Physics Area - PhD course in Theory
and Numerical Simulation
of Condensed Matter Physics

Pushing the boundaries of Matrix Product States in quantum many-body physics and computing

Candidate:
Guglielmo Lami

Advisor:
Prof. Mario Collura

Academic Year 2022-2023

SISSA

ABSTRACT

The Thesis explores the application of Matrix Product States (MPS) in the domain of quantum many-body systems and quantum computing, showcasing the versatility and effectiveness of MPS for addressing complex computational challenges. The opening chapter introduces essential tools for Tensor Networks, addressing the curse of dimensionality and elucidating the main characteristics of Matrix Product States and Matrix Product Operators. The discussion encompasses also important numerical techniques, including Density Matrix Renormalization Group and Time Evolving Block Decimation. Furthermore, the relation between MPS and quantum circuits is discussed. The subsequent Chapters delve into specific applications of MPS. In the second Chapter, we introduce a Tensor Network ansatz inspired by the backflow transformation for correlated systems. This extension of the MPS representation ensures an area law for entanglement in dimensions one or greater. We employ an optimization scheme combining DMRG and variational Monte Carlo algorithms for efficient ground-state search. Benchmarking against spin models demonstrates high accuracy. The third Chapter explores quantum annealing for optimizing complex classical spin Hamiltonians, as the Hopfield model and the binary perceptron. We introduce an efficient Tensor Network representation for the adiabatic time evolution of quantum annealing, enabling scalable classical simulations. The use of MPS in mitigating Trotter errors and mapping to quantum circuits is also explored. In the fourth Chapter we present a novel method for evaluating the amount of nonstabilizerness (also known as quantum magic) contained in an MPS. We overcome the exponentially hard evaluation of this quantity by employing a simple perfect sampling technique of Pauli string. Our innovative MPS approach enables efficient computation. Benchmarked on magic states and the quantum Ising chain ground state, this method offers easy access to the non-equilibrium dynamics of nonstabilizerness following a quantum quench. The Appendices provide additional insights, including: an introduction to variational Monte Carlo, technical details on MPS simulations for the quantum annealing and a comprehensive overview on stabilizer formalism and nonstabilizerness.

CONTENTS

PREFACE

Quantum mechanics is approximately one hundred years old. At the beginning of the twentieth century, while the European powers were preparing for the impending apocalyptic carnage of the first world war and a generation of talented avant-gardists, like Stravinskij and Picasso, were revolutionizing the arts, a series of groundbreaking works by Max Planck, Albert Einstein and Niels Bohr remodeled physics [1]. Their pioneering efforts marked a major departure from the classical physics and revealed for the first time a fundamental aspect of nature, namely that certain physical quantities, such as energy or angular momentum, can take only a *quantized* set of values, at least within the realm of atoms and elementary particles. The war could not stop the driving force of such tremendous innovations. A new and more comprehensive formulation of quantum mechanics was developed in the mid 1920's thanks to the achievements of some young minds. Werner Heisenberg, Max Born, Pascual Jordan were the first to come up with the idea of representing physical observables using matrices, leading to the development of the fundamental matrix formalism. Concurrently, Erwin Schrödinger formulated a wave equation describing the time evolution of any quantum mechanical particle. Max Born was the first to suggest the interpretation of the square of the wave function amplitude as a *probability density*, i.e. the probability of detecting a system in a specific state upon measurements [2]. Other seminal contributions were made by John von Neumann, who formalized quantum mechanics in a rigorous mathematical framework, and Paul Dirac, who made substantial advancements in the quantum matrix theory and wrote down the relativistic wave equation for the electron. In this period, several enigmatic aspects of quantum mechanics came to light, including wave-function collapse and *entanglement*. These phenomena appeared so puzzling that even some of the historical founders of the quantum theory started to consider it as unphysical or incomplete [3]. Nevertheless, numerous experiments conducted over the decades have consistently corroborated all the core elements of the theory. As a result, quantum mechanics is now universally recognized as one of the main achievement in humanity's understanding of the natural world [4]. In a practical perspective, it can be viewed as a very powerful tool enabling us to make systematic predictions about the physical phenomena with an unprecedented level of accuracy[1]. However, certain aspects of the quantum mechanics have been experimentally inaccessible for long time. As Schrödinger wrote in 1952 [5]:

> *[...] we never experiment with just one electron or atom or (small) molecule. In thought-experiments we sometimes assume that we do;*

1 As an example, calculations in the framework of the quantum theory of the electromagnetic field yields to a prediction for the gyromagnetic ratio of the electron that has been confirmed by experimental measurements with a precision of less that one part over 10^{12}. To appreciate how incredibly accurate is this result, it should be noted that is like being able to predict the distance between the earth and the moon with an error of only a fraction of mm!

> *this invariably entails ridiculous consequences [...] we are not experimenting with single particles, any more than we can raise Ichthyosauria in the zoo.*

In general, the experimental observation and replication of any phenomenon that relies on preserving quantum coherence proved to be almost unfeasible. Indeed, the delicate nature of this quantum property leads to rapