

Федеральное государственное автономное образовательное учреждение
высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»
Физтех-школа физики и исследований им. Ландау
Образовательная программа «Фундаментальные проблемы физики квантовых
технологий»

Направление подготовки / специальность: 03.04.01 Прикладные математика и физика
(магистратура)

Направленность (профиль) подготовки: Квантовая оптика и лазерная физика

РАСПОЗНАВАНИЕ ФАЗОВЫХ ПЕРЕХОДОВ В КВАНТОВЫХ ЦЕПОЧКАХ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

(магистерская диссертация)

Студент:

Березутский Александр Владимирович



(подпись студента)

Научный руководитель:

Филиппов Сергей Николаевич,
канд. физ.-мат. наук, доц.

(подпись научного руководителя)

Консультант (при наличии):

(подпись консультанта)

Москва 2020



MOSCOW INSTITUTE OF PHYSICS AND TECHNOLOGY
MASTER'S THESIS

**Probing Phase Transitions in Quantum Spin Chains
With Neural Networks**

Landau Phystech-School of Physics and Research
Educational Program: "Fundamental Problems of Physics of Quantum
Technologies"

Student: _____
Aleksandr V. Berezutskii

Research Advisor: _____
Sergey N. Filippov
Associate Professor

Moscow 2020

All rights reserved.©

The author hereby grants to Moscow Institute of Physics and Technology permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole and in part in any medium now known or hereafter created.

Probing Phase Transitions in Quantum Spin Chains With Neural Networks

Aleksandr V. Berezutskii

Submitted to Moscow Institute of Physics and Technology

Abstract

The numerical emulation of quantum systems often requires an exponential number of degrees of freedom which translates to a computational bottleneck. Methods of machine learning have been used in adjacent fields for effective feature extraction and dimensionality reduction of high-dimensional datasets. Recent studies have revealed that neural networks are further suitable for the determination of macroscopic phases of matter and associated phase transitions as well as efficient quantum state representation. In this work, we address quantum phase transitions in quantum spin chains, namely the transverse field Ising chain (both with and without the longitudinal field) and the anisotropic XY chain, and show that even neural networks with no hidden layers can be effectively trained to distinguish between magnetically ordered and disordered phases. Our neural network acts to predict the corresponding crossovers finite-size systems undergo.

Research Advisor:
Name: Sergey N. Filippov
Degree: Ph.D.
Title: Associate Professor

Contents

1	Introduction	8
1.1	Quantum phase transitions	8
1.2	Motivation for using neural networks in condensed matter physics and quantum information theory	12
1.3	Purpose statement	12
1.4	Problem statement	13
2	Quantum spin models: theoretical background	14
2.1	Transverse field Ising model	14
2.2	Transverse field Ising model in the longitudinal field	16
2.3	Anisotropic XY model	16
3	Methods for finding ground states of quantum spin Hamiltonians	18
3.1	Analytical approaches	18
3.2	Numerical approaches	19
3.3	Machine learning approaches	20
3.4	Methods for detecting quantum phase transitions	20
4	Classification with neural networks	23
4.1	Architecture	23
4.2	Activation function	24
4.3	Computation graph	24
4.4	Our model	25
4.5	Training	26
5	The setup	28
5.1	Sampling	28
5.2	Methodology	32

6	Results and discussion	34
6.1	Transverse field Ising model	34
6.2	Tilted field Ising model	41
6.3	Anisotropic XY model	44
7	Summary and conclusion	46
A	Tables	47

List of Figures

1.1	Low eigenvalues of a Hamiltonian $H(\tau)$ on a finite lattice, as a function of some dimensionless coupling τ in a generic case.	9
2.1	One-dimensional chain of $L = 10$ sites.	14
2.2	Zero-temperature phase diagram of the transverse Ising model.	15
4.1	A common neural network architecture with 1 hidden layer.	23
4.2	Nonlinear activation functions typically used for machine learning problems.	24
4.3	Neural network architecture used in the present study.	25
5.1	Visualization of matrix elements of the TFIM Hamiltonian of 8 spins at $\tau = 1$	29
5.2	Visualization of the coin toss sampling.	30
5.3	Distribution of the basis bitstring configurations for the TFIM Hamiltonian on $L = 7$ spins at $\tau = 1$	31
5.4	Distribution of the basis bitstring configurations for the TFIM Hamiltonian on $L = 7$ spins at $\tau = 1$ and $N = 10000$ configurations sampled by our method.	31
5.5	Distribution of the basis bitstring configurations for the TFIM Hamiltonian on $L = 15$ spins at $\tau = 1$ and $N = 10000$ configurations sampled by our method.	32
5.6	Visualization of the setup.	33
6.1	Energy gap Δ against the transverse field τ for various lengths of the chain.	34
6.2	Energy gap Δ and the linear fitting function against the transverse field τ for $L = 20$ exhibiting the QPT point at $\tau_c = 0.94$	35
6.3	Transverse magnetization m_x against the transverse field τ for various lengths of the chain.	36
6.4	Transverse magnetic susceptibility χ_x against the transverse field τ for various lengths of the chain.	36
6.5	Neural network output c plotted against the transverse field strength τ for a chain of 20 spins.	37

6.6	Visualized weights \mathbf{W} for every spin in a chain of 20 spins against the transverse field τ	38
6.7	Critical points obtained by the neural-network method and by the explicit order parameter for finite-size chains.	39
6.8	Critical points obtained by the neural-network method and by the explicit order parameter for finite-size chains.	39
6.9	Relative entropy D between the ground state at τ_j and τ_{j+1} against the transverse field τ for various lengths of the chain.	41
6.10	Crossover widths obtained by the neural-network method, magnetic susceptibility and relative entropy for finite-size chains.	41
6.11	Energy gap Δ for the TFIM in longitudinal field on a chain of $L = 10$ spins versus τ and λ	42
6.12	The transverse magnetization m_x as a function of τ and λ for a chain of $L = 10$	43
6.13	An example of a typical neural network output c for $L = 12$ and $\tau = 0.5$ versus longitudinal magnetic field λ	44
6.14	The phase diagram of the Ising model in external longitudinal magnetic field obtained by the neural-network method for $L = 12$	44
6.15	The neural network output as a function of the anisotropy parameter γ for $L = 20$ spins on an anisotropic XY chain with open boundary conditions.	45

List of Tables

A.1 Variables and definitions	48
---	----

Chapter 1

Introduction

This chapter aims to give an introduction into the physics of quantum phase transitions. Besides that, we introduce some definitions and concepts used throughout this work.

1.1 Quantum phase transitions

Physicists call a phase transition between quantum phases of matter, i.e., phases of matter at zero temperature, a quantum phase transition (QPT). Quantum phase transitions are accessed by varying an external physical parameter. The latter can be, e.g., an external magnetic field or pressure. A quantum phase transition is characterized by a sharp change in the ground state of a system. The reason for this is the presence of quantum fluctuations. Despite that the absolute zero temperature cannot be reached physically, this regime is realizable within the following conditions: the typical quantum fluctuations' energy $\hbar\omega$ (being the order of the energy difference between the vacuum and the first excited state, i.e., the lowest eigenvalue and the next in the ascending order) should be much bigger than the system's thermal fluctuations $k_B T$ [1]. For convenience, hereafter we consider $\hbar = k_B = 1$.

Consider a Hamiltonian, $H(\tau)$, which describes quantum spins positioned on the sites of a lattice and which is controlled by a parameter τ . We are interested in the dependence of the physical observables (e.g., the ground state energy) with respect to τ . If the lattice is finite, then, in general, the ground state energy will be an analytic function of τ [1]. In addition, in most cases, the system does not exhibit level crossing (see Figure 1.1). However, the energy gap, i.e., the energy difference between the first excited state and the ground state, could become non-analytic at τ_c in the infinite-size (thermodynamic) limit [1]. Also, in this limit, some other observables, such as the magnetic susceptibility, can show non-analytic behavior

at the quantum phase transition point.

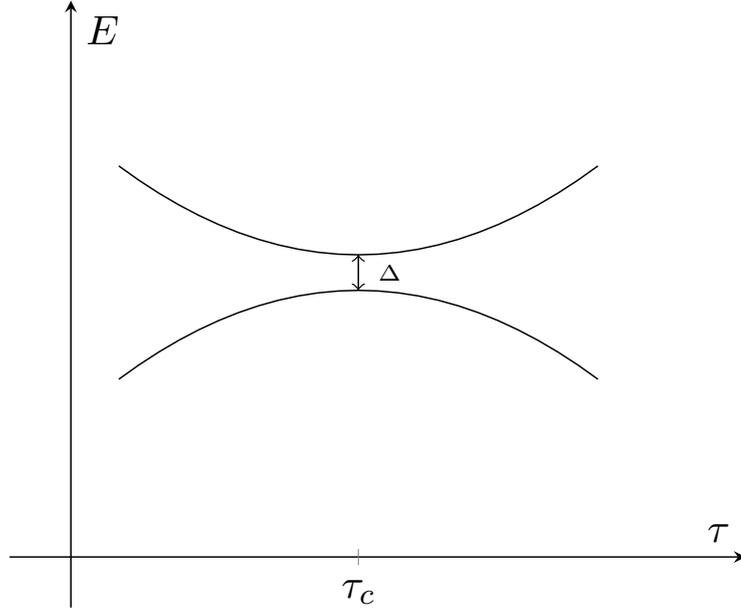


Figure 1.1: Low eigenvalues of a Hamiltonian $H(\tau)$ on a finite lattice, as a function of some dimensionless coupling τ in a generic case.

Let us call Δ the energy gap. In most cases, it is known that as τ gets closer to τ_c , Δ vanishes obeying the power law [1],

$$\Delta \sim |\tau - \tau_c|^{z\nu}, \quad (1.1)$$

where z and ν are the critical exponents related to the critical point τ_c and defined as follows. In a phase transition, one can define a length scale ξ , an important quantity in the field of quantum phase transitions, which determines the exponential decay of the spatial two-point correlation function in the ground state at a fixed point in time. For example, the correlation function $G(r)$ of an observable O between two points separated by distance r can be written as

$$G(r) = \langle O(0)O(r) \rangle - \langle O(0) \rangle \langle O(r) \rangle \propto e^{-r/\xi}, \quad (1.2)$$

where averaging is performed over the quantum state, that is $\langle O \rangle = \langle \psi | O | \psi \rangle$. As the critical point τ_c is being approached, ξ diverges with a critical exponent ν as

$$\xi \propto |\tau - \tau_c|^{-\nu}. \quad (1.3)$$

At the critical point, $\xi \rightarrow \infty$ and the correlation function $G(r)$ decays as a power law. Similarly, one can define a two-point temporal correlation function for a fixed point in space and the time scale ξ_t for the decay of the equal-space correlations, which diverges as

$$\xi_t \propto \xi^z \propto |g - g_c|^{-\nu z} \propto \Delta^{-1}. \quad (1.4)$$

The values of critical exponents are often universal, i.e., they do not depend on most of the microscopic details of the Hamiltonian $H(\tau)$, such as microscopic exchange constants, for example. The behavior mentioned above holds both for $\tau < \tau_c$ and $\tau > \tau_c$ with the same value of the exponent $z\nu$, but with different non-universal constants of proportionality. z is usually called the dynamical critical exponent. An illustrative example of a system, for which the critical exponents are known, is the well-known two-dimensional ferromagnetic Ising model, which has $\gamma = 7/4$ and $z \approx 2.2$ [2].

Magnetic properties are also important to discuss when talking about QPTs. Hereafter, we will be considering spin- $\frac{1}{2}$ systems, we set $\mu_B = 1$. Quantum magnetic properties and/or magnetic ordering usually originates from the distribution of unpaired spins in the system. For future analysis, here we define the average magnetization in the α -direction

$$m_\alpha = \frac{1}{L} \left\langle \sum_{i=1}^L \sigma_i^\alpha \right\rangle, \quad (1.5)$$

where σ^α is a corresponding Pauli matrix, ($\alpha = x, y, z$) and averaging is performed over a quantum state which implicitly depends on the control parameter τ . We omit the tensor products \otimes between the Pauli matrices for brevity. Thus, in our notation, σ_i^x stands for the following expression, where I denotes the 2×2 identity matrix:

$$\sigma_i^x = \underbrace{I \otimes I \cdots \otimes I}_{i-1} \otimes \sigma^x \otimes \underbrace{1 \otimes I \cdots \otimes I}_{L-i} \equiv I^{\otimes(i-1)} \otimes \sigma^x \otimes I^{\otimes(L-i)}. \quad (1.6)$$

Next, we define the average magnetic susceptibility in the α -direction:

$$\chi_\alpha = \frac{\partial}{\partial \tau} m_\alpha, \quad (1.7)$$

where $\alpha = x, y, z$ and averaging is again performed over a quantum state.

Let us now briefly itemize some experimental systems, where quantum criticality has been

observed. We provide some examples to show that quantum phase transitions arise in numerous systems, ranging from wires interacting with quantum dots to 3-dimensional materials. This serves as a good enough reason why it is cardinal to give insights into such systems in both analytical and numerical ways. Here we present a consistent albeit not exhaustive review of experimental results.

- Heavy fermion materials [3]: in this paper by Lohneysen et al. from 1994, it was shown that non-Fermi-liquid behaviour of $\text{CeCu}_{6-x}\text{Au}_x$ occurs at low temperatures which was evidenced by sharp maxima in specific heat and magnetic susceptibility.
- Bilayer ^3He [4]: in 2007 a continuous quantum phase transition was observed by showing that the bilayer helium-3 system can be put in a quantum critical point at which the effective mass diverges while interband coupling goes to zero.
- LiHoF_4 – a model magnet [5]: this example is very important for us, because it allows experimental realization of the Ising chain in a transverse field, which we will address later in this work. This crystal is uniaxial and Ho ions behave exactly as the Ising model prescribes at low temperatures. By applying a magnetic field perpendicular to the easy axis (axis parallel to the direction in which it is the easiest to magnetize the material), Bitko et al. demonstrated in 1996 that a quantum phase transition happens at a certain magnetic field strength, where the ferromagnetic order is destroyed.
- The two-channel Kondo effect in a single electron transistor [6]: in 2007 the group of Goldhaber-Gordon at Stanford observed an unusual effect in the two channel Kondo model (where an $1/2$ impurity spin interacts with two independent baths of electrons thus suffering competition between two couplings) – when the difference in the couplings is small a two-channel Kondo effect emerges showing non-Fermi-liquid behaviour.
- Ultracold atoms in magnetic traps and optical lattices [7]: a vibrant field in experimental physics investigates atomic vapors cooled to the quantum limit. By now, in both bosonic and fermionic systems, which can be trapped routinely, external static potentials, interactions and even disorder can be tuned in a flexible way, thus ultracold gases offer a new exciting possibility for simulating the properties of correlated condensed matter systems. The realization of bosonic Mott insulators and the observation of the fermionic BEC-BCS transition are known among the most famous achievements of this field.

