



Technische Universität München  
Max-Planck-Institut für Quantenoptik



---

# Tensor networks, conformal fields and machine learning: applications in the description of quantum many-body systems

---

Ivan H. M. Glasser

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

*Vorsitzender* Prof. Dr. Alexander Holleitner

*Prüfer der Dissertation* Hon.-Prof. Dr. Ignacio Cirac  
Prof. Dr. Nora Brambilla

Die Dissertation wurde am 30.08.2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 23.11.2018 angenommen.





Technische Universität München  
Max-Planck-Institut für Quantenoptik



---

**Tensor networks, conformal fields and machine  
learning: applications in the description of  
quantum many-body systems**

---

Ivan H. M. Glasser

*Supervisor*

Hon.-Prof. J. Ignacio Cirac, PhD



# Contents

<b>Abstract</b>	<b>ix</b>
<b>Zusammenfassung</b>	<b>xi</b>
<b>List of publications</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Tensor networks states and entanglement . . . . .	3
1.2 Quantum phase transitions . . . . .	5
1.3 The fractional quantum Hall effect . . . . .	7
1.4 Neural-network quantum states . . . . .	9
1.5 Classical and quantum machine learning . . . . .	10
1.6 Thesis structure . . . . .	11
<b>2 Infinite dimensional Matrix Product States and quantum criticality</b>	<b>15</b>
2.1 Introduction to tensor networks and Matrix Product States . . . . .	16
2.1.1 Tensor networks . . . . .	16
2.1.2 Matrix Product States . . . . .	17
2.1.3 Properties of Matrix Product States . . . . .	18
2.2 Infinite dimensional Matrix Product States for critical systems . . . . .	20
2.2.1 Critical states and Matrix Product States . . . . .	20
2.2.2 Construction of an infinite dimensional Matrix Product States from a free massless boson . . . . .	21
2.2.3 Computing properties of the state with Monte Carlo techniques	24
2.3 A quantum phase transition with infinite dimensional Matrix Product States . . . . .	26
2.3.1 Construction of an infinite dimensional Matrix Product State from a massive field . . . . .	26
2.3.2 Properties of the phase transition at zero temperature . . . . .	28
2.4 Infinite dimensional Matrix Product States at finite temperature . . . . .	29
2.5 From a model wave function to a model Hamiltonian . . . . .	32
2.6 Conclusion . . . . .	35
<b>3 Fractional quantum Hall lattice models from infinite dimensional tensor   networks</b>	<b>37</b>
3.1 The fractional quantum Hall effect and Laughlin's wave function . . . . .	38

3.2	Lattice effects on Laughlin wave functions and parent Hamiltonians . . . . .	41
3.2.1	Lattice Laughlin states at different lattice filling factors . . . . .	42
3.2.2	Properties of the lattice Laughlin states on the square lattice . . . . .	46
3.2.3	States at half-filling: from long-range order on the square lattice to topological order on frustrated lattices . . . . .	50
3.2.4	Edge states from a charge at infinity . . . . .	54
3.2.5	Derivation of parent Hamiltonians . . . . .	57
3.3	Quasielectrons and quasiholes in lattice fractional quantum Hall states . . . . .	63
3.3.1	Wave functions of quasiholes . . . . .	63
3.3.2	Braiding statistics of quasiholes . . . . .	64
3.3.3	Wave function, charge and density profile of the quasielectron . . . . .	65
3.3.4	Quasiparticles on the torus . . . . .	67
3.3.5	Exact parent Hamiltonians for states with quasiparticles . . . . .	68
3.3.6	Local Hamiltonians . . . . .	69
3.4	Moore-Read states on lattices . . . . .	71
3.4.1	Definition of lattice Moore-Read states . . . . .	71
3.4.2	The lattice states become Moore-Read states in the continuum limit . . . . .	73
3.4.3	Properties of the lattice Moore-Read states . . . . .	76
3.4.4	Parent Hamiltonians for the lattice Moore-Read states . . . . .	79
3.4.5	Local Hamiltonians in one and two dimensions . . . . .	83
3.5	Conclusion . . . . .	86
<b>4</b>	<b>Neural-network quantum states and tensor networks: relationship and application to chiral states</b> . . . . .	<b>89</b>
4.1	Variational Monte Carlo with tensor networks and neural-network quantum states . . . . .	90
4.1.1	Variational Monte Carlo and Stochastic Reconfiguration . . . . .	91
4.1.2	Variational Monte Carlo method with tensor networks . . . . .	93
4.1.3	Variational Monte Carlo method with neural networks . . . . .	96
4.2	Relationship between tensor-network and neural-network states . . . . .	98
4.2.1	Jastrow wave functions, restricted Boltzmann machines and the Majumdar-Gosh model . . . . .	98
4.2.2	Short-range restricted Boltzmann machines are Entangled Plaquette States . . . . .	100
4.2.3	Fully-connected restricted Boltzmann machines are String-Bond States . . . . .	101
4.2.4	Generalizing restricted Boltzmann machines to non-local String-Bond States . . . . .	102
4.3	Application to chiral topological states . . . . .	105
4.3.1	Restricted Boltzmann machines can describe a Laughlin state exactly . . . . .	105
4.3.2	Numerical approximation of a Laughlin state . . . . .	106

4.3.3	Numerical approximation of a chiral spin liquid . . . . .	107
4.4	Conclusion . . . . .	110
<b>5</b>	<b>Machine learning with generalized tensor networks</b>	<b>113</b>
5.1	Introduction to machine learning . . . . .	114
5.2	Probabilistic graphical models . . . . .	115
5.3	Generalized tensor networks . . . . .	117
5.4	Supervised learning algorithm . . . . .	121
5.5	Learning feature vectors of data . . . . .	124
5.6	Application to image and sound classification . . . . .	127
5.6.1	Image classification . . . . .	127
5.6.2	Environmental sound classification . . . . .	129
5.7	Conclusion . . . . .	131
<b>6</b>	<b>Conclusion and outlook</b>	<b>133</b>
<b>A</b>	<b>Parent Hamiltonians for lattice Laughlin states</b>	<b>137</b>
A.1	Operators annihilating the $\psi_q^1[p]_\infty$ wave functions . . . . .	137
A.2	Operators annihilating the $\psi_4^2$ wave function . . . . .	139
A.3	Conformal transformations of the parent Hamiltonians . . . . .	141
<b>B</b>	<b>Parent Hamiltonians for lattice Moore-Read states</b>	<b>143</b>
B.1	Operators annihilating the $(q = 1, \eta = 1)_1$ lattice Moore-Read state .	143
B.2	Operators annihilating the $(q, \eta = 1)_1, q \geq 2$ lattice Moore-Read state	144
B.3	Operators annihilating the $(q, \eta = 1)_{1/2}, q \geq 2$ lattice Moore-Read state	148
	<b>Bibliography</b>	<b>151</b>
	<b>List of Acronyms</b>	<b>167</b>
	<b>List of Figures</b>	<b>169</b>
	<b>List of Tables</b>	<b>179</b>
	<b>Acknowledgments</b>	<b>181</b>





# Abstract

This thesis is devoted to the application of tensor-network methods to problems in quantum many-body physics in one and two dimensions and in machine learning.

In the first part we contribute to the theory of infinite dimensional Matrix Product States, for which the wave function of a lattice system takes the form of a correlator of field operators. We show that these states can be defined not only from conformal fields, but also from fields of a massive field theory. These states have analytical wave functions, which allows us to construct a model of a quantum phase transition and to study its properties using Monte Carlo simulations. We extend the application of infinite dimensional Matrix Product States to density operators and investigate the finite temperature phase diagram above the quantum phase transition.

We then turn to the application of infinite dimensional tensor networks in two dimensions. The corresponding states are lattice versions of fractional quantum Hall wave functions and allow us to build exact models realizing the same physics. We investigate lattice effects on these wave functions, with a focus on generalizations of the Kalmeyer-Laughlin and Moore-Read states, and characterize when topological order is present along an interpolation between lattice and continuum limit. We construct both quasihole and quasielectron excitations for these states and study their anyonic properties on the lattice. Exact non-local parent Hamiltonians for which these states are ground states are derived, and in some cases we find local Hamiltonians for which the ground state is well approximated by the corresponding infinite dimensional tensor network.

Studying these local Hamiltonians on large sizes is challenging. For this reason we investigate the power of different classes of states, such as states for which the wave function is represented by an artificial neural network, in approximating these chiral spin liquids. We show that neural-network quantum states taking the form of restricted Boltzmann machines are a subclass of String-Bond States, a kind of tensor network, and discuss the implications of this connection with respect to the underlying architecture of restricted Boltzmann machines. We demonstrate that neural-network quantum states and their tensor-network generalizations are able to represent certain lattice fractional quantum Hall states exactly and we numerically compare their ability at approximating a chiral spin liquid.

Finally we show that the connection between Boltzmann machines and tensor networks generalizes to probabilistic graphical models and has applications in machine learning. It motivates the definition of generalized tensor networks, which include String-Bond States as special cases and can be combined with neural-network architectures. We provide an algorithm to train these networks in the context of supervised learning and discuss its application to real-valued data. This algorithm is benchmarked for several generalized tensor-network architectures on the task of classifying images and recognizing environmental sounds, and we show that it outperforms previously introduced tensor-network algorithms.

# Zusammenfassung

Diese Dissertation widmet sich der Anwendung von Tensornetzwerkmethoden auf Fragen der Quantenvielteilchentheorie in ein und zwei Dimensionen sowie des maschinellen Lernens.

Im ersten Teil dieser Arbeit beschäftigen wir uns mit der Theorie der unendlich-dimensionalen Matrixproduktzustände, bei denen die Wellenfunktion eines Spinsystems auf einem Gitter die Form eines Korrelators von Feldoperatoren annimmt. Wir zeigen, dass sich diese Zustände nicht nur durch konforme Felder, sondern auch durch Felder einer massiven Feldtheorie definieren lassen. Diese Zustände haben analytische Wellenfunktionen, die es uns ermöglichen, ein Modell eines Quantenphasenübergangs zu erstellen und seine Eigenschaften mit Hilfe von Monte-Carlo-Simulationen zu untersuchen. Wir erweitern die Anwendung von unendlich-dimensionalen Matrixproduktzuständen auf Dichteoperatoren und studieren das Phasendiagramm bei endlicher Temperatur oberhalb des Quantenphasenübergangs.

Anschließend widmen wir uns der Anwendung von unendlich-dimensionalen Tensornetzwerken auf Systeme in zwei Dimensionen. Diese Zustände erlauben es uns, die Wellenfunktionen fraktionaler Quanten-Hall (FQH) Systeme auf Gittersysteme zu verallgemeinern und exakte Modelle mit denselben wesentlichen Eigenschaften wie im Kontinuum zu erhalten. Wir untersuchen Gittereffekte dieser Wellenfunktionen, mit einem Fokus auf Verallgemeinerungen der Kalmeyer-Laughlin und Moore-Read Zustände, und charakterisieren, wann eine Interpolation zwischen Gitter und Kontinuum topologisch geordnet ist. Wir konstruieren sowohl Quasi-Loch- als auch Quasi-Elektron-Anregungen für diese Zustände und untersuchen deren anyonische Eigenschaften auf dem Gitter. Exakte nicht-lokale Hamiltonoperatoren, für welche diese Zustände Grundzustände sind, werden abgeleitet. In einigen Fällen finden wir sogar lokale Hamiltonoperatoren, deren Grundzustand durch das entsprechende unendlich-dimensionale Tensornetzwerk gut approximiert wird.

Dennoch stellt es eine große Herausforderung dar, diese lokalen Hamiltonoperatoren auf größeren Gittern zu studieren. Deshalb untersuchen wir die Eignung verschiedener Klassen von Zuständen für die Beschreibung solcher Quantenspinflüssigkeiten, wie zum Beispiel die Klasse der Zustände, deren Wellenfunktion durch ein

künstliches neuronales Netz repräsentiert wird. Wir zeigen, dass durch beschränkte Boltzmann-Maschinen definierte neuronale Netzzustände eine Unterklasse von String-Bond States, einer Art Tensornetzwerk, bilden. Die Implikationen dieser Verbindung in Bezug auf die zugrundeliegende Architektur beschränkter Boltzmann-Maschinen wird diskutiert. Wir beweisen, dass neuronale Netzzustände und ihre Tensornetzwerk-Verallgemeinerungen in der Lage sind, bestimmte FQH Zustände auf dem Gitter darzustellen und vergleichen ihre Fähigkeit, eine chirale Spinflüssigkeit numerisch zu approximieren.

Schließlich zeigen wir, dass die Verbindung zwischen Boltzmann-Maschinen und Tensornetzwerken auf probabilistische graphische Modelle verallgemeinert werden kann und Anwendungen im Bereich des maschinellen Lernens hat. Diese Verbindung motiviert die Definition von generalisierten Tensornetzwerken, welche String-Bond-States als Sonderfälle beinhalten und mit neuronalen Netzwerkarchitekturen kombiniert werden können. Wir präsentieren einen Algorithmus, um diese Netzwerke im Kontext des überwachten Lernens zu trainieren und diskutieren seine Anwendung auf reellwertige Daten. Dieser Algorithmus wird für mehrere verallgemeinerte Tensornetzwerkarchitekturen zur Klassifizierung von Bildern und zur Erkennung von Umgebungsgeräuschen bewertet und wir zeigen, dass er bereits bekannte Tensornetzwerkalgorithmen in seiner Leistungsfähigkeit übertrifft.

# List of publications

## Publications covered in this thesis

- [1] : Ivan Glasser, J. Ignacio Cirac, Germán Sierra and Anne E. B. Nielsen, ‘Construction of spin models displaying quantum criticality from quantum field theory’, *Nuclear Physics B* **886**, 63-74 (2014)  
*See chapter 2*
- [2] : Ivan Glasser, J. Ignacio Cirac, Germán Sierra and Anne E. B. Nielsen, ‘Exact parent Hamiltonians of bosonic and fermionic Moore-Read states on lattices and local models’, *New Journal of Physics* **17**, 082001 (2015)  
*See chapter 3*
- [3] : Ivan Glasser, J. Ignacio Cirac, Germán Sierra and Anne E. B. Nielsen, ‘Lattice effects on Laughlin wave functions and parent Hamiltonians’, *Physical Review B* **94**, 245104 (2016)  
*See chapter 3*
- [4] : Ivan Glasser, Nicola Pancotti, Moritz August, Ivan D. Rodriguez and J. Ignacio Cirac, ‘Neural-Network Quantum States, String-Bond States, and Chiral Topological States’, *Physical Review X* **8**, 011006 (2018)  
*See chapter 4*
- [5] : Anne E. B. Nielsen, Ivan Glasser and Ivan D. Rodriguez, ‘Quasielectrons as inverse quasiholes in lattice fractional quantum Hall models’, *New Journal of Physics* **20**, 033029 (2018)  
*See chapter 3*
- [6] : Ivan Glasser, Nicola Pancotti and J. Ignacio Cirac, ‘Supervised learning with generalized tensor networks’, arXiv:1806.05964 (2018)  
*See chapter 5*

## Other publications not covered in this thesis

- [7] : Laura Bernard, Alon E. Faraggi, Ivan Glasser, John Rizos and Hasan Sonmez, ‘String derived exophobic  $SU(6) \times SU(2)$  GUTs’, Nuclear Physics B **868**, 1-15 (2013)

# Introduction

” Observe whenever the rays are let in and pour the sunlight through the dark chambers of houses: you will see many minute bodies in many ways through the apparent void mingle in the midst of the light of the rays, and as in never-ending conflict skirmish and give battle combating in troops and never halting, driven about in frequent meetings and partings; so that you may guess from this what it is for the first-beginnings of things to be ever tossing about in the great void. So far as it goes, a small thing may give an illustration of great things and put you on the track of knowledge.

— **Titus Lucretius Carus**

(De Rerum Natura, 1<sup>st</sup> century BCE,  
translation by H. A. J. Munro)

Physics is pushing the boundaries of our knowledge, from the very small scale associated with the properties of elementary particles to the cosmological scale of the universe. But there is another frontier, ubiquitous in all areas of science: the complexity frontier. This complexity often manifests itself when the system consists of many interacting entities, such as molecules in fluid dynamics and meteorology, neurons in neurobiology or humans in sociology. In these cases collective phenomena that are not properties of the individual entities may play a crucial role. Such a complexity may appear even when all the single entities are simple and the interactions between them are governed by well defined physical laws, because we may not be able to solve the equations describing the system or identify the relevant emerging degrees of freedom.

One area where this property stands out is quantum physics. Consider a system of  $N$  classical particles. Each configuration of the system is fully determined by the positions and velocities of the particles. However, at atomic length scales the motion of the particles is no longer governed by classical mechanics, but by quantum mechanics. The position and momenta no longer have fixed values, but follow probability distributions resulting from the Schrödinger equation. Instead of considering one configuration of the system, all possible configurations must be taken into account, which renders the exact description of the system intractable for more

than a few particles[8]. Many interesting phenomena in condensed matter physics nevertheless arise through the interplay of interacting quantum particles, such as superconductivity or the fractional quantum Hall effect, rendering approaches to tackle the quantum many-body problem necessary. A quantum computer is another example of a complex many-body quantum system, and understanding which quantum systems can be efficiently simulated with classical resources is therefore of practical interest.

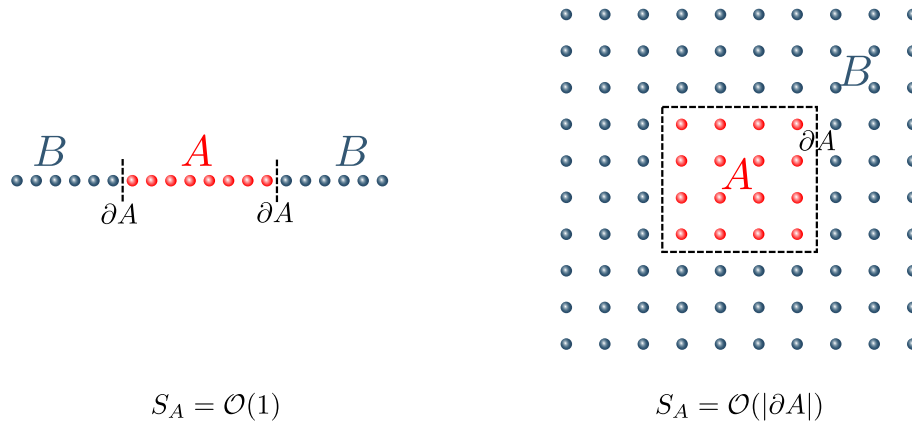
In this thesis we will be mainly concerned with the classical simulation of many-body quantum systems of spins or electrons placed on a lattice. Such systems can be seen as approximations of materials, but are also of independent interest since they can be realized artificially, for example with ultracold atoms in optical lattices[9]. Simulating a general system of  $N$  spin-1/2 particles on a classical computer requires storing  $2^N$  coefficients, rendering simulation of more than a few dozens of spins impossible even on the largest computers. Different techniques have been introduced to tackle larger systems. On the one hand, solvable models that are non-interacting or integrable provide examples for which the properties can be computed exactly even for large sizes[10]. These models have contributed to our understanding of both equilibrium and non-equilibrium quantum phenomena, but many systems of interest are interacting and not exactly solvable. On the other hand, numerical simulation techniques have been developed, such as Quantum Monte Carlo methods which sample the exponentially large space of configurations[11]. Quantum Monte Carlo methods have been highly successful but suffer from a fundamental limitation: the sign problem[12], which commonly appears in the simulation of fermions or frustrated spin systems.

Other methods thus have to be studied. One approach which enjoyed tremendous success, in particular in the simulation of one-dimensional quantum systems, is the Density Matrix Renormalization Group algorithm[13, 14]. The understanding of this algorithm in terms of tensor networks has since led to many insights about the structure of quantum states[15] as well as efficient algorithms to study them[16]. Tensor-network algorithms have the advantage, compared to Quantum Monte Carlo methods, that they do not suffer from a sign problem and can thus be used for studying problems that are not amenable to Quantum Monte Carlo. Moreover they also make simulation of dynamical systems possible and give direct access to a parametrization of the quantum wave function, which can be used to compute its properties. This thesis will be concerned with the development of new methods which extend the applicability of tensor networks to challenging problems in condensed matter physics as well as machine learning.



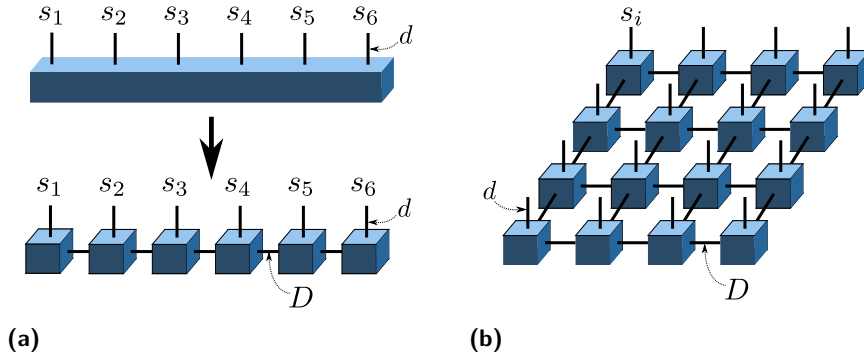
## 1.1 Tensor networks states and entanglement

We are interested in the properties of quantum states at low temperatures, where quantum effects play the most important role. These states live in a Hilbert space which grows exponentially with the size of the system. Nevertheless it is still expected that not all states in this Hilbert space are physical. Indeed many interactions in nature are local, thus resulting in Hamiltonians which are not arbitrary. The ground states of local Hamiltonians are not randomly distributed in the Hilbert space but form a tiny subset of it, the physical corner of the Hilbert space. Tensor networks are an attempt at parametrizing this subset.



**Fig. 1.1.:** Area law in one and two dimensions.

To characterize the locality of quantum states, a system is divided between a region  $A$  and a region  $B$  (Fig. 1.1). Intuitively, we expect that if interactions are local then the correlations between regions  $A$  and  $B$  are concentrated on the boundary between the two regions, while spins in the bulk of region  $A$  bring little contribution to such correlations because they are far from any spin in  $B$ . This statement can be made more rigorous. A measure of the amount of entanglement (quantum correlations) between regions  $A$  and  $B$  is the entanglement entropy  $S_A = -\text{Tr} \rho_A \log \rho_A$ , where  $\rho_A$  is the reduced density matrix of the system  $A$ . While increasing the size of region  $A$ , the entanglement entropy of a random state in the Hilbert space typically scales as the volume (the number of spins) in region  $A$ . But ground states of local Hamiltonians display a different behavior: indeed, in many cases one finds that the entanglement entropy  $S_A$  of the ground state of a local Hamiltonian scales as the boundary between systems  $A$  and  $B$ , a property known as area law:  $S_A = \mathcal{O}(|\partial A|)$ . Such an area law has been proven in many cases, such as for models for which there is an energy gap between the ground state and the excited states in 1D[17], for such gapped models satisfying some spectral conditions in arbitrary dimension[18], for free bosonic and fermionic models[19, 20], for frustration-free spin models[21] as well as for states that are in the same phase as a state satisfying the area law[22].



**Fig. 1.2.:** a) Decomposition of the general tensor representing the wave function into a one-dimensional tensor network known as Matrix Product State (MPS).  
b) In two dimensions, the corresponding tensor network is a Projected Entangled Pair State (PEPS).

Let us now see how this property is related to tensor networks. In general, a wave function of a system of  $N$  spins with  $d$  degrees of freedom can be written as

$$|\psi\rangle = \sum_{s_1, \dots, s_N=1}^d \psi^{s_1, \dots, s_N} |s_1\rangle \otimes |s_2\rangle \otimes \dots \otimes |s_N\rangle, \quad (1.1)$$

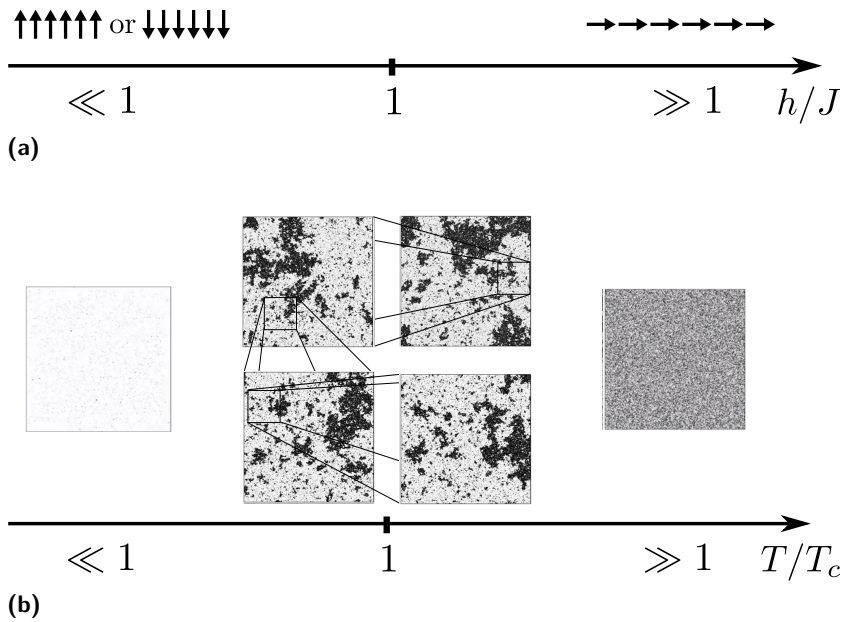
where  $\psi^{s_1, \dots, s_N}$  is a tensor containing  $d^N$  complex coefficients. Tensor network states[15, 23, 24] are quantum states for which this tensor has been decomposed into a network of local tensors that are contracted (Fig. 1.2). The idea of tensor network states is therefore to construct a global object (the wave function) from simple building blocks (tensors) which are local. Links between the tensors correspond to a sum over the corresponding indexes of the tensors. The maximum dimension of the tensor indexes which are contracted is called the bond dimension  $D$ . In one dimension, the resulting tensor network is a Matrix Product State (MPS) (Fig. 1.2a), while in two dimension one obtains a Projected Entangled Pair State (PEPS)[25](Fig. 1.2b). These tensor network states satisfy the area law when  $D$  is bounded. More importantly, it has been proven that any quantum state satisfying an area law in one dimension can be efficiently approximated by a MPS[26, 27] (this statement depends on the exact measure of entanglement considered), by which we mean that the corresponding MPS will have a number of parameters that scales polynomially, and not exponentially, with the system size. Together with the previous results on ground states of local Hamiltonians satisfying an area law, this means that MPS are the right parametrization for one-dimensional ground states of gapped Hamiltonians. In two dimensions, satisfying an area law is not a sufficient condition for having an efficient parametrization such as PEPS[28], but many states of interest can nevertheless be approximated by a PEPS, such as thermal states of local Hamiltonians[29].

## 1.2 Quantum phase transitions

The fact that tensor networks satisfy an area law also means that states with high entanglement can be difficult to approximate with tensor network states. One area where such states naturally appear are quantum phase transitions[30], which occur at zero temperature. Changing a parameter of the Hamiltonian, for example by applying a magnetic field, can bring the ground state of the system from one phase of matter to another. The transition between the two phases happens at a critical point corresponding to an abrupt change in the ground state wave function. In a conventional continuous phase transition, the properties at the critical point are associated with the symmetries on both sides of the transition. A local order parameter that is exactly zero in one phase and non-zero in the other phase can be defined, and the energy gap between the ground state and first excited state vanishes at the phase transition. The characteristic length scale of the system diverges as the critical point is approached, leading to power-law behavior of the correlation functions and scale invariance at criticality (Fig. 1.3). Because of scale invariance, the microscopic details of the system become irrelevant, which leads to the universality of phase transitions: different physical systems can behave in a similar way at the critical point and can be studied using the same scale-invariant theory.

In many cases, this theory is a conformal field theory (CFT)[32], which, in addition to being scale invariant, is invariant under transformations that preserve angles. The conformal invariance provides analytical tools to study these theories in great details. Consider for example a one-dimensional quantum system at a phase transition described by a CFT. Then one can compute the entanglement entropy of a reduced region  $A$  of the system and show that it scales as the logarithm of the size of region  $A$ [33–35]. This is a violation to the previously introduced area law. A MPS with fixed bond dimension satisfies the area law and has exponentially decaying correlations at long distance. Therefore it cannot be used to study such a critical system in the thermodynamic limit. Nevertheless, one can show that by increasing the bond dimension polynomially with the system size, MPS can in fact approximate critical systems efficiently[26]. Another approach is to build a tensor network with a different network structure as the original lattice. The multi-scale entanglement renormalization ansatz [36, 37] is such a tensor network which allows scale invariance to be satisfied and critical states to be constructed.

The limitation of MPS to represent critical systems stems from the finite bond dimension which limits the amount of entanglement that can be captured by the MPS. Another approach to represent critical systems that will be developed in this thesis are infinite dimensional MPS[38], in which the finite dimensional tensors have been replaced with infinite dimensional operators, such as field operators in a quantum field theory or conformal field theory. The wave function of such a state can



**Fig. 1.3.:** a) Phase diagram of the 1D transverse field Ising model at zero temperature. For low values of the transverse field  $h$  with respect to the Ising coupling  $J$  the ground state is a ferromagnet, while for high values of the transverse field the ground state is a paramagnet. In between there is a quantum phase transition at a particular value of the transverse field.  
 b) This model can be mapped to a 2D classical Ising model at finite temperature, which displays a phase transition at a critical temperature  $T_c$ . The spin configurations display scale invariance at the critical point. Monte Carlo simulation results by Douglas Ashton[31].

then be written as a correlator of field operators, and its properties can be studied using Monte Carlo simulations. Infinite dimensional MPS can support logarithmic violations of the area law as well as polynomially decaying correlations even for large systems, and are thus particularly well suited for describing critical systems.

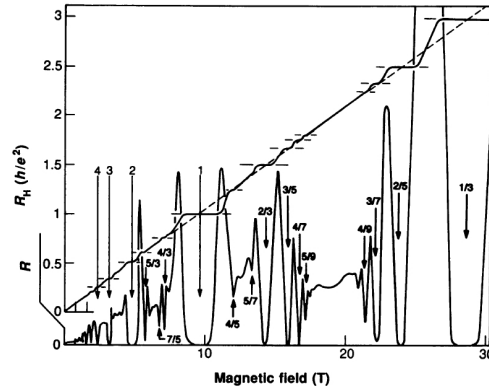
In chapter 2, we will construct a toy model of a quantum phase transition using infinite dimensional MPS. The construction of an infinite dimensional MPS from conformal fields gives rise to a critical state. We will extend it to fields which are not conformal and introduce a length scale in the system, and study the corresponding transition. Quantum phase transitions happen at strictly zero temperature and it may seem that they are not relevant to experiments taking place at small but non-zero temperature. We will show that infinite dimensional MPS can also be defined at finite temperature and use them to study how the quantum phase transition at zero temperature affects the properties of the system at finite temperature, a phenomenon known as quantum criticality[30].

## 1.3 The fractional quantum Hall effect

Not all phase transitions admit a local order parameter and can be fully characterized by symmetries. Indeed, some quantum systems display a new kind of order that goes beyond the conventional classification of phases of matter: topological order[39]. Topologically ordered states are long-range entangled: they cannot be transformed into a product state through a local unitary evolution (quantum circuit with finite depth), and their entanglement entropy displays a finite subleading correction to the area law known as topological entanglement entropy[40, 41]. Such systems are associated with a ground state degeneracy that depends on the topology of the space they are defined on. This degeneracy cannot be lifted by any local perturbation, so is not due to symmetry. This renders such states particularly suitable for storing quantum information in a way that is robust to noise[42]. The finite energy quasiparticle excitations over these ground states are neither bosons nor fermions, but anyons with fractional statistics that can be employed to perform fault-tolerant quantum computation[43, 44].

Physical systems realizing topological order are thus widely sought after, both for the fundamental physics they display as well as for their potential applications in quantum information processing. The fractional quantum Hall (FQH) effect, discovered in 1982[45], provides such an experimental realization in a 2D electron gas. By applying a strong magnetic field, the electrons form an incompressible quantum fluid that gives rise to fractionally charged excitations which are anyons. The magnetic field breaks time-reversal and parity symmetry, so the ground state is known as a chiral topological state and is associated with chiral edge states which have a preferred direction along the edge of the sample. This effect manifests itself in the quantization of the Hall conductance observed in the experiments at particular rational values of the ratio between the number of electrons and the number of magnetic fluxes (Fig. 1.4).

Recently, there has been a lot of interest in finding fractional quantum Hall physics in other systems. Indeed, experimental realizations of fractional quantum Hall states and manipulation of their quasiparticles remain a challenge. Finding other systems exhibiting the same physics is an important step towards a better understanding of these phenomena as well as practical applications in quantum computing. Lattice models hosting the same physics have been developed, such as for example fractional Chern insulators which mimic the physical interactions leading to the fractional quantum Hall effect on a lattice[47–51], or chiral spin liquids, which consists in an interacting system of spins on a lattice for which the ground state displays similar properties[52–62].



**Fig. 1.4.:** Hall resistance and longitudinal resistance of a 2D electron gas subject to a magnetic field. From [46], reprinted with permission from AAAS.

The understanding of the fractional quantum Hall effect was in large part made possible through the discovery of wave functions describing the electrons, such as Laughlin’s wave function[63]. To describe chiral spin liquids on a lattice exact model wave functions as well as numerical methods also need to be developed. Efforts to construct chiral topological states with tensor networks such as PEPS have been undertaken recently[64–67], but the resulting states are critical and are the ground state of a gapless local parent Hamiltonian, unlike what happens in the fractional quantum Hall effect.

Another approach is to construct chiral lattice wave functions from the fractional quantum Hall continuum wave functions, the paradigmatic example being the Kalmeyer-Laughlin wave function[52]. Other continuum wave functions which can be extended to lattices are wave functions written in terms of correlators of a CFT[68]. These lead to lattice wave functions that are infinite dimensional MPS[69]. The resulting quantum states can then be studied in great details and exact Hamiltonians for which they are ground states can be derived[70]. We note here that this approach is unlike traditional approaches in condensed matter physics: instead of being given a Hamiltonian and looking for the properties of its ground state, we first construct a state with desirable properties and then look for a Hamiltonian for which this state would be the ground state.

In chapter 3, we provide several contributions to the theory of infinite dimensional MPS realizing fractional quantum Hall physics. We study properties of generalizations of the Kalmeyer-Laughlin states on several lattices and provide a characterization of when chiral topological order is present. The excitations over these states are studied both in terms of localized quasiparticles and in terms of edge excitations. In particular we construct both quasihole and quasielectron excitations and study their anyonic properties on the lattice. A family of chiral spin liquids with non-Abelian excitations is discovered, which generalizes the Moore-Read[68] continuum

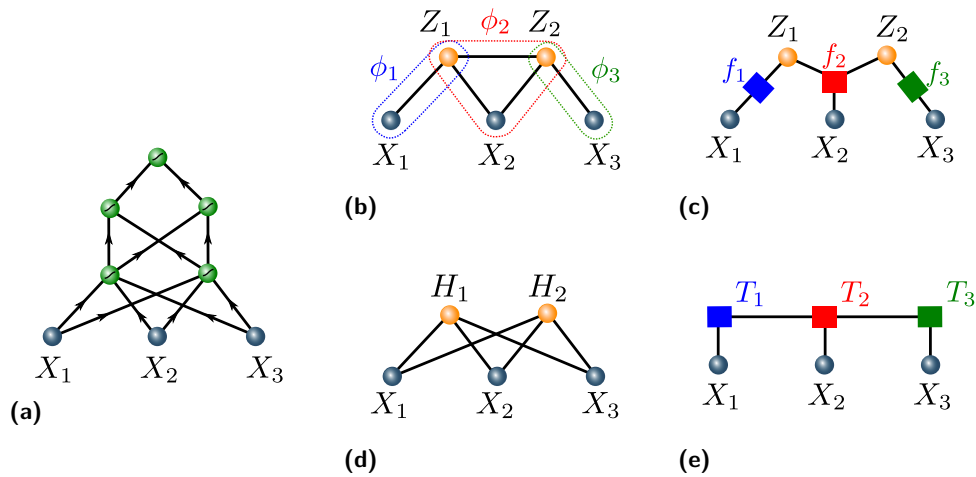
wave functions. Both exact non-local Hamiltonians as well as approximate local Hamiltonians that realize some of these states are constructed.

## 1.4 Neural-network quantum states

Infinite dimensional MPS allow us to construct many exact chiral topological models, but possess only few parameters and are thus not suitable as a variational ansatz to the ground state of a generic Hamiltonian. In particular, once an interesting exact model has been constructed it is often necessary to study other Hamiltonians that realize the same physics but that might be simpler or more suitable to experiments. Other numerical methods to study such systems on large sizes are therefore necessary to complement the use of infinite dimensional tensor networks. Recently, another class of states has been introduced as an ansatz for a many-body quantum state: neural-network quantum states[71] represent a wave function using an artificial neural network. Artificial neural networks have been introduced as an approximation to a function in the context of machine learning and have recently been at the center of tremendous advances in computer vision, natural language processing or robotics[72]. Inspired from biological neural networks, they consist of simple building blocks, neurons, which take as input the value of several variables and return as output a non-linear function of a weighted sum of their inputs. Arranged together in a network, they are able to model complex dependencies. The versatility of machine learning with neural networks has allowed scientists to employ it in a number of problems which span from quantum control[73–75] and error correcting codes[76] to tomography[77]. In condensed matter physics it has been applied to recognize quantum phases[78–81] or to enhance Quantum Monte Carlo methods [82, 83].

Particularly suited to the problem of representing wave functions of spin systems are tools designed to represent probability distributions with discrete variables, such as probabilistic graphical models[84]. These models include Boltzmann machines[85], which represent a probability distribution as the Boltzmann weight of an Ising Hamiltonian on a graph, as special cases. Boltzmann machines have recently been shown to be able to represent many quantum states of interest[71, 77, 86–89]. The structure of neural networks or graphical models, consisting of small local blocks linked together in a complex network, is reminiscent of the structure of tensor networks (Fig. 1.5). In fact, we will show that this connection can be made precise in a number of cases, and that this connection has implications both for quantum physics as well as for machine learning.

In chapter 4, we will discuss the connection between Boltzmann machines and classes of tensor networks known as Entangled Plaquette States (EPS)[90–92] and String-Bond States (SBS)[93, 94]. We will show that these tensor networks



**Fig. 1.5.:** Different network architectures based on graphs that can be used to define a function of the variables  $X_1$ ,  $X_2$  and  $X_3$ : a) a feed-forward neural network, b) an undirected probabilistic graphical model, c) a factor graph, d) a restricted Boltzmann machine, e) a tensor network (Matrix Product State).

provide a natural generalization of restricted Boltzmann machines, a special class of Boltzmann machines. We will apply these different classes of states to the problem of representing a chiral spin liquid previously discovered with infinite dimensional MPS methods. Through both analytical wave functions as well as numerical simulations, we will show that these networks allow us to get an accurate description of the ground state and to compute its properties.

## 1.5 Classical and quantum machine learning

So far we have discussed the application of machine learning techniques to study quantum systems, but applying quantum techniques to machine learning has also received a lot of attention[95, 96]. In supervised learning, data with labels is provided and one tries to find a function that maps the input data to its label. This function should generalize to previously unseen data, so it is not sufficient to memorize the training data and its labels. Performance is then evaluated on a test set which has not been used during training of the model. In unsupervised learning, data without labels is provided and one tries to find some structure in the data, for example through clustering, or by approximating its probability distribution. For large scale datasets, one needs to rely on approximations and an ansatz function is chosen to approximate the function or probability distribution given by the data. This ansatz function can for example be a neural network or a probabilistic graphical model.



Quantum circuits that may be implemented on near-term quantum devices can implement functions that cannot be computed efficiently on a classical computer, and these functions could be used as approximation to the desired function or probability distribution, in both supervised and unsupervised learning. Understanding how these circuits may be useful in machine learning attracted a lot of recent interest[97–102]. In this setting, the function evaluation is performed by a quantum device, but the optimization of its variational parameters is performed classically. Such algorithms are therefore known as variational hybrid quantum-classical algorithms.

Since tensor networks are tools to simulate quantum systems, they are a natural candidate to perform the classical simulation of quantum circuits. The circuits that can be simulated efficiently using classical methods do not require a quantum computer to run them. It is therefore of interest to understand how tensor networks can be used in machine learning, both in order to develop new classical machine learning algorithms[103–106] as well as to guide the development of future quantum machine learning architectures[107, 108].

In chapter 5, we will show that the previously observed connection between restricted Boltzmann machines and tensor networks generalizes to probabilistic graphical models and motivates the definition of generalized tensor networks. These networks rely on the copy and reuse of local tensor information and remain efficient to contract. We provide an algorithm to train these networks in a supervised learning context and show that they overcome the limitations of regular tensor networks in higher dimensions. Our algorithm is benchmarked for several generalized tensor network architectures on the task of classifying images and recognizing environmental sounds, and we show that it outperforms previously introduced tensor-network algorithms.

## 1.6 Thesis structure

The rest of this thesis is organized as follows:

### Chapter 2

In this chapter we first introduce the formalism of tensor networks with a focus on Matrix Product States (MPS) in section 2.1. We discuss the applicability of MPS to critical systems and motivate the definition of infinite dimensional MPS, where the tensor network is replaced by a correlator of field operators. A simple infinite dimensional MPS introduced in Ref. [38] to represent the ground state of a critical system is presented in section 2.2. The resulting wave function arises from the correlator of operators in a free massless boson quantum field theory. Our first contribution consists in the generalization of this infinite dimensional MPS to massive fields, which allows us to construct a toy model of a quantum phase

transition in section 2.3. In section 2.4 we then provide a way to represent these states at finite temperature by writing a purification of the system as an infinite dimensional MPS, and study how the quantum phase transition at zero temperature influences the phase diagram at finite temperature. We finally show in section 2.5 that our model could be realized in a realistic setting by constructing a Hamiltonian whose thermal state is close to our mixed state ansatz.

### Chapter 3

We then turn to the two dimensional case, where infinite dimensional MPS arising from operators in a conformal field theory lead to the definition of fractional quantum Hall lattice models. After introducing the fractional quantum Hall effect and properties of Laughlin states in the continuum, we study in section 3.2 generalizations of the Kalmeyer-Laughlin states on several lattices, with different number of particles per flux and particles per lattice site. We show that these states have the same topological properties as the corresponding continuum Laughlin states except on some particular lattices when the lattice is half-filled. We construct analytical wave functions which correspond to edge states for these models and derive exact parent Hamiltonians for which the infinite dimensional MPS is a ground state for different fillings of the lattice. In section 3.3, we turn to the investigation of the quasiparticle excitations of these states. While the construction of a quasielectron wave function in the continuum remains more difficult than for the quasihole, we show that the quasielectron wave function can be defined on the lattice in a similar way as for the quasihole. This allows us to study properties of the quasiholes and quasielectrons in a vast range of models that interpolate between the continuum and the lattice limit. Exact Hamiltonians having quasiparticle states as their ground states are constructed. In these models, braiding of quasiparticles can be realized simply by changing the value of the couplings in the Hamiltonian. Section 3.4 finally extends the lattice models displaying the fractional quantum Hall effect to models having non-Abelian anyons as their quasiparticles. We introduce a family of states which correspond to lattice versions of the Moore-Read wave functions. We provide numerical evidence that the topological properties of these states remain the same along an interpolation between the continuum and the lattice, and derive exact parent Hamiltonians for these states on lattices of arbitrary size. By deforming these parent Hamiltonians, we construct local Hamiltonians that stabilize some of the states we introduce in one and two dimensions.

### Chapter 4

In this chapter we focus on the variational optimization of tensor-network states and neural-network quantum states. We first introduce in section 4.1 the Variational Monte Carlo method that can be used to optimize both tensor networks and neural networks. We then present in section 4.2 our results showing that there are strong connections between neural-network quantum states in the form of restricted Boltzmann machines and some classes of tensor-network states in arbitrary

dimensions. We demonstrate that short-range restricted Boltzmann machines are Entangled Plaquette States, while fully connected restricted Boltzmann machines are String-Bond States with a non-local geometry. These results allow us to understand the underlying architecture of restricted Boltzmann machines and their efficiency at representing many-body quantum states. In section 4.3 we turn to the application of these classes of states to the problem of representing the ground state of lattice Hamiltonians introduced in chapter 3. We give analytical evidence that neural-network quantum states and their non-local tensor-network generalizations are able to represent lattice fractional quantum Hall states exactly. Furthermore we numerically compare these different classes of states and show that some of them are able to approximate a chiral spin liquid with high accuracy.

## Chapter 5

After having used machine learning techniques to approximate quantum states, we turn to the application of tensor network techniques introduced in quantum physics to perform machine learning. In section 5.2 we introduce probabilistic graphical models and discuss their connections with tensor networks. We show that this connection motivates the definition of generalized tensor networks, which include Entangled Plaquette States and String-Bond States as special cases. We prove that these generalized tensor networks can represent some functions with exponentially less parameters than regular tensor networks, and show how they can be used in conjunction with neural networks. An algorithm for training these networks in a supervised learning context is provided in section 5.4, and its applicability to data which is real and not discrete is discussed in section 5.5. Finally we benchmark our algorithm in section 5.6 for several generalized tensor network architectures on the task of classifying images and recognizing environmental sounds, showing that it outperforms previously introduced tensor-network algorithms.



# Infinite dimensional Matrix Product States and quantum criticality

Quantum phase transitions are phase transitions happening at zero temperature and driven by quantum instead of thermal fluctuations. They can happen for example by tuning an external magnetic field, the pressure or the composition of a material. At a conventional critical point, the ground state displays fluctuations over all length scales, is highly entangled and its correlations decay polynomially with the system size. While the quantum phase transition happens at zero temperature, it is still relevant for experiments at finite temperature. Indeed the properties of the system at finite temperature are influenced by the presence of the quantum phase transition [30, 109, 110], a behaviour that has been observed in several experiments [111–116]. The complete region affected by the critical point is in general very hard to study, since one has to take into account the strongly-coupled dynamics of the quantum critical point as well as its excitations.

In this chapter we discuss the application of infinite dimensional tensor networks to quantum phase transitions in one dimension. We start by introducing the general formalism of tensor networks in section 2.1, with a focus on Matrix Product States (MPS). We review the properties of MPS and their applicability to critical systems. In section 2.2 we present the notion of infinite dimensional MPS, where the tensors are replaced by infinite dimensional field operators, through the example of a critical state arising from massless fields introduced in Ref. [38]. We propose to use this construction as a starting point for building models with a quantum critical region in their low temperature phase diagram. We explain how an additional parameter can drive the model away from criticality at zero temperature in section 2.3, thus constructing a toy model of a quantum phase transition. We provide a way to study the properties of the state at finite temperature in section 2.4, which is done by doubling the number of spins and considering the reduced state of half of the spins. Because the analytical expression of the wave functions is known, the properties of the states can be computed using Monte Carlo techniques for large system sizes, which allows us to investigate how the quantum phase transition influences the finite temperature phase diagram. Finally we show in section 2.5 that our model of

a mixed state ansatz corresponds to the thermal state of a local 2-body Hamiltonian. Sections 2.3, 2.4 and 2.5 of this chapter are based on and reuse parts of Ref. [1]:

- [1] : Ivan Glasser, J. Ignacio Cirac, Germán Sierra and Anne E. B. Nielsen, ‘Construction of spin models displaying quantum criticality from quantum field theory’, Nuclear Physics B **886**, 63-74 (2014), under CC BY 3.0 license, <http://dx.doi.org/10.1016/j.nuclphysb.2014.06.016>

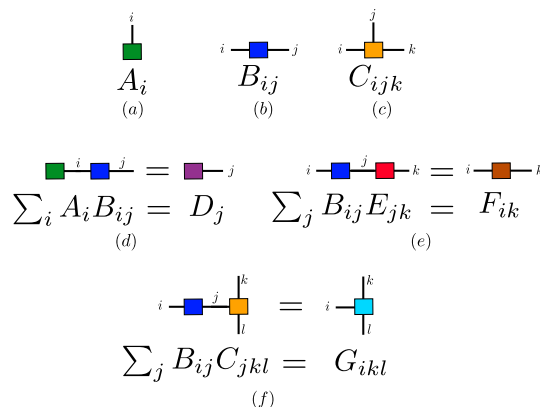
## 2.1 Introduction to tensor networks and Matrix Product States

### 2.1.1 Tensor networks

Consider a generic quantum state of  $N$  spins  $s_1, \dots, s_N$  on a lattice. The wave function can be written as

$$|\psi\rangle = \sum_{s_1, \dots, s_N=1}^d \psi^{s_1, \dots, s_N} |s_1\rangle \otimes |s_2\rangle \otimes \dots \otimes |s_N\rangle \quad (2.1)$$

where  $\psi^{s_1, \dots, s_N}$  is a tensor containing  $d^N$  complex coefficients and  $d$  is the dimension of the local Hilbert space. A tensor network state [15, 23, 24] is a state for which this tensor has been decomposed in a network of contracted tensors over a graph. An introduction to the graphical notation of tensor networks is presented in Fig. 2.1. A leg between two tensors corresponds to a summation of the corresponding indices of the tensors. Open legs index the resulting tensor and correspond to the physical indices. The dimension of the indices corresponding to open legs is therefore called



**Fig. 2.1.:** Graphical notation for tensor networks: (a) vector, (b) matrix, (c) order 3 tensor, (d) vector-matrix multiplication (e) matrix-matrix multiplication (f) matrix-tensor contraction. Figure reproduced from [6].

the physical dimension  $d$ , while the maximum dimension of the contracted indices is called the bond dimension and will be denoted  $D$ . Tensor networks provide both a representation of quantum states, as well as algorithms based on tensor network contractions to compute their properties. They can be defined in different geometries and in this chapter we will focus on the one-dimensional case.

## 2.1.2 Matrix Product States

In one dimension, a natural choice of network is a Matrix Product State (MPS) (Fig. 2.2), in which the wave function is decomposed as

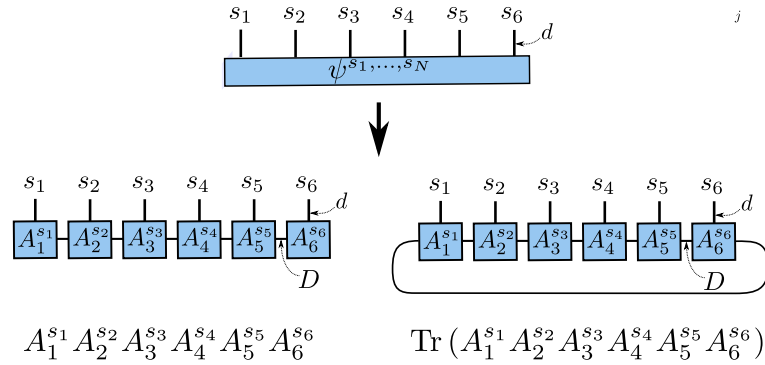
$$\psi^{s_1, \dots, s_N} = \sum_{\alpha, \beta, \dots, \omega=1}^D A_1^{s_1, \alpha} A_2^{s_2, \alpha, \beta} \cdots A_N^{s_N, \omega} \quad (2.2)$$

$$= A_1^{s_1} A_2^{s_2} \cdots A_N^{s_N}, \quad (2.3)$$

where  $A_1^{s_1}$  and  $A_N^{s_N}$  are  $D$ -dimensional vectors, and  $A_j^{s_j}$ ,  $j \in \{2, \dots, N-1\}$  are  $D \times D$  matrices. In periodic boundary conditions, all the  $A_j^{s_j}$  are matrices and the coefficients of the wave function are

$$\psi^{s_1, \dots, s_N} = \sum_{\alpha, \beta, \dots, \omega=1}^D A_{1, \alpha, \beta}^{s_1} A_{2, \beta, \gamma}^{s_2} \cdots A_{N, \omega, \alpha}^{s_N} \quad (2.4)$$

$$= \text{Tr} (A_1^{s_1} A_2^{s_2} \cdots A_N^{s_N}). \quad (2.5)$$



**Fig. 2.2.:** Decomposition of a tensor into a MPS in open boundary conditions (left) or periodic boundary conditions (right).

The number of coefficients in the MPS is  $\mathcal{O}(dD^2N)$ , which is linear in the system size, instead of the original exponential scaling  $d^N$  of the number of coefficients of  $\psi$ . By successive application of the singular value decomposition to the tensor  $\psi^{s_1, \dots, s_N}$ , one can show that any quantum state can be written as a MPS. In general

the corresponding bond dimension  $D$  will however scale exponentially with the system size. Nevertheless, already for small bond dimensions many interesting states can be represented by a MPS. Consider for example the AKLT model, with Hamiltonian for a one dimensional spin-1 system given by[117]

$$H = \sum_{i=1}^N \left[ \frac{1}{2} \mathbf{S}_i \cdot \mathbf{S}_{i+1} + \frac{1}{6} (\mathbf{S}_i \cdot \mathbf{S}_{i+1})^2 + \frac{1}{3} \right]. \quad (2.6)$$

This model is gapped and its ground state in periodic boundary conditions can be written as a MPS, with matrices of bond dimension 2 given by

$$A_j^{s_j=-1} = \sqrt{\frac{2}{3}} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, A_j^{s_j=0} = -\sqrt{\frac{1}{3}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, A_j^{s_j=1} = -\sqrt{\frac{2}{3}} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (2.7)$$

Given a Hamiltonian, one can use variational methods to find an approximation to its ground state in the form of a MPS. These algorithms rely on the minimization of the energy of the MPS. The variational principle ensures that the ground state is the only state minimizing the energy (up to potential degeneracies if the ground state manifold contains more than one state), so it can be approximated by minimizing the energy of a wave function:

$$\psi_{\text{GS}} = \min_{\psi} \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle}. \quad (2.8)$$

The Density Matrix Renormalization Group algorithm[16] performs this variational search on the class of MPS. It fixes the value of all tensors in the MPS except for one and optimizes this tensor to minimize the energy, which can be computed using tensor network contractions. By then sweeping along the MPS to update all tensors and to minimize the energy, an approximation to the ground state can be found. Another algorithm to optimize MPS based on Monte Carlo sampling will be presented later in chapter 4. Conversely, given a MPS one can always find a local Hamiltonian, denoted as parent Hamiltonian, for which this state is a ground state.

### 2.1.3 Properties of Matrix Product States

Let us now discuss an important property of MPS: they satisfy an area law. Consider a one dimensional spin chain with open boundary conditions and  $|\psi\rangle$  a quantum state that is a matrix product state. Let us denote  $A$  the region including the first  $L$  spins, and  $B$  the region containing the remaining  $N - L$  spins (Fig. 1.1), and write the Schmidt decomposition of  $|\psi\rangle$  as

$$|\psi\rangle = \sum_{\alpha=1}^{D_r} \lambda_{\alpha} |\psi_{A,\alpha}\rangle \otimes |\psi_{B,\alpha}\rangle, \quad (2.9)$$



where  $|\psi_{A,\alpha}\rangle$  (resp.  $|\psi_{B,\alpha}\rangle$ ) are orthonormal states living in region  $A$  (resp.  $B$ ) and  $\lambda_\alpha$  are positive numbers. For a MPS, we can compute explicitly the Schmidt values  $\lambda_\alpha$  by performing successive singular value decompositions which bring the MPS into a canonical form[118], we then have that the Schmidt rank  $D_r$  is bounded by the bond dimension  $D$  of the matrices. The reduced density matrix over system  $A$  can be written as

$$\rho_A = \text{Tr}_B |\psi\rangle\langle\psi| = \sum_{\alpha=1}^D \lambda_\alpha^2 |\psi_{A,\alpha}\rangle\langle\psi_{A,\alpha}|. \quad (2.10)$$

Let us introduce the entanglement entropy  $S_A$ , which is the Von Neumann entropy of  $\rho_A$ , and the  $\alpha$ -Renyi entropies  $S_A^\alpha$  ( $\alpha \geq 0$ ):

$$S_A = -\text{Tr} \rho_A \log \rho_A, \quad (2.11)$$

$$S_A^\alpha = \frac{1}{1-\alpha} \log \text{Tr} \rho_A^\alpha. \quad (2.12)$$

We have that  $S_A = -\sum_{\alpha=1}^D \lambda_\alpha^2 \log \lambda_\alpha^2$ , which is maximum when all the Schmidt values are equal to  $1/D$ . Finally we obtain that

$$S_A \leq \log D, \quad (2.13)$$

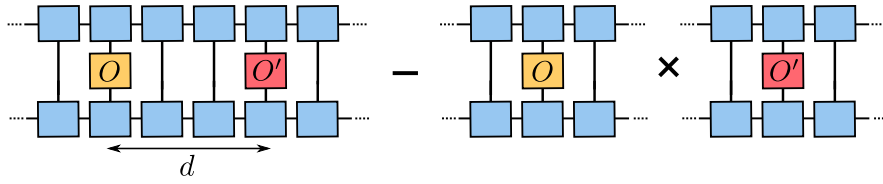
which is the area law for a MPS: for a bounded value of  $D$ , the entanglement entropy is bounded by a constant, which is proportional to the size of the boundary between regions  $A$  and  $B$ . Notice that if region  $A$  consists of a region in the middle of the chain, then two bounds are cut at the boundary between  $A$  and  $B$ , and we have that  $S_A \leq 2 \log D$ .

Moreover, any quantum state satisfying an area law in one dimension for  $\alpha$ -Renyi entropies,  $\alpha < 1$ , can be efficiently approximated by a MPS[26, 27]. This means that the number of parameters in the MPS scales polynomially with the system size for a fixed accuracy. Moreover it has been shown that the area law holds for one-dimensional Hamiltonians with local interactions and a finite non-zero energy gap between the ground state and the first excited state[17]. This means that MPS can approximate ground states of one-dimensional gapped Hamiltonians efficiently.

Let us now look at the computation of expectation values of observables for a MPS. Consider an infinite translational invariant MPS and two local observables  $O_r$  and  $O'_{r+d}$  acting at two sites separated by a distance  $d$ . The correlation function is defined as

$$C(d) = \langle\psi|O_r O'_{r+d}|\psi\rangle - \langle\psi|O_r|\psi\rangle\langle\psi|O'_{r+d}|\psi\rangle, \quad (2.14)$$

which in tensor diagrams notation is represented in Fig. 2.3 (we assume here that the MPS is normalized). The correlation function can be computed by contracting



**Fig. 2.3.:** Correlation function of two observables for a MPS.

the corresponding tensor network. For an infinite translational invariant MPS, it can be shown that at long distances

$$|C(d)| \sim C e^{-d/\xi}, \quad (2.15)$$

where  $C$  is a constant and  $\xi$  is the correlation length. Correlations in a MPS therefore decay exponentially with the distance.

## 2.2 Infinite dimensional Matrix Product States for critical systems

### 2.2.1 Critical states and Matrix Product States

Consider now a system at a continuous quantum phase transition. At the critical point, the gap closes and the correlation length diverges. Scale invariance emerges because there is no finite characteristic length in the system, so the correlations cannot decay exponentially with the distance. Instead, the correlation function takes the form of a polynomial  $C(d) = 1/d^\nu$ , where  $\nu$  is a critical exponent. Because of scale invariance, the critical point can be described by a scale invariant massless theory. This theory is very often a Conformal Field Theory (CFT). Tools from CFT allow in addition to compute the entanglement and Renyi entropies of an interval of length  $L$ , obtaining the universal formulas[33–35]:

$$S_L \sim \frac{c}{3} \log \frac{L}{a}, \quad (2.16)$$

$$S_L^\alpha \sim \frac{c}{6} \left(1 - \frac{1}{\alpha}\right) \log \frac{L}{a}, \quad (2.17)$$

where  $c$  is the central charge, a number characterizing the CFT. These formulas show that critical systems violate the area law, although only logarithmically.

