation function

$$C(j, r) = \frac{\langle \hat{\sigma}_j^z \hat{\sigma}_{j+r}^z \rangle}{\langle \hat{\sigma}_j^z \rangle \langle \hat{\sigma}_{j+r}^z \rangle}. \quad (1.39)$$

as done in [11], where the authors simulated a 1D chain of $N = 61$ sites with open boundary conditions using a matrix product based method and studied the NESS correlations at the centre of the system (i.e., for $j = \lceil N/2 \rceil$), denoted as $C(r)$, as a function of Δ/γ , as shown in Figure 1.3. As the detuning increases to a certain value, the $C(1)$ correlation drops below 1, which implies a transition from a regime where neighbouring excitations tend to cluster together to a configuration where such clustering of excitations is disfavoured. In other words, the system undergoes a bunching-antibunching transition as the detuning is changed. Note that the mean field approximation will fail to capture this transition as it ignores the correlations between sites and the function $C(r)$ will always have value 1.

1.2.3 Review of recent works on dissipative phase transitions

Let us briefly review some of the other theoretical and experimental works in dissipative phase transitions.

Phase diagram of the dissipative quantum Ising model on a square lattice [12]. The authors studied the dissipative transverse field Ising model

$$\hat{H} = \frac{V}{4} \sum_{\langle i,j \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z + \frac{g}{2} \sum_i \hat{\sigma}_i^x, \quad (1.40)$$

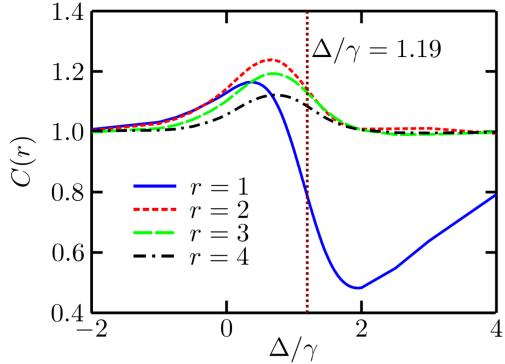


Figure 1.3: Two-site correlations $C(r)$ from the central site of the system and $r = 1, 2, 3, 4$ as a function of Δ/γ , for $J/\gamma = 2$, $F/\gamma = 1$, $N = 61$ and $\chi = 200$ (bound dimension in the matrix product ansatz). the vertical line indicates the point of fastest decay of correlations $C(1)$. Figure taken from [11].

using cluster methods and quantum trajectories for square lattice systems in two cases: 1) the dissipation is along the interaction direction with $\hat{L}_i^{(z)} = \frac{\sqrt{\gamma}}{2}(\hat{\sigma}_i^x - i\hat{\sigma}_i^y)$; 2) the dissipation is along the field direction with $\hat{L}_i^{(x)} = \frac{\sqrt{\gamma}}{2}(\hat{\sigma}_i^z - i\hat{\sigma}_i^y)$. They found that in the first case, while the mean field approximation predicts bistability, when short-range interactions are taken into account, the systems shows a first-order phase transition. In the second case, an exact treatment of short-range correlations leads to the presence of second-order transitions, which were thought as first order in previous literature.

Bose-Hubbard Model: Relation Between Driven-Dissipative Steady-States and Equilibrium Quantum Phases [13]. The authors studied the driven-dissipative BH model in the limit of weak pumping and weak dissipation, and obtained analytical solutions for the mean-field master equation. They studied the process of n -photon resonance, where the absorption of photons is resonantly enhanced when n incident driving photons have the same energy as n photons in the cavity:

$$n\omega_p = n\omega_c + \frac{U}{2}n(n-1) \Leftrightarrow \frac{U}{-\Delta} = \frac{2}{n-1}, \quad (1.41)$$

and found that if the intensity of the pump is sufficiently large when compared with the dissipation rate, a single cavity can be driven into a statistical mixture that has the same second-order correlation function $g^{(2)}(0) = \langle \hat{b}^\dagger \hat{b} \hat{b}^\dagger \hat{b} \rangle / \langle \hat{b}^\dagger \hat{b} \rangle^2$ as a pure Fock state with n photons, and a mean photon number of $n/2$. They also showed that when the coupling strength between sites is above a critical value, the system will be driven into a

classical coherent state.

Fermionized photons in an array of driven-dissipative nonlinear cavities [14]. The authors studied the driven-dissipative BH model in the hard core boson limit, which allows a maximum of 1 photon per cavity, mimicking the Pauli exclusion principle. In this limit, bosonic operators are mapped to fermionic ones by a Jordan-Wigner transformation. The authors predict in this regime a fermionisation of photons and the onset of a Tonks-Girardeau phase, as confirmed by the value of second-order cross and autocorrelation functions.

Polariton crystallization in driven arrays of lossy nonlinear resonators [15]. The authors studied the driven-dissipative BH model as well but with a site-dependent alternating drive described by the term $\sum_j (\Omega_j \hat{a}_j^\dagger + \Omega_j^* \hat{a}_j)$ with $\Omega_j = \Omega e^{-i\phi_j}$. They found that for strong cavity nonlinearities, polariton crystallisation can be identified from density-density correlations.

Photon solid phases in driven arrays of nonlinearly coupled cavities [16]. This work extended the idea of photon crystal to arrays of cavities with cross-Kerr nonlinearities by adding the term $V \sum_{\langle i,j \rangle} \hat{b}_i^\dagger \hat{b}_j^\dagger \hat{b}_j \hat{b}_i$ to the BH Hamiltonian. Through mean field and numerical studies, they identified a photon crystal phase in the NESS of the system.

Observation of a dissipative phase transition in a one-dimensional circuit QED lattice [17]. This work studied a 1D chain of 72 coupled microwave cavities, as shown in Figure 1.4, each coupled to a superconducting qubit, with coherent driving. They found experimental evidence for a dissipative phase transition in the NESS of the system. Bistability has also been measured near the boundary of two phases.

A dissipatively stabilised Mott insulator of photons [18]. The authors constructed a 8-site driven-dissipative BH lattice for microwave photons in the strong interaction regime using superconducting circuits. They developed a dissipative scheme to populate and stabilise the Mott insulator phase of photons.

1.2.4 Numerical techniques for probing dissipative phases of matter

Before we end this chapter, we briefly summarise the typical numerical techniques that have been developed for probing phases of matter in open quantum systems.

When the system size is small, the master equation can be explicitly constructed and we can find the NESS by a brute force integration to obtain the time evolution of the system until it reaches the steady state, or alternatively, we can perform an exact diagonalisation of the superoperator \mathcal{L} and find the eigen-state corresponding to the 0 eigenvalue. This approach provides exact solutions with little approximation, and yet it

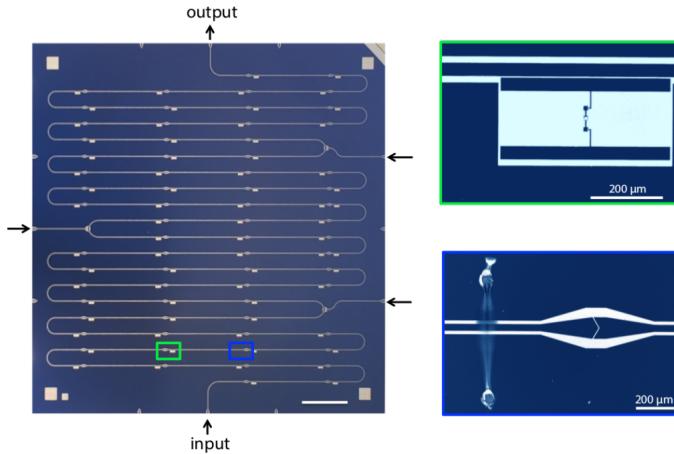


Figure 1.4: The 72-site circuit QED lattice used in [17].

cannot scale up to large systems due to the exponential growth of the dimension of the Hilbert space.

Another common approach is the mean field theory, where one ignores the correlations and solves the master equation for the mean field density matrix, that has the same dimension as in the case of a single site. However, this approximation ignores local fluctuations and is inaccurate in low dimensions.

More sophisticated numerical methods have been developed aiming to efficiently study open quantum systems, such as matrix product operators [19], Monte Carlo wave function method [20] and corner-space renormalisation method [21]. However, these methods are mainly limited to specific regimes of problems. For example, matrix product operator techniques are mainly restricted to one-dimensional systems and are adequate for systems with short-range correlations.

In recent years, artificial neural networks are emerging as a new technique for probing quantum phases of matter in both closed systems and open systems, which is still in its very infancy. This approach is the focus of our study, that will be introduced and discussed in detail in the following chapters.

1.3 Conclusion

In this chapter, we have introduced the key concepts to understanding this thesis, with focus on the driven-dissipative phase transitions of lattice models, whose time evolution can be described by the Lindblad master equation. The rest of this manuscript will be focusing on finding the NESS from the master equation using machine learning algorithms.

Chapter 2

Machine learning quantum phases of matter

Having seen the concept of dissipative phase transitions in open quantum systems, we now present how we can solve for the NESS of the system using machine learning techniques. As we have mentioned in the Introduction, the exact solution for a many-body system lies in a Hilbert space whose dimension grows exponentially with the number of particles in the system. Therefore, a feasible solution given today’s computer technology requires us to represent the state using less information. The key concept in this chapter is the representation of quantum states using artificial neural networks, as well as the algorithm for training the network in order to “learn” the steady state of the system. We will start with a brief introduction to the concepts in machine learning and then review recent works that studied quantum systems using artificial neural networks. We will also propose a network structure that is adapted to quantum systems with translational symmetry, which proves to be more efficient and to give better performance, as we will show in the next chapter with numerical results.