1.2 Motivation for using neural networks in condensed matter physics and quantum information theory

Big and complex ensembles of particles are hard to study because of the exponential growth of the phase space, classical or quantum. Despite the curse of dimensionality, machine learning is routinely used for analyzing giant and non-trivial sets of data. In the past decade, machine learning and artificial intelligence has achieved tremendous success, e.g., in the text, voice, and computer vision technologies and also in complex games such as Atari, Chess, and Go [8]. It is often argued that machine learning and artificial intelligence will also offer a strong helping hand in making scientific discoveries [9, 10, 11], and in particular, there is a hope that future techniques of this kind will be capable of automated discovery of robust characteristic features that enable a more efficient investigation of physical effects. Such methods would be most welcome for systems that are either highly non-integrable and even more for systems where even a clear order parameter is missing. Moreover, the physics which serves as the background for the quantum phase transitions discussed above is quite complex, and in many cases, not completely understood.

The last decade has been marked by growing interest to applications of various machine learning techniques in the field of condensed matter physics and quantum information science. Recent works include learning phase transition of the classic Ising model [12, 13, 14], solving the transverse field Ising model with the Boltzmann machine [15, 16, 17], super-resolving of the Ising model with convolutional neural networks [18] and machine learning topological invariants in one-dimensional insulator models [19], as well as quantum state tomography for both pure and mixed states [20, 21, 22, 23, 24]. In 2018 a group of physicists from Japan developed a deep neural network-based approach to detect phases of the quantum Ising chain [25]. This work in addition to [12] served as the primary research resource for this thesis project.

Noteworthy, all of the works mentioned above employ neural-network setups with various numbers of hidden layers and neurons, which is usually computationally demanding, meaning the models are expensive to train and evaluate.

1.3 Purpose statement

The **goal of the study** is to formulate an algorithm which uses **artificial neural networks** and results in a quantity capable of serving as a **qualitative order parameter, compare the latter to physical observables**, such as the entanglement entropy and the actual

order parameter of a system studied. Finally, **examine the algorithm on quantum spin models** for which the phase diagram has been known.

1.4 Problem statement

To achieve the goal of this study the following problems have been posed and successfully solved.

1. **Study the properties** of one-dimensional Ising model with both transverse and longitudinal fields and one-dimensional anisotropic XY model.
2. **Generate samples**, i.e., local measurements of spins in a certain basis, for different values of the corresponding control parameter, to train the algorithm on – **due to lack of experimental data**.
3. Calculate the **observables**.
4. **Program the algorithm**, based on comparison of ground-state probability distributions to detect quantum phase transitions, using modern machine learning libraries in Python.
5. **Test the algorithm** on the studied models.
6. **Show how useful the results of this research can be** for the field of condensed matter physics and quantum information theory.

Chapter 2

Quantum spin models: theoretical background

In this chapter, we review the quantum Ising model both with and without transverse and longitudinal fields and the anisotropic XY model, describe what is already known in the literature concerning their properties and explain the notations.

2.1 Transverse field Ising model

The transverse field Ising model (TFIM) represents a unique example of many-body interacting spin systems which can be rigorously approached in both equilibrium [26] and non-equilibrium [27] regimes. This model naturally appears upon approaching a classical two-dimensional Ising ferromagnet with nearest neighbor interaction and its exact solution is dated back to the original work [28]. Generally, the one-dimensional TFIM of L spins with nearest neighbor coupling is specified by the Hamiltonian

$$H = -J \sum_{i=1}^{L-1} \sigma_i^z \sigma_{i+1}^z - \tau \sum_{i=1}^L \sigma_i^x, \quad (2.1)$$

which represents a $2^L \times 2^L$ matrix, while a vector of Pauli matrices $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ acts in spin space, and J and $\tau \geq 0$ stand for the strength of the exchange coupling and the external magnetic field respectively (in the following we put $J = 1$ without loss of generality).



Figure 2.1: One-dimensional chain of $L = 10$ sites.

Interestingly, despite its relative simplicity this model was discussed to mimic the order-disorder transitions in ferroelectric crystals of KH_2PO_4 [29]. In the ordered phase described by the Hamiltonian (2.1) a rather strong magnetic field $\tau \geq 1$ destroys magnetic order even at $T = 0$, making thus this system an ideal playground for studying quantum phase transitions. At zero temperature, quantum fluctuations may lead to dramatic ground state restructure which is manifested by a non-analyticity in the ground state energy as a function of τ , that is, the derivatives of the energy are discontinuous. The QPT in the TFIM belongs to the class of continuous phase transition as the order parameter vanishes continuously at the critical value $\tau_c = 1$ of the magnetic field.

With no magnetic field present in the system ($\tau = 0$), the ground state configuration is purely determined by the exchange interaction, the first term in (2.1), which favors collinear magnetic ordering. For $J > 0$ ferromagnetic state is energetically preferable, meaning that all magnetic moments point in the same direction $\langle \sigma_i^z \rangle = +1$ (or -1), signalling the double degeneracy. Increasing the transverse field higher than $\tau = \tau_c$ makes the system unstable towards spin flip and all the spins aligned in x direction in the limit $\tau \rightarrow \infty$, i.e., disordered in σ_z basis.

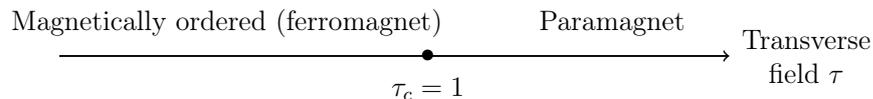


Figure 2.2: Zero-temperature phase diagram of the transverse Ising model.

One-dimensional TFIM can be worked out analytically by virtue of Jordan-Wigner transformation [2]. Indeed, the latter makes it possible to map an interacting spin model onto that of free spin polarized fermions [1]. Remarkably, the exact solution unambiguously demonstrates a continuous QPT upon passing through the critical field $\tau_c = 1$, separating magnetically ordered ferromagnet ($\tau < \tau_c$) and disordered paramagnet states ($\tau > \tau_c$). Although, there is no exact analytical solution in higher dimensional systems a phase transition can be clearly detected [1].

It is worth noting that the phase diagram of a one-dimensional TFIM is very similar to that of a two-dimensional classical Ising model at finite temperature with thermal driven phase transition. Interestingly, this dualism has a strict mathematical form corresponding to the so-called Suzuki-Trotter decomposition and which maps a d -dimensional quantum model to a $d + 1$ dimensional classical one [30].

2.2 Transverse field Ising model in the longitudinal field

It is quite natural to address the properties of the Hamiltonian (2.1) in a homogeneous longitudinal magnetic field $\lambda > 0$,

$$H = -J \sum_{i=1}^{L-1} \sigma_i^z \sigma_{i+1}^z - \tau \sum_{i=1}^L \sigma_i^x - \lambda \sum_{i=1}^L \sigma_i^z. \quad (2.2)$$

In fact, recent experimental investigations of low-lying spectrum of spin excitations in CoNb_2O_6 , described by this Hamiltonian, allow for direct probe of QPT [31]. Unlike the Hamiltonian (2.1) the one given by (2.2) belongs to the class of non-integrable systems, which means that it can not be diagonalized by analytic means.

Besides that, the phase diagram still stays under thorough investigation [32]. The main difference between this case and the case without longitudinal field is that since the ground state in the ferromagnetic regime is not degenerate anymore due to the longitudinal field and hence there is no spontaneous Z_2 symmetry breaking between the phases [33, 34].

The ZZ -interaction and transverse terms are competing in this system. This is parametrized when we scale everything by the Ising coupling J . We consider the model at $T = 0$. In case of the vanishing transverse field, the ground state is $\otimes_{i=1}^L |1\rangle_i$ with all spins lined up in the z -direction.

2.3 Anisotropic XY model

The XY model is yet another well-known quantum spin lattice model of magnetism. One can arrive to the isotropic version of this model by switching off the ZZ couplings in the Heisenberg Hamiltonian. In its turn, the anisotropic XY model is a generalization of it in the sense that the interaction strength in the XY plane is not isotropic anymore. In this study, we limit ourselves to the case when there is no field transverse to the interaction plane. The Hamiltonian of the model is thus given by

$$H = -J \sum_{i=1}^{L-1} \left(\frac{1+\gamma}{2} \sigma_i^x \sigma_{i+1}^x + \frac{1-\gamma}{2} \sigma_i^y \sigma_{i+1}^y \right), \quad (2.3)$$

where γ is the anisotropy parameter that is usually restricted to $-1 \leq \gamma \leq 1$ and J is the coupling strength which we set to 1 hereafter. If one sets $\gamma = 0$ the fully isotropic case, which possesses an additional symmetry $[H, \sigma_i^z] = 0$, is restored. On the other hand, it is also well-known that in the opposite case, i.e. $\gamma = 1$, the ground state possesses a long-range

Neel order which yields

$$\sigma_i^x |\sigma\rangle = (-1)^i |\sigma\rangle \quad (2.4)$$

and

$$\sigma_i^y |\sigma\rangle = (-1)^i |\sigma\rangle \quad (2.5)$$

for $\gamma = -1$ accordingly, as is described in detail in [35]. It is clear that as γ decreases from 1 to -1, the x - and y -components begin to compete. Its phase diagram is thus given by an x - and y -ferromagnetic states for $\gamma = 1$ and -1 accordingly. The model is fully isotropic at $\gamma = 0$ and undergoes a second-order phase transition at this point while the gap continuously vanishes [35, 36].

Chapter 3

Methods for finding ground states of quantum spin Hamiltonians

In this chapter, we review and briefly explain the existing techniques of tackling quantum many-body problems and quantum phase transitions.

3.1 Analytical approaches

Studying exact analytical and numerical approaches for the transverse field models has a long history and is summarized in a number of excellent monographs [37, 38, 39]. In this section, we briefly discuss the main strategies with which to tackle TFIM in the vicinity of the QPT. The presence of the transverse magnetic field τ in the Hamiltonian (2.1) leads to the tendency towards the flip of z components of spins, being thus the source of quantum fluctuations in the system. Upon imposing periodic boundary conditions, i.e., $\sigma_{L+1}^z = \sigma_1^z$ the system allows an exact analytical solution by means of Jordan-Wigner transformation. The latter makes it possible to reduce the Hamiltonian (2.1) to that of non-interacting spinless fermions with the quasiparticle energy

$$\omega_q = J\sqrt{1 + 2\kappa \cos q + \kappa^2}, \quad (3.1)$$

where q specifies the quasimomentum, and $\kappa = J/\tau$ is the relative strength of the external magnetic field. It is therefore not surprising that the quasiparticle energy gap $\Delta = 2|1 - \kappa|$ collapses at $\tau = J$, manifesting QPT.

At criticality $\tau_c = J$ the ground state energy E of a one-dimensional TFIM of L spins with the Hamiltonian (2.1) and open boundary conditions is determined by the exact analytical

expression,

$$E = J \left[1 - \operatorname{cosec} \left(\frac{\pi}{2(2L + 1)} \right) \right]. \quad (3.2)$$

Both order parameters, i.e., the one in magnetically ordered state $m_z = \frac{1}{L} \sum_{i=1}^L \langle \sigma_i^z \rangle$ and the other in disordered paramagnet phase $m_x = \frac{1}{L} \sum_{i=1}^L \langle \sigma_i^x \rangle$, continuously vanish at the QPT point.

3.2 Numerical approaches

For a long time, a proper theoretical description of interacting quantum systems had been based on various approximate solutions, in which the microscopic interaction was considered either perturbatively or was simplified. In the meantime, rapid progress in technological development has made it possible to exactly approach small and moderate-size interacting quantum systems by means of numerical exact diagonalization. Most of the numerical techniques exhibit an exponential scaling with the size of a system, which blows up the allocated computer memory to store a quantum state or a Hamiltonian. However, a variety of approximations can be made to speed up the simulations beyond this issue. In particular, numerical methods "throwing out" high-energy properties of the system, the most irrelevant to condensed matter experiments, dramatically reduce the dimension of the corresponding Hilbert space leaving us with the problem of polynomial complexity. The latter makes it possible to directly estimate low lying states by means of diagonalization routine based, e.g., on Lanczos algorithms or iterative solvers, which yields a better scaling as compared to brute force approach [40]. More involved methods, e.g., the density matrix renormalization group (DMRG), can be also successfully applied and involve sophisticated decimation schemes based on the entanglement structure of the wave function [41].

Another main strategy for solving many-body interacting systems is stochastic approaches based on Markov chain Monte Carlo. They require statistical sampling of the configurations that are distributed according to the wave function. In these methods, configurations are sampled in a Markov process [42]. The problem with these stochastic methods is that the coefficient of the wave function for each one of the samples has to be interpreted as a probability distribution. Only then can one update configurations stochastically giving access to observables. There are many cases where stochastic approaches do not work in many-body problems, typically when the so-called sign problem (which is basically the presence of negative probabilities) appears [43]. Where these methods do work, one can get the number of spins L in the thousands, say for DMRG [41]. And for stochastic quantum Monte Carlo methods, in its turn, one can reach L in the millions or more [42].

When these methods work, we should definitely use them. But there are many cases, where we know they do not work [44, 45]. In these cases what people are doing right now in physics, computer science, quantum chemistry, academia and industry is turning to new approaches based on machine learning.

3.3 Machine learning approaches

Keeping in mind that machine learning approaches are also numerical, the motivation behind using them can be justified in a simple manner. Spin configurations from the snapshots of the system can be mapped to either binary numbers or pixels, which immediately maps a physics problem to an image processing problem for a large number of samples. There are various strategies that people use to employ machine learning so that one can turn to quantum many-body problems [12, 14, 46, 24].

Supervised learning is when one has data with labels. So if we can take configurations and label them, say, ferromagnetic or paramagnetic, then we can essentially do many tasks in classification with supervised learning techniques. This, of course, requires labels, which is a nontrivial task itself.

In many cases, in quantum many-body problems, where we do not have labels, we would turn to unsupervised learning. Unsupervised learning might be looking at a quantum wave function and trying to find clustering or associative rules that govern its behavior. Unsupervised learning can also be taking measurement data and essentially reconstructing the wave function from individual images or snapshots. That is like a reconstruction of, say, a probability distribution in the classical case or the full wave function in the quantum case. These reconstruction techniques based on machine learning are now being studied and compared to traditional techniques based on quantum state and quantum process tomography [47, 48, 49, 50, 24].