If we compare these properties with the properties of MPS of fixed bond dimension, we see that they are incompatible: in a MPS with fixed bond dimension, the correlations will decay exponentially, and not polynomially at large distances, and the area law will be satisfied. Nevertheless it was shown that MPS can approximate critical systems efficiently even if the entropy scales logarithmically with the size of the subsystem[26, 27]. In this case the bond dimension  $D$  must increase polynomially with the system size, as can be seen from (2.13). One may ask if there is another way of defining MPS such that they would represent critical systems in the thermodynamic limit, which requires having an infinite bond dimension. In general this would require an infinite number of parameters and the wave function could not be computed. Nevertheless, in some cases it is possible to deal with infinite operators which have few parameters and compute an MPS wave function. The resulting states are called infinite dimensional MPS and were introduced in [38].

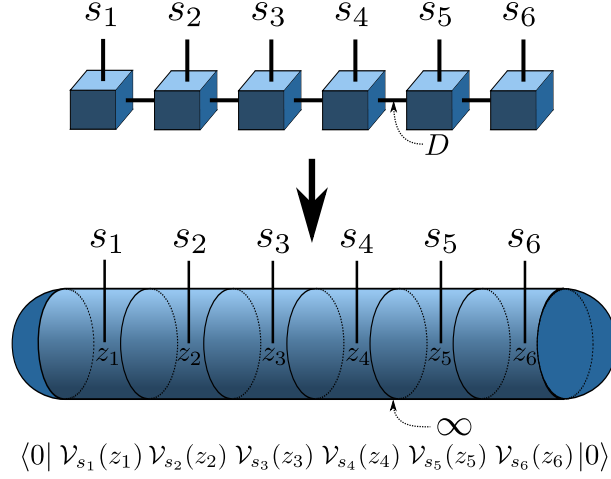
## 2.2.2 Construction of an infinite dimensional Matrix Product States from a free massless boson

An infinite dimensional MPS is a state for which the coefficients of the wave function have the form of a vacuum expectation value of a product of field operators:

$$\psi(s_1, \dots, s_N) \propto \langle 0 | \mathcal{V}_{s_1}(z_1) \dots \mathcal{V}_{s_N}(z_N) | 0 \rangle. \quad (2.18)$$

Here we consider a field theory with vacuum  $|0\rangle$  as well as field operators  $\mathcal{V}_{s_j}$ , which depend on the value of the spin variable and act at positions  $z_i$  which will be specified later. In general such a correlator cannot be evaluated, but as we will see an appropriate choice of the field operators leads to analytical expressions for the wave function. This expression is reminiscent from the definition of a MPS in (2.3), but the matrices have been replaced with field operators, and the boundary vectors by a vacuum expectation value (Fig. 2.4). For a bosonic field for example the field operators can be expressed in terms of bosonic creation and annihilation operators, acting on an infinite dimensional space which is the bosonic Fock space. The virtual degrees of freedom in a MPS now correspond to bosonic modes in this Fock space.

In the following we will review the construction for the first infinite dimensional MPS introduced in [38]. Consider a field theory in 1+1 space-time dimensions, with coordinates  $x$  and  $t$ . Since we are interested in describing critical systems, it is natural to choose this theory to be conformal invariant, and the operators in (2.18) to be scale invariant field operators of this theory. Because we later want to extend our construction to other field theories which may not be conformal field theories, we will present the construction independently of the CFT structure. This can be



**Fig. 2.4.:** MPS (top) and infinite dimensional MPS (bottom) in which the tensors have been replaced by field operators depending on the spin  $s_j$  and acting at positions  $z_j$ .

done as follows. We choose the free massless boson field theory, which is a CFT[32], with field  $\phi(x, t)$  and Lagrangian density

$$\mathcal{L} = \frac{1}{8\pi} \left( (\partial_t \phi(x, t))^2 - (\partial_x \phi(x, t))^2 \right). \quad (2.19)$$

One can already see that this theory will be invariant under scaling transformations of the form  $(x, t) \rightarrow (\lambda x, \lambda t)$ . We consider a cylinder geometry with circumference  $R$  such that  $\phi(x + R, t) = \phi(x, t)$  and work in the transformed coordinates  $z = e^{-2\pi i(x-it)/R}$  and  $\bar{z} = e^{2\pi i(x+it)/R}$ . The theory can be quantized and the mode expansion for the field can be expressed as

$$\phi(z, \bar{z}) = \varphi_0 - i\pi_0 \ln(z\bar{z}) + \sum_{n=1}^{\infty} \frac{i}{\sqrt{n}} \left( a_n z_j^{-n} - a_n^\dagger z_j^n \right) + \sum_{n=1}^{\infty} \frac{i}{\sqrt{n}} \left( \bar{a}_n \bar{z}_j^{-n} - \bar{a}_n^\dagger \bar{z}_j^n \right), \quad (2.20)$$

where the non-zero commutators are  $[a_n, a_m^\dagger] = [\bar{a}_n, \bar{a}_m^\dagger] = \delta_{nm}$ ,  $[\varphi_0, \pi_0] = i$ . In the massless case, this field can be divided into chiral fields  $\varphi_L(\bar{z})$  and  $\varphi_R(z)$  which depend only on  $z$  and  $\bar{z}$ :  $\phi(z) = \varphi_L(\bar{z}) + \varphi_R(z)$ , with

$$\varphi_R(z) = Q - iP \ln(z) + \sum_{n=1}^{\infty} \frac{i}{\sqrt{n}} \left( a_n z_j^{-n} - a_n^\dagger z_j^n \right), \quad (2.21)$$

$$\varphi_L(\bar{z}) = \bar{Q} - i\bar{P} \ln(\bar{z}) + \sum_{n=1}^{\infty} \frac{i}{\sqrt{n}} \left( \bar{a}_n \bar{z}_j^{-n} - \bar{a}_n^\dagger \bar{z}_j^n \right), \quad (2.22)$$

where  $[Q, P] = [\bar{Q}, \bar{P}] = i$  and  $[Q, \bar{Q}] = [Q, \bar{P}] = [P, \bar{Q}] = [P, \bar{P}] = 0$ .

Now that we know the expression of the field in terms of bosonic creation and annihilation operators, we have to make a choice for the operators  $\mathcal{V}_{s_j}$  appearing in the wave function. To obtain a scale-invariant wave function a natural choice is to take operators which transform appropriately under such transformations. More precisely, since we are working with a conformal theory, consider conformal transformations. These are transformations which preserve angles, locally they are equivalent to rotations and dilations. On the complex plane global conformal transformations are defined as

$$z \rightarrow \frac{az + b}{cz + d}, \quad ad - bc = 1. \quad (2.23)$$

Any mapping  $z \rightarrow w = f(z)$ ,  $\bar{z} \rightarrow \bar{w} = \bar{f}(\bar{z})$ , where  $f$  is analytic and  $\bar{f}$  is antianalytic, is locally conformal. A primary field is a field which transforms under such a mapping as

$$\Phi(z, \bar{z}) \rightarrow \Phi'(w, \bar{w}) = \left( \frac{\partial w}{\partial z} \right)^{-h} \left( \frac{\partial \bar{w}}{\partial \bar{z}} \right)^{-\bar{h}} \Phi(z, \bar{z}), \quad (2.24)$$

where  $h$  and  $\bar{h}$  are called the conformal dimensions of the field  $\Phi$ . Fields transforming accordingly under global conformal transformations are called quasi-primary field. Invariance under rotations, translations and scale invariance completely fixes the correlations of such fields: two-point correlations take the form

$$\langle \Phi_1(z_1, \bar{z}_1) \Phi_2(z_2, \bar{z}_2) \rangle \propto \frac{1}{(z_1 - z_2)^{h_1+h_2} (\bar{z}_1 - \bar{z}_2)^{\bar{h}_1+\bar{h}_2}}, \quad (2.25)$$

which is reminiscent of the similar form for scale-invariant decay of correlations in a critical system.

The boson field  $\phi$  is not a primary field, and so not a suitable candidate to construct a scale invariant infinite dimensional MPS. However it can be used to construct primary fields in the free massless boson theory which are vertex operators, normal ordered exponentials of the field defined as  $V_\beta(z, \bar{z}) =: e^{i\beta\phi(z, \bar{z})} :.$  The normal ordering denoted as  $: \cdot :$  reorders operators such that creation operators are on the left of annihilation operators. As we have seen we can decompose the field  $\phi$  into chiral fields depending only on  $z$  or  $\bar{z}$ , which allows us to decouple holomorphic and non-holomorphic operators, so here we can also use only the chiral part of the field in the expression of the vertex operator:

$$V_{R,\beta}(z) =: e^{i\beta\varphi_R(z)} :. \quad (2.26)$$

These operators have for correlations

$$\langle V_{R,\beta(z_1)} V_{R,\beta'(z_2)} \rangle = \delta_{\beta+\beta'=0} (z_1 - z_2)^{\beta\beta'}. \quad (2.27)$$

Ref. [38] constructed an infinite dimensional MPS using these operators by making the choice

$$\mathcal{V}_{s_j}(z_j) = \chi_{s_j} V_{R, \sqrt{\alpha} s_j}(z_j), \quad (2.28)$$

$$= \chi_{s_j} : e^{i\sqrt{\alpha} s_j \varphi_R(z_j)} :, \quad (2.29)$$

where  $\alpha$  is a positive real number that is a free parameter,  $z_j$  is a position on the complex plane and  $\chi_{s_j}$  is a phase that is chosen to be

$$\chi_j = (-1)^{(j-1)(s_j+1)/2}. \quad (2.30)$$

It remains to evaluate the correlator of the field operators in (2.18), leading to a wave function, that we denote  $\psi_0$ , equal to

$$\psi_0(s_1, \dots, s_N) \propto \langle \chi_1 : e^{i\sqrt{\alpha} s_1 \varphi_R(z_1)} : \dots \chi_N : e^{i\sqrt{\alpha} s_N \varphi_R(z_N)} : \rangle, \quad (2.31)$$

$$\propto \prod_{l=1}^N \chi_l \delta_s \prod_{j < k} (z_j - z_k)^{\alpha s_j s_k}, \quad (2.32)$$

where  $\delta_s = 1$  if  $\sum_{j=1}^N s_j = 0$  and zero otherwise. The  $z_j$  are free parameters that can (but need not) be taken as position of the spins on the complex plane. For example a translational invariant chain leads to positions  $z_j = e^{2\pi i j/N}$  on the unit circle.

## 2.2.3 Computing properties of the state with Monte Carlo techniques

Now that we have defined a quantum state using an infinite dimensional MPS, one may ask what are its properties. Correlations can no longer be computed using tensor-network contractions, but one has access to the analytical expression of the wave function  $\psi_0$  (2.32). This allows the use of powerful Monte Carlo techniques to compute expectation values of observables. Indeed consider for example the expectation value of  $\sigma_j^z \sigma_k^z$ , which can be written as

$$\langle \sigma_j^z \sigma_k^z \rangle = \frac{\langle \psi_0 | \sigma_j^z \sigma_k^z | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle}, \quad (2.33)$$

$$= \sum_{s_1, \dots, s_N=1}^d \frac{|\psi_0(s_1, \dots, s_N)|^2}{\langle \psi_0 | \psi_0 \rangle} s_j s_k. \quad (2.34)$$

Now observe that  $p(s_1, \dots, s_N) = \frac{|\psi_0(s_1, \dots, s_N)|^2}{\langle \psi_0 | \psi_0 \rangle}$  is a probability distribution: it is positive and normalized. The expectation value  $\langle \sigma_j^z \sigma_k^z \rangle$  can be computed as an average value of  $s_j s_k$  over samples from probability distribution  $p$ . Using the Metropolis-Hastings algorithm[119, 120], it is possible to generate samples from a distribution as long as one can compute a function  $f$  proportional to the probability

distribution  $p$ . In our case we know the analytical expression for the wave function up to a normalization factor and can use this expression as function  $f$ . The procedure constructs a Markov chain which asymptotically reaches a stationary distribution equal to  $p$ , using the following algorithm:

- Initialization:
  - Set the time  $t = 0$ .
  - Choose a random configuration  $\mathbf{s}^0 = s_1, \dots, s_N$ .
  - Choose a proposal distribution  $g(\mathbf{s}'|\mathbf{s})$  which generates a new configuration from a previous configuration, for example by flipping a spin.
- Iterate at time  $t$ :
  - Generate a new candidate configuration  $\mathbf{s}'$  randomly according to  $g(\mathbf{s}'|\mathbf{s}^t)$ .
  - Compute an acceptance probability  $A = \min\left(1, \frac{f(\mathbf{s}')}{f(\mathbf{s}^t)} \frac{g(\mathbf{s}^t|\mathbf{s}')}{g(\mathbf{s}'|\mathbf{s}^t)}\right)$
  - Generate a uniform random number  $r \in [0, 1]$ .
    - \* If  $r \leq A$  (which happens with probability  $A$ ), accept the new configuration  $\mathbf{s}^{t+1} = \mathbf{s}'$
    - \* If  $r > A$ , reject the new configuration and set  $\mathbf{s}^{t+1} = \mathbf{s}^t$
    - \* Increment the time  $t$  by 1.

By repeating this procedure one obtains (correlated) samples  $\mathbf{s}^0, \mathbf{s}^1, \dots, \mathbf{s}^t$  from  $p$ . It is straightforward to parallelize this algorithm by running multiple copies of the program starting with different initial configurations. The mean of  $s_j s_k$  over these samples converges to the desired expectation value and can be computed with hundred of spins, which is a much larger system than can be studied with exact diagonalization techniques.

Another quantity that can be computed efficiently with Monte Carlo techniques is the second Renyi entropy  $S_A^{(2)} = -\log \text{Tr} \rho_A^2$ , where  $\rho_A$  is the density matrix of a subsystem  $A$ . This can be done by rewriting [38, 121–123]

$$e^{-S_A^{(2)}} = \sum_{n,n',m,m'} |\langle n, m | \psi_0 \rangle|^2 |\langle n', m' | \psi_0 \rangle|^2 \frac{\langle \psi_0 | n', m \rangle \langle \psi_0 | n, m' \rangle}{\langle \psi_0 | n, m \rangle \langle \psi_0 | n', m' \rangle} \left[ \sum_{n,m} |\langle n, m | \psi_0 \rangle|^2 \right]^{-2}, \quad (2.35)$$

where  $|n\rangle$  (and  $|n'\rangle$ ) is an orthonormal basis in the space of the spins in region  $A$  and  $|m\rangle$  (and  $|m'\rangle$ ) another basis corresponding to the rest of the spins. The right-hand side of this expression is an expectation value so that the sum can be performed again by using the previous algorithm, this time with two independent spin chains.

Using these techniques, it was shown in [38] that the two-point correlations of the state  $\psi_0$  decay polynomially with the distance and that the Renyi entropy of a reduced part of the system of size  $L$  scales logarithmically with  $L$ , for all values of  $0 < \alpha \leq 1/2$ . These results will be reproduced and extended on in the next

section. These properties confirm that  $\psi_0$  is a critical state. Other critical models have also been constructed using a different CFT as the starting point[124, 125]. It has moreover been demonstrated that a related infinite dimensional MPS including fewer bosonic modes can display even power law violations of the area law[126].

So far we have considered a wave function but not a physical system. One may ask whether such a state is the ground state of a physical Hamiltonian. It was found that  $\psi_0$  has a high overlap with the ground state of the critical XXZ chain with periodic boundary conditions for a suitable choice of the anisotropic coupling. Moreover at  $\alpha = 1/2$ ,  $\psi_0$  is exactly equal to the ground state of the Haldane-Shastry model[127, 128]. This confirms that the state  $\psi_0$  is critical. An exact parent Hamiltonian for which  $\psi_0$  is the ground state at  $\alpha = 1/2$  for any choice of positions  $z_j$  has also been derived[70].

## 2.3 A quantum phase transition with infinite dimensional Matrix Product States

### 2.3.1 Construction of an infinite dimensional Matrix Product State from a massive field

We now would like to introduce a parameter that takes the state away from criticality, to create a model of a phase transition. This can be done naturally by generalizing the previous construction arising from a free massless boson to the case of a free boson with mass  $m$ , breaking scale invariance.

The mode expansion of a free boson  $\phi_m$  with mass  $m$  is

$$\begin{aligned} \phi_m(z_j, \bar{z}_j) = & i\sqrt{\frac{2\pi}{Rm}} \left( a_0 z_j^{-\frac{mR}{4\pi}} \bar{z}_j^{-\frac{mR}{4\pi}} - a_0^\dagger z_j^{\frac{mR}{4\pi}} \bar{z}_j^{\frac{mR}{4\pi}} \right) \\ & + i \sum_{n=1}^{\infty} \sqrt{\frac{2\pi}{Rw_n}} \left( a_n z_j^{-\frac{n}{2} - \frac{w_n R}{4\pi}} \bar{z}_j^{\frac{n}{2} - \frac{w_n R}{4\pi}} - a_n^\dagger z_j^{\frac{n}{2} + \frac{w_n R}{4\pi}} \bar{z}_j^{-\frac{n}{2} + \frac{w_n R}{4\pi}} \right) \\ & + i \sum_{n=1}^{\infty} \sqrt{\frac{2\pi}{Rw_n}} \left( \bar{a}_n z_j^{+\frac{n}{2} - \frac{w_n R}{4\pi}} \bar{z}_j^{-\frac{n}{2} - \frac{w_n R}{4\pi}} - \bar{a}_n^\dagger z_j^{-\frac{n}{2} + \frac{w_n R}{4\pi}} \bar{z}_j^{\frac{n}{2} + \frac{w_n R}{4\pi}} \right), \quad (2.36) \end{aligned}$$

where the frequency is  $w_n = \sqrt{m^2 + \left(\frac{2\pi n}{R}\right)^2}$ . Unlike in the massless case the zero mode cannot be separated into two chiral parts, so instead of working with the chiral field  $\varphi_R(\bar{z})$  we will work with the total field  $\phi_m(z, \bar{z})$ . We can observe that already



in the massless case the diagonal elements of the density operator only depend on the full correlator:

$$|\psi_0(s_1, \dots, s_N)|^2 \propto \langle : e^{i\sqrt{\alpha}s_1\phi(z_1, \bar{z}_1)} : \dots : e^{i\sqrt{\alpha}s_N\phi(z_N, \bar{z}_N)} : \rangle. \quad (2.37)$$

This motivates us to choose a wave function which satisfies the same property with a field of mass  $m$ . In this case, the full correlator takes the form

$$\begin{aligned} & \langle : e^{i\sqrt{\alpha}s_1\phi_m(z_1, \bar{z}_1)} : \dots : e^{i\sqrt{\alpha}s_N\phi_m(z_N, \bar{z}_N)} : \rangle = \\ & \exp\left(-\alpha \sum_{j<k} s_j s_k \frac{2\pi}{Rm} \left(\frac{|z_k|}{|z_j|}\right)^{\frac{Rm}{2\pi}}\right) \exp\left(-\alpha \sum_{j<k} s_j s_k \sum_{n=1}^{\infty} \frac{2\pi}{Rw_n} \left(\frac{z_k}{z_j}\right)^{\frac{n}{2} + \frac{w_n R}{4\pi}} \left(\frac{\bar{z}_j}{\bar{z}_k}\right)^{\frac{n}{2} - \frac{w_n R}{4\pi}}\right) \\ & \times \exp\left(-\alpha \sum_{j<k} s_j s_k \sum_{n=1}^{\infty} \frac{2\pi}{Rw_n} \left(\frac{\bar{z}_k}{\bar{z}_j}\right)^{\frac{n}{2} + \frac{w_n R}{4\pi}} \left(\frac{z_j}{z_k}\right)^{\frac{n}{2} - \frac{w_n R}{4\pi}}\right). \end{aligned} \quad (2.38)$$

To obtain a wave function satisfying (2.37), one has to divide this correlator. This can be done by choosing a wave function proportional to

$$\begin{aligned} & \exp\left(-\alpha \sum_{j<k} s_j s_k \frac{\pi}{Rm} \left(\frac{|z_k|}{|z_j|}\right)^{\frac{Rm}{2\pi}}\right) \\ & \times \exp\left(-\alpha \sum_{j<k} s_j s_k \sum_{n=1}^{\infty} \frac{2\pi}{Rw_n} \left(\frac{z_k}{z_j}\right)^{\frac{n}{2} + \frac{w_n R}{4\pi}} \left(\frac{\bar{z}_j}{\bar{z}_k}\right)^{\frac{n}{2} - \frac{w_n R}{4\pi}}\right). \end{aligned} \quad (2.39)$$

This expression is not translational invariant when the spins are placed on a circle at positions  $z_j = e^{2\pi i j/N}$ , as was the case in the massless case. We therefore take the limit  $R \rightarrow \infty$  in the previous expression, which maps the cylinder to a plane and a periodic spin chain to an open spin chain. The resulting wave function is

$$\psi_m(s_1, \dots, s_N) \propto \prod_{l=1}^N \chi_l \exp\left(-\alpha \sum_{j<k} s_j s_k \int_0^{\infty} \frac{e^{-ip(x_k - x_j)}}{\sqrt{m^2 + p^2}} e^{-\sqrt{m^2 + p^2}(t_k - t_j)} dp\right). \quad (2.40)$$

The diagonal elements of the density operator are now equal to the full correlator for the field with mass (2.38). Moreover, in the limit  $m \rightarrow 0$ , the wave function reduces to

$$\psi_0(s_1, \dots, s_N) \propto \delta_s \prod_{j=1}^N \chi_j \prod_{j<k} (v_j - v_k)^{\alpha s_j s_k}, \quad (2.41)$$

where  $v_j = x_j - it_j$ . Thus we recover the massless wave function (2.32) introduced in the previous section, but in the planar geometry. For this reason we choose the same particular choice of phases (2.30), which will be assumed in the rest of this work.

### 2.3.2 Properties of the phase transition at zero temperature

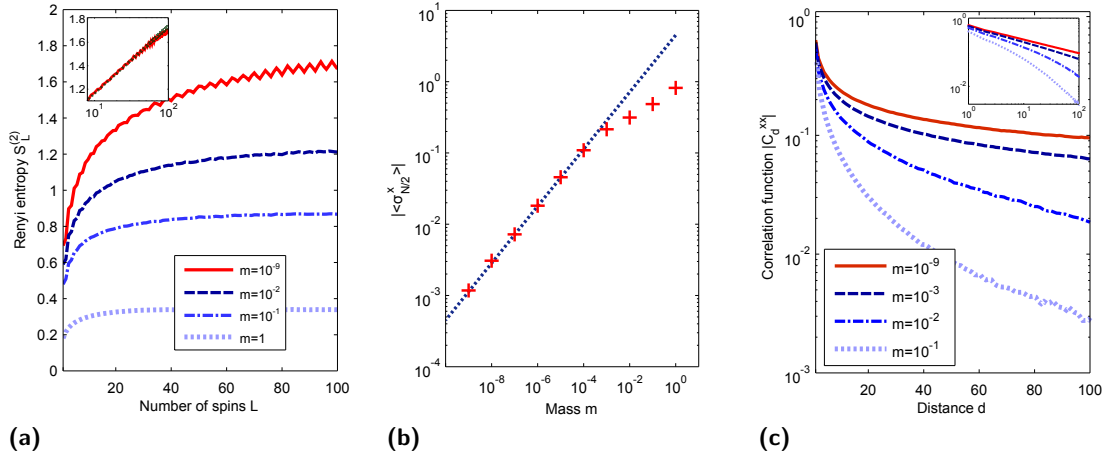
Let us now consider a one-dimensional chain of  $N$  spins at positions  $x_j = j$ ,  $t_j = 0$  with open boundary conditions and suppose that the wave function is given by (2.40). The integral in the wave function can be expressed as:

$$\int_0^\infty \frac{e^{-ip(x_k - x_j)}}{\sqrt{m^2 + p^2}} dp = K_0(m|x_k - x_j|) - \frac{i\pi}{2} \frac{k - j}{|k - j|} (I_0(m|x_k - x_j|) - L_0(m|x_k - x_j|)), \quad (2.42)$$

where  $L_0$  is a modified Struve function and  $K_0$  and  $I_0$  are modified Bessel functions. This wave function depends only on the distance between two spins. In the thermodynamic limit, it is invariant under a translation  $x_k \rightarrow x_k + 2$ . We shall only consider a value of  $\alpha$  in the range  $(0, 1/2]$ . In the limit of zero mass and when the spins are on a chain with periodic boundary conditions, the wave function corresponds to the massless wave function studied in [38]: the state is critical and close to the ground state of a critical XXZ chain. Here we study a chain with open boundary conditions but we expect the state to have similar properties in the massless limit.

Let us first compute the second Renyi entropy (2.12) of a reduced part of the system of size  $L$  in the middle of the chain. This can be done using the Monte Carlo algorithm introduced in 2.2.3. The results for different values of the mass are shown in Fig. 2.5a. When the mass is close to zero, the second Renyi entropy scales as  $\frac{1}{4} \log(L)$ , which is the expected result for an infinite critical chain with central charge  $c = 1$  [33–35], such as the massless free boson theory. For higher values of the mass, the entropy saturates to a constant that is independent of  $L$  and the state is no longer critical.

We can also compute the expectation value of single spin operators  $\langle \sigma_n^a \rangle$  ( $a = x, y, z$ ) and two-point correlation functions  $C_d^{aa} = \langle \sigma_n^a \sigma_{n+d}^a \rangle - \langle \sigma_n^a \rangle \langle \sigma_{n+d}^a \rangle$ . Since we are interested in the thermodynamic limit, we compute the correlators between spins at positions  $(N - d)/2$  and  $(N + d)/2$  that sit in the middle of the chain. We check that these quantities do not depend on the total number of spins  $N$  as long as  $d < N/2$ , so that the behaviour in the thermodynamic limit can be extracted from these measurements. The results in Fig. 2.5b and 2.5c show that in the massless case  $C_d^{xx}$  decays polynomially and the expectation value  $\langle \sigma_n^x \rangle$  is zero. In the massive case however this expectation value is no longer zero, but shows long-range anti-ferromagnetic order in the  $x$  direction, so the expectation value  $\langle \sigma_n^x \rangle$  can serve as order parameter. The correlation function  $C_d^{xx}$  decays exponentially at large distances when there is a mass: this defines a finite correlation length  $\lambda$  such that  $C_d^{xx} \propto e^{-d/\lambda}$ . This length diverges when the mass goes to zero, while in the limit  $m \rightarrow \infty$ , the state is a Néel state in the  $x$  direction, which is invariant under



**Fig. 2.5.:** All quantities are evaluated from Monte Carlo simulations for  $\alpha = 0.2$ .

(a) Scaling of the Renyi entropy  $S_L^{(2)}$  for different values of the mass,  $N = 200$  spins. The inset shows with a logarithmic scale for the abscissa that the scaling is logarithmic in  $L$  when the mass is close to zero. The dotted line in the inset is a fit of the form  $\frac{1}{4} \log(L) + \text{constant}$ , confirming that the central charge of the critical point is 1. For higher masses, the entropy saturates to a constant. The error bars, estimated from the standard error of a mean of Monte Carlo trajectories with different initial conditions, are of the order  $10^{-3}$  for all points.

(b) Absolute value of the magnetization in the  $x$  direction for one site  $\langle \sigma_{N/2}^x \rangle$  as a function of the mass for a chain with  $N = 600$  spins. The error bars are smaller than 5% of the values for all points. The dotted line is a fit of the first 6 points of the form  $\gamma m^\tau$ , from which the critical exponent can be extracted. Here  $\tau = 0.4$ .

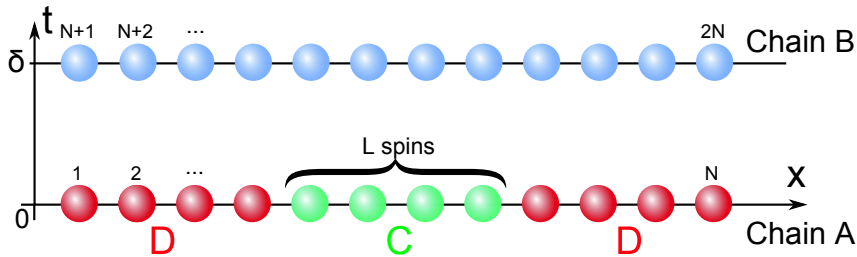
(c) Absolute value of the connected correlation function  $C_d^{xx}$ ,  $N = 200$  spins. When the mass is zero this quantity decays polynomially (red solid line), while in the massive case the decay for large  $d$  is exponential. The inset shows the same quantity in a log-log scale. The error bars are smaller than  $2 \times 10^{-4}$  for all points. Figure reproduced from [1].

translations  $x_k \rightarrow x_k + 2$ . We can conclude that the mass introduces a length scale in the system and breaks the criticality of the state.

## 2.4 Infinite dimensional Matrix Product States at finite temperature

So far we have used a pure state description at zero temperature. Experiments are however at small but non-zero temperature and we may ask what are the properties above the quantum phase transition at finite temperature. For this we need to introduce a mixed state ansatz to describe a spin chain at finite temperature. Consider two chains  $A$  and  $B$  of  $N$  spins each, with coordinates  $(x_j = j, t_j = 0)$  and  $(x_j = j, t_j = \delta)$  respectively (Fig. 2.6). Let us describe the state of the complete

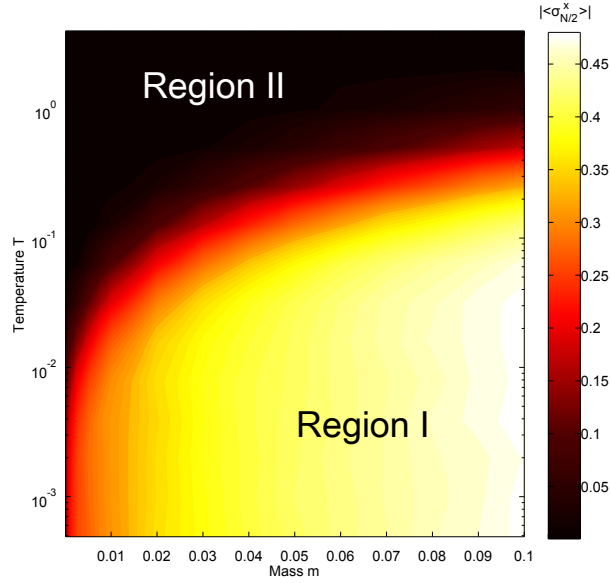
system  $\psi_{AB}$  by the previously introduced wave function (2.40). The state of the first spin chain can be computed by tracing out the degrees of freedom of the second spin chain, resulting in the reduced density matrix  $\rho_A = \text{Tr}_B |\psi_{AB}\rangle\langle\psi_{AB}|$ . In the limit where  $\delta \rightarrow \infty$ , the two chains decouple. System A is then in a pure state at zero temperature and we recover two copies of the state described in the previous section. In the limit  $\delta \rightarrow 0$ , each spin from the first chain is very close to a spin from the second chain and the two spins form a singlet, so that the effective temperature for each chain is infinite. For a finite  $\delta$ , this construction therefore introduces an effective temperature for the chain A. We define  $T \equiv 1/\delta$  as a representation of the temperature of the system A. Note that the effective temperature may depend differently on  $\delta$ .



**Fig. 2.6.:** The two spin chains A and B are separated by a distance  $\delta$ . The one dimensional state that is studied is the reduced state of the chain A. To compute the mutual information we separate the chain into the systems C of  $L$  spins in the middle of the chain and the complementary system D of  $N - L$  spins. Figure reproduced from [1].

It is not possible to compute the complete wave function and take the partial trace for a large system, but it is not necessary in this case: Renyi entropies and spin-spin correlators between spins on the first chain can be computed using the wave function of the two chains in the same way as in the zero temperature case, so that the trace never needs to be performed in actual computations.

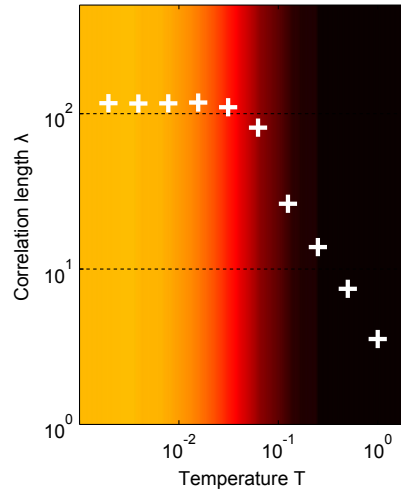
At zero temperature and in the thermodynamic limit, the expectation value  $\langle\sigma_i^x\rangle$  is zero at the critical point but does not vanish when there is a non-zero mass. The absolute value of this quantity, computed on a site at position  $i = \frac{N}{2}$  in the middle of a chain of 600 spins, is used to draw the phase diagram at finite temperature (Fig. 2.7). We observe that two distinct regions appear in the phase diagram. For non-zero mass and small temperature,  $|\langle\sigma_{N/2}^x\rangle|$  is non-zero and independent of the temperature. There is still long-range order in the  $x$  direction (region I). If we increase the temperature, the magnetization starts to decrease rapidly with the temperature, before reaching a very small value, which may disappear in the thermodynamic limit (region II). These two distinct regions also appear when looking at other observables. We show in Fig. 2.8 the behavior of the correlation length  $\lambda$  computed from the  $C_d^{xx}$  correlation function. For small temperatures, the correlation length is independent of the temperature, whereas in the second region the correlation length decreases



**Fig. 2.7.:** Absolute value of the magnetization in the  $x$  direction for one site  $\langle \sigma_{N/2}^x \rangle$  as a function of mass and temperature from Monte Carlo simulations with  $N = 600$ ,  $\alpha = 0.2$ . Figure reproduced from [1].

with the temperature. We can compare this behaviour with the phase diagram of an Ising model with a transverse magnetic field, which is the prototype model of quantum criticality [30, 129, 130]: this model has a quantum paramagnetic phase at low temperatures in which the correlation length is independent of the temperature, while at higher temperatures it reaches a region of quantum criticality in which the correlation length decays as  $1/T$ . Our mixed state ansatz displays qualitatively similar properties.

Another quantity that can be used to probe a state at finite temperature is the mutual information  $I = S_C + S_D - S_{C \cup D}$ , where  $\{C, D\}$  is a partition of the spin chain and  $S_C$  is the von Neumann entropy of subsystem  $C$ . The mutual information is a measure of both classical and quantum correlations between two parts of the system [131]. It fulfils an area law when the system has a finite correlation length [132]. Here we consider instead the Renyi mutual information, obtained by replacing von Neumann entropies by second Renyi entropies in the definition of the mutual information:  $I_L^{(2)} = S_C^{(2)} + S_D^{(2)} - S_A^{(2)}$ . We take  $C$  to be the system of  $L$  spins in the middle of the chain and  $D$  to be the system of the  $N - L$  complementary spins in the chain  $A$ . In this case  $A = C \cup D$  (Fig. 2.6). The Renyi mutual information has been measured for spin systems at finite temperature [133–135] and is expected to have a similar behaviour as the mutual information. At zero temperature, the chain is in a pure state and the entropy of the complete chain is zero. The Renyi mutual information therefore reduces to  $2S_L^{(2)}$ , which scales logarithmically with  $L$  at the critical point. We observe that the Renyi mutual information saturates to a constant

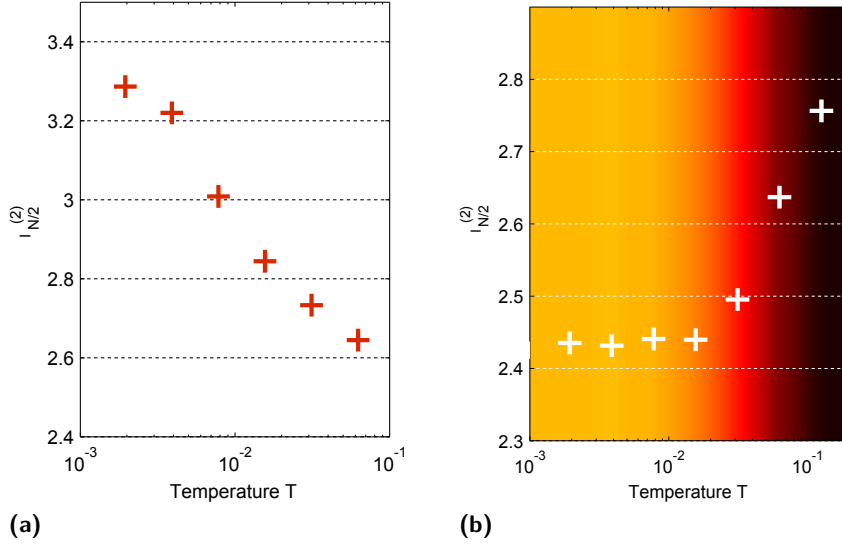


**Fig. 2.8.:** Correlation length  $\lambda$  as a function of the temperature for a spin chain with  $N = 200$  and  $\alpha = 0.2$ , at  $m = 0.01$ . The background colours represent the value of  $|\langle\sigma_{N/2}^x\rangle|$  taken from the phase diagram (Fig. 2.7) for the same value of the mass. Figure reproduced from [1].

at large  $L$  except when we are close to the critical point, which is compatible with the observation that the correlation length is finite everywhere except at the critical point. In region I this constant is independent of the temperature. In region II however the Renyi mutual information converges to a value that has some non trivial dependence on the temperature. Above the critical point we observe that the Renyi mutual information between two halves of the system decays with the temperature (Fig. 2.9a). A similar behavior is expected for an XXZ chain at finite temperature [136], since the mutual information diverges at zero temperature. In the massive case this quantity increases when the region of quantum criticality is reached (Fig. 2.9b), while in the limit of infinite temperatures there would be no correlations and the mutual information would vanish again. A similar behavior of the mutual information has been observed near regions of quantum criticality at finite temperature in different models [133, 135, 137]. These results confirm the presence of a region of quantum criticality above the critical point in the phase diagram, in which strong correlations are present.

## 2.5 From a model wave function to a model Hamiltonian

So far we have considered only wave functions, but not physical systems. One may ask whether these wave functions correspond to the ground state or thermal state of some physical Hamiltonian, which might realize a system with the properties we



**Fig. 2.9.:** (a) Mutual information between two halves of the chain  $I_{N/2}^{(2)}$  for  $N = 200$ ,  $\alpha = 0.2$  and  $m = 10^{-6}$ : for this value of the mass we are almost above the quantum critical point and the mutual information decreases as we go away from the critical point by increasing the temperature. The error bars, estimated from the standard error of a mean of Monte Carlo trajectories with different initial conditions, are smaller than  $2 \times 10^{-2}$  for all points. (b) Mutual information between two halves of the chain  $I_{N/2}^{(2)}$  for  $N = 200$ ,  $\alpha = 0.2$  and  $m = 0.01$ . The background colours represent the value of  $|\langle \sigma_{N/2}^x \rangle|$  taken from the phase diagram (Fig. 2.7) for the same value of the mass: the mutual information increases as we enter the region of quantum criticality. The error bars are smaller than  $5 \times 10^{-3}$  for all points. Figure reproduced from [1].

have computed. It was already shown in [38] that  $\psi_0$  has a high overlap with the ground state of the critical XXZ chain with periodic boundary conditions:

$$H_{XXZ} = \sum_{j=1}^N \left( S_j^x S_{j+1}^x + S_j^y S_{j+1}^y + \Delta S_j^z S_{j+1}^z \right), \quad (2.43)$$

where  $S_i^\alpha = \frac{1}{2} \sigma_i^\alpha$  for  $\alpha \in \{x, y, z\}$  and the  $\sigma^\alpha$  are Pauli matrices. In this section we show that we can also find a Hamiltonian for the massive case.

For a general Hamiltonian, a thermal state is given by

$$\rho_{th} = e^{-\beta H} / \text{Tr} \left[ e^{-\beta H} \right]. \quad (2.44)$$

Since we know the density matrix but not the Hamiltonian, we can first define

$$H_{\rho_A} = -\log \rho_A, \quad (2.45)$$

where  $\rho_A$  is the thermal state describing the chain  $A$  at a finite temperature. Note that  $H_{\rho_A}$  may have a non-trivial dependence on the temperature and be non local, and so would not be a suitable Hamiltonian.

We observe that when the mass is close to zero and the temperature is high,  $H_{\rho_A}$  restricted to two-body interactions has the form of an  $H_{XXZ}$  Hamiltonian, up to some non translational invariant corrections. This suggests to look at  $H_{\rho_A}$  at high temperatures in the massive case as well.  $H_{\rho_A}$  restricted to two-body interactions has, up to some non translational invariant terms, the form

$$H_m = H_{XXZ} + H_m^0, \quad (2.46)$$

$$H_m^0 = \lambda \sum_{i=1}^N (-1)^i S_i^x + \mu \sum_{j=1}^{N-1} S_j^z S_{j+1}^y. \quad (2.47)$$

Let us take this Hamiltonian as a guess for a Hamiltonian realizing the thermal state. In the thermodynamic limit, this Hamiltonian is invariant under translations  $x_k \rightarrow x_k + 2$ , as is the wave function. Let us now define a thermal state  $\rho_H$  for the Hamiltonian  $H_m$  by:

$$\rho_H = e^{-\beta H_m} / \text{Tr} \left[ e^{-\beta H_m} \right]. \quad (2.48)$$

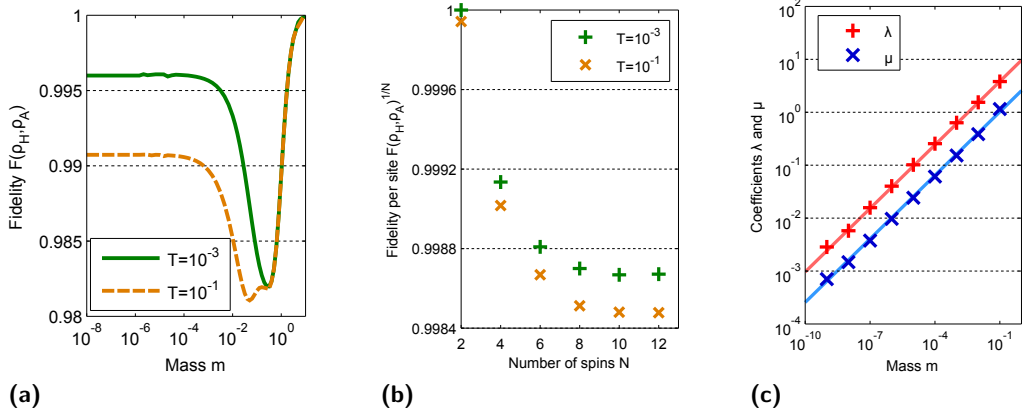
This state depends on the parameters  $\beta$ ,  $\Delta$ ,  $\lambda$  and  $\mu$ . A way to check whether this Hamiltonian corresponds to our system is to compute the fidelity [138, 139] between this state and the state  $\rho_A$ :

$$F(\rho_H, \rho_A) = \text{Tr} \left[ \sqrt{\sqrt{\rho_H} \rho_A \sqrt{\rho_H}} \right]. \quad (2.49)$$

For different values of  $\alpha$ ,  $m$  and  $T$  we optimize over the parameters  $\beta$ ,  $\Delta$ ,  $\lambda$  and  $\mu$  in such a way that the fidelity is maximized. The maximum for  $N = 12$  spins (24 spins in total for the two chains) is shown in Fig. 2.10a for different values of the mass and the temperature. When the mass is zero and the temperature is small, we recover the result from [38]: the fidelity is above 99% when  $\lambda$  and  $\mu$  are zero, so that the state is close to the ground state of the XXZ chain. In the massive case the thermal state from Hamiltonian  $H_m$  has a fidelity higher than 98% with the state of chain A for all values of the mass and  $T$  smaller than 0.1, which is very high considering the size of the Hilbert space: the fidelity per lattice site  $F^{1/N}$  is higher than 99.85% (Fig. 2.10b).

We can reduce the number of free parameters in the Hamiltonian by using the result in the massless case to fix the parameters  $\beta$  and  $\Delta$  with respect to  $T$  and  $\alpha$ . We then assume that these parameters do not depend on the mass  $m$ , and that all the dependence in the mass is captured by parameters  $\lambda$  and  $\mu$ . By optimizing over these





**Fig. 2.10.:** (a) Maximum of the fidelity  $F(\rho_H, \rho_A)$  between a thermal state obtained from a Hamiltonian of the form  $H_m$  and the state of chain A as a function of the mass, at  $\alpha = 0.2$  and different temperatures ( $N = 12$  spins). (b) Corresponding fidelity per site  $F(\rho_H, \rho_A)^{1/N}$  for different numbers of spins, at  $\alpha = 0.2$  and  $m = 10^{-1}$ . (c) Parameters  $\lambda$  and  $\mu$  in the Hamiltonian  $H_m$  for which the fidelity  $F(\rho_H, \rho_A)$  is maximal as a function of the mass, for  $\alpha = 0.2$ ,  $T = 10^{-3}$ ,  $N = 10$ . The two straight lines are fits of these data of the form  $\lambda = \lambda_0 m^{2\alpha}$ ,  $\mu = \mu_0 m^{2\alpha}$ . Figure reproduced from [1].

two remaining parameters, we find that there exist two functions  $\lambda(m)$  and  $\mu(m)$  for which the previous result still holds. In the regime  $m \in (0, 10^{-1}]$ ,  $\alpha < 1/4$ , these functions can be written as  $\lambda(m) = \lambda_0 m^{2\alpha}$  and  $\mu(m) = \mu_0 m^{2\alpha}$ , where  $\lambda_0$  and  $\mu_0$  are constants independent of the mass (Fig. 2.10c). The ansatz (2.40) therefore can be used in this regime to describe the thermal state of a spin chain governed by a Hamiltonian of the form:

$$H_m = \sum_{i=1}^{N-1} \left( S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + \Delta S_i^z S_{i+1}^z \right) + (-1)^i m^{2\alpha} \left( \lambda S_i^x + \mu S_i^z S_{i+1}^y \right) + m^{2\alpha} \lambda S_N^x. \quad (2.50)$$

This two-body Hamiltonian is local and contains only nearest neighbour interactions, so might be implemented in experiments.

## 2.6 Conclusion

Matrix Product States do not capture all properties of critical systems at a phase transition. This motivated the definition of infinite dimensional MPS, arising from the correlator of products of quantum fields. These ansatz wave functions can naturally capture the properties of critical systems. We have shown that this formalism is not limited to conformal fields, but that it also applies to other field theories,

which allowed us to construct a model wave function displaying a quantum phase transition. This model has an analytical wave function and can be investigated by Monte Carlo simulations, which enable us to study the entanglement properties and correlators of the spin chain. These quantities show that it presents a critical point as well as a non-critical phase at zero temperature. We have shown that by doubling the size of the system and considering the reduced state of one half of the system, one can construct a thermal state for a one-dimensional spin chain at finite temperature. The phase diagram of this model has then been investigated and it was shown that the finite temperature properties of the state reflect the phase transition at zero temperature. We also provided a Hamiltonian with only nearest neighbour interactions whose thermal state is very close to this model, showing that such an analytical construction is relevant for the description of realistic physical Hamiltonians.

## Fractional quantum Hall lattice models from infinite dimensional tensor networks

There exist other phase transitions which can arise in strongly correlated quantum systems and which do not admit a local order parameter. These occur between topological states of matter, which have attracted a lot of attention, both for the fundamental physics they display as well as for their potential practical applications in quantum computing. They were realized experimentally with the discovery of the fractional quantum Hall (FQH) effect[45], in which electrons of a two-dimensional electron gas subject to a strong magnetic field form an incompressible quantum fluid giving rise to fractionally charged excitations. A large understanding of the FQH effect was made possible by the discovery of analytical wave functions describing the electrons, such as Laughlin's wave function[63]. Experimental realizations of FQH states and manipulation of their quasiparticles however remain a challenge, which motivated a lot of research devoted to the realization of the FQH effect in different systems, such as lattice systems.

This chapter is concerned with the use of infinite dimensional tensor networks to build lattice models displaying FQH physics. Arising from correlators of field operators, the infinite dimensional tensor networks that we consider are natural lattice versions of continuum FQH wave functions such as Laughlin's wave function. Because the analytical expressions for the wave functions are known, their properties can be studied even for large sizes using Monte Carlo techniques. Wave functions for the quasiparticle excitations above these states can also be constructed, allowing for the study of the properties of anyons on lattices. This approach furthermore enables us to construct exact parent Hamiltonians for which these states are ground states. These Hamiltonians are not local, but in a few cases we are able to find local Hamiltonians that approximately realize the same states.

In section 3.1, we start by reviewing some properties of the FQH effect and of Laughlin wave functions in the continuum. We then show in section 3.2 how generalizations of the Kalmeyer-Laughlin states arise from infinite dimensional MPS. We study these states on several lattices and investigate when these states have the same topological properties as the corresponding continuum Laughlin states. We derive exact non-local parent Hamiltonians for which the infinite dimensional MPS are ground states for different fillings of the lattice, and discuss the properties of

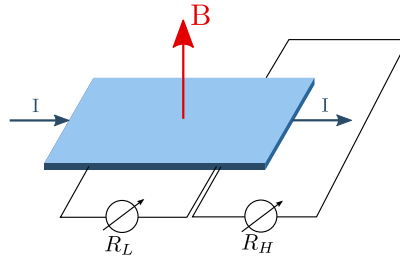
related edge states wave functions that are also eigenstates of the same Hamiltonians. In section 3.3, we turn to the investigation of the quasiparticle excitations of these states, which are anyons. We show that we can define analytical wave functions for both the quasihole and quasielectron excitations, which allows us to study their properties on the lattice as well as to derive exact Hamiltonians for which the quasiparticle wave functions are ground states. Section 3.4 finally generalizes the previous results to models of Moore-Read states with spin-1/2 or spin-1 lattices, which have non-Abelian anyons as their quasiparticles. We study the properties of these states along an interpolation between the continuum and the lattice, and derive exact parent Hamiltonians for these states. While these parent Hamiltonians are not local, we show that we can construct local Hamiltonians that stabilize some of the states we introduce. This chapter is based on and reuses parts of Refs. [2, 3, 5]:

- [2] : Ivan Glasser, J. Ignacio Cirac, Germán Sierra and Anne E. B. Nielsen, ‘Exact parent Hamiltonians of bosonic and fermionic Moore-Read states on lattices and local models’, *New Journal of Physics* **17**, 082001 (2015), under CC BY 3.0 license, <http://dx.doi.org/10.1088/1367-2630/17/8/082001>
- [3] : Ivan Glasser, J. Ignacio Cirac, Germán Sierra and Anne E. B. Nielsen, ‘Lattice effects on Laughlin wave functions and parent Hamiltonians’, *Physical Review B* **94**, 245104 (2016), ©2016 by the American Physical Society, <http://dx.doi.org/10.1103/PhysRevB.94.245104>
- [5] : Anne E. B. Nielsen, Ivan Glasser and Ivan D. Rodriguez, ‘Quasielectrons as inverse quasiholes in lattice fractional quantum Hall models’, *New Journal of Physics* **20**, 033029 (2018), under CC BY 3.0 license, <http://dx.doi.org/10.1088/1367-2630/aab5d0>

## 3.1 The fractional quantum Hall effect and Laughlin’s wave function

The fractional quantum Hall (FQH) effect[140, 141], discovered in 1982[45], arises in a 2D electron gas subject to a strong perpendicular magnetic field. One drives a current through the sample and measures the longitudinal as well as transverse Hall resistance (Fig. 3.1). Whereas classical physics predicts a linear dependence of the transverse Hall resistivity with the magnetic field, quantized plateaux are observed in experiments at values  $R_H = \frac{1}{\nu} \frac{h}{e^2}$  (Fig. 1.4), while the longitudinal resistance vanishes at the same points.  $\nu$  is a rational number known as the filling fraction. This quantization is so precise that it is nowadays used in the definition

of the electrical resistance standard. When  $\nu$  is an integer, this effect is known as the integer quantum Hall effect[142] and can be understood without taking into account interactions between the electrons. If we consider non-interacting electrons in a magnetic field, the quantized energy levels are known as Landau levels and are degenerate, with degeneracy given by the number of flux quanta through the surface of the sample. The filling fraction  $\nu$  is equal to the ratio of the number of electrons and the number of flux quanta. The integer quantum Hall effect appears when an integer number of Landau levels is filled, so when  $\nu$  is an integer. The incompressibility of the system is due to the energy gap needed to add an electron once a Landau level is filled.



**Fig. 3.1.:** Measurement of the Hall resistance  $R_H$  and longitudinal resistance  $R_L$  when a current  $I$  is driven through a 2D electron gas subject to a perpendicular magnetic field.

As such it was a surprise to find also an incompressible fluid when  $\nu = 1/3$ [45], and later at many more rational numbers. In this case, the interactions between the electrons have to be taken into account and a perturbative treatment is not possible, which renders a microscopic solution to the FQH effect an intricate problem. Effective descriptions of the state of the electrons, through trial wave functions, have contributed a lot to our understanding of the FQH effect.

These approaches started with the introduction of Laughlin's wave function[63]:

$$\psi_{\text{Laughlin}}(Z_1, \dots, Z_M) \propto \prod_{k < l} (Z_k - Z_l)^q e^{-\frac{1}{4} \sum_j |Z_j|^2}, \quad (3.1)$$

which represents the wave function of the electrons in the lowest Landau level, where  $Z_j$  are the positions of the  $M$  electrons on the complex plane and  $q$  is an integer related to the filling fraction by  $\nu = 1/q$ . The gaussian factor comes from the single particle wave functions, and the Jastrow factor in front vanishes when two electrons are at the same position. In this expression the magnetic length has been set to one. This wave function has very high overlap with the exact ground state of the original Hamiltonian with Coulomb interactions at  $q = 3$  for small number of electrons[143]. It can be shown to be the exact ground state of a gapped Hamiltonian in which the interactions have been modified to pseudopotentials[144].

A natural next step is therefore to ask what are the gapped excitations above this ground state. Such excitations can appear when slightly varying the filling factor away from  $\nu = 1/q$ , which can be done by changing the magnetic field slightly. This led Laughlin to suggest an ansatz for a localized quasiparticle at position  $w_0$ :

$$\psi_{\text{Laughlin}}^{qh}(Z_1, \dots, Z_M) \propto \prod_i (w_0 - z_i) \prod_{k < l} (Z_k - Z_l)^q e^{-\frac{1}{4} \sum_j |Z_j|^2}, \quad (3.2)$$

which is the wave function for an excitation when a flux quantum has been added to the system, thus slightly lowering the filling fraction. The probability of finding an electron at or around position  $w_0$  has been lowered, so the wave function creates a quasihole at this position. In order to compensate for the additional flux, this excitation needs to have a charge  $e/q$ . This shows that the quasihole carries a fractional charge, a property which distinguishes it from known elementary particles. A quasielectron with the same properties can also be defined, but it has turned out to be much harder to describe theoretically. This is because it is easier to modify the wave function to reduce the electron density locally than to increase it due to the Pauli exclusion principle[63, 141, 145]. Another fascinating property of these quasiparticles is that they are neither bosons nor fermions. In three dimensions, wrapping one particle adiabatically around another is equivalent to interchanging the particles twice, which can be continuously deformed to the identity operation. The wave function should be left unchanged, so under a single exchange the wave function can change only by a minus sign (for fermions) or not at all (for bosons). In two dimensions, however, when two particles are interchanged twice in a clockwise manner, their trajectory involves a nontrivial winding. The system does not necessarily come back to the same state, in which case the particles are called anyons[44]. Laughlin quasiholes are an example of such anyons, which acquire a phase, or statistics, of  $2\pi/q$  when braiding one quasihole around another. Systems with anyonic quasiparticles are topologically ordered, which also manifests itself in patterns of long-range entanglement. Indeed, one way to detect the presence of anyonic quasiparticles directly from the ground state wave function is to compute the topological entanglement entropy. Consider a region of the system with boundary of size  $L$ , its entanglement entropy for large sizes satisfies[40, 41]

$$S_L = \alpha L - \gamma, \quad (3.3)$$

where the first part of this expression is the area law, and the constant correction is the topological entanglement entropy, which depends on the number and properties of the anyons. Laughlin states have a topological entanglement entropy of  $\gamma_0(q) = \ln(\sqrt{q})$ [146].

Laughlin wave functions describe the state of the electrons at filling fractions  $1/q$ , but not at other filling fractions which have been observed experimentally. They have been generalized to states at other filling fractions through a hierarchical

construction describing states at filling factors with odd denominators[144, 147] or using composite fermions[148]. One way of generalizing the Laughlin wave function that is of particular interest for the rest of this chapter is the observation that the Laughlin state can be written as a correlator from CFT fields[68]. Indeed,

$$\psi_{\text{Laughlin}}(Z_1, \dots, Z_M) \propto \langle 0 | \prod_{j=1}^M : e^{i\sqrt{q}\varphi_R(Z_j)} :: e^{-i\sqrt{q}\rho_0 \int d^2Z \varphi_R(Z)} : | 0 \rangle, \quad (3.4)$$

where we recognize the vertex operators of a chiral free massless boson introduced in (2.26),  $\rho_0 = \frac{1}{2\pi q}$  is the electron density and the last term ensures that the charge neutrality condition is satisfied. One can also construct a quasihole operator from a vertex operator  $: e^{i\frac{1}{\sqrt{q}}\varphi_R(w_0)} :$ , and inserting this operator in the correlator leads to the Laughlin quasihole wave function.

Since its discovery in 1987 [149], one FQH state has attracted a lot of attention: the  $\nu = 5/2$  FQH state with electrons occupying the second Landau level cannot be explained by a hierarchical construction based on the Laughlin states. The leading candidate for its description is the Moore-Read ‘‘Pfaffian’’ state at filling  $1/2$  [68, 150, 151], describing the wave function of the electrons in the second Landau level. It can be constructed as a correlator of other CFT fields, leading to a wave function written as

$$\psi_{\text{MR}}(Z_1, \dots, Z_M) \propto \text{Pfaffian} \left( \frac{1}{Z_i - Z_j} \right) \prod_{k < l} (Z_k - Z_l)^2 e^{-\frac{1}{4} \sum_j |Z_j|^2}, \quad (3.5)$$

for which the quasiparticle excitations are fractionally charged anyons displaying non-Abelian braiding statistics[68, 152–155], which can be used for topological quantum computation[43, 44].

## 3.2 Lattice effects on Laughlin wave functions and parent Hamiltonians

To realize FQH physics on lattices, several approaches have been suggested. In one approach, one tries to mimic the interactions which give rise to the FQH effect in a lattice system. This leads to the proposal of flat-band models which mimic the Landau levels[47–51], where the magnetic field is replaced by a complex hopping phase on the lattice. Another approach are chiral spin liquids, in which spins on a lattice form a collective state that has the same topological properties as FQH states. These states were first introduced with the Kalmeyer-Laughlin state[52], a bosonic Laughlin state at filling fraction  $\nu = 1/2$  which has the same expression as the Laughlin state with  $q = 2$  (3.1), but with positions of particles limited to the lattice sites. Over the last decades, there has been a lot of interest in finding spin

Hamiltonians that have such a state as their ground state. Parent Hamiltonians have been obtained for the Kalmeyer-Laughlin state and some of its generalizations [54, 55, 69, 156–160]. Recently several local Hamiltonians on the square or Kagome lattice have been shown to exhibit such a chiral spin liquid as their ground state [54–62].

The previous examples of the realization of the Kalmeyer-Laughlin state on spin- $1/2$  lattices typically occur either at half-filling of the lattice (half of the spins are pointing up in the  $z$ -direction) or at low filling of the lattice (so the system is close to a continuum limit). As we will see we can in principle define such a state also on lattices with different fillings. Indeed it was suggested that this state can emerge on a lattice at filling  $1/3$  [161], which was recently observed numerically for a Bose-Hubbard model on the Kagome lattice [162]. It is also of interest to understand whether the lattice may destroy the physics of the FQH effect. Indeed, it is not a priori clear that a wave function of a FQH state in the continuum still displays the same topological effects once discretized and placed on a lattice, and we will see in this section that while in most cases this happens [54, 69, 163–165], there are cases where the lattice destroys the FQH physics.