2.1 Machine learning basics

The term “machine learning” can be formally defined as follows [22]: A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T , as measured by P , improves with experience E . In practice, it is a way of designing algorithms without giving explicit instructions to the machine. For example, the task of recognising hand-written numbers can be viewed as fitting a highly nonlinear function, whose inputs are the pixels in the images and the outputs are the digits recognised by the machine. Instead of explicitly coding this nonlinear function by hand, we set a general ansatz with tunable parameters, such as an artificial neural network, that is

expected to be good at approximating other functions, and let the machine find the right parameters itself through a process called *training*, after which this nonlinear function can output the correct values on input images that it has never seen before. In our case of studying open quantum systems, we want to train a neural network to represent the density matrix of the NESS of the system.

In general, machine learning algorithms can be categorised as follows:

Supervised learning: the machine tries to build a mathematical model through a set of labelled data [23], known as training data, where each training example in the data contains both the input and the desired output. After training on those data, the machine is capable of predicting the output from new input data. Such algorithm is widely used in regression and classification [24]. The example of image recognition we presented above falls in this category.

Unsupervised learning: the machine takes a set of unlabelled training data consisting of only input data without the desired output. The algorithm identifies patterns in the data, as commonly used in data clustering.

Reinforcement learning: this type of learning algorithms does not require training data. Instead, the machine learns how to take actions in a certain environment, based on a penalty-and-reward mechanism, so as to optimise an *objective*. It can be, for example, used to train a game playing software against a human opponent. This mechanism may remind us of the standard variational method used in quantum mechanics for finding the ground state for a given Hamiltonian, where we try to adjust the variational parameters in the wave function to minimise the energy (the objective, or the *cost function*). Indeed, when we have a complicated ansatz function for the state, instead of solving for the parameters analytically by hand, we can implement reinforcement learning algorithms to learn the state.//

In parallel with these algorithms, the machine learning community uses a variety of artificial neural networks, among which a simple yet powerful model is the *restricted Boltzmann machine* (RBM), that was initially invented in 1986 [25] and boosted the development of machine learning. The RBMs can be trained to perform various tasks, including dimensionality reduction [26], classification [27], etc., and are recently applied to the study of quantum many-body systems.

2.2 Recent works applying machine learning to open quantum systems

The idea of solving quantum many-body problems with artificial neural networks and machine learning was first proposed and carried out in 2017 by Carleo and Troyer [6], where they showed that the restricted Boltzmann

machine (RBM) can well approximate the wave function for a closed interacting spin system both in one and two dimensions. This work sparked tremendous interest at the interface between machine learning and quantum many-body physics.

In early 2019, four research groups [28] adapted Carleo and Troyer’s approach to open quantum systems. We briefly summarise their works in Table 2.1, together with our work done in the scope of this thesis. Their representations of the density matrix are based on the RBM, and they explored different methods of training the neural networks.

They benchmarked the learning algorithms against various spin models of small lattice sizes, whose exact solutions are accessible, namely the transverse-field Ising model in 1D [29, 30], the *XYZ* model in 1D [29] and 2D [31], and the anisotropic Heisenberg model in 1D [32]. All the models they tested have periodic boundary conditions and the sizes range from 4 to 16 spins. They all found good agreement between the machine learning results and the exact solutions.

In this thesis, we will adopt the neural network presented in [33, 31] and the cost function in [34, 30], which is a novel combination to our knowledge. In addition, we propose a new network ansatz that is optimised for lattice models with symmetries and we will test them on toy models with periodic boundary conditions that have translational invariance.

2.3 Artificial neural networks for open quantum systems

An artificial neural network (ANN) [35], as the name suggests, is inspired from the immense neural connections in our brains. It is one of the most important models used in machine learning algorithms for building mathematical models and approximating functions. They stand behind many technologies that serve us today, such as pattern recognition, product recommendation, data mining, game playing, etc.

An ANN is a set of connected nodes (units, or neurons), where each node can have a bias and each connection is represented by a weight, that are to be adjusted in the training process. ANNs can be categorised into different types depending on the connections between the nodes. An ANN with two layers of nodes and with connections only between the two layers is called a restricted Boltzmann machine (RBM), that is mathematically proven to be able to well approximate many-body wave functions. We will now construct a neural network representation for the density matrix of an open quantum system using an RBM, which contains much fewer free parameters compared to the dimension of the Hilbert space of the physical

Team	ANN ansatz	Training method	Toy model
Hartmann and Carleo [32]	RBM with ancillary units	Time evolution generated by \mathcal{L}	1D anisotropic Heisenberg model
Nagy and Savona [31]	RBM with ancillary units	Time evolution generated by \mathcal{L}	2D XYZ model
Vicentini <i>et al.</i> [30]	2 RBMs with ancillary units, for phase and amplitude respectively	Minimisation based on $\mathcal{L}^\dagger \mathcal{L}$	1D Transverse field Ising model
Yoshioka and Hamazaki [29]	RBM without ancillary units	Minimisation based on $\mathcal{L}^\dagger \mathcal{L}$	1D transverse field Ising model and 1D XYZ model
Author of this thesis	RBM with ancillary units respecting lattice symmetries	Minimisation based on $\mathcal{L}^\dagger \mathcal{L}$	1D and 2D transverse field Ising model, 1D XY model and 1D Bose-Hubbard model

Table 2.1: Summary of recent works applying artificial neural networks to open quantum systems, in comparison with our work presented in this thesis.

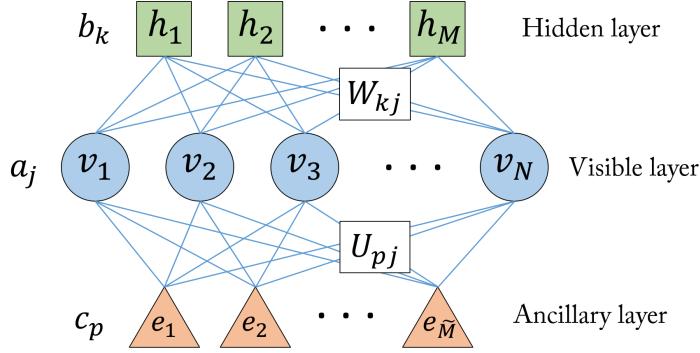


Figure 2.1: Graphical representation of the RBM wave function ansatz with N visible units, M hidden units and \tilde{M} ancillary units. The physical quantum numbers $\{v_j\}$ are encoded in the visible layer with biases $\{a_j\}$. The ancillary units $\{e_p\}$ have biases $\{c_p\}$ and the hidden units $\{h_k\}$ have biases $\{b_k\}$. The interaction weights between the visible layer and the hidden layer are denoted as $\{W_{kj}\}$ and those between the visible layer and the ancillary layer are denoted as $\{U_{pj}\}$.

system, thus avoiding the problem of exponential scaling.

For an open quantum system, we can add ancillary degrees of freedom into consideration to mathematically represent the effect of the environment. For simplicity, we choose the ancillary units to be spin-1/2 units. The total state of the system and the environment can be expanded in the basis of quantum numbers:

$$|\psi\rangle = \sum_{v,e} \psi(v,e) |v,e\rangle, \quad (2.1)$$

where

$$v = (v_1, v_2, \dots, v_N) \quad (2.2)$$

are the physical quantum numbers and

$$e = (e_1, e_2, \dots, e_{\tilde{M}}) \quad (2.3)$$

are the ancillary ones for the \tilde{M} spin-1/2 ancillary units. We can write an ansatz for the wave function $\psi(v,e)$ using the RBM graphically represented in Figure 2.1, with the expression

$$\begin{aligned} \psi_W(v,e) = & \sum_{\{h_p=\pm 1\}} \exp \left(\sum_{j=1}^N a_j v_j + \sum_{k=1}^M b_k h_k + \sum_{p=1}^{\tilde{M}} c_p e_p \right. \\ & \left. + \sum_{j=1}^N \sum_{k=1}^M W_{kj} h_k v_j + \sum_{j=1}^N \sum_{p=1}^{\tilde{M}} U_{pj} e_p v_j \right), \end{aligned} \quad (2.4)$$

where the v_j 's are the values taken by the visible units, each having a bias a_j , the h_k 's are the values of the hidden units that can be either $+1$ or -1 , each having a bias b_k , the e_p 's are the values of the ancillary units that can be either $+1$ or -1 , each having a bias c_p , the W_{kj} 's are the weights between the visible layer and the hidden layer and the U_{pj} 's are the weights¹ between the visible layer and the ancillary layer. This structure can be intuitively understood as follows: the hidden layer is responsible for generating correlation and entanglement between the sites represented by the visible layer, and the ancillary layer can mathematically represent the environment, allowing us to simulate an open system. The weights and the biases are complex-valued in order to represent a general wave function. \mathcal{W} denotes the set of all variational parameters, i.e.

$$\mathcal{W} = \{a_j, b_k, c_p, W_{kj}, U_{pj} \mid j = 1, \dots, N; k = 1, \dots, M; p = 1, \dots, \tilde{M}\} \quad (2.5)$$

for an RBM with N visible units, M hidden units and \tilde{M} ancillary units. The density matrix can be obtained by performing a partial trace over the ancillary degrees of freedom of the ansatz state:

$$\begin{aligned} \hat{\rho}_{\mathcal{W}} &= \text{Tr}_e[|\psi_{\mathcal{W}}\rangle\langle\psi_{\mathcal{W}}|] \\ &= \sum_e \langle e | \left[\sum_{l,e'} \psi_{\mathcal{W}}(l, e') |l, e'\rangle \right] \left[\sum_{r,e''} \langle r, e'' | \psi_{\mathcal{W}}^*(r, e'') \right] |e\rangle \\ &= \sum_{l,r} \left[\sum_e \psi_{\mathcal{W}}(l, e) \psi_{\mathcal{W}}^*(r, e) \right] |l\rangle\langle r| \\ &= \sum_{l,r} \rho_{\mathcal{W}}(l, r) |l\rangle\langle r|, \end{aligned} \quad (2.6)$$

where

$$\begin{aligned} \rho_{\mathcal{W}}(l, r) &= \sum_e \psi_{\mathcal{W}}(l, e) \psi_{\mathcal{W}}^*(r, e) \\ &= \sum_{\{e_p=\pm 1\}} \psi_{\mathcal{W}}(l, e) \psi_{\mathcal{W}}^*(r, e). \end{aligned} \quad (2.7)$$