So, as we go into the future, we have a toolbox made up of traditional analytical and numerical approaches, which have been developed for decades, but we also have a new toolbox based on machine learning that might be crucial in the investigation of future quantum materials and devices.

3.4 Methods for detecting quantum phase transitions

Suppose for now that we managed to find the ground state of the model by any of the methods listed above and know how to extract any type of information from it. The open

question is then how one can determine the signature of a quantum phase transition or its absence. To elaborate, we are gradually varying the control parameter and at any given time would like to know which phase our system is in. As an answer, several methods can be proposed, which nowadays tend to be used the most often.

- First of all, one can look at the behavior of an observable, typically for this role one employs the average magnetization. In this particular method what one can try to catch is nothing but a finite region of sudden change that shrinks in the thermodynamic limit to a single point of non-analyticity [51].
- Secondly, one can examine the critical behavior of the energy gap Δ or the correlation length ξ , that were mentioned in Chapter 1, since the vanishing gap and the diverging correlation length are a clear sign of the QPT [1, 52].
- And last but not least, is the fidelity approach, which is focusing on studying the amount of change of the wave-function while changing the governing parameter. More precisely, the fidelity approach to QPTs uses the overlap between ground states of the system to gain some information about its quantum phases [53, 54].

Quantum many-body problems belong to the class of problems whose full solution is exponentially hard. Albeit, approximate methods have been developed, many of these are also extremely computationally demanding. As opposed to the latter, machine learning methods can be less computationally demanding, however catching all physics, that lies beyond, e.g., non-local correlations.

To the best of our knowledge, the quantum transverse field Ising model and the QPT it undergoes was addressed in the following papers:

1. In 1979, Bambi (see [55]) used the Renormalization Group (RG) approach and revealed the QPT point and the critical exponents.
2. In 2008, Jordan et al. tackled the quantum Ising criticality with the help of the Projected Entangled-Pair State (PEPS) algorithm [56], a tensor-network based approach.
3. In 2015, tensor networks were employed in the form of the Multi-Scale Entanglement Renormalization Ansatz (MERA) by Evenbly and Vidal to find out the critical properties of the 2D classical Ising model, which is mappable to the quantum model [57].
4. In 2018, a group from Japan [25] used deep neural networks to detect the QPT in 1D quantum Ising model.

As for the anisotropic XY model:

1. In 1961, Lieb et al. (see [35]) proposed the model as a model for ferromagnetic materials.
2. In 1971, Barouch and McCoy [58] solved the model exactly with the help of the usual technique of Jordan-Wigner and Bogoluibov transformations.
3. In 2018, a group of Chinese physicists employed the fidelity approach to find the critical point of the model [36].

In all of the works listed above, the best accuracy (relative to the exact theoretical result) achieved was less or equal than 1%.

Chapter 4

Classification with neural networks

4.1 Architecture

Let us begin with the description of a feed-forward neural network (we will hereafter omit the word "feed-forward" since this is the only type of neural networks we consider in this study) by the architecture diagram, which is presented in Figure 4.1.

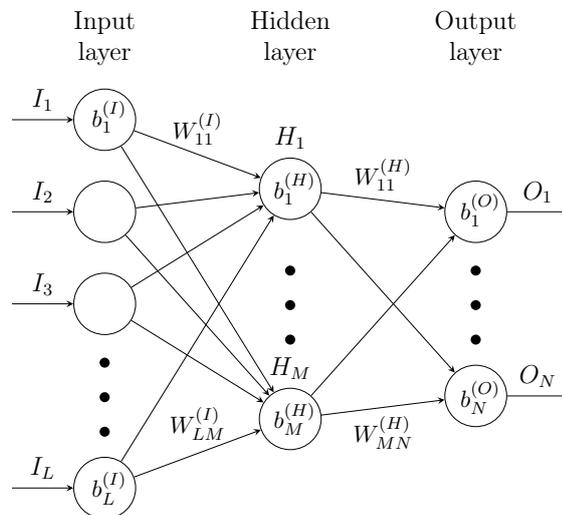


Figure 4.1: A common neural network architecture with 1 hidden layer.

Such neural networks have the following characteristics.

1. They consist of neurons and connections between them (white circles and black arrows accordingly in Figure 4.1).
2. The neurons are arranged in layers such that the first layer $I = (I_1, \dots, I_L)$ takes in the input and the last layer $O = (O_1, \dots, O_N)$ produces the output.

3. Each layer has a corresponding set of numbers $b^{(I)}, b^{(H)}, b^{(O)}$ which we will refer to as biases.
4. The middle layer $H = (H_1, \dots, H_M)$ is not being observed during computation and hence is called the hidden layer.
5. Each neuron in a layer is connected to every neuron of the next layer by the corresponding set of weights W , which means the information is constantly flowing forward from one layer to the next.
6. There is no connection among neurons in the same layer.
7. The number of hidden layers is not limited and can be as large as possible, resulting in a more expressive model, but more expensive computation.

4.2 Activation function

During the computation, we will use non-linear functions, which are usually called activation functions. Most common choice includes one of the three functions, shown in Figure 4.2.

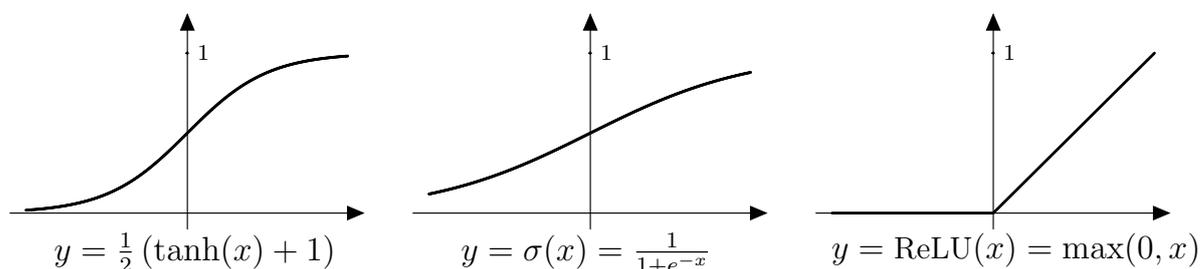


Figure 4.2: Nonlinear activation functions typically used for machine learning problems.

These three are referred to as hyperbolic tangent, sigmoid (or logistic), and ReLU – rectified linear unit. The first two provide continuous maps $\mathbb{R} \rightarrow (0, 1)$ and so their output is interpreted as the probability of a neuron to be turned on. $\text{ReLU}(x)$ however has unbounded output and a discontinuous derivative at zero.

4.3 Computation graph

Now, let us explain how the whole computation is handled. The calculation of the output proceeds as follows: first, the state of the hidden layer $H_i, i = 1, \dots, M$ is computed as

$$H_i = f_H \left(\sum_{j=1}^L W_{ij}^{(I)} I_j + b_i^{(H)} \right), \quad (4.1)$$

or, in the vector-matrix product form

$$H = f_H (W^{(I)} I + b^{(H)}), \quad (4.2)$$

then the values of the hidden neurons are passed to the output layer, where we again apply the activation function to the linear transform of the hidden neuron values:

$$O = f_O (W^{(O)} H + b^{(O)}) = g (W^{(O)} f_H (W^{(I)} I + b^{(H)}) + b^{(O)}). \quad (4.3)$$

Note, that the activation functions are applied element-wise.

4.4 Our model

Even though some problems require really deep neural networks (with the number of the hidden layers order of 100), they are usually computationally hard to train and evaluate. Let us proceed to the simplification of the model depicted in Figure 4.1, that we introduce in our setup by removing the hidden layer of the network. In the present form (see Figure 4.3) this setup represents the simplest neural network possible.

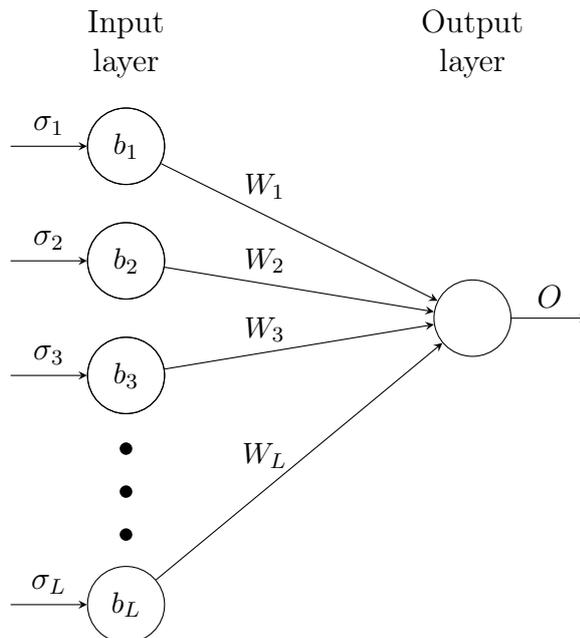


Figure 4.3: Neural network architecture used in the present study.

In Figure 4.3, σ_i 's stand for the spin projection values, i.e., $(\sigma_1, \dots, \sigma_L)$ is a bitstring of 0's and 1's, which we obtain during the sampling procedure, described in Chapter 5, and feed to the neural network. At the same time we will denote the values each input neuron takes during the computations as σ_i as there is straight analogy of this set of neurons to one-dimensional spin chain.

When the input neurons take certain values σ_i , the values of the output neuron O is calculated in the form of

$$O = f \left(\sum_{i=1}^L W_i \sigma_i + b_i \right), \quad (4.4)$$

where W_i is a vector of weights, which are real numbers, connecting visible neurons to hidden neurons and b_i is a bias of the i^{th} hidden neuron.

A spin configuration of a quantum spin chain is fed to the input layer. When a configuration is fed to the visible layer of the network N_i , the output neuron O is prescribed to take a value of either 1 or 0, depending on the parameter value τ , the spin configuration was drawn from $-\tau_0$ or τ_i (meaning that we have two probability distributions in front of us at two values of the control parameter). The fact that we have only one output neuron can be explained by the nature of binary classification. Should we have multiclass classification, then the number of classes would correspond to the number of output neurons.

4.5 Training

What then happens is a so-called supervised learning phase: a network is shown a large set of configurations belonging to either τ_0 or τ_i . Each configuration is known to belong to the corresponding probability distribution — and thus this configuration is labeled with a correct answer.

For a given configuration and a correct answer, an estimate of the classifier's mistake is given by the so-called binary cross-entropy, which is sometimes also called the log-loss function, defined as:

$$H_p(q) = -\frac{1}{N} \sum_{i=1}^N [y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i))], \quad (4.5)$$

where N is the number of samples used during the training stage, y_i is the label of i^{th}

configuration (either 0 or 1), and $p(y_i)$ is the predicted probability of a configuration to belong to one of the classes.

The value of $H_p(q)$ is zero for a perfect classifier. For a given training set of configurations with each one assigned a proper class, binary cross-entropy is a function of network parameters – the weights W_i and the biases b_i . The network is trained by minimizing the log-loss in the space of weights and biases with the help of RMSprop optimization algorithm [59].

Chapter 5

The setup

The aim of this chapter is to introduce the reader to the machine learning setup and manipulations with the Hamiltonians and the ground states we perform throughout this work.

5.1 Sampling

Having large amount of data, a so-called data set, is crucial for applications of machine learning by its nature. As a method to sample spin configurations we decided to employ the exact diagonalization routine instead of sophisticated quantum Monte-Carlo techniques, which are nontrivial to realize and still remain a subject of investigations themselves [60]. In the end, for well-studied systems, where all the properties are known in the thermodynamic limit, we do not need tremendous chain's size to demonstrate the success of machine learning in detecting phase boundaries and the transitions between them.

Let us describe the sampling procedure by example of the TFIM without any longitudinal field. Interestingly, adding the longitudinal field does not change the procedure qualitatively. For all of the following calculations, numerics and visualization we use Python programming language [61] with the following libraries: NumPy [62] and SciPy [63] for basic operations and numerics, Matplotlib [64] for visualization, scikit-learn [65] and Keras [66] for machine learning subroutines. All the code is reproducible and can be found in [67]. Detailed description of the machine learning part is given in Appendix A.

Note that in the canonical basis with σ^z being diagonal, which is also sometimes called the standard or the computational basis, the TFIM Hamiltonian matrix is a sparse one, i.e., it is a matrix with most of the elements being zeros. Adding the longitudinal field does not change this, as σ_i^z is diagonal in the standard basis for every i in the chain. In Figure 5.1 one

can see the visualized matrix elements and make sure that most of the elements are zeros. The property that the majority of the elements are zero does not change while varying the control parameter τ .

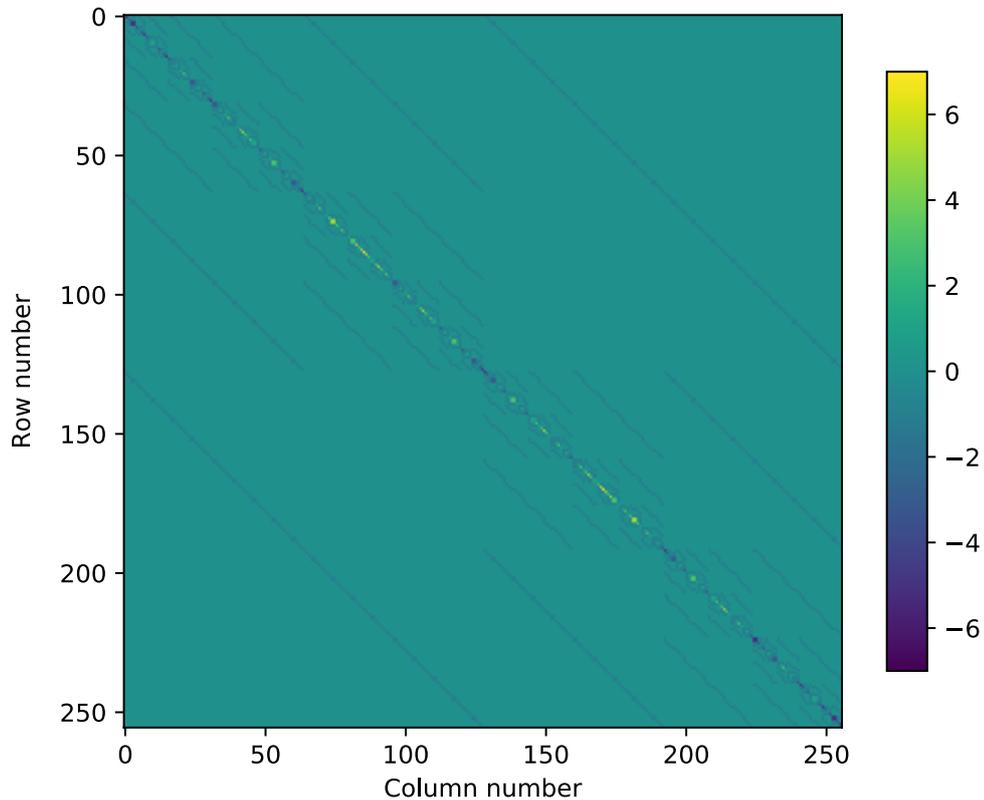


Figure 5.1: Visualization of matrix elements of the TFIM Hamiltonian of 8 spins at $\tau = 1$.