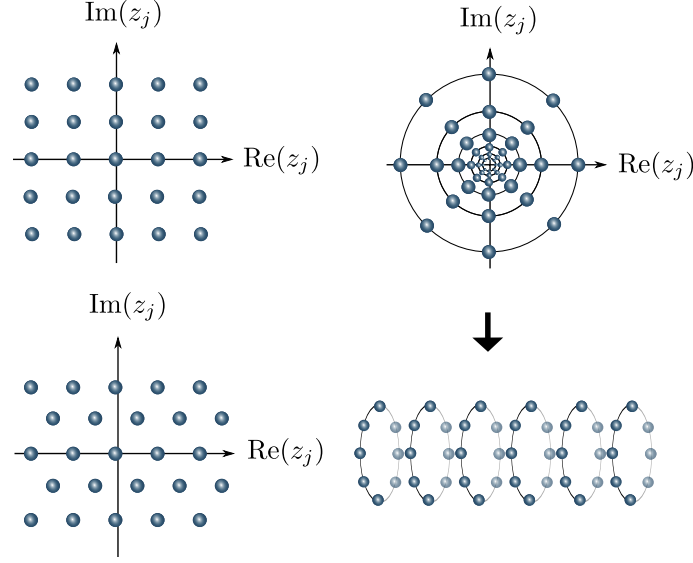
In this section we define lattice versions of Laughlin wave functions on arbitrary lattices, with lattice filling factor not necessarily equal to  $\nu$ , and investigate the phase diagram using measures of correlations and entanglement entropy. We find evidence that in a large part of the phase diagram the states on a square lattice have the same topological properties as the continuum Laughlin states, but that particular particle-hole symmetric states are topologically trivial. This behavior does not subsist if the lattice is deformed to a frustrated lattice, like the triangular and the Kagome lattice. We study edge states wave functions obtained by adding a charge outside the boundary of the lattice, and parent Hamiltonians are constructed for both edge states as well as the original lattice Laughlin states. This section is based on Ref. [3] and is a modification thereof.

### 3.2.1 Lattice Laughlin states at different lattice filling factors

In this subsection we start by defining Laughlin wave functions with number of particles per flux  $\nu = 1/q$  ( $q \in \mathbb{N}$ ) on lattices with arbitrary lattice filling factor. Let us consider a lattice with  $N$  sites at positions  $z_j$ ,  $j \in \{1, 2, \dots, N\}$  in the complex plane. Particular choices of the coordinates  $z_j$  will for example yield a square or triangular lattice on the plane or on the cylinder, as pictured in Fig. 3.2.

The systems we consider in this section can be spin systems with spin  $1/2$  or fermionic systems. In order to provide a joint treatment of these cases, we map the spin  $1/2$  to hardcore bosons, where a spin down corresponds to the absence of a boson, and





**Fig. 3.2.:** Example of lattices considered in this work: square lattice (upper left), triangular lattice (bottom left), square lattice on a cylinder (right), obtained by mapping the plane to the cylinder. The coordinates for the square lattice on the cylinder are  $z_j = \exp(2\pi((x_j - L_x/2 + 1/2)/L_y + (y_j - L_y/2 + 1/2)i/L_y))$ ,  $x_j \in \{0, \dots, L_x - 1\}$ ,  $y_j \in \{0, \dots, L_y - 1\}$ .

a spin up to the presence of a boson. We can then consider a local Hilbert space  $|n_i\rangle$  such that the lattice site can be empty ( $n_i = 0$ ) or occupied ( $n_i = 1$ ), either by a hardcore boson or a fermion. In the next sections we will also consider spin-1 systems, which are in this case mapped to a bosonic system with at most two bosons per site, corresponding to  $n_i = 2$ . We consider  $M$  particles hopping on this lattice and define  $\mu \equiv \frac{M}{N}$  as the lattice filling factor, which shouldn't be confused with the filling factor  $\nu$  in the FQH effect.

In general, a wave function defined on this lattice will have the form

$$|\psi\rangle = \sum_{n_1, \dots, n_N} \psi(n_1, \dots, n_N) |n_1, \dots, n_N\rangle, \quad (3.6)$$

where  $\psi(n_1, \dots, n_N) = 0$  unless the number of particles is  $\sum_i n_i = M$  (if the particles are fermions, the Fock states are defined using the same ordering as for the lattice sites).

Let  $q$  be a positive integer. We would like to define Laughlin wave functions at  $\nu = 1/q$  on this lattice. For this we merge the ideas of writing the Laughlin wave function as a correlator of CFT operators[68] with the ideas presented in section 2.2

of infinite dimensional MPS also arising from correlators of field operators but on a lattice[38], making the choice of an infinite dimensional MPS written as

$$\psi(n_1, \dots, n_N) \propto \langle \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_N}(z_N) \rangle, \quad (3.7)$$

where  $\mathcal{V}_{n_j}(z_j)$  are operators attached at position  $z_j$ . We choose as operators modified versions of the vertex operators appearing in the continuum description of the Laughlin state:

$$\mathcal{V}_{n_j}(z_j) = \begin{cases} : e^{-i\frac{\eta}{\sqrt{q}}\varphi_R(z_j)} : & \text{if } n_j = 0, \\ e^{i\pi\eta'(j-1)} : e^{i\frac{q-\eta}{\sqrt{q}}\varphi_R(z_j)} : & \text{if } n_j = 1, \end{cases} \quad (3.8)$$

where  $\varphi_R(z)$  is the chiral bosonic field introduced in (2.21) and the operators are vertex operators as introduced in (2.26),  $\eta$  is a positive rational number such that  $\eta N/q$  is an integer and in all this work  $\eta'$  denotes  $\min(\eta, q - \eta)$ . If  $\eta$  is not an integer, a choice of branch cuts can be made consistently for all the formulas in this work. We note here that an alternative way of implementing MPS ideas to 2D FQH systems has been developed in [166, 167], where MPS ideas apply to continuum systems.

Evaluating the correlator in (3.7) yields[32, 160] a wave function, that we denote as  $\psi_q^\eta$ , such that

$$\psi_q^\eta(n_1, \dots, n_N) \propto \delta_n \xi_\eta \prod_{i < k} (z_i - z_k)^{q n_i n_k} \prod_{j \neq l} (z_l - z_j)^{-\eta n_l}, \quad (3.9)$$

where  $\delta_n$  is zero unless the number of particles is  $M = \sum_i n_i = \eta \frac{N}{q}$  and  $\xi_\eta$  is 1 if  $\eta \leq q - \eta$  or  $(-1)^q \sum_j (j-1)n_j$  otherwise. The last factor in (3.9) represents the effect of the lattice on a given particle, that in the continuum is generated by the background charge. Note that no background charge needs to be inserted to evaluate the correlator and charge neutrality is directly ensured by the choice of operators (3.8).

It is important to note that the filling factor of the lattice is  $\mu = \frac{M}{N} = \frac{\eta}{q}$ , which can be different from the Laughlin filling fraction  $\nu = 1/q$ . Here  $\mu$  is the number of particles per lattice site, whereas  $\nu$  is the number of particles per number of fluxes. In this work we will be particularly interested in lattice effects in the cases where  $\eta \neq 1$ , so that  $\mu \neq \nu$ .

These wave functions have been studied in detail for particular values of  $q$  and  $\eta$ . It was shown that the state at  $q = 2, \eta = 1$  can be realized with a local Hamiltonian that can be implemented in optical lattices[54, 69, 168], and non-local parent Hamiltonians have been obtained for the states at  $\eta = 1$ [160]. It was shown for lattices defined on the plane that in the thermodynamic limit and when the area

$a$  of each site is the same the state reduce to the Kalmeyer-Laughlin states in the continuum limit. This can be seen by observing that

$$\prod_{j \neq l} (z_l - z_j)^{-\eta n_l} \propto e^{-i \sum_l g_l} e^{-\frac{1}{4} \frac{2\pi\eta}{a} \sum_l |z_l|^2 n_l}, \quad (3.10)$$

where  $g_l \equiv \text{Im}[\eta \sum_{j(\neq l)} n_l \ln(z_l - z_j)]$  is a real number. In the continuum, the wave functions are expressed in the basis spanned by the position of the particles  $Z_i$ . Here the  $Z_i$  are the positions  $z_i$  where  $n_i = 1$  and the  $(Z_1, \dots, Z_M)$  form a basis of the Hilbert space. The wave function written in this basis then becomes, in the thermodynamic limit (here the phase factors, which can be transformed away if desired, are omitted)

$$\Psi_{\text{Laughlin}}(Z_1, \dots, Z_M) \propto \prod_{i < j} (Z_i - Z_j)^q e^{-\frac{1}{4} \frac{2\pi\eta}{a} \sum_l |Z_l|^2}, \quad (3.11)$$

which is the Kalmeyer-Laughlin wave function at filling fraction  $\nu = \frac{1}{q}$  for  $M$  particles with positions restricted to the lattice sites. This means that the wave functions we introduce are equivalent to the Kalmeyer-Laughlin wave functions in the thermodynamic limit, but slightly different on finite lattices.

Let us show that this property is still valid on a cylinder geometry. To define the states on the cylinder, the coordinates are taken to be of the form  $z_j = e^{x_j + iy_j}$ , where  $x_j + iy_j$  define the corresponding lattice on the plane. The previous factor can then be computed by writing

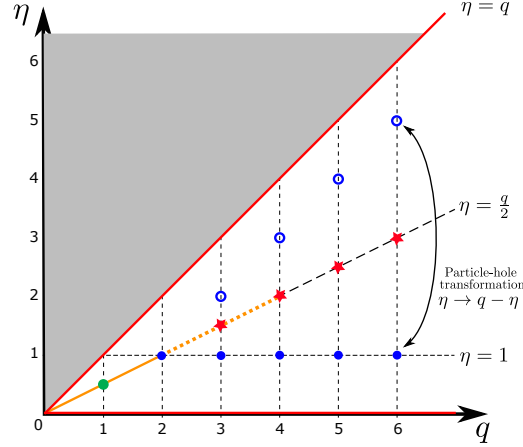
$$\prod_{j \neq l} (z_l - z_j)^{-\eta n_l} \propto e^{-\sum_l n_l \sum_{j(\neq l)} \eta \log(z_l - z_j)}, \quad (3.12)$$

and in the thermodynamic limit the real part of the sum can be replaced by an integral on the cylinder:

$$\sum_{j(\neq l)} \eta \log(|z_l - z_j|) \propto \frac{\eta}{a} \int_{x=0}^R \int_{y=0}^{2\pi} \log(|z_l - e^x e^{iy}|) dy dx, \quad (3.13)$$

which can be evaluated to yield  $-\frac{2\pi\eta}{a} \frac{|x_l|^2}{2} + \text{constant}$ , where the constant does not depend on  $z_l$ . In the thermodynamic limit on the cylinder the wave function therefore also reduces to the Kalmeyer-Laughlin wave function with positions restricted to the lattice sites.

Let us first discuss a few particular values of the parameters  $q$  and  $\eta$ . In Fig. 3.3, we draw a diagram of the  $\psi_q^\eta$  states for the different values of the parameters that will be considered in this work. When  $\mu = 0$  (resp.  $\mu = 1$ ), all the lattice sites are empty (resp. occupied), so the state is a trivial ferromagnet. This happens when  $\eta = 0$  or when  $\eta = q$  (red lines in Fig. 3.3). When  $\eta > q$ , the number of particles



**Fig. 3.3.:** Diagram of the  $\psi_q^\eta$  states for the values of  $q$  and  $\eta$  considered in this work.  $q$  is integer while  $\eta$  is a rational number in the interval  $[0, q]$ , since the states for  $\eta > q$  (grey region) are not defined. When  $\eta \rightarrow 0$  or  $\eta \rightarrow q$  (red lines), the state is a ferromagnet if  $N$  is fixed, or a Laughlin state in the continuum if  $N$  goes to infinity and  $N \propto 1/\eta$  (if  $\eta \rightarrow 0$ ) or  $N \propto 1/(q - \eta)$  (if  $\eta \rightarrow q$ ). The solid blue disks on the line  $\eta = 1$  correspond to lattice Laughlin states that satisfy  $\mu = \nu = 1/q$ . The green disk is an integer quantum Hall state. On the orange line ( $q$  not necessarily integer), the state in 1D is critical, while on the dotted orange line it has long range order. Figure reproduced from [3].

in the lattice has to be larger than the number of lattice sites, which is not possible since only single occupancy is allowed. The state  $\psi_q^\eta$  is therefore not defined for  $\eta > q$  (grey region in Fig. 3.3). In the limit where  $\eta \rightarrow 0$  but the number of particles  $M$  is kept fixed while  $N \propto 1/\eta \rightarrow \infty$ , the states become Laughlin states in the continuum[160].

One particularly interesting case is the choice  $\eta = q/2$ , which corresponds to half-filling of the lattice. In one dimension, the states defined at half-filling (and  $q$  not necessarily integer) have been shown[38] to be critical states for  $q \leq 2$  (orange line in Fig. 3.3) and to exhibit antiferromagnetic long-range order for  $2 < q < 4$  (dotted orange line in Fig. 3.3). At  $q = 1$ ,  $\eta = 1/2$  (green disk), the state in 2D corresponds to an integer quantum Hall state[38]. Note that the states are bosonic for even  $q$  and fermionic for odd  $q$ .

### 3.2.2 Properties of the lattice Laughlin states on the square lattice

In this subsection we explore the properties of the lattice Laughlin States to determine the phase diagram of the states in the  $(q, \eta)$  plane. Particle-hole transformation shows that it is symmetric along the  $\eta = q/2$  line, on which the states are particle-hole symmetric. We then give numerical evidence that the states have exponentially decaying correlations on the square lattice unless  $\eta = q/2$ ,  $q \geq 5$ , and that when

this is the case the states have the same topological entanglement entropy as the Laughlin states in the continuum.

### Particle-hole transformation

The different wave functions  $\psi_q^\eta$  are related under particle-hole transformation. This transformation has the effect of exchanging  $n_i$  and  $1 - n_i$ , so the vertex operators used in the definition of the wave function are transformed as

$$\mathcal{V}_{n_j}(z_j) = e^{i\pi\eta'(j-1)n_j} : e^{i\frac{qn_j - \eta}{\sqrt{q}}\phi(z_j)} : \longrightarrow \mathcal{V}'_{n_j}(z_j), \quad (3.14)$$

where

$$\mathcal{V}'_{n_j}(z_j) \equiv e^{i\pi\eta'(j-1)} e^{-i\pi\eta'(j-1)n_j} : e^{-i\frac{qn_j - (q-\eta)}{\sqrt{q}}\phi(z_j)} : . \quad (3.15)$$

Now let us define  $\tilde{\mathcal{V}}_{n_i}(z_j)$  the operators used to define the state  $\psi_q^{q-\eta}$  and observe that

$$\langle \mathcal{V}'_{n_1}(z_1) \dots \mathcal{V}'_{n_N}(z_N) \rangle \propto \langle \tilde{\mathcal{V}}_{n_1}(z_1) \dots \tilde{\mathcal{V}}_{n_N}(z_N) \rangle. \quad (3.16)$$

This shows that the states  $\psi_q^\eta$  and  $\psi_q^{q-\eta}$  are exchanged under particle-hole transformation (see Fig. 3.3). Note that this transformation also changes the number of particles, so that it relates states at lattice filling factor  $1/q$  and  $1 - 1/q$ . This can be compared to the situation in the continuum, where a particle-hole transformation can be defined[169] to relate the FQH states at filling fraction  $1/q$  and  $1 - 1/q$ . On the lattice the particle-hole transformation is however different, because it only involves exchanging the  $|0\rangle$  and  $|1\rangle$  states, so there is no separate treatment of an electron or a hole. The  $\psi_q^\eta$  and the  $\psi_q^{q-\eta}$  states are therefore related by a simple change of basis, and as such all properties of the wave function that are symmetric with respect to particle-hole transformation will be the same for the  $\psi_q^\eta$  and the  $\psi_q^{q-\eta}$  states. This is in particular the case of the connected particle-particle correlation function, since  $\langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle = \langle (1 - n_i)(1 - n_j) \rangle - \langle 1 - n_i \rangle \langle 1 - n_j \rangle$ .

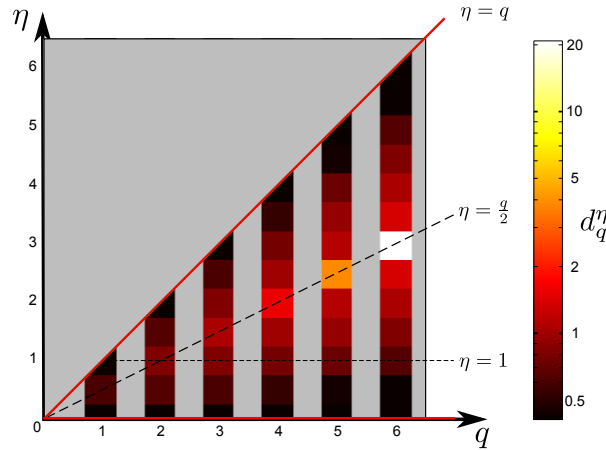
The states at  $\eta = q/2$  are special with respect to this transformation, since these states are particle-hole symmetric: the operators used to define these states are  $: e^{\pm i\sqrt{q}/2} : .$  This is in particular the case for the bosonic  $\psi_2^1$  state, as is the case for the bosonic  $\nu = 1/2$  Laughlin state in the continuum. For larger values of  $q$ , the Laughlin states in the continuum are not particle-hole symmetric. However on the lattice we can change independently the lattice filling factor  $\mu$  to  $1/2$ , which amounts to considering the states  $\psi_q^{q/2}$  (red stars in Fig. 3.3). The properties of these states will be investigated in more details in the following.

## Correlations in the parameter diagram

To investigate possible different phases in the  $(q, \eta)$ -diagram, we compute the connected particle-particle correlation function  $C_{ij} \equiv \langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle$  in the bulk of the system for a square lattice on the cylinder, following the algorithm introduced in section 2.2.3. We find that the correlation function decays exponentially with the distance for all values of  $q$  and  $\eta$ , except for  $q \geq 5$  at  $\eta = q/2$ . In Fig. 3.4, we explore the parameter range by using as parameter the correlation length estimated from nearest-neighbours and next-nearest neighbours correlations in the bulk of the system. More precisely we compute a characteristic length, defined as

$$d_q^\eta = \frac{1}{\ln(C_{i(i+1)}) - \ln(C_{i(i+2)})}, \quad (3.17)$$

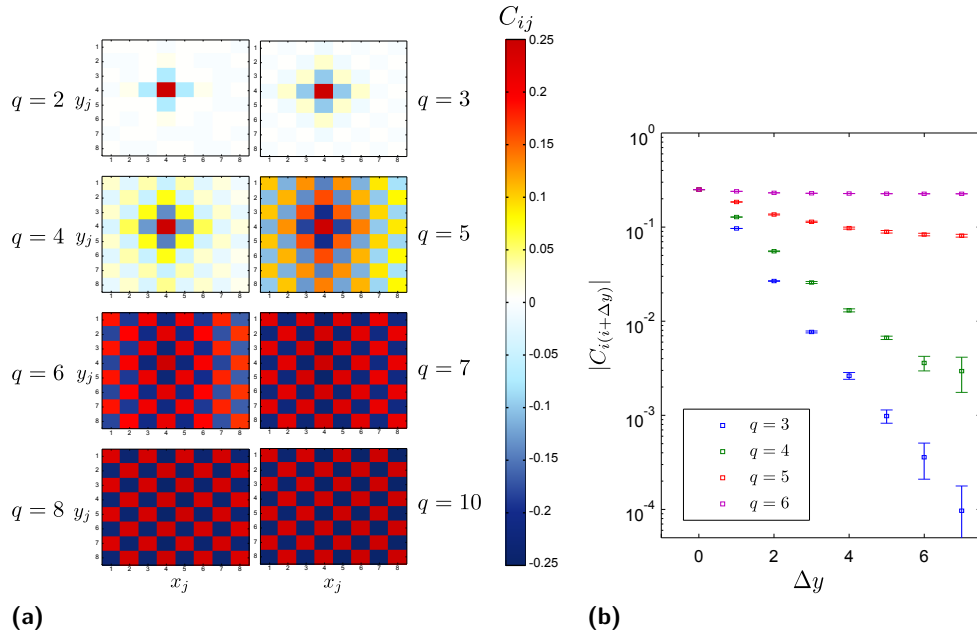
where  $i, i + 1$  and  $i + 2$  index neighbouring lattice sites in the periodic direction that are located on a ring in the middle of the cylinder. This parameter is an estimate on how fast the correlations decay. When the decay is exponential with a short correlation length, this quantity almost coincides with the correlation length. When this characteristic length is larger, further investigations are performed to characterize the decay of correlations.



**Fig. 3.4.:** Characteristic length  $d_q^\eta$  (3.17) estimating the correlation length of the  $\psi_q^\eta$  states. The characteristic length is computed for  $q$  integer and  $\eta$  integer and half-integer using a Metropolis-Hastings algorithm on a square  $12 \times 12$  (or  $10 \times 12$  for  $q = 5$ ) lattice on the cylinder. The decay is found to be exponential everywhere, except when  $\eta = q/2, q \geq 5$ . Figure reproduced from [3].

We first observe that for a given  $q$  the correlation length decreases when  $\eta$  goes away from the half-filling point  $\eta = q/2$ . At half-filling, we find that the states  $\psi_5^{2.5}$  and  $\psi_6^3$  do not display exponentially decaying correlations, although for  $\eta$  slightly different than  $q/2$  the corresponding  $\psi_q^\eta$  have exponentially decaying correlations. In Fig. 3.5, we show the decay of correlations in two dimensions and along the periodic direction of the cylinder for the half-filled states at different values of  $q$ . It

is observed that for  $q \leq 4$  the correlations decay exponentially. For  $q \geq 6$ , the states display clear long range anti-ferromagnetic correlations. The properties of the states at these particular points will be investigated in more details in section 3.2.3. For  $q = 5$ , the correlations decay and results on larger sizes suggest that long range anti-ferromagnetic order is also present at larger scales. In addition, we observe that Monte Carlo simulations at  $\eta = q/2$  need more computational effort to converge, in part due to the larger number of particles and in part due to the structure of the wave functions. The computations can be improved by exploiting the particle-hole symmetry at these points to allow global change of configurations respecting this symmetry in the Monte Carlo trajectories.

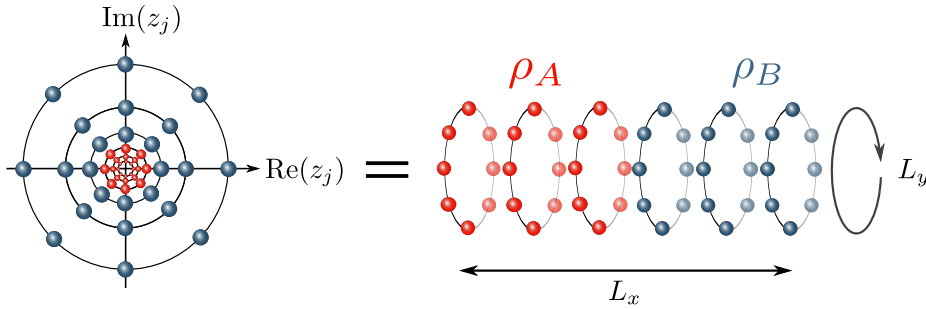


**Fig. 3.5.:** (a) Correlation function  $C_{ij}$  between a fixed lattice site  $i$  in the middle of the lattice and all other lattice sites  $j$  in the state  $\psi_q^{q/2}$  on a square  $8 \times 8$  lattice on the cylinder. (b) Absolute value of the correlation function  $|C_{i(i+\Delta y)}|$  between two lattice sites separated by a distance  $\Delta y$  in the periodic direction in the state  $\psi_q^{q/2}$  on a square  $16 \times 16$  lattice on the cylinder. There is long range anti-ferromagnetic order for  $q \geq 6$ , while the correlations decay exponentially for  $q \leq 4$ . Figure reproduced from [3].

### Topological entanglement entropy of the $\psi_q^\eta$ states

So far we have only checked that the states display exponentially decaying correlations in most of the parameter range. We now turn to the investigation of their topological properties. To compute the topological entanglement entropy (TEE) of the states, we start with a lattice on the cylinder, cut the cylinder into two halves and compute the second Renyi entropy of the first half (Fig. 3.6). The size along the

cut is  $L_y$  and we use the behaviour of the entanglement entropy (3.3) to extract the topological entanglement entropy  $\gamma$ , as has already been done using Monte Carlo simulations for several chiral spin liquids for which the wave function is known [2, 69, 160, 170]. The results for the square lattice at  $q = 4$  and different values of  $\eta$ , shown in Fig. 3.7(a), confirm that the  $\psi_4^\eta$  have the same topological entanglement entropy as the continuum Laughlin state at filling  $1/4$ , independently of the value of the lattice filling factor  $\mu$ . A particularly interesting case is the state  $\psi_4^2$ , defined on a half-filled lattice: this state is particle-hole symmetric, but has the topological entanglement entropy of the continuum Laughlin state at filling fraction  $\nu = 1/4$ , which itself is not particle-hole symmetric. The same observation remains true at  $q = 2$  and  $q = 3$ , where the particle-hole symmetric lattice state have the same topological entanglement entropy as the continuum Laughlin state at  $\nu = 1/q$  (see Fig. 3.7(b)). However, as we have seen previously the states  $\psi_q^{q/2}$  for  $q$  larger than 5 on the square lattice have long range anti-ferromagnetic correlations. They do not display non-trivial topological behaviour, as will be clear from more investigations in the next subsection.



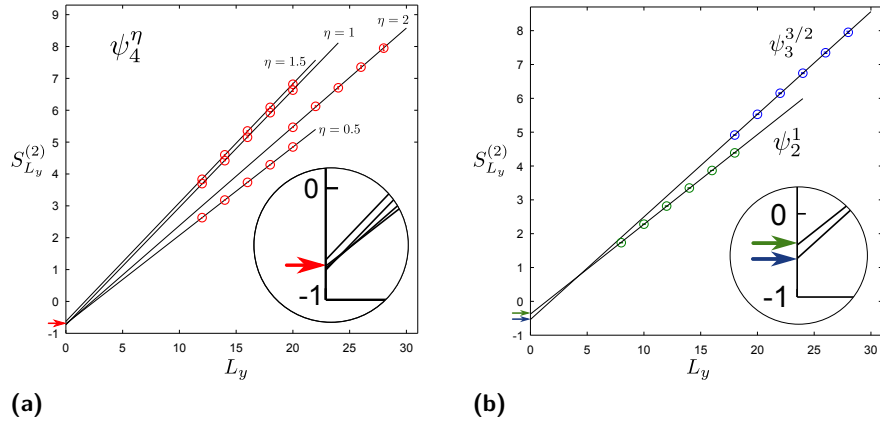
**Fig. 3.6.:** To compute the entanglement entropy of the state, the cylinder is cut into two halves and the Renyi entropy of the first half is computed using a Metropolis-Hastings algorithm. The topological entanglement entropy is extracted by varying the size  $L_y$  of the cylinder. Figure reproduced from [3].

Another tool that we can use to characterize the topological properties of these states is the braiding properties of their anyonic excitations and the localization of quasihole wave functions. This will be studied for these states in section 3.3.

### 3.2.3 States at half-filling: from long-range order on the square lattice to topological order on frustrated lattices

So far we have investigated the properties of the states on a square lattice. As we have seen, the states on this lattice are lattice versions of the continuum Laughlin states, except at half-filling when  $q \geq 5$  where the states display long-range anti-ferromagnetic order. We now focus on these particular values of the parameters and investigate this effect in more details. We explain the behaviour of the correlations





**Fig. 3.7.:** (a) Scaling of the Renyi entropy  $S_{L_y}^{(2)}$  of the state  $\psi_4^\eta$  on a  $L_x \times L_y$  square lattice on the cylinder. The topological entanglement entropy of the Laughlin state at filling  $1/4$  ( $\gamma_0(4) = \ln(2) \approx 0.693$ ) is indicated with a red arrow. The size  $L_x$  is taken to be 12, unless  $\eta = 2$  in which case  $L_x$  is 20. Larger sizes are taken when  $\eta = q/2$ , to account for the longer correlation length and to get rid of finite size effects. The black lines are linear fits and the values found for the topological entanglement entropy when  $\eta$  equals to 0.5, 1, 1.5 and 2 are respectively 0.698, 0.734, 0.643 and 0.718. (b) Scaling of the Renyi entropy  $S_{L_y}^{(2)}$  for the states at half-filling  $\psi_2^1$  and  $\psi_3^{3/2}$  on the square lattice on the cylinder. Here  $L_x$  is 12 for  $q = 2$  and 20 for  $q = 3$ . The topological entanglement entropy of the Laughlin state at filling  $1/2$ ,  $\gamma_0(2) \approx 0.346$  (resp. at filling  $1/3$ ,  $\gamma_0(3) \approx 0.549$ ) is indicated with a green (resp. blue) arrow. The values found for the topological entanglement entropy are respectively 0.375 and 0.536.

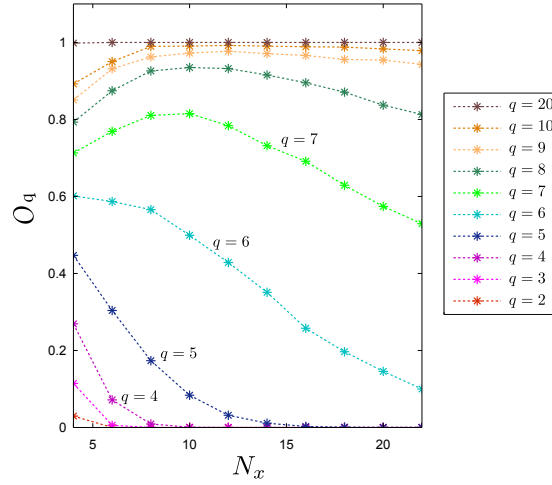
Figure reproduced from [3].

on the square lattice and show that long-range order is not present on frustrated lattices, where the expected topological behaviour is recovered.

### Long range antiferromagnetic order and Néel states

The correlations previously computed show that for large values of  $q$  and  $\eta = q/2$ , the state on the square lattice has long range anti-ferromagnetic correlations. To understand this, let us look at the behaviour in the limit of infinite  $q$ . The dominating term in the wave function has the form  $\prod_{i < j} (z_i - z_j)^{q n_i n_j}$ , and only terms where the two positions  $z_i$  and  $z_j$  are occupied by a particle contribute to the wave function. The main contribution to the wave function therefore comes from states that maximize the product of the distances between pairs of particles on the lattice. In the infinite  $q$  limit, only the states maximizing this product contribute to the wave function and the contribution of the other terms is suppressed. Since the lattice is half-filled, it is not possible to put particles too far apart to maximize this product. For the square lattice, the maximum is obtained when lattice sites alternate between empty

and occupied sites in a checkerboard pattern. There are two such possibilities, corresponding to the two Néel states, that we denote  $\psi_{\text{Néel}}^1$  and  $\psi_{\text{Néel}}^2$  and that are related by particle-hole transformation.



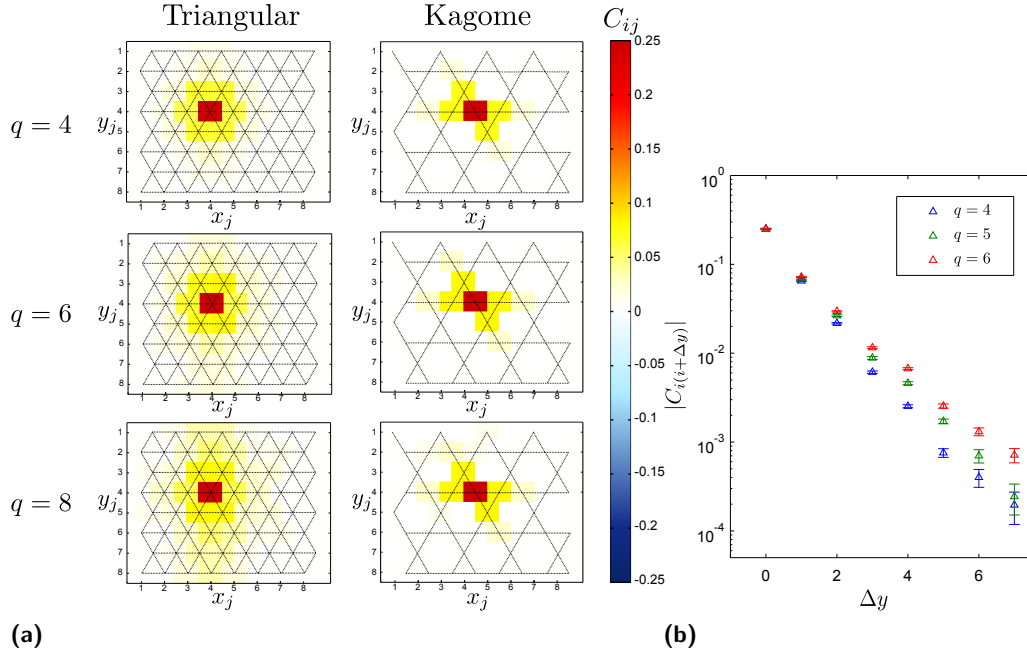
**Fig. 3.8.:** Overlap, defined as  $O_q \equiv |\langle \psi_{\text{Néel}}^1 | \psi_q^{q/2} \rangle|^2 + |\langle \psi_{\text{Néel}}^2 | \psi_q^{q/2} \rangle|^2$ , between the state at half-filling  $\psi_q^{q/2}$  and the two Néel states, computed using a Metropolis-Hastings algorithm on a square lattice of size  $N_x \times N_x$  on the cylinder. The errors from the Monte Carlo simulations are below 0.02 for all points. Figure reproduced from [3].

To investigate how much these two Néel states contribute to the wave function, we compute the overlaps between each Néel state and the state  $\psi_q^{q/2}$  for different lattice sizes (see Fig. 3.8). For large values of  $q$ , it is found that  $O_q \equiv |\langle \psi_{\text{Néel}}^1 | \psi_q^{q/2} \rangle|^2 + |\langle \psi_{\text{Néel}}^2 | \psi_q^{q/2} \rangle|^2$  is close to one, which shows that the wave function is almost a simple superposition of  $\psi_{\text{Néel}}^1$  and  $\psi_{\text{Néel}}^2$ . For smaller values of  $q$ , this quantity goes rapidly to zero, while there is a transition in the range  $5 \leq q \leq 7$  where the overlap goes to zero but remains high, especially considering the size of the Hilbert space for the lattices considered. This explains the behaviour of the correlations, which decay exponentially for small  $q$  but have long range order for large  $q$ . For other bipartite lattices which can support Néel states the argument given for infinite  $q$  remains valid, so we expect a behaviour similar as for the square lattice. However the transition range might happen for different values of  $q$ .

### Frustration destroys the anti-ferromagnetic order

The two Néel states can arise because of the geometry of the square lattice. In this particular case, strong lattice effects at half-filling can give rise to anti-ferromagnetic behaviour. On frustrated lattices, we cannot define two Néel states respecting the symmetries of the lattice. We therefore now turn to investigations of the  $\psi_q^{q/2}$  on the triangular and Kagome lattices. Fig. 3.9 shows evidence that the particle-particle

correlations on the triangular and Kagome lattices decay exponentially with the distance even for  $q = 6$ , in sharp contrast with the behavior on the square lattice.

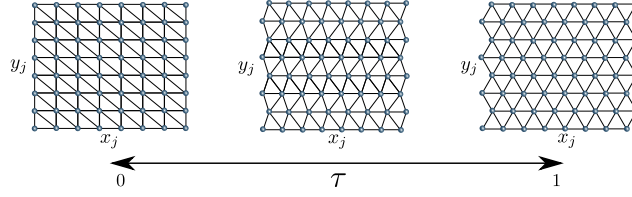


**Fig. 3.9.:** (a) Correlation function  $C_{ij}$  between a fixed lattice site  $i$  in the middle of the lattice and all other lattice sites  $j$  in the state  $\psi_q^{q/2}$  on a triangular and Kagome lattice on the cylinder. (b) Absolute value of the correlation function  $|C_{i(i+\Delta y)}|$  between two lattice sites separated by a distance  $\Delta y$  in the periodic direction in the state  $\psi_q^{q/2}$  on a triangular  $16 \times 16$  lattice on the cylinder. Unlike on the square lattice for the same parameters, there is no long range order. Figure reproduced from [3].

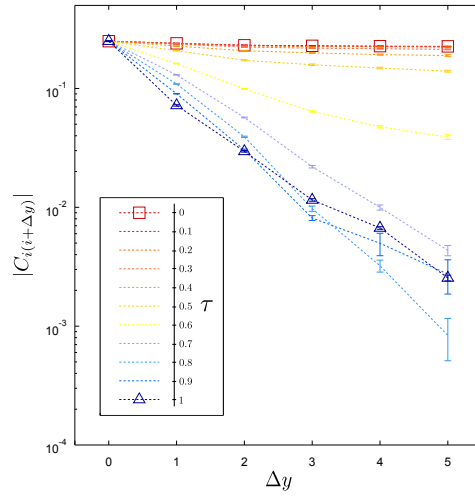
Since the states are defined as a function of the positions  $z_i$  of the lattice points, we can study the transition between the square and the triangular lattice by changing continuously the coordinates  $z_i$  to interpolate linearly between a square lattice and a triangular lattice on the cylinder (Fig. 3.10). The decay of correlations for  $q = 6$  is shown in Fig. 3.11 for different values of the interpolation parameter  $\tau$ . We observe that the correlations decay exponentially when the lattice is close to a triangular lattice ( $0.7 \leq \tau \leq 1$ ), while there is a transition towards anti-ferromagnetic order when the lattice gets closer to the square lattice. This confirms the importance of the geometry of the lattice in the study of the wave functions  $\psi_q^{q/2}$ .

### Topological entanglement entropy on the triangular lattice

It was shown previously that the long range order was destroyed on frustrated lattices. We might therefore wonder whether these half-filled states on the triangular



**Fig. 3.10.:** Positions of the coordinates of the lattice sites along the interpolation between a square and a triangular lattice. The coordinates along the periodic direction  $y_j$  are kept fixed to keep the periodicity, while the coordinates along the other direction  $x_j$  are linearly interpolated between the square ( $\tau = 0$ ) and the triangular ( $\tau = 1$ ) lattice. The coordinates on the plane are then  $z_j = e^{\frac{2\pi}{Ly}(x_j + iy_j)}$ , as in Fig. 3.6. Figure reproduced from [3].

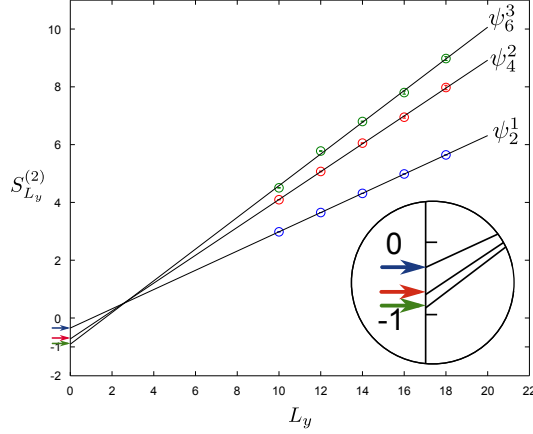


**Fig. 3.11.:** Absolute value of the correlation function  $|C_{i(i+\Delta y)}|$  between two lattice sites separated by a distance  $\Delta y$  in the periodic direction in the state  $\psi_6^3$  on a  $16 \times 16$  lattice interpolating between the square ( $\tau = 0$ ) and the triangular ( $\tau = 1$ ) lattice on the cylinder. Figure reproduced from [3].

lattice have the same topological properties as the continuum Laughlin states at filling fraction  $\nu = 1/q$ . We compute in Fig. 3.12 the topological entanglement entropy of the state  $\psi_q^{q/2}$  on the triangular lattice and show that its value is compatible with the value for the corresponding continuum Laughlin states at  $q = 2, 4$  and  $6$ . Particle-hole symmetric lattice Laughlin states with topological order on the triangular lattice may therefore be described by these wave functions.

### 3.2.4 Edge states from a charge at infinity

We now define wave functions for edge states of the lattice Laughlin states. These states are constructed using a charge operator placed at infinity and their wave functions have the same expression as the lattice Laughlin states, except that the number of particles is different. They have the same correlations as the  $\psi_q^\eta$  states in the bulk but a different density at the edge, as we shall see below.



**Fig. 3.12.:** (b) Scaling of the Renyi entropy  $S_{L_y}^{(2)}$  for the states at half-filling  $\psi_2^1$ ,  $\psi_4^2$  and  $\psi_6^3$  on a triangular lattice on the cylinder. Here  $L_x$  is taken to be 16. The topological entanglement entropy of the continuum Laughlin state at  $q = 2$  (resp.  $q = 4$ ,  $q = 6$ ) is indicated with a blue (resp. red, green) arrow. The values found for the topological entanglement entropy are respectively 0.347, 0.723 and 0.907. Figure reproduced from [3].

Let us investigate the effect of inserting in the correlator defining the wave function a single operator  $W_p(\infty) =: e^{i\frac{p}{\sqrt{q}}\phi(\infty)}$  : of charge  $p/q$ ,  $p$  integer, when the position of the operator is taken to infinity. When the lattice is mapped to the cylinder, this amounts to having a charge at infinity on the axis of the cylinder (Fig. 3.13). The wave function is now given by

$$\psi_q^\eta[p(\infty)](n_1, \dots, n_N) \propto \langle W_p(\infty) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_N}(z_N) \rangle. \quad (3.18)$$

We suppose here that we start with a state at  $\eta = 1$  and the resulting evaluation of the correlator leads to

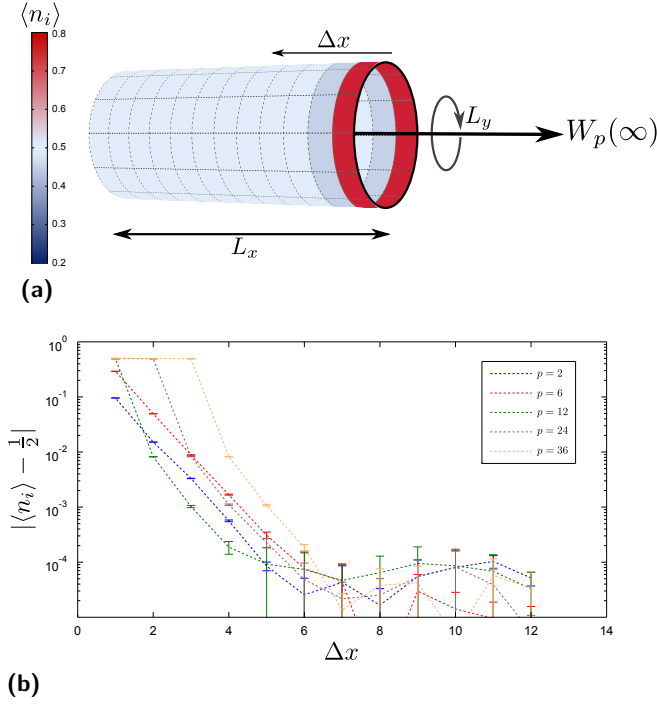
$$\begin{aligned} \psi_q^1[p(\infty)] &\propto \lim_{w \rightarrow \infty} \delta'_n \prod_j (w - z_j)^{pn_j} \prod_{i < j} (z_i - z_j)^{qn_i n_j} \prod_{i \neq j} (z_i - z_j)^{-n_i}, \\ &\propto \lim_{w \rightarrow \infty} \delta'_n w^{p(N-p)/q} \prod_{i < j} (z_i - z_j)^{qn_i n_j} \prod_{i \neq j} (z_i - z_j)^{-n_i}, \\ &\propto \delta'_n \prod_{i < j} (z_i - z_j)^{qn_i n_j} \prod_{i \neq j} (z_i - z_j)^{-n_i}, \end{aligned} \quad (3.19)$$

where  $\delta'_n$  is zero unless the number of particles is

$$M = \sum_i n_i = \frac{N-p}{q}. \quad (3.20)$$

We denote this state  $\psi_q^1[p]_\infty$ . It has exactly the same expression for the wave function as the state  $\psi_q^1$ , except that the number of particles has been modified to  $\frac{N-p}{q}$ . In

the particular case  $p = N(1 - \eta)$ , then the wave functions  $\psi_q^1[p]_\infty$  and  $\psi_q^\eta$  have the same lattice filling factor  $\mu = \eta/q$  and differ only by an  $\eta$  exponent in the last term of the wave function.

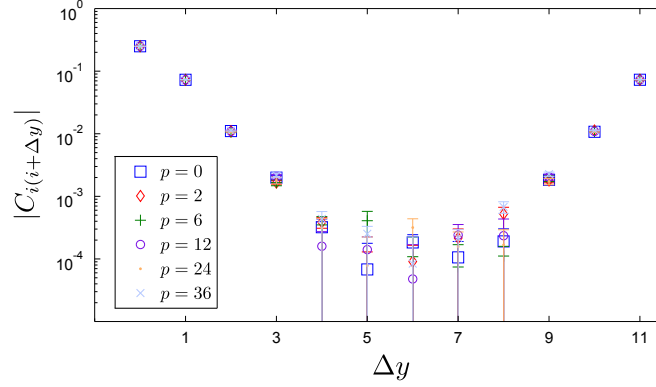


**Fig. 3.13.:** (a) Square  $12 \times 12$  lattice on the cylinder with a charge (here  $p = -6$ ) at infinity. The color shows the density of the state  $\psi_2^1[p]_\infty$ . The state with zero charge at infinity  $\psi_2^1$  has uniform density  $1/2$  but here the density is modified at the edge to account for the change of total particle number. (b) Difference of density between the  $\psi_2^1[p]_\infty$  states and the  $\psi_2^1$  state with respect to the distance to the edge  $\Delta x$ . Values below  $2 \times 10^{-4}$  are not converged. The change of density is exponentially localized at the edge. Figure reproduced from [3].

We show in Fig. 3.13 that these two states differ in their density: the density of the  $\psi_2^1[p]_\infty$  state on the cylinder is the same as the density of the  $\psi_2^1$  state ( $1/2$ ) in the bulk, but is modified on the edge to account for the different total number of particles. This modification of the density is exponentially decaying from the edge towards the bulk of the system. This is in contrast with the density of the  $\psi_2^\eta$  state with the same number of particles, which has a roughly uniform density of  $\eta/q$ .

In addition, we give numerical evidence in Fig. 3.14 that the states  $\psi_2^1[p]_\infty$  and  $\psi_2^1$  have the same particle-particle correlations in the bulk. Therefore the  $\psi_2^1[p]_\infty$  states can be understood as edge states with respect to the  $\psi_2^1$  wave function, since they differ from it only at the edge.

The properties investigated here are similar to the properties of the edge states defined in Ref. [171], but their construction and wave functions are different.

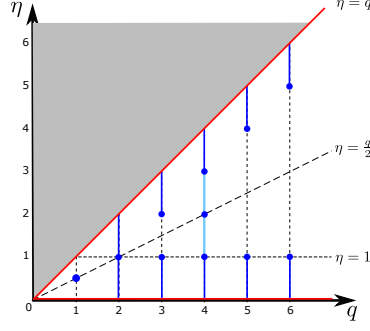


**Fig. 3.14.:** Absolute value of the correlation function  $|C_{i(i+\Delta y)}|$  between two lattice sites in the bulk (middle of the cylinder) separated by a distance  $\Delta y$  in the periodic direction on a square  $12 \times 12$  lattice on the cylinder. The states considered here are the  $\psi_q^1[p]_\infty$  state for different values of  $p$ . When  $p = 0$ , the state is simply the  $\psi_2^1$  state, while it is found that the other states have the same correlations. Figure reproduced from [3].

Note that on the cylinder it is also possible to describe edge states on the other edge by choosing the position of the charge to be 0 instead of infinity. Since the coordinates of the lattice on the cylinder as well as the coefficients of the wave functions are invariant under a transformation  $z \rightarrow 1/z$ , this simply amounts to flipping the two sides of the cylinder and as such will not change the properties of the states.

### 3.2.5 Derivation of parent Hamiltonians

We now derive exact parent Hamiltonians of the lattice Laughlin states. We start by deriving a parent Hamiltonian for which the  $\psi_q^1[p]_\infty$  edge states are ground states with different number of particles. The lattice Laughlin state  $\psi_q^1$  is the unique ground state of this Hamiltonian with lattice filling factor  $1/q$ . We then obtain parent Hamiltonians of the  $\psi_q^\eta$  states when  $\eta \leq 1$  or  $\eta \geq q - 1$ . Compared to the approach used to connect Hamiltonians in the lattice and in the continuum limit in [172], our Hamiltonians do not increase in complexity when we go from the lattice to the continuum limit, since only the coupling strengths need to be changed with  $\eta$ . A parent Hamiltonian is also derived for the  $\psi_4^2$  state. Fig. 3.15 shows the diagram of the states and the values of the parameters for which parent Hamiltonians are obtained. These Hamiltonians involve few-body interactions, but are not local.



**Fig. 3.15.:** Diagram of the  $\psi_q^\eta$  states. The blue lines and dots represent values of the parameters for which exact parent Hamiltonians are derived in this section. On the light blue line the parent Hamiltonians derived have a degenerate ground state on the plane. Figure reproduced from [3].

### Parent Hamiltonian for the edge state $\psi_q^1[p]_\infty$

Let us start by deriving parent Hamiltonians for the edge states  $\psi_q^1[p]_\infty$ . This state is defined by introducing a charge operator in the wave function and taking the position of this operator going to infinity:

$$\psi_q^1[p]_\infty(n_1, \dots, n_N) \propto \langle W_p(\infty) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_N}(z_N) \rangle. \quad (3.21)$$

In [173], an exact parent Hamiltonian was derived for states written in this form when  $p > 0$ . We follow here the same procedure and extend this result to  $p > -q$ . The starting point is to define a null field  $\chi(w)$  [160, 173] such that the correlator with this field inserted vanishes:

$$\langle W_p(w) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_{i-1}}(z_{i-1}) \chi(z_i) \mathcal{V}_{n_{i+1}}(z_{i+1}) \dots \mathcal{V}_{n_N}(z_N) \rangle = 0. \quad (3.22)$$

This is done in details in Appendix A.1, where it is shown that this equation can be rewritten, when  $p > -q$ , as

$$\Lambda_i |\psi_q^1[p(w)]\rangle = 0, \quad (3.23)$$

where

$$\Lambda_i = \sum_{j(\neq i)} \frac{1}{z_i - z_j} [d_j - d_i(qn_j - 1)] - \frac{p}{z_i - w} d_i. \quad (3.24)$$

Here  $d_j$  is the hardcore boson (resp. fermion) annihilation operator for  $q$  even (resp. odd), and  $n_j = d_j^\dagger d_j$  is the number of particles at site  $j$ . Taking the limit  $w \rightarrow \infty$ , so that the charge is at infinity, leads to

$$\Lambda_i |\psi_q^1[p]_\infty\rangle = 0, \quad (3.25)$$



where

$$\Lambda_i = \sum_{j(\neq i)} \frac{1}{z_i - z_j} [d_j - d_i(qn_j - 1)]. \quad (3.26)$$

This leads to  $H|\psi_q^1[p]_\infty\rangle = 0$ , where

$$H = \sum_i \Lambda_i^\dagger \Lambda_i \quad (3.27)$$

is a positive semi-definite operator annihilating the wave function  $\psi_q^1[p]_\infty$ . Here we note that there is no dependence on  $p$  in this Hamiltonian. This means that the Hamiltonian  $H$  has a degenerate ground space containing all the  $\psi_q^1[p]_\infty$  edge states for  $p > -q$ . These states have however different total number of particles, and in physical settings where the number of particles is fixed only one of these states would be ground state of the corresponding Hamiltonian. A particular ground state of this Hamiltonian is the state at  $p = 0$ , which is the  $\psi_q^1$  Laughlin state, corresponding to the ground state at filling factor  $\mu = 1/q$ . Moreover we have checked numerically for small lattice sizes that  $H$  has only one ground state for each subspace with fixed number of particles.

### Parent Hamiltonians for $\psi_q^\eta$ states

Let us now turn to the derivation of parent Hamiltonians for the  $\psi_q^\eta$  states. The first observation is that the  $\psi_q^\eta$  states differ from the  $\psi_q^1[p]_\infty$  states with  $p = N(1 - \eta)$  only by a factor  $\prod_{i \neq j}^N (z_i - z_j)^{(\eta-1)n_i}$ . We will use our previous result and account for this factor at the level of the Hamiltonian. Let us denote  $T$  the operator

$$\begin{aligned} T &= \prod_i \left( \prod_{j(\neq i)}^N (z_i - z_j)^{(\eta-1)} \right)^{n_i}, \\ &= \prod_i \beta_i^{n_i}, \end{aligned} \quad (3.28)$$

where we have defined  $\beta_i = \prod_{j(\neq i)}^N (z_i - z_j)^{(\eta-1)}$ . Then

$$T|\psi_q^\eta\rangle = |\psi_q^1[p]_\infty\rangle, \quad (3.29)$$

and using (3.25),

$$T^{-1} \Lambda_i T |\psi_q^\eta\rangle = 0. \quad (3.30)$$

Notice first that

$$T^{-1} d_j T = \beta_j d_j. \quad (3.31)$$

This means that we can rewrite  $T^{-1}\Lambda_i T$  as

$$\begin{aligned}\Lambda'_i &\equiv T^{-1}\Lambda_i T, \\ &= \sum_{j(\neq i)} \frac{1}{z_i - z_j} [\beta_j d_j - \beta_i d_i (q n_j - 1)].\end{aligned}\quad (3.32)$$

Finally we get that the state  $\psi_q^\eta$  is a ground state of the Hamiltonian

$$H_q^\eta = \sum_i \Lambda_i'^\dagger \Lambda'_i, \quad (3.33)$$

and exact diagonalization on small systems shows that it is its only ground state of total number of particle  $M = \eta \frac{N}{q}$ . Note that in this derivation we assumed  $p > -q$  to be able to use the Hamiltonian for the edge state  $\psi_q^1[p]_\infty$ . This implies that this derivation is valid whenever  $N(1 - \eta) > -q$ , in other words  $\eta < 1 + \frac{q}{N}$ . In the thermodynamic limit, the Hamiltonian is therefore valid for  $\eta \leq 1$  (see Fig. 3.15).

### Parent Hamiltonians for the states at $q = 4$

The parent Hamiltonians derived in the previous section are valid for  $\eta < 1 + \frac{q}{N}$ , and are therefore not valid for all the values of  $\eta$ , unless  $q = 2$ . In this section we derive a Hamiltonian valid for all values of  $\eta$  at  $q = 4$ . We first concentrate on the half-filled case  $q = 4$ ,  $\eta = 2$ .

Using a similar procedure as in the previous case we derive in Appendix A.2 an operator  $\Omega_i$  annihilating the wave function  $\psi_4^2$ :

$$\Omega_i = \sum_{j(\neq i)} \sum_{k(\neq j)} \frac{1}{z_k - z_j} n_i d_j s_k, \quad (3.34)$$

where  $d_j$  is the bosonic annihilation operator at site  $j$ ,  $n_i = d_i^\dagger d_i$  and  $s_k = 2n_k - 1$  is the corresponding spin-1/2 degree of freedom at site  $k$ . In the rest of this section we will use both the notations  $n_k$  and  $s_k$  for brevity.

In addition, the wave-function is particle-hole symmetric and is also annihilated by the particle-hole transformed operator

$$\bar{\Omega}_i = - \sum_{j(\neq i)} \sum_{k(\neq j)} \frac{1}{z_k - z_j} (1 - n_i) d_j^\dagger s_k, \quad (3.35)$$

where we have replaced all  $d_l$  by  $d_l^\dagger$ ,  $d_l^\dagger$  by  $d_l$  and  $n_l$  by  $1 - n_l$  (or  $s_l$  by  $-s_l$ ).

Finally we define a Hamiltonian

$$H_4^2 = \sum_i \left[ \Omega_i^\dagger \Omega_i + \bar{\Omega}_i^\dagger \bar{\Omega}_i \right], \quad (3.36)$$

which can be expanded and simplified, up to a global irrelevant real factor, as

$$H_4^2 = \sum_{j,l} \sum_{\substack{m(\neq l) \\ k(\neq j)}} \frac{1}{(z_k^* - z_j^*)(z_m - z_l)} s_k \left( d_j^\dagger d_l + d_j d_l^\dagger \right) s_m, \quad (3.37)$$

which is a 4-body Hamiltonian for the state  $\psi_4^2$ . Note that since the state  $\psi_4^2$  is equivalent to the  $SO(2)$  spin state defined in [124], this derivation also provides a parent Hamiltonian for this case which was left open.

Replacing the operators  $d_j$  in (3.34) by  $\beta_j d_j$  as in 3.2.5 now leads to operators annihilating the  $\psi_4^\eta$  wave function for any value of  $\eta$ , and thus to parent Hamiltonians for these states. We observe however numerically that this construction does not always give a single ground state of the parent Hamiltonian when  $\eta \neq 2$  (see Fig. 3.15).

### Conformal transformations of the parent Hamiltonians

The wave functions defined from correlators of conformal fields are invariant under general Möbius transformations of the lattice coordinates:

$$M : z \rightarrow \frac{az + b}{cz + d}, \quad ad - bc = 1. \quad (3.38)$$

This is not the case of the Hamiltonians derived previously. In fact there is a class of Hamiltonians annihilating the lattice Laughlin states. These Hamiltonians are related by conformal transformations and some of them have the same symmetries as the lattices considered in this work, but there is no non-zero Hamiltonian that is invariant under all conformal transformations.

The first observation is that the operators annihilating the wave functions have the form  $\Gamma_i = \sum_{j(\neq i)} \lambda_{ij} f_{ij}$ , where  $f_{ij}$  does not depend on the coordinates  $z_i$  and  $\lambda_{ij} = 1/(z_i - z_j)$ . As explained in more details in Appendix A.3, the space of operators annihilating the wave functions that can be obtained by applying Möbius transformations and multiplying by constant factors the operator  $\Gamma_i$  is the space of operators

$$\Gamma_i^{\alpha\beta\gamma} = \sum_{j(\neq i)} (\alpha\lambda_{ij} + \beta\kappa_{ij} + \gamma\rho_{ij}) f_{ij}, \quad (3.39)$$

where  $\alpha$ ,  $\beta$  and  $\gamma$  are complex numbers and  $\kappa_{ij} = \frac{z_i z_j}{z_i - z_j}$ ,  $\rho_{ij} = \frac{z_i + z_j}{z_i - z_j}$ . In the class of Hamiltonians obtained from  $\Gamma_i^{\alpha\beta\gamma}$ , one can check that there is no non-zero Hamiltonian invariant under all conformal transformations, while the Hamiltonians of the form  $\sum_i \sum_{k(\neq i)} \sum_{j(\neq i)} \rho_{kj}^* \rho_{ij} f_{kj}^\dagger f_{ij}$  are invariant under the symmetries of the lattice on the cylinder (Fig. 3.6). This result can be applied to all the parent Hamiltonians derived in this chapter to make them invariant under the symmetry transformations of the cylinder.