Note that the summation over h_p in Equation (2.4) and that the summation over e_p in Equation (2.7) can be carried out analytically, giving the final

¹Not to be confused with the time evolution operator \hat{U} or the nonlinearity strength U in the Bose-Hubbard model. When referred to as the weight, we always keep its subscripts.

expression for the ANN ansatz of the density matrix element:

$$\begin{aligned}
\rho_{\mathcal{W}}(l, r) &= \exp \left[\sum_{j=1}^N (a_j l_j + a_j^* r_j) \right] \\
&\times \prod_{k=1}^M 2 \cosh \left(b_k + \sum_{j=1}^N W_{kj} l_j \right) \\
&\times \prod_{k=1}^M 2 \cosh \left(b_k^* + \sum_{j=1}^N W_{kj}^* r_j \right) \\
&\times \prod_{p=1}^{\tilde{M}} 2 \cosh \left[c_p + c_p^* + \sum_{j=1}^N (U_{pj} l_j + U_{pj}^* r_j) \right].
\end{aligned} \tag{2.8}$$

The total number of complex variational parameters is therefore

$$|\mathcal{W}| = (N+1)(M+\tilde{M}) + N, \tag{2.9}$$

which does not scale exponentially with N like the dimension of the Hilbert space. Since Equation (2.8) is, in general, a non-analytic function of the variational parameters, the real parts and the imaginary parts should be regarded as independent real parameters. We denote this ensemble of real parameters as \mathcal{R} . We do not include the imaginary parts of the c_p 's in \mathcal{R} since they are not involved in the ansatz. We further define

$$\alpha = \frac{M}{N}, \quad \beta = \frac{\tilde{M}}{N} \tag{2.10}$$

to be the hidden unit density and the ancillary unit density respectively. The log-derivatives of the ansatz in Equation (2.8) with respect to the real parameters can be easily calculated, as shown in Appendix A, and will be used in a gradient-based optimisation algorithm that we will introduce later in this chapter.

RBM with symmetries

When the Hamiltonian and the Lindbladian of the system exhibit intrinsic symmetries, they must also be satisfied by the NESS of the system. This property can be exploited to reduce the number of variational parameters in the ansatz, by enforcing the ansatz to respect those symmetries of the system. Therefore, we propose a symmetrised RBM ansatz inspired from [6] that implemented this idea when solving ground states for closed systems, in order to reduce the complexity of the RBM ansatz.

Consider, for example, a Lindbladian with translational symmetry, i.e. the state remains invariant under the site index permutations

$$i \mapsto P_k(i), \tag{2.11}$$

where P_k is an element in the translation group and $k = 1, \dots, N_p$, with N_p being the number of permutations in the group. In the case of translational invariance, the orbit has the same size as the lattice, i.e. $N_p = N$. We can enforce our ansatz for the NESS density matrix element to be invariant under these permutations by defining

$$\begin{aligned} \rho_{\mathcal{W}^s}(l, r) &= \exp \left(a^s \sum_{i=1}^N l_i + a^{s*} \sum_{i=1}^N r_i \right) \\ &\times \prod_{f=1}^{\alpha} \prod_{k=1}^{N_p} 2 \cosh \left(b_f^s + \sum_{i=1}^N W_{f,P_k(i)}^s l_i \right) \\ &\times \prod_{f=1}^{\alpha} \prod_{k=1}^{N_p} 2 \cosh \left(b_f^{s*} + \sum_{i=1}^N W_{f,P_k(i)}^{s*} r_i \right) \\ &\times \prod_{g=1}^{\beta} \prod_{k=1}^{N_p} 2 \cosh \left[c_g^s + c_g^{s*} + \sum_{i=1}^N \left(U_{g,P_k(i)}^s l_i + U_{g,P_k(i)}^{s*} r_i \right) \right], \end{aligned} \quad (2.12)$$

where the symmetrised parameters now read

$$\mathcal{W}^s = \{a^s, b_f^s, c_g^s, W_{fi}^s, U_{gi}^s \mid i = 1, \dots, N; f = 1, \dots, \alpha; g = 1, \dots, \beta\}. \quad (2.13)$$

They can be mapped to the parameters in the standard ansatz without symmetry defined in equation (2.8) by:

$$a_j = a^s, \quad (2.14)$$

$$b_j = b_f^s, \quad f = \lfloor j/N_p \rfloor, \quad (2.15)$$

$$c_m = c_g^s, \quad g = \lfloor m/N_p \rfloor, \quad (2.16)$$

$$W_{ji} = W_{f,P_k(i)}^s, \quad k \equiv j \pmod{N_p}, \quad (2.17)$$

$$U_{mi} = U_{g,P_k(i)}^s. \quad (2.18)$$

We can see that this ansatz corresponds to a standard RBM with α and β being the hidden unit density and the ancillary unit density respectively, while the number of variational parameters is now

$$|\mathcal{W}^s| = (\alpha + \beta)(N + 1) + 1, \quad (2.19)$$

that scales linearly with the system size for fixed α and β , whereas in the ansatz without symmetry the scaling is quadratic.

2.4 Variational Monte Carlo for finding the NESS

Having defined the ANN ansatz for the NESS density matrix, we now present how to train it to learn the NESS solution. In this section, we will

drop the subscripts \mathcal{W} and \mathcal{W}^s for the ansatz when there is no ambiguity, for a simpler notation. The training process is based on the variational Monte Carlo (VMC) method, which is developed for finding the ground state of closed systems [36] but can be directly generalised to open systems.

2.4.1 Variational principle for open quantum systems

We will adopt a reinforcement learning scheme to solve for the NESS, which requires a cost function, which is a functional of the ansatz, to be minimised. The choice of the cost function is inspired from the variational principle for closed quantum systems. Recall that for a given Hamiltonian \hat{H} , its ground state $|\epsilon_0\rangle$ satisfies the time-independent Schrödinger equation,

$$\hat{H}|\epsilon_0\rangle = E_0|\epsilon_0\rangle, \quad (2.20)$$

where the ground state energy E_0 is the smallest eigenvalue of \hat{H} . It then follows that for an arbitrary state $|\psi\rangle$ (which we no not assume to be normalised), if we define a functional of $|\psi\rangle$ to be

$$E[|\psi\rangle] = \langle\hat{H}\rangle = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle}, \quad (2.21)$$

we always have the inequality

$$E[|\psi\rangle] \geq E_0, \quad (2.22)$$

where equality is satisfied if and only if $|\psi\rangle$ is the ground state of the system. This method has been successfully exploited in machine learning the ground state for closed quantum systems with a neural network ansatz, as originally proposed in [6]. Mathematically, this amounts to finding the eigenvector of a Hermitian operator that has the smallest eigenvalue. As the NESS density matrix satisfies

$$\mathcal{L}[\hat{\rho}_{\text{NESS}}] = 0, \quad (2.23)$$

it is an eigenvector of the (super-)operator \mathcal{L} with eigenvalue 0. Therefore, it is also an eigenvector of the Hermitian operator $\mathcal{L}^\dagger\mathcal{L}$ since

$$\mathcal{L}^\dagger\mathcal{L}[\hat{\rho}_{\text{NESS}}] = 0. \quad (2.24)$$

Note that $\mathcal{L}^\dagger\mathcal{L}$ is obviously positive-semidefinite, this means that $\hat{\rho}_{\text{NESS}}$ is the eigenvector with the smallest eigenvalue of $\mathcal{L}^\dagger\mathcal{L}$. This allows us to define a cost function by drawing a direct analogy from equation (2.21), as proposed in [37]:

$$\mathcal{C}[\hat{\rho}] = \langle\mathcal{L}^\dagger\mathcal{L}\rangle = \frac{\|\mathcal{L}[\hat{\rho}]\|_2^2}{\|\hat{\rho}\|_2^2} = \frac{\text{Tr}[\hat{\rho}^\dagger\mathcal{L}^\dagger\mathcal{L}\hat{\rho}]}{\text{Tr}[\hat{\rho}^\dagger\hat{\rho}]} \geq 0, \quad (2.25)$$

where $\|\cdot\|_2$ denotes the 2-norm, which, when minimised, gives the NESS density matrix $\hat{\rho}_{\text{NESS}}$. As a remark on the practical implementation, the relevant quantities can be calculated by considering the vectorised density matrix defined by the map

$$\hat{\rho} = \sum_{l,r} \rho(l,r) |l\rangle \langle r| \mapsto |\rho\rangle = \sum_{l,r} \rho(l,r) |l,r\rangle, \quad (2.26)$$

that exploits the isomorphism

$$\mathbb{C}^m \otimes \mathbb{C}^n \cong \mathbb{C}^{mn}. \quad (2.27)$$

The action of operators acting on $\hat{\rho}$ can be mapped as follows:

$$\hat{A}\hat{\rho}\hat{B} \mapsto |A\rho B\rangle = \hat{A} \otimes \hat{B}^T |\rho\rangle. \quad (2.28)$$

The Lindbladian can then be mapped to

$$\mathcal{L} \mapsto \hat{\mathcal{L}} = -i\left(\hat{H} \otimes \mathbb{1} - \mathbb{1} \otimes \hat{H}^T\right) + \sum_k \left(\hat{L}_k \otimes \hat{\bar{L}}_k - \frac{1}{2} \hat{L}_k^\dagger \hat{L}_k \otimes \mathbb{1} - \frac{1}{2} \mathbb{1} \otimes \hat{L}_k^T \hat{\bar{L}}_k \right), \quad (2.29)$$

which can be written as a matrix acting on the vector $|\rho\rangle$. Note that the full matrix \mathcal{L} need not be calculated and stored during the optimisation, as can be seen from the algorithm described in the next section. For a Hamiltonian that is a sum of localised terms, the matrix elements of \mathcal{L} can be efficiently computed. The cost function can then be written as

$$\mathcal{C}[|\rho\rangle] = \frac{\langle \rho | \hat{\mathcal{L}}^\dagger \hat{\mathcal{L}} | \rho \rangle}{\langle \rho | \rho \rangle} \quad (2.30)$$

in vectorised notations, and the NESS solution $|\rho_{\text{NESS}}\rangle$ can be regarded as the “ground state” of the effective Hamiltonian $\hat{\mathcal{L}}^\dagger \hat{\mathcal{L}}$ with “energy” $E_0 = 0$.