Next, we use exact diagonalization to obtain the ground state of the system $|g\rangle$, which is the eigenvector of a hermitian matrix $H = H^\dagger$ corresponding to the smallest eigenvalue E_0 (which is actually the energy of the system in the state $|g\rangle$) defined as follows:

$$H|g\rangle = E_0|g\rangle. \quad (5.1)$$

In general, the ground state vector $|g\rangle$ is 2^L -dimensional vector with complex entries g_i .

$$|g\rangle = \begin{pmatrix} g_0 \\ \vdots \\ g_{2^L-1} \end{pmatrix}. \quad (5.2)$$

Thus, having the ground state vector, we can obtain, following Born's rule [68], the probability distribution of spin positions in the σ_z -basis. The distribution looks the following way:

$$\mathbf{p}(g) = \begin{pmatrix} p_0(g) \\ \vdots \\ p_{2^L-1}(g) \end{pmatrix}, \quad (5.3)$$

where $p_i(g) = |g_i|^2 = \bar{g}_i g_i$ and the bar sign $\bar{}$ means the complex conjugation.

Thus $p_i(g)$ is the probability of observing the i -th bitstring (by a bitstring we mean a configuration of σ_z observables of magnetizations on sites of the lattice) if the system is prepared and stays in the ground state.

The task is now to sample such bitstrings with the corresponding probabilities. If one has an access to the ground state wavefunction (the ground state vector $|g\rangle$), then one simple way to do this is the following: Consider the interval $(0, 1)$ partitioned into intervals with respect to the probabilities p_i . Then we can toss a coin – that is, a uniformly-distributed (over $(0, 1)$) random number is generated, and the string that corresponds to the interval where the coin fell is thus sampled. On the image below, $L = 2$ – there are two spins, thus there are four basis bitstrings: $\{s_1 = |00\rangle, s_2 = |01\rangle, s_3 = |10\rangle, s_4 = |11\rangle\}$ and the coin fell on the $|01\rangle$ interval, thus the $|01\rangle$ string is returned. This sampling process is visualized in Figure 5.2.

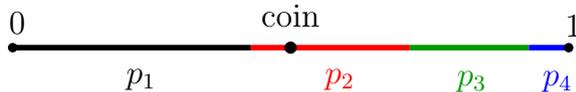


Figure 5.2: Visualization of the coin toss sampling.

Luckily, such a sampling procedure can be easily performed with the help of only one function from the NumPy package [62] – `numpy.random.choice`, which generates a random sample from a given one-dimensional array.

Finally, to obtain the bitstrings from the sampled integers, one has simply to convert them

to binary, which is trivial.

Let us now show the procedure described above with a TFIM example. The problem could be that the major number components of the ground state vector are tiny numbers (they have to sum up to 1 as the probability distribution properties prescribe) – and due to this fact such corresponding bitstrings are very unlikely to be sampled for us.

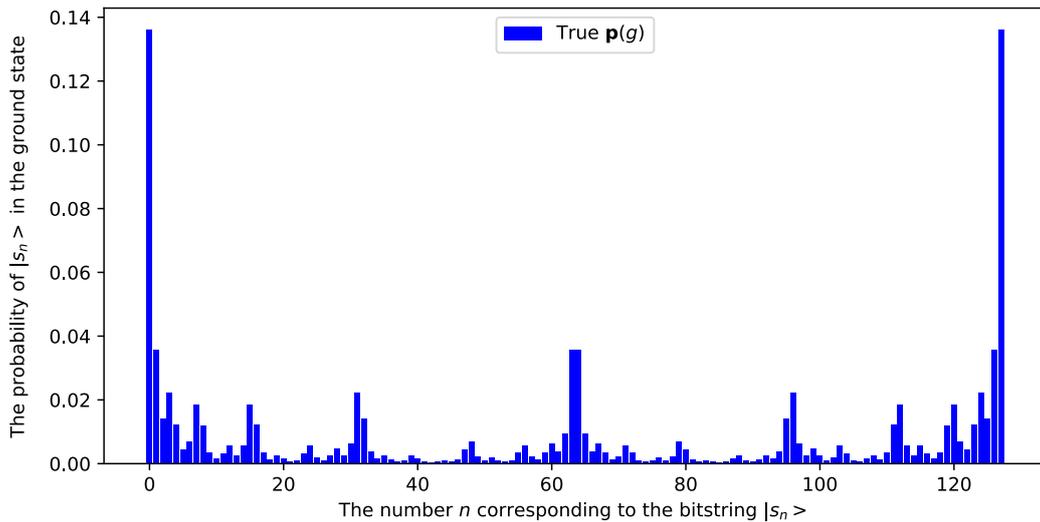


Figure 5.3: Distribution of the basis bitstring configurations for the TFIM Hamiltonian on $L = 7$ spins at $\tau = 1$.

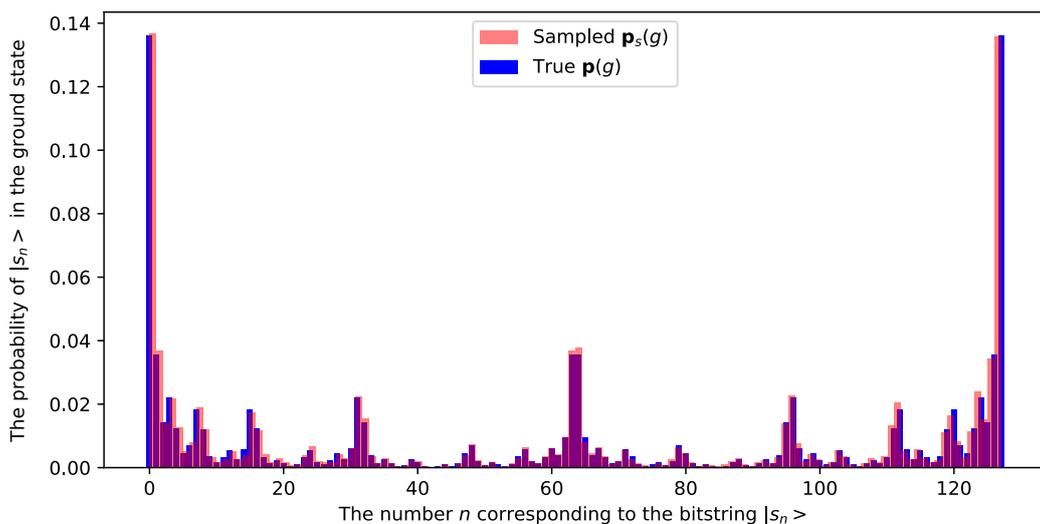


Figure 5.4: Distribution of the basis bitstring configurations for the TFIM Hamiltonian on $L = 7$ spins at $\tau = 1$ and $N = 10000$ configurations sampled by our method.

In Figure 5.4 one can see that the distributions coincide with each other with remarkable accuracy with the classical fidelity between two probability distributions defined as

$$F(p, q) = \left(\sum_i \sqrt{p_i q_i} \right)^2, \quad (5.4)$$

being 0.99.

For a larger system, the probability distribution of the ground state over the basis states is much harder to visualize (see Figure 5.5) due to the fact that the number of possible bitstrings 2^L is exponentially big, but the fidelity (being 0.94 for two probability distributions in Figure 5.5) is still acceptable.

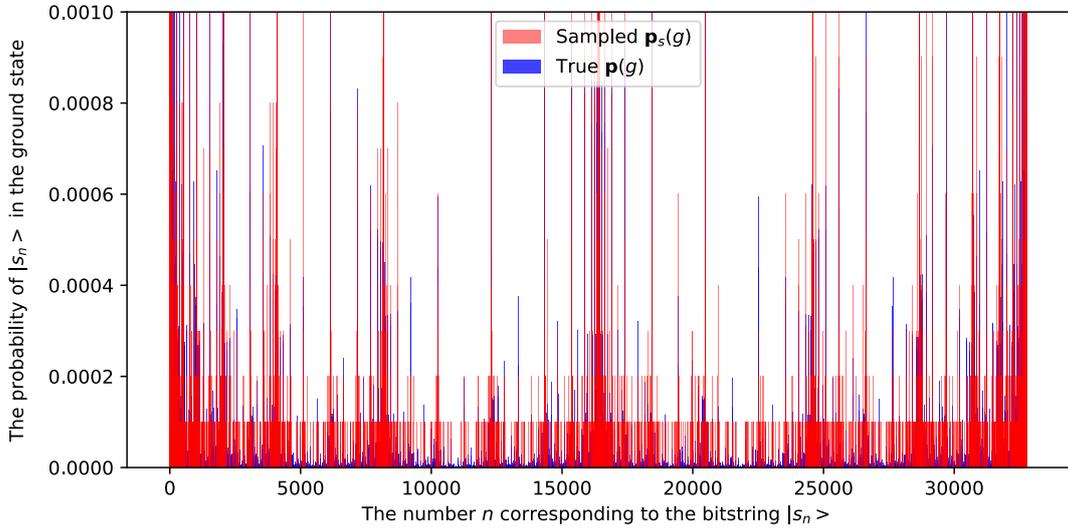


Figure 5.5: Distribution of the basis bitstring configurations for the TFIM Hamiltonian on $L = 15$ spins at $\tau = 1$ and $N = 10000$ configurations sampled by our method.

5.2 Methodology

Now we will describe the main part of our work.

First of all, we set the discretized values (in quantity of D) of the Hamiltonian control parameter $\tau = \{\tau_k\}_{k=0}^{D-1}$.

Subsequently, for every τ_k , we find the ground state vector $|g(\tau_k)\rangle$ and transform it to the probability distribution $\mathbf{p}(g(\tau_k))$ in the way that was described in the previous section.

After that, we sample N bitstrings for each τ_k using the sampling procedure described above. The boundary values were taken as follows: $\tau_0 = 0$ and $\tau_{D-1} = 2$. In our work we take $N = 20000$ and $D = 40$.

Afterwards, a feed-forward neural network \mathbb{N}_k is trained to classify the bitstrings sampled from $\mathbf{p}(g(\tau_0))$ and $\mathbf{p}(g(\tau_k))$. Finally, we end up with $D - 1$ pairs of (c_i, τ_i) with $c_i \in (0, 1)$ being the mean output of the neural network evaluated on the samples drawn from the probability distribution given by the ground state of $H(\tau_i)$. The split ratio between the train and the test data S was kept 0.25 during all of the simulations. In what follows, we show that the value of c with respect to τ dramatically changes signalling a phase transition. We apply a similar procedure to the anisotropic XY model with the anisotropy parameter $-1 \leq \gamma \leq 1$ starting with $\gamma_1 = -0.99$. This scanning subroutine is demonstrated in Figure 5.6.

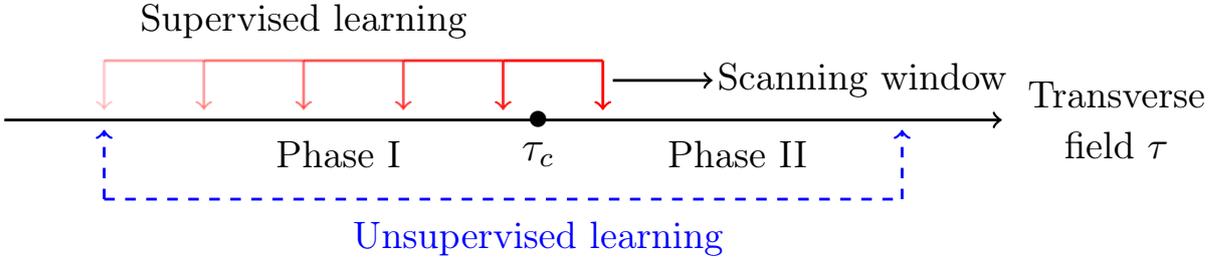


Figure 5.6: Visualization of the setup.

Performing a numerical experiment which implies scanning this way through the whole reasonable region of the corresponding parameter axis, we obtain the behavior of mean output c with respect to τ . It is natural to believe that c is going to have critical behavior at the phase transition point and thus would signal the phase transition region, because the ground states of the system should be dramatically different in distinct phases.

Chapter 6

Results and discussion

In this chapter, we provide and discuss the main results of this work.

6.1 Transverse field Ising model

For the TFIM it first makes sense to examine the behavior of the energy gap Δ , as we know from the literature, that it is an explicit order parameter obtained from exact theory [1]. Figure 6.1 shows its dependence of the transverse field value τ . The values of Δ were obtained by exact diagonalization of the model's Hamiltonian.

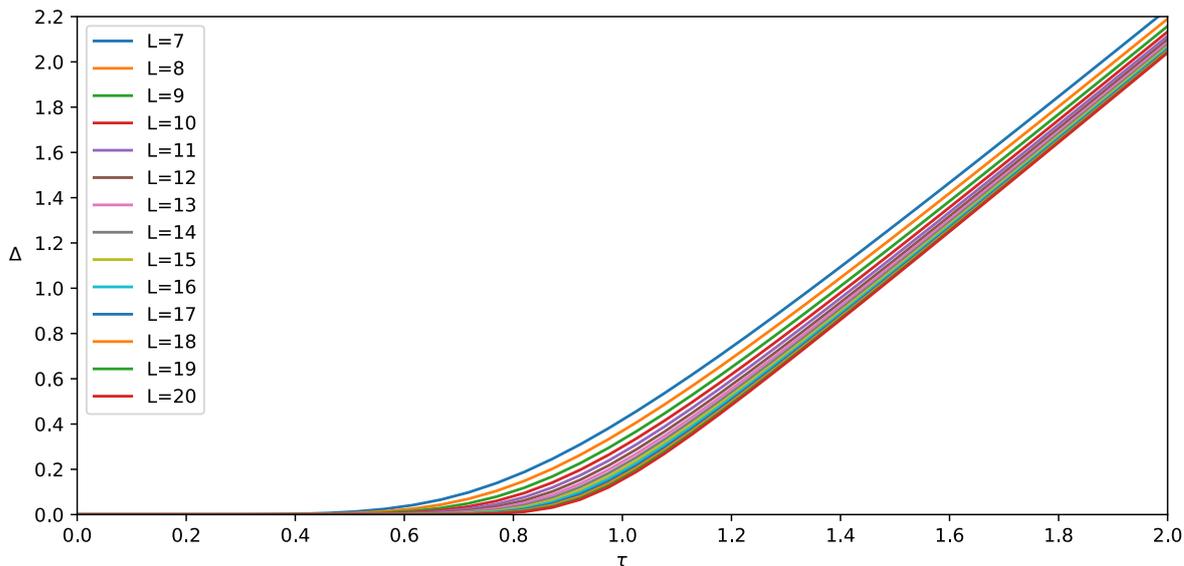


Figure 6.1: Energy gap Δ against the transverse field τ for various lengths of the chain.