### Local Hamiltonian for the $\psi_2^1$ Laughlin state

At the particular values  $q = 2$ ,  $\eta = 1$ , the state has an additional  $SU(2)$  symmetry. The parent Hamiltonian previously obtained for this state is not  $SU(2)$  invariant, but it is possible to construct another parent Hamiltonian invariant under  $SU(2)$  transformations for this state. This was performed in [70] to obtain the Hamiltonian

$$H_{\text{parent}} = \frac{2}{3} \sum_{i \neq j} |w_{ij}|^2 \mathbf{S}_i \cdot \mathbf{S}_j + \frac{2}{3} \sum_{i \neq j \neq k} \bar{w}_{ij} w_{ik} \mathbf{S}_j \cdot \mathbf{S}_k - \frac{2i}{3} \sum_{i \neq j \neq k} \bar{w}_{ij} w_{ik} \mathbf{S}_i \cdot (\mathbf{S}_j \times \mathbf{S}_k), \quad (3.40)$$

where  $w_{ij} = \frac{z_i + z_j}{z_i - z_j}$  and  $\mathbf{S}_j = (S_j^x, S_j^y, S_j^z)$  is the spin-1/2 operator at site  $j$ .

It was moreover shown in [54] that a local Hamiltonian approximately stabilizing the  $\psi_2^1$  state can be obtained on the square lattice by restricting  $H_{\text{parent}}$  to local terms and setting the long-range interactions to zero. This leads to the Hamiltonian

$$H_l = J \sum_{\langle i, j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j + J_\chi \sum_{\langle i, j, k \rangle_\circ} \mathbf{S}_i \cdot (\mathbf{S}_j \times \mathbf{S}_k), \quad (3.41)$$

where  $\langle i, j \rangle$  indicates indices of nearest neighbours on the lattice and  $\langle i, j, k \rangle_\circ$  indicates indices of all triangles of neighboring spins, with vertices labelled in the counter clockwise direction. When  $J = J_\chi$ , the ground state of  $H_l$  has above 98% overlap with the Laughlin wave function  $\psi_2^1$  on a  $4 \times 4$  lattice, and in a large range of parameters the ground state stays in the same phase[54]. A protocol to implement this Hamiltonian experimentally in an optical lattice has been developed[168]. The ground state of this Hamiltonian will be studied using other methods in chapter 4.

### 3.3 Quasielectrons and quasiholes in lattice fractional quantum Hall states

So far we have considered lattice versions of Laughlin wave functions. In this section we extend this construction to the quasiparticle excitations of these states. In the continuum we have already seen that we can construct the quasihole wave function by inserting a quasihole operator in the correlator, and we will see that this construction generalizes to the lattice. This amounts to inserting a flux tube with positive flux, but if one instead inserts a flux tube with negative flux to get a quasielectron, one gets a singularity in the continuum wave function[63, 141]. A lot of work has therefore been trying to construct suitable quasielectron wave functions[63, 68, 141, 145, 174–182]. The resulting wave functions are however significantly more complicated than the quasihole wave function and computing the properties of quasielectrons is difficult.

Explicitly computing braiding properties of various types of quasiholes and quasielectrons has recently been performed using different methods[173, 183–191], and we will see in this section that it can be done for both quasiholes and quasielectrons on a lattice. One advantage of the lattice construction is that because electrons can only be placed at fixed lattice sites, there is no singularity if one tries to construct a quasielectron. The quasielectron wave function is therefore not more complicated than the quasihole wave function and we can compute its shape and braiding properties. It is even possible to define models where there is an exact symmetry between quasielectrons and quasiholes. We also show that parent Hamiltonians for states containing quasiholes and quasielectrons can be constructed. In these Hamiltonians, the positions of the quasiparticles are parameters that can be tuned for creating or braiding the anyons. We also find that some of the wave functions we introduce have high overlap with states realized in a fractional Chern insulator model. This section is based on and reuses parts of Ref. [5].

#### 3.3.1 Wave functions of quasiholes

Let us first consider the wave function for quasiholes (3.2), which can be defined on the lattice by using the expression of a quasihole operator inserted at particular lattice positions in the CFT correlator[68, 172, 173]. Consider  $Q$  quasiholes at positions  $w_j$ ,  $j \in \{1, 2, \dots, Q\}$ , with charge  $p_j/q$ , where  $p_j$  is a positive integer. The positions  $w_j$  are not restricted to be on the lattice sites. We denote  $P = \sum_i^Q p_i/q$  the total charge of the quasiholes and  $\vec{p} \equiv (p_1, p_2, \dots, p_Q)$ . At the position of each quasihole, we attach a vertex operator

$$W_{p_j}(w_j) =: e^{i\frac{p_j}{\sqrt{q}}\phi(w_j)} : \quad (3.42)$$

and we consider the wave function with coefficients given by

$$\psi_q^\eta[p_1(w_1), \dots, p_Q(w_Q)](n_1, \dots, n_N) \propto \langle W_{p_1}(w_1) \dots W_{p_Q}(w_Q) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_N}(z_N) \rangle. \quad (3.43)$$

This expression evaluates to [173]

$$\begin{aligned} & \delta'_n \xi_\eta \prod_{i,j} (w_i - z_j)^{p_i n_j} \prod_{i < j} (z_i - z_j)^{q n_i n_j} \prod_{i \neq j} (z_i - z_j)^{-\eta n_i} \\ & \times \prod_{i < j} (w_i - w_j)^{p_i p_j / q} \prod_{i,j} (w_i - z_j)^{-p_i / q}, \end{aligned} \quad (3.44)$$

where  $\delta'_n$  is zero unless the number of particles is

$$M = \sum_i n_i = \eta \frac{N}{q} - P. \quad (3.45)$$

The total charge therefore modifies the number of particles, and quasiholes have to be inserted in such a way that  $M$  is an integer. Moreover the number of particles removed by a single quasihole is  $p_i/q$ , which confirms that the charge of a quasihole is fractional and equal to  $p_i/q$ .

### 3.3.2 Braiding statistics of quasiholes

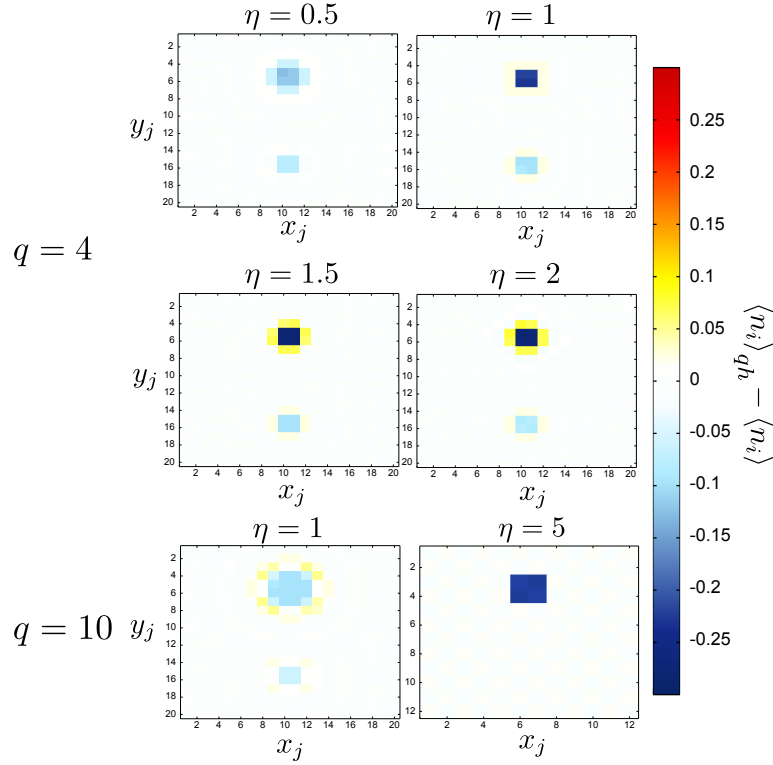
The Berry phase, defined as  $\theta_k = i \oint_c \langle \psi | \frac{\partial \psi}{\partial w_k} \rangle dw_k + c.c.$  [192, 193], that arises when moving the  $k$ th quasihole around a closed curve  $c$  evaluates to

$$\theta_k = i \frac{p_k}{2} \oint_c \sum_i \frac{\langle n_i \rangle_{\vec{p}}}{w_k - z_i} dw_k + c.c. \quad (3.46)$$

The statistics of the quasiholes can be computed by evaluating  $\theta_{k,(w_j \text{ inside})} - \theta_{k,(w_j \text{ outside})}$ , where  $\theta_{k,(w_j \text{ inside})}$  (resp.  $\theta_{k,(w_j \text{ outside})}$ ) is the Berry phase when the  $j$ th quasihole is inside (resp. outside)  $c$  and not close to  $c$ . As observed in [172, 173], when the change of density due to the quasihole is localized around the position of the quasihole, then the Berry phase is zero and therefore the statistics is governed by the monodromy, i.e. the change obtained from analytical continuation of the wave function when the quasiholes move around each other. Indeed in this case we have that  $\langle n_i \rangle_{\vec{p},(w_j \text{ inside})} - \langle n_i \rangle_{\vec{p},(w_j \text{ outside})}$  is zero except close to the position of the  $j$ th quasihole and does not depend on  $w_k$ . Since  $\sum_{(i \text{ inside } c)} (\langle n_i \rangle_{\vec{p},(w_j \text{ inside})} - \langle n_i \rangle_{\vec{p},(w_j \text{ outside})})$  is the total excess charge of the quasihole  $-p_j/q$ , we obtain that

$$\theta_{k,(w_j \text{ inside})} - \theta_{k,(w_j \text{ outside})} = 2\pi p_j p_k / q, \quad (3.47)$$

which is the expected statistics for Laughlin anyons with charges  $p_j/q$  and  $p_k/q$ .



**Fig. 3.16.:** Difference in the density  $\langle n_i \rangle_{qh} - \langle n_i \rangle$  between the state  $\psi_q^\eta$  with one quasihole of charge  $1/q$  and one quasihole of charge  $(q-1)/q$  and the state  $\psi_q^\eta$  without quasiholes on a square  $20 \times 20$  (or  $12 \times 12$  for  $q = 10, \eta = 5$ ) lattice on the cylinder. The coordinates  $w_j$  of the quasiholes are placed in the center between 4 lattice sites and are visible in blue as a lack of density on the neighbouring sites. At  $q = 4$ , it is found for all values of  $\eta$  that the quasiholes are localized. For  $q = 10$  however, the quasiholes are localized when  $\eta = 1$ , but at half-filling ( $\eta = 5$ ) we observe that there is no splitting of the charge between a quasihole of charge  $1/q$  and a quasihole of charge  $(q-1)/q$  and thus no screening of the quasiholes. Figure reproduced from [3].

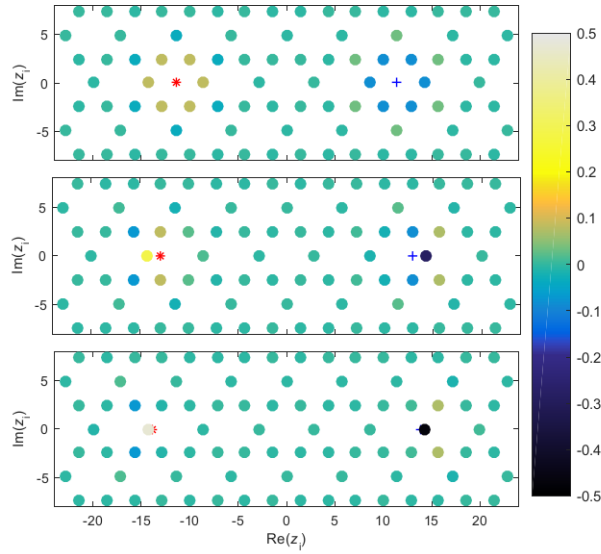
We give numerical evidence in Fig. 3.16 that the change of density due to a quasihole is indeed localized around the position of the quasihole on the square lattice, except for the values of  $q$  and  $\eta$  where the state is not topological. This is an additional indication that for large  $q$  and at half-filling of the lattice the states do not display the same topological properties as the Laughlin states, while for other values of  $q$  and  $\eta$  the states are topological and display the properties of the Laughlin states.

### 3.3.3 Wave function, charge and density profile of the quasielectron

So far we have considered only quasiholes. On the lattice, we define a wave function for quasielectrons by simply setting some of the  $p_i$  in (3.44) to be negative integers.

While this would lead to a singularity in the continuum, the wave function is still well defined on the lattice.

We numerically investigate the change of particle density in the presence of both quasiholes and quasielectrons in Figs. 3.17 and 3.18, showing that for  $q = 3$  on the Kagome lattice the anyons only alter the particle density in a small region close to their position. This confirms that the statistics of quasielectrons is the same as the statistics of quasiholes.



**Fig. 3.17.:** Modification of the particle density due to the presence of anyons in the lattice Laughlin state at  $q = 3$  and half lattice filling ( $\eta = q/2$ ). The lattice is chosen to be a kagome lattice defined on a disc with radius 27.9. A quasielectron (resp. quasihole) with charge  $-1/3$  (resp.  $+1/3$ ) is placed at the position  $*$  ( $+$ ), and the color of the  $j$ th lattice site shows  $\langle n_j \rangle_{[-1,+1]} - \langle n_j \rangle_{[0,0]}$ . Figure reproduced from [5].

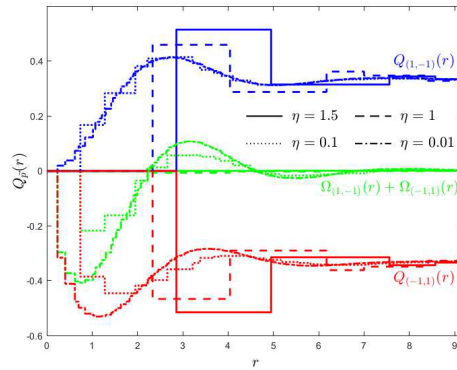
Let us now take  $\eta = q/2$  (i.e., lattice filling  $\sum_j n_j/N = 1/2 - \sum_i p_i/(Nq)$ ), which corresponds to the particle-hole symmetric case investigated in section 3.2.2. In this case, the wave function with quasiparticles is invariant under the transformation  $n_i \rightarrow 1 - n_i, p_i \rightarrow -p_i$ , which exchanges quasiholes and quasielectrons. The density profile, defined as the difference of density in the presence and absence of an anyon, of a quasihole of charge  $p_i/q$  is then the opposite of the density profile of a quasielectron with charge  $-p_i/q$ . This quantity is presented in Fig. 3.17, which shows how the shape of the density profile varies with the positions of the anyons. When a quasihole or quasielectron approaches a lattice site, no singularity occur and the probability that the site is empty or occupied approaches one.



Away from the symmetric point  $\eta = q/2$ , we observe that the excess charges for quasiholes and quasielectrons, defined as

$$Q_{\bar{p}}(r) = - \sum_{\{i \in \{1, 2, \dots, N\} \mid |z_i - w_1| \leq r\}} (\langle n_i \rangle_{\bar{p}} - \langle n_i \rangle_{\bar{0}}), \quad (3.48)$$

are still close to being opposite for a large range of values of  $\eta$  (Fig. 3.18). When  $\eta$  is small we approach the continuum limit. In this case the excess charge of the quasihole converges to a fixed function, while the charge distribution of the quasielectron becomes more and more narrow, corresponding to a singularity in the continuum limit. We also observe that the radius of the quasihole is only a bit larger in the lattice than in the continuum, which is in agreement with the results in [188].



**Fig. 3.18.:** Excess charge 3.48 of a quasihole/quasielectron (blue/red) and their sum (green) for  $q = 3$  for different densities of the lattice sites. The quasihole/quasielectron is placed at the origin at the center of a hexagon in a kagome lattice. The lattice is defined on a disc with radius 27.9 for  $\eta = 1.5, 1, 0.1$  and 18.2 for  $\eta = 0.01$ , and an anyon of the opposite charge is placed at infinity. Figure reproduced from [5].

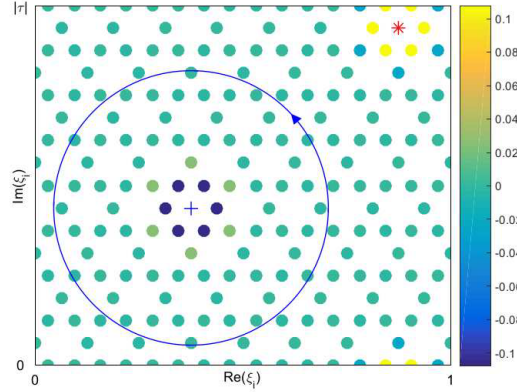
### 3.3.4 Quasiparticles on the torus

In the continuum, it is particularly difficult to construct quasielectrons wave functions, a task which has only been solved recently[182]. On the lattice however, such a construction can be done in a simple manner. Let us consider a torus, defined on the complex plane by identifying all points separated by  $nr_1 + mr_2$  with  $n, m \in \mathbb{Z}$ , where  $r_1$  is real and  $r_2$  is a complex number such that  $\text{Im}(r_2) > 0$ . Let us define  $\tau = r_2/r_1$ ,  $\xi_j = z_j/r_1$ , and  $\zeta_k = w_k/r_1$ . Lattice Laughlin states on the torus have

been defined in [194], and we can generalize this construction to include quasiholes and quasielectrons, leading to a wave function

$$\begin{aligned} \psi_q^\eta[p_1(w_1), \dots, p_Q(w_Q)](n_1, \dots, n_N) &\propto \delta_n \prod_i \chi_{n_i} \prod_{i,j} \theta \left[ \begin{matrix} 1/2 \\ 1/2 \end{matrix} \right] (\zeta_i - \zeta_j, \tau)^{p_i n_j} \\ &\times \theta \left[ \begin{matrix} a + l/q \\ b \end{matrix} \right] \left( \sum_{i=1}^N \xi_i (q n_i - \eta) + \sum_{j=1}^K \zeta_j p_j, q\tau \right) \\ &\times \prod_{i < j} \theta \left[ \begin{matrix} 1/2 \\ 1/2 \end{matrix} \right] (\xi_i - \xi_j, \tau)^{q n_i n_j - n_i \eta - n_j \eta}, \end{aligned} \quad (3.49)$$

where  $\theta \left[ \begin{matrix} a \\ b \end{matrix} \right] (\xi, \tau) \equiv \sum_{n \in \mathbb{Z}} e^{i\pi\tau(n+a)^2 + 2\pi i(n+a)(\xi+b)}$  and  $l \in \{0, 1, \dots, q-1\}$ . For  $q$  even  $(a, b) = (0, 0)$ , and for  $q$  odd  $(a, b)$  can be either  $(0, 0)$ ,  $(0, 1/2)$ ,  $(1/2, 0)$ , or  $(1/2, 1/2)$ . For  $\eta = q/2$  and  $l = 0$  or  $l = q/2$ , the density profiles for quasielectrons and quasiholes are still opposite of each other. In Fig. 3.19 we show numerically for  $q = 2$  that the anyons are localized and obey the expected statistics.



**Fig. 3.19.:** Quasihole (+) and quasielectron (\*) on the torus in the state  $\psi_{\frac{1}{2}}^1[1, -1]$ . The color of the lattice sites shows  $\langle n_j \rangle_{[1, -1]} - \langle n_j \rangle_{[0, 0]}$ . When moving the quasihole around the blue curve we find numerically, using Monte Carlo simulations, that the difference in Berry phase when the quasielectron is at + and at \*, respectively, is  $\phi = -3.145$  with a statistical error of order 0.003. This is in agreement with the expected result  $-\pi$ . Figure reproduced from [5].

### 3.3.5 Exact parent Hamiltonians for states with quasiparticles

So far we have considered wave functions for quasiparticles and one may ask if there are Hamiltonians for which these states are ground states. In section 3.2.5, we showed that the operator

$$\Lambda'_i = \sum_{j(\neq i)} \frac{1}{z_i - z_j} [\beta_j d_j - \beta_i d_i (q n_j - 1)] \quad (3.50)$$

annihilates the lattice Laughlin state without quasiparticles. Using a procedure similar as in section 3.2.5 we can obtain an operator annihilating the state with anyons simply by redefining  $\beta_k = e^{i\phi_k} e^{-i\pi(k-1)} \prod_i (w_i - z_k)^{p_i} \prod_j (z_j - z_k)^{1-\eta}$ , provided the inequality  $\eta - \sum_i p_i/N < 1 + q/N$  is fulfilled. A 3-body Hamiltonian for these states is then given by  $\sum_i \Lambda_i^\dagger \Lambda_i$ , and we have checked numerically that this Hamiltonian has a single ground state on small systems. Note that this only changes the strengths of the terms in the Hamiltonian compared to the original Hamiltonian for Laughlin states without quasiparticles.

This construction is very convenient for manipulating anyons, since the state with anyons are ground states, and not excited states, of the model. By simply changing the strengths of the interactions in the Hamiltonian one can create a pair of quasi-electron and quasihole and separate them, as well as perform braiding of anyons or fuse them by bringing them to the same point.

### 3.3.6 Local Hamiltonians

The parent Hamiltonians we have introduced previously are non-local, but in some cases it has been possible to find also a local Hamiltonian for which the ground state is approximately the analytical states defined as correlators from CFT [2, 38, 54, 195]. It was already shown in [163, 196] that a bosonic fractional quantum Hall state with  $q = 2$  can be realized on a square lattice using a Bose-Hubbard model with complex hopping terms and interactions for low lattice fillings. This model is defined on an  $N_x \times N_y$  lattice on the torus. We define the coordinates  $n \in \{0, 1, \dots, N_x - 1\}$  and  $m \in \{0, 1, \dots, N_y - 1\}$  and take  $z_{n+mN_x+1} = (n + mi)\sqrt{2\pi\eta}$ , such that the area of one lattice site is  $2\pi\eta$  (Fig. 3.20). In the Landau gauge, the Hamiltonian takes the form

$$H = -J \sum_{n,m} \left( \hat{d}_{n+1,m}^\dagger \hat{d}_{n,m} e^{-2\pi i \eta m} + \hat{d}_{n,m+1}^\dagger \hat{d}_{n,m} + h.c. \right) + U \sum_{n,m} \hat{n}_{n,m} (\hat{n}_{n,m} - 1), \quad (3.51)$$

where  $\hat{d}$  is the boson annihilation operator,  $\hat{n}$  is the number operator,  $J$  is the strength of the hopping term, and  $U$  is the strength of the interaction term. In the following, we take the limit of hardcore interactions  $U \rightarrow \infty$ . The amount of magnetic flux going through one plaquette is  $\eta$ , and in the absence of quasiparticles we fix the number of particles in the system to  $\eta N/2$ , so that the number of particles per unit flux is  $1/2$ .

For a square lattice, the factor  $\prod_{i<j} \theta \left[ \frac{1/2}{1/2} \right] (\xi_i - \xi_j, \tau)^{-n_i \eta - n_j \eta}$  in (3.49) is proportional to a Gaussian up to single particle phase factors [197], and in this section we choose the phase factors  $\chi_{n_i}$  such that

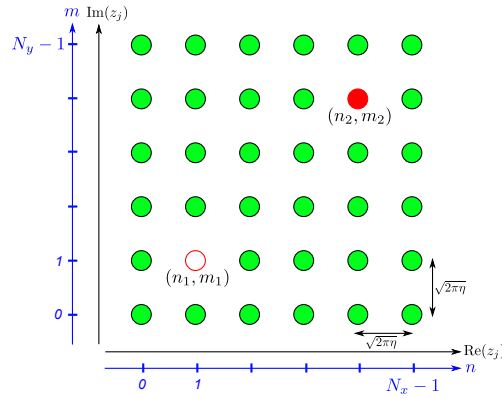
$$\prod_i \chi_{n_i} \prod_{i<j} \theta \left[ \frac{1/2}{1/2} \right] (\xi_i - \xi_j, \tau)^{-n_i \eta - n_j \eta} \propto \exp \left( -\frac{1}{2} \sum_i n_i \text{Im}(z_i)^2 \right). \quad (3.52)$$

We take  $N_x = N_y = 6$  and consider the case of 3 particles on the lattice, so that  $\eta = 1/6$ . This ensures magneto-periodic boundary conditions of the Hamiltonian on the torus. The Hamiltonian  $H$  has two ground states and each of them has an overlap of 0.991 with a combination of the two Laughlin states on the torus without quasiparticles.

We now add a potential to localize a quasihole and a quasielectron excitation by giving an energy penalty to one lattice site to be occupied and an energy penalty to another lattice site to be empty:

$$V = Q (\hat{n}_{n_1, m_1} - \hat{n}_{n_2, m_2}). \quad (3.53)$$

The lattice sites with coordinates  $(n_1, m_1)$  and  $(n_2, m_2)$  are taken to be as far as possible from each other on the torus (Fig. 3.20). Since we are trying to create both a quasihole and a quasielectron, the total number of particles is unchanged.



**Fig. 3.20.:** Lattice coordinates on the torus.  $(n_1, m_1)$  indicates the position of a quasihole, and  $(n_2, m_2)$  indicates the position of a quasielectron. Figure reproduced from [5].

$H + V$  has two almost degenerate ground states. We compare these two states with combinations of the two analytical states on the torus with a quasihole at position  $(n_1, m_1)$  and a quasielectron at position  $(n_2, m_2)$  and find that they have overlap 0.994 and 0.989, respectively, for large  $Q$ . This shows that the analytical states considered above are relevant for fractional Chern insulator models.

## 3.4 Moore-Read states on lattices

So far we have only been concerned with Laughlin wave functions, which can be used when the filling fraction  $\nu$  is  $1/q$ . As we have seen in section 3.1, there is another FQH state which has attracted a lot of attention: the leading candidate to describe the  $\nu = 5/2$  FQH state is a Moore-Read state at filling  $1/2$  [68, 150, 151], describing the wave function of the electrons in the second Landau level and supporting non-Abelian excitations.

Non-Abelian FQH states were also found in lattice models with topological flat bands [198–201], and Hamiltonians were derived for non-Abelian chiral spin liquids with excitations with  $SU(2)_k$  statistics [53, 159]. For  $k = 2$  this corresponds to a bosonic lattice Moore-Read state at filling fraction 1 on a spin-1 lattice [53]. Moore-Read states of bosons have also been considered on one dimensional lattices, where parent Hamiltonians have been obtained [202].

In this section we extend the construction of lattice wave functions from CFT correlators introduced in the previous sections to non-Abelian FQH states. We construct a family of lattice versions of the Moore-Read state at filling fraction  $1/q$  on arbitrary spin-1/2 and spin-1 lattices in one or two dimensions. This family of states allows us to interpolate between the continuum limit, where all states become continuum Moore-Read wave functions, and a lattice limit. We investigate the properties of the states and provide numerical evidence that they are critical states in one dimension and that they are topologically ordered states in two dimensions. Parent Hamiltonians for which the lattice Moore-Read states are ground states are derived on any lattice of arbitrary size. These parent Hamiltonians have long-range interactions, so we provide a method to deform them to local Hamiltonians and show numerical evidence that the state at filling fraction 1 on a spin-1 square lattice can be stabilized by a local Hamiltonian in one and two dimensions, while the state at filling fraction  $1/2$  on a spin-1/2 square lattice can be stabilized by a local Hamiltonian in one dimension. This section is based on Ref. [2] and is a modification thereof.

### 3.4.1 Definition of lattice Moore-Read states

The lattices we now consider are chosen such that the average area per site satisfies  $\eta = \frac{a}{2\pi}$ . We will study both models with local degree of freedom  $n_i \in \{0, 1\}$  (case denoted as  $S = 1/2$ , which can represent a spin-1/2, a number of hardcore bosons or a number of fermions at each lattice site) and with degree of freedom  $n_i \in \{0, 1, 2\}$  (case denoted as  $S = 1$ , which corresponds to a spin-1 model, a bosonic model with at most two bosons per lattice site or a fermionic model with up to two fermions of different type per lattice site).

We consider an infinite dimensional MPS as in (3.7), but now with the different field operators:

$$V_{n_j}(z_j) = \chi(z_j)^{\bar{\delta}_{n_j}} : e^{i(qn_j - \eta)\varphi_R(z_j)/\sqrt{q}} :, \quad (3.54)$$

where  $\varphi_R(z)$  is the chiral bosonic field introduced in (2.21),  $\chi$  is a Majorana fermion field and  $\bar{\delta}_{n_j}$  is 1 if  $n_j = 1$  and 0 otherwise. As in the rest of this chapter,  $q$  is an integer and  $\eta$  is a real number. We also define a phase coefficient

$$U_{n_j}(z_j) = \xi_j^{n_j} e^{i\pi(j-1)\eta n_j}, \quad (3.55)$$

where the  $\xi_j$  are phase factors to be specified.

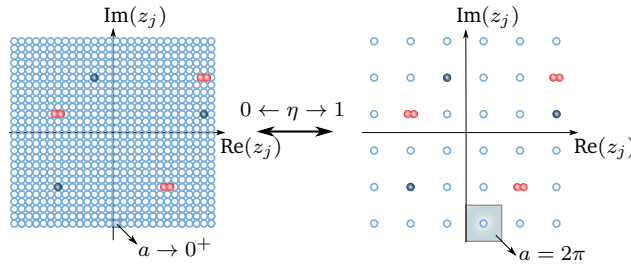
The infinite dimensional MPS defined with the previous operators is

$$\psi_{q,S}^\eta(n_1, \dots, n_N) \propto \langle \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_N}(z_N) \rangle, \quad (3.56)$$

where

$$\mathcal{V}_{n_j}(z_j) = U_{n_j}(z_j)V_{n_j}(z_j). \quad (3.57)$$

Note that depending on whether we consider a model with spin-1/2 ( $S=1/2$ ) or spin-1 ( $S=1$ ), this expression defines two different wave functions, which we will refer to as the  $(q, \eta)_S$  lattice Moore-Read states.



**Fig. 3.21.:** Illustration of a square lattice on the complex plane in the continuum limit ( $\eta \rightarrow 0, N \rightarrow \infty$ ) and in the lattice limit ( $\eta \rightarrow 1$ ). At each site there can be 0 (blue circle), 1 (blue disk) or 2 (red disks) particles. The interpolation is performed by fixing the number of particles  $M = \eta \frac{N}{q}$  and by varying  $\eta = \frac{a}{2\pi}$  between 0 and 1, which changes the number of lattice sites per particle between infinity and  $q$ . Figure reproduced from [2].

Let us now evaluate the correlator. The correlator is zero unless  $\sum_{i=1}^N n_i = \eta N/q$ . This condition fixes the total number of particles in the system to  $M \equiv \sum_{i=1}^N n_i = \eta N/q$ . In this section we will fix the value of  $M$ , so that  $\eta$  interpolates between the continuum limit ( $\eta \rightarrow 0, N \rightarrow \infty$ ) with infinitely many lattice sites per particle and the lattice limit ( $\eta = 1$ ) at which the lattice filling fraction is equal to  $1/q$ , which corresponds to the Landau level filling fraction in the FQH effect (Fig. 3.21). Let us

write  $\delta_n = 1$  if  $\sum_{i=1}^N n_i = M$  and  $\delta_n = 0$  otherwise. The evaluation of the correlator yields [32]

$$\psi_{q,S}^\eta(n_1, \dots, n_N) \propto \delta_n \prod_{i < j} (z_i - z_j)^{qn_i n_j} \text{Pf}_{n_i=n_j=1} \left[ \frac{1}{z_i - z_j} \right] \prod_l f_N(z_l)^{n_l}, \quad (3.58)$$

where  $f_N(z_l) \equiv \xi_l \prod_{j(\neq l)} (z_l - z_j)^{-\eta}$  and the Pfaffian is evaluated at the coordinates where  $n_i = 1$ . The Pfaffian is antisymmetric, so these are bosonic states when  $q$  is odd and fermionic states when  $q$  is even. The states  $(q, \eta)_{1/2}$  are projections of the states  $(q, \eta)_1$  onto the Hilbert space allowing only for single occupancy at each site, while states with  $S = 1$  and  $q$  odd (resp. even) can have sites with two bosons (resp. fermions of different types). The state  $(q = 1, \eta = 1)_{1/2}$  is trivial since the number of particles is fixed to  $N$  and there are  $N$  sites, so in the following we restrict ourselves to states  $(q, \eta)_{1/2}$  with  $q \geq 2$  and  $(q, \eta)_1$  with  $q \geq 1$  (Table 3.1).

Lattice limit $\eta = 1$		
q	S	
1		$\frac{1}{2}$
2		1
⋮		⋮
1		×
2	fermionic	$(q = 2, \eta = 1)_{1/2}$
⋮		⋮
Continuum limit $\eta \rightarrow 0^+$		
q	S	
1		$\frac{1}{2}$
2		1
⋮		⋮
1		×
2	fermionic	Moore-Read
⋮		⋮
1		bosonic Moore-Read
2		fermionic Moore-Read
⋮		⋮

Tab. 3.1.: First  $(q, \eta)_S$  Moore-Read states.

### 3.4.2 The lattice states become Moore-Read states in the continuum limit

In this section we consider a two dimensional lattice defined on a disk  $\mathcal{D}$  of radius  $R \rightarrow \infty$  and show that the lattice Moore-Read states we have introduced reduce to Moore-Read states of particles in the continuum, that is (3.5), when  $\eta \rightarrow 0$ ,  $N \rightarrow \infty$  and the number of particles  $M$  is fixed. We restrict ourselves to lattices where the area per site  $a_i$  is constant equal to  $a$ , but the derivation remains true for any lattice if we make  $\eta$  position dependent [160].

Let us first compute  $\prod_l f_N(z_l)^{n_l}$ . Notice that  $|f_N| = \exp(-\sum_{j(\neq l)} \eta \ln(|z_l - z_j|))$  and since  $\eta = \frac{a}{2\pi}$ , in the continuum limit this sum can be replaced by an integral

$\int_{\mathcal{D}} \ln(|z_l - z_j|) dz^2 / 2\pi$ . This integral evaluates, in the thermodynamic limit, to  $|z_l|^2 + \text{constant}$  [68], so that

$$f_{N \rightarrow \infty}(z_l) \propto \xi_l e^{-ig_l} e^{-|z_l|^2/4}, \quad (3.59)$$

where  $g_l = \text{Im} \left( \sum_{j(\neq l)} \eta \ln(z_l - z_j) \right)$  is a real number. It was found numerically in [160] and [69] that this formula was an accurate approximation even for moderately large  $N$ . We thus get that

$$\prod_l f_{N \rightarrow \infty}(z_l)^{n_l} \propto \left( \prod_l \xi_l^{n_l} e^{-in_l g_l} \right) e^{-\frac{1}{4} \sum_l |z_l|^2 n_l}. \quad (3.60)$$

In the rest of this section we set the phase factors such that  $\xi_l = e^{ig_l}$  to get rid of the overall gauge factor, which does however not change properties like the particle-particle correlation function and entanglement entropy of the state.

### Continuum limit of the $S = \frac{1}{2}$ states

Let us now write the complete wave function in the continuum limit:

$$\psi_S(n_1, \dots, n_N) \propto \delta_n \prod_{i < j} (z_i - z_j)^{q n_i n_j} \text{Pf}_{n_i = n_j = 1} \left[ \frac{1}{z_i - z_j} \right] e^{-\frac{1}{4} \sum_l |z_l|^2 n_l}, \quad (3.61)$$

where the gauge factor has been set to one. It is not straightforward to take the continuum limit in this basis, since one has to define the limit of the Hilbert space on which the wave functions are defined. However, since the number of particles is conserved, we can rewrite the wave function in the basis spanned by the positions  $Z_1, \dots, Z_M$  of the particles. For  $S = 1/2$  there is at most one particle per site so the wave function can be simply expressed as

$$\psi_{\frac{1}{2}}(Z_1, \dots, Z_M) \propto \prod_{i < j} (Z_i - Z_j)^q \text{Pf} \left[ \frac{1}{Z_i - Z_j} \right] e^{-\frac{1}{4} \sum_l |Z_l|^2}, \quad (3.62)$$

where the  $Z_i$  are restricted to positions in the lattice. In the limit of infinitely many lattice sites per particle the lattice becomes a continuous plane and the positions  $Z_i$  become positions in the plane. This state then coincides with the Moore-Read state (3.5). The number of particles on the lattice is  $M = \eta \frac{N}{q} = \frac{a}{2\pi} \frac{N}{q}$ , so if the flux is  $\frac{\text{Area}}{2\pi}$ , then we can express  $\frac{1}{q} = \frac{M}{\text{flux}}$ . This explains that this quantity corresponds to the filling fraction in the continuum, defined as the number of particles per magnetic flux.



### Continuum limit of the $S = 1, q = 1$ state

For  $S = 1, q = 1$ , the state also has the form (3.61), however since the  $n_i$  can take the value 2, it is not straightforward to take the continuum limit. We first have to write the wave function in the basis spanned by the position of the particles. For a basis element  $|n_1, \dots, n_N\rangle$ , let  $Z_r, r \in \{1, \dots, M\}$  be the positions of the particles. Since we interpret the state  $|2\rangle$  as the presence of two particles, positions  $Z_i$  where  $n_i = 2$  are listed with two different indices  $r$  in the set  $\{Z_r\}$ . We now write the wave function in the basis given by the sets  $\{Z_r\}$ . As a starting point, observe that

$$\prod_{r=1}^M e^{-|Z_r|^2/4} = \prod_{i=1}^n e^{-n_i |z_i|^2/4}. \quad (3.63)$$

We can also prove that[2]:

$$\prod_{r<s} (Z_r - Z_s) \text{Pf} \left[ \frac{1}{Z_r - Z_s} \right] = \prod_{i<j} (z_i - z_j)^{n_i n_j} \text{Pf}_{n_i=n_j=1} \left[ \frac{1}{z_i - z_j} \right]. \quad (3.64)$$

The wave function can therefore be written as

$$\psi_1^{q=1}(Z_1, \dots, Z_M) \propto \prod_{i<j} (Z_i - Z_j) \text{Pf} \left[ \frac{1}{Z_i - Z_j} \right] e^{-\frac{1}{4} \sum_l |Z_l|^2}, \quad (3.65)$$

where the  $Z_r$  can be repeated twice to allow for states with double occupation, in which case this expression does not vanish because of a cancellation between the Jastrow factor and the Pfaffian. If we now take the continuum limit, the positions of the particles can be anywhere on the plane and this becomes the bosonic Moore-Read state at filling fraction 1, which also does not vanish when two particles are at the same site. The contribution of configurations with two particles at one or more sites can be neglected in the continuum limit. Note that the previous derivation also shows that the  $(q = 1, \eta = 1)_1$  Moore-Read state in the thermodynamic limit is equivalent to the spin-1 non-Abelian chiral spin liquid introduced in Ref. [53], but the two states are different on finite lattices.

### Continuum limit of the $S = 1, q \geq 2$ state

When  $q \geq 2$  the state can be written as

$$\psi_1^{q \geq 2}(n_1, \dots, n_N) \propto \prod_{i<j} (z_i - z_j)^{(q-1)n_i n_j} \psi_1^{q=1}(n_1, \dots, n_N) \quad (3.66)$$

We have already derived the continuum limit of the state on the right at  $q = 1$  and the remaining factor can be expressed as

$$\prod_{i < j} (z_i - z_j)^{(q-1)n_i n_j} = \prod_{r < s} g(Z_r, Z_s)^{(q-1)}, \quad (3.67)$$

where

$$g(Z_r, Z_s) = \begin{cases} Z_r - Z_s & \text{if } w_r \neq w_s \\ 1 & \text{otherwise} \end{cases} \quad (3.68)$$

In the continuum limit, the wave function can therefore be written as

$$\psi_1^{q \geq 2}(Z_1, \dots, Z_m) \propto \prod_{r < s} g(Z_r - Z_s)^q \text{Pf} \left[ \frac{1}{Z_r - Z_s} \right] \prod_r e^{-|Z_r|^2/4}. \quad (3.69)$$

When no two particles are at the same site, this wave function is the same as the  $(q, \eta)_{1/2}$  Moore-Read state in the continuum limit. Because configurations with two particles at the same site do not contribute to the wave function in the continuum limit[2], the  $(q, \eta)_1$  Moore-Read states for  $q \geq 2$  have the same continuum limit as the  $(q, \eta)_{1/2}$  Moore-Read states, which is the continuum Moore-Read state at filling fraction  $1/q$ .

### One dimensional continuum limit

So far we have focused on two dimensional states. In the one dimensional setting, when  $z_j = e^{2\pi i j/N}$ , the same results enable us to perform an interpolation between the lattice and the continuum. The only difference with the 2D case is that it is now possible to compute analytically  $f_N(z_l) \propto \xi_l z_l^\eta$ , so the wave functions in the continuum can be expressed as

$$\psi_S(Z_1, \dots, Z_M) \propto \prod_{i < j} (Z_i - Z_j)^q \text{Pf} \left[ \frac{1}{Z_i - Z_j} \right] \prod_j \xi_j, \quad (3.70)$$

which is a one-dimensional version of the Moore-Read state.

### 3.4.3 Properties of the lattice Moore-Read states

In two dimensions, the Moore-Read states in the continuum are topological states which support non-Abelian quasi-particle excitations. It is of high interest to check whether the lattice Moore-Read states we have introduced share these properties. In one dimension, we expect that the lattice Moore-Read states display critical behaviour related to the conformal operators used to construct the wave function.

Here we focus on the states with  $q = 1$  and  $q = 2$ , and numerically compute some of the properties of the states we have introduced.

### One dimensional critical states

We now look at one-dimensional chains such that  $z_j = e^{2\pi i j/N}$ . Since we will find local Hamiltonians for the states  $(q = 1, \eta)_1$  and  $(q = 2, \eta)_{1/2}$  in one dimension in the lattice limit in section 3.4.5, we focus on these states. First we compute the Renyi entropy (2.12) of a subsystem of size  $L$  of the chain, following the algorithm detailed in section 2.2.3.

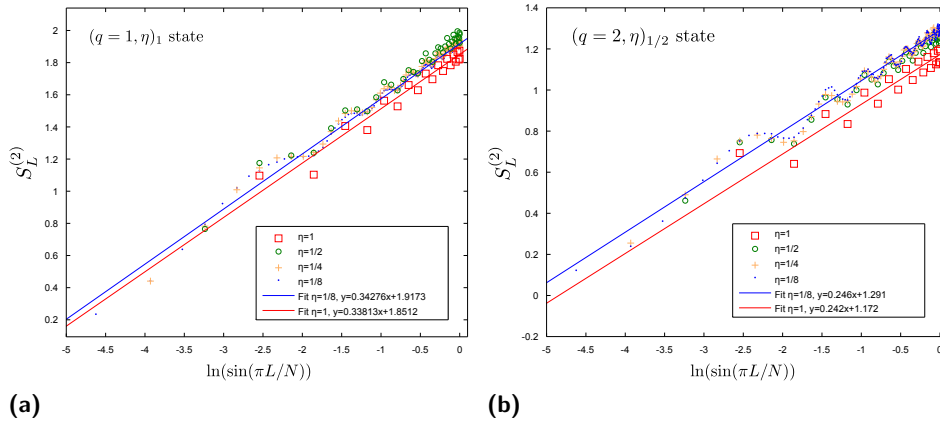
The results are shown in Fig. 3.22. The entropy scales logarithmically with the size of the subchain for all values of  $\eta$ . Moreover, the scaling is approximately the same for different values of  $\eta$  and fits of the form

$$S_L^{(2)} = \frac{c}{4} \ln(\sin(\frac{\pi L}{N})) + b \quad (3.71)$$

yield a value of the central charge  $c$  approximately equal to 1.36 for the  $(q = 1, \eta)_1$  Moore-Read states (the main source of errors is here the finite size of the lattice considered). This value is in agreement with the value of 1.395 found for the state in the lattice limit in [70], where it was also shown that a value of 1.5 for the central charge, as expected for the  $SU(2)_2$  WZW model, could not be excluded. For the  $(q = 2, \eta)_{1/2}$  Moore-Read states we find a value of 0.98 which is compatible with a central charge equal to 1. Another quantity that can be computed using Monte-Carlo techniques is the particle-particle correlation function  $C_L = \langle n_1 n_L \rangle - \langle n_1 \rangle \langle n_L \rangle$ . Results in Fig. 3.23 confirm that the states are critical since the correlation functions decay polynomially with the distance  $L$ . For the  $(q = 1, \eta = 1)_1$  Moore-Read state, the critical exponent is found to be 0.70, which is in agreement with the value of 0.69 found in [70], where it was observed that such a value can be influenced by a multiplicative logarithmic correction which could explain the difference with the expected value of 0.75 [203–205]. Moreover for the  $(q = 2, \eta)_{1/2}$  states at different values of  $\eta$ , the correlations are very close once rescaled by a factor of  $1/\eta^2$ , which confirms that properties of the state do not change along the interpolation. For a Tomonaga-Luttinger liquid, the expected behaviour of the particle-particle correlation function is [206]

$$C_L = \frac{A \cos(2Lk_F)}{\left| \sin(\frac{\pi L}{N}) \frac{N}{\pi} \right|^{2K}} + \frac{K}{2\pi^2 \left| \sin(\frac{\pi L}{N}) \frac{N}{\pi} \right|^2}, \quad (3.72)$$

where  $K$  is the Luttinger parameter,  $k_F = \eta\pi/q$  is the Fermi momentum and  $A$  is a non-universal constant. For the  $(q = 2, \eta)_{1/2}$  state we find a good agreement of this formula for  $K = 0.494$ ,  $A = 0.123$ . This suggests that this state in one dimension



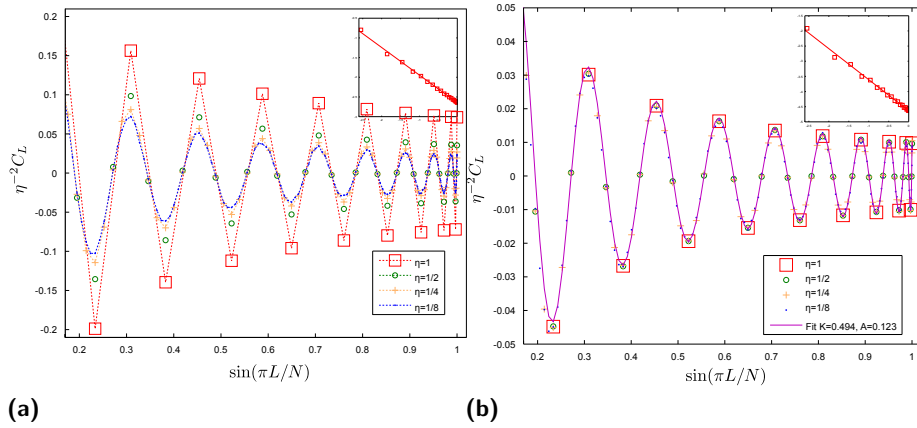
**Fig. 3.22.:** Renyi entropy  $S_L^{(2)}$  of a subsystem of  $L$  consecutive sites for the 1D  $(q = 1, \eta)_1$  Moore-Read states (a) and  $(q = 2, \eta)_{1/2}$  Moore-Read states (b) for different values of  $\eta$ . The number of particles  $M = \eta N/q$  is fixed so the sizes of the chain are  $N = 40, 80, 160, 320$  for  $\eta = 1, 1/2, 1/4, 1/8$  respectively. The lines are linear fits of the points for  $\eta = 1/8$  (blue) and  $\eta = 1$  (red). Figure reproduced from [2].

is well described by a Tomonaga-Luttinger liquid with central charge  $c = 1$  and Luttinger parameter  $K = 0.5$ , which corresponds to the properties of a free-boson CFT with radius  $\sqrt{2}$ , as was the case for the corresponding one-dimensional Laughlin state[160].

## Two dimensional topological states

In the continuum, the Moore-Read state at filling fraction  $1/q$  has a topological entanglement entropy (TEE), as defined in (3.3), of [146]  $\gamma_0(q) = \frac{1}{2} \ln(4q)$ . We compute the TEE as in section 3.2.2, see Fig. 3.6. The results in Fig. 3.24 for the states  $(q = 1, \eta)_1$  (resp.  $(q = 2, \eta)_{1/2}$ ) are in agreement with the topological entanglement entropy of a Moore-Read state at filling fraction 1 (resp.  $1/2$ ),  $\gamma_0(1) = \frac{1}{2} \ln(4) \approx 0.69$  (resp.  $\gamma_0(2) = \frac{1}{2} \ln(8) \approx 1.04$ ). Moreover the value of the TEE does not change with  $\eta$ , so the topological properties of the states remain the same along the interpolation between the continuum and the lattice limit.

We observe however that the state  $(q = 2, \eta)_1$  has a TEE close to zero and different than  $\gamma_0(2)$ . The TEE does not stay constant when  $\eta$  is changed, which is compatible with the expectation that its value is  $\gamma_0(2)$  in the continuum limit: there must be a phase transition along the interpolation between the continuum and the lattice. The states with  $S = 1$ ,  $q \geq 2$  can therefore define distinct lattice states from the states with  $S = 1/2$ ,  $q \geq 2$ , while having the same continuum limit.



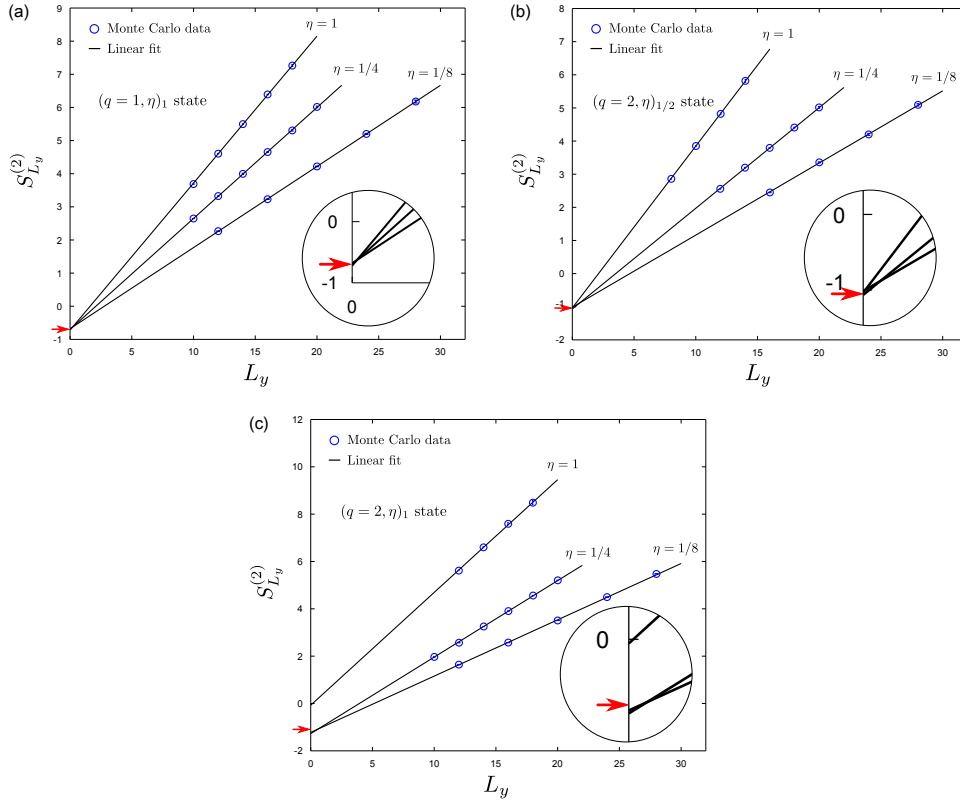
**Fig. 3.23.:** Rescaled correlation function  $\eta^{-2}C_L$  as a function of the distance between the sites for the 1D  $(q = 1, \eta)_1$  Moore-Read states (a) and  $(q = 2, \eta)_{1/2}$  Moore-Read states (b) for different values of  $\eta$ . The number of particles  $M = \eta N/q$  is fixed so the sizes of the chain are  $N = 40, 80, 160, 320$  for  $\eta = 1, 1/2, 1/4, 1/8$  respectively. The data for  $\eta = 1$  is shown in the insets in log-log scale, confirming the polynomial decay of correlations, and the line in the insets is a linear fit yielding critical exponents 0.70 ( $(q = 1, \eta = 1)_1$  state) and 1.02 ( $(q = 2, \eta = 1)_{1/2}$  state). In (b), the line is a fit of the form (3.72) with parameters  $K=0.494$  and  $A=0.123$ . Figure reproduced from [2].

### 3.4.4 Parent Hamiltonians for the lattice Moore-Read states

So far we have considered wave functions for lattice states. It is also relevant to ask whether these states are ground states of some Hamiltonians and whether these Hamiltonians can be realized in nature or implemented in experiments. We now turn to the construction of parent Hamiltonians for which the lattice Moore-Read states are ground states. The procedure is similar as the one introduced in section 3.2.5: null fields are found such that the correlators vanish. These equations are transformed in a set of operators  $\Lambda_i$  annihilating the wave function and a Hamiltonian  $\sum_i \Lambda_i^\dagger \Lambda_i$  is defined for which the lattice states are ground states. The details can be found in Appendix B and in the following we summarize the expressions found for the operators  $\Lambda_i$ . In the rest of this chapter the phase factors are fixed to  $\xi_l = 1$ .

#### Parent Hamiltonians for the $SU(2)_2$ $(q = 1, \eta = 1)_1$ Moore-Read state

The special case of the  $(q = 1, \eta = 1)_1$  state has a wave function constructed from the spin 1 primary fields of the  $SU(2)_2$  Wess-Zumino-Witten conformal field theory [70, 207] and has already been considered partially in Ref. [70], where, however, the focus was on 1D systems. This  $SU(2)_2$  symmetry can be used to construct



**Fig. 3.24.:** Linear behaviour of the Renyi entropy with the size of the cut  $L_y$  for the  $(q = 1, \eta)_1$  (a),  $(q = 2, \eta)_{1/2}$  (b) and  $(q = 2, \eta)_1$  (c) Moore-Read states on a  $L_x \times L_y$  lattice. The topological entanglement entropy of the continuum Moore-Read states at filling 1 (a),  $\gamma_0(1) \approx 0.69$ , and at filling  $1/2$  (b,c),  $\gamma_0(2) \approx 1.04$ , are indicated with a red arrow. The values of  $\eta$  are 1,  $1/4$  and  $1/8$  and the corresponding sizes  $L_x$  are respectively 12, 16 and 16. The insets are enlarged views confirming that the topological entanglement entropy stays the same when  $\eta$  is varied and that its value corresponds to  $\gamma_0(1)$  (resp.  $\gamma_0(1/2)$ ) in the first two cases, while the topological entanglement entropy of the  $(q = 2, \eta)_1$  Moore-Read state is close to zero in the lattice limit and close to  $\gamma_0(1/2)$  in the continuum limit. Figure reproduced from [2].

parent Hamiltonians invariant under  $SU(2)$  transformations. Let us define the spin-1 operators  $t_i^a$ , written in the spin basis at site  $i$  as

$$t_i^x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad t_i^y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad t_i^z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad (3.73)$$

the Levi-Civita symbol  $\epsilon_{abc}$  and rewrite the coordinates as  $\omega_{ij} = \frac{z_i + z_j}{z_i - z_j}$ . Then the results of Appendix B.1 lead to operators annihilating the wave function

$$\Lambda_i^a = \sum_{j(\neq i)}^N \omega_{ij} \left[ \frac{2}{3} t_j^a - \frac{5}{12} i \epsilon_{abc} t_j^b t_i^c - \frac{1}{12} (t_i^a t_i^b + t_i^b t_i^a) t_j^b \right], \quad (3.74)$$

from which we can derive a Hamiltonian

$$\begin{aligned} H = & \frac{4}{3} \sum_{i \neq j}^N \omega_{ij}^* \omega_{ij} + \frac{1}{3} \sum_{i \neq j}^N \left( \omega_{ij}^* \omega_{ij} + 2 \sum_{k(\neq i, j)}^N \omega_{ki}^* \omega_{kj} \right) t_i^a t_j^a \\ & - \frac{1}{6} \sum_{i \neq j}^N \omega_{ij}^* \omega_{ij} (t_i^a t_j^a)^2 + \sum_{i \neq j \neq k}^N \left( \frac{1}{3} \omega_{ik}^* \omega_{ij} - \frac{1}{2} \omega_{ik} \omega_{ij}^* \right) t_i^a t_j^a t_k^b t_k^b. \end{aligned} \quad (3.75)$$

This Hamiltonian is  $SU(2)$  invariant and numerical diagonalization on small systems confirm that it has the  $(q = 1, \eta = 1)_1$  Moore-Read state as a unique ground state. This Hamiltonian is similar to the one obtained in [159] for the spin-1 non-Abelian chiral spin liquid state introduced in [53]. However the Hamiltonian we just constructed is valid for any choice of lattice and not only in the thermodynamic limit as is the case in [159].

### Parent Hamiltonians for the $(q, \eta = 1)_1, q \geq 2$ Moore-Read states

For  $q \geq 2$ , the  $(q, \eta = 1)_1$  Moore-Read states do not display an  $SU(2)$  symmetry. It is still possible to find null fields and construct operators annihilating the wave function. Let us define the operators  $d^\dagger, d$  to be creation and annihilation operators (bosonic for  $q$  odd, fermionic for  $q$  even) acting between states  $|0\rangle$  and  $|1\rangle$ , and  $d'^\dagger, d'$  to be creation and annihilation operators acting between states  $|1\rangle$  and  $|2\rangle$ . The number of particles at site  $i$  is thus  $n_i = n_i^{(1)} + 2n_i^{(2)}$ , where  $n_i^{(1)} = d^\dagger d$  and  $n_i^{(2)} = d'^\dagger d'$ . In Appendix B.2 we derive the following operators annihilating the wave function:

$$\Lambda^0 = \sum_i d_i, \quad (3.76)$$

$$\Lambda_i^p = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^p} d_j d_i^\dagger, \quad (3.77)$$

$$\Lambda_i^{q-1} = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^{q-1}} d_j d_i^\dagger + n_i^{(2)}, \quad (3.78)$$

$$\Lambda_i^q = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^q} d_j d_i^\dagger - \sum_{j(\neq i)} \frac{qn_j - 1}{z_i - z_j} n_i^{(2)}. \quad (3.79)$$

This leads to a three-body Hamiltonian annihilating the wave function

$$H = \sum_{i=1}^N \sum_{a=0}^q \Lambda_i^{a\dagger} \Lambda_i^a + \left( \sum_i n_i - \frac{N}{q} \right)^2, \quad (3.80)$$

where the last term fixes the number of particles. This Hamiltonian annihilates the wave function  $(q, \eta = 1)_1$ , however we find numerically that the ground space of this Hamiltonian is degenerate when  $q \geq 3$  and that the degeneracy does not depend on the number of sites. Other simple null fields of the theory constructed with the same current operators do not lead to operators acting on the wave function that would reduce this degeneracy.

### Parent Hamiltonians for the $(q, \eta = 1)_{1/2}$ , $q \geq 2$ Moore-Read states

We can now use the previous results to construct parent Hamiltonians for the  $(q, \eta = 1)_{1/2}$ ,  $q \geq 2$  Moore-Read states. They are projections of the  $(q, \eta = 1)_1$  Moore-Read states in the subspace allowing only for single occupation at each site, that we denote  $\mathcal{H}^1$ . Let us also define the Hilbert space  $\mathcal{H}^2$  spanned by basis elements containing at least one site with two particles. We will now project the operators annihilating the  $(q, \eta = 1)_1$  Moore-Read states onto  $\mathcal{H}^1$  in order to get operators annihilating the  $(q, \eta = 1)_{1/2}$  Moore-Read states (see Appendix B.3 for the detailed derivation):

$$\Lambda'^0 = \sum_i d_i, \quad (3.81)$$

$$\Lambda_i'^p = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^p} d_j n_i^{(1)}, \quad (3.82)$$

$$\Lambda_i'^{q-1} = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^q} d_j n_i^{(1)} + \sum_{j(\neq i)} \sum_{h(\neq i)} \frac{1}{(z_i - z_h)^{q-1}} \frac{q n_j^{(1)} - 1}{z_i - z_j} d_h n_i^{(1)}. \quad (3.83)$$

This leads to a five-body Hamiltonian

$$H = \sum_{i=1}^N \sum_{a=0}^{q-1} \Lambda_i^{a\dagger} \Lambda_i^a + \left( \sum_i n_i^{(1)} - \frac{N}{q} \right)^2. \quad (3.84)$$

As in the previous case, this parent Hamiltonian has a single ground state only when  $q = 2$  and this ground state is the  $(q = 2, \eta = 1)_{1/2}$  Moore-Read state.