2.4.2 Evaluating the cost function and its gradient

To minimise the cost function, we adopt a gradient descent scheme, where the variational parameters are adjusted according to the gradient of the cost function until convergence. This requires us to measure the cost function and its partial derivatives with respect to the variational parameters.

Although the number of variational parameters we use to represent the density matrix is much smaller than the dimension of the Hilbert space, the full expression of the ansatz defined by Equation (2.6) still contains exponentially many terms, and so does the cost function and its gradient. Luckily, as we will show below, the cost function and its gradient can be measured through a stochastic sampling scheme, allowing us to estimate

these quantities using a sample drawn from the full space, which avoids calculating the exponentially many terms in the full expansion.

To begin with, it is convenient to rewrite the cost function in the following way:

$$\begin{aligned}
\mathcal{C}[|\rho\rangle] &= \frac{\langle\rho|\hat{\mathcal{L}}^\dagger\hat{\mathcal{L}}|\rho\rangle}{\langle\rho|\rho\rangle} \\
&= \frac{\langle\rho|(\sum_{l,r}|l,r\rangle\langle l,r|)\hat{\mathcal{L}}^\dagger\hat{\mathcal{L}}(\sum_{l',r'}|l',r'\rangle\langle l',r'|)|\rho\rangle}{\langle\rho|(\sum_{l,r}|l,r\rangle\langle l,r|)|\rho\rangle} \\
&= \frac{\sum_{l,r}\sum_{l',r'}\rho^*(l,r)\hat{\mathcal{L}}^\dagger\hat{\mathcal{L}}(l,r,l',r')\rho(l',r')}{\sum_{l,r}\rho^*(l,r)\rho(l,r)} \\
&= \frac{\sum_{l,r}\rho^*(l,r)\rho(l,r)\sum_{l',r'}\hat{\mathcal{L}}^\dagger\hat{\mathcal{L}}(l,r,l',r')\frac{\rho(l',r')}{\rho(l,r)}}{\sum_{l,r}\rho^*(l,r)\rho(l,r)} \\
&= \sum_{l,r}\Pi(l,r)\mathcal{C}^{\text{loc}}(l,r) \\
&= \langle\langle\mathcal{C}^{\text{loc}}\rangle\rangle,
\end{aligned} \tag{2.31}$$

where

$$\Pi(l,r) = \frac{\rho^*(l,r)\rho(l,r)}{\sum_{l,r}\rho^*(l,r)\rho(l,r)}, \tag{2.32}$$

$$\mathcal{C}^{\text{loc}}(l,r) = \sum_{l',r'}\hat{\mathcal{L}}^\dagger\hat{\mathcal{L}}(l,r,l',r')\frac{\rho(l',r')}{\rho(l,r)}, \tag{2.33}$$

and $\langle\langle\mathcal{C}^{\text{loc}}\rangle\rangle$ denotes the statistical expectation of $\mathcal{C}^{\text{loc}}(l,r)$ over the probability distribution $\Pi(l,r)$. This interpretation allows us to measure the cost function via stochastic sampling, as we will introduce later. Using similar notations, we can show that the gradient of the cost function can also be written as the expectation of some estimator. By denoting

$$D_{p_i}(l,r) = \frac{\partial \ln \rho(l,r)}{\partial p_i} \tag{2.34}$$

with $p_i \in \mathcal{R}$ being a real variational parameter of the ansatz, and defining

$$G_{p_i}(l,r) = 2\text{Re}[(\mathcal{C}^{\text{loc}}(l,r) - \langle\langle\mathcal{C}\rangle\rangle)D_{p_i}^*(l,r)] \tag{2.35}$$

to be the gradient estimator, we can verify that

$$\frac{\partial}{\partial p_i}\mathcal{C} = \langle\langle G_{p_i}\rangle\rangle. \tag{2.36}$$

2.4.3 Stochastic sampling

The estimation of the cost function and its gradient can be achieved via stochastic sampling over the probability distribution $\Pi(x)$, where x denotes the string (l, r) , which can be done by a Markov chain Monte Carlo (MCMC) process, where we generate a Markov chain of the configurations $x^{(1)} \rightarrow x^{(2)} \rightarrow \dots x^{(k)} \rightarrow \dots$, sampling the distribution $\Pi(x)$. This can be done using the Metropolis-Hastings algorithm [38], where at each step of the Markov chain, an update of configuration from x to x' is randomly proposed with probability $Q(x'|x)$ and the new configuration is accepted with probability

$$\begin{aligned} A(x \rightarrow x') &= \min \left\{ 1, \frac{\Pi(x')}{\Pi(x)} \frac{Q(x|x')}{Q(x'|x)} \right\} \\ &= \min \left\{ 1, \frac{|\rho(x')|^2}{|\rho(x)|^2} \frac{Q(x|x')}{Q(x'|x)} \right\}. \end{aligned} \quad (2.37)$$

The estimation of the quantities can be achieved by calculating the estimators on each of the configurations in the Markov chain and then taking the average.

We can also measure the physical observables of interest by stochastic sampling via another Markov chain. The expectation value of an observable \hat{O} can be calculated as

$$\begin{aligned} \langle \hat{O} \rangle &= \frac{\text{Tr}[\hat{\rho}\hat{O}]}{\text{Tr}[\hat{\rho}]} \\ &= \sum_l \frac{\rho(l,l)}{\text{Tr}[\hat{\rho}]} \sum_r \frac{\rho(l,r)O(r,l)}{\rho(l,l)} \\ &= \sum_l \Pi^{\text{obs}}(l)O^{\text{loc}}(l) \\ &= \langle\langle O^{\text{loc}} \rangle\rangle, \end{aligned} \quad (2.38)$$

where

$$\Pi^{\text{obs}}(l) = \frac{\rho(l,l)}{\text{Tr}[\hat{\rho}]}, \quad (2.39)$$

$$O^{\text{loc}}(l) = \sum_r \frac{\rho(l,r)O(r,l)}{\rho(l,l)}. \quad (2.40)$$

This allows us to estimate the expectation value of observable \hat{O} with the estimator $O^{\text{loc}}(l)$ on the probability distribution $\Pi^{\text{obs}}(l)$, using the same algorithm as described above.

2.4.4 Natural gradient adaptation

Now that we know how to estimate the cost function and its gradient, we are left with the parameter updating rules that will optimise the cost function towards the minimum.

In a standard *steepest gradient descent* method [39], the parameters $p = (p_1, p_2, \dots, p_{N_{\text{par}}})^T$ are updated according to

$$p \leftarrow p - \eta \nabla \mathcal{C}, \quad (2.41)$$

where η is called the *learning rate* and

$$\nabla \mathcal{C} = \left(\frac{\partial \mathcal{C}}{\partial p_1}, \frac{\partial \mathcal{C}}{\partial p_2}, \dots, \frac{\partial \mathcal{C}}{\partial p_{N_{\text{par}}}} \right)^T \quad (2.42)$$

is the gradient of the cost function. Inspired by [6], for better convergence of the cost function towards the global minimum, we add a correction to this method known as the natural gradient adaptation [40, 41], which is commonly used in the machine learning community. Intuitively, this necessity can be understood as follows. The steepest gradient descent works best in the case where the space of the variational parameters is Euclidean, which is not the general case [40], implying that the updating rule given above may not follow the steepest descent direction of the cost function in the parameter space. This can be corrected by considering the distance defined by

$$\delta s^2 = \left\| \frac{\hat{\rho}_{p+\delta p}}{\|\hat{\rho}_{p+\delta p}\|_2} - \frac{\hat{\rho}_p}{\|\hat{\rho}_p\|_2} \right\|_2^2 = \sum_{i,j} \delta p_i S_p^{i,j} \delta p_j, \quad (2.43)$$

where $S_p^{i,j}$ is the (local) metric tensor in the parameter space. By considering the first-order Taylor expansion of $\hat{\rho}_{p+\delta p}$ in terms of δp_i , the metric tensor can be written in terms of the log-derivatives of the ansatz :

$$S_p^{i,j} = \langle\langle D_{p_i} D_{p_j} \rangle\rangle - \langle\langle D_{p_i} \rangle\rangle \langle\langle D_{p_j} \rangle\rangle, \quad (2.44)$$

where $\langle\langle \cdot \rangle\rangle$ denotes the statistical expectation over the probability distribution $\Pi(x)$ introduced previously. The natural gradient method then defines the updating rule to be

$$p \leftarrow p - \eta S_p^{-1} \nabla \mathcal{C}, \quad (2.45)$$

where S_p^{-1} is the inverse metric tensor. Note that S_p may be non-invertible in practice. Therefore, to improve the numerical stability, we added a small diagonal shift before solving for the inverse, i.e.

$$S_p^{-1} = (S_p + \lambda \mathbb{1})^{-1}, \quad (2.46)$$

with $\lambda = 10^{-3}$ in our implementation, and we used a learning rate $\eta \in [0.005, 0.01]$, leading to the results in the next chapter.