Indeed, one can see that the plot shows two distinct regions. In the first, the quantity Δ is zero, in the second – nonzero. This behaviour can be explained in an intuitive way:

the ground state is double-degenerate (thus the two lowest eigenvalues are equal) in the ferromagnetic phase and non-degenerate in the paramagnetic phase. The critical point and its neighbourhood moves closer to 1 while increasing the chain size.

Now that we can extract a critical point by fitting the plot in the paramagnetic phase with a linear function (keeping in mind that for TFIM the critical exponent value for the energy gap Δ is 1, which is common knowledge available, e.g., in [60]) and looking at the x -coordinate of the intersection with the x axis, as displayed in Figure 6.2, we have an etalon to compare our machine learning-based method with.

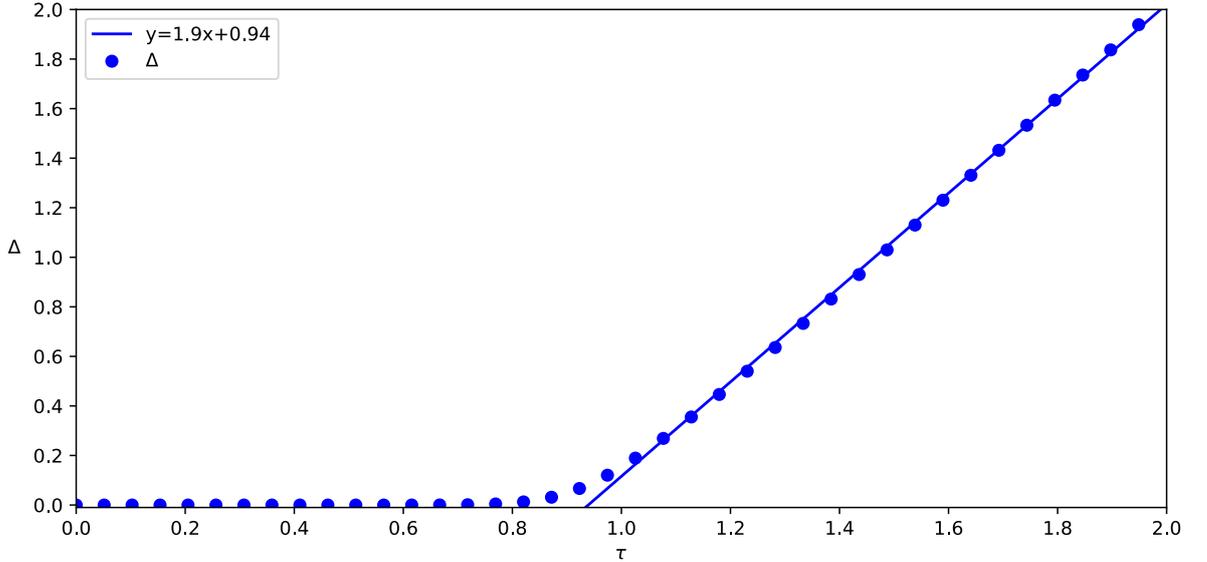


Figure 6.2: Energy gap Δ and the linear fitting function against the transverse field τ for $L = 20$ exhibiting the QPT point at $\tau_c = 0.94$.

Besides the energy gap Δ , one can also take a look at the magnetic properties of the system. Originally, the Ising model was used to obtain macroscopic magnetic features from microscopic interactions in the system [69]. Thus, one can expect that the phase boundaries might be signalled by magnetic observables, for example, the transverse magnetization m_x , which we define below in the following manner:

$$m_x = \frac{1}{L} \left\langle \sum_{i=1}^L \sigma_i^x \right\rangle, \quad (6.1)$$

where averaging is performed over the quantum state.

Figure 6.3 shows our calculations of the transverse magnetization m_x via exact diagonalization with respect to the transverse field τ . The curve displays a typical tangentoid form,

exactly as we could have expected from [1].

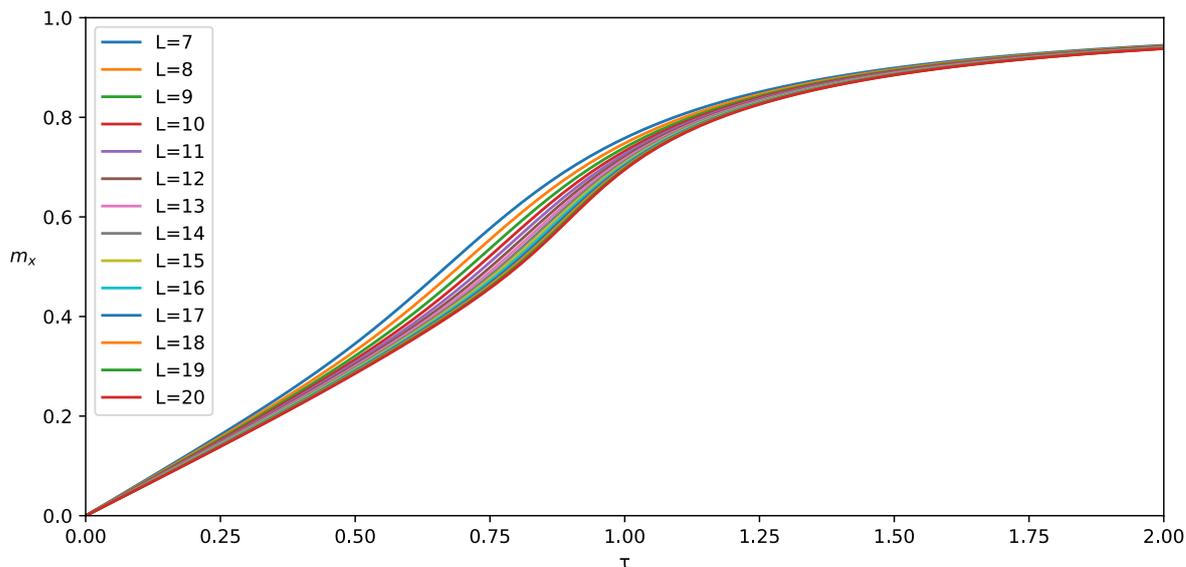


Figure 6.3: Transverse magnetization m_x against the transverse field τ for various lengths of the chain.

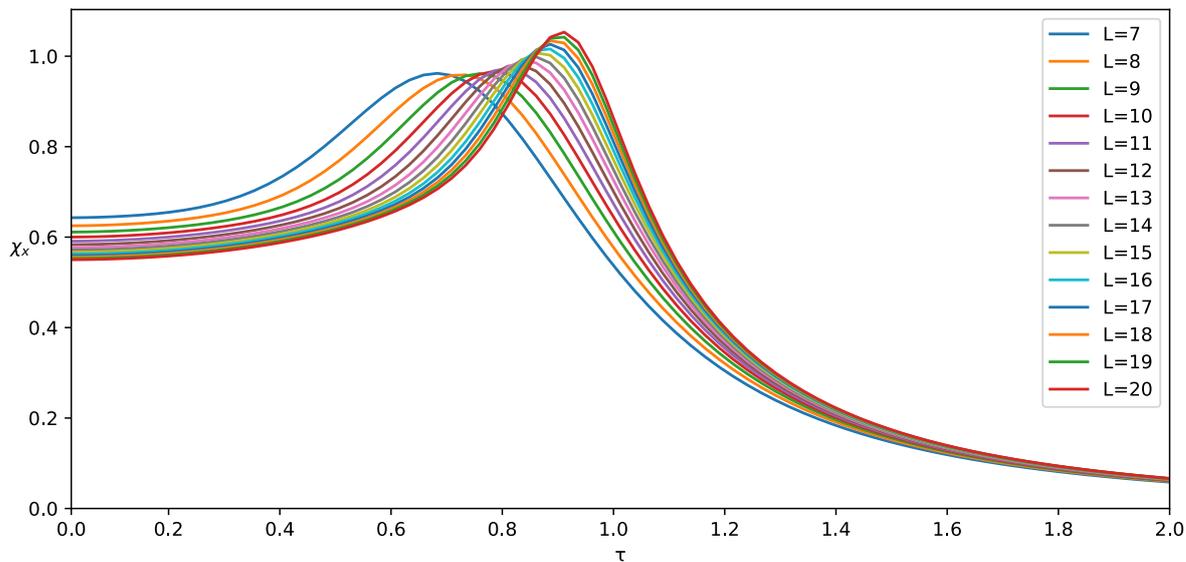


Figure 6.4: Transverse magnetic susceptibility χ_x against the transverse field τ for various lengths of the chain.

Moving further, to get clearer visualization of the QPT crossover region, one might be interested in the behavior of the average transverse magnetic susceptibility χ_x which is defined below:

$$\chi_x = \frac{\partial}{\partial \tau} m_x. \tag{6.2}$$

In Figure 6.4 we plot the susceptibility calculated by first-order finite difference of the average transverse magnetization m_x .

Loosely speaking, the plot in Figure 6.4 can be explained by a simple idea: it shows that the spins are more sensitive to the change of an external transverse field τ in the vicinity of the QPT point. Of course, the x -coordinate of the hill corresponds to the point where the energy gap Δ vanishes.

Now, following the logic of our work, let us look at how the method, which was derived and explained in Chapter 5, tackles the problem of defining the critical point.

In Figure 6.5, one can see a typical plot from a large number of figures we have obtained during our work with chains of different sizes. This particular one displays the crossover for $L = 20$ at $\tilde{\tau}_c = 0.96$. We have performed experiments with chains from $L = 7$, which appeared to be the lower limit of displaying anything distinguishable from the sampling noise, to $L = 20$, which bounds the memory overflow on a typical laptop we have used.

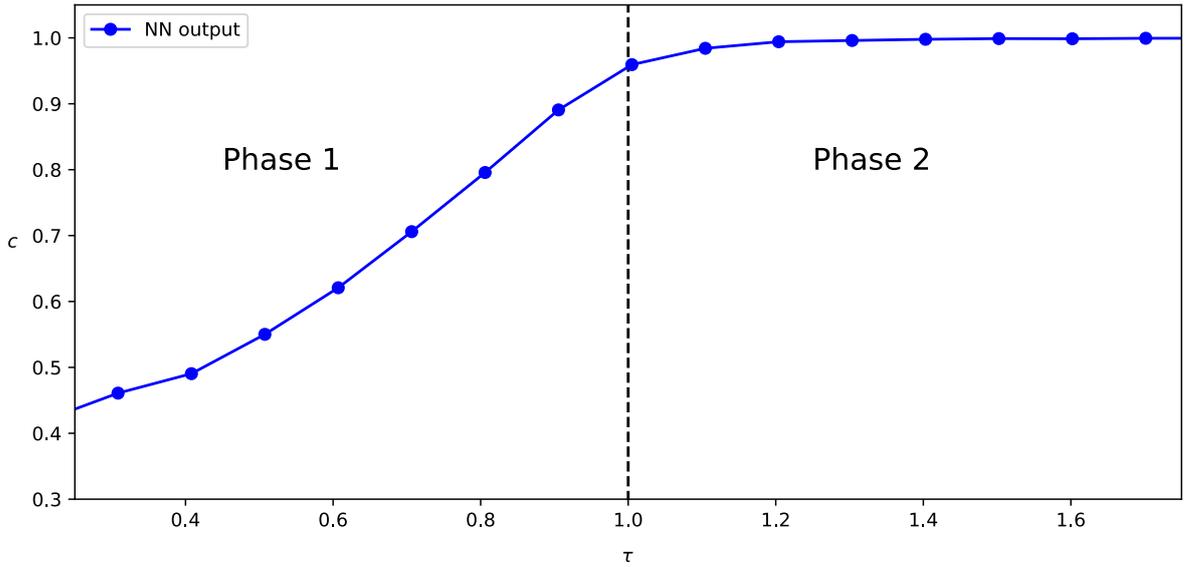


Figure 6.5: Neural network output c plotted against the transverse field strength τ for a chain of 20 spins.

To extract the pseudocritical transverse magnetic field value $\tilde{\tau}_c$ from the set of points our setup returns, we fit the obtained curve by a hyperbolic tangentoid defined in the following form:

$$A \cdot \tanh(B \cdot (\tau - \tilde{\tau}_c)) + C \tag{6.3}$$

where A , B , C and $\tilde{\tau}_c$ are parameters of the function.

The fitting procedure was performed with the help of the standard mean square fitting technique.

Note, that while the resulting curve is typical of a transverse magnetization curve for TFIM, there was no information about the x -projections of the spin measurements in our setup, but only the measurements in the z -basis.

Moreover, the simplicity of a neural-network architecture we are using gives us the opportunity to visualize the weight vector \mathbf{W} of a neural network \mathbb{N}_i without extra matrix manipulations (as it was done, for example, in [25]), for details see Appendix A. As a result, Figure 6.6 clearly displays the crossover in the neighborhood of the pseudocritical point $\tilde{\tau}_c$.

Remarkably, this picture appears to be interpretable, which is, to the best of our knowledge, unusual for typical machine learning models such as deep neural networks [70, 71]. The explanation might proceed in such a manner: our architecture implies only coefficients by which the spins' z -projection values are multiplied before feeding the whole sum to the output neuron. Consequently, one might suggest that the model tries to mimic the spin configurations given the transverse magnetic field value τ . This would be a plausible explanation of why the columns on the color map, which correspond to coefficients to be multiplied by spin projections, are uniform in the ferromagnetic limit and look random in the disordered phase. The boundary spins' coefficients might be different because of the open boundary conditions.