### 3.4.5 Local Hamiltonians in one and two dimensions





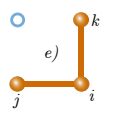
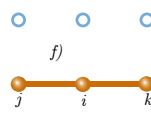
The parent Hamiltonians we have derived involve three or five-body interactions between all sites on the lattice. These Hamiltonians would therefore be very difficult to implement in experiments. However in some cases it has turned out that states constructed from correlators of conformal fields had very high overlaps with ground states of local Hamiltonians [1, 38, 54, 195]. This has led to a protocol to implement one of these states in experiments [54, 168]. In this section we show that there is a local Hamiltonian for which the ground state is close to the  $(q = 1, \eta = 1)_1$  Moore-Read state in one and in two dimensions and that this result is also true for the  $(q = 2, \eta = 1)_{1/2}$  Moore-Read state in one dimension.

#### Local Hamiltonians for the $(q = 1, \eta = 1)_1$ Moore-Read state

In one dimension the case of the  $(q = 1, \eta = 1)_1$  Moore-Read state was studied in Ref. [70]. It was shown that this state has a high overlap with the ground state of the bilinear-biquadratic spin 1 Hamiltonian

$$H_{1D}^{(1)} = \sum_{i=1}^N \left[ \cos(\beta) t_i^a t_{i+1}^a + \sin(\beta) (t_i^a t_{i+1}^a)^2 \right], \quad (3.85)$$

with periodic boundary conditions, when  $\beta = -0.3213$ . Note that this Hamiltonian includes the 2-body terms present in the parent Hamiltonian (3.75).

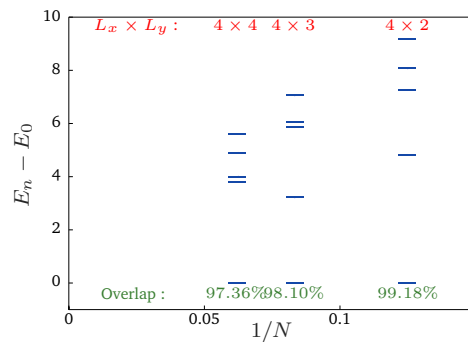
Operator	Configuration	Coefficient
$t_i^a t_j^a$	 a)	a) 1
	 b)	b) 0.6227
$(t_i^a t_j^a)^2$	 c)	c) -0.1762
	 d)	d) 0.3226
$t_i^a t_j^a t_i^b t_k^b$	 e)	e) 0.4637i f) 0.0208 + c.c. + j ↔ k
	 f)	

**Tab. 3.2.:** Terms in the Hamiltonian  $H_{2D}$  and coefficients obtained after numerical optimization on a  $4 \times 4$  lattice.

We now study the two dimensional case and build a local Hamiltonian from the parent Hamiltonian (3.75). The operators  $\Lambda_i^a$  contain 2-body interactions between sites  $i$  and  $j$ . We cut these operators by keeping only terms for which the sites  $i$  and  $j$  are nearest-neighbours on the square lattice. This leads to a local Hamiltonian with three-body interactions. In addition to these terms, we include the two-body interactions between next-nearest neighbours present in the parent Hamiltonian. All six terms included in our trial Hamiltonian, that we denote  $H_{2D}$ , are shown in Table 3.2. Note that the coefficients of these terms in the exact parent Hamiltonian are position-dependent. In our local Hamiltonian, however, we choose them to be position-independent and invariant under rotations.

By exact diagonalization and optimization on these coefficients, we find that there is a local Hamiltonian for which the overlap  $|\langle \psi_H | \psi_{\text{Moore-Read}} \rangle|$  between the ground state and the  $(q = 1, \eta = 1)_1$  Moore-Read state on a  $4 \times 4$  square lattice on the plane is 97.36%. Considering the size of the Hilbert space  $3^{16} \approx 4 \times 10^7$ , this overlap is very high. Note that with the same parameters, the overlap is also above 98% on a  $4 \times 3$  or on a  $4 \times 2$  lattice. On a cylinder geometry, i.e. periodic boundary conditions in one direction, the overlap on a  $4 \times 4$  square lattice is 97.21%.

Compared to the local Hamiltonian found in Ref. [53], which is for a state that is equivalent to the  $(q = 1, \eta = 1)_1$  Moore-Read state in the thermodynamic limit, but different on finite lattices, the Hamiltonian we find has less free parameters (5 instead of 11) to fine-tune, which might make it easier to implement. Moreover, the very good scaling with lattice sizes lets us expect that a good agreement will persist on larger lattices. In Fig. 3.25, we show the low-energy spectrum of this local Hamiltonian. This figure is compatible with having a gap in the thermodynamic limit, but the limitations on the system sizes that we can consider prevent us from making a reliable extrapolation.



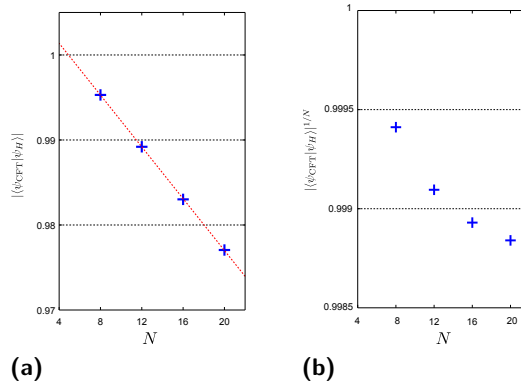
**Fig. 3.25.:** Energy difference to the ground state energy for the first excited states of the Hamiltonian  $H_{2D}$  for different sizes of lattices. The overlap between the ground state of this Hamiltonian and the  $(q = 1, \eta = 1)_1$  Moore-Read state is indicated below each ground state. Figure reproduced from [2].

### Local Hamiltonian for the 1D $(q = 2, \eta = 1)_{1/2}$ Moore-Read state

The state  $(q = 2, \eta = 1)_{1/2}$  Moore-Read state has a five-body parent Hamiltonian. Let us consider the one dimensional case. If we cut the  $\Lambda_i^a$  operators by keeping only terms for which the sites  $i$  and  $j$  are nearest-neighbours, we get a local Hamiltonian with several terms. We find however numerically that even a smaller number of terms is already sufficient to get a good overlap. Specifically, we choose to keep only the simplest two-body and the simplest three-body terms, to obtain a local Hamiltonian with periodic boundary conditions

$$H_{1D}^{(2)} = \sum_{i=1}^N \left( n_i n_{i+1} n_{i+2} + \kappa d_i^\dagger d_{i+1} \right) + c.c., \quad (3.86)$$

where  $d_i^\dagger, d_i$  are fermionic creation and annihilation operators at site  $i$  and  $n_i = d_i^\dagger d_i$ . For  $\kappa = 0.274 + 0.052i$ , we find that the overlap between the ground state of  $H_{1D}^{(2)}$  and the  $(q = 2, \eta = 1)_{1/2}$  Moore-Read state is 97.71% for a chain with 20 spins (Fig. 3.26a).



**Fig. 3.26.:** (a) Overlap  $|\langle \psi_{\text{CFT}} | \psi_H \rangle|$  between the  $(q = 2, \eta = 1)_{1/2}$  Moore-Read state and the ground state of Hamiltonian  $H_{1D}^{(2)}$  with  $\kappa$  fixed to  $0.274 + 0.052i$ , as a function of the number  $N$  of lattice sites. The dotted line is a linear fit with equation  $y = 1.0075 - 0.00152N$ . If the overlap continues to follow this behavior at larger sizes, it will still be above 85% for a spin chain with 100 lattice sites. (b) Overlap per site  $|\langle \psi_{\text{CFT}} | \psi_H \rangle|^{1/N}$  between the same two states. Figure reproduced from [2].

In two dimensions, since the  $SU(2)$  symmetry is not present in this model, cutting the parent Hamiltonian leads to a local Hamiltonian with up to five-body interactions with many different coefficients. In addition, the fact that the five-body terms involve more sites means that each of them stretches over a larger part of the lattice. With the limited lattice sizes that we can consider with exact diagonalization, this is problematic because the local regions need to be small compared to the total size

of the lattice (otherwise it would not be expected that the same local Hamiltonian would also work for other lattice sizes). This suggests that even if a local Hamiltonian that is related to the exact Hamiltonian exists, we may not be able to find it with exact diagonalization. Instead of cutting the exact parent Hamiltonian, we therefore asked whether, by chance, a local Hamiltonian can be obtained if we restrict the range of the interactions to all interactions preserving the number of particles on all possible configurations inside a plaquette of the lattice. Optimizing the coefficients in this Hamiltonian, we did not, however, find a set of coefficients for these interactions for which the ground state of this Hamiltonian is close to the  $(q = 2, \eta = 1)_{1/2}$  Moore-Read state. Whether there exists a more complicated, but still local, Hamiltonian stabilizing this state therefore remains an interesting open problem.

### 3.5 Conclusion

In this section we have shown that lattice fractional quantum Hall states can be defined by writing a wave function as a correlator of conformal fields acting at particular lattice sites. The resulting infinite dimensional MPS can be studied for large sizes using Monte Carlo simulations and exact parent Hamiltonians having these states as ground states can be constructed.

We first studied lattice versions of Laughlin wave functions with number of particles per flux  $\nu = 1/q$  on arbitrary lattices with filling factor  $\mu = \eta/q$ , which allowed us to investigate lattice effects all the way along the interpolation between the continuum limit and the completely filled lattice. We found by computing correlations and topological entanglement entropies that the lattice Laughlin states on the square lattice have the same topological properties as the continuum Laughlin state in most of the parameter range, except at the particle-hole symmetric point  $\eta = q/2$  when  $q$  is larger than five. We found that this effect is only present on the square lattice and does not persist if the lattice is deformed to a triangular lattice or to the Kagome lattice, in which case frustration destroys the long-range order and the topological properties of the state are recovered. We defined edge states wave functions which have a modified density at the edge of the lattice, but share the same correlations in the bulk as the lattice Laughlin states. This then allowed us to construct few-body non-local parent Hamiltonians for the lattice Laughlin states on arbitrary lattices in a wide range of filling factors.

We then turned to the local excitations of these states and showed that lattice versions of the Laughlin quasihole and quasielectron wave functions can be defined. Unlike in the continuum, the quasielectrons on the lattice can be constructed as inverse quasiholes. This allowed us to characterize the braiding properties, size and charge of both quasiholes and quasielectrons. In particular we constructed models for which there is a symmetry between quasiholes and quasielectrons, and

studied how this symmetry is broken along the interpolation between the lattice and the continuum. Parent hamiltonians for various states containing quasiholes and quasielectrons have been constructed. These allow creation, braiding and fusion of anyons simply by changing the interaction strengths in the Hamiltonian. It was also found that the states we introduced are good approximations of particular fractional Chern insulator states.

Finally we investigated how the previous constructions could be extended to states having non-Abelian anyons as their ground state. We introduced a family of strongly-correlated spin states on arbitrary lattices in one and two dimensions which correspond to Moore-Read states of bosons and fermions in the continuum limit. We provided numerical evidence that these states are critical states in one dimension and topological states in two dimensions, and that their properties remain unchanged along the interpolation between lattice and continuum. Non-local parent Hamiltonians for these states were derived in the lattice limit, and we found that in some cases the states could be stabilized by local Hamiltonians in one and two dimensions.

Our results show that models in which at least the ground state can be found analytically allow to study with high precision phenomena appearing in quantum many-body systems, such as topology. The models we introduced can also be used as a starting point to identify simpler models that realize the same physics, but that are more realistic to realize in experiments.



## Neural-network quantum states and tensor networks: relationship and application to chiral states

In the previous chapters we have considered infinite dimensional tensor networks with few variational parameters that allowed us to construct models with interesting properties. The exact parent Hamiltonians we obtained were not local and difficult to realize experimentally. It is therefore of interest to study other Hamiltonians that might be simpler to implement but that realize the same physics. In general, the ground state of these Hamiltonians cannot be found exactly and it is desirable to be able to approximate it with more sophisticated variational wave functions. In this chapter we investigate the power of different classes of states to represent ground states of many-body Hamiltonians, with a focus on chiral topological states introduced in the previous chapter.

Neural-network quantum states[71] were recently introduced as an ansatz for a many-body wave function based on an artificial neural network. While a few different neural-network architectures have been used[71, 208–210], the most promising results so far have been obtained with Boltzmann machines[85], which were originally designed for representing probability distributions. In particular, state of the art numerical results have been obtained on popular models with restricted Boltzmann machines (RBM) and it has been demonstrated that deep Boltzmann machines can represent gapped ground states of many-body Hamiltonians and quantum states generated by polynomial size quantum circuits efficiently[87, 88]. Neural-network quantum states can be optimized to approximate a ground state wave function using the Variational Monte Carlo (VMC) method[211, 212]. Seemingly unrelated classes of states that are optimized using the same method are classes of tensor-network states such as Entangled Plaquette States (EPS)[90–92] and String-Bond States (SBS)[93, 94], which construct a wave function as a product of tensor-network wave functions over overlapping clusters of spins.

In this chapter we will show that these classes of tensor-network states generalize neural-network quantum states taking the form of restricted Boltzmann machines. We will then exploit this connection to investigate the power of these different classes of states in approximating a chiral spin liquid introduced in the previous chapter.

We start in section 4.1 by introducing the Variational Monte Carlo method as an algorithm to optimize both tensor networks and neural networks to approximate the ground state of a many-body system. In section 4.2 we then provide a mapping between restricted Boltzmann machines, Entangled Plaquette States and String-Bond States and demonstrate that these classes of tensor networks generalize RBM states in arbitrary dimension. This relation provides new insights over the geometric structure of RBM and their efficiency and motivates us to define non-local String-Bond States which provide a way of combining RBM and SBS together, while taking leverage of both the entanglement structure of tensor networks and the efficiency of neural networks. In section 4.3 we then apply these methods to the problem of approximating chiral topological states. We first provide a way to represent the Kalmeyer-Laughlin wave function analytically as a RBM, and use its parent Hamiltonian to evaluate the accuracy that can be achieved numerically. We then compare the power of EPS, SBS and RBM in approximating the ground state of a chiral spin liquid and give evidence that the properties of the ground state of the model can be accurately reproduced using our method. This chapter is based on Ref. [4] and is a modification thereof:

- [4] : Ivan Glasser, Nicola Pancotti, Moritz August, Ivan D. Rodriguez and J. Ignacio Cirac, ‘Neural-Network Quantum States, String-Bond States, and Chiral Topological States’, *Physical Review X* **8**, 011006 (2018), <http://dx.doi.org/10.1103/PhysRevX.8.011006>, under CC BY 4.0 license.

## 4.1 Variational Monte Carlo with tensor networks and neural-network quantum states

In the previous chapters we have mostly discussed the construction of exact models from infinite dimensional tensor networks. Given a general Hamiltonian  $H$ , we would also like to be able to obtain its ground state in the form of a tensor network. The algorithms that one can use for solving this task in general depend on the class of tensor network states that are considered. The density matrix renormalization group algorithm[13] is an example of such an algorithm for Matrix Product States (MPS). In this section we introduce the Variational Monte Carlo (VMC) method[211, 212], which is a general algorithm for finding an approximation to a ground state wave function in a class of states for which the wave function coefficients can be computed efficiently. In particular we introduce the Stochastic Reconfiguration method, a powerful optimization method for performing VMC. We show that it can be used to optimize MPS, but also other classes of tensor-network states such as EPS and SBS. We then introduce neural-network quantum states based on Boltzmann machines and explain how they are also optimized using the same method.



### 4.1.1 Variational Monte Carlo and Stochastic Reconfiguration

Finding the ground state wave function  $|\psi_0\rangle$  satisfying the Schrödinger equation  $H|\psi_0\rangle = E_0|\psi_0\rangle$  is an eigenvalue problem, but this problem can be mapped to an optimization problem. Indeed the variational principle states that the energy of any quantum state is higher than the energy of the ground state, with equality only for the ground state. Finding the ground state therefore amounts to finding the wave function of lowest energy. Consider a generic quantum state defined on a lattice with  $N$  spins, written in the computational basis as

$$|\psi\rangle = \sum_{s_1, \dots, s_N} \psi(s_1, \dots, s_N) |s_1, \dots, s_N\rangle. \quad (4.1)$$

Instead of finding the exponentially many parameters  $\psi(s_1, \dots, s_N)$  such that the energy is minimized, one may restrict the search to a specific class of states specified by a particular choice of function  $\psi_{\mathbf{w}}(s_1, \dots, s_N)$  depending on polynomially many variational parameters  $\mathbf{w} = \{w_1, \dots, w_{N_p}\}$ . The energy of such a state can be expressed as

$$E_{\mathbf{w}} = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} = \sum_{\mathbf{s}} p(\mathbf{s}) E_{\text{loc}}(\mathbf{s}), \quad (4.2)$$

where  $\mathbf{s} = s_1, \dots, s_N$  is a spin configuration,  $p(\mathbf{s}) = \frac{|\psi_{\mathbf{w}}(\mathbf{s})|^2}{\sum_{\mathbf{s}} |\psi_{\mathbf{w}}(\mathbf{s})|^2}$  and the local energy is  $E_{\text{loc}}(\mathbf{s}) = \sum_{\mathbf{s}'} \langle \mathbf{s} | H | \mathbf{s}' \rangle \frac{\psi_{\mathbf{w}}(\mathbf{s}')}{\psi_{\mathbf{w}}(\mathbf{s})}$ . Note that if the function  $\psi_{\mathbf{w}}$  is given in a way that is efficient to evaluate on a spin configuration, then  $p(\mathbf{s})$  can be evaluated efficiently, and the local energy can also be evaluated efficiently for Hamiltonians involving few-body interactions.  $p$  is a positive and normalized function over the Hilbert space, it therefore defines a probability distribution. The energy is then the expectation value of the local energy with respect to a probability distribution, and can be evaluated using Markov Chain Monte Carlo sampling techniques as explained in section 2.2.3.

To minimize the energy, several algorithms can be used. In gradient based methods, one also needs to compute its gradient  $\mathbf{f}$ , which can be expressed using

$$\frac{\partial E_{\mathbf{w}}}{\partial w_i} = 2 \sum_{\mathbf{s}} p(\mathbf{s}) \Delta_{w_i}(\mathbf{s})^* (E_{\text{loc}}(\mathbf{s}) - E_{\mathbf{w}}), \quad (4.3)$$

where we have defined  $\Delta_{w_i}(\mathbf{s}) = \frac{1}{\psi_{\mathbf{w}}(\mathbf{s})} \frac{\partial \psi_{\mathbf{w}}(\mathbf{s})}{\partial w_i}$  as the log-derivative of the wave function with respect to some parameter  $w_i$ . The gradient is also an expectation value with respect to the probability distribution  $p$  and can therefore be sampled at the same time. The simplest possible algorithm is gradient descent, which consist in the following steps:

1. Initialize the variational parameters  $\mathbf{w}^0$  at iteration 0 to random numbers.
2. At iteration  $t$ , the parameters take the values  $\mathbf{w}^t$ . Run a Markov Chain Monte Carlo algorithm to compute the energy of the corresponding wave function, as well as its gradient with respect to the variational parameter.
3. Update the parameters by small steps in the direction of negative energy derivative:  $\mathbf{w}^{t+1} \leftarrow \mathbf{w}^t + \gamma$ , where  $\gamma = -\alpha \mathbf{f}$ .  $\alpha$  is called the step size or learning rate.
4. Repeat steps [2-3] until convergence of the energy.

This method only requires the efficient computation of  $\frac{\psi_{\mathbf{w}}(s')}{\psi_{\mathbf{w}}(s)}$  for two spin configurations  $s$  and  $s'$ , as well as the log-derivative of the wave function  $\Delta_{\mathbf{w}}(s)$ . More efficient optimization methods can be used, such as conjugate-gradient descent which uses information from the gradient at previous steps or the linear method[213–215] which use second-order derivatives of the energy. Newton method[216] would result in a change of parameters given instead by

$$\gamma = -\alpha \mathbf{H}^{-1} \mathbf{f}, \quad (4.4)$$

where  $\mathbf{H}$  is the Hessian of the energy.

Small changes of the variational parameters may however lead to big changes in the wave function. Taking into account the metric of changes of the wave function leads to the Stochastic Reconfiguration[217, 218] method, which is equivalent to the natural gradient descent[219] and replaces the Hessian in (4.4) by the covariance matrix of the derivatives of the wave function, avoiding the computation of the second-order derivatives of the energy. The Stochastic Reconfiguration method can also be viewed as an approximate imaginary-time evolution in the tangent space of the wave function. Consider the normalized wave function  $|\bar{\psi}_0\rangle$  and its derivatives

$$|\bar{\psi}_0\rangle = \frac{|\psi_0\rangle}{\sqrt{\langle\psi_0|\psi_0\rangle}}, \quad (4.5)$$

$$|\bar{\psi}_i\rangle = \frac{|\psi_i\rangle}{\sqrt{\langle\psi_0|\psi_0\rangle}} - \frac{\langle\psi_0|\psi_i\rangle}{\langle\psi_0|\psi_0\rangle} \frac{|\psi_0\rangle}{\sqrt{\langle\psi_0|\psi_0\rangle}}, \quad (4.6)$$

defining a non-orthogonal basis set  $\Omega$ . Expanding the wave function to linear order around some parameters  $\mathbf{w}$  leads to

$$|\bar{\psi}(\mathbf{w} + \gamma)\rangle = \sum_{i=0}^{N_w} \gamma_i |\bar{\psi}_i\rangle. \quad (4.7)$$

To minimize the energy, one can apply the imaginary-time evolution operator  $e^{-\alpha H}$ , which for small  $\alpha$  can be extended as  $1 - \alpha H$ . The change of coefficients  $\gamma$  is found

by applying this operator to  $|\bar{\psi}(\mathbf{w} + \boldsymbol{\gamma})\rangle$  and projecting in the set  $\Omega$ , which leads to the equation

$$-\alpha \langle \bar{\psi}_i | H | \bar{\psi}_0 \rangle = \sum_{j=1}^M \langle \bar{\psi}_i | \bar{\psi}_j \rangle \gamma_j, \quad (4.8)$$

which can be rewritten as

$$\boldsymbol{\gamma} = -\alpha \mathbf{S}^{-1} \mathbf{f}, \quad (4.9)$$

where  $S_{ij} = \langle \bar{\psi}_i | \bar{\psi}_j \rangle$  and  $f_i = \langle \bar{\psi}_i | H | \bar{\psi}_0 \rangle$ . If we expand these expressions as expectation values over the probability distribution  $p(\mathbf{s}) = \frac{|\psi_{\mathbf{w}}(\mathbf{s})|^2}{\sum_{\mathbf{s}} |\psi_{\mathbf{w}}(\mathbf{s})|^2}$ , we obtain

$$f_i = \langle \Delta_i^* E_{\text{loc}} \rangle - \langle \Delta_i^* \rangle \langle E_{\text{loc}} \rangle, \quad (4.10)$$

$$S_{ij} = \langle \Delta_i^* \Delta_j \rangle - \langle \Delta_i^* \rangle \langle \Delta_j \rangle. \quad (4.11)$$

Finally, the complete algorithm of the Stochastic Reconfiguration method is as follows:

1. Initialize the variational parameters  $\mathbf{w}^0$  at iteration 0 to random numbers.
2. At iteration  $t$ , the parameters take the values  $\mathbf{w}^t$ . Run a Markov Chain Monte Carlo algorithm to compute stochastic estimates for the expectation values  $\langle \Delta_j \rangle$ ,  $\langle E_{\text{loc}} \rangle$ ,  $\langle \Delta_i^* E_{\text{loc}} \rangle$ ,  $\langle \Delta_i^* \Delta_j \rangle$ .
3. Construct the vector  $\mathbf{f}$  and matrix  $\mathbf{S}$ .
4. Update the parameters according to:  $\mathbf{w}^{t+1} \leftarrow \mathbf{w}^t + \boldsymbol{\gamma}$ , where  $\boldsymbol{\gamma} = -\alpha \mathbf{S}^{-1} \mathbf{f}$ .
5. Repeat steps [2-4] until convergence of the energy.

In practice we randomly select a subset of the parameters at each iteration of the algorithm and update only these parameters, to reduce the computational cost associated with the operations dealing with  $\mathbf{f}$  and  $\mathbf{S}$ . Moreover we avoid forming the full matrix  $\mathbf{S}$  by instead solving (4.9) with a conjugate-gradient solver[220]. The Stochastic Reconfiguration method might lead to numerical instabilities in the solution of (4.9). These are prevented by adding a small constant  $\epsilon$  to the diagonal elements of the matrix  $\mathbf{S}$ , rotating the direction of change towards the steepest descent direction.

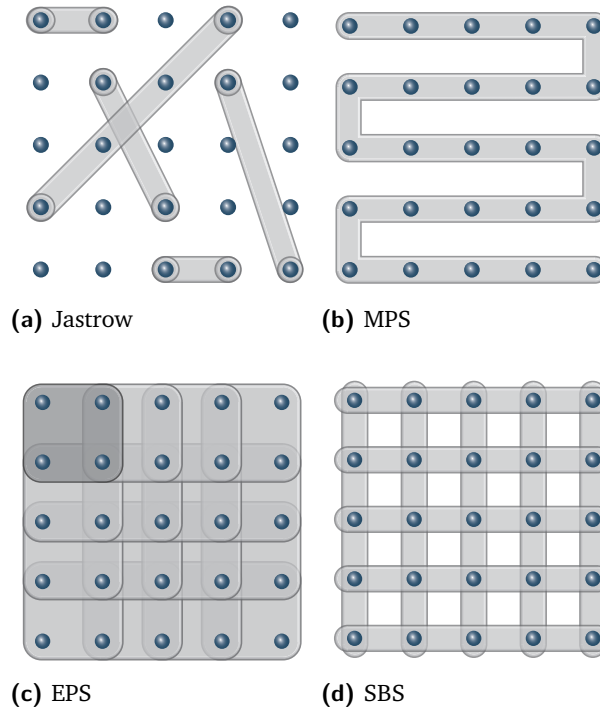
## 4.1.2 Variational Monte Carlo method with tensor networks

At this point one has to choose a parametrization for the wave function. One natural choice in one dimension is a MPS, defined in (2.5). MPS can be efficiently optimized using the density matrix renormalization group[13], but the previously described VMC method can also be applied[93, 221] by observing that the ratio of

two configurations is straightforward to compute, since one just needs to contract the MPS, and that the log-derivative with respect to some matrix  $A_k^{s'_k}$  is given by

$$\Delta_{A_k^{s'_k}}(\mathbf{s}) = \frac{\delta_{s_k, s'_k} \left( A_{k+1}^{s_k} \cdots A_N^{s_N} A_1^{s_1} A_{k-1}^{s_{k-1}} \right)^\top}{\text{Tr}(A_1^{s_1} \cdots A_N^{s_N})}. \quad (4.12)$$

In some cases, this method is less likely to be trapped in a local minimum than DMRG, since all coefficients can be updated at once. In addition, the cost only scales as  $O(D^3)$  in the bond dimension for periodic boundary conditions.



**Fig. 4.1.:** Geometry of ansatz wave functions: (a) Jastrow wave function include correlations within all pairs of spins. (b) Matrix Product States (MPS) in 2D cover the lattice in a snake pattern. (c) Entangled Plaquette States (EPS) include all spin correlations within each plaquette (2x2 on the figure) and mediate correlations between distant spins through overlapping plaquettes. (d) String-Bond States (SBS) cover the lattice with many 1D strings on which the interactions within spins are captured by a MPS. Figure reproduced from [4].

In this chapter we will be mostly concerned with two dimensional systems. In this case, a MPS can be defined by mapping the system to a line (Fig. 4.1b). Spins which sit close to each other might however be separated by a long distance on the line, so MPS in 2D fail to reproduce the local structure of the state, which leads to an exponential scaling of the computational resources needed with the system size[222]. Instead, one could use Projected Entangled Pair State (PEPS) (Fig. 1.2), but their exact contraction is  $\#P$  hard[223] and their optimization cannot rely on the standard VMC method without approximations. Instead we will consider other

classes of tensor-network states in more than one dimension for which the exact computation of the wave function is efficient, which allows for the direct use of the VMC method.

One way of defining such tensor networks relies on the definition of  $P$  small clusters of  $n_p$  spins, or plaquettes, on the lattice. These clusters may overlap, and because the clusters are small the wave function on each plaquette can be written by a tensor depending on the values of the spins in this plaquette. The wave function of the full quantum system is then taken to be the product of the wave functions in each plaquette. Because the plaquettes overlap, correlations take place between the different plaquettes (Fig. 4.1c). The wave function of such an Entangled Plaquette State (EPS, also called a Correlated Product State) is written as[90–92]:

$$\psi_{\mathbf{w}}(\mathbf{s}) = \prod_{p=1}^P C_p^{\mathbf{s}_p}, \quad (4.13)$$

where a coefficient  $C_p^{\mathbf{s}_p}$  is assigned to each of the  $2^{n_p}$  (for spin-1/2 particles) configurations  $\mathbf{s}_p = s_{a_1}, \dots, s_{a_{n_p}}$  of the spins on the plaquette  $p$ . Each  $C_p$  can be seen as the most general function on the Hilbert space corresponding to the spins in plaquette  $p$ . The accuracy can be improved by enlarging the size of the plaquettes. Moreover, once the spin configuration  $\mathbf{s}_p$  is fixed, the log-derivative of the wave function with respect to the variational parameters is simply

$$\Delta_{C_p^{\mathbf{s}_p}}(\mathbf{s}) = \frac{1}{C_p^{\mathbf{s}_p}}, \quad (4.14)$$

which is efficient to compute. The VMC method can therefore be applied to optimize EPS. Another choice of clusters of spins can be all pairs of spins on the lattice. The corresponding EPS is a Jastrow wave function[211, 224], with the form

$$\psi_{\mathbf{w}}(\mathbf{s}) = \prod_{i<j} C_{ij}^{s_i, s_j}. \quad (4.15)$$

This wave function does not presuppose a particular geometry: in general it is non-local due to the correlations between all pairs of spins (Fig. 4.1a). A local structure can be introduced by choosing a form for  $C_{ij}$  which decays with the distance between position  $i$  and  $j$ .

EPS are limited to small plaquettes since for each plaquette the number of coefficients scales exponentially with the size of the plaquette. To be able to use larger clusters, one can describe the state on each cluster of spins by a MPS, avoiding the exponentially many coefficients needed to describe the complete state of each cluster. The lattice is now cut in overlapping 1D strings which can mediate correlations on longer distances compared to local plaquettes (Fig. 4.1d). The resulting ansatz is

a String-Bond State (SBS)[93] defined by a set of strings  $i \in S$  (each string  $i$  is an ordered subset of the set of spins) and a MPS for each string:

$$\psi_{\mathbf{w}}(\mathbf{s}) = \prod_i \text{Tr} \left( \prod_{j \in i} A_{i,j}^{s_j} \right). \quad (4.16)$$

The choice of strings depends of the geometry of the problem considered. By using small strings covering small plaquettes and a large bond dimension, SBS generalize EPS. Instead, a single long string in a snake pattern includes a MPS in 2D. In 3D, strings parallel to the axes of the lattice have been chosen[94]. Since the form of the wave function is a product of MPS, the log-derivative with respect to some elements present in one of the MPS is simply the log-derivative for the corresponding MPS (4.12). The VMC procedure for optimizing SBS and MPS thus have the same cost. In addition, the ratio of two configurations which differ only by a few spins can be computed by considering only the strings including these spins, which speeds up the computation considerably. One can map a SBS to a MPS analytically, but the resulting MPS has a bond dimension exponential in the number of strings.

### 4.1.3 Variational Monte Carlo method with neural networks

It was recently realized that the VMC method can be viewed as a form of reinforcement learning, in which an agent interacts with an environment through actions and tries to maximize a reward, corresponding in this case to a low energy. This motivated the use of neural-network quantum states[71] in which the wave function has the structure of an artificial neural network. A few different networks have been investigated[71, 77, 208, 209], but the most promising results so far have been obtained with Boltzmann machines[85]. Boltzmann machines are a kind of generative stochastic artificial neural networks that can learn a probability distribution over the set of their inputs. In quantum many-body physics, the inputs are spin configurations and the wave function is interpreted as a (complex) probability distribution that the networks tries to approximate. Boltzmann machines consist of two sets of binary units (classical spins): the visible units  $v_i$ ,  $i \in \{1, \dots, N\}$ , corresponding to the configurations of the original spins in a chosen basis, and the hidden units  $h_j$ ,  $j \in \{1, \dots, M\}$ , which introduce correlations between the visible units. The whole system interacts through an Ising interaction which defines a joint probability distribution over the visible and hidden units as the Boltzmann weight of this Hamiltonian:

$$P(\mathbf{v}, \mathbf{h}) = \frac{1}{Z} e^{\mathcal{H}(\mathbf{v}, \mathbf{h})}, \quad (4.17)$$

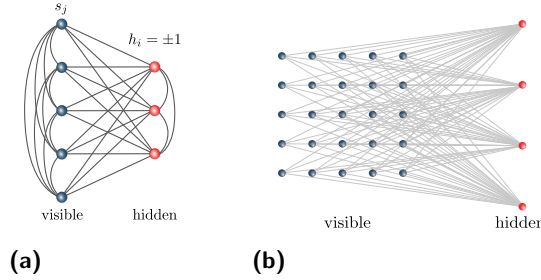
where the Hamiltonian  $\mathcal{H}$  is defined as

$$\mathcal{H} = \sum_j a_j v_j + \sum_i b_i h_i + \sum_{i<j} c_{ij} v_i v_j + \sum_{i,j} w_{ij} h_i v_j + \sum_{i<j} d_{ij} h_i h_j,$$

and  $Z$  is the partition function. The marginal probability of a visible configuration is then given by summing over all possible hidden configurations:

$$P(\mathbf{v}) = \sum_{\mathbf{h}} \frac{1}{Z} e^{\mathcal{H}(\mathbf{v},\mathbf{h})}. \quad (4.18)$$

This quantity, when the inputs are spin configurations in a given basis, is chosen as the wave function  $\psi_{\mathbf{w}}(\mathbf{s}) = P(\mathbf{s})$ . The variational parameters are the complex parameters of the Ising Hamiltonian. In the case where there are interactions between the hidden units (Fig. 4.2a), the Boltzmann machine is called a deep (or unrestricted) Boltzmann machine. It has been shown that deep Boltzmann machines



**Fig. 4.2.:** (a) Boltzmann machines approximate a probability distribution by the Boltzmann weights of an Ising Hamiltonian on a graph including visible units (corresponding to the spins  $s_j$ ) and hidden units  $h_i$  which are summed over. (b) Restricted Boltzmann machines (here in 2D) only include interactions between the visible and the hidden units. Figure reproduced from [4].

can efficiently represent ground states of many-body Hamiltonians with polynomial-size gap, local tensor-network state and quantum states generated by any polynomial size quantum circuits[87–89]. On the other hand, computing the wave function  $\psi_{\mathbf{w}}(\mathbf{s})$  of such a deep Boltzmann machine in the general case is intractable, due to the exponential sum over the hidden variables, so the VMC method cannot be applied to deep Boltzmann machines without approximations. Restricted Boltzmann machines (RBM), which only include interactions between the visible and hidden units (as well as the one-body interaction terms which correspond to biases), do not suffer from such a limitation. In this case, the sum over the hidden units can be performed analytically and the resulting wave function can be written as (here we take the hidden units to have values  $\pm 1$ ):

$$\psi_{\mathbf{w}}(\mathbf{s}) = e^{\sum_j a_j s_j} \prod_i \cosh \left( b_i + \sum_j w_{ij} s_j \right). \quad (4.19)$$

Because this expression, as well as its derivative, can be evaluated efficiently, RBM can be optimized using the VMC method. RBM can represent many quantum states of interest, such as the toric code[86], any graph state, cluster states and coherent thermal states[87]; the possibility of computing efficiently  $\psi_{\mathbf{w}}(\mathbf{s})$  prevents it however to approximate all PEPS and ground states of local Hamiltonians, since there are ground states of local Hamiltonians for which computing the coefficients of the wave function is intractable[87].

## 4.2 Relationship between tensor-network and neural-network states

Boltzmann machines and tensor networks are seemingly unrelated, but we will see in this section that there are in fact strong connections between the two ansatz classes. It was recently shown that while fully connected RBM can exhibit volume-law entanglement, contrary to local tensor networks, all short-range RBM satisfy an area law[225]. Moreover short-range and sufficiently sparse RBM can be written as a MPS[89], but doing so for a fully-connected RBM would require an exponential scaling of the bond dimension with the size of the system. In this section we will show that short-range RBM are special cases of EPS, while fully-connected RBM are special cases of SBS in arbitrary dimension. We will also prove that any Jastrow wave function can be written as a RBM. These connections have implications for the geometry of RBM and motivate the definition of non-local SBS.

### 4.2.1 Jastrow wave functions, restricted Boltzmann machines and the Majumdar-Gosh model

Let us first consider the simple case of the Jastrow wave function (4.15). Boltzmann machines including only interactions between the visible units have a wave function

$$\psi_{\mathbf{w}}(\mathbf{s}) = \prod_k e^{a_k s_k} \prod_{i < j} e^{c_{ij} s_i s_j}, \quad (4.20)$$

which has the form of a product between functions of pairs of spins, and is thus a Jastrow wave function. More generally, semi-restricted Boltzmann machines including interactions between visible units as well as between hidden and visible units are a product of a RBM and a Jastrow factor. Nevertheless, one may ask



whether a RBM alone is enough to describe a Jastrow factor. Let us first rewrite the RBM as

$$\psi_{\mathbf{w}}(\mathbf{s}) = \prod_j A_j^{s_j} \prod_i \left( B_i \prod_j W_{ij}^{s_j} + \frac{1}{B_i \prod_j W_{ij}^{s_j}} \right), \quad (4.21)$$

where we have redefined the parameters with uppercase letters as the exponential of the original parameters, thus removing the exponentials in the hyperbolic cosine. This form will be convenient for the numerical simulations presented later. Since Jastrow wave functions are a product of functions of all pairs of spins, let us show that a RBM with one hidden unit can represent any function of two spins. It then follows that a RBM with  $M = N(N - 1)/2$  hidden units, each representing a function of one pair of spins, can represent a Jastrow wave function with polynomial resources. We have to solve for a system of four non-linear equations with  $s_1, s_2 \in \{-1, 1\}$  and  $f$  the most general function of two spins, parametrized by its values:

$$F_{11} = A_1 A_2 \left( W_1 W_2 + \frac{1}{W_1 W_2} \right) \quad (4.22)$$

$$F_{-1-1} = \frac{1}{A_1 A_2} \left( W_1 W_2 + \frac{1}{W_1 W_2} \right) \quad (4.23)$$

$$F_{1-1} = \frac{A_1}{A_2} \left( \frac{W_1}{W_2} + \frac{W_2}{W_1} \right) \quad (4.24)$$

$$F_{-11} = \frac{A_2}{A_1} \left( \frac{W_2}{W_1} + \frac{W_1}{W_2} \right), \quad (4.25)$$

where we have set  $B_1 = B_2 = 1$ . The RBM is well defined when all parameters are non-zero and we change of variables by defining  $X = W_1 W_2$ ,  $Y = \frac{W_1}{W_2}$ ,  $A = A_1 A_2$ ,  $B = \frac{A_1}{A_2}$ , obtaining a new set of equations:

$$F_{-1-1} A^2 = F_{11} \quad (4.26)$$

$$F_{-11} B^2 = F_{1-1} \quad (4.27)$$

$$X^2 - \frac{1}{A} X + 1 = 0 \quad (4.28)$$

$$Y^2 - \frac{1}{B} Y + 1 = 0. \quad (4.29)$$

We first suppose that the values  $F_{s_i s_j}$  are non-zero. These quadratic equations all have non-zero analytical solutions in the complex plane, that we denote  $A_0, B_0, X_0, Y_0$ . The original parameters are then the solutions of

$$W_1^2 = X_0 Y_0 \quad (4.30)$$

$$W_2^2 = X_0 / Y_0 \quad (4.31)$$

$$A_1^2 = A_0 B_0 \quad (4.32)$$

$$A_2^2 = A_0 / B_0, \quad (4.33)$$

which is again a set of quadratic equations with non-zero analytical solutions. If  $F_{11} = F_{-1-1} = 0$  (resp.  $F_{1-1} = F_{-11} = 0$ ), the exact solution is given directly by  $A_0 = 1, X_0 = i$  (resp.  $B_0 = 1, Y = i$ ). In the remaining cases where some  $F_{s_i s_j}$  are zeros, the equations do not always have an exact solution, but the function can still be approximated to arbitrary precision. This case corresponds to strong restrictions on the part of the Hilbert space which is used to write the wave function and these constraints can also be imposed on the states directly by adding a delta function to the wave function which is equal to 1 only when the constraints on the spins are satisfied.

As an application, we use this result to write the ground state of the Majumdar-Ghosh model[226] exactly as a RBM. The Majumdar-Ghosh model is defined by the following spin-1/2 Hamiltonian:

$$H = J \sum_{i=1}^{N-1} \mathbf{S}_i \cdot \mathbf{S}_{i+1} + \frac{J}{2} \sum_{i=1}^{N-2} \mathbf{S}_i \cdot \mathbf{S}_{i+2} \quad (4.34)$$

The ground state wave function is a product of singlets formed by neighboring pairs of spins:

$$|\psi\rangle \propto \prod_{n=1}^{N/2} |\uparrow_{2n-1}\rangle |\downarrow_{2n}\rangle - |\downarrow_{2n-1}\rangle |\uparrow_{2n}\rangle, \quad (4.35)$$

This wave function can also be expanded in the computational basis as

$$\psi(s_1, \dots, s_N) \propto \prod_{n=1}^{N/2} (-1)^{(s_{2n-1}+3)/2} \delta_{s_{2n-1} \neq s_{2n}}, \quad (4.36)$$

$$\propto \prod_{n=1}^{N/2} f(s_{2n-1}, s_{2n}). \quad (4.37)$$

Using the previous result, each function of two spins  $f$  can be written as a RBM using one hidden unit, which leads to a RBM representation of the ground states with  $M = N/2$  hidden units. We also find numerically on small systems that a RBM using less than  $M = N/2$  has higher energy than the ground state, which suggests that  $M = N/2$  could be optimal.

## 4.2.2 Short-range restricted Boltzmann machines are Entangled Plaquette States

We now turn to the specific case of RBM with short-range connections (sRBM). This includes all quantum states that have previously been written exactly as a RBM, such as for example the toric code or the 1D symmetry-protected topological

cluster state[86]. Such states have weights connections between visible and hidden units that are local. Each hidden unit is connected to a local region with at most  $d$  neighboring spins. If we divide the lattice into  $M$  subsets  $p_i$ ,  $i \in \{1, \dots, M\}$ , the wave function can be rewritten as (we omit here the biases  $a_j$  which are local one-body terms):

$$\psi_{\mathbf{w}}(\mathbf{s}) = \prod_{i=1}^M \cosh \left( b_i + \sum_{j \in p_i} w_{ij} s_j \right) \quad (4.38)$$

$$= \prod_{i=1}^M C_i^{s_i}, \quad (4.39)$$

where  $s_i$  is the spin configuration in the subset  $p_i$ . This is the form of an EPS (Fig. 4.3a), so short-range RBM are EPS. For translational invariant systems, the short-range RBM becomes a convolutional RBM, which corresponds to a translational invariant EPS. The main difference between a short-range RBM and an EPS is that the RBM considers a very specific function among all possible functions of the spins inside a plaquette, hence EPS are more general than short-range RBM. This also implies that the entanglement of short-range RBM follows an area law. The main advantage of short-range RBM over EPS is that due to the exponential scaling of EPS with the size of the plaquettes, larger plaquettes can be used in short-range RBM than in EPS. In practice fully-connected RBM should therefore be preferred to short-range RBM. Other results relating EPS and RBM have been obtained in [227].

### 4.2.3 Fully-connected restricted Boltzmann machines are String-Bond States

Fully-connected RBM do not necessarily satisfy an area law[225], and hence cannot always be approximated by local tensor networks. As we will see, they can nevertheless be represented by SBS by writing the RBM wave function as (here we omit the bias  $a_j$ ):

$$\psi_{\mathbf{w}}(\mathbf{s}) = \prod_i \cosh \left( b_i + \sum_j w_{ij} s_j \right) \quad (4.40)$$

$$\propto \prod_i \left( e^{b_i + \sum_j w_{ij} s_j} + e^{-b_i - \sum_j w_{ij} s_j} \right) \quad (4.41)$$

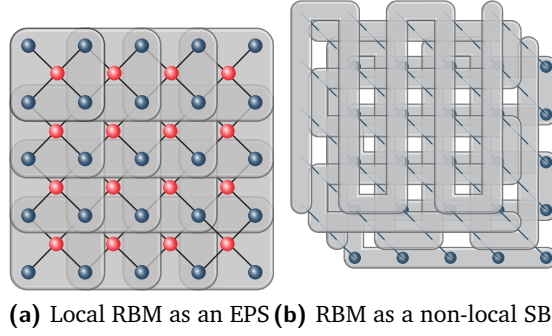
$$\propto \prod_i \text{Tr} \begin{pmatrix} e^{b_i + \sum_j w_{ij} s_j} & 0 \\ 0 & e^{-b_i - \sum_j w_{ij} s_j} \end{pmatrix} \quad (4.42)$$

$$\propto \prod_i \text{Tr} \left( \prod_{j \in i} A_{i,j}^{s_j} \right), \quad (4.43)$$

where

$$A_{i,j}^{s_j} = \begin{pmatrix} e^{b_i/N + w_{ij}s_j} & 0 \\ 0 & e^{-b_i/N - w_{ij}s_j} \end{pmatrix} \quad (4.44)$$

are diagonal matrices of bond dimension 2. This shows that RBM are SBS, as the wave function can be written as a product of MPS over strings, where each hidden unit corresponds to one string. The difference between SBS as depicted in Fig. 4.1d and a RBM is the geometry of the strings. In a fully-connected RBM, each string goes over the full lattice.



**Fig. 4.3.:** (a) A locally connected RBM is an EPS where each plaquette encodes the local connections to a hidden unit. (b) Once expressed as a SBS a fully-connected RBM can be represented by many strings on top of each other. Enlarging the RBM by using non-commuting matrices to non-local SBS induces a geometry in each string. Figure reproduced from [4].

#### 4.2.4 Generalizing restricted Boltzmann machines to non-local String-Bond States

We have seen that RBM consists in SBS with many strings overlapping on the full lattice. The matrices in each string of a RBM are diagonal, hence commute. Each string therefore does not have a fixed geometry and can adapt to stronger correlations in different parts of the lattice, even over long distances. This motivates the definition of SBS with diagonal matrices in which each string covers the full lattice (Fig. 4.3b). We denote these states as non-local dSBS. They generalize RBM by relaxing the constraints on the RBM parameters to the most general diagonal matrix, with a bond dimension which may be larger than two. For example taking the matrices

$$A_{i,j}^{s_j} = \begin{pmatrix} a_{i,j}^{s_j} & 0 & 0 \\ 0 & b_{i,j}^{s_j} & 0 \\ 0 & 0 & c_{i,j}^{s_j} \end{pmatrix}, \quad (4.45)$$

with different parameters  $a_{i,j}^{s_j}$  for each string, lattice site and spin direction, leads to the wave function (here  $D = 3$ ):

$$\psi_{\mathbf{w}}(\mathbf{s}) = \prod_i \left( \prod_j a_{i,j}^{s_j} + \prod_j b_{i,j}^{s_j} + \prod_j c_{i,j}^{s_j} \right). \quad (4.46)$$

We can now generalize such a wave function to larger spins than spin-1/2, since the values taken by  $s_i$  are just indexing the matrices. This provides a way of defining a natural generalization of RBM which can handle systems with larger physical dimension.

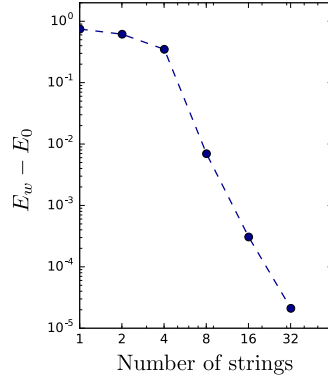
Another way in which RBM can be generalized is by including non-commuting matrices. This fixes the geometry of each string by defining an order and also enables to represent more complicated interactions. In the following we will refer to SBS in such a geometry as non-local SBS. The advantage of this approach is that it can capture more complex correlations within each string, while introducing additional geometric information about the problem at hand. It comes however at a greater numerical cost than non-local dSBS or RBM due to the additional number of parameters. Note that one can use an already optimized RBM or dSBS as a way of initializing a non-local SBS.

It is not always clear which representation should be preferred. In some cases, the SBS representation is more compact than the RBM/dSBS representation, but the opposite can be true in different cases. Let us consider again the ground state of the Majumdar-Gosh Hamiltonian, which we previously wrote as a RBM with  $M = N/2$  hidden units. The ground state of the Majumdar-Gosh Hamiltonian can also be written as a simple MPS with bond dimension 3 and periodic boundary conditions, with matrices [16]

$$A_n^{s_n=-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 \end{pmatrix}, \quad A_n^{s_n=1} = \begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \quad (4.47)$$

Since this state is a MPS, it is also a SBS with 1 string. The RBM representation of the same state requires  $N/2$  strings. In practice the number of non-zero coefficients are comparable, since in both cases the representation is sparse, but for numerical purposes a fully-connected RBM needs of the order  $O(N^2)$  parameters before finding the exact ground state, while a MPS or SBS with one string will need  $O(N)$  parameters for both open and periodic boundary conditions. Another example is the AKLT model[117], previously defined in (2.6). Its ground state can be written as a MPS of bond dimension 2, as given in (2.7). We have numerically optimized the spin-1 extension of a RBM with form (4.46) and found that already for small sizes of the chain a much higher number of parameters is required to approach the ground state energy as compared to a SBS with non-commuting matrices, which is

exact with one string of bond dimension 2 (Fig. 4.4). We will also show in section 4.3 that in some other cases the RBM needs less parameters than a SBS to obtain a similar energy. This demonstrates that both RBM and SBS have advantages and that their efficiency depends on the particular model that is investigated.



**Fig. 4.4.:** Energy difference with the exact ground state energy of a spin-1 extension of a RBM (4.46) with  $D = 2$  and different number of strings for the AKLT model on a spin-1 chain with 8 spins. A non-local SBS with non-commuting matrices and one string is exact within numerical accuracy. Figure reproduced from [4].

In order to combine the advantages of both RBM (efficient to compute, few parameters) and SBS (complex representation, geometric interpretation), we suggest to use SBS in which some strings have a full MPS, while some strings have the form of an RBM. Optimizing a wave function starting from a random state is in general difficult, but in many cases an initial approximation of the ground state can be obtained. This initial approximation can then be used in conjunction with the previous ansatz classes by simply multiplying a SBS with the initial approximation. For the resulting wave function

$$\psi_{\mathbf{w}}(\mathbf{s}) = \psi_{\mathbf{w}}^{\text{init}}(\mathbf{s})\psi_{\mathbf{w}}^{\text{SBS}}(\mathbf{s}), \quad (4.48)$$

the ratio of the wave function on two configurations as well as the log-derivatives depend only on the respective ratio and log-derivatives of each separate wave function, making the application of the VMC method straightforward. This procedure has the advantage of reducing the number of parameters necessary for obtaining a good approximation to the ground state and making the optimization procedure more stable, since the initial state is not a completely random state. A similar technique has been used to construct tensor-product projected states with tensor networks in [228]. More generally it can be used to project the wave function of an initial reference state in a Fock space and is thus also suitable to describe fermionic systems.

## 4.3 Application to chiral topological states

So far we have discussed the relations between neural-network quantum states in the form of restricted Boltzmann machines and tensor networks. We now turn to a practical application on a challenging problem for traditional tensor-network methods, namely the approximation of a state with chiral topological order. We have seen in the previous chapter that infinite dimensional MPS can be used to represent chiral topological states, but we would like to find a more general ansatz class which would also be able to approximate such states. While chiral topological PEPS have been constructed, the resulting states are critical[64–67]. Moreover the local parent Hamiltonian of a chiral fermionic Gaussian PEPS has to be gapless[65]. In this section we show that, because of their non-locality, this obstruction does not carry on to the tensor-network and neural-network states that we have introduced previously.

### 4.3.1 Restricted Boltzmann machines can describe a Laughlin state exactly

Let us consider the Laughlin state  $\psi_2^1$  defined in (3.9). The wave function can be rewritten, in terms of the spin-1/2 degrees of freedom, as

$$\psi_{\text{Laughlin}}(\mathbf{s}) = \delta_{\mathbf{s}} \prod_k \chi_k^{s_k} \chi \prod_{i < j} (z_i - z_j)^{\frac{1}{2} s_i s_j}, \quad (4.49)$$

where  $\delta_{\mathbf{s}}$  fixes the total spin to 0, the  $z_i$  are the complex coordinates of the positions of the lattice sites and the phase factor are defined as  $\chi_k^{s_k} = e^{i\pi(k-1)(s_k+1)/2}$ . We know from chapter 3 that this state is topological and admits an exact parent Hamiltonian given in (3.40).

The Laughlin wave function consists of a product of function of pairs of spins, so it has the structure of a Jastrow wave function and we have shown in section 4.2.1 that any Jastrow wave function can be written as a RBM with  $M = N(N - 1)/2$  hidden units. It follows that RBM and non-local SBS can represent a gapped chiral topological state exactly. This is in sharp contrast to local tensor-network states for which there is no exact description of a (non-critical) chiral topological state known. This difference is due to the non-local connections in the RBM and Jastrow wave function which allow them to easily describe a Laughlin state. We note that a chiral p-wave superconductor is another example of a gapped chiral topological state which has been recently written as a (fermionic) quasi-local Boltzmann machine[88].

The previous construction is however not satisfactory in the sense that the RBM requires a number of hidden units scaling as  $O(N^2)$  to represent the Laughlin state

exactly. Although the representation is sparse, if the structure of the state is not known in advance it will require  $O(N^3)$  parameters to find the Laughlin state, which is too high for numerical purposes on lattices which are not extremely small. We thus turn to the approximate representation of the Laughlin wave function using a RBM.

### 4.3.2 Numerical approximation of a Laughlin state

We consider a  $6 \times 6$  square lattice with open boundary conditions and the parent Hamiltonian (3.40) of the Laughlin state. We minimize the energy of different wave functions with respect to this Hamiltonian by applying the VMC method. Results are presented in Table 4.1.

It is found that EPS with plaquettes of size up to  $3 \times 3$  have an energy difference with the Laughlin state of the order  $10^{-2}$ , which is better than a short-range RBM (denoted sRBM) on  $3 \times 3$  plaquettes and up to  $M' = 4$  hidden units per plaquette, while the energy of a fully connected RBM with  $M = 2N$  hidden units is within  $10^{-5}$  of the energy of the ground state. The resulting RBM uses much less hidden units than would be required for it to be exact, yet reaches an overlap of 99.99% with the Laughlin wave function. Similar results have been obtained on a related Hamiltonian in [229]. This result shows that the fully-connected structure of the RBM is an advantage to describe this state and that EPS can be used instead of short-range RBM. We have moreover found that EPS are easier to optimize numerically than a short-range RBM: they are more stable, since each coefficient is considered separately, no exponentials or products that lead to unstable behavior are present and the derivatives have a very simple form.

Ansatz	$(E_w - E_0)/N$	$ \langle \psi_w   \psi_{\text{Laughlin}} \rangle $
EPS $2 \times 2$	$4.3 \times 10^{-2}$	46.10%
EPS $3 \times 3$	$2.2 \times 10^{-2}$	75.79%
sRBM $M' = 1$	$8.3 \times 10^{-2}$	0.01%
sRBM $M' = 2$	$3.1 \times 10^{-2}$	46.32%
sRBM $M' = 4$	$2.5 \times 10^{-2}$	59.07%
RBM $M = N$	$5.8 \times 10^{-4}$	99.7%
RBM $M = 2N$	$1.1 \times 10^{-5}$	99.99%

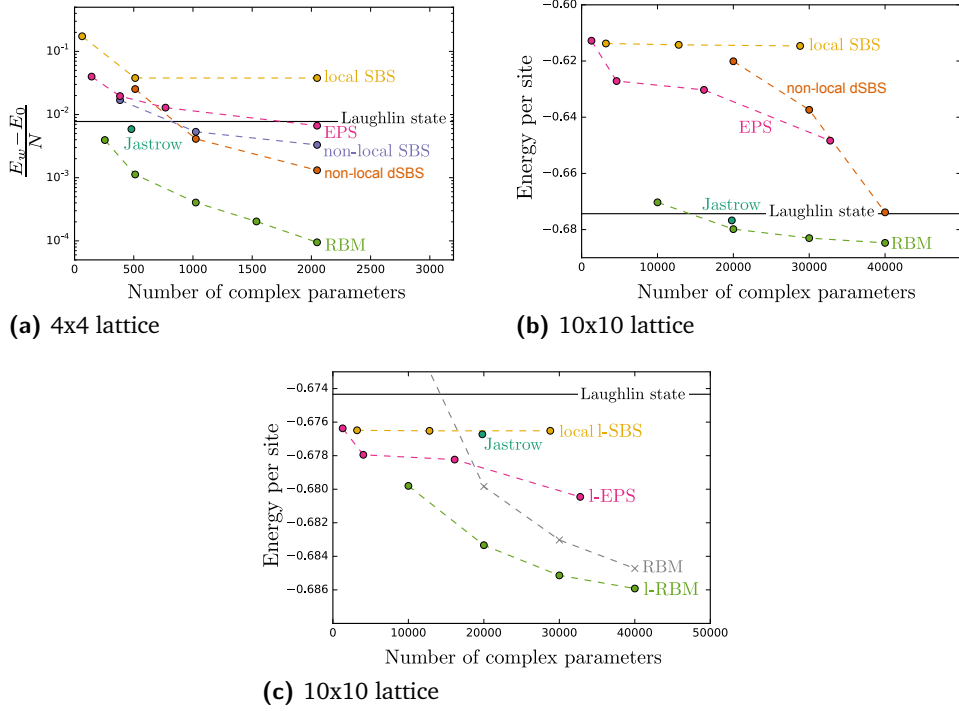
**Tab. 4.1.:** Energy per site difference with the ground state energy and overlap with the Laughlin state of different ansatz wave functions optimized with respect to the Hamiltonian  $H_{\text{parent}}$  on a  $6 \times 6$  square lattice with open boundary conditions. sRBM have  $M'$  hidden units connected to all spins in each plaquette of size  $3 \times 3$ , while RBM have  $M$  hidden units connected to all spins of the lattice.



### 4.3.3 Numerical approximation of a chiral spin liquid

The previous results indicate that RBM might be useful for approximating chiral topological states numerically, but the exact Laughlin state has a particularly simple structure and one may wonder whether the RBM would be able to find a state in the same phase which is not exactly a Laughlin state. We now consider the Hamiltonian  $H_l$ , obtained in [54] and introduced in (3.41). This Hamiltonian is defined on the square lattice, and its ground state for  $J = 1, J_\chi = 1$  has above 98% overlap with the Laughlin wave function on a  $4 \times 4$  lattice, but its exact ground state is different than the Laughlin state, although in the same phase. We minimize the energy of different classes of states with respect to this Hamiltonian on a  $4 \times 4$  and  $10 \times 10$  square lattices with open boundary conditions. We consider several ansatz wave functions including EPS with plaquettes of size  $2 \times 2, 3 \times 2, 4 \times 2$  and  $3 \times 3$ , local SBS covering the lattice with horizontal, vertical and diagonal strings and increasing bond dimension, RBM with increasing number of hidden units, non-local SBS with diagonal matrices (dSBS) or with non-commuting matrices of bond dimension 2 and different number of strings covering the full lattice. We observe that while the optimization of EPS and SBS is particularly stable, the optimization of RBM can lead to numerical instabilities that are resolved by writing the RBM in the form presented in (4.21). Since we use the same optimization procedure for all ansatz wave functions and since the required time and memory to perform the optimization is mainly a function of the number of parameters and of the accuracy, we can compare these classes by comparing how many parameters are needed to obtain a similar energy.

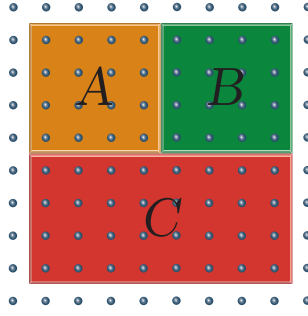
We first focus on a  $4 \times 4$  lattice for which the exact ground state can be obtained using exact diagonalization and the results are presented in Fig. 4.5a. Local SBS have an energy higher than the Laughlin state and the energy is saturated with increasing bond dimension, which means that the pattern of horizontal, vertical and diagonal strings is not enough to capture all correlations in the ground state. While a large  $4 \times 4$  plaquette would make EPS exact on this small lattice, this would require  $2^{16}$  parameters. The energy of the Laughlin state is already reached for  $3 \times 3$  plaquettes. RBM with a number of hidden units larger than  $N$  and non-local SBS with a corresponding number of strings have lower energy than the Laughlin state or the Jastrow wave function. When the number of strings grows, the energy decreases even further. On a larger  $10 \times 10$  lattice (Fig. 4.5b) the exact ground state energy is unknown but we can compare the energy of the different wave functions and observe similar results. Only the Jastrow wave function, non-local SBS and RBM have an energy comparable to the Laughlin state. Notice that non-local SBS have a constant factor more parameters than a RBM for the same number of strings. On the one side this allows SBS to achieve better energy than RBM with the same number of strings. On the other side this comes with the drawback that we can only optimize fewer strings and on the large lattice we are numerically limited to non-local dSBS with up to  $N$  strings. We can conclude that RBM are particularly



**Fig. 4.5.:** Energy of  $H_l$  per site for different optimized ansatz wave functions on a square lattice. The number of parameters ( $N_p$ ) is modified by increasing the bond dimension  $D$  (local SBS,  $N_p \propto D^2$ ), the size of the plaquettes (EPS,  $N_p \propto M_P 2^P$ , where  $M_P$  is the number of plaquettes and  $P$  is the number of spins in one plaquette), the number of strings  $M_S$  (non-local SBS and dSBS,  $N_p \propto M_S$ ) or the number of hidden units  $M_h$  (RBM,  $N_p \propto M_h$ ). (a) 4x4 lattice for which the energy difference with the exact ground state energy is plotted. (b) 10x10 lattice for which the exact ground state energy is unknown and the reference energy of the Laughlin state is indicated as a black line. (c) Optimization of wave functions that have been multiplied by the Laughlin wave function on a 10x10 lattice. The original RBM results are indicated for reference as grey crosses. Figure reproduced from [4].

efficient in this example since they require significantly less parameters than SBS for attaining the same energy. This has to be contrasted with the previous examples of the Majumdar-Gosh and AKLT models where the opposite was true. Therefore each class of states has advantages and drawbacks depending on the model we are looking at. We note in addition that a non-local SBS can be initialized with the results of a previous optimization with a RBM, which could provide a way of minimizing the difficulties of optimizing large number of parameters.

As we have previously noticed, we can also use an initial approximation of the ground state in combination with the previous ansatz classes. In the case of the Hamiltonian  $H_l$ , the analytical Laughlin wave function can be used as our initial approximation in (4.48). We denote l-EPS (resp. l-SBS, l-RBM) a wave function that consists in a product of the Laughlin wave function and an EPS (resp. SBS,



**Fig. 4.6.:** Partition of the lattice used to compute the topological entanglement entropy. Figure reproduced from [4].

RBM) and minimize the energy of the resulting states. This allows us to obtain lower energies for each ansatz class (Fig. 4.5c). Once the wave functions are optimized, their properties can be computed using Monte Carlo sampling. To check that the ground state is indeed in the same class as the Laughlin state, we compute the topological entropy of some of the optimized states by dividing the lattice into four regions (Fig. 4.6) and computing the Renyi entropy (2.12) of each subregion using the method presented in section 2.2.3. The topological entanglement entropy can then be computed as[40, 41]

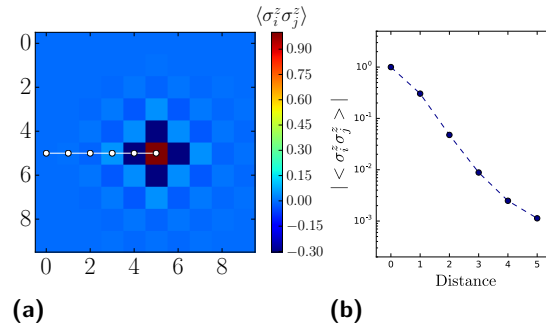
$$S_{\text{topo}} = S_A^{(2)} + S_B^{(2)} + S_C^{(2)} - S_{AB}^{(2)} - S_{AC}^{(2)} - S_{BC}^{(2)} + S_{ABC}^{(2)}, \quad (4.50)$$

and is expected to be equal to  $-\ln 2 \approx -0.347$  for the Laughlin state[146].

Ansatz	TEE
Laughlin	-0.339(3)
l-EPS $3 \times 3$	-0.36(1)
RBM $M = 4N$	-0.34(1)
l-RBM $M = 4N$	-0.34(1)

**Tab. 4.2.:** Topological entanglement entropy (TEE) of the analytical Laughlin state and optimized l-EPS, RBM and l-RBM.

The results we obtain are presented in Table 4.2 and provide additional evidence that the ground state of  $H_l$  has the same topological properties as the Laughlin state. The Hamiltonian  $H_l$  was recently investigated on an infinite lattice using infinite-PEPS[230] and further evidence was provided that the ground state is chiral. The PEPS results suggest the presence of long-range algebraically decaying correlations that may be a feature of the model or a restriction of PEPS to study chiral systems. The correlations on short distances agree with the correlations that we can compute on our finite system (Fig. 4.7a) but our method does not allow us to make claims about the long-distance behavior of the correlation function. In [225] it was observed that the entanglement entropy of some specific short-range RBM can be computed



**Fig. 4.7.:** (a) The spin-spin correlation function between one lattice site (in red) and all other spins on the lattice measured on the optimized l-RBM with lowest energy reveals the antiferromagnetic behavior of the correlations. (b) Decay of the correlations with the distance across the direction indicated in (a) as a white solid line. The error bars are within dot size and finite size effects can already be seen for the last point. Figure reproduced from [4].

analytically from the weights of the RBM. The method we use here works in the general case and also for a fully-connected RBM, but requires Monte Carlo sampling of the wave function. The optimized RBM weights encode every information about the wave function, it would thus be interesting to understand more precisely which quantities can be extracted directly from them. Whether direct information about the phase of the system can be obtained in this way without requiring Monte Carlo sampling remains an interesting open problem for future work.

## 4.4 Conclusion

In this chapter we have introduced the Variational Monte Carlo method as an algorithm to approximate a ground state wave function with both neural-network quantum states and tensor-network states. We have then shown that these classes of states are in fact intimately related. In particular we proved that short-range restricted Boltzmann machines are a subclass of Entangled Plaquette States, while fully connected restricted Boltzmann machines are a subclass of String-Bond States. This connection allows us to understand the underlying structure of restricted Boltzmann machines and to generalize them on larger local Hilbert spaces and with additional geometric flexibility using non-local String-Bond States. We compared the advantages and drawbacks of these different classes of states and suggested a way to combine them together.

We then applied these methods to the problem of describing states with chiral topological order introduced in the previous chapter. We showed that every Jastrow wave function, and thus a Laughlin wave function, can be written as an exact

restricted Boltzmann machine. By performing variational optimization on the parent Hamiltonian of this state, we provided evidence that already a small number of hidden units in the Boltzmann machine is enough to obtain a good approximation of the Laughlin state. Finally we turned to the approximation of the ground state of a chiral spin liquid that is in the same phase as the Laughlin state, but for which the exact ground state is unknown. We showed that restricted Boltzmann machines and non-local String-Bond States are able to achieve lower energy than the approximate Laughlin state, and that they allow us to characterize the ground state and compute its topological entanglement entropy.



## Machine learning with generalized tensor networks

In the previous chapter we have used functions introduced in machine learning to represent wave functions of quantum many body systems. In this chapter we will take the opposite perspective and use functions and algorithms introduced in quantum many-body physics to perform machine learning.

Tensor networks such as Matrix Product States can be used not only for representing wave functions of quantum many-body systems, but also tensors that appear in machine learning[231, 232]. They have been used for example to compress weights of neural networks[233–236], to study model expressivity[237–241] or to parametrize complex dependencies between variables[103–105, 242–244].

In this chapter we show that the connection introduced in the previous chapter between restricted Boltzmann machines and tensor networks generalizes to more general probabilistic graphical models and motivates the definition of generalized tensor networks, which are still efficient to contract and can be used to perform machine learning in higher dimensions where Matrix Product States are not efficient.

We start in section 5.1 by providing an introduction to machine learning with a focus on supervised and unsupervised learning algorithms. In section 5.2 we then review the framework of probabilistic graphical models, and how Boltzmann machines appear in this framework. Section 5.3 discusses the connections between probabilistic graphical models and tensor networks, and introduces generalized tensor networks in which local tensor information can be copied and reused in different parts of the network. We provide a proof, relying on the entanglement properties of tensor networks, that generalized tensor networks are exponentially more efficient at approximating some functions than regular tensor networks. In section 5.4 we then provide an algorithm for training such generalized tensor networks in a supervised learning context. Since we want to apply our algorithm not only to discrete data, but also to continuous inputs, we discuss several ways to use tensor networks with real data and introduce methods to combine tensor networks and neural networks in section 5.5. Finally we benchmark our algorithm for different network architectures on the tasks of classifying images and recognizing

environmental sounds in section 5.6. This chapter is based on Ref. [6] and is a modification thereof:

- [6] : Ivan Glasser, Nicola Pancotti and J. Ignacio Cirac, ‘Supervised learning with generalized tensor networks’, arXiv:1806.05964 (2018).

## 5.1 Introduction to machine learning

In this section we introduce simple notions and notations appearing in the context of supervised and unsupervised learning[84].

Consider a dataset  $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}$ , where each  $\mathbf{x}_i$  is the vector corresponding to the  $i$ -th training example in some input space  $X$  and  $y_i$  is its label in an output space  $Y$ . A supervised learning algorithm learns a function  $g : X \rightarrow Y$  which maps an input vector to its label. If the label takes discrete values, this corresponds to a problem of classification, while if it is a continuous variable then the task becomes a regression. If the input space  $X$  is large, it is not possible to evaluate all possible functions. Consider for examples images with  $N$  black and white pixels, then there are  $2^N$  possible input images and it is not possible to parametrize a function by its values over each image as it would require too much memory. This curse of dimensionality, that we have already encountered for a quantum many-body wave function, leads to a choice of parametrized functions  $f$  in an hypothesis space of smaller size. This space can be for example the space of linear functions, or the space of neural-network functions with some architecture. In order to find a function in this space that approximates the mapping between input vectors and labels we define a loss function  $L : Y, Y \rightarrow \mathbb{R}_{\geq 0}$ , such as the square error  $L(y, y') = (y - y')^2$ , and minimize a cost function estimated from the training data

$$\mathcal{C} = \frac{1}{|\mathcal{D}|} \sum_{i=1}^{|\mathcal{D}|} L(y_i, f(x_i)), \quad (5.1)$$

with respect to the parameters in the function  $f$ , for example using a steepest descent algorithm. This procedure may lead to overfitting: if the function memorizes the mapping  $\mathbf{x} \rightarrow y$  on the training examples, it will not necessarily generalize well to unseen examples. The performance is therefore evaluated on a test set of examples that are not used during training. The structure of the function  $f$  and its complexity can lead to overfitting, and one way to prevent it is to add a regularization term in the cost function such as a penalty on the norm of the parameters in  $f$ .

In unsupervised learning, we are presented with a dataset  $\mathcal{D} = \{(\mathbf{x}_i)\}$ , where no labels are present. We assume the  $\mathbf{x}_i$  to be independent and identically distributed. The task is then to learn some structure present in the data, for example by finding



clusters of related data points. A central problem in unsupervised learning is the problem of density estimation, in which one constructs an estimate  $p(\mathbf{x})$  of the unobservable underlying probability density function of the data. One method to perform such a task is maximum likelihood estimation. The estimate of the probability distribution is chosen in a parametrized hypothesis space and one finds the value of the parameters which maximizes the likelihood given the values taken by the data, or more often the (average) log-likelihood:

$$\mathcal{L} = \frac{1}{|\mathcal{D}|} \sum_{i=1}^{|\mathcal{D}|} \log p(\mathbf{x}_i). \quad (5.2)$$

Such a model, once obtained, can be used for example to generate new samples following the same distribution.

Generative training can be used to perform supervised learning in a classification setting. In this case one learns the joint probability distribution  $p(\mathbf{x}, y)$  as in the unsupervised case, by maximizing the log-likelihood

$$\mathcal{L} = \frac{1}{|\mathcal{D}|} \sum_{i=1}^{|\mathcal{D}|} \log p(\mathbf{x}_i, y_i). \quad (5.3)$$

Once the model has been trained, one can compute the conditional probability

$$p(y|\mathbf{x}) = \frac{p(\mathbf{x}, y)}{\sum_{y_j} p(\mathbf{x}, y_j)}, \quad (5.4)$$

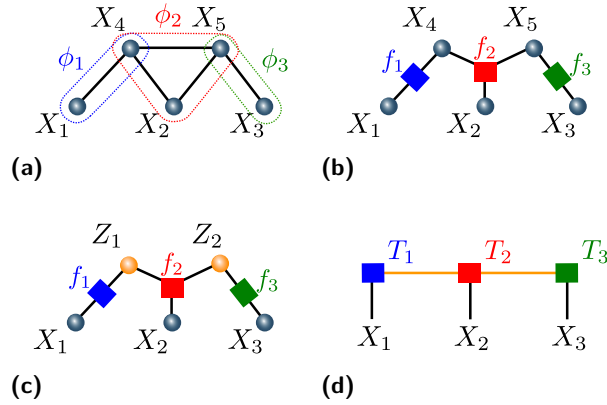
and the label predicted by the model for a feature vector  $\mathbf{x}_i$  is  $f(\mathbf{x}_i) = \arg \max_y p(y|\mathbf{x}_i)$ . On the other hand in discriminative training one directly optimizes the conditional probability  $p(y_i|\mathbf{x}_i)$ , since one is ultimately interested in classification performance, which can be done by maximizing

$$\mathcal{L}_{\text{discriminative}} = \frac{1}{|\mathcal{D}|} \sum_{i=1}^{|\mathcal{D}|} \log p(y_i|\mathbf{x}_i), \quad (5.5)$$

This shows that functions designed for representing probability distributions can in principle be used naturally in both supervised and unsupervised learning.

## 5.2 Probabilistic graphical models

Let us consider the case where the data examples are samples from a set of discrete random variables  $\mathbf{X} = \{X_1, \dots, X_N\}$ . A common choice of parametrized models for  $p$  are graphical models[84], which correspond to a factorization of the probability distribution over a graph. Consider a graph  $G = (V, E)$ . Here  $V$  is a set of vertices



**Fig. 5.1.:** (a) Undirected graphical model with three maximal cliques depicted in colors (b) Corresponding factor graph (c) Factor graph with hidden units in orange that are marginalized (d) Equivalent tensor network, which is a Matrix Product State. Figure reproduced from [6].

corresponding to the random variables,  $E$  a set of edges between these vertices and  $\text{cl}(G)$  is the set of maximal cliques of the graph. An undirected graphical model or a Markov random field defines a factorization of the joint probability of all random variables as

$$p(\mathbf{x}) = \frac{1}{Z} \prod_{C \in \text{cl}(G)} \phi_C(\mathbf{x}_C), \quad (5.6)$$

where  $\mathbf{x}_C$  are the values of the random variables in clique  $C$ ,  $\phi_C$  are the clique potentials which are positive functions and  $Z$  is a constant that ensures normalization of the probability (Fig. 5.1a). Graphical models can be converted to factor graphs[245] defined on a bipartite graph of factors and variable vertices: one factor node  $f_s$  is created for each maximal clique and the factor is connected to the variables in the corresponding clique (Fig. 5.1b). The factorization of the probability distribution still reads

$$p(\mathbf{X} = \mathbf{x}) = \frac{1}{Z} \prod_s f_s(\mathbf{x}_c), \quad (5.7)$$

and inference can be performed through belief propagation and the sum-product algorithm on factor graphs. To increase the set of distributions which can be represented we can add additional dependencies by introducing ancillary hidden variables (which are unobserved, i.e. their values are not supplied in the data)  $\mathbf{Z} = \{Z_1, \dots, Z_M\}$  (Fig. 5.1c). The resulting probability distribution is obtained by marginalizing these hidden variables:

$$p(\mathbf{X} = \mathbf{x}) = \frac{1}{Z} \sum_{\mathbf{z}} \prod_s f_s(\mathbf{x}_c, \mathbf{z}_c). \quad (5.8)$$

An example of graphical model that we have encountered in the previous chapter are Boltzmann machines. Boltzmann machines are graphical models for which the cliques potentials take a particular form. Because the potentials are positive, we can write them as the exponential of some real function. In a Boltzmann machine this function is taken to be a classical Ising Hamiltonian with two-body Ising interaction between all pairs of variables in a clique. For a restricted Boltzmann machines (RBM) [246, 247], which is defined on a bipartite graph with visible variables  $\mathbf{X}$  and hidden variables  $\mathbf{Z}$  (Fig. 5.3a), the resulting probability distribution once the hidden variables have been marginalized is

$$P(\mathbf{x}) = \frac{1}{Z} \sum_{\mathbf{z}} e^{\mathcal{H}(\mathbf{x}, \mathbf{z})}, \quad (5.9)$$

$$= \frac{1}{Z} \prod_i (1 + e^{\sum_j w_{ij} x_j}). \quad (5.10)$$

where the Hamiltonian  $\mathcal{H}$  is a classical Ising Hamiltonian defined as (we omit here the bias terms for simplicity)

$$\mathcal{H} = \sum_{i,j} w_{ij} z_i x_j.$$

Note that the formula for the RBM is different than equation (4.19) obtained in the previous chapter, since here we assume the variables to take values 0 and 1, instead of  $-1$  and  $1$ . Because the normalization  $Z$  is unknown, computing the likelihood (5.2) for a restricted Boltzmann machine cannot be done exactly. Approximate algorithms relying on Monte Carlo (Gibbs) sampling can nevertheless be used, such as contrastive divergence [85, 247]. Discriminative training can on the other hand be performed exactly, since the normalizations cancel when computing the conditional probabilities (5.4).

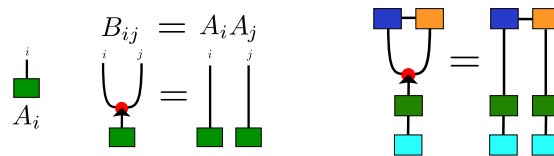
## 5.3 Generalized tensor networks

In the case where all variables are discrete, the functions  $f_s(\mathbf{x}_c, \mathbf{z}_c)$  appearing in (5.8) are tensors indexed by the values of the variables  $\mathbf{x}_c$  and  $\mathbf{z}_c$ . Since we sum over the values of the variables  $\mathbf{z}_c$ , this correspond to a contraction of the corresponding indexes of the tensors. The probability distribution defined by a probabilistic graphical models can therefore be written as a tensor network (or in some cases hypernetwork). This connection, depicted in Fig. 5.1c and 5.1d for a MPS, has been previously observed in Refs. [4, 89, 248, 249] in particular models, and we refer to Ref. [250] for a more detailed analysis of this duality.

There is still an important difference between the frameworks of tensor networks and graphical models: tensors of a factor graph coming from a graphical model have non-

negative elements, while tensor networks are usually studied in the context of real (or complex) elements. This has important consequences for the optimization algorithms that can be used with them. Graphical models can be used in conjunction with expectation-maximization algorithms, which rely on the computation of conditional probabilities of some of the variables. Tensor networks instead are not interpreted probabilistically, but their optimization algorithms can rely on the singular-value decomposition of matrices.

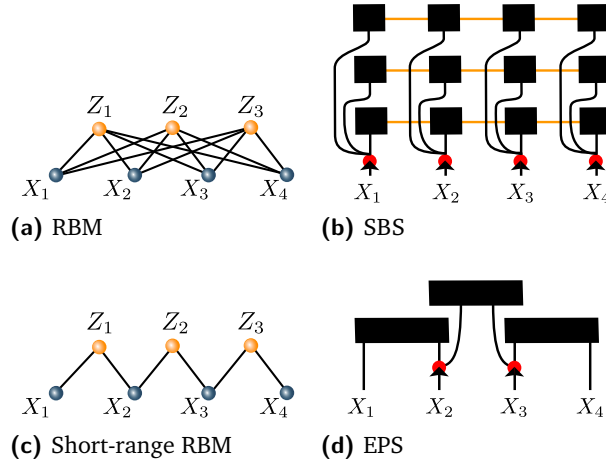
It is interesting to note that there are simple classes of graphical models which do not share the properties of tensor networks that all visible variables are only connected to one factor. This is the case for example in a restricted Boltzmann machine, where each visible variable is connected to all hidden variables. Despite the graph having loops,  $p(\mathbf{x})$  or  $p(\mathbf{h}|\mathbf{x})$  can be computed analytically for arbitrary sizes. This relies on the fact that once we fix the value of the visible variables, the contraction of the network from bottom to top becomes efficient.



**Fig. 5.2.:** Copy operation of a vector input  $A_i$ , resulting in a new tensor  $B_{ij} = A_i A_j$ , or of a tensor network. Figure reproduced from [6].

This leads us to the definition of a new class of tensor networks, that we call generalized tensor networks, where the value of tensors can be copied and reused in different parts of the network. Defined on an appropriate graph, the tensor network can still be contracted if the value of the visible variables are fixed. We graphically depict this through a red dot between edges of the graph and an arrow which marks the incoming edge. This copy operation copies vector inputs, resulting in two copies of the original vector (Fig. 5.2). More generally one can apply this copy operation to a larger tensor network, which is then copied. In practice one would first contract the incoming network, resulting in a vector. This vector is then copied and sent to the remaining edges. The rest of the tensor network can then be independently contracted. We impose that there are no directed loops containing dots and ensure that the tensor network can be efficiently contracted as long as it is contracted in the right order, which amounts to choosing a tree structure of the graph obtained after copying all parts of the network. Contrary to regular tree tensor networks, this network includes several tensors which share the same values. In general, it is not possible to write this dot as a tensor[251]. In particular cases, when all the inputs are discrete and in a fixed basis, it is possible to write the copy operation as a COPY-dot tensor, as introduced in [252]. For discrete inputs and when the copy operation only applies to the inputs, the generalized tensor network can therefore be written as a tensor network including COPY-dot tensors. In the more general case,

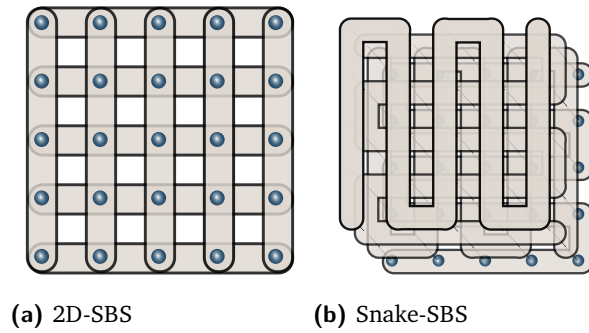
while it is not possible to represent the copy operation as a tensor, duplicating a vector and sending it to two different parts of the calculation is easily achieved in practice.



**Fig. 5.3.:** (a) Restricted Boltzmann Machine (RBM) consisting of visible and hidden variables (b) String-Bond State with 1D geometry generalizing RBM. The legs corresponding to contracted indices in each MPS are depicted in orange for visibility. (c) Short-range RBM with local connections between visible and hidden variables (d) Entangled Plaquette State (EPS) generalizing the short-range RBM. Figure reproduced from [6].

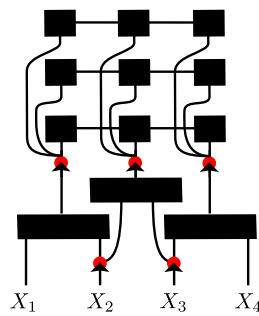
Examples of such tensor networks with copy of the input states have been used in machine learning in the form of tree tensor networks with copy of the inputs[239, 251]. They have also been used in quantum physics, as we have seen in chapter 4. Entangled Plaquette States (EPS), defined in (4.13), are tensor networks where the inputs are copied and fed into overlapping plaquettes (Fig. 5.3d). EPS also share some similarity with Convolutional Neural Networks (CNN): a convolutional layer with discrete inputs is a particular case of EPS with weight sharing between the tensors. The EPS realizes all possible convolutional filters over a discrete input space, instead of selecting a small number of filters as CNN do. String-Bond States, defined in (4.16) are another example in which the inputs are copied to different MPS (Fig. 5.3b). As we have proven in section 4.2, EPS generalize short-range RBM and SBS generalize RBM. They provide a way to interpolate between a MPS (large bond dimension, only one string) and a RBM (bond dimension 2, diagonal matrices, many strings). In this chapter we will use the following string geometries: horizontal and vertical strings covering the 2D lattice (Fig. 5.4a) will be denoted as 2D-SBS. We will also also consider the choice of 4 strings, each covering the whole lattice in a snake pattern, but in a different order (Fig. 5.4b). We denote these SBS as snake-SBS.

More generally, one can think of complex networks built using the copy operation for gluing different networks together. As example we will later consider the case



**Fig. 5.4.:** Possible geometries of SBS: (a) 2D-SBS consisting of horizontal and vertical overlapping strings. (b) Snake-SBS consisting of 4 overlapping strings in a snake pattern. Figure reproduced from [6].

of an EPS, whose output are copied and taken as input into a SBS (Fig. 5.5). The input variables are first copied and fed into overlapping clusters parametrized by tensors. The output leg of each of these tensors is a vector which is copied a few times. Each of these copies can then be contracted with the open legs of a different MPS, forming together a SBS. In 2D, we choose  $2 \times 2$  overlapping plaquettes in the first layer, and 4 strings forming a snake-SBS in the second layer. In the following we will call this generalized tensor network EPS-SBS, but we observe that more complex networks based on trees or hierarchical designs with more than two layers can also be constructed in the same way.



**Fig. 5.5.:** EPS-SBS consisting of a first layer of EPS, followed by a copy operation and a second layer of SBS. Figure reproduced from [6].

These generalized tensor networks have the advantage, compared to standard tensor networks, that they can be easily defined in arbitrary dimension and geometry while remaining efficient to contract, as long as the input is fixed. This is in contrast to Projected Entangled Pair States (Fig. 1.2b), which are naturally defined in higher dimensions but cannot be contracted exactly efficiently. Moreover, the reuse of input information in generalized tensor network is more similar to state-of-the-art CNN, and weights can also be shared between different tensors.

Another advantage of generalized tensor networks is that they can represent some functions with exponentially fewer parameters than regular tensor networks. Indeed, regular tensor network functions satisfy the area law, so they cannot represent functions with volume law entanglement entropy.

One such function, defined for binary variables on an open chain  $x_1, \dots, x_N$  with  $N$  even, is

$$f(\mathbf{x}) = \prod_{l=1}^{N/2} g(x_l, x_{N-l+1}), \quad (5.11)$$

where  $g(x_i, x_j) = \frac{1}{\sqrt{2}}(-1)^{x_i \text{ XOR } x_j}$ . This function is the wave function of a quantum state in which there are singlets between site  $l$  and site  $N - l + 1$ . The entanglement entropy of the first half of the chain is therefore  $\frac{N}{2} \log(2)$ , and a MPS representing  $f$  will need a bond dimension exponential in  $N$ .

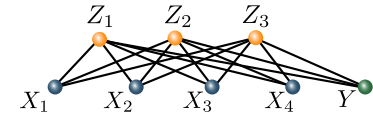
A SBS defined on the same chain, with each string corresponding to the whole chain, can represent this function by having  $N/2$  strings with bond dimension 2, since each string can capture the function  $g$  with a MPS of bond dimension 2. The function  $f$  can then be represented as a SBS with a polynomial number of parameters.

## 5.4 Supervised learning algorithm

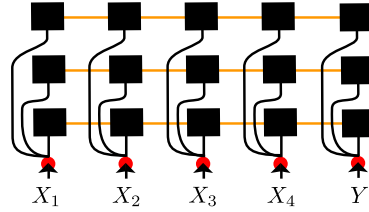
Graphical models are designed to represent probability distributions, so are mostly used in the context of unsupervised learning. For a graphical model or a tensor network on a tree, it is possible to compute the normalization  $Z$ , which gives exact access to the likelihood (5.2). MPS can be used for unsupervised learning in this way [242]. In the more general case, which includes RBM, the normalization  $Z$  cannot be computed efficiently and Monte Carlo sampling has to be used. Generalized tensor networks suffer from the same issue, which makes unsupervised learning computationally expensive. Since these networks correspond to quantum states, it might be possible to implement them on a quantum computer and sample from them efficiently. In the rest of this chapter we focus instead on supervised learning, where access to the normalization  $Z$  is not necessary.

As explained in section 5.1, an ansatz designed for a probability distribution can be naturally used in a discriminative setting by maximizing  $\mathcal{L}_{\text{discriminative}}$  (5.5). We therefore approximate the joint probability distribution of the variables and labels as a generalized tensor network (GTN):

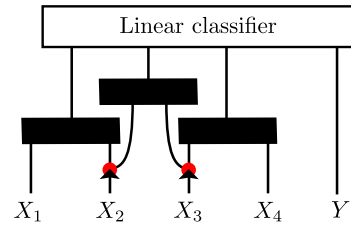
$$p(\mathbf{x}, y) \propto \text{GTN}(\mathbf{x}, y). \quad (5.12)$$



(a) Discriminative RBM



(b) Discriminative SBS



(c) Discriminative EPS

**Fig. 5.6.:** (a) A classification RBM turns the label into an additional visible unit. (b) The same procedure can be defined for a SBS by adding a node corresponding to the label, and corresponding tensors which connect it to the rest of the tensor network. (c) Generalized tensor networks can be combined with additional layers of neural networks. For example an EPS output is a tensor that can be combined with a linear classifier. Figure reproduced from [6].

The label is now seen as an additional input which corresponds to the index of one tensor (Fig. 5.6a). For generalized tensor networks, once the network inputs are fixed, the network factorizes in several tensor networks in the last layer. Each of these tensor networks can have a tensor indexed by the label (Fig. 5.6b). In order for  $\mathcal{L}_{\text{discriminative}}$  to be well defined,  $p$  should be positive. This can be done by ensuring that the tensor elements are positive, as in a RBM or graphical model, or by choosing instead

$$p(\mathbf{x}, y) \propto (\text{GTN}(\mathbf{x}, y))^2, \quad (5.13)$$

or

$$p(\mathbf{x}, y) \propto e^{\text{GTN}(\mathbf{x}, y)}. \quad (5.14)$$

In the following we will adopt this last choice, for which training is more efficient.



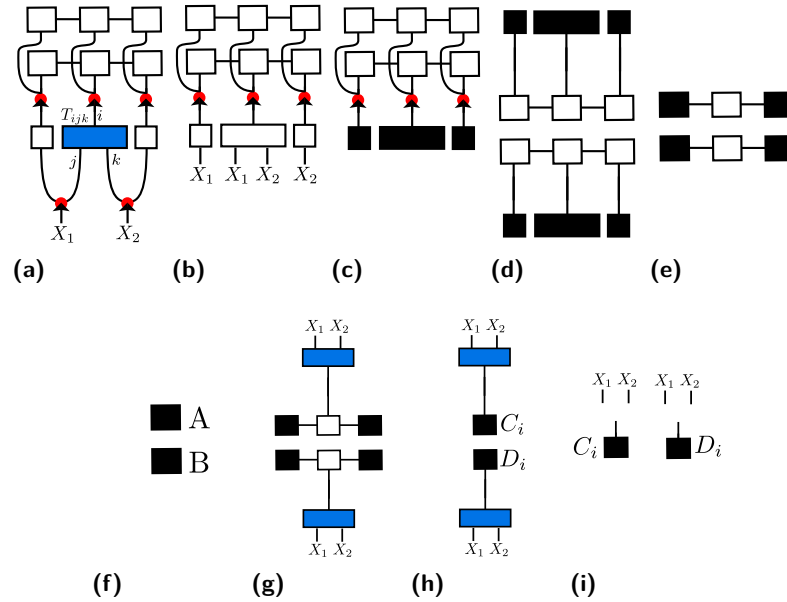
We then define

$$\text{GTN}(y_k|\mathbf{x}_i) = \frac{\text{GTN}(\mathbf{x}_i, y_k)}{\sum_{y_j} \text{GTN}(\mathbf{x}_i, y_j)}. \quad (5.15)$$

The maximization of  $\mathcal{L}_{\text{discriminative}}$  can be performed by observing that one can compute its gradient, which depends on

$$\frac{\partial \log \text{GTN}(y_i|\mathbf{x}_i)}{\partial w} = \frac{\partial \log \text{GTN}(\mathbf{x}_i, y_i)}{\partial w} - \sum_{y_j} \text{GTN}(y_j|\mathbf{x}_i) \frac{\partial \log \text{GTN}(\mathbf{x}_i, y_j)}{\partial w}. \quad (5.16)$$

$\text{GTN}(\mathbf{x}_i, y_j)$  can be computed exactly by fixing the value of the input units and labels and contracting the network. The derivatives with respect to the parameters in each tensor can be computed by observing that the copy operations give rise to a tree network in which the derivative with respect to a tensor is obtained by contracting the rest of the network with the corresponding tensor removed, just as in a standard tensor network (Fig. 5.7a). From the point of view of supervised learning there is



**Fig. 5.7.:** (a)-(f) Forward pass of contracting an EPS-SBS generalized tensor network. We denote scalar and vectors that are the result of a tensor contraction as black boxes, while other tensors are denoted as empty boxes. The result is a scalar  $A \times B$ . (g)-(i) We compute as an example the derivative with respect to the tensor  $T_{ijk}$  denoted in blue in (a). The computation can start from stage (d), obtained during the forward propagation. Remaining additional tensors are contracted until we are left with the log-derivative, equal to  $\partial \log \text{GTN}(X_1, X_2) / \partial T_{ijk} = \delta_{X_1=j} \delta_{X_2=k} (C_i + D_i) / (AB)$ . Figure reproduced from [6].

no difference between SBS and MPS in terms of the optimization algorithms: the cost of optimizing a SBS is only a constant factor (the number of strings) more than that of optimizing a MPS, and this procedure can be straightforwardly parallelized.

This is unlike in quantum physics where Monte Carlo sampling is necessary to optimize a SBS. To further regularize the tensor network, we adopt the procedure suggested in Ref. [105] to randomly drop tensor elements to 0 with probability  $\delta$  during training.

The tensor networks we have constructed so far have no open legs when an input and a label are given. We can also construct networks with open legs and use tensor networks in combination with other machine learning techniques. In this case the tensor network maps the input to a tensor, which can for example be used as input in a neural network. In the case of EPS where each tensor over overlapping plaquettes has an open leg, an input is mapped to a tensor with an extra dimension as output. This is similar to the role of convolutional filters in a CNN. The simplest way to combine EPS with other neural networks is to place a linear classifier on top of the EPS (Fig. 5.6c). The backpropagation algorithm used to compute derivatives of the neural network is in this case combined with the algorithm for computing derivatives of a tensor network, and the joint network can be optimized using stochastic gradient descent.

## 5.5 Learning feature vectors of data

In the previous discussion we have always considered discrete input data. In practice, one may want to apply these techniques to real data. In this section we explore several strategies that can be used for this purpose. We suggest to learn relevant tensor features as part of the tensor network and discuss how tensor features can also be learned as part of a deep learning architectures which combines a neural network extracting features with a tensor network.

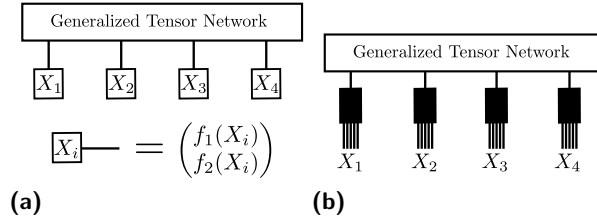
A naive way of applying tensor networks with real data would be to discretize data or use its binary representation. This is not a suitable approach, because that would amount to increasing the size of the data a lot, rendering learning very slow, and would also lead to big tensor networks which would be prone to overfitting. Another approach, as suggested in [104], is to map the real data to a higher dimensional feature space. Each variable is first independently mapped to a vector of length at least two and these vectors are then contracted with the open legs of the tensor network (Fig.5.8a). Choices of feature maps that have been used in [104, 105, 243, 244] include

$$x \rightarrow \begin{pmatrix} 1 \\ x \end{pmatrix} \text{ or } \begin{pmatrix} \cos(\frac{\pi}{2}x) \\ \sin(\frac{\pi}{2}x) \end{pmatrix}, \quad (5.17)$$

and generalizations to higher dimensions. A choice which is suitable with our algorithm, assuming that the data is normalized between 0 and 1, is to use

$$x \rightarrow \begin{pmatrix} \cos^2(\frac{\pi}{2}x) \\ \sin^2(\frac{\pi}{2}x) \end{pmatrix}, \quad (5.18)$$

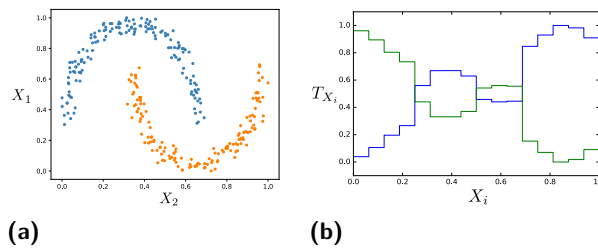
because this ensures that the vectors are positive and the normalization prevents numerical instabilities.



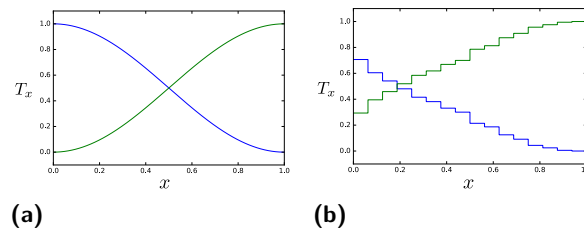
**Fig. 5.8.:** (a) Real inputs  $X_i$  are mapped to a feature vector (here with length two). This vector can then be used as input to a generalized tensor network by contracting it with the open legs of the generalized tensor network. (b) Feature tensors can compress the discretized representation of the inputs  $X_i$  to a smaller dimensional space. These tensors can share weights and can be learned as part of the tensor network. Figure reproduced from [6].

These choices however limit the functions that can be learned. Consider for example the dataset with just two variables presented in Fig. 5.9a. It cannot be separated by a MPS of bond dimension 2 with one of these feature choices, since the boundary decision will be a polynomial of degree two of the features. A different feature choice could distinguish the two classes, even with bond dimension 2. We therefore suggest to learn the appropriate features as part of the learning algorithm. This can be done by parametrizing the feature functions and learning them at the same time as the rest of the network. A choice of parametrization can be itself a tensor network. In the simplest case, we discretize the real data and use a tensor to compress the large dimensional input into a smaller dimensional vector of suitable length. This tensor can be learned as part of the whole tensor network and prevents the size of the rest of the tensor network to increase when the discretization size changes. The feature tensor can be the same for all variables, for example image pixels, but can be different in the case where the variables are of different nature. Using this procedure, a MPS of bond dimension 2 is able to get perfect accuracy on the dataset presented in Fig. 5.9a. The two features that the network has learned are presented in Fig. 5.9b. We note that starting from random features on more complex datasets makes learning difficult, but the feature tensor can be pretrained using a linear classifier, before being trained with the rest of the network.

In comparison, we also show in Fig. 5.10b the features learned while classifying MNIST with greyscale pixels and a snake-SBS (see section 5.6). These features are not very different from the choice in (5.18) (Fig. 5.10a), and we could not



**Fig. 5.9.:** (a) Dataset with two features  $X_1$  and  $X_2$  and two classes (depicted in different colors) that cannot be learned by a MPS of bond dimension 2 with features in (5.18). (b) Two normalized features learned by a tensor while classifying the previous dataset with a MPS of bond dimension 2. The features have been discretized in 16 intervals. Using this choice of features the MPS can classify the dataset perfectly. Figure reproduced from [6].

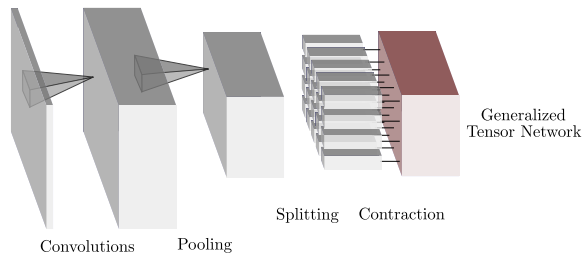


**Fig. 5.10.:** (a) Choice of two features in (5.18) for an input taking real values between 0 and 1. (b) Two normalized features learned by a tensor with output dimension 2 combined with a snake-SBS classifying the MNIST dataset. The input features  $x$  are the greyscale value of pixels, normalized between 0 and 1 and discretized in 16 intervals. Figure reproduced from [6].

distinguish performance with this choice or with learned features on this dataset. Moreover, the size of the feature vector provides a regularization of the model, and higher sizes might be necessary for more complex datasets. More generally this tensor could be itself represented with a small tensor network, to prevent the number of parameters to increase too much with a very small discretization interval. It is interesting to note that the features learned in our examples are almost continuous even if we use smaller discretization intervals. This means that two real inputs that are close to each other will lead to the same predictions by the network. Our approach of learning the features as part of the tensor network may be especially relevant in the context of quantum machine learning, where the tensor network is replaced by a quantum circuit and it might be suitable to have the full network as part of the same quantum machine learning architecture.

As an alternative way of choosing the features, we can combine the feature choice with other machine learning techniques. If the input data represents images, it is a natural choice to use Convolutional Neural Networks as feature extractors, since these have been highly successful in image classification. CNN consist in convolution

filters, which use convolutional kernels to transform an image into a set of filtered images, and pooling layers which downsize the images (Fig. 5.11). The different filters can be seen as different features of the corresponding pixel or region of the image and preserve locality. Therefore it is natural to consider the vector of applied filters associated with each location in the image as a feature vector that can be used in conjunction with generalized tensor networks. The CNN and the tensor network can be trained together, since the derivatives of the tensor network can be used in the backpropagation algorithm which computes the gradient of the cost function.



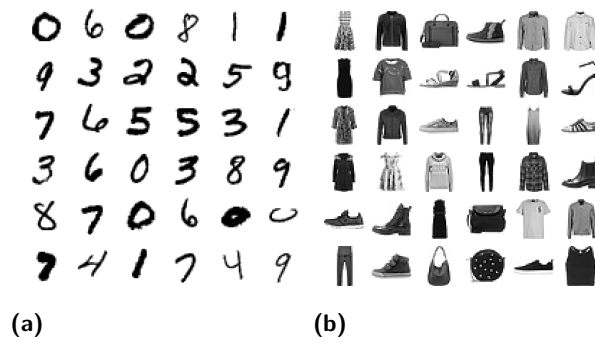
**Fig. 5.11.:** Using convolutional Neural Networks as feature vector extractors from real data: the output of the CNN is seen as an image with a third dimension collecting the different features. For each pixel of this image, the vector of features is contracted with the open legs of a tensor network. Figure reproduced from [6].

## 5.6 Application to image and sound classification

In this section we test the generalized tensor network approach on the task of image classification, where a natural two-dimensional geometry is present, as well as on the task of urban sound recognition, where the time dimension provides a one-dimensional geometry.

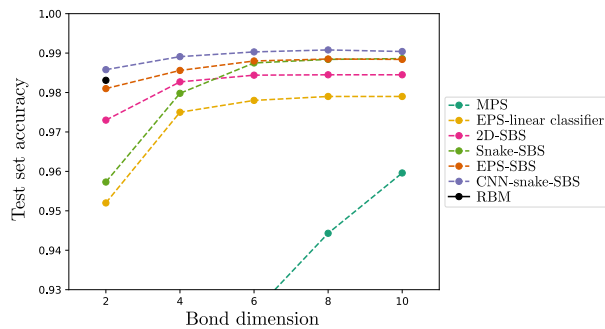
### 5.6.1 Image classification

We first consider the MNIST dataset[253], which consists of  $28 \times 28$  greyscale images of digits. There are 10 classes and we adopt a multiclass classification procedure in which one tensor of the tensor network is parametrized by the ten possible labels. The original training set is split into training and validation sets of 55000 and 5000 examples and the performance of the different models is evaluated on the test set of 10000 examples. We consider the following generalized tensor networks: a snake-SBS with 4 strings (Fig. 5.4b), a 2D-SBS (Fig. 5.4a), an EPS with a  $2 \times 2$  translational-invariant plaquette combined with a linear classifier (Fig. 5.6c), an EPS-SBS with translational-invariant plaquette combined with a snake-SBS (Fig. 5.5) and a CNN-snake-SBS which uses a 1-layer CNN as input features (Fig. 5.11). The CNN considered here uses a convolutional layer applying  $6 \times 5 \times 5$  filters (stride 1) with ReLU activation function and a pooling layer performing max pooling with a



**Fig. 5.12.:** Examples of images from the MNIST (a) and fashion MNIST (b) dataset. Figure reproduced from [6].

$2 \times 2$  filter. All other networks use the choice of features presented in (5.18) and the greyscale values are normalized between 0 and 1. We compare the performance of these networks with a MPS and a RBM. Hyperparameters such as the learning rate, the regularization rate and the number of iterations over the training set are determined through a grid search while evaluating the performance on the validation set.



**Fig. 5.13.:** Test set accuracy of different generalized tensor networks on the MNIST dataset. Figure reproduced from [6].

The test set accuracy, presented in Fig. 5.13, shows that even with a very small bond dimension generalized tensor network are able to accurately classify the dataset. Their performance is significantly better than that of a tree tensor network[244] or a MPS trained in frequency space[243], and while a MPS can also achieve 99.03% accuracy with a bond dimension of 120 [104], the cost of optimizing very large tensors has prohibited the use of this method for larger problems so far. The snake-SBS with bond dimension larger than 6 has also better performance than a RBM. Since the snake-SBS provides an interpolation between RBM and MPS, the choice of number of strings and geometry can be seen as additional parameters which could be tuned further to improve over the performance of both methods. All networks have a training set accuracy very close to 100% when the bond dimension is larger than 6, and we expect that better regularization techniques or network architectures

have to be developed to significantly increase the test set performances obtained here. We also optimized a snake-SBS with positive elements (by parametrizing each element in a tensor as the exponential of the new parameters), which is a graphical model. Using the same algorithm, we were not able to achieve better performance than 93% classification accuracy with bond dimensions up to 10. This shows that while having a structure closely related to graphical models, tensor networks may provide different advantages.

Method	Accuracy
Support Vector Machine	84.1%
EPS + linear classifier	86.3%
Multilayer perceptron	87.7%
EPS-SBS	88.6%
Snake-SBS	89.2%
AlexNet	89.9%
CNN-snake-SBS	92.3%
GoogLeNet	93.7%

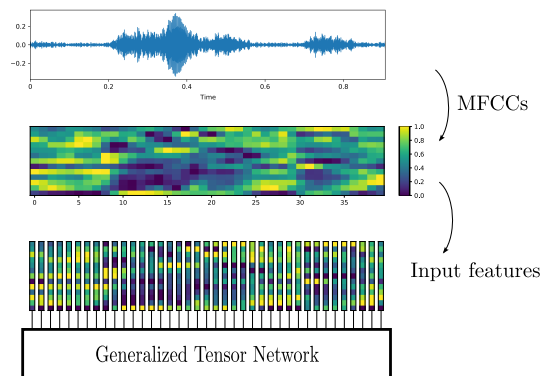
**Tab. 5.1.:** Test set accuracy of generalized tensor networks and other approaches[254] on the fashion MNIST dataset.

We then turn to the fashion MNIST dataset[254], consisting of  $28 \times 28$  greyscale images of clothes. While having the same size as the original MNIST dataset, it is significantly harder to classify. We report the best accuracy obtained with different generalized tensor networks with bond dimension up to 10 in Table 5.1. It is found that these networks are competitive with other approaches such as Support Vector Machines, AlexNet and GoogLeNet Convolutional Neural Networks or a multilayer perceptron neural network.

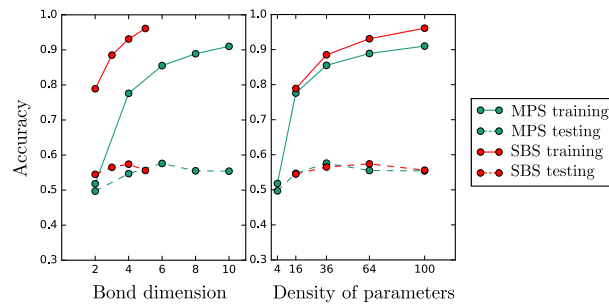
## 5.6.2 Environmental sound classification

So far we have considered greyscale images, but it is also interesting to study how generalized tensor networks work with other types of data. In the following we consider the task of classifying environmental sounds. The UrbanSound8K dataset[255] is a collection of 8732 audio clips (4s or less) divided into 10 classes of urban sounds: air conditioner, car horn, children playing, dog barking, drilling, engine idling, gun shot, jackhammer, siren and street music. The dataset is divided into 10 folds and we use folds 1-9 for training and fold 10 for testing. The one-dimensional structure of sounds allows us to compare MPS and SBS with the same 1D string geometry. Preprocessing of the data takes place as follows: clips shorter than 4s are repeated to reach a fixed length of 4s. The first 13 Mel-frequency cepstral coefficients (MFCCs) are extracted for each clip (sampled at 22050Hz) using a window size of 2048 and hop length of 512, resulting in a sequence of length 173 and dimension 13 (Fig. 5.14). The corresponding 13-dimensional vectors are used as

input feature vectors for the tensor network, and the time dimension of the sequence corresponds to the 1-dimensional structure of the MPS, or the strings of the SBS. Note that we do not perform any data augmentation nor split the training examples to enlarge the size of the dataset set, since we are interested in comparing MPS and SBS, rather than achieving the best possible accuracy on this dataset. The training and testing accuracies are reported in Fig. 5.15 for a MPS with bond dimension up to 10 and a SBS with 4 strings and bond dimension up to 5. Since we are interested in comparing the expressivity of the different networks, no regularization is used and training is performed until the training accuracy does not improve anymore. Note that a MPS with bond dimension  $D$  has as many variational parameters as a SBS with 4 strings and bond dimension  $D/2$ .



**Fig. 5.14.:** From the raw audio signal, Mel-frequency cepstral coefficients (MFCCs) are extracted over short overlapping windows, resulting in a sequence of high dimensional vectors. These vectors are taken as input to a generalized tensor network. Figure reproduced from [6].



**Fig. 5.15.:** Training and testing accuracy of a MPS and a SBS with 4 strings on the Urban-Sound8K dataset. The density of parameters is the total number of parameters divided by 174 (the length of the strings). Figure reproduced from [6].

We observe that the SBS has slightly higher training accuracy than a MPS with larger bond dimension and the same number of parameters. The test set performance is not significantly different between the different architectures and in both cases we find that the networks overfit, which is not surprising given the small number of training examples. Higher accuracies have been reported with other methods on the same dataset. For example Convolutional Neural Networks can reach above 70% test set accuracy[256], but use much more input features and rely on data augmentation.



Nevertheless our results show that SBS should also be considered along with MPS when considering one-dimensional data, and may be applied in other settings such as natural language processing[257, 258].

## 5.7 Conclusion

In this chapter we have shown that tensor networks defined to represent wave functions of quantum many-body systems can also be used in the context of machine learning. We have discussed the relationship between tensor networks and probabilistic graphical models, and shown that it motivated the definition of generalized tensor networks, which enlarge the class of tensor networks by introducing a reuse of information taking the form of a copy operation of tensor elements. We focused on particular architectures of generalized tensor networks, such as Entangled Plaquette States and String-Bond States, and discussed their relationship to other networks studied in machine learning. We then provided an algorithm to train these models to perform supervised learning and discussed how they can be used not only with discrete input, but also with real-valued data. We demonstrated that generalized tensor networks that can be contracted exactly can perform accurate image classification with much smaller bond dimension than regular tensor networks, that they can be used in other settings such as sound recognition and that they can be combined with neural-network architectures.

Tensor networks can also be seen as a tool to simulate quantum circuits. The functions we introduced can be implemented on a quantum circuit by using several copies of the input features, and thus show that such quantum circuits are able to perform accurate image classification. Moreover the results we have obtained show that finding suitable regularizations for quantum circuits will be important in order to achieve better accuracies. Generalized tensor networks which originate from the classical simulation of quantum states may thus have applications in classical machine learning, as well as serve as a testing and benchmarking platform of near-term quantum machine learning architectures.



## Conclusion and outlook

Throughout this thesis we studied the application of novel tensor-network methods to challenging problems in both quantum many-body physics and machine learning. In these two areas, the curse of dimensionality prevents an exact treatment from being possible on even moderate system sizes, and tensor networks provide powerful function approximators with desirable properties.

The first part of this thesis contributed to our understanding of infinite dimensional tensor networks, for which the wave function takes the form of a correlator of field operators. In chapter 2 we reviewed the construction of critical models in one dimension arising from conformal fields, and showed that infinite dimensional matrix product states are not limited to conformal fields but can also be defined using operators coming from a massive field theory. This allowed us to build a toy model of a phase transition for which the wave function has an analytical expression. The properties of the system could then be studied on large systems using Monte Carlo methods. Our construction was extended to a mixed state ansatz, for which the finite temperature properties were characterized. We demonstrated numerically that this model density operator approximates the thermal state of a Hamiltonian with nearest neighbour interactions, thus demonstrating that the analytical model we constructed is physically relevant. Our construction is quite general and could be performed using other field theories. Since we used the second dimension of the field theory to introduce a temperature, it remains an interesting problem whether similar techniques could be used to study phase transitions and their finite temperature effects in two dimensions.

In chapter 3 we explored the properties of infinite dimensional tensor networks taking the form of a correlator of conformal fields in two dimensions. These states correspond to lattice versions of fractional quantum Hall wave functions. We investigated the properties of these states on different lattices and along an interpolation between the continuum and the lattice limit, with a focus on Laughlin and Moore-Read wave functions. We showed that in most (but not all) cases, the states have the same topological properties as the continuum wave functions. We constructed lattice wave functions for the localized quasiholes and quasielectrons of these states and showed that their braiding properties, size and charge could be computed. Unlike in the continuum, the quasielectrons on the lattice can be constructed as inverse quasiholes, and in some models there is an exact symmetry between quasiholes and quasielectrons. We derived exact parent Hamiltonians for which the ground states

are lattice Laughlin and Moore-Read states at different filling fractions, as well as for which the ground states contain quasiparticles. These Hamiltonians have long-range interactions but allow for the manipulation of quasiparticles simply by changing the coupling strengths of these interactions. In a few cases, we were able to obtain local Hamiltonians for which the ground states are close to the analytical wave functions we introduced. This was possible by truncating the non-local parent Hamiltonians and numerically optimizing the resulting Hamiltonian to keep the same ground state. This procedure is however limited to small systems, and it would be highly desirable to find a systematic way of obtaining local Hamiltonians for the states we introduce. An experimental realization of such Hamiltonians, whether in natural materials or in quantum simulators, would provide new platforms for studying topological systems and using them in quantum computing.

The local Hamiltonians realizing lattice fractional quantum Hall effects do not admit an exact description of the ground state, and other methods are needed to study them on large system sizes. We showed in chapter 4 that this could be done by minimizing the energy of other variational wave functions generalizing tensor networks in higher dimension, as well as using neural-network quantum states for which the wave function takes the form of a restricted Boltzmann machine. We demonstrated that these classes of states are related, and in particular that restricted Boltzmann machines are a subclass of String-Bond States. This connection sheds light on the underlying architecture and geometry of restricted Boltzmann machines and allowed us to generalize them to larger local Hilbert spaces using non-local String-Bond States. We showed that lattice Laughlin states can be written exactly as a restricted Boltzmann machine, and compared the ability of different classes of states to represent a Laughlin wave function by variationally optimizing the energy of these states with respect to the parent Hamiltonian of the Laughlin state. Finally we applied these techniques to the local Hamiltonian realizing a related chiral spin liquid, and showed that they allowed us to characterize the ground state of this model. Whether similar techniques can be used to systematically approximate ground states of more complicated topological systems remains an interesting future research direction.

This connection between restricted Boltzmann machines and tensor networks reveals a bridge between classical approaches to the quantum many-body problem and machine learning techniques in high-dimensional settings. In chapter 5 we discussed these connections in the more general context of probabilistic graphical models. We showed that these connections motivate the definition of generalized tensor networks, which include String-Bond States and Entangled Plaquette States and can be used both in quantum physics as well as in machine learning. These networks can moreover be combined with traditional neural-network architectures. We provided an algorithm to train these models to perform supervised learning, which does not require Monte Carlo sampling. We studied how these models can be used with real-

valued input data, and showed that they could perform accurate image classification and sound recognition with much smaller cost than regular tensor networks. The connections we discussed might have an impact on both classical as well as quantum machine learning. On the one hand tensor-network algorithms can be used in conjunction with classical machine learning techniques. On the other hand these tensor networks correspond to quantum circuits that can be simulated classically, and may serve as a benchmarking platform of near-term quantum machine learning architectures.

Machine learning is emerging as a tool that can be used to help physicist understand and manipulate quantum systems. As such, it might be used to build quantum computers, for example through the design of novel quantum error correcting codes[76]. The building blocks of such a computer may be topological systems that can be used to perform topological quantum computation[43, 44], and such a computer might then perform quantum machine learning algorithms[95, 96] or quantum simulations of condensed matter systems[259]. Future research directions in quantum computing, topological matter and classical and quantum machine learning are therefore naturally interconnected, and it is of no surprise that tensor networks play a role in all these areas of science.



## Parent Hamiltonians for lattice Laughlin states

The procedure for constructing parent Hamiltonians is as follows: first a null field  $\chi(w)$  is defined, such that the correlator used to define the wave function with this null field inserted is zero:

$$\langle \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_{i-1}}(z_{i-1}) \chi(z_i) \mathcal{V}_{n_{i+1}}(z_{i+1}) \dots \mathcal{V}_{n_N}(z_N) \rangle = 0. \quad (\text{A.1})$$

This expression can then be transformed into operators annihilating the wave function

$$\Lambda_i |\psi\rangle = 0. \quad (\text{A.2})$$

One can then define a positive semi-definite operator annihilating the original wave function as

$$H = \sum_i \Lambda_i^\dagger \Lambda_i. \quad (\text{A.3})$$

In the following sections we derive the operators  $\Lambda_i$  for different wave functions. This appendix reproduces derivations first presented in [3].

### A.1 Operators annihilating the $\psi_q^1[p]_\infty$ wave functions

In this section we obtain operators annihilating the  $\psi_q^1[p]_\infty$  wave functions by extending the procedure of [173] to  $p > -q$ ,  $p$  integer. It was shown in Ref. [160] that

$$\chi(w) = \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} [G^+(z) \mathcal{V}_-(w) - qJ(z) \mathcal{V}_+(w)], \quad (\text{A.4})$$

is a null field, where  $G^+(z) =: e^{i\sqrt{q}\phi(z)}$ ,  $J(z) = i\partial_z \phi / \sqrt{q}$ ,  $\mathcal{V}_-(z) =: e^{-i\frac{1}{\sqrt{q}}\phi(z)}$  and  $\mathcal{V}_+(z) =: e^{i\frac{q-1}{\sqrt{q}}\phi(z)}$ . The correlator with this field inserted vanishes :

$$\langle W_w(\infty) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_{i-1}}(z_{i-1}) \chi(z_i) \mathcal{V}_{n_{i+1}}(z_{i+1}) \dots \mathcal{V}_{n_N}(z_N) \rangle = 0. \quad (\text{A.5})$$

This expression is the sum of two terms,

$$\oint_{z_i} \frac{dz}{2\pi i} \frac{1}{z - z_i} \langle W_p(w) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_{i-1}}(z_{i-1}) G^+(z) \mathcal{V}_-(z_i) \mathcal{V}_{n_{i+1}}(z_{i+1}) \dots \mathcal{V}_{n_N}(z_N) \rangle, \quad (\text{A.6})$$

and

$$-q \oint_{z_i} \frac{dz}{2\pi i} \frac{1}{z - z_i} \langle W_p(w) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_{i-1}}(z_{i-1}) J(z) \mathcal{V}_+(z_i) \mathcal{V}_{n_{i+1}}(z_{i+1}) \dots \mathcal{V}_{n_N}(z_N) \rangle. \quad (\text{A.7})$$

The second term, multiplied by  $(-1)^{i-1} |n_1, \dots, n_{i-1}, 0, n_{i+1}, \dots, n_N\rangle$  and after summation of all  $n_k$ , gives [173]

$$- \left( \sum_{j(\neq i)} \frac{1}{z_i - z_j} d_i (q n_j - 1) + \frac{p}{z_i - w} d_i \right) |\psi_q^1[p(w)]\rangle, \quad (\text{A.8})$$

where  $d_j$  is the hardcore boson (resp. fermion) annihilation operator for  $q$  even (resp. odd), and  $n_j = d_j^\dagger d_j$  is the number of particles at site  $j$ . The first term instead becomes, after deforming the contour integral around each  $z_j$ ,

$$\begin{aligned} & - \oint_w \frac{dz}{2\pi i} \frac{1}{z - z_i} \langle W_p(w) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_{i-1}}(z_{i-1}) G^+(z) \mathcal{V}_-(z_i) \mathcal{V}_{n_{i+1}}(z_{i+1}) \dots \mathcal{V}_{n_N}(z_N) \rangle \\ & - \sum_{j(\neq i)} \oint_{z_j} \frac{dz}{2\pi i} \frac{1}{z - z_i} \langle W_p(w) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_{i-1}}(z_{i-1}) G^+(z) \\ & \quad \times \mathcal{V}_-(z_i) \mathcal{V}_{n_{i+1}}(z_{i+1}) \dots \mathcal{V}_{n_N}(z_N) \rangle. \end{aligned} \quad (\text{A.9})$$

The sum in the second term has been computed in Ref. [173] and gives, after multiplication by  $(-1)^{i-1} |n_1, \dots, 0, \dots, n_N\rangle$ ,

$$\sum_{j(\neq i)} \frac{1}{z_i - z_j} d_j |\psi_q^1[p(w)]\rangle. \quad (\text{A.10})$$

The last remaining term is

$$-(-1)^{i-1+p} \sum_{n_i} \delta_{n_i=0} \oint_w \frac{dz}{2\pi i} \frac{(-1)^q \sum_{k=1}^{i-1} n_k}{z - z_i} \langle G^+(z) W_p(w) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_N}(z_N) \rangle, \quad (\text{A.11})$$

where we have commuted  $G^+(z)$  in front of the correlator. Since we know the expression of the correlator, we can compute the contour integral. The expression then simplifies to

$$-(-1)^{i-1+p} \delta_{p<0} \sum_{n_i} n_i \frac{1}{(-p-1)!} (w - z_i)^{-p} \prod_{k(\neq i)} (z_i - z_k)^{-(q n_k - 1)}$$



$$\times \lim_{z \rightarrow w} \frac{d^{-p-1}}{dz^{-p-1}} \prod_j (z - z_j)^{qn_j - 1 - (q+1)\delta_{ij}} \langle W_p(w) \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_N}(z_N) \rangle. \quad (\text{A.12})$$

Now observe that this expression is zero when  $p > 0$  (due to the first delta factor), but also when  $p > -q$  (due to the derivative and the exponent of the polynomial). This shows that when  $p > -q$ , this term does not contribute and the resulting expression is given by summing Eq. (A.8) and Eq. (A.10), which leads to

$$\Lambda_i |\psi_q^1[p(w)]\rangle = 0, \quad (\text{A.13})$$

where

$$\Lambda_i = \sum_{j(\neq i)} \frac{1}{z_i - z_j} [d_j - d_i(qn_j - 1)] - \frac{p}{z_i - w} d_i. \quad (\text{A.14})$$

## A.2 Operators annihilating the $\psi_4^2$ wave function

In this section we obtain operators annihilating the  $\psi_4^2$  wave function. Using the fields defined in section 3.2.5, we have for  $q = 4$  and  $\eta = 2$  the following operator product expansions :

$$\begin{aligned} G^+(z) \mathcal{V}_-(w) &= \frac{1}{(z-w)^2} e^{i(2\phi(z) - \phi(w))}, \\ &\sim \frac{1}{(z-w)^2} e^{i\phi(w)} + \frac{1}{(z-w)} 2i\partial\phi(w) e^{i\phi(w)} \\ &\sim \frac{1}{(z-w)^2} \mathcal{V}_+(w) + \frac{1}{(z-w)} 2\partial\mathcal{V}_+(w), \end{aligned} \quad (\text{A.15})$$

so that

$$G^+(z) \mathcal{V}_{n_j=0}(w) \sim \frac{1}{(z-w)^2} \mathcal{V}_{n_j=1}(w) + \frac{1}{(z-w)} 2\partial\mathcal{V}_{n_j=1}(w), \quad (\text{A.16})$$

$$G^+(z) \mathcal{V}_{n_j=1}(w) \sim 0. \quad (\text{A.17})$$

The field

$$\chi_2(w) = \oint_w \frac{dz}{2\pi i} G^+(z) \mathcal{V}_+(w) \quad (\text{A.18})$$

is therefore a null field. We can now use the fact that the correlator with the field inserted vanishes :

$$\langle \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n_{i-1}}(z_{i-1}) \chi(z_i) \mathcal{V}_{n_{i+1}}(z_{i+1}) \dots \mathcal{V}_{n_N}(z_N) \rangle = 0. \quad (\text{A.19})$$

This can be rewritten as

$$\begin{aligned}
& \oint_{z_i} \frac{dz}{2\pi i} \langle \mathcal{V}_{n_1}(z_1) \dots G^+(z) \mathcal{V}_+(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle \\
&= \oint_{\infty} \frac{dz}{2\pi i} \langle \mathcal{V}_{n_1}(z_1) \dots G^+(z) \mathcal{V}_+(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle \\
&\quad - \sum_{j(\neq i)} \oint_{z_j} \frac{dz}{2\pi i} \langle \mathcal{V}_{n_1}(z_1) \dots G^+(z) \mathcal{V}_+(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle, \tag{A.20}
\end{aligned}$$

where we have deformed the contour integral around the positions  $z_j$ . The contour integral at infinity can be computed by evaluating the correlator, which leads to a zero contribution. The remaining part can be transformed by observing that the operators  $G^+(z)$  and  $\mathcal{V}_{\pm}(z_j)$  commute, so that

$$\begin{aligned}
& - \sum_{j(\neq i)} \oint_{z_j} \frac{dz}{2\pi i} \langle \mathcal{V}_{n_1}(z_1) \dots G^+(z) \mathcal{V}_+(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle \\
&= - \sum_{j(\neq i)} \oint_{z_j} \frac{dz}{2\pi i} \langle \mathcal{V}_{n_1}(z_1) \dots G^+(z) \mathcal{V}_{n_j}(z_j) \dots \mathcal{V}_+(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle \\
&= - \sum_{j(\neq i)} \oint_{z_j} \frac{dz}{2\pi i} \frac{\sum_{n'_j} d_{n_j n'_j}}{(z - z_j)^2} \langle \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n'_j}(z_j) \dots \mathcal{V}_+(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle \\
&\quad - \sum_{j(\neq i)} \oint_{z_j} \frac{dz}{2\pi i} 2 \frac{\sum_{n'_j} d_{n_j n'_j}}{z - z_j} \langle \mathcal{V}_{n_1}(z_1) \dots \partial \mathcal{V}_{n'_j}(z_j) \dots \mathcal{V}_+(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle, \tag{A.21}
\end{aligned}$$

where we have moved the operator  $G^+(z)$  to position  $j$  and applied the operator product expansion in Eq. (A.16).  $d_{n_j n'_j}$  are the matrix elements of the bosonic annihilation operator  $d_j$  at site  $j$  ( $d_{00} = d_{11} = d_{10} = 0$ ,  $d_{01} = 1$ ), so that only terms having a non-zero operator product expansion contribute to this expression. The contour integral evaluated in the first term gives zero, so that we are left with

$$\begin{aligned}
& - 2 \sum_{j(\neq i)} \sum_{n'_j} d_{n_j n'_j} \langle \mathcal{V}_{n_1}(z_1) \dots \partial \mathcal{V}_{n'_j}(z_j) \dots \mathcal{V}_+(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle \\
&= -2 \sum_{j(\neq i)} \sum_{n'_j} d_{n_j n'_j} \frac{\partial}{\partial z_j} \langle \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_{n'_j}(z_j) \dots \mathcal{V}_+(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle \\
&= -2 \sum_{j(\neq i)} \sum_{n'_j} d_{n_j n'_j} \frac{\partial}{\partial z_j} \psi_4^2(n_1, \dots, n'_j, \dots, n_i = 1, \dots, n_N), \tag{A.22}
\end{aligned}$$

where we have used the expression of the wave function  $\psi_4^2$  as a correlator of fields. Let us now compute the derivative of the wave function :

$$\begin{aligned}
& \frac{\partial}{\partial z_j} \psi_4^2(n_1, \dots, n'_j, \dots, n_i = 1, \dots, n_N) \\
&= \psi(n_1, \dots, n_N) \frac{\partial}{\partial z_j} \ln(\psi_4^2(n_1, \dots, n_N))
\end{aligned}$$

$$\begin{aligned}
&= \psi(n_1, \dots, n_N) \frac{\partial}{\partial z_j} \left[ \sum_{k(<j)} s_k s'_j \ln(z_k - z_j) + \sum_{k(>j)} s_k s'_j \ln(z_j - z_k) \right] \\
&= \psi(n_1, \dots, n_N) \left[ \sum_{k(\neq j)} s_k s'_j \frac{-1}{z_k - z_j} \right], \tag{A.23}
\end{aligned}$$

where  $s_k = 2n_k - 1$ ,  $s'_j = 2n'_j - 1$  are the corresponding spin-1/2 degree of freedom at site  $k$  and  $j$ . In the rest of this section we will use both the notations  $n_k$  and  $s_k$  for brevity. Note that  $k$  can be equal to  $i$  in this sum, in which case  $s_i = 1$  since  $n_i = 1$ . The previous expression becomes

$$2 \sum_{j(\neq i)} \sum_{n'_j} d_{n_j n'_j} \sum_{k(\neq j)} s_k s'_j \frac{1}{z_k - z_j} \psi_4^2(n_1, \dots, n'_j, \dots, n_i = 1, \dots, n_N). \tag{A.24}$$

Since this expression is zero unless  $n_j = 0$  and  $n'_j = 1$ , we can replace  $s'_j$  by 1. To take into account the fact that this expression evaluates to 0 when  $n_i = 0$ , we introduce the number operator  $n_i$  and write the expression as

$$2 \sum_{j(\neq i)} \sum_{n'_j} d_{n_j n'_j} n_i \sum_{k(\neq j)} s_k \frac{1}{z_k - z_j} \psi_4^2(n_1, \dots, n'_j, \dots, n_i, \dots, n_N), \tag{A.25}$$

which when multiplied by the basis elements  $|n_1, \dots, n_N\rangle$  and summed over all  $n_k$  leads to

$$\sum_{j(\neq i)} \sum_{k(\neq j)} \frac{1}{z_k - z_j} n_i d_j s_k |\psi_4^2\rangle. \tag{A.26}$$

We started with the fact that this expression was zero, so the operator

$$\Omega_i = \sum_{j(\neq i)} \sum_{k(\neq j)} \frac{1}{z_k - z_j} n_i d_j s_k \tag{A.27}$$

annihilates the wave function  $\psi_4^2$ .

### A.3 Conformal transformations of the parent Hamiltonians

In this section we discuss the properties of the parent Hamiltonians obtained in this work under conformal transformations of the coordinates  $z_i$ . We will use the notations :

$$\lambda_{ij} = \frac{1}{z_i - z_j}, \quad \kappa_{ij} = \frac{z_i z_j}{z_i - z_j}, \quad \rho_{ij} = \frac{z_i + z_j}{z_i - z_j}. \tag{A.28}$$

If an operator  $\Gamma_i$  has the form  $\Gamma_i = \sum_{j(\neq i)} \lambda_{ij} f_{ij}$ , where  $f_{ij}$  does not depend on the coordinates  $z_i$ , then it is transformed under a Möbius transformation  $M$  as

$$\Gamma_i \rightarrow \sum_{j(\neq i)} \left( d^2 \lambda_{ij} + c^2 \kappa_{ij} + cd \rho_{ij} \right) f_{ij}. \quad (\text{A.29})$$

If  $\Gamma_i$  annihilates a wave function that is invariant under  $M$ , then  $\sum_{j(\neq i)} \kappa_{ij} f_{ij}$  annihilates the same wave function (choice  $c = 1, d = 0$  in the previous equation). Since  $\rho_{ij} = z_i \lambda_{ij} + \frac{1}{z_i} \kappa_{ij}$ , then  $\sum_{j(\neq i)} \rho_{ij} f_{ij}$  also annihilates the wave function. Noting that the composition of Möbius transformations is a Möbius transformation, the space of operators annihilating the wave function that can be obtained by applying Möbius transformations and multiplying by constant factors the operator  $\Gamma_i$  is therefore the space of operators

$$\Gamma_i^{\alpha\beta\gamma} = \sum_{j(\neq i)} (\alpha \lambda_{ij} + \beta \kappa_{ij} + \gamma \rho_{ij}) f_{ij}, \quad (\text{A.30})$$

where  $\alpha, \beta$  and  $\gamma$  are complex numbers. An operator in this space is invariant under  $M$  when the following conditions are satisfied :

$$\begin{aligned} \alpha &= d^2 \alpha + b^2 \beta + 2bd\gamma, \\ \beta &= c^2 \alpha + a^2 \beta + 2ac\gamma, \\ \gamma &= cd\alpha + ab\beta + (ad + bc)\gamma. \end{aligned} \quad (\text{A.31})$$

Note that since the Hamiltonians we construct have the form  $\sum_i \Gamma_i^\dagger \Gamma_i$ , these conditions only need to be satisfied up to a phase for the Hamiltonian to be invariant under  $M$ . A particular case is the translation  $z \rightarrow z + 1$  ( $a = 1, b = 1, c = 0, d = 1$ ), which leads to the conditions  $\beta = 0, \gamma = 0$ . Another particular case are the rotations along the periodic direction of the cylinder  $z \rightarrow e^{2\pi i/N_y} z$ , ( $a = e^{2\pi i/N_y}, b = 0, c = 0, d = 1$ ), which lead to the conditions  $\alpha = 0, \beta = 0$ . Therefore  $\Gamma_i$  cannot be invariant under translations and rotations at the same time (unless  $\Gamma_i = 0$ ), hence it cannot be invariant (up to a phase) under all conformal transformations. We note moreover that for the operators  $f_{ij}$  that we have used before (in particular the case  $f_{ij} = d_j - d_i(qn_j - 1)$ ) no further simplification in  $H = \sum_i \Gamma_i^\dagger \Gamma_i$  appears, so that in the space of Hamiltonians we constructed there is no non-zero Hamiltonian that would be invariant under all conformal transformations.

However the choice  $\alpha = 0, \beta = 0$  leads to an operator that is invariant under rotations along the periodic direction of the cylinder. In addition the operator  $\Gamma_i = \sum_{j(\neq i)} \rho_{ij} f_{ij}$  is then invariant, up to a phase, under a transformation  $z \rightarrow 1/z$ . This means that the Hamiltonians of the form  $\sum_i \sum_{k(\neq i)} \sum_{j(\neq i)} \rho_{kj}^* \rho_{ij} f_{kj}^\dagger f_{ij}$  are invariant under the symmetries of the lattice on the cylinder (Fig. 3.6).

## Parent Hamiltonians for lattice Moore-Read states

The procedure followed here is the same as in appendix A, but this time we use the operators used to construct the Moore-Read wave function (3.58). This appendix reproduces derivations first presented in [2].

### B.1 Operators annihilating the $(q = 1, \eta = 1)_1$ lattice Moore-Read state

Let us consider first the case  $q = 1, \eta = 1$  for a spin-1 model. This model arises from the spin 1 primary fields of the  $SU(2)_2$  Wess-Zumino-Witten conformal field theory [70, 207]. As shown in [70], we can parametrize a null field in this case by

$$\chi^a(z_i) = \left(K^{(i)}\right)_b^a (J_{-1}^b \varphi_n)(z_i), \quad (\text{B.1})$$

where repeated indices are summed over,  $\varphi_n$  is a chiral spin-1 primary field used to define the wave function,  $J_{-1}^b$  are the  $-1$  modes of the  $SU(2)_2$  current operators and

$$\left(K^{(i)}\right)_b^a = \frac{2}{3} \delta_{ab} - \frac{5}{12} i \epsilon_{abc} t_i^c - \frac{1}{12} (t_i^a t_i^b + t_i^b t_i^a), \quad (\text{B.2})$$

where  $\epsilon_{abc}$  is the Levi-Civita symbol and  $t_i^a$  are the spin-1 operators acting on site  $i$ . These operators can be written in the spin basis at site  $i$  as

$$t_i^x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad t_i^y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad t_i^z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (\text{B.3})$$

We then exploit that the correlator with a null field inserted is zero :

$$\langle \varphi_{n_1}(z_1) \dots \chi^a(z_i) \dots \varphi_{n_N}(z_N) \rangle = 0. \quad (\text{B.4})$$

We can transfer the action of a current operator to fields at other positions :

$$\langle \varphi_{n_1}(z_1) \dots (J_{-1}^b \varphi_{n_i})(z_i) \dots \varphi_{n_N}(z_N) \rangle =$$

$$\sum_{j(\neq i)}^N \frac{t_j^a}{z_1 - z_j} \langle \varphi_{n_1}(z_1) \dots \varphi_{n_i}(z_i) \dots \varphi_{n_N}(z_N) \rangle. \quad (\text{B.5})$$

Using (B.5) it is possible to rewrite (B.4) as [70]

$$\Lambda_i^a |\psi\rangle = 0, \quad (\text{B.6})$$

where

$$\Lambda_i^a = \sum_{j(\neq i)}^N \left( K^{(i)} \right)_b^a \omega_{ij} t_j^b, \quad (\text{B.7})$$

$$= \sum_{j(\neq i)}^N \omega_{ij} \left[ \frac{2}{3} t_j^a - \frac{5}{12} i \epsilon_{abc} t_j^b t_i^c - \frac{1}{12} (t_i^a t_i^b + t_i^b t_i^a) t_j^b \right], \quad (\text{B.8})$$

where we have defined  $\omega_{ij} = \frac{z_i + z_j}{z_i - z_j}$ . This can be used to construct a parent Hamiltonian  $H = \sum_{a,i} \Lambda_i^{a\dagger} \Lambda_i^a$ , which gives

$$\begin{aligned} H &= \frac{4}{3} \sum_{i \neq j}^N \omega_{ij}^* \omega_{ij} + \frac{1}{3} \sum_{i \neq j}^N \left( \omega_{ij}^* \omega_{ij} + 2 \sum_{k(\neq i,j)}^N \omega_{ki}^* \omega_{kj} \right) t_i^a t_j^a \\ &\quad - \frac{1}{6} \sum_{i \neq j}^N \omega_{ij}^* \omega_{ij} (t_i^a t_j^a)^2 + \sum_{i \neq j \neq k} \left( \frac{1}{3} \omega_{ik}^* \omega_{ij} - \frac{1}{2} \omega_{ik} \omega_{ij}^* \right) t_i^a t_j^a t_i^b t_k^b. \end{aligned} \quad (\text{B.9})$$

## B.2 Operators annihilating the $(q, \eta = 1)_1$ , $q \geq 2$ lattice Moore-Read state

We restrict ourselves to the case  $q \geq 2$ , since the construction of the parent Hamiltonian when  $q = 1$  has already been obtained. Note however that the derivation presented here can also be used to obtain a Hamiltonian without the  $SU(2)$  symmetry for the  $(q = 1, \eta = 1)_1$  Moore-Read state. The Moore-Read states are defined from the operators  $\mathcal{V}_{n_j}(z_j)$  given in Eq. (3.57). Additional operators that are needed to construct parent Hamiltonians are the operators  $G^\pm(z) = \chi(z) e^{\pm i\sqrt{q}\phi(z)}$ ,  $J(z) = \frac{i}{\sqrt{q}} \partial\phi(z)$ . The operator product expansion (OPE) of  $G^+(z)$  and  $J(z)$  with the operators used to build the wave function are, for  $q \geq 2$ ,

$$G^+(z) \mathcal{V}_{n_j}(w) \sim (-1)^{j-1} \frac{\delta_{n_j,0}}{z-w} \mathcal{V}_1(w), \quad (\text{B.10})$$

$$J(z) \mathcal{V}_{n_j}(w) \sim \frac{1}{q} \frac{qn_j - 1}{z-w} \mathcal{V}_{n_j}(w). \quad (\text{B.11})$$

We first show that  $\chi^q(w)$  is a null field, where

$$\chi^q(w) = \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} \frac{1}{(z-w)^{q-1}} G^+(z) V_1(w) - \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} qJ(z) V_2(w), \quad (\text{B.12})$$

$$= \Omega_2^q(w) - \Omega_3^q(w), \quad (\text{B.13})$$

where the contour of the integration is a circle around  $w$  traversed counter-clockwise. We have that

$$\begin{aligned} \Omega_2^q(w) &= \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} \left[ \frac{1}{(z-w)^{q-1}} G^+(z) V_1(w) \right], \\ &= \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} \left[ \frac{1}{(z-w)^{q-1}} \chi(z) \chi(w) e^{+i\sqrt{q}\phi(z)} e^{i(q-1)\phi(w)/\sqrt{q}} \right], \\ &= \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} \left[ \frac{(z-w)^{q-1}}{(z-w)^{q-1}} \chi(z) \chi(w) e^{i\sqrt{q}\phi(z)+i(q-1)\phi(w)/\sqrt{q}} \right], \\ &= \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} \left[ \left( \frac{1}{z-w} + (z-w)A(w) + \dots \right) e^{i\sqrt{q}\phi(z)+i(q-1)\phi(w)/\sqrt{q}} \right], \\ &= \oint_w \frac{dz}{2\pi i} \left[ \left( \frac{1}{(z-w)^2} + A(w) + \dots \right) e^{i\sqrt{q}\phi(z)+i(q-1)\phi(w)/\sqrt{q}} \right], \\ &= \oint_w \frac{dz}{2\pi i} \left[ \frac{1}{(z-w)^2} e^{i\sqrt{q}\phi(z)+i(q-1)\phi(w)/\sqrt{q}} \right], \\ &= \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} \left[ i\sqrt{q}\partial\phi(w) e^{i(2q-1)\phi(w)/\sqrt{q}} \right], \end{aligned} \quad (\text{B.14})$$

and

$$\Omega_3^q(w) = \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} [qJ(z) V_2(w)] \quad (\text{B.15})$$

$$= \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} \left[ \sqrt{qi}\partial\phi(z) e^{i(2q-1)\phi(w)/\sqrt{q}} \right] \quad (\text{B.16})$$

$$= \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} \left[ \sqrt{qi}\partial\phi(w) e^{i(2q-1)\phi(w)/\sqrt{q}} \right] \quad (\text{B.17})$$

$$= \Omega_2^q(w), \quad (\text{B.18})$$

which shows that  $\chi^q(w)$  is a null field. Similarly, there are other null fields

$$\chi^p(w) = \oint_w \frac{dz}{2\pi i} \frac{1}{(z-w)^p} G^+(z) V_1(w), \quad (\text{B.19})$$

$$\chi^{q-1}(w) = \oint_w \frac{dz}{2\pi i} \left[ \frac{1}{(z-w)^{q-1}} G^+(z) V_1(w) - \frac{1}{(z-w)} V_2(w) \right]. \quad (\text{B.20})$$

Let us now use the fact that by replacing the field at site  $i$  by a null field, the correlator vanishes :

$$0 = \langle \mathcal{V}_{n_1}(z_1) \cdots \chi^a(z_i) \cdots \mathcal{V}_{n_N}(z_N) \rangle. \quad (\text{B.21})$$

We will transform this equation into an equation involving the wave function by deforming the contour integral and moving the operators in the null fields at different positions. Let us do it for the null field  $\chi^q(w)$ . We will use the OPEs as well as the commutation relations :

$$\mathcal{V}_{n_j}(z_j)G^+(z) = (-1)^{(q+1)n_j-1}G^+(z)\mathcal{V}_{n_j}(z_j), \quad (\text{B.22})$$

where we have used :  $e^{i\alpha\phi(z)} :: e^{i\beta\phi(w)} : = (-1)^{\alpha\beta} : e^{i\beta\phi(w)} :: e^{i\alpha\phi(z)} :$  and  $\chi(z)\chi(w) = (-1)\chi(w)\chi(z)$ , which adds a minus sign only when  $n_j = 1$ . We then have, starting with the first term involving  $\chi_2^q(w)$  :

$$\begin{aligned} & \oint_{z_i} \frac{dz}{2\pi i} \frac{1}{(z-z_i)^q} \langle \mathcal{V}_{n_1}(z_1) \dots G^+(z) V_1(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle, \\ &= - \sum_{j(\neq i)} \oint_{z_j} \frac{dz}{2\pi i} \frac{1}{(z-z_j)^q} \langle \mathcal{V}_{n_1}(z_1) \dots G^+(z) V_1(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle, \\ &= -(-1)^{i-1} \sum_{j=1}^{i-1} \oint_{z_j} \frac{dz}{2\pi i} \frac{(-1)^{(q+1)\sum_{k=j}^{i-1} n_k}}{(z-z_i)^q} \frac{\delta_{n_j,0}}{z-z_j} \\ & \quad \times \langle \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_1(z_j) \dots V_1(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle \\ & - (-1)^{i-1} \sum_{j=i+1}^N \oint_{z_j} \frac{dz}{2\pi i} \frac{(-1)^{(q+1)}(-1)^{(q+1)\sum_{k=i+1}^{j-1} n_k}}{(z-z_i)^q} \frac{\delta_{n_j,0}}{z-z_j} \\ & \quad \times \langle \mathcal{V}_{n_1}(z_1) \dots V_1(z_i) \dots \mathcal{V}_1(z_j) \dots \mathcal{V}_{n_N}(z_N) \rangle, \\ &= -(-1)^{i-1} \sum_{j=1}^{i-1} \frac{(-1)^{(q+1)\sum_{k=j}^{i-1} n_k}}{(z_j-z_i)^q} \delta_{n_j,0} \langle \mathcal{V}_{n_1}(z_1) \dots \mathcal{V}_1(z_j) \dots V_1(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle \\ & - (-1)^{i-1} \sum_{j=i+1}^N \frac{(-1)^{(q+1)}(-1)^{(q+1)\sum_{k=i+1}^{j-1} n_k}}{(z_j-z_i)^q} \delta_{n_j,0} \\ & \quad \times \langle \mathcal{V}_{n_1}(z_1) \dots V_1(z_i) \dots \mathcal{V}_1(z_j) \dots \mathcal{V}_{n_N}(z_N) \rangle, \\ &= - \sum_{j=1}^{i-1} \frac{(-1)^{(q+1)\sum_{k=j+1}^{i-1} n_k}}{(z_j-z_i)^q} \delta_{n_j,0} \psi_{(q,\eta=1)_1}(n_1, \dots, 1, \dots, 1, \dots, n_N) \\ & - \sum_{j=i+1}^N \frac{(-1)^{(q+1)}(-1)^{(q+1)\sum_{k=i+1}^{j-1} n_k}}{(z_j-z_i)^q} \delta_{n_j,0} \psi_{(q,\eta=1)_1}(n_1, \dots, 1, \dots, 1, \dots, n_N), \\ &= \sum_{j=1}^{i-1} (-1)^{(q+1)} \frac{(-1)^{(q+1)\sum_{k=j+1}^{i-1} n_k}}{(z_i-z_j)^q} \delta_{n_j,0} \psi_{(q,\eta=1)_1}(n_1, \dots, 1, \dots, 1, \dots, n_N) \\ & + \sum_{j=i+1}^N \frac{(-1)^{(q+1)\sum_{k=i+1}^{j-1} n_k}}{(z_i-z_j)^q} \delta_{n_j,0} \psi_{(q,\eta=1)_1}(n_1, \dots, 1, \dots, 1, \dots, n_N), \quad (\text{B.23}) \end{aligned}$$



where  $\psi_{(q,\eta=1)_1}$  is the wave function of the  $(q, \eta = 1)_1$  Moore-Read state. Let us now define the creation and annihilation operators  $d_j, d_j^\dagger, d'_j, d'_j{}^\dagger$  acting on the Hilbert space at site  $j$  as

$$d_j |n_j\rangle = (-1)^{(q+1)\sum_{k=1}^{j-1} n_k} \begin{cases} 0 & n_j = 0 \\ |0\rangle & n_j = 1 \\ 0 & n_j = 2 \end{cases} \quad (\text{B.24})$$

$$d_j^\dagger |n_j\rangle = (-1)^{(q+1)\sum_{k=1}^{j-1} n_k} \begin{cases} |1\rangle & n_j = 0 \\ 0 & n_j = 1 \\ 0 & n_j = 2 \end{cases} \quad (\text{B.25})$$

$$d'_j |n_j\rangle = (-1)^{(q+1)\sum_{k=1}^{j-1} n_k} \begin{cases} 0 & n_j = 0 \\ 0 & n_j = 1 \\ |1\rangle & n_j = 2 \end{cases} \quad (\text{B.26})$$

$$d'_j{}^\dagger |n_j\rangle = (-1)^{(q+1)\sum_{k=1}^{j-1} n_k} \begin{cases} 0 & n_j = 0 \\ |2\rangle & n_j = 1 \\ 0 & n_j = 2 \end{cases} \quad (\text{B.27})$$

We also define the particle number operators corresponding to these operators as  $n_j^{(1)} = d_j^\dagger d_j$  and  $n_j^{(2)} = d'_j{}^\dagger d'_j$ , such that  $n_j = n_j^{(1)} + 2n_j^{(2)}$ . We multiply Eq. (B.23) by

$$|n_1, \dots, n_{j-1}, n_j, n_{j+1}, \dots, n_{i-1}, 2, n_{i+1}, \dots, n_N\rangle, \quad (\text{B.28})$$

and sum over all  $n_k, k \neq i$ , to get an expression involving the wave function

$$\sum_{j(\neq i)} \frac{1}{(z_i - z_j)^q} d_j d'_j{}^\dagger |\psi_{(q,\eta=1)_1}\rangle. \quad (\text{B.29})$$

Let us now look at the second term involving  $\chi_3^q(w)$ :

$$-q \oint_{z_i} \frac{dz}{2\pi i} \frac{1}{z - z_i} \langle \mathcal{V}_{n_1}(z_1) \dots J(z) V_2(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle \quad (\text{B.30})$$

$$= q \sum_{j(\neq i)} \oint_{z_j} \frac{dz}{2\pi i} \frac{1}{z - z_i} \langle \mathcal{V}_{n_1}(z_1) \dots J(z) V_2(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle, \quad (\text{B.31})$$

$$= \sum_{j(\neq i)} \frac{qn_j - 1}{z_j - z_i} \langle \mathcal{V}_{n_1}(z_1) \dots V_2(z_i) \dots \mathcal{V}_{n_N}(z_N) \rangle. \quad (\text{B.32})$$

Let us multiply this expression by

$$|n_1, \dots, n_{i-1}, 2, n_{i+1}, \dots, n_N\rangle = \sum_{n'_i} n_i^{(2)} |n_1, \dots, n'_i, \dots, n_N\rangle, \quad (\text{B.33})$$

which leads after summing over all  $n_k$ ,  $k \neq i$  to

$$- \sum_{j(\neq i)} \frac{qn_j - 1}{z_i - z_j} n_i^{(2)} |\Psi\rangle. \quad (\text{B.34})$$

Summing the two terms together, we get an operator annihilating the wave function:

$$\Lambda_i^q = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^q} d_j d_i^\dagger - \sum_{j(\neq i)} \frac{qn_j - 1}{z_i - z_j} n_i^{(2)}. \quad (\text{B.35})$$

The same procedure applied to the other null fields gives the following operators annihilating the  $(q, \eta = 1)_1$  CFT wave functions :

$$\Lambda^0 = \sum_i d_i, \quad (\text{B.36})$$

$$\Lambda_i^p = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^p} d_j d_i^\dagger, \quad (\text{B.37})$$

$$\Lambda_i^{q-1} = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^{q-1}} d_j d_i^\dagger + n_i^{(2)}. \quad (\text{B.38})$$

### B.3 Operators annihilating the $(q, \eta = 1)_{1/2}$ , $q \geq 2$ lattice Moore-Read state

To obtain operators annihilating the  $(q, \eta = 1)_{1/2}$  Moore-Read wave functions, we can follow the same procedure and use the following null field instead of  $\chi^q(w)$  :

$$\oint_w \frac{dz}{2\pi i} \frac{1}{(z-w)^q} G^+(z) \left( \oint_w \frac{dx}{2\pi i} G^+(x) V_0(w) \right) - q \oint_w \frac{dz}{2\pi i} \frac{1}{z-w} J(z) \left( \oint_w \frac{dx}{2\pi i} \frac{1}{(x-w)^{q-1}} G^+(x) V_1(w) \right). \quad (\text{B.39})$$

However this procedure does not work when  $q = 2$  since this is not a null field when  $q = 2$ . Instead of following this approach, we present a different way to obtain operators annihilating the  $(q, \eta = 1)_{1/2}$  Moore-Read wave functions, using the operators already obtained in Appendix B.2. The resulting operators are the same that would be obtained directly using null fields when  $q > 2$  but this approach allows us to also construct a parent Hamiltonian when  $q = 2$ .

We start by multiplying the previously obtained operators on the left by  $d'_i$ . Since  $d'_i d_i^\dagger = n_i^{(1)}$ , this leads to new operators

$$\Lambda_i''^p = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^p} d_j n_i^{(1)}, \quad (\text{B.40})$$

$$\Lambda_i''^{q-1} = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^{q-1}} d_j n_i^{(1)} + d'_i, \quad (\text{B.41})$$

$$\Lambda_i''^q = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^q} d_j n_i^{(1)} - \sum_{j(\neq i)} \frac{qn_j - 1}{z_i - z_j} d'_i. \quad (\text{B.42})$$

The operator  $d'_i$  in  $\Lambda_i''^q$  can be replaced by  $-\sum_{h(\neq i)} \frac{1}{(z_i - z_h)^{q-1}} d_h n_i^{(1)}$  since  $\Lambda_i''^{q-1}$  annihilates the  $(q, \eta = 1)_1$  wave function :

$$\Lambda_i''^q = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^q} d_j n_i^{(1)} + \sum_{j(\neq i)} \sum_{h(\neq i)} \frac{1}{(z_i - z_h)^{q-1}} \frac{qn_j^{(1)} - 1}{z_i - z_j} d_h n_i^{(1)}. \quad (\text{B.43})$$

$\Lambda_i''^q$  then acts separately on  $\mathcal{H}^1$  and on  $\mathcal{H}^2$  and the operators  $\Lambda_i''^1, \dots, \Lambda_i''^{q-2}$  give zero on  $\mathcal{H}^2$ . By keeping only the terms acting on  $\mathcal{H}^1$  and since the  $(q, \eta = 1)_{1/2}$  states are projections of the  $(q, \eta = 1)_1$  states on  $\mathcal{H}^1$ , we get operators annihilating the  $(q, \eta = 1)_{1/2}$  wave functions:

$$\Lambda'^0 = \sum_i d_i, \quad (\text{B.44})$$

$$\Lambda_i'^p = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^p} d_j n_i^{(1)}, \quad (\text{B.45})$$

$$\Lambda_i'^{q-1} = \sum_{j(\neq i)} \frac{1}{(z_i - z_j)^q} d_j n_i^{(1)} + \sum_{j(\neq i)} \sum_{h(\neq i)} \frac{1}{(z_i - z_h)^{q-1}} \frac{qn_j^{(1)} - 1}{z_i - z_j} d_h n_i^{(1)}. \quad (\text{B.46})$$

The Hamiltonian can finally be constructed as

$$H = \sum_{i=1}^N \sum_{a=0}^{q-1} \Lambda_i'^{a\dagger} \Lambda_i'^a + \left( \sum_i n_i^{(1)} - \frac{N}{q} \right)^2. \quad (\text{B.47})$$



# Bibliography

- [1] I. Glasser, J. I. Cirac, G. Sierra, and A. E. B. Nielsen. „Construction of spin models displaying quantum criticality from quantum field theory“. *Nucl. Phys. B* 886 (2014), pp. 63–74.
- [2] I. Glasser, J. I. Cirac, G. Sierra, and A. E. B. Nielsen. „Exact parent Hamiltonians of bosonic and fermionic Moore-Read states on lattices and local models“. *New J. Phys.* 17 (2015), p. 082001.
- [3] I. Glasser, J. I. Cirac, G. Sierra, and A. E. B. Nielsen. „Lattice effects on Laughlin wave functions and parent Hamiltonians“. *Phys. Rev. B* 94 (2016), p. 245104.
- [4] I. Glasser, N. Pancotti, M. August, I. D. Rodríguez, and J. I. Cirac. „Neural-Network Quantum States, String-Bond States, and Chiral Topological States“. *Phys. Rev. X* 8 (2018), p. 011006.
- [5] A. E. B. Nielsen, I. Glasser, and I. D. Rodríguez. „Quasielectrons as inverse quasiholes in lattice fractional quantum Hall models“. *New J. Phys.* 20 (2018), p. 033029.
- [6] I. Glasser, N. Pancotti, and J. I. Cirac. „Supervised learning with generalized tensor networks“. *arXiv:1806.00000* (2018).
- [7] L. Bernard, A. E. Faraggi, I. Glasser, J. Rizos, and H. Sonmez. „String derived exophobic  $SU(6) \times SU(2)$  GUTs“. *Nucl. Phys. B* 868 (2013), pp. 1–15.
- [8] R. P. Feynman. „Simulating physics with computers“. *Int. J. Theor. Phys.* 21 (1982), pp. 467–488.
- [9] I. Bloch, J. Dalibard, and W. Zwerger. „Many-body physics with ultracold gases“. *Rev. Mod. Phys.* 80 (2008), pp. 885–964.
- [10] L. Faddeev. „Algebraic aspects of the Bethe ansatz“. *Int. J. Mod. Phys. A* 10 (1995), pp. 1845–1878.
- [11] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. „Quantum Monte Carlo simulations of solids“. *Rev. Mod. Phys.* 73 (2001), pp. 33–83.
- [12] M. Troyer and U.-J. Wiese. „Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations“. *Phys. Rev. Lett.* 94 (2005), p. 170201.

- [13] S. R. White. „Density matrix formulation for quantum renormalization groups“. *Phys. Rev. Lett.* 69 (1992), pp. 2863–2866.
- [14] S. R. White. „Density-matrix algorithms for quantum renormalization groups“. *Phys. Rev. B* 48 (1993), pp. 10345–10356.
- [15] F. Verstraete, V. Murg, and J. Cirac. „Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems“. *Adv. Phys.* 57 (2008), pp. 143–224.
- [16] U. Schollwöck. „The density-matrix renormalization group in the age of matrix product states“. *Ann. Phys.* 326 (2011), pp. 96–192.
- [17] M. B. Hastings. „An area law for one-dimensional quantum systems“. *J. Stat. Mech.* 2007 (2007), P08024.
- [18] M. B. Hastings. „Entropy and entanglement in quantum ground states“. *Phys. Rev. B* 76 (2007), p. 035114.
- [19] M. B. Plenio, J. Eisert, J. Dreißig, and M. Cramer. „Entropy, Entanglement, and Area: Analytical Results for Harmonic Lattice Systems“. *Phys. Rev. Lett.* 94 (2005), p. 060503.
- [20] M. Cramer and J. Eisert. „Correlations, spectral gap and entanglement in harmonic quantum systems on generic lattices“. *New J. Phys.* 8 (2006), p. 71.
- [21] N. de Beaudrap, M. Ohliger, T. J. Osborne, and J. Eisert. „Solving Frustration-Free Spin Systems“. *Phys. Rev. Lett.* 105 (2010), p. 060504.
- [22] K. Van Acoleyen, M. Mariën, and F. Verstraete. „Entanglement Rates and Area Laws“. *Phys. Rev. Lett.* 111 (2013), p. 170501.
- [23] R. Orús. „A practical introduction to tensor networks: Matrix product states and projected entangled pair states“. *Ann. Phys.* 349 (2014), pp. 117–158.
- [24] J. Eisert. „Entanglement and tensor network states“. *Modeling and Simulation* 3 (2013), p. 520.
- [25] F. Verstraete and J. I. Cirac. „Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions“. *arXiv:0407066* (2004).
- [26] F. Verstraete and J. I. Cirac. „Matrix product states represent ground states faithfully“. *Phys. Rev. B* 73 (2006), p. 094423.
- [27] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac. „Entropy Scaling and Simulability by Matrix Product States“. *Phys. Rev. Lett.* 100 (2008), p. 030504.
- [28] Y. Ge and J. Eisert. „Area laws and efficient descriptions of quantum many-body states“. *New J. Phys.* 18 (2016), p. 083026.
- [29] A. Molnar, N. Schuch, F. Verstraete, and J. I. Cirac. „Approximating Gibbs states of local Hamiltonians efficiently with projected entangled pair states“. *Phys. Rev. B* 91 (2015), p. 045138.
- [30] S. Sachdev. *Quantum phase transitions*. Cambridge University Press, 2011.

- [31] D. Ashton. *Scale invariance in the critical Ising model*.  
<http://www.kineticallyconstrained.com/2009/05/critical-point.html>.  
 2009.
- [32] P. Di Francesco, P. Mathieu, and D. Sénéchal. *Conformal field theory*. Springer New York, 1997.
- [33] C. Holzhey, F. Larsen, and F. Wilczek. „Geometric and renormalized entropy in conformal field theory“. *Nucl. Phys. B* 3 (1994), pp. 443–467.
- [34] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev. „Entanglement in quantum critical phenomena“. *Phys. Rev. Lett.* 90 (2003), p. 227902.
- [35] P. Calabrese and J. L. Cardy. „Entanglement entropy and quantum field theory“. *J. Stat. Mech.* 2004 (2004), P06002.
- [36] R. N. Pfeifer, G. Evenbly, and G. Vidal. „Entanglement renormalization, scale invariance, and quantum criticality“. *Phys. Rev. A* 79 (2009), p. 040301.
- [37] S. Montangero, M. Rizzi, V. Giovannetti, and R. Fazio. „Critical exponents with a multiscale entanglement renormalization Ansatz channel“. *Phys. Rev. B* 80 (2009), p. 113103.
- [38] J. I. Cirac and G. Sierra. „Infinite matrix product states, conformal field theory, and the Haldane-Shastry model“. *Phys. Rev. B* 81 (2010), p. 104431.
- [39] X. G. Wen. „Topological Order in Rigid States“. *Int. J. Mod. Phys. B* 4 (1990), p. 239.
- [40] A. Kitaev and J. Preskill. „Topological Entanglement Entropy“. *Phys. Rev. Lett.* 96 (2006), p. 110404.
- [41] M. Levin and X.-G. Wen. „Detecting Topological Order in a Ground State Wave Function“. *Phys. Rev. Lett.* 96 (2006), p. 110405.
- [42] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill. „Topological quantum memory“. *J. Math. Phys.* 43 (2002), pp. 4452–4505.
- [43] A. Kitaev. „Fault-tolerant quantum computation by anyons“. *Ann. Phys.* 303 (2003), pp. 2–30.
- [44] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma. „Non-Abelian anyons and topological quantum computation“. *Rev. Mod. Phys.* 80 (2008), pp. 1083–1159.
- [45] D. C. Tsui, H. L. Stormer, and A. C. Gossard. „Two-Dimensional Magnetotransport in the Extreme Quantum Limit“. *Phys. Rev. Lett.* 48 (1982), pp. 1559–1562.
- [46] J. P. Eisenstein and H. L. Stormer. „The Fractional Quantum Hall Effect“. *Science* 248 (1990), pp. 1510–1516.
- [47] D. Sheng, Z.-C. Gu, K. Sun, and L. Sheng. „Fractional quantum Hall effect in the absence of Landau levels“. *Nat. Commun.* 2 (2011), p. 389.
- [48] T. Neupert, L. Santos, C. Chamon, and C. Mudry. „Fractional Quantum Hall States at Zero Magnetic Field“. *Phys. Rev. Lett.* 106 (2011), p. 236804.

- [49] Y.-F. Wang, Z.-C. Gu, C.-D. Gong, and D. N. Sheng. „Fractional Quantum Hall Effect of Hard-Core Bosons in Topological Flat Bands“. *Phys. Rev. Lett.* 107 (2011), p. 146803.
- [50] K. Sun, Z. Gu, H. Katsura, and S. Das Sarma. „Nearly Flatbands with Nontrivial Topology“. *Phys. Rev. Lett.* 106 (2011), p. 236803.
- [51] N. Regnault and B. A. Bernevig. „Fractional Chern Insulator“. *Phys. Rev. X* 1 (2011), p. 021014.
- [52] V. Kalmeyer and R. B. Laughlin. „Equivalence of the resonating-valence-bond and fractional quantum Hall states“. *Phys. Rev. Lett.* 59 (1987), pp. 2095–2098.
- [53] M. Greiter and R. Thomale. „Non-Abelian Statistics in a Quantum Antiferromagnet“. *Phys. Rev. Lett.* 102 (2009), p. 207203.
- [54] A. E. B. Nielsen, G. Sierra, and J. I. Cirac. „Local models of fractional quantum Hall states in lattices and physical implementation“. *Nat. Commun.* 4 (2013), p. 2864.
- [55] B. Bauer, L. Cincio, B. P. Keller, M. Dolfi, G. Vidal, S. Trebst, and A. W. W. Ludwig. „Chiral spin liquid and emergent anyons in a Kagome lattice Mott insulator“. *Nat. Commun.* 5 (2014), p. 5137.
- [56] S.-S. Gong, W. Zhu, and D. N. Sheng. „Emergent Chiral Spin Liquid: Fractional Quantum Hall Effect in a Kagome Heisenberg Model“. *Sci. Rep* 4 (2014), p. 6317.
- [57] Y.-C. He, D. N. Sheng, and Y. Chen. „Chiral Spin Liquid in a Frustrated Anisotropic Kagome Heisenberg Model“. *Phys. Rev. Lett.* 112 (2014), p. 137202.
- [58] S.-S. Gong, W. Zhu, L. Balents, and D. N. Sheng. „Global phase diagram of competing ordered and quantum spin-liquid phases on the kagome lattice“. *Phys. Rev. B* 91 (2015), p. 075112.
- [59] Y.-C. He and Y. Chen. „Distinct Spin Liquids and Their Transitions in Spin-1/2  $XXZ$  Kagome Antiferromagnets“. *Phys. Rev. Lett.* 114 (2015), p. 037201.
- [60] A. Wietek, A. Sterdyniak, and A. M. Läuchli. „Nature of chiral spin liquids on the kagome lattice“. *Phys. Rev. B* 92 (2015), p. 125122.
- [61] W.-J. Hu, W. Zhu, Y. Zhang, S. Gong, F. Becca, and D. N. Sheng. „Variational Monte Carlo study of a chiral spin liquid in the extended Heisenberg model on the kagome lattice“. *Phys. Rev. B* 91 (2015), p. 041124.
- [62] K. Kumar, K. Sun, and E. Fradkin. „Chiral spin liquids on the kagome lattice“. *Phys. Rev. B* 92 (2015), p. 094433.
- [63] R. B. Laughlin. „Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations“. *Phys. Rev. Lett.* 50 (1983), pp. 1395–1398.
- [64] T. B. Wahl, H.-H. Tu, N. Schuch, and J. I. Cirac. „Projected Entangled-Pair States Can Describe Chiral Topological States“. *Phys. Rev. Lett.* 111 (2013), p. 236805.
- [65] J. Dubail and N. Read. „Tensor network trial states for chiral topological phases in two dimensions and a no-go theorem in any dimension“. *Phys. Rev. B* 92 (2015), p. 205307.



- [66] S. Yang, T. B. Wahl, H.-H. Tu, N. Schuch, and J. I. Cirac. „Chiral Projected Entangled-Pair State with Topological Order“. *Phys. Rev. Lett.* 114 (2015), p. 106803.
- [67] D. Poilblanc, J. I. Cirac, and N. Schuch. „Chiral topological spin liquids with projected entangled pair states“. *Phys. Rev. B* 91 (2015), p. 224431.
- [68] G. Moore and N. Read. „Nonabelions in the fractional quantum Hall effect“. *Nucl. Phys. B* 360 (1991), pp. 362–396.
- [69] A. E. B. Nielsen, J. I. Cirac, and G. Sierra. „Laughlin Spin-Liquid States on Lattices Obtained from Conformal Field Theory“. *Phys. Rev. Lett.* 108 (2012), p. 257206.
- [70] A. E. B. Nielsen, J. I. Cirac, and G. Sierra. „Quantum spin Hamiltonians for the  $SU(2)_k$  WZW model“. *J. Stat. Mech.* 2011 (2011), P11014.
- [71] G. Carleo and M. Troyer. „Solving the quantum many-body problem with artificial neural networks“. *Science* 355 (2017), pp. 602–606.
- [72] Y. LeCun, Y. Bengio, and G. Hinton. „Deep learning“. *Nature* 521 (2015), pp. 436–444.
- [73] E. Zahedinejad, J. Ghosh, and B. C. Sanders. „Designing High-Fidelity Single-Shot Three-Qubit Gates: A Machine-Learning Approach“. *Phys. Rev. Applied* 6 (2016), p. 054005.
- [74] M. August and X. Ni. „Using recurrent neural networks to optimize dynamical decoupling for quantum memory“. *Phys. Rev. A* 95 (2017), p. 012335.
- [75] L. Banchi, N. Pancotti, and S. Bose. „Quantum gate learning in qubit networks: Toffoli gate without time-dependent control“. *npj Quant. Inf.* 2 (2016), p. 16019.
- [76] G. Torlai and R. G. Melko. „Neural Decoder for Topological Codes“. *Phys. Rev. Lett.* 119 (2017), p. 030501.
- [77] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo. „Many-body quantum state tomography with neural networks“. *arXiv:1703.05334* (2017).
- [78] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber. „Learning phase transitions by confusion“. *Nat. Phys.* 13 (2017), pp. 435–439.
- [79] J. Carrasquilla and R. G. Melko. „Machine learning phases of matter“. *Nat. Phys.* 13 (2017), pp. 431–434.
- [80] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst. „Machine learning quantum phases of matter beyond the fermion sign problem“. *Sci. Rep* 7 (2017), p. 8823.
- [81] L. Wang. „Discovering phase transitions with unsupervised learning“. *Phys. Rev. B* 94 (2016), p. 195105.
- [82] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu. „Self-learning Monte Carlo method“. *Phys. Rev. B* 95 (2017), p. 041101.
- [83] L. Huang and L. Wang. „Accelerated Monte Carlo simulations with restricted Boltzmann machines“. *Phys. Rev. B* 95 (2017), p. 035105.
- [84] D. Koller and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*. The MIT Press, 2009.

- [85] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski. „A Learning Algorithm for Boltzmann Machines“. *Cognitive Science* 9 (1985), pp. 147–169.
- [86] D.-L. Deng, X. Li, and S. Das Sarma. „Exact Machine Learning Topological States“. *arXiv:1609.09060* (2016).
- [87] X. Gao and L.-M. Duan. „Efficient representation of quantum many-body states with deep neural networks“. *Nat. Commun.* 8 (2017), p. 662.
- [88] Y. Huang and J. E. Moore. „Neural network representation of tensor network and chiral states“. *arXiv:1701.06246* (2017).
- [89] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang. „Equivalence of restricted Boltzmann machines and tensor network states“. *Phys. Rev. B* 97 (2018), p. 085104.
- [90] A. Gendiar and T. Nishino. „Latent heat calculation of the three-dimensional  $q = 3, 4,$  and 5 Potts models by the tensor product variational approach“. *Phys. Rev. E* 65 (2002), p. 046702.
- [91] F. Mezzacapo, N. Schuch, M. Boninsegni, and J. I. Cirac. „Ground-state properties of quantum many-body systems: entangled-plaquette states and variational Monte Carlo“. *New J. Phys.* 11 (2009), p. 083026.
- [92] H. J. Changlani, J. M. Kinder, C. J. Umrigar, and G. K.-L. Chan. „Approximating strongly correlated wave functions with correlator product states“. *Phys. Rev. B* 80 (2009), p. 245116.
- [93] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac. „Simulation of Quantum Many-Body Systems with Strings of Operators and Monte Carlo Tensor Contractions“. *Phys. Rev. Lett.* 100 (2008), p. 040501.
- [94] A. Sfondrini, J. Cerrillo, N. Schuch, and J. I. Cirac. „Simulating two- and three-dimensional frustrated quantum systems with string-bond states“. *Phys. Rev. B* 81 (2010), p. 214426.
- [95] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. „Quantum machine learning“. *Nature* 549 (2017), pp. 195–202.
- [96] V. Dunjko and H. J. Briegel. „Machine learning and artificial intelligence in the quantum domain: a review of recent progress“. *Rep. Prog. Phys.* (2018).
- [97] E. Farhi, J. Goldstone, and S. Gutmann. „A Quantum Approximate Optimization Algorithm“. *arXiv:1411.4028* (2014).
- [98] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik. „The theory of variational hybrid quantum-classical algorithms“. *New J. of Phys.* 18 (2016), p. 023023.
- [99] E. Farhi and H. Neven. „Classification with quantum neural networks on near term processors“. *arXiv:1802.0600* (2014).
- [100] M. Schuld and N. Killoran. „Quantum machine learning in feature hilbert spaces“. *arXiv:1803.07128* (2018).
- [101] M. Benedetti, D. Garcia-Pintos, Y. Nam, and A. Perdomo-Ortiz. „A generative modeling approach for benchmarking and training shallow quantum circuits“. *arXiv:1801.07686* (2018).

- [102] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. „Quantum Circuit Learning“. *arXiv:1803.00745* (2018).
- [103] A. Anandkumar, R. Ge, D. Hsu, S. M. Kakade, and M. Telgarsky. „Tensor decompositions for learning latent variable models“. *J. Mach. Learn. Res.* 15 (2014), p. 2773.
- [104] E. M. Stoudenmire and D. J. Schwab. „Supervised Learning with Quantum-Inspired Tensor Networks“. *Advances in Neural Information Processing Systems* 29 (2016), p. 4799.
- [105] A. Novikov, M. Trofimov, and I. Oseledets. „Exponential Machines“. *arXiv:1605.03795* (2016).
- [106] M. Blondel, M. Ishihata, A. Fujino, and N. Ueda. „Polynomial networks and factorization machines: new insights and efficient training algorithms“. *33rd International Conference on Machine Learning*. 2018, pp. 850–858.
- [107] W. Huggins, P. Patel, K. B. Whaley, and E. M. Stoudenmire. „Towards Quantum Machine Learning with Tensor Networks“. *arXiv:1803.11537* (2018).
- [108] E. Grant, M. Benedetti, S. Cao, A. Hallam, J. Lockhart, V. Stojevic, A. G. Green, and S. Severini. „Hierarchical quantum classifiers“. *arXiv:1804.03680* (2018).
- [109] P. Coleman and A. J. Schofield. „Quantum criticality“. *Nature* 433 (2005), pp. 226–229.
- [110] S. Sachdev and B. Keimer. „Quantum criticality“. *Physics Today* 64 (2011), p. 29.
- [111] 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 (2010), pp. 177–180.
- [112] C. Rüegg, B Normand, M Matsumoto, A Furrer, D. F. McMorrow, K Krämer, H.-U. Güdel, S. N. Gvasaliya, H Mutka, and M Boehm. „Quantum magnets under pressure: controlling elementary excitations in  $TlCuCl_3$ “. *Phys. Rev. Lett.* 100 (2008), p. 205701.
- [113] J. Custers, P. Gegenwart, H. Wilhelm, K. Neumaier, Y. Tokiwa, O. Trovarelli, C. Geibel, F. Steglich, C. Pépin, and P. Coleman. „The break-up of heavy electrons at a quantum critical point“. *Nature* 424 (2003), pp. 524–527.
- [114] H. v. Löhneysen, T Pietrus, G Portisch, H. Schlager, A Schröder, M Sieck, and T Trappmann. „Non-Fermi-liquid behavior in a heavy-fermion alloy at a magnetic instability“. *Phys. Rev. Lett.* 72 (1994), p. 3262.
- [115] S. A. Grigera, R. S. Perry, A. J. Schofield, M Chiao, S. R. Julian, G. G. Lonzarich, S. I. Ikeda, Y Maeno, A. J. Millis, and A. P. Mackenzie. „Magnetic field-tuned quantum criticality in the metallic ruthenate  $Sr_3Ru_2O_7$ “. *Science* 294 (2001), pp. 329–332.
- [116] S. R. Julian, C Pfeleiderer, F. M. Grosche, N. D. Mathur, G. J. McMullan, A. J. Diver, I. R. Walker, and G. G. Lonzarich. „The normal states of magnetic d and f transition metals“. *J. Phys.: Condens. Matter* 8 (1996), p. 9675.
- [117] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki. „Rigorous results on valence-bond ground states in antiferromagnets“. *Phys. Rev. Lett.* 59 (1987), pp. 799–802.

- [118] G. Vidal. „Efficient Classical Simulation of Slightly Entangled Quantum Computations“. *Phys. Rev. Lett.* 91 (2003), p. 147902.
- [119] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. „Equation of State Calculations by Fast Computing Machines“. *J. Chem. Phys.* 21 (1953), p. 1087.
- [120] W. K. Hastings. „Monte Carlo sampling methods using Markov chains and their applications“. *Biometrika* 57 (1970), 97–109.
- [121] M. B. Hastings, I. González, A. B. Kallin, and R. G. Melko. „Measuring Renyi Entanglement Entropy in Quantum Monte Carlo Simulations“. *Phys. Rev. Lett.* 104 (2010), p. 157201.
- [122] M. Bajdich, L. Mitas, L. K. Wagner, and K. E. Schmidt. „Pfaffian pairing and backflow wavefunctions for electronic structure quantum Monte Carlo methods“. *Phys. Rev. B* 77 (2008), p. 115112.
- [123] M. Wimmer. „Algorithm 923“. *ACM Trans. on Math. Softw.* 38 (2012), 1–17.
- [124] H.-H. Tu. „Projected BCS states and spin Hamiltonians for the  $SO(n)_1$  Wess-Zumino-Witten model“. *Phys. Rev. B* 87 (2013), p. 041103.
- [125] R. Bondesan and T. Quella. „Infinite matrix product states for long-range spin models“. *Nucl. Phys. B* 886 (2014), pp. 483–523.
- [126] A. E. B. Nielsen, G. Sierra, and J. I. Cirac. „Violation of the area law and long-range correlations in infinite-dimensional-matrix product states“. *Phys. Rev. A* 83 (2011), p. 053807.
- [127] F. D. M. Haldane. „Exact Jastrow-Gutzwiller resonating-valence-bond ground state of the spin-(1/2 antiferromagnetic Heisenberg chain with  $1/r^2$  exchange“. *Phys. Rev. Lett.* 60 (1988), pp. 635–638.
- [128] B. S. Shastry. „Exact solution of an  $S=1/2$  Heisenberg antiferromagnetic chain with long-ranged interactions“. *Phys. Rev. Lett.* 60 (1988), pp. 639–642.
- [129] P. Pfeuty. „The one-dimensional Ising model with a transverse field“. *Ann. Phys.* 57 (1970), pp. 29–90.
- [130] R. Elliott, P. Pfeuty, and C. Wood. „Ising Model with a Transverse Field“. *Phys. Rev. Lett.* 25 (1970), p. 2370.
- [131] B. Groisman, S. Popescu, and A. Winter. „Quantum, classical, and total amount of correlations in a quantum state“. *Phys. Rev. A* 72 (2005), p. 032317.
- [132] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. Cirac. „Area Laws in Quantum Systems: Mutual Information and Correlations“. *Phys. Rev. Lett.* 100 (2008), p. 070502.
- [133] R. G. Melko, A. B. Kallin, and M. B. Hastings. „Finite-size scaling of mutual information in Monte Carlo simulations: Application to the spin-1  $2 \times 2$  XXZ model“. *Phys. Rev. B* 82 (2010), p. 100409.
- [134] R. R. P. Singh, M. B. Hastings, A. B. Kallin, and R. G. Melko. „Finite-Temperature Critical Behavior of Mutual Information“. *Phys. Rev. Lett.* 106 (2011), p. 135701.

- [135] L. Bonnes, H. Pichler, and A. M. Läuchli. „Entropy perspective on the thermal crossover in a fermionic Hubbard chain“. *Phys. Rev. B* 88 (2013), p. 155103.
- [136] M. Žnidarič, T. Prosen, and I. Pižorn. „Complexity of thermal states in quantum spin chains“. *Phys. Rev. A* 78 (2008), p. 022103.
- [137] J. Wilms, J. Vidal, F. Verstraete, and S. Dusuel. „Finite-temperature mutual information in a simple phase transition“. *J. Stat. Mech.* 2012 (2012), P01023.
- [138] A. Uhlmann. „The "Transition Probability" in the State Space of a  $*$ -Algebra“. *Rep. Math. Phys.* 9 (1976), p. 273.
- [139] R. Jozsa. „Fidelity for mixed quantum states“. *J. Mod. Opt.* 41 (1994), pp. 2315–2323.
- [140] M. O. Goerbig. *Quantum Hall Effects*. Lecture notes of Les Houches Summer School “Ultracold Gases and Quantum Information”. 2009.
- [141] R. B. Laughlin. „Elementary Theory: the Incompressible Quantum Fluid“. *The Quantum Hall Effect*. Ed. by R. E. Prange and S. M. Girvin. New York, NY: Springer New York, 1990, pp. 233–301.
- [142] K. v. Klitzing, G. Dorda, and M. Pepper. „New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance“. *Phys. Rev. Lett.* 45 (1980), pp. 494–497.
- [143] F. D. M. Haldane and E. H. Rezayi. „Finite-Size Studies of the Incompressible State of the Fractionally Quantized Hall Effect and its Excitations“. *Phys. Rev. Lett.* 54 (1985), pp. 237–240.
- [144] F. D. M. Haldane. „Fractional Quantization of the Hall Effect: A Hierarchy of Incompressible Quantum Fluid States“. *Phys. Rev. Lett.* 51 (1983), pp. 605–608.
- [145] T. H. Hansson, M. Hermanns, and S. Viefers. „Quantum Hall quasielectron operators in conformal field theory“. *Phys. Rev. B* 80 (2009), p. 165330.
- [146] O. S. Zozulya, M. Haque, K. Schoutens, and E. H. Rezayi. „Bipartite entanglement entropy in fractional quantum Hall states“. *Phys. Rev. B* 76 (2007), p. 125310.
- [147] B. I. Halperin. „Statistics of Quasiparticles and the Hierarchy of Fractional Quantized Hall States“. *Phys. Rev. Lett.* 52 (1984), pp. 2390–2390.
- [148] J. K. Jain. „Composite-fermion approach for the fractional quantum Hall effect“. *Phys. Rev. Lett.* 63 (1989), pp. 199–202.
- [149] R. Willett, J. P. Eisenstein, H. L. Störmer, D. C. Tsui, A. C. Gossard, and J. H. English. „Observation of an even-denominator quantum number in the fractional quantum Hall effect“. *Phys. Rev. Lett.* 59 (1987), pp. 1776–1779.
- [150] M. Greiter, X.-G. Wen, and F. Wilczek. „Paired Hall state at half filling“. *Phys. Rev. Lett.* 66 (1991), pp. 3205–3208.
- [151] M. Greiter, X.-G. Wen, and F. Wilczek. „Paired Hall states in double-layer electron systems“. *Phys. Rev. B* 46 (1992), pp. 9586–9589.
- [152] C. Nayak and F. Wilczek. „ $2n$ -quasihole states realize  $2n-1$ -dimensional spinor braiding statistics in paired quantum Hall states“. *Nucl. Phys. B* 479 (1996), pp. 529–553.

- [153] N. Read and D. Green. „Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries and the fractional quantum Hall effect“. *Phys. Rev. B* 61 (2000), pp. 10267–10297.
- [154] D. A. Ivanov. „Non-Abelian Statistics of Half-Quantum Vortices in  $p$ -Wave Superconductors“. *Phys. Rev. Lett.* 86 (2001), pp. 268–271.
- [155] P. Bonderson, V. Gurarie, and C. Nayak. „Plasma analogy and non-Abelian statistics for Ising-type quantum Hall states“. *Phys. Rev. B* 83 (2011), p. 075303.
- [156] D. F. Schroeter, E. Kapit, R. Thomale, and M. Greiter. „Spin Hamiltonian for which the Chiral Spin Liquid is the Exact Ground State“. *Phys. Rev. Lett.* 99 (2007), p. 097202.
- [157] R. Thomale, E. Kapit, D. F. Schroeter, and M. Greiter. „Parent Hamiltonian for the chiral spin liquid“. *Phys. Rev. B* 80 (2009), p. 104406.
- [158] E. Kapit and E. Mueller. „Exact Parent Hamiltonian for the Quantum Hall States in a Lattice“. *Phys. Rev. Lett.* 105 (2010), p. 215303.
- [159] M. Greiter, D. F. Schroeter, and R. Thomale. „Parent Hamiltonian for the non-Abelian chiral spin liquid“. *Phys. Rev. B* 89 (2014), p. 165125.
- [160] H.-H. Tu, A. E. B. Nielsen, J. I. Cirac, and G. Sierra. „Lattice Laughlin states of bosons and fermions at filling fractions  $1/q$ “. *New J. Phys.* 16 (2014), p. 033025.
- [161] K. Kumar, K. Sun, and E. Fradkin. „Chern-Simons theory of magnetization plateaus of the spin- $\frac{1}{2}$  quantum XXZ Heisenberg model on the kagome lattice“. *Phys. Rev. B* 90 (2014), p. 174409.
- [162] W. Zhu, S. S. Gong, and D. N. Sheng. „Interaction-driven fractional quantum Hall state of hard-core bosons on kagome lattice at one-third filling“. *arXiv:1509.05509* (2015).
- [163] M. Hafezi, A. S. Sørensen, E. Demler, and M. D. Lukin. „Fractional quantum Hall effect in optical lattices“. *Phys. Rev. A* 76 (2007), p. 023613.
- [164] Y. Zhang, T. Grover, and A. Vishwanath. „Topological entanglement entropy of  $\mathbb{Z}_2$  spin liquids and lattice Laughlin states“. *Phys. Rev. B* 84 (2011), p. 075128.
- [165] Y. Zhang, T. Grover, A. Turner, M. Oshikawa, and A. Vishwanath. „Quasiparticle statistics and braiding from ground-state entanglement“. *Phys. Rev. B* 85 (2012), p. 235151.
- [166] M. P. Zaletel and R. S. K. Mong. „Exact matrix product states for quantum Hall wave functions“. *Phys. Rev. B* 86 (2012), p. 245305.
- [167] B. Estienne, Z. Papic, N. Regnault, and B. A. Bernevig. „Matrix product states for trial quantum Hall states“. *Phys. Rev. B* 87 (2013), p. 161112.
- [168] A. E. B. Nielsen, G. Sierra, and J. I. Cirac. „Optical-lattice implementation scheme of a bosonic topological model with fermionic atoms“. *Phys. Rev. A* 90 (2014), p. 013606.
- [169] S. M. Girvin. „Particle-hole symmetry in the anomalous quantum Hall effect“. *Phys. Rev. B* 29 (1984), pp. 6012–6014.

- [170] J. Wildeboer and N. E. Bonesteel. „Spin correlations and topological entanglement entropy in a non-Abelian spin-one spin liquid“. *Phys. Rev. B* 94 (2016), p. 045125.
- [171] B. Herwerth, G. Sierra, H.-H. Tu, J. I. Cirac, and A. E. B. Nielsen. „Edge states for the Kalmeyer-Laughlin wave function“. *Phys. Rev. B* 92 (2015), p. 245111.
- [172] I. D. Rodríguez and A. E. B. Nielsen. „Continuum limit of lattice models with Laughlin-like ground states containing quasiholes“. *Phys. Rev. B* 92 (2015), p. 125105.
- [173] A. E. B. Nielsen. „Anyon braiding in semianalytical fractional quantum Hall lattice models“. *Phys. Rev. B* 91 (2015), p. 041106.
- [174] G. S. Jeon and J. K. Jain. „Nature of quasiparticle excitations in the fractional quantum Hall effect“. *Phys. Rev. B* 68 (2003), p. 165346.
- [175] T. H. Hansson, C.-C. Chang, J. K. Jain, and S. Viefers. „Composite-fermion wave functions as correlators in conformal field theory“. *Phys. Rev. B* 76 (2007), p. 075347.
- [176] B. A. Bernevig and F. D. M. Haldane. „Clustering Properties and Model Wave Functions for Non-Abelian Fractional Quantum Hall Quasielectrons“. *Phys. Rev. Lett.* 102 (2009), p. 066802.
- [177] T. H. Hansson, M. Hermanns, N. Regnault, and S. Viefers. „Conformal Field Theory Approach to Abelian and Non-Abelian Quantum Hall Quasielectrons“. *Phys. Rev. Lett.* 102 (2009), p. 166805.
- [178] J. Suorsa, S. Viefers, and T. H. Hansson. „A general approach to quantum Hall hierarchies“. *New J. Phys.* 13 (2011), p. 075006.
- [179] G. J. Sreejith, A. Wójs, and J. K. Jain. „Unpaired Composite Fermion, Topological Exciton, and Zero Mode“. *Phys. Rev. Lett.* 107 (2011), p. 136802.
- [180] I. D. Rodríguez, A. Sterdyniak, M. Hermanns, J. K. Slingerland, and N. Regnault. „Quasiparticles and excitons for the Pfaffian quantum Hall state“. *Phys. Rev. B* 85 (2012), p. 035128.
- [181] B. Yang and F. D. M. Haldane. „Nature of Quasielectrons and the Continuum of Neutral Bulk Excitations in Laughlin Quantum Hall Fluids“. *Phys. Rev. Lett.* 112 (2014), p. 026804.
- [182] M. Greiter, V. Schnells, and R. Thomale. „Laughlin states and their quasiparticle excitations on the torus“. *Phys. Rev. B* 93 (2016), p. 245156.
- [183] D. Arovas, J. R. Schrieffer, and F. Wilczek. „Fractional Statistics and the Quantum Hall Effect“. *Phys. Rev. Lett.* 53 (1984), pp. 722–723.
- [184] B. Paredes, P. Fedichev, J. I. Cirac, and P. Zoller. „ $\frac{1}{2}$ -Anyons in Small Atomic Bose-Einstein Condensates“. *Phys. Rev. Lett.* 87 (2001), p. 010402.
- [185] P. Bonderson, V. Gurarie, and C. Nayak. „Plasma analogy and non-Abelian statistics for Ising-type quantum Hall states“. *Phys. Rev. B* 83 (2011), p. 075303.
- [186] E. Kapit, P. Ginsparg, and E. Mueller. „Non-Abelian Braiding of Lattice Bosons“. *Phys. Rev. Lett.* 108 (2012), p. 066802.

- [187] Y.-L. Wu, B. Estienne, N. Regnault, and B. A. Bernevig. „Braiding Non-Abelian Quasiholes in Fractional Quantum Hall States“. *Phys. Rev. Lett.* 113 (2014), p. 116801.
- [188] Z. Liu, R. N. Bhatt, and N. Regnault. „Characterization of quasiholes in fractional Chern insulators“. *Phys. Rev. B* 91 (2015), p. 045126.
- [189] G. S. Jeon, K. L. Graham, and J. K. Jain. „Fractional Statistics in the Fractional Quantum Hall Effect“. *Phys. Rev. Lett.* 91 (2003), p. 036801.
- [190] G. S. Jeon, K. L. Graham, and J. K. Jain. „Berry phases for composite fermions: Effective magnetic field and fractional statistics“. *Phys. Rev. B* 70 (2004), p. 125316.
- [191] G. S. Jeon and J. K. Jain. „Thermodynamic behavior of braiding statistics for certain fractional quantum Hall quasiparticles“. *Phys. Rev. B* 81 (2010), p. 035319.
- [192] „Quantal phase factors accompanying adiabatic changes“. *Proc. Royal Soc. Lond. A* 392 (1984), pp. 45–57.
- [193] N. Read. „Non-Abelian adiabatic statistics and Hall viscosity in quantum Hall states and  $p_x + ip_y$  paired superfluids“. *Phys. Rev. B* 79 (2009), p. 045308.
- [194] A. Deshpande and A. E. B. Nielsen. „Lattice Laughlin states on the torus from conformal field theory“. *J. Stat. Mech.* 2016 (2016), p. 013102.
- [195] H.-H. Tu, A. E. B. Nielsen, and G. Sierra. „Quantum spin models for the Wess–Zumino–Witten model“. *Nucl. Phys. B* 886 (2014), pp. 328–363.
- [196] A. S. Sørensen, E. Demler, and M. D. Lukin. „Fractional Quantum Hall States of Atoms in Optical Lattices“. *Phys. Rev. Lett.* 94 (2005), p. 086803.
- [197] A. E. B. Nielsen and G. Sierra. „Bosonic fractional quantum Hall states on the torus from conformal field theory“. *J. Stat. Mech.* 2014 (2014), P04007.
- [198] Y.-F. Wang, H. Yao, Z.-C. Gu, C.-D. Gong, and D. N. Sheng. „Non-Abelian Quantum Hall Effect in Topological Flat Bands“. *Phys. Rev. Lett.* 108 (2012), p. 126805.
- [199] Z. Liu and E. J. Bergholtz. „From fractional Chern insulators to Abelian and non-Abelian fractional quantum Hall states: Adiabatic continuity and orbital entanglement spectrum“. *Phys. Rev. B* 87 (2013), p. 035306.
- [200] Z. Liu, E. J. Bergholtz, and E. Kapit. „Non-Abelian fractional Chern insulators from long-range interactions“. *Phys. Rev. B* 88 (2013), p. 205101.
- [201] E. J. Bergholtz and Z. Liu. „Topological flat band models and fractional Chern insulators“. *Int. J. Mod. Phys. B*, 27 (2013), p. 1330017.
- [202] B. Paredes. „Non-Abelian fractional quantum Hall states for hard-core bosons in one dimension“. *Phys. Rev. B* 85 (2012), p. 195150.
- [203] I. Affleck. „Exact critical exponents for quantum spin chains, non-linear  $\sigma$ -models at  $\Theta = \Pi$  and the quantum Hall effect“. *Nucl. Phys. B* 265 (1986), pp. 409–447.
- [204] I. Affleck, D. Gepner, H. J. Schulz, and T. Ziman. „Critical behaviour of spin- $s$  Heisenberg antiferromagnetic chains: analytic and numerical results“. *J. Phys. A: Math. Gen.* 22 (1989), p. 511.



- [205] O. Narayan and B. S. Shastry. „Spin- $s$  wave functions with algebraic order“. *Phys. Rev. B* 70 (2004), p. 184440.
- [206] D. C. Cabra and P Pujol. *Lect. Notes Phys.* 645 (2004), p. 253.
- [207] E. Ardonne and G. Sierra. „Chiral correlators of the Ising conformal field theory“. *J. Phys. A: Math. Theor.* 43 (2010), p. 505402.
- [208] Z. Cai. „Approximating quantum many-body wave-functions using artificial neural networks“. *arXiv:1704.05148* (2017).
- [209] H. Saito. „Solving the Bose-Hubbard Model with Machine Learning“. *J. Phys. Soc. Japan* 86 (2017), p. 093001.
- [210] G. Torlai and R. G. Melko. „Learning thermodynamics with Boltzmann machines“. *Phys. Rev. B* 94 (2016), p. 165134.
- [211] W. L. McMillan. „Ground State of Liquid He<sup>4</sup>“. *Phys. Rev.* 138 (1965), A442–A451.
- [212] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. „Quantum Monte Carlo simulations of solids“. *Rev. Mod. Phys.* 73 (2001), pp. 33–83.
- [213] M. P. Nightingale and V. Melik-Alaverdian. „Optimization of Ground- and Excited-State Wave Functions and van der Waals Clusters“. *Phys. Rev. Lett.* 87 (2001), p. 043401.
- [214] J. Toulouse and C. J. Umrigar. „Optimization of quantum Monte Carlo wave functions by energy minimization“. *J. Chem. Phys.* 126 (2007), p. 084102.
- [215] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig. „Alleviation of the Fermion-Sign Problem by Optimization of Many-Body Wave Functions“. *Phys. Rev. Lett.* 98 (2007), p. 110201.
- [216] C. J. Umrigar and C. Filippi. „Energy and Variance Optimization of Many-Body Wave Functions“. *Phys. Rev. Lett.* 94 (2005), p. 150201.
- [217] S. Sorella. „Generalized Lanczos algorithm for variational quantum Monte Carlo“. *Phys. Rev. B* 64 (2001), p. 024512.
- [218] S. Sorella. „Wave function optimization in the variational Monte Carlo method“. *Phys. Rev. B* 71 (2005), p. 241103.
- [219] S. Amari. „Natural Gradient Works Efficiently in Learning“. *Neural Comput.* 10 (1998), pp. 251–276.
- [220] E. Neuscamman, C. J. Umrigar, and G. K.-L. Chan. „Optimizing large parameter sets in variational quantum Monte Carlo“. *Phys. Rev. B* 85 (2012), p. 045103.
- [221] A. W. Sandvik and G. Vidal. „Variational Quantum Monte Carlo Simulations with Tensor-Network States“. *Phys. Rev. Lett.* 99 (2007), p. 220602.
- [222] S. Liang and H. Pang. „Approximate diagonalization using the density matrix renormalization-group method: A two-dimensional-systems perspective“. *Phys. Rev. B* 49 (1994), pp. 9214–9217.
- [223] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac. „Computational Complexity of Projected Entangled Pair States“. *Phys. Rev. Lett.* 98 (2007), p. 140506.

- [224] R. Jastrow. „Many-Body Problem with Strong Forces“. *Phys. Rev.* 98 (1955), pp. 1479–1484.
- [225] D.-L. Deng, X. Li, and S. Das Sarma. „Quantum Entanglement in Neural Network States“. *Phys. Rev. X* 7 (2017), p. 021021.
- [226] C. K. Majumdar and D. K. Ghosh. „On Next Nearest Neighbor Interaction in Linear Chain. I“. *J. Math. Phys.* 10 (1969), pp. 1388–1398.
- [227] S. R. Clark. „Unifying Neural-network Quantum States and Correlator Product States via Tensor Networks“. *arXiv:1710.03545* (2017).
- [228] O. Sikora, H.-W. Chang, C.-P. Chou, F. Pollmann, and Y.-J. Kao. „Variational Monte Carlo simulations using tensor-product projected states“. *Phys. Rev. B* 91 (2015), p. 165113.
- [229] R. Kaubruegger, L. Pastori, and J. C. Budich. „Chiral Topological Phases from Artificial Neural Networks“. *arXiv:1710.04713* (2017).
- [230] D. Poilblanc. „Investigation of the chiral antiferromagnetic Heisenberg model using PEPS“. *arXiv:1707.07844* (2017).
- [231] A. Cichocki, N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, and D. P. Mandic. „Tensor Networks for Dimensionality Reduction and Large-scale Optimization: Part 1 Low-Rank Tensor Decompositions“. *Foundations and Trends® in Machine Learning* 9 (2016), pp. 249–429.
- [232] A. Cichocki, A.-H. Phan, Q. Zhao, N. Lee, I. Oseledets, M. Sugiyama, and D. P. Mandic. „Tensor Networks for Dimensionality Reduction and Large-scale Optimization: Part 2 Applications and Future Perspectives“. *Foundations and Trends® in Machine Learning* 9 (2017), pp. 431–673.
- [233] A. Novikov, D. Podoprikhin, A. Osokin, and D. Vetrov. „Tensorizing Neural Networks“. *arXiv:1509.06569* (2015).
- [234] A. Hallam, E. Grant, V. Stojevic, S. Severini, and A. G. Green. „Compact neural networks based on the multiscale entanglement renormalization ansatz“. *arXiv:1711.03357* (2017).
- [235] R. Yu, S. Zheng, A. Anandkumar, and Y. Yue. „Long-term forecasting using tensor-train RNNs“. *arXiv:1711.00073* (2017).
- [236] J. Kossaifi, Z. C. Lipton, A. Khanna, T. Furlanello, and A. Anandkumar. „Tensor contraction and regression networks“. *arXiv:1707.08308* (2017).
- [237] N. Cohen, O. Sharir, and A. Shashua. „On the expressive power of deep learning: a tensor analysis“. *29th Annual Conf. on Learning Theory*. 2016, pp. 698–728.
- [238] N. Cohen and A. Shashua. „Inductive bias of deep convolutional networks through pooling geometry“. *5th International Conference on Learning Representations*. 2016.
- [239] N. Cohen, O. Sharir, and A. Shashua. „Convolutional rectifier networks as generalized tensor decompositions“. *33rd International Conference on Machine Learning*. 2016, pp. 955–963.

- [240] N. Cohen, O. Sharir, Y. Levine, R. Tamari, D. Yakira, and A. Shashua. „Analysis and Design of Convolutional Networks via Hierarchical Tensor Decompositions“. *arXiv:1705.02302* (2017).
- [241] Y. Levine, D. Yakira, N. Cohen, and A. Shashua. „Deep learning and quantum entanglement: Fundamental connections with implications to network design“. *6th International Conference on Learning Representations*. 2018.
- [242] Z.-Y. Han, J. Wang, H. Fan, L. Wang, and P. Zhang. „Unsupervised Generative Modeling Using Matrix Product States“. *arXiv:1709.01662* (2017).
- [243] Y. Liu, X. Zhang, M. Lewenstein, and S.-J. Ran. „Learning architectures based on quantum entanglement: a simple matrix product state algorithm for image recognition“. *arXiv:1803.09111* (2018).
- [244] D. Liu, S.-J. Ran, P. Wittek, C. Peng, R. B. García, G. Su, and M. Lewenstein. „Machine Learning by Two-Dimensional Hierarchical Tensor Networks: A Quantum Information Theoretic Perspective on Deep Architectures“. *arXiv:1710.04833* (2017).
- [245] B. J. Frey. „Extending Factor Graphs so as to Unify Directed and Undirected Graphical Models“. *Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence*. 2003, pp. 257–264.
- [246] P. Smolensky. „Information processing in dynamical systems: Foundations of harmony theory“. *D. Rumelhart and J. McClelland (Eds.), Parallel distributed processing, vol. 1, chapter 6, MIT Press* (1986).
- [247] G. E. Hinton. „Training products of experts by minimizing contrastive divergence.“ *Neural Comput.* 14(8) (2002), 1771–1800.
- [248] A. Critch and J. Morton. „Algebraic Geometry of Matrix Product States“. *SIGMA Symmetry Integrability Geom. Methods Appl.* (2014).
- [249] M. Kliesch, D. Gross, and J. Eisert. „Matrix-Product Operators and States: NP-Hardness and Undecidability“. *Phys. Rev. Lett.* 113 (2014), p. 160503.
- [250] E. Robeva and A. Seigal. „Duality of Graphical Models and Tensor Networks“. *arXiv:1710.01437* (2017).
- [251] Y. Levine, O. Sharir, N. Cohen, and A. Shashua. „Bridging Many-Body Quantum Physics and Deep Learning via Tensor Networks“. *arXiv:1803.09780* (2018).
- [252] J. D. Biamonte, S. R. Clark, and D. Jaksch. „Categorical Tensor Network States“. *AIP Advances* 1 (2011), p. 042172.
- [253] Y. LeCun, C. Cortes, and C. J. Burges. *MNIST handwritten digit database*. <http://yann.lecun.com/exdb/mnist/>. 1998.
- [254] H. Xiao, K. Rasul, and R. Vollgraf. „Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms“. *arXiv:1708.07747* (2017).
- [255] J. Salamon, C. Jacoby, and J. P. Bello. „A Dataset and Taxonomy for Urban Sound Research“. *22nd ACM International Conference on Multimedia (ACM-MM'14)*. Orlando, USA, 2014, pp. 1041–1044.

- [256] K. J. Piczak. „Environmental sound classification with convolutional neural networks“. *IEEE 25th International Workshop on Machine Learning for Signal Processing*. 2015, pp. 1–6.
- [257] V. Pestun and Y. Vlassopoulos. „Tensor network language model“. *arXiv:1710.10248* (2017).
- [258] V. Pestun, J. Terilla, and Y. Vlassopoulos. „Language as a matrix product state“. *arXiv:1711.01416* (2017).
- [259] I. M. Georgescu, S. Ashhab, and F. Nori. „Quantum simulation“. *Rev. Mod. Phys.* 86 (2014), pp. 153–185.

# List of Acronyms

- **CNN** : Convolutional Neural Network
- **CFT** : Conformal Field Theory
- **DMRG** : Density Matrix Renormalization Group
- **dSBS** : diagonal String-Bond State
- **EPS** : Entangled Plaquette State
- **FQH** : Fractional Quantum Hall
- **GTN** : generalized tensor network
- **MPS** : Matrix Product State
- **PEPS** : Projected Entangled Pair State
- **RBM** : Restricted Boltzmann Machine
- **SBS** : String-Bond State
- **sRBM** : short-range Restricted Boltzmann Machine
- **TEE** : Topological Entanglement Entropy
- **VMC** : Variational Monte Carlo



# List of Figures

1.1	Area law in one and two dimensions. . . . .	3
1.2	a) Decomposition of the general tensor representing the wave function into a one-dimensional tensor network known as Matrix Product State (MPS). b) In two dimensions, the corresponding tensor network is a Projected Entangled Pair State (PEPS). . . . .	4
1.3	a) Phase diagram of the 1D transverse field Ising model at zero temperature. For low values of the transverse field $h$ with respect to the Ising coupling $J$ the ground state is a ferromagnet, while for high values of the transverse field the ground state is a paramagnet. In between there is a quantum phase transition at a particular value of the transverse field. b) This model can be mapped to a 2D classical Ising model at finite temperature, which displays a phase transition at a critical temperature $T_c$ . The spin configurations display scale invariance at the critical point. Monte Carlo simulation results by Douglas Ashton[31]. . . . .	6
1.4	Hall resistance and longitudinal resistance of a 2D electron gas subject to a magnetic field. From [46], reprinted with permission from AAAS.	8
1.5	Different network architectures based on graphs that can be used to define a function of the variables $X_1$ , $X_2$ and $X_3$ : a) a feed-forward neural network, b) an undirected probabilistic graphical model, c) a factor graph, d) a restricted Boltzmann machine, e) a tensor network (Matrix Product State). . . . .	10
2.1	Graphical notation for tensor networks: (a) vector, (b) matrix, (c) order 3 tensor, (d) vector-matrix multiplication (e) matrix-matrix multiplication (f) matrix-tensor contraction. Figure reproduced from [6]. . . . .	16
2.2	Decomposition of a tensor into a MPS in open boundary conditions (left) or periodic boundary conditions (right). . . . .	17
2.3	Correlation function of two observables for a MPS. . . . .	20
2.4	MPS (top) and infinite dimensional MPS (bottom) in which the tensors have been replaced by field operators depending on the spin $s_j$ and acting at positions $z_j$ . . . . .	22

- 2.5 All quantities are evaluated from Monte Carlo simulations for  $\alpha = 0.2$ .  
 (a) Scaling of the Renyi entropy  $S_L^{(2)}$  for different values of the mass,  $N = 200$  spins. The inset shows with a logarithmic scale for the abscissa that the scaling is logarithmic in  $L$  when the mass is close to zero. The dotted line in the inset is a fit of the form  $\frac{1}{4} \log(L) + \text{constant}$ , confirming that the central charge of the critical point is 1. For higher masses, the entropy saturates to a constant. The error bars, estimated from the standard error of a mean of Monte Carlo trajectories with different initial conditions, are of the order  $10^{-3}$  for all points.  
 (b) Absolute value of the magnetization in the  $x$  direction for one site  $\langle \sigma_{N/2}^x \rangle$  as a function of the mass for a chain with  $N = 600$  spins. The error bars are smaller than 5% of the values for all points. The dotted line is a fit of the first 6 points of the form  $\gamma m^\tau$ , from which the critical exponent can be extracted. Here  $\tau = 0.4$ .  
 (c) Absolute value of the connected correlation function  $C_d^{xx}$ ,  $N = 200$  spins. When the mass is zero this quantity decays polynomially (red solid line), while in the massive case the decay for large  $d$  is exponential. The inset shows the same quantity in a log-log scale. The error bars are smaller than  $2 \times 10^{-4}$  for all points.  
 Figure reproduced from [1]. . . . . 29
- 2.6 The two spin chains A and B are separated by a distance  $\delta$ . The one dimensional state that is studied is the reduced state of the chain A. To compute the mutual information we separate the chain into the systems C of  $L$  spins in the middle of the chain and the complementary system D of  $N - L$  spins. Figure reproduced from [1]. . . . . 30
- 2.7 Absolute value of the magnetization in the  $x$  direction for one site  $\langle \sigma_{N/2}^x \rangle$  as a function of mass and temperature from Monte Carlo simulations with  $N = 600$ ,  $\alpha = 0.2$ . Figure reproduced from [1]. . . . . 31
- 2.8 Correlation length  $\lambda$  as a function of the temperature for a spin chain with  $N = 200$  and  $\alpha = 0.2$ , at  $m = 0.01$ . The background colours represent the value of  $|\langle \sigma_{N/2}^x \rangle|$  taken from the phase diagram (Fig. 2.7) for the same value of the mass. Figure reproduced from [1]. . . . . 32



2.9	<p>(a) Mutual information between two halves of the chain <math>I_{N/2}^{(2)}</math> for <math>N = 200</math>, <math>\alpha = 0.2</math> and <math>m = 10^{-6}</math>: for this value of the mass we are almost above the quantum critical point and the mutual information decreases as we go away from the critical point by increasing the temperature. The error bars, estimated from the standard error of a mean of Monte Carlo trajectories with different initial conditions, are smaller than <math>2 \times 10^{-2}</math> for all points.</p> <p>(b) Mutual information between two halves of the chain <math>I_{N/2}^{(2)}</math> for <math>N = 200</math>, <math>\alpha = 0.2</math> and <math>m = 0.01</math>. The background colours represent the value of <math> \langle \sigma_{N/2}^x \rangle </math> taken from the phase diagram (Fig. 2.7) for the same value of the mass: the mutual information increases as we enter the region of quantum criticality. The error bars are smaller than <math>5 \times 10^{-3}</math> for all points.</p> <p>Figure reproduced from [1]. . . . . 33</p>	33
2.10	<p>(a) Maximum of the fidelity <math>F(\rho_H, \rho_A)</math> between a thermal state obtained from a Hamiltonian of the form <math>H_m</math> and the state of chain A as a function of the mass, at <math>\alpha = 0.2</math> and different temperatures (<math>N = 12</math> spins).</p> <p>(b) Corresponding fidelity per site <math>F(\rho_H, \rho_A)^{1/N}</math> for different numbers of spins, at <math>\alpha = 0.2</math> and <math>m = 10^{-1}</math>.</p> <p>(c) Parameters <math>\lambda</math> and <math>\mu</math> in the Hamiltonian <math>H_m</math> for which the fidelity <math>F(\rho_H, \rho_A)</math> is maximal as a function of the mass, for <math>\alpha = 0.2</math>, <math>T = 10^{-3}</math>, <math>N = 10</math>. The two straight lines are fits of these data of the form <math>\lambda = \lambda_0 m^{2\alpha}</math>, <math>\mu = \mu_0 m^{2\alpha}</math>.</p> <p>Figure reproduced from [1]. . . . . 35</p>	35
3.1	<p>Measurement of the Hall resistance <math>R_H</math> and longitudinal resistance <math>R_L</math> when a current <math>I</math> is driven through a 2D electron gas subject to a perpendicular magnetic field. . . . . 39</p>	39
3.2	<p>Example of lattices considered in this work: square lattice (upper left), triangular lattice (bottom left), square lattice on a cylinder (right), obtained by mapping the plane to the cylinder. The coordinates for the square lattice on the cylinder are <math>z_j = \exp(2\pi((x_j - L_x/2 + 1/2)/L_y + (y_j - L_y/2 + 1/2)i/L_y))</math>, <math>x_j \in \{0, \dots, L_x - 1\}</math>, <math>y_j \in \{0, \dots, L_y - 1\}</math>. . . 43</p>	43
3.3	<p>Diagram of the <math>\psi_q^\eta</math> states for the values of <math>q</math> and <math>\eta</math> considered in this work. <math>q</math> is integer while <math>\eta</math> is a rational number in the interval <math>[0, q]</math>, since the states for <math>\eta &gt; q</math> (grey region) are not defined. When <math>\eta \rightarrow 0</math> or <math>\eta \rightarrow q</math> (red lines), the state is a ferromagnet if <math>N</math> is fixed, or a Laughlin state in the continuum if <math>N</math> goes to infinity and <math>N \propto 1/\eta</math> (if <math>\eta \rightarrow 0</math>) or <math>N \propto 1/(q - \eta)</math> (if <math>\eta \rightarrow q</math>). The solid blue disks on the line <math>\eta = 1</math> correspond to lattice Laughlin states that satisfy <math>\mu = \nu = 1/q</math>. The green disk is an integer quantum Hall state. On the orange line (<math>q</math> not necessarily integer), the state in 1D is critical, while on the dotted orange line it has long range order. Figure reproduced from [3]. . . . 46</p>	46

3.4	<p>Characteristic length <math>d_q^\eta</math> (3.17) estimating the correlation length of the <math>\psi_q^\eta</math> states. The characteristic length is computed for <math>q</math> integer and <math>\eta</math> integer and half-integer using a Metropolis-Hastings algorithm on a square <math>12 \times 12</math> (or <math>10 \times 12</math> for <math>q = 5</math>) lattice on the cylinder. The decay is found to be exponential everywhere, except when <math>\eta = q/2</math>, <math>q \geq 5</math>. Figure reproduced from [3]. . . . .</p>	48
3.5	<p>(a) Correlation function <math>C_{ij}</math> between a fixed lattice site <math>i</math> in the middle of the lattice and all other lattice sites <math>j</math> in the state <math>\psi_q^{q/2}</math> on a square <math>8 \times 8</math> lattice on the cylinder. (b) Absolute value of the correlation function <math> C_{i(i+\Delta y)} </math> between two lattice sites separated by a distance <math>\Delta y</math> in the periodic direction in the state <math>\psi_q^{q/2}</math> on a square <math>16 \times 16</math> lattice on the cylinder. There is long range anti-ferromagnetic order for <math>q \geq 6</math>, while the correlations decay exponentially for <math>q \leq 4</math>. Figure reproduced from [3]. . . . .</p>	49
3.6	<p>To compute the entanglement entropy of the state, the cylinder is cut into two halves and the Renyi entropy of the first half is computed using a Metropolis-Hastings algorithm. The topological entanglement entropy is extracted by varying the size <math>L_y</math> of the cylinder. Figure reproduced from [3]. . . . .</p>	50
3.7	<p>(a) Scaling of the Renyi entropy <math>S_{L_y}^{(2)}</math> of the state <math>\psi_4^\eta</math> on a <math>L_x \times L_y</math> square lattice on the cylinder. The topological entanglement entropy of the Laughlin state at filling <math>1/4</math> (<math>\gamma_0(4) = \ln(2) \approx 0.693</math>) is indicated with a red arrow. The size <math>L_x</math> is taken to be 12, unless <math>\eta = 2</math> in which case <math>L_x</math> is 20. Larger sizes are taken when <math>\eta = q/2</math>, to account for the longer correlation length and to get rid of finite size effects. The black lines are linear fits and the values found for the topological entanglement entropy when <math>\eta</math> equals to 0.5, 1, 1.5 and 2 are respectively 0.698, 0.734, 0.643 and 0.718. (b) Scaling of the Renyi entropy <math>S_{L_y}^{(2)}</math> for the states at half-filling <math>\psi_2^1</math> and <math>\psi_3^{3/2}</math> on the square lattice on the cylinder. Here <math>L_x</math> is 12 for <math>q = 2</math> and 20 for <math>q = 3</math>. The topological entanglement entropy of the Laughlin state at filling <math>1/2</math>, <math>\gamma_0(2) \approx 0.346</math> (resp. at filling <math>1/3</math>, <math>\gamma_0(3) \approx 0.549</math>) is indicated with a green (resp. blue) arrow. The values found for the topological entanglement entropy are respectively 0.375 and 0.536. Figure reproduced from [3]. . . . .</p>	51
3.8	<p>Overlap, defined as <math>O_q \equiv  \langle \psi_{\text{Néel}}^1   \psi_q^{q/2} \rangle ^2 +  \langle \psi_{\text{Néel}}^2   \psi_q^{q/2} \rangle ^2</math>, between the state at half-filling <math>\psi_q^{q/2}</math> and the two Néel states, computed using a Metropolis-Hastings algorithm on a square lattice of size <math>N_x \times N_x</math> on the cylinder. The errors from the Monte Carlo simulations are below 0.02 for all points. Figure reproduced from [3]. . . . .</p>	52

- 3.9 (a) Correlation function  $C_{ij}$  between a fixed lattice site  $i$  in the middle of the lattice and all other lattice sites  $j$  in the state  $\psi_q^{q/2}$  on a triangular and Kagome lattice on the cylinder. (b) Absolute value of the correlation function  $|C_{i(i+\Delta y)}|$  between two lattice sites separated by a distance  $\Delta y$  in the periodic direction in the state  $\psi_q^{q/2}$  on a triangular  $16 \times 16$  lattice on the cylinder. Unlike on the square lattice for the same parameters, there is no long range order. Figure reproduced from [3]. . . . . 53
- 3.10 Positions of the coordinates of the lattice sites along the interpolation between a square and a triangular lattice. The coordinates along the periodic direction  $y_j$  are kept fixed to keep the periodicity, while the coordinates along the other direction  $x_j$  are linearly interpolated between the square ( $\tau = 0$ ) and the triangular ( $\tau = 1$ ) lattice. The coordinates on the plane are then  $z_j = e^{\frac{2\pi}{L_y}(x_j + iy_j)}$ , as in Fig. 3.6. Figure reproduced from [3]. . . . . 54
- 3.11 Absolute value of the correlation function  $|C_{i(i+\Delta y)}|$  between two lattice sites separated by a distance  $\Delta y$  in the periodic direction in the state  $\psi_6^3$  on a  $16 \times 16$  lattice interpolating between the square ( $\tau = 0$ ) and the triangular ( $\tau = 1$ ) lattice on the cylinder. Figure reproduced from [3]. 54
- 3.12 (b) Scaling of the Renyi entropy  $S_{L_y}^{(2)}$  for the states at half-filling  $\psi_2^1, \psi_4^2$  and  $\psi_6^3$  on a triangular lattice on the cylinder. Here  $L_x$  is taken to be 16. The topological entanglement entropy of the continuum Laughlin state at  $q = 2$  (resp.  $q = 4, q = 6$ ) is indicated with a blue (resp. red, green) arrow. The values found for the topological entanglement entropy are respectively 0.347, 0.723 and 0.907. Figure reproduced from [3]. . . 55
- 3.13 (a) Square  $12 \times 12$  lattice on the cylinder with a charge (here  $p = -6$ ) at infinity. The color shows the density of the state  $\psi_2^1[p]_\infty$ . The state with zero charge at infinity  $\psi_2^1$  has uniform density  $1/2$  but here the density is modified at the edge to account for the change of total particle number. (b) Difference of density between the  $\psi_2^1[p]_\infty$  states and the  $\psi_2^1$  state with respect to the distance to the edge  $\Delta x$ . Values below  $2 \times 10^{-4}$  are not converged. The change of density is exponentially localized at the edge. Figure reproduced from [3]. . . . . 56
- 3.14 Absolute value of the correlation function  $|C_{i(i+\Delta y)}|$  between two lattice sites in the bulk (middle of the cylinder) separated by a distance  $\Delta y$  in the periodic direction on a square  $12 \times 12$  lattice on the cylinder. The states considered here are the  $\psi_q^1[p]_\infty$  state for different values of  $p$ . When  $p = 0$ , the state is simply the  $\psi_2^1$  state, while it is found that the other states have the same correlations. Figure reproduced from [3]. . 57
- 3.15 Diagram of the  $\psi_q^\eta$  states. The blue lines and dots represent values of the parameters for which exact parent Hamiltonians are derived in this section. On the light blue line the parent Hamiltonians derived have a degenerate ground state on the plane. Figure reproduced from [3]. . . 58

- 3.16 Difference in the density  $\langle n_i \rangle_{qh} - \langle n_i \rangle$  between the state  $\psi_q^\eta$  with one quasihole of charge  $1/q$  and one quasihole of charge  $(q-1)/q$  and the state  $\psi_q^\eta$  without quasiholes on a square  $20 \times 20$  (or  $12 \times 12$  for  $q = 10$ ,  $\eta = 5$ ) lattice on the cylinder. The coordinates  $w_j$  of the quasiholes are placed in the center between 4 lattice sites and are visible in blue as a lack of density on the neighbouring sites. At  $q = 4$ , it is found for all values of  $\eta$  that the quasiholes are localized. For  $q = 10$  however, the quasiholes are localized when  $\eta = 1$ , but at half-filling ( $\eta = 5$ ) we observe that there is no splitting of the charge between a quasihole of charge  $1/q$  and a quasihole of charge  $(q-1)/q$  and thus no screening of the quasiholes. Figure reproduced from [3]. . . . . 65
- 3.17 Modification of the particle density due to the presence of anyons in the lattice Laughlin state at  $q = 3$  and half lattice filling ( $\eta = q/2$ ). The lattice is chosen to be a kagome lattice defined on a disc with radius 27.9. A quasielectron (resp. quasihole) with charge  $-1/3$  (resp.  $+1/3$ ) is placed at the position  $*$  ( $+$ ), and the color of the  $j$ th lattice site shows  $\langle n_j \rangle_{[-1,+1]} - \langle n_j \rangle_{[0,0]}$ . Figure reproduced from [5]. . . . . 66
- 3.18 Excess charge 3.48 of a quasihole/quasielectron (blue/red) and their sum (green) for  $q = 3$  for different densities of the lattice sites. The quasihole/quasielectron is placed at the origin at the center of a hexagon in a kagome lattice. The lattice is defined on a disc with radius 27.9 for  $\eta = 1.5, 1, 0.1$  and 18.2 for  $\eta = 0.01$ , and an anyon of the opposite charge is placed at infinity. Figure reproduced from [5]. . . . . 67
- 3.19 Quasihole ( $+$ ) and quasielectron ( $*$ ) on the torus in the state  $\psi_2^1[1, -1]$ . The color of the lattice sites shows  $\langle n_j \rangle_{[1,-1]} - \langle n_j \rangle_{[0,0]}$ . When moving the quasihole around the blue curve we find numerically, using Monte Carlo simulations, that the difference in Berry phase when the quasielectron is at  $+$  and at  $*$ , respectively, is  $\phi = -3.145$  with a statistical error of order 0.003. This is in agreement with the expected result  $-\pi$ . Figure reproduced from [5]. . . . . 68
- 3.20 Lattice coordinates on the torus.  $(n_1, m_1)$  indicates the position of a quasihole, and  $(n_2, m_2)$  indicates the position of a quasielectron. Figure reproduced from [5]. . . . . 70
- 3.21 Illustration of a square lattice on the complex plane in the continuum limit ( $\eta \rightarrow 0, N \rightarrow \infty$ ) and in the lattice limit ( $\eta \rightarrow 1$ ). At each site there can be 0 (blue circle), 1 (blue disk) or 2 (red disks) particles. The interpolation is performed by fixing the number of particles  $M = \eta \frac{N}{q}$  and by varying  $\eta = \frac{a}{2\pi}$  between 0 and 1, which changes the number of lattice sites per particle between infinity and  $q$ . Figure reproduced from [2]. . . . . 72

- 3.22 Renyi entropy  $S_L^{(2)}$  of a subsystem of  $L$  consecutive sites for the 1D  $(q = 1, \eta)_1$  Moore-Read states (a) and  $(q = 2, \eta)_{1/2}$  Moore-Read states (b) for different values of  $\eta$ . The number of particles  $M = \eta N/q$  is fixed so the sizes of the chain are  $N = 40, 80, 160, 320$  for  $\eta = 1, 1/2, 1/4, 1/8$  respectively. The lines are linear fits of the points for  $\eta = 1/8$  (blue) and  $\eta = 1$  (red). Figure reproduced from [2]. . . . . 78
- 3.23 Rescaled correlation function  $\eta^{-2}C_L$  as a function of the distance between the sites for the 1D  $(q = 1, \eta)_1$  Moore-Read states (a) and  $(q = 2, \eta)_{1/2}$  Moore-Read states (b) for different values of  $\eta$ . The number of particles  $M = \eta N/q$  is fixed so the sizes of the chain are  $N = 40, 80, 160, 320$  for  $\eta = 1, 1/2, 1/4, 1/8$  respectively. The data for  $\eta = 1$  is shown in the insets in log-log scale, confirming the polynomial decay of correlations, and the line in the insets is a linear fit yielding critical exponents 0.70 ( $(q = 1, \eta = 1)_1$  state) and 1.02 ( $(q = 2, \eta = 1)_{1/2}$  state). In (b), the line is a fit of the form (3.72) with parameters  $K=0.494$  and  $A=0.123$ . Figure reproduced from [2]. . . . . 79
- 3.24 Linear behaviour of the Renyi entropy with the size of the cut  $L_y$  for the  $(q = 1, \eta)_1$  (a),  $(q = 2, \eta)_{1/2}$  (b) and  $(q = 2, \eta)_1$  (c) Moore-Read states on a  $L_x \times L_y$  lattice. The topological entanglement entropy of the continuum Moore-Read states at filling 1 (a),  $\gamma_0(1) \approx 0.69$ , and at filling 1/2 (b,c),  $\gamma_0(2) \approx 1.04$ , are indicated with a red arrow. The values of  $\eta$  are 1, 1/4 and 1/8 and the corresponding sizes  $L_x$  are respectively 12, 16 and 16. The insets are enlarged views confirming that the topological entanglement entropy stays the same when  $\eta$  is varied and that its value corresponds to  $\gamma_0(1)$  (resp.  $\gamma_0(1/2)$ ) in the first two cases, while the topological entanglement entropy of the  $(q = 2, \eta)_1$  Moore-Read state is close to zero in the lattice limit and close to  $\gamma_0(1/2)$  in the continuum limit. Figure reproduced from [2]. . . . . 80
- 3.25 Energy difference to the ground state energy for the first excited states of the Hamiltonian  $H_{2D}$  for different sizes of lattices. The overlap between the ground state of this Hamiltonian and the  $(q = 1, \eta = 1)_1$  Moore-Read state is indicated below each ground state. Figure reproduced from [2]. . . . . 84
- 3.26 (a) Overlap  $|\langle \psi_{\text{CFT}} | \psi_H \rangle|$  between the  $(q = 2, \eta = 1)_{1/2}$  Moore-Read state and the ground state of Hamiltonian  $H_{1D}^{(2)}$  with  $\kappa$  fixed to  $0.274 + 0.052i$ , as a function of the number  $N$  of lattice sites. The dotted line is a linear fit with equation  $y = 1.0075 - 0.00152N$ . If the overlap continues to follow this behavior at larger sizes, it will still be above 85% for a spin chain with 100 lattice sites. (b) Overlap per site  $|\langle \psi_{\text{CFT}} | \psi_H \rangle|^{1/N}$  between the same two states. Figure reproduced from [2]. . . . . 85

4.1	Geometry of ansatz wave functions: (a) Jastrow wave function include correlations within all pairs of spins. (b) Matrix Product States (MPS) in 2D cover the lattice in a snake pattern. (c) Entangled Plaquette States (EPS) include all spin correlations within each plaquette (2x2 on the figure) and mediate correlations between distant spins through overlapping plaquettes. (d) String-Bond States (SBS) cover the lattice with many 1D strings on which the interactions within spins are captured by a MPS. Figure reproduced from [4]. . . . .	94
4.2	(a) Boltzmann machines approximate a probability distribution by the Boltzmann weights of an Ising Hamiltonian on a graph including visible units (corresponding to the spins $s_j$ ) and hidden units $h_i$ which are summed over. (b) Restricted Boltzmann machines (here in 2D) only include interactions between the visible and the hidden units. Figure reproduced from [4]. . . . .	97
4.3	(a) A locally connected RBM is an EPS where each plaquette encodes the local connections to a hidden unit. (b) Once expressed as a SBS a fully-connected RBM can be represented by many strings on top of each other. Enlarging the RBM by using non-commuting matrices to non-local SBS induces a geometry in each string. Figure reproduced from [4]. . . . .	102
4.4	Energy difference with the exact ground state energy of a spin-1 extension of a RBM (4.46) with $D = 2$ and different number of strings for the AKLT model on a spin-1 chain with 8 spins. A non-local SBS with non-commuting matrices and one string is exact within numerical accuracy. Figure reproduced from [4]. . . . .	104
4.5	Energy of $H_l$ per site for different optimized ansatz wave functions on a square lattice. The number of parameters ( $N_p$ ) is modified by increasing the bond dimension $D$ (local SBS, $N_p \propto D^2$ ), the size of the plaquettes (EPS, $N_p \propto M_P 2^P$ , where $M_P$ is the number of plaquettes and $P$ is the number of spins in one plaquette), the number of strings $M_S$ (non-local SBS and dSBS, $N_p \propto M_S$ ) or the number of hidden units $M_h$ (RBM, $N_p \propto M_h$ ). (a) 4x4 lattice for which the energy difference with the exact ground state energy is plotted. (b) 10x10 lattice for which the exact ground state energy is unknown and the reference energy of the Laughlin state is indicated as a black line. (c) Optimization of wave functions that have been multiplied by the Laughlin wave function on a 10x10 lattice. The original RBM results are indicated for reference as grey crosses. Figure reproduced from [4]. . . . .	108
4.6	Partition of the lattice used to compute the topological entanglement entropy. Figure reproduced from [4]. . . . .	109

4.7	(a) The spin-spin correlation function between one lattice site (in red) and all other spins on the lattice measured on the optimized l-RBM with lowest energy reveals the antiferromagnetic behavior of the correlations. (b) Decay of the correlations with the distance across the direction indicated in (a) as a white solid line. The error bars are within dot size and finite size effects can already be seen for the last point. Figure reproduced from [4]. . . . .	110
5.1	(a) Undirected graphical model with three maximal cliques depicted in colors (b) Corresponding factor graph (c) Factor graph with hidden units in orange that are marginalized (d) Equivalent tensor network, which is a Matrix Product State. Figure reproduced from [6]. . . . .	116
5.2	Copy operation of a vector input $A_i$ , resulting in a new tensor $B_{ij} = A_i A_j$ , or of a tensor network. Figure reproduced from [6]. . . . .	118
5.3	(a) Restricted Boltzmann Machine (RBM) consisting of visible and hidden variables (b) String-Bond State with 1D geometry generalizing RBM. The legs corresponding to contracted indices in each MPS are depicted in orange for visibility. (c) Short-range RBM with local connections between visible and hidden variables (d) Entangled Plaquette State (EPS) generalizing the short-range RBM. Figure reproduced from [6]. . . . .	119
5.4	Possible geometries of SBS: (a) 2D-SBS consisting of horizontal and vertical overlapping strings. (b) Snake-SBS consisting of 4 overlapping strings in a snake pattern. Figure reproduced from [6]. . . . .	120
5.5	EPS-SBS consisting of a first layer of EPS, followed by a copy operation and a second layer of SBS. Figure reproduced from [6]. . . . .	120
5.6	(a) A classification RBM turns the label into an additional visible unit. (b) The same procedure can be defined for a SBS by adding a node corresponding to the label, and corresponding tensors which connect it to the rest of the tensor network. (c) Generalized tensor networks can be combined with additional layers of neural networks. For example an EPS output is a tensor that can be combined with a linear classifier. Figure reproduced from [6]. . . . .	122
5.7	(a)-(f) Forward pass of contracting an EPS-SBS generalized tensor network. We denote scalar and vectors that are the result of a tensor contraction as black boxes, while other tensors are denoted as empty boxes. The result is a scalar $A \times B$ . (g)-(i) We compute as an example the derivative with respect to the tensor $T_{ijk}$ denoted in blue in (a). The computation can start from stage (d), obtained during the forward propagation. Remaining additional tensors are contracted until we are left with the log-derivative, equal to $\partial \log \text{GTN}(X_1, X_2) / \partial T_{ijk} = \delta_{X_1=j} \delta_{X_2=k} (C_i + D_i) / (AB)$ . Figure reproduced from [6]. . . . .	123

5.8	(a) Real inputs $X_i$ are mapped to a feature vector (here with length two). This vector can then be used as input to a generalized tensor network by contracting it with the open legs of the generalized tensor network. (b) Feature tensors can compress the discretized representation of the inputs $X_i$ to a smaller dimensional space. These tensors can share weights and can be learned as part of the tensor network. Figure reproduced from [6]. . . . .	125
5.9	(a) Dataset with two features $X_1$ and $X_2$ and two classes (depicted in different colors) that cannot be learned by a MPS of bond dimension 2 with features in (5.18). (b) Two normalized features learned by a tensor while classifying the previous dataset with a MPS of bond dimension 2. The features have been discretized in 16 intervals. Using this choice of features the MPS can classify the dataset perfectly. Figure reproduced from [6]. . . . .	126
5.10	(a) Choice of two features in (5.18) for an input taking real values between 0 and 1. (b) Two normalized features learned by a tensor with output dimension 2 combined with a snake-SBS classifying the MNIST dataset. The input features $x$ are the greyscale value of pixels, normalized between 0 and 1 and discretized in 16 intervals. Figure reproduced from [6]. . . . .	126
5.11	Using convolutional Neural Networks as feature vector extractors from real data: the output of the CNN is seen as an image with a third dimension collecting the different features. For each pixel of this image, the vector of features is contracted with the open legs of a tensor network. Figure reproduced from [6]. . . . .	127
5.12	Examples of images from the MNIST (a) and fashion MNIST (b) dataset. Figure reproduced from [6]. . . . .	128
5.13	Test set accuracy of different generalized tensor networks on the MNIST dataset. Figure reproduced from [6]. . . . .	128
5.14	From the raw audio signal, Mel-frequency cepstral coefficients (MFCCs) are extracted over short overlapping windows, resulting in a sequence of high dimensional vectors. These vectors are taken as input to a generalized tensor network. Figure reproduced from [6]. . . . .	130
5.15	Training and testing accuracy of a MPS and a SBS with 4 strings on the UrbanSound8K dataset. The density of parameters is the total number of parameters divided by 174 (the length of the strings). Figure reproduced from [6]. . . . .	130



## List of Tables

3.1	First $(q, \eta)_S$ Moore-Read states. . . . .	73
3.2	Terms in the Hamiltonian $H_{2D}$ and coefficients obtained after numerical optimization on a $4 \times 4$ lattice. . . . .	83
4.1	Energy per site difference with the ground state energy and overlap with the Laughlin state of different ansatz wave functions optimized with respect to the Hamiltonian $H_{\text{parent}}$ on a $6 \times 6$ square lattice with open boundary conditions. sRBM have $M'$ hidden units connected to all spins in each plaquette of size $3 \times 3$ , while RBM have $M$ hidden units connected to all spins of the lattice. . . . .	106
4.2	Topological entanglement entropy (TEE) of the analytical Laughlin state and optimized l-EPS, RBM and l-RBM. . . . .	109
5.1	Test set accuracy of generalized tensor networks and other approaches[254] on the fashion MNIST dataset. . . . .	129



# Acknowledgments

First and foremost, I am deeply indebted to my supervisor Ignacio Cirac for the opportunity of doing research in his group. This thesis has greatly benefited from his guidance and insights, and his extraordinary scientific knowledge and creativity have been a source of inspiration throughout these years. I am especially grateful for his kindness, generous support and his encouragements to develop my own ideas in an outstanding research environment.

I would like to thank Anne Nielsen for her supervision during the first years of my PhD, for introducing me to the physics of strongly correlated quantum many-body systems and for the opportunity to visit her group in Dresden. The many discussions we had on both theoretical aspects as well as technical details of this work, her guidance and patience have been invaluable to me and this thesis benefited from many of her ideas.

I am profoundly grateful to my collaborators, without whom this work would not have been possible. Germán Sierra provided me with the opportunity to learn from his knowledge of the vast landscape of field theories, Moritz August introduced me to the practical implementation of machine learning algorithms and Iván Rodríguez taught me a lot about the fractional quantum Hall effect. My sincere thanks to Nicola Pancotti for the nice atmosphere in our office, for the innumerable discussions we had and for his many contributions to the present work.

These years at the Max Planck Institute of Quantum Optics have been filled with many inspiring interactions and I am very grateful to all former and present members of the theory division for all the things they have taught me and for the time spent together. In particular I have to acknowledge helpful discussions with Hendrik Dreyer, Vedran Dunjko, Alessandro Farace, Yimin Ge, Andrew Goldsborough, Anna Hackenbroich, Benedikt Herwerth, Johannes Kofler, Stefan Kühn, András Molnár, Julian Roos, Tao Shi, Antoine Sterdyniak, Antoine Tilloy, Hong-Hao Tu, Jordi Tura, Thorsten Wahl and Erez Zohar, who all contributed in one way or another to this thesis. I also would like to thank Lucas Clemente for being a great office mate and Andrea Kluth for her help in all administrative matters.

Over the last years, I have travelled to many places and this work was influenced by discussions with great scientists all over the world, I am grateful to all of them.

Most importantly, I owe my deepest gratitude to my family and friends for their support. I would especially like to thank my parents and grand parents for nurturing my curiosity about the world. Finally, I thank Anna Medvedeva for her unconditional love and endless support.