2.4.5 Summary of the VMC training protocol

To conclude this section, we briefly summarise the key steps in the training protocol. For a given Lindbladian, the following steps are performed:

1. Randomly initialise the variational parameters $p \in \mathcal{R}$.
2. Repeat until convergence of the cost function \mathcal{C} :
 - (a) Evaluate the cost function \mathcal{C} by stochastic sampling.
 - (b) Evaluate the observables of interest by stochastic sampling.
 - (c) Evaluate the gradient $\nabla\mathcal{C}$ by stochastic sampling.
 - (d) Evaluate the metric tensor S_p by stochastic sampling, apply a small diagonal shift and calculate its inverse S_p^{-1} .
 - (e) Update the variational parameters according to the gradient. The method of our choice is natural gradient, i.e. the parameters are updated by $p \leftarrow p - \eta S_p^{-1} \nabla\mathcal{C}$.

In the procedure presented above, the stochastic sampling can be easily distributed into multiple processes, while there do not exist yet, to the best of our knowledge, efficient distributed solvers for dense linear systems, making the calculation of S_p^{-1} the bottleneck of our algorithm.

2.5 Conclusion

In this chapter, we have presented how machine learning can be applied to finding the NESS of an open quantum system, where an RBM is used as the ansatz for the density matrix, which is then optimised by VMC to obtain the NESS solution. We also proposed a symmetrised RBM structure adapted to translational-invariant lattice models. In the next chapter, we will see the results obtained from the implementation of the algorithms presented in this chapter.

Chapter 3

Artificial neural network results on driven-dissipative lattice models

With the ANN ansatz and the training protocol introduced in the last chapter in mind, we now present our results on the 1D and 2D transverse-field Ising model, the 1D driven-dissipative XY model and the 1D driven-dissipative Bose-Hubbard model. The largest number of spins that we have simulated is 18, which is higher compared to all the existing publications we have reviewed in Section 2.2. The results are benchmarked against the exact solutions obtained from brute force numerical integration of the master equation¹. We found a good overall performance of the ANN approach, while the accuracy is higher for regimes far from dissipative phase transitions.

3.1 Dissipative transverse-field Ising model

First, we test our method on the dissipative TFI model. As we have introduced in Section 1.2.3, the Hamiltonian can be written as

$$\hat{H} = \frac{V}{4} \sum_{\langle i,j \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z + \frac{g}{2} \sum_i \hat{\sigma}_i^x, \quad (3.1)$$

and we consider the jump operators to be

$$\hat{L}_i = \sqrt{\gamma} \hat{\sigma}_i^-, \quad (3.2)$$

the effect of which is to flip a spin towards the down direction along z .

¹In our implementation this is done with QuTiP [42].

We trained the symmetrised RBM for a 1D spin chain with $N = 5$ (number of sites), $V/\gamma = 2$, $\alpha = \beta = 1$ and the periodic boundary condition. In each iteration, 1000 samples are taken in the MCMC procedure. In Figure 3.1, we present the typical evolution of the cost function and the observables, which are the spatial spin components, with respect to the iteration steps, for a fixed value of $g/\gamma=1$. The cost function reaches a level between 10^{-3} and 10^{-2} after roughly 10^3 steps, where the observables also stabilise and oscillate around the exact values, indicating the convergence towards the global optimum of the ansatz. This allows us to estimate the number of steps required for the convergence of the algorithm.

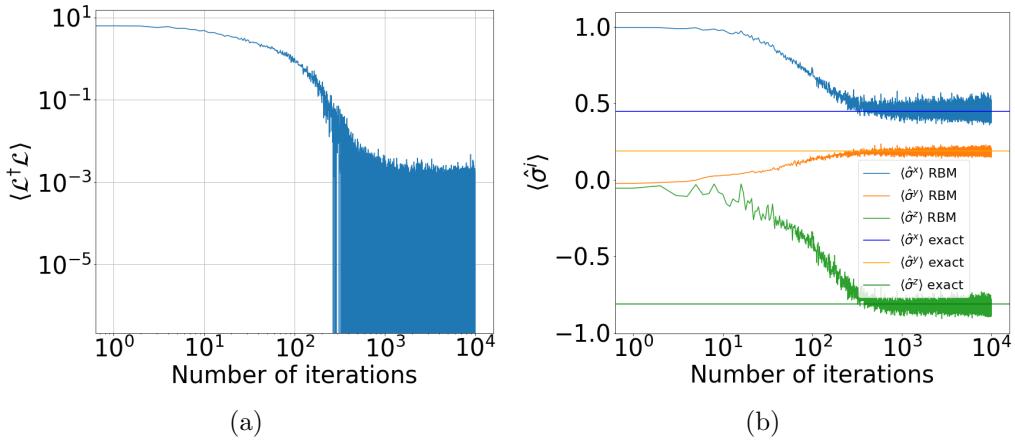


Figure 3.1: (a) The evolution of the cost function with respect to the iteration steps. (b) The corresponding evolution of the expectation value of the observables $\hat{\sigma}^x$, $\hat{\sigma}^y$ and $\hat{\sigma}^z$. The horizontal lines indicate the exact values. Results obtained for the 1D TFI model with $N = 5$, $V/\gamma = 2$, $g/\gamma = 1$, $\alpha = \beta = 1$ and the periodic boundary condition. The symmetrised RBM is used as the ansatz. In each iteration, 1000 samples are taken for the cost function and the observables respectively.

The numerical results of the NESS observables for different values of g/γ are reported in Figure 3.2(a), together with the corresponding values of the cost function in Figure 3.2(b). The machine learning result for each value of g is obtained after 10^4 iteration steps². Due to the oscillation observed in Figure 3.1, the final result is taken to be the average value over the last 100 iterations and the error bar is taken to be standard deviation. This convention is also applied to the results for the other models presented in this chapter. We can notice a good overall agreement between the machine learning results and the exact values for the entire range of the parameter

²Each data point takes between 20 and 30 minutes to produce using 24 processes on a cluster installed with 2-socket Intel Xeon 2.27GHz E5-2650 v4 processors, with 12-cores in each CPU socket.

g , where a dissipative phase transition from $\langle \hat{\sigma}^z \rangle = -1$ to $\langle \hat{\sigma}^z \rangle = 0$ is also captured by the ANN approach. For values of g/γ around 2, noticeable errors appeared in the observables (especially in $\langle \hat{\sigma}^x \rangle$), accompanied by a peak in the cost function. This can be understood by looking at the spectral gap³ of the super-operator $\mathcal{L}^\dagger \mathcal{L}$, as plotted in Figure 3.2(c) for the same range of g/γ . The region with higher errors corresponds to the smaller spectral gaps of $\mathcal{L}^\dagger \mathcal{L}$, which may make it harder for the ansatz to converge to the exact steady state. This comparison also shows that the value of the cost function can be used to measure how good our result is without knowing the exact solution beforehand: since we know that the global minimum is 0, the lower the cost function is, the closer we get to the NESS of the system.

We also tested our method on the 2D dissipative model for a 3×3 square lattice, using the standard RBM as the ansatz with $\alpha = \beta = 2$. In each iteration, 2000 samples are taken, and the other parameters are the same as above. The results for the NESS observables and the corresponding cost function values are reported in Figure 3.3.

We can observe patterns similar to the 1D case: the errors for values of g close to the phase transition are more significant than others. Otherwise, the observables showed good agreement with the exact values.

3.2 Driven-dissipative XY model

Next, we consider another open spin model, which is the driven-dissipative XY model, whose Hamiltonian is given in section 1.2.2:

$$\hat{H} = -J \sum_{\langle i,j \rangle} \hat{\sigma}_i^+ \hat{\sigma}_j^- + \sum_i \Delta \hat{\sigma}_i^+ \hat{\sigma}_i^- + \sum_i F(\hat{\sigma}_i^+ + \hat{\sigma}_i^-), \quad (3.3)$$

with the same jump operators as for the TFI model:

$$\hat{L}_i = \sqrt{\gamma} \hat{\sigma}_i^-, \quad (3.4)$$

which fully determine the Lindbladian of the system. We also recall the definition for the two-site correlation function defined as

$$C(1) = \frac{\langle \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z \rangle}{\langle \hat{\sigma}_j^z \rangle \langle \hat{\sigma}_{j+1}^z \rangle}, \quad (3.5)$$

where j is fixed to be the index of the middle site.

We trained the symmetrised RBM for a 1D spin chain with $N = 8$, $J/\gamma = 1$, $F/\gamma = 1$ and the periodic boundary condition, for different values of α , β and Δ . 1000 samples are taken in each iteration and 10^4 iterations

³Calculated in MATLAB [43].

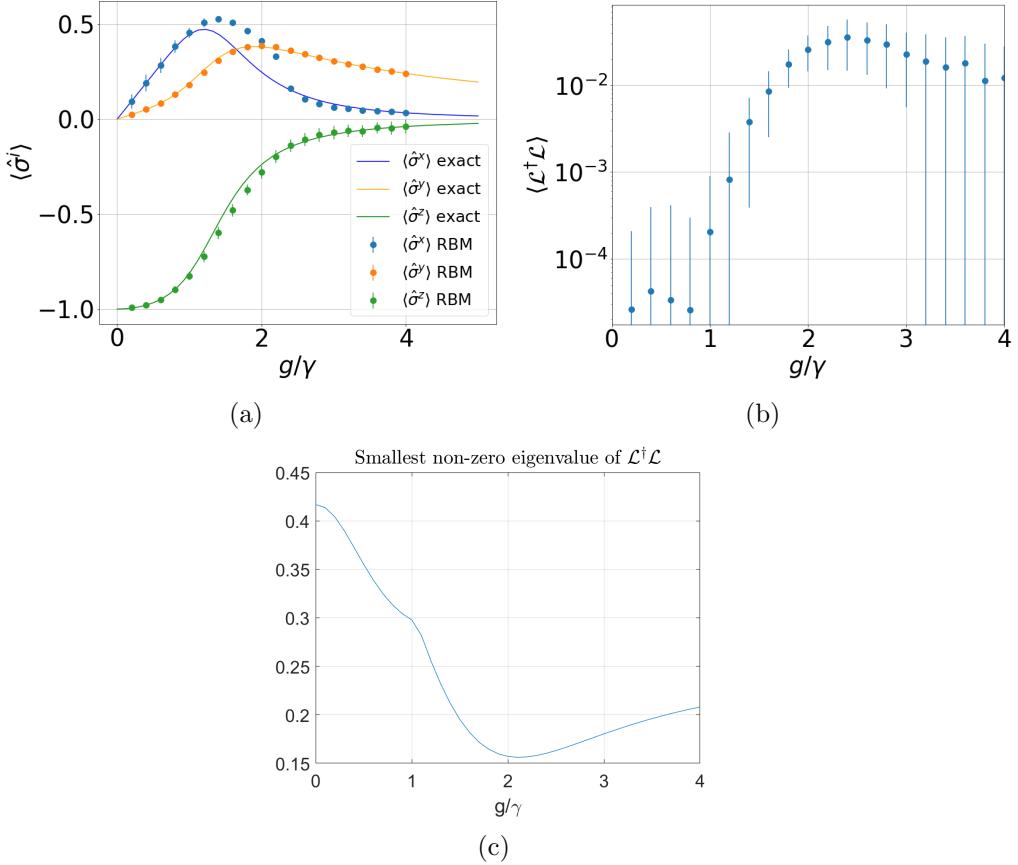


Figure 3.2: (a) Steady state observables learned by the ANN compared to the exact values for different values of g . (b) Cost function after optimisation for different values of g . (c) The spectral gap of the super-operator $\mathcal{L}^\dagger \mathcal{L}$ as a function of g , obtained from exact diagonalisation. Same parameters as in Figure 3.1. For each data point in (a) and (b) 10^4 iteration steps are performed and the mean value over the last 100 iterations is taken to be the final result. The error bars are taken to be the standard deviation over the last 100 iterations.

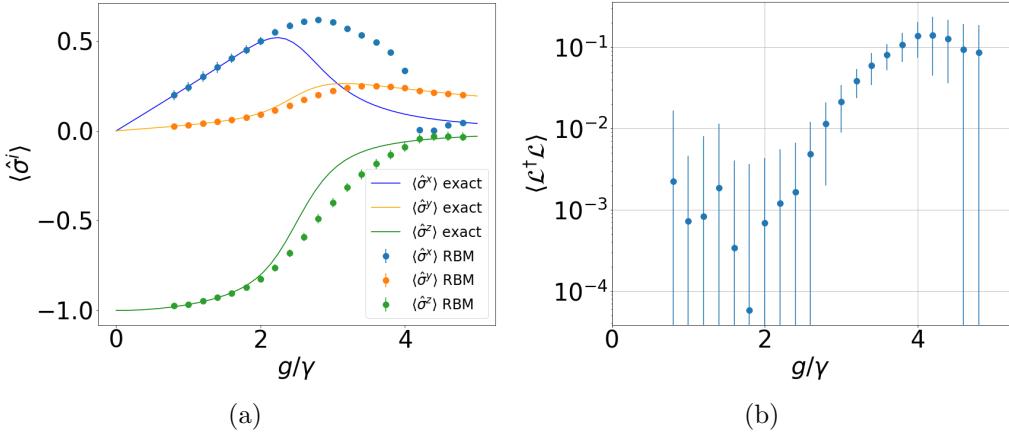


Figure 3.3: (a) Steady state observables learned by the ANN and the exact values for different values of g . (b) The corresponding cost function values. The results are obtained for the 2D TFI model on a 3×3 square lattice with $V/\gamma = 2$, $g/\gamma = 1$, $\alpha = \beta = 2$ and the periodic boundary condition. The standard RBM is used as the ansatz. In each iteration, 2000 samples are taken and for each data point 10^4 iterations are performed. Same convention for taking the average values and the error bars as in Figure 3.2(a).

are performed for each value of Δ . The numerical results obtained from machine learning for the two-site correlation $C(1)$ and the expectation value $\langle \hat{\sigma}^z \rangle$ are presented in Figure 3.4(a) and 3.4(b) respectively, together with the cost function values in Figure 3.4(c). In the first run, we used $\alpha = \beta = 1$ for all values of Δ , and noticed significant errors in both $C(1)$ and $\langle \hat{\sigma}^z \rangle$ and significantly high values of the cost function for a certain range of Δ . We then increased the hidden unit density and the ancillary unit density for these values of Δ and repeated this process until all the results lie on the line of the exact solution. Figure 3.4 gives us an intuitive understanding of how increasing α and β improves the representing power of the ANN: the curves traced out by the machine learning results get closer and closer to the exact solution as α and β are increased. As a result, we found that $\alpha = \beta = 8$ was sufficient to make all points roughly converge to the correct value. Note that the bunching-antibunching transition is captured by the ANN solution. Even in the worst case of $\alpha = \beta = 1$, the signature of the transition is qualitatively displayed.

Since we have introduced two structures of RBM in the last chapter, the standard one and the symmetrised one, we would like to compare their performances as well as the error scaling with respect to the hidden and ancillary unit densities and the system size N . The optimised cost function as a function of α, β and N is shown in Figure 3.5.

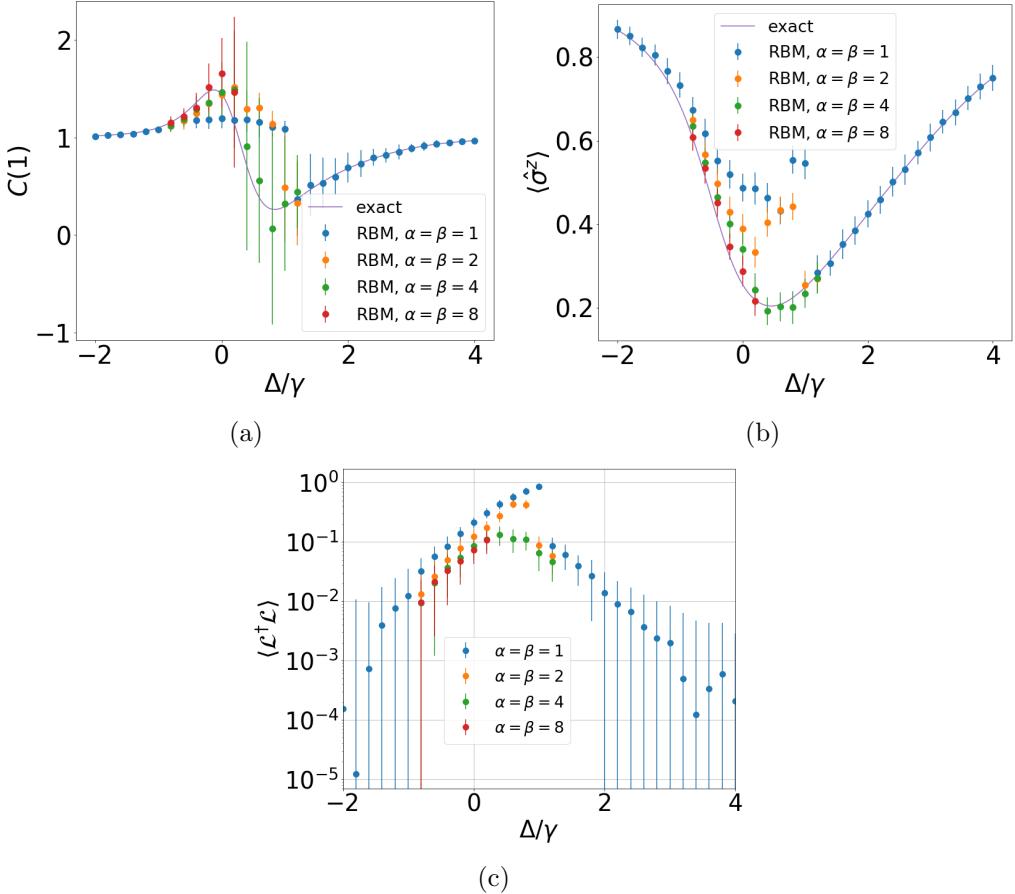


Figure 3.4: (a) The two-site correlation function $C(1)$ learned by the ANN for different values of Δ , α and β compared to the exact value. (b) The corresponding expectation values of $\hat{\sigma}^z$. (c) The corresponding values of the cost function. Results obtained for the 1D XY model with $N = 8$, $J/\gamma = 1$, $F/\gamma = 1$ and the periodic boundary condition. The symmetrised RBM is used as the ansatz. 1000 samples are taken in each iteration and 10^4 iterations are performed for each data point. Same convention for taking the average values and the error bars as in Figure 3.2(a).