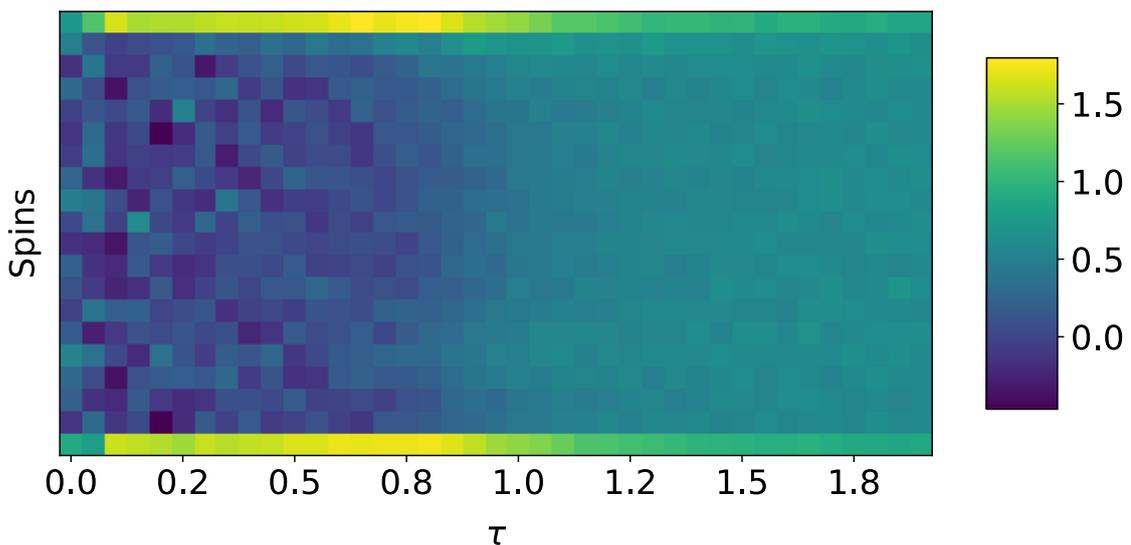


Figure 6.6: Visualized weights \mathbf{W} for every spin in a chain of 20 spins against the transverse field τ .

We can also compare the obtained critical points to estimate the accuracy of our method and track the finite-size critical exponent γ defined with by the following equation:

$$|\tau_c - \tilde{\tau}_c(L)| = \frac{1}{L^\gamma}. \quad (6.4)$$

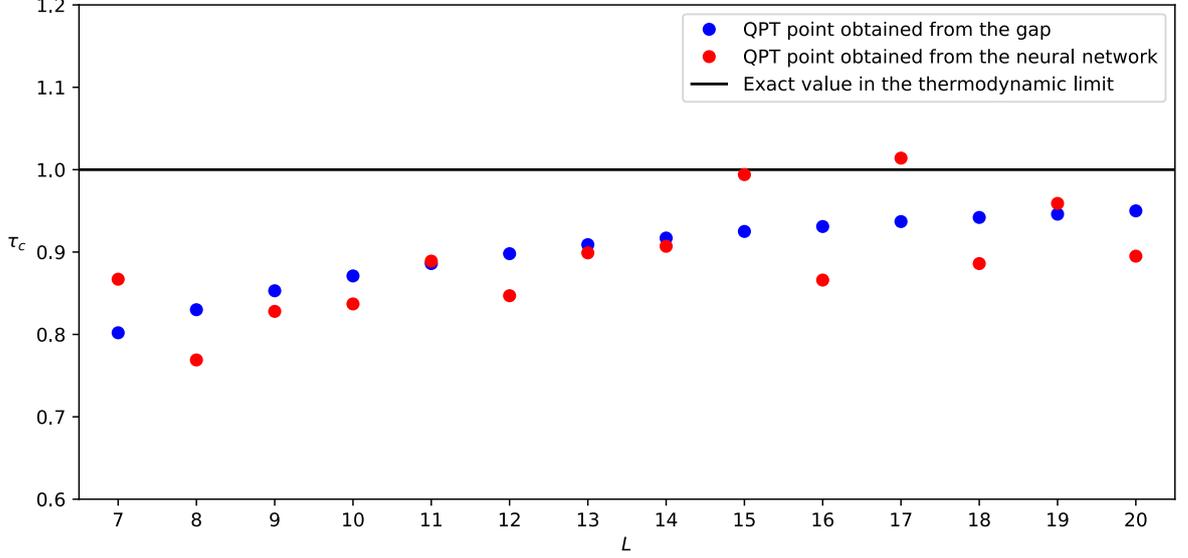


Figure 6.7: Critical points obtained by the neural-network method and by the explicit order parameter for finite-size chains.

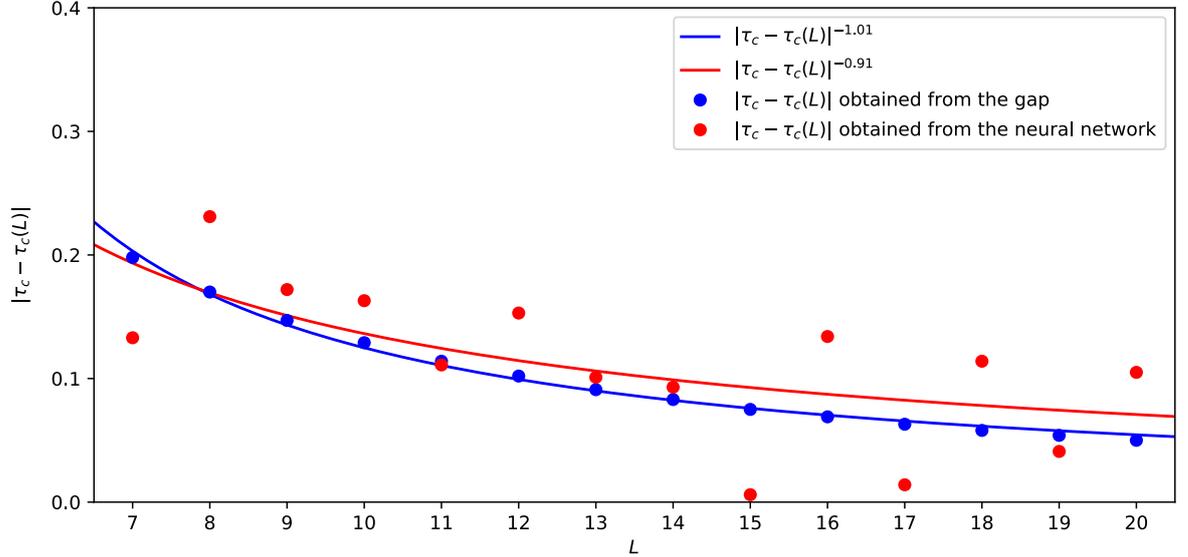


Figure 6.8: Critical points obtained by the neural-network method and by the explicit order parameter for finite-size chains.

From Figure 6.8 one can see that our method reaches accuracy of approximately 10%. To improve this one could reasonably suggest to run the program, obtaining the point multiple

times to obtain the value of statistical mean. However, this is very computational- and thus time-demanding. To the best of our knowledge, people usually take the best score of the machine learning subroutine, but not the mean value obtained from, say, 100 software runs.

Another way to address our machine learning-based approach is to compare the widths of the crossovers obtained by different methods. First of all, one can track the widths of the hill the transverse magnetic susceptibility χ_x shows in the vicinity of the critical point.

To obtain a better understanding of how the probability distribution given by the ground state $|g(\tau)\rangle$ changes with respect to τ we shall use the relative entropy (also called Kullback-Leibler divergence) which is often employed as a measure of similarity between probability distributions.

For discrete probability distributions on the same probability space \mathcal{X} the relative entropy is defined as

$$D_{\text{KL}}(P \parallel Q) = - \sum_{x \in \mathcal{X}} P(x) \log \left(\frac{Q(x)}{P(x)} \right). \quad (6.5)$$

In terms of our model, the ground state $|g(\tau)\rangle$ gives rise to the probability distribution $\mathbf{p}(g)$, which is a stochastic vector with entries $p_i(g)$, where i varies from 0 to $2^L - 1$. Remember that the τ -axis is discretized, so we have the set of parameters $\{\tau_i\}_{i=1}^D$. We would like to look at the behaviour of the following function with respect to τ :

$$D_{\text{KL}}(\tau_j, \tau_k) = - \sum_{i=1}^{2^L} p_i(\tau_j) \log \left(\frac{p_i(\tau_k)}{p_i(\tau_j)} \right). \quad (6.6)$$

In what follows, we use $k = j + 1$ for simulations. To the best of our knowledge, there is no widely-known definition of the quantum phase transition width. That is why for susceptibilities and the relative entropies we looked at the width of the hill at the 0.9 for the susceptibility and 0.7 for the entropy of the peak level, see Figures 6.5, 6.9 and 6.10.

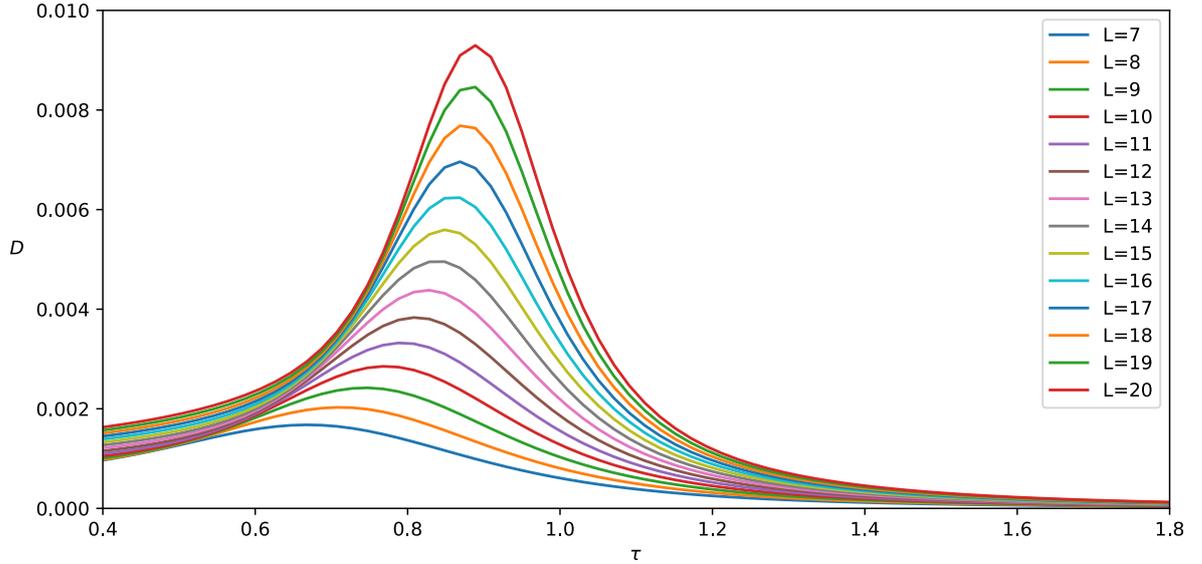


Figure 6.9: Relative entropy D between the ground state at τ_j and τ_{j+1} against the transverse field τ for various lengths of the chain.

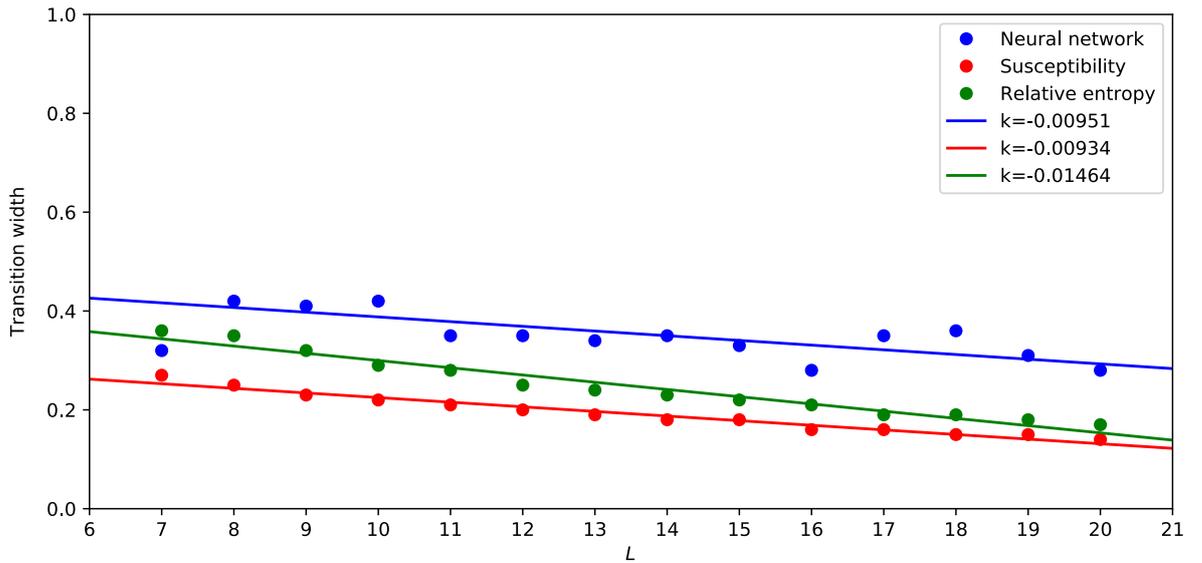


Figure 6.10: Crossover widths obtained by the neural-network method, magnetic susceptibility and relative entropy for finite-size chains.

6.2 Tilted field Ising model

Concerning the transverse field Ising model in the presence of a longitudinal field we examined the behaviour of a chain of $L = 10$ spins. We calculated the energy gap Δ as a function of both transverse and longitudinal external magnetic fields τ and λ . Furthermore, we looked at the magnetic properties of the system, i.e., the transverse magnetization m_x and

the longitudinal magnetization m_z . As for the neural-network approach to this model, we took the following strategy: we fixed the transverse field τ and scanned the ground state distributions $g(\lambda)$ with respect to λ in order to obtain $c(\lambda)$. We varied the boundaries of scanning, i.e., experimented with them until the output displayed a tangentoid curve, as for the quantum phase transition in TFIM.

All the calculations have been performed using exact diagonalization, the code is available in [67].

In Figure 6.11 one might spot the phase boundaries by observing the crossover of the energy gap Δ from 0 to a non-zero value. Indeed, the color map seems to mimic the phase diagram obtained in [32].

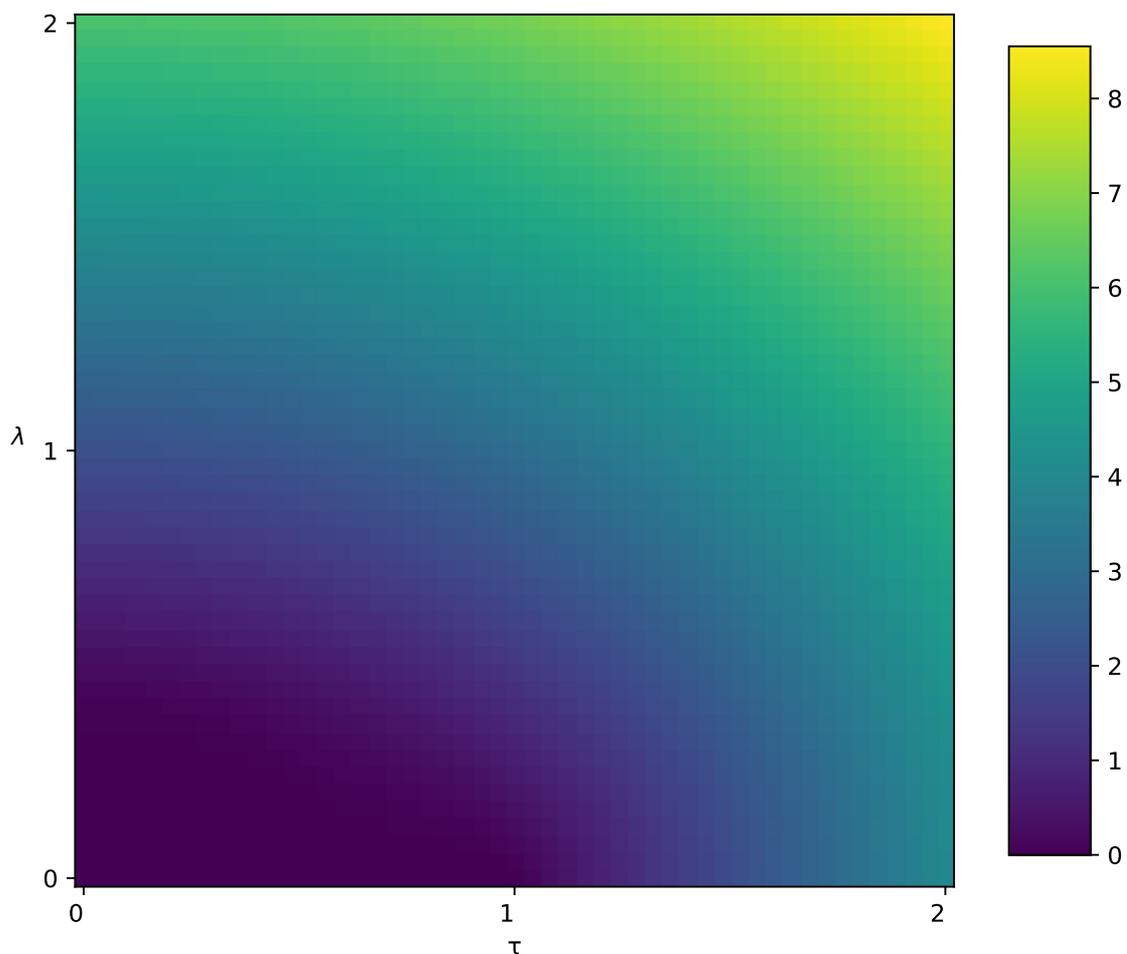


Figure 6.11: Energy gap Δ for the TFIM in longitudinal field on a chain of $L = 10$ spins versus τ and λ .

As for the magnetizations, it is clear that no phase boundaries can be observed from Figures 6.12.

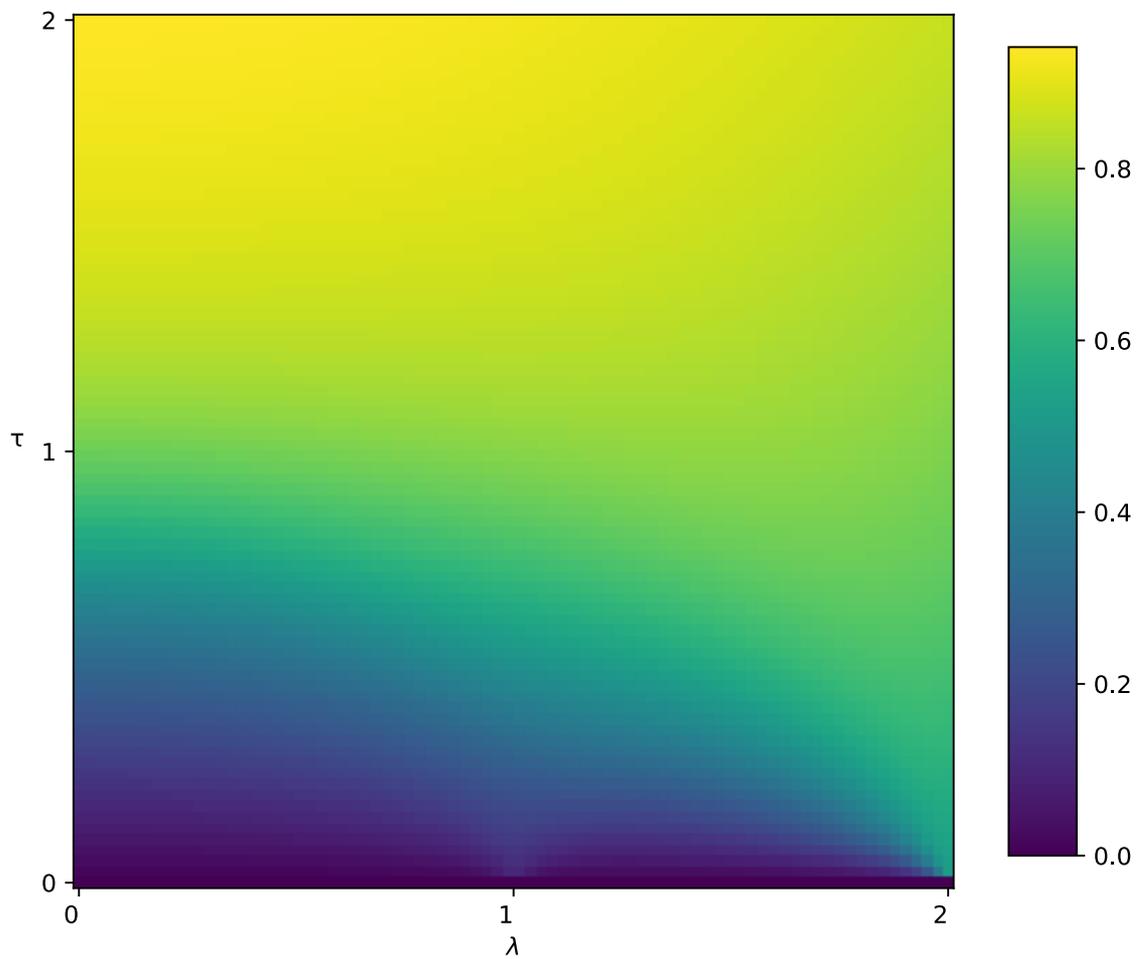


Figure 6.12: The transverse magnetization m_x as a function of τ and λ for a chain of $L = 10$.

In Figure 6.13 we show a typical neural-network output for $L = 12$ and $\tau = 0.5$.

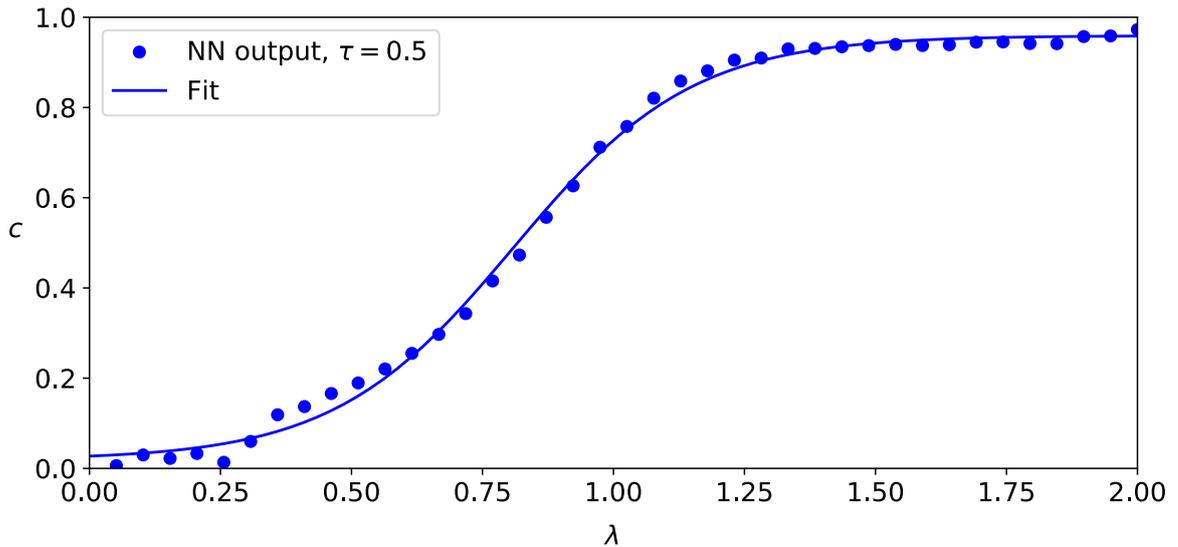


Figure 6.13: An example of a typical neural network output c for $L = 12$ and $\tau = 0.5$ versus longitudinal magnetic field λ .

By picking different values of τ from 0.1 to 1.0 with step size of 0.1, we obtained the following phase diagram, depicted in Figure 6.14.

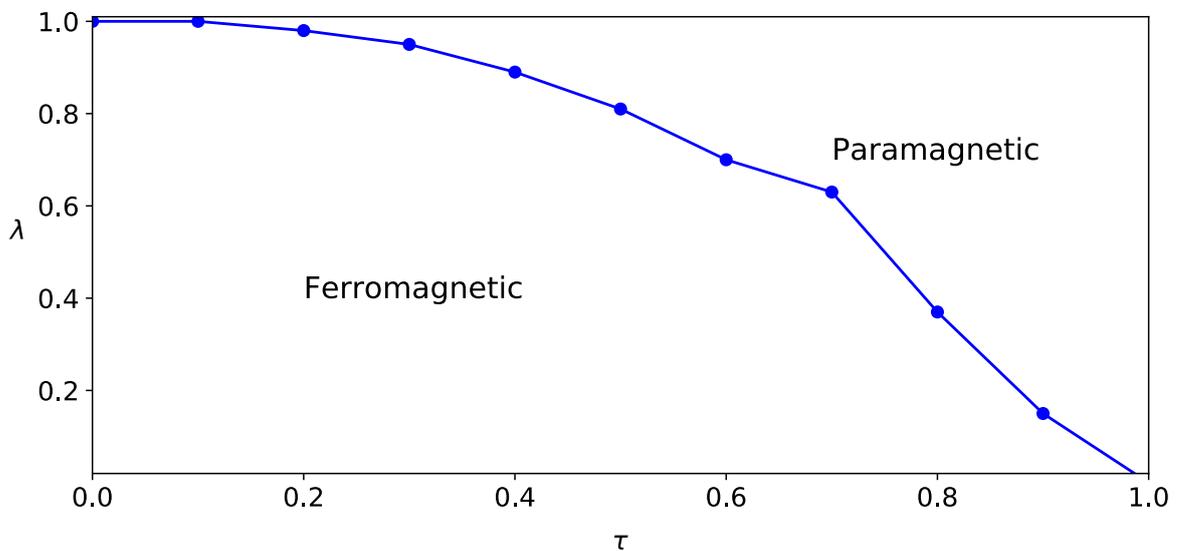


Figure 6.14: The phase diagram of the Ising model in external longitudinal magnetic field obtained by the neural-network method for $L = 12$.

6.3 Anisotropic XY model

In Figure 6.15, we show the result for an anisotropic XY chain of $L = 20$ spins. In this plot, one can clearly see the phase crossover induced by the change of γ which is a sign of a

well-studied anisotropy-induced phase transition in an infinite system [72], similarly to the phase transition induced by the critical value of the magnetic field. Noteworthy, while our algorithm is given information about the z -components of spins, it is capable of exposing a phase crossover induced by the anisotropy in the x - y plain.

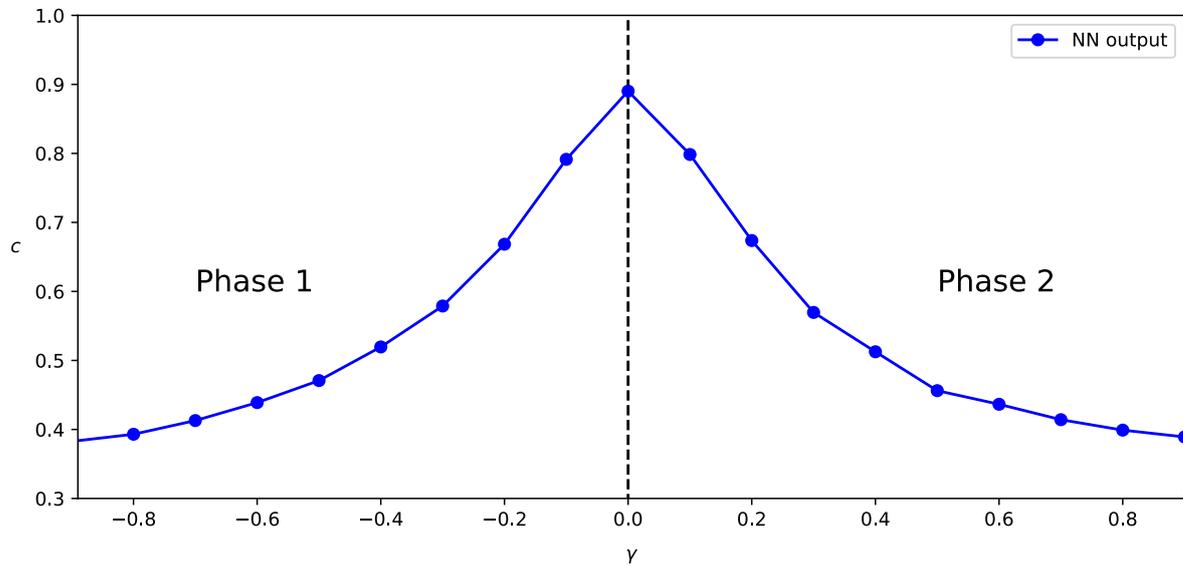


Figure 6.15: The neural network output as a function of the anisotropy parameter γ for $L = 20$ spins on an anisotropic XY chain with open boundary conditions.

Chapter 7

Summary and conclusion

In this work, we considered a simple neural network architecture with no hidden layers present and applied this to study finite-size phase crossovers in transverse field interacting spin models and the anisotropic XY model on a one-dimensional chain. We were able to distinguish the regions of different phases using neural networks *without* prior knowledge of the phase diagram by observing the corresponding phase boundary crossover in a finite-size system. We managed to almost exactly extract the critical value of the transverse magnetic field separating regions of magnetically ordered and disordered phases for the transverse spin models. Relative simplicity of the machine learning setup allowed us to visualize the weights of the corresponding neural network and unambiguously relate this plot to configuration of different spin orderings. In the presence of the longitudinal magnetic field we were capable of correctly reproducing the phase diagram, while observing the quantum criticality which is in line with the previous results. Besides that, we proved the presence of quantum criticality via several methods. First of all, it is the energy gap, which serves as the explicit order parameter for the transverse field Ising model. Secondly, we looked at the transverse magnetization and the transverse magnetic susceptibility. The latter displays a smooth crossover (hill) near the quantum phase transition region. Finally and most importantly, our neural-network method exhibited a tangentoid-type curve which made us able to define the quantum critical point and the width of its neighborhood using a proper fitting function.

The results of this study have been published as a preprint [73] which is currently awaiting publication in *Journal of Physics: Complexity*.

Appendix A

Tables

Table A.1: Variables and definitions

Variable	Definition	Description
L	Size of the spin chain.	Number of sites of the graph on which the Hamiltonian is defined on.
Δ	Energy gap.	The difference between the ground state energy and the energy of the first excited state.
τ	Transverse magnetic field.	A parameter for transverse field Ising model.
λ	Longitudinal magnetic field.	A parameter for transverse field Ising with longitudinal magnetic field.
$ g(p)\rangle$	Ground state.	A vector that corresponds to the lowest eigenvalue of a Hamiltonian controlled by a parameter p .
σ_i^x	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	The operator for an x -projection of the spin on site i .
σ_i^y	$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$	The operator for a y -projection of the spin on site i .
σ_i^z	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	The operator for a z -projection of the spin on site i .
m_α	$\frac{1}{L} \langle \sum_{i=1}^L \sigma_i^\alpha \rangle$	Average magnetization in the α -direction, where $\alpha = x, y, z$ and averaging is performed over a quantum state.
χ_α	$\frac{\partial}{\partial \tau} m_\alpha$	Average magnetic susceptibility in the α -direction, where $\alpha = x, y, z$.
D	Discretization value.	The total number of equidistant points that we define on a parameter axis.
N_{samples}	Total number of samples.	$N \times L$ array of bitstrings used for training and testing a machine learning model.
S	$N_{\text{test}}/N_{\text{samples}}$	A split ratio between testing and training data.
N_{train}	$(1 - S)N_{\text{samples}}$	Number of samples used for training a machine learning model.
N_{test}	SN_{samples}	Number of samples used for testing a machine learning model.
c	Neural network output.	The mean output of the neural network evaluated on the samples drawn from the probability distribution given by the ground state of the corresponding Hamiltonian.
\mathbf{W}	Weights.	Learnable parameters of a machine learning model.
\mathbf{b}	Biases.	Learnable parameters of a machine learning model.
i	$i^2 = -1$	The imaginary unit.

Bibliography

- [1] S. Sachdev. Quantum phase transitions. *Handbook of Magnetism and Advanced Magnetic Materials*, 2007.
- [2] L. Onsager. Crystal statistics I: a two-dimensional model with an order-disorder transition. *Physical Review*, 65(3-4):117, 1944.
- [3] H. V. Löhneysen, T. Pietrus, G. Portisch, H. G. Schlager, A. Schröder, M. Sieck, and T. Trappmann. Non-Fermi-liquid behavior in a heavy-fermion alloy at a magnetic instability. *Physical Review Letters*, 72(20):3262, 1994.
- [4] M. Neumann, J. Nyéki, B. Cowan, and J. Saunders. Bilayer 3He: a simple two-dimensional heavy-fermion system with quantum criticality. *Science*, 317(5843):1356–1359, 2007.
- [5] D. Bitko, T. F. Rosenbaum, and G. Aeppli. Quantum critical behavior for a model magnet. *Physical Review Letters*, 77(5):940, 1996.
- [6] R. M. Potok, I. G. Rau, H. Shtrikman, Y. Oreg, and D. Goldhaber-Gordon. Observation of the two-channel Kondo effect. *Nature*, 446(7132):167, 2007.
- [7] I. Bloch, J. Dalibard, and W. Zwerger. Many-body physics with ultracold gases. *Reviews of Modern Physics*, 80(3):885, 2008.
- [8] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, and M. Lanctot. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484, 2016.
- [9] V. Dunjko and H. J. Briegel. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Reports on Progress in Physics*, 81(7):074001, 2018.
- [10] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová. Machine learning and the physical sciences. *Reviews of Modern Physics*, 91(4):045002, 2019.
- [11] J. Carrasquilla. Machine learning for quantum matter. *arXiv preprint arXiv:2003.11040*, 2020.
- [12] J. Carrasquilla and R. G. Melko. Machine learning phases of matter. *Nature Physics*, 13(5):431, 2017.
- [13] D. Kim and D.-H. Kim. Smallest neural network to learn the Ising criticality. *Physical Review E*, 98(2):022138, 2018.

- [14] A. Tanaka and A. Tomiya. Detection of phase transition via convolutional neural networks. *Journal of the Physical Society of Japan*, 86(6):063001, 2017.
- [15] G. Carleo and M. Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.
- [16] G. Carleo, Y. N., and M. Imada. Constructing exact representations of quantum many-body systems with deep neural networks. *Nature Communications*, 9(1):1–11, 2018.
- [17] R. G. Melko, G. Carleo, J. Carrasquilla, and J. I. Cirac. Restricted Boltzmann machines in quantum physics. *Nature Physics*, 15(9):887–892, 2019.
- [18] S. Efthymiou, M. J. S. Beach, and R. G. Melko. Super-resolving the Ising model with convolutional neural networks. *Physical Review B*, 99(7):075113, 2019.
- [19] P. Zhang, H. Shen, and H. Zhai. Machine learning topological invariants with neural networks. *Physical Review Letters*, 120(6):066401, 2018.
- [20] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko, and J. Carrasquilla. Recurrent neural network wavefunctions. *arXiv preprint arXiv:2002.02973*, 2020.
- [21] M. J. S. Beach, I. De Vlucht, A. Golubeva, P. Huembeli, B. Kulchytskyy, R. Luo, R. G. Melko, E. Merali, and G. Torlai. Qucumber: wavefunction reconstruction with neural networks. *SciPost Physics*, 7(1):009, 2019.
- [22] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita. Reconstructing quantum states with generative models. *Nature Machine Intelligence*, 1(3):155–161, 2019.
- [23] G. Torlai and R. G. Melko. Latent space purification via neural density operators. *Physical Review Letters*, 120(24):240503, 2018.
- [24] I. A. Luchnikov, A. Ryzhov, P.-J. Stas, S. N. Filippov, and H. Ouerdane. Variational autoencoder reconstruction of complex many-body physics. *Entropy*, 21(11):1091, 2019.
- [25] S. Arai, M. Ohzeki, and K. Tanaka. Deep neural network detects quantum phase transition. *Journal of the Physical Society of Japan*, 87(3):033001, 2018.
- [26] S. A. Pikin and V. M. Tsukernik. The thermodynamics of linear spin chains in a transverse magnetic field. *Journal of Theoretical and Experimental Physics*, 23(5), 1966.
- [27] U. Brandt and K. Jacoby. The transverse correlation function of anisotropic XY chains: exact results at $T = \infty$. *Zeitschrift für Physik B Condensed Matter*, 26(3):245–252, 1977.
- [28] S. Katsura. Statistical mechanics of the anisotropic linear Heisenberg model. *Physical Review*, 127(5):1508, 1962.
- [29] P. G. De Gennes. Collective motions of hydrogen bonds. *Solid State Communications*, 1(6):132–137, 1963.
- [30] M. Suzuki. Generalized Trotter’s formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Communications in Mathematical Physics*, 51(2):183–190, 1976.

- [31] R. Coldea, D. A. Tennant, E. M. Wheeler, E. Wawrzynska, D. Prabhakaran, M. Telling, K. Habicht, P. Smeibidl, and K. Kiefer. Quantum criticality in an Ising chain: experimental evidence for emergent E8 symmetry. *Science*, 327(5962):177–180, 2010.
- [32] O. F. de Alcantara Bonfim, B. Boechat, and J. Florencio. Ground-state properties of the one-dimensional transverse Ising model in a longitudinal magnetic field. *Physical Review E*, 99(1):012122, 2019.
- [33] S. Czischek, M. Gärttner, and T. Gasenzer. Quenches near Ising quantum criticality as a challenge for artificial neural networks. *Physical Review B*, 98(2):024311, 2018.
- [34] A. A. Ovchinnikov, D. V. Dmitriev, V. Y. Krivnov, and V. O. Chervanovskii. Antiferromagnetic Ising chain in a mixed transverse and longitudinal magnetic field. *Physical Review B*, 68(21):214406, 2003.
- [35] E. Lieb, T. Schultz, and D. Mattis. Two soluble models of an antiferromagnetic chain. *Annals of Physics*, 16(3):407–466, 1961.
- [36] Q. Luo, J. Zhao, and X. Wang. Fidelity susceptibility of the anisotropic XY model: the exact solution. *Physical Review E*, 98(2):022106, 2018.
- [37] A. Dutta, G. Aeppli, B. K. Chakrabarti, U. Divakaran, T. F. Rosenbaum, and D. Sen. *Quantum phase transitions in transverse field spin models: from statistical physics to quantum information*. Cambridge University Press, 2015.
- [38] M. Continentino. *Quantum scaling in many-body systems*. Cambridge University Press, 2017.
- [39] F. Becca and S. Sorella. *Quantum Monte Carlo approaches for correlated systems*. Cambridge University Press, 2017.
- [40] C. Lanczos. *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*. United States Governm. Press Office Los Angeles, CA, 1950.
- [41] S. R. White. Density matrix formulation for quantum renormalization groups. *Physical Review Letters*, 69(19):2863, 1992.
- [42] B. L. Hammond, W. A. Lester, and P. J. Reynolds. *Monte Carlo methods in ab initio quantum chemistry*, volume 1. World Scientific, 1994.
- [43] E. Y. Loh Jr, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar. Sign problem in the numerical simulation of many-electron systems. *Physical Review B*, 41(13):9301, 1990.
- [44] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst. Machine learning quantum phases of matter beyond the fermion sign problem. *Scientific Reports*, 7(1):8823, 2017.
- [45] S. Greydanus. *Approximating matrix product states with machine learning*. PhD thesis, Dartmouth college, 2017.
- [46] A. Morningstar and R. G. Melko. Deep learning the Ising model near criticality. *The Journal of Machine Learning Research*, 18(1):5975–5991, 2017.

- [47] L. Wang. Discovering phase transitions with unsupervised learning. *Physical Review B*, 94(19):195105, 2016.
- [48] P. Broecker, F. F. Assaad, and S. Trebst. Quantum phase recognition via unsupervised machine learning. *arXiv preprint arXiv:1707.00663*, 2017.
- [49] P. Huembeli, A. Dauphin, and P. Wittek. Identifying quantum phase transitions with adversarial neural networks. *Physical Review B*, 97(13):134109, 2018.
- [50] E. Van Nieuwenburg, Y.-H. Liu, and S. D. Huber. Learning phase transitions by confusion. *Nature Physics*, 13(5):435, 2017.
- [51] M. Vojta. Quantum phase transitions. *Reports on Progress in Physics*, 66(12):2069, 2003.
- [52] J. Tsuda, Y. Yamanaka, and H. Nishimori. Energy gap at first-order quantum phase transitions: an anomalous case. *Journal of the Physical Society of Japan*, 82(11):114004, 2013.
- [53] L. C. Venuti and P. Zanardi. Quantum critical scaling of the geometric tensors. *Physical Review Letters*, 99(9):095701, 2007.
- [54] B. Damski. Fidelity approach to quantum phase transitions in quantum Ising model. In *Quantum Criticality in Condensed Matter: Phenomena, Materials and Ideas in Theory and Experiment*, pages 159–182. World Scientific, 2016.
- [55] B. Hu. The classical Ising model: A quantum renormalization group approach. *Physics Letters A*, 71(1):83–86, 1979.
- [56] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac. Classical simulation of infinite-size quantum lattice systems in two spatial dimensions. *Physical Review Letters*, 101(25):250602, 2008.
- [57] G. Evenbly and G. Vidal. Tensor network renormalization. *Physical Review Letters*, 115(18):180405, 2015.
- [58] E. Barouch and B. M. McCoy. Statistical mechanics of the XY model II: spin-correlation functions. *Physical Review A*, 3(2):786, 1971.
- [59] G. Hinton, N. Srivastava, and K. Swersky. Neural networks for machine learning: overview of mini-batch gradient descent (lecture 6a). *Coursera lecture slides*, 2012.
- [60] M. Suzuki. *Quantum Monte Carlo methods in condensed matter physics*. World scientific, 1993.
- [61] G. Rossum. Python reference manual. Technical report, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, The Netherlands, 1995.
- [62] Travis O. NumPy: A guide to NumPy. USA: Trelgol Publishing, 2006. [Online].
- [63] Eric J., Travis O., Pearu P., et al. SciPy: Open source scientific tools for Python, 2001. [Online].
- [64] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.

- [65] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, 12(Oct):2825–2830, 2011.
- [66] F. Chollet et al. Keras. <https://keras.io>, 2015.
- [67] A. V. Berezutskii. ML_Ising_XY github repository. https://github.com/meandmytram/ML_Ising_XY, 2019.
- [68] M. Born. Quantenmechanik der stoßvorgänge. *Zeitschrift für Physik*, 38(11-12):803–827, 1926.
- [69] E. Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik A Hadrons and Nuclei*, 31(1):253–258, 1925.
- [70] G. Montavon, W. Samek, and K.-R. Müller. Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 73:1–15, 2018.
- [71] K. Liu, J. Greitemann, and L. Pollet. Learning multiple order parameters with interpretable machines. *Physical Review B*, 99(10):104410, 2019.
- [72] H. T. Quan. Finite-temperature scaling of magnetic susceptibility and the geometric phase in the XY spin chain. *Journal of Physics A: Mathematical and Theoretical*, 42(39):395002, 2009.
- [73] A. V. Berezutskii, M. E. Bektov, D. Yudin, Z. Zimborás, and J. D. Biamonte. Probing criticality in quantum spin chains with neural networks. *arXiv preprint arXiv:2005.02104*, 2020.