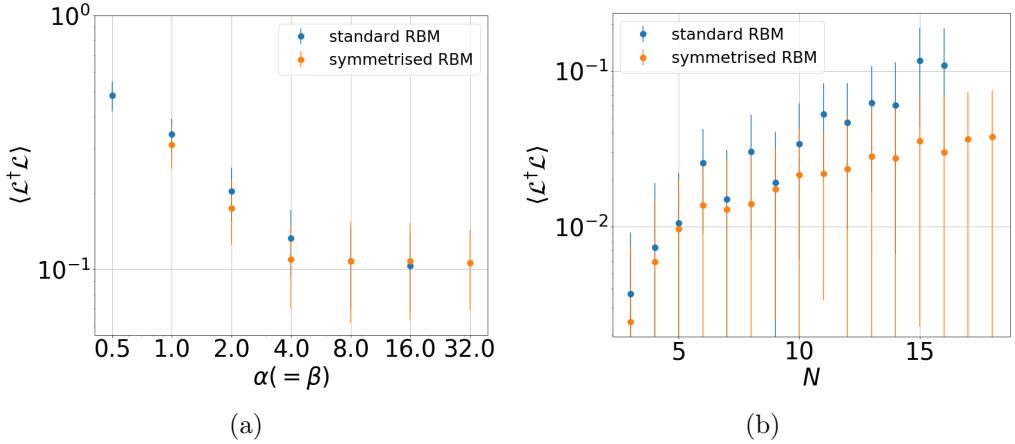


Figure 3.5: (a) Cost function after optimisation for different values of α and β ($\alpha, \beta \in [0.5, 16]$ For the standard RBM and $\alpha, \beta \in [1, 32]$ for the symmetrised RBM), with a fixed value of $\Delta/\gamma = 0.2$ and $N = 8$. (b) Cost function after optimisation for different values of N ($N \in [3, 16]$ for the standard RBM and $N \in [3, 18]$ for the symmetrised RBM), with a fixed value of $\Delta/\gamma = 2$ and $\alpha = \beta = 1$. The other parameters are the same as in Figure 3.4(a). Both the standard and the symmetrised RBMs are tested.

Figure 3.5(a) shows that the error (as measured by the value of the cost function) for both two structures can be reduced by increasing α and β . For $\alpha, \beta \leq 4$, the symmetrised RBM has smaller error. However, for larger values of α and β errors for both structures saturate at around 10^{-1} . Considering that the standard RBM has quadratically more parameters than the symmetrised one, this graph shows that the symmetrised RBM has a better overall performance. A similar statement can be made from Figure 3.5(b), where the error for the symmetrised RBM always stays lower than the standard one, for the entire range of N tested. The points corresponding to the symmetrised RBM also suggest a better error scaling as they show a smaller slope.

3.3 Driven-dissipative Bose-Hubbard model

Finally, we applied our method to the driven-dissipative BH model. The Hamiltonian is given in Section 1.2.2:

$$\hat{H} = -J \sum_{\langle i,j \rangle} \hat{b}_i^\dagger \hat{b}_j + \sum_i \Delta \hat{b}_i^\dagger \hat{b}_i + \sum_i \frac{U}{2} \hat{b}_i^\dagger \hat{b}_i^\dagger \hat{b}_i \hat{b}_i + \sum_i F(\hat{b}_i^\dagger + \hat{b}_i), \quad (3.6)$$

and we consider the jump operators

$$\hat{L}_i = \sqrt{\gamma} \hat{b}_i \quad (3.7)$$

having an effect of reducing the number of bosons in the lattice.

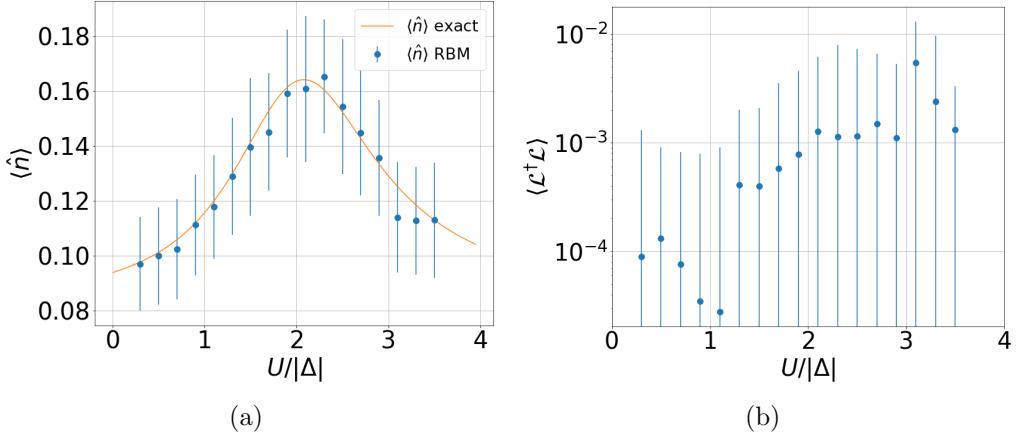


Figure 3.6: (a) Steady state observable \hat{n} learned by the ANN and the exact values for different values of U . (b) The corresponding values of the cost function. The results are obtained for the 1D BH model with $N = 5$, $N_{\max} = 2$ (maximum number of bosons allowed per site), $J/|\Delta| = 0.1$, $F/|\Delta| = 0.4$, $\alpha = \beta = 1$ and the periodic boundary condition. The symmetrised RBM is used. 1000 samples are taken in each iteration and 10^4 iterations are performed for each data point. Same convention for taking the average values and the error bars as in Figure 3.2(a).

We used the symmetrised RBM as the ansatz and trained it for a 1D lattice with $N = 5$, $N_{\max} = 2$ (maximum number of bosons allowed per site), $J/|\Delta| = 0.1$, $F/|\Delta| = 0.4$, $\alpha = \beta = 1$ and the periodic boundary condition, for different values of U . 1000 samples are taken in each iteration and a total of 10^4 iterations are performed for each value of U . The results for the NESS mean boson density $\langle \hat{n} \rangle = \langle \hat{b}^\dagger \hat{b} \rangle$ and the corresponding cost function values are presented in Figure 3.6.

In Figure 3.6(a), all points from the machine learning result showed good agreement with the exact values. In particular, the peak around $U/|\Delta| = 2$ is captured by the ANN solution. Physically, this corresponds to the two-photon resonance defined in Equation (1.41), and our case corresponds to $n = 2$.

3.4 Discussion

In the results presented above, we have examined different aspects of the performance of the ANN approach by applying it to various models and by benchmarking against exact solutions of the models. One of the interesting points in the results is the relation between the error, the spectral gap of $\mathcal{L}^\dagger \mathcal{L}$ and the dissipative phase transition.

3.4.1 Spectral gap of $\mathcal{L}^\dagger \mathcal{L}$

As we have discussed in the results for the 1D TFI model, the quality of the solution is closely related to the spectral gap of the super-operator $\mathcal{L}^\dagger \mathcal{L}$, where a smaller gap makes the algorithm harder to converge, which corresponds to a dissipative phase transition in the models we investigated. According to the theory presented in [44], when a dissipative phase transition takes place, the spectral gap of the Lindbladian (defined to be the modulus of the real part of the non-zero eigenvalue of \mathcal{L} having the smallest modulus of the real part) closes, and the imaginary part of the corresponding eigen value of \mathcal{L} should also be 0. This implies that the gap should also close for $\mathcal{L}^\dagger \mathcal{L}$. Physically, this means that there exists a state that is close to the steady state (decaying slowly with time), making it more difficult for algorithm to converge to the exact steady state, and that the final solution we obtain may depend on the initial state, corresponding to the bistability regime that we often encounter in the mean field theory [11, 12].

To end this discussion, we can make the following general remarks:

- The cost function can be used as a measure of the quality of the machine learning result. The smaller it is, the better the outcome.
- The RBM is capable of accurately representing the NESS of open systems in the regimes where the spectral gap of $\mathcal{L}^\dagger \mathcal{L}$ is large, which usually corresponds to a regime far from the phase transition. In most cases, the RBM is also capable of capturing the dissipative phase transition, or at least qualitatively.
- The accuracy of the result can be improved by using more hidden units and ancillary units. However, the error may saturate at a certain point, which may be due to the limitation of the representing power of the ansatz.
- The errors for both standard and symmetrised RBMs increase with the system size, while the symmetrised RBM has a better overall performance, lower complexity and better error scaling.

3.5 Conclusion

In this chapter, we have applied the ANN approach of finding the NESS to various toy models of open quantum systems, and the results agree well with the exact solutions in general. The RBM managed to capture the dissipative phases in these models. The method we have implemented is generic and can be scaled up to larger systems, allowing us to probe the rich physics in quantum many-body problems.

Conclusion and outlook

The remarkable advances in computer technologies and algorithm designs have opened up endless new possibilities for physicist to learn about the nature of matter, enabling us to simulate real-life material with higher accuracy. This thesis is written on the exciting interface between machine learning and physical sciences, exploring the artificial neural network as an ansatz when resolving phases of matter at the steady state in driven dissipative many-body quantum systems. As we have seen in the first chapter, finding the NESS involves solving the Lindblad master equation, which is a hard task in general, due to the exponentially growing size of the Hilbert space.

In the second chapter, we presented a newly emerging method for finding the NESS for open systems, where the RBM is used to construct the density matrix ansatz. After reviewing the works that proposed and tested this idea, we propose a novel RBM structure adapted to systems with translational symmetry and a combination of the ansatz and training method that are novel to the best of our knowledge by the completion of this work.

In the last chapter, we implemented the ansatz and the training algorithm we have presented and obtained fruitful results. The ANN approach has a overall good performance when benchmarked on the models we considered, and the symmetrised RBM that we proposed exhibits better performance than the standard one. In addition, the applications of this approach to the 2D dissipative transverse-field Ising model, the 1D driven-dissipative *XY* model and the 1D driven-dissipative Bose-Hubbard model are also novel to the best of our knowledge.

In this thesis, we pave the way to a promising avenue of research with exciting challenges. The method we presented here has the potential to be generalised to more classes of systems, for example, fermionic systems, hybrid systems, etc. Moreover, as the algorithm applies not only to the RBM but also to any general ansatz for which we can calculate the gradient, more ANN structures can be explored, such as deep neural networks with different types of connections between nodes. Having witnessed and experimented with the successful cross-fertilisation between machine learning and many-body physics, this direction of research is expected to sparkle more interest and success in the future.

Bibliography

- [1] S Sachdev. “Quantum Phase Transitions (Cambridge University Press, Cambridge, 1999).” In: () .
- [2] Erich Strohmaier et al. *TOP500 list*. 2015.
- [3] David P Landau and Kurt Binder. *A guide to Monte Carlo simulations in statistical physics*. Cambridge university press, 2014.
- [4] Jos Thijssen. *Computational physics*. Cambridge university press, 2007.
- [5] Ulrich Schollwöck. “The density-matrix renormalization group in the age of matrix product states”. In: *Annals of Physics* 326.1 (2011), pp. 96–192.
- [6] Giuseppe Carleo and Matthias Troyer. “Solving the quantum many-body problem with artificial neural networks”. In: *Science* 355.6325 (2017), pp. 602–606.
- [7] Subir Sachdev. “Quantum phase transitions”. In: *Handbook of Magnetism and Advanced Magnetic Materials* (2007).
- [8] Eric M Kessler et al. “Dissipative phase transition in a central spin system”. In: *Physical Review A* 86.1 (2012), p. 012116.
- [9] Michael Albiez et al. “Direct observation of tunneling and nonlinear self-trapping in a single bosonic Josephson junction”. In: *Physical review letters* 95.1 (2005), p. 010402.
- [10] Changsuk Noh and Dimitris G Angelakis. “Quantum simulations and many-body physics with light”. In: *Reports on Progress in Physics* 80.1 (2016), p. 016401.
- [11] JJ Mendoza-Arenas et al. “Beyond mean-field bistability in driven-dissipative lattices: Bunching-antibunching transition and quantum simulation”. In: *Physical Review A* 93.2 (2016), p. 023821.
- [12] Jiasen Jin et al. “Phase diagram of the dissipative quantum Ising model on a square lattice”. In: *Physical Review B* 98.24 (2018), p. 241108.

- [13] Alexandre Le Boité, Giuliano Orso, and Cristiano Ciuti. “Bose-hubbard model: Relation between driven-dissipative steady states and equilibrium quantum phases”. In: *Physical Review A* 90.6 (2014), p. 063821.
- [14] Iacopo Carusotto et al. “Fermionized photons in an array of driven dissipative nonlinear cavities”. In: *Physical review letters* 103.3 (2009), p. 033601.
- [15] Michael J Hartmann. “Polariton crystallization in driven arrays of lossy nonlinear resonators”. In: *Physical review letters* 104.11 (2010), p. 113601.
- [16] Jiasen Jin et al. “Photon solid phases in driven arrays of nonlinearly coupled cavities”. In: *Physical review letters* 110.16 (2013), p. 163605.
- [17] Mattias Fitzpatrick et al. “Observation of a dissipative phase transition in a one-dimensional circuit QED lattice”. In: *Physical Review X* 7.1 (2017), p. 011016.
- [18] Ruichao Ma et al. “A dissipatively stabilized Mott insulator of photons”. In: *Nature* 566.7742 (2019), p. 51.
- [19] Eduardo Mascarenhas, Hugo Flayac, and Vincenzo Savona. “Matrix-product-operator approach to the nonequilibrium steady state of driven-dissipative quantum arrays”. In: *Physical Review A* 92.2 (2015), p. 022116.
- [20] Klaus Mølmer, Yvan Castin, and Jean Dalibard. “Monte Carlo wavefunction method in quantum optics”. In: *JOSA B* 10.3 (1993), pp. 524–538.
- [21] S Finazzi et al. “Corner-space renormalization method for driven-dissipative two-dimensional correlated systems”. In: *Physical review letters* 115.8 (2015), p. 080604.
- [22] T. Mitchell. *Machine Learning*. McGraw Hill, 1997.
- [23] Stuart J Russell and Peter Norvig. *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited, 2016.
- [24] Ethem Alpaydin. *Introduction to machine learning*. MIT press, 2009.
- [25] Paul Smolensky. *Information processing in dynamical systems: Foundations of harmony theory*. Tech. rep. Colorado Univ at Boulder Dept of Computer Science, 1986.
- [26] Geoffrey E Hinton and Ruslan R Salakhutdinov. “Reducing the dimensionality of data with neural networks”. In: *science* 313.5786 (2006), pp. 504–507.
- [27] Hugo Larochelle and Yoshua Bengio. “Classification using discriminative restricted Boltzmann machines”. In: *Proceedings of the 25th international conference on Machine learning*. ACM. 2008, pp. 536–543.

- [28] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. “Neural Networks Take on Open Quantum Systems”. In: *Physics* 12 (2019), p. 74.
- [29] Nobuyuki Yoshioka and Ryusuke Hamazaki. “Constructing neural stationary states for open quantum many-body systems”. In: *arXiv preprint arXiv:1902.07006* (2019).
- [30] Filippo Vicentini et al. “Variational neural network ansatz for steady states in open quantum systems”. In: *arXiv preprint arXiv:1902.10104* (2019).
- [31] Alexandra Nagy and Vincenzo Savona. “Variational quantum Monte Carlo method with a neural-network ansatz for open quantum systems”. In: *Physical Review Letters* 122.25 (2019), p. 250501.
- [32] Michael J Hartmann and Giuseppe Carleo. “Neural-Network Approach to Dissipative Quantum Many-Body Dynamics”. In: *arXiv preprint arXiv:1902.05131* (2019).
- [33] Giuseppe Carleo et al. “Machine learning and the physical sciences”. In: *arXiv preprint arXiv:1903.10563* (2019).
- [34] Nobuyuki Yoshioka, Yutaka Akagi, and Hoshio Katsura. “Learning disordered topological phases by statistical recovery of symmetry”. In: *Physical Review B* 97.20 (2018), p. 205110.
- [35] Mohamad H Hassoun et al. *Fundamentals of artificial neural networks*. MIT press, 1995.
- [36] Sandro Sorella. “Wave function optimization in the variational Monte Carlo method”. In: *Physical Review B* 71.24 (2005), p. 241103.
- [37] Jian Cui, J Ignacio Cirac, and Mari Carmen Banuls. “Variational matrix product operators for the steady state of dissipative quantum systems”. In: *Physical review letters* 114.22 (2015), p. 220601.
- [38] Nicholas Metropolis et al. “Equation of state calculations by fast computing machines”. In: *The journal of chemical physics* 21.6 (1953), pp. 1087–1092.
- [39] Léon Bottou. “Large-scale machine learning with stochastic gradient descent”. In: *Proceedings of COMPSTAT’2010*. Springer, 2010, pp. 177–186.
- [40] Shun-Ichi Amari. “Natural gradient works efficiently in learning”. In: *Neural computation* 10.2 (1998), pp. 251–276.
- [41] Sandro Sorella, Michele Casula, and Dario Rocca. “Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods”. In: *The Journal of chemical physics* 127.1 (2007), p. 014105.

- [42] J Robert Johansson, Paul D Nation, and Franco Nori. “QuTiP 2: A Python framework for the dynamics of open quantum systems”. In: *Computer Physics Communications* 184.4 (2013), pp. 1234–1240.
- [43] MATLAB. *version 9.4.0 (R2018a)*. Natick, Massachusetts: The MathWorks Inc., 2018.
- [44] Fabrizio Minganti et al. “Spectral theory of Liouvillians for dissipative phase transitions”. In: *Physical Review A* 98.4 (2018), p. 042118.

Appendix A

Log-derivatives of the RBM density matrix ansatz

We explicitly present the expressions for the log-derivatives of the RBM density matrix ansatz below:

$$\frac{\partial \ln \rho_{\mathcal{W}}}{\partial \text{Re}[a_j]} = l_j + r_j, \quad (\text{A.1})$$

$$\frac{\partial \ln \rho_{\mathcal{W}}}{\partial \text{Im}[a_j]} = i(l_j - r_j), \quad (\text{A.2})$$

$$\frac{\partial \ln \rho_{\mathcal{W}}}{\partial \text{Re}[b_k]} = \tanh \theta_k^l + \tanh \theta_k^r, \quad (\text{A.3})$$

$$\frac{\partial \ln \rho_{\mathcal{W}}}{\partial \text{Im}[b_k]} = i(\tanh \theta_k^l - \tanh \theta_k^r), \quad (\text{A.4})$$

$$\frac{\partial \ln \rho_{\mathcal{W}}}{\partial \text{Re}[c_p]} = 2 \tanh \theta_p^m, \quad (\text{A.5})$$

$$\frac{\partial \ln \rho_{\mathcal{W}}}{\partial \text{Re}[W_{kj}]} = l_j \tanh \theta_k^l + r_j \tanh \theta_j^r, \quad (\text{A.6})$$

$$\frac{\partial \ln \rho_{\mathcal{W}}}{\partial \text{Im}[W_{kj}]} = i(l_j \tanh \theta_k^l - r_j \tanh \theta_j^r), \quad (\text{A.7})$$

$$\frac{\partial \ln \rho_{\mathcal{W}}}{\partial \text{Re}[U_{pj}]} = (l_j + r_j) \tanh \theta_p^m, \quad (\text{A.8})$$

$$\frac{\partial \ln \rho_{\mathcal{W}}}{\partial \text{Im}[U_{pj}]} = i(l_j - r_j) \tanh \theta_p^m, \quad (\text{A.9})$$

where we have defined

$$\theta_k^l = b_k + \sum_{j=1}^N W_{kj} l_j, \quad (\text{A.10})$$

$$\theta_k^r = b_k^* + \sum_{j=1}^N W_{kj}^* r_j, \quad (\text{A.11})$$

$$\theta_p^m = c_p + c_p^* + \sum_{j=1}^N (U_{pj} l_j + U_{pj}^* r_j). \quad (\text{A.12})$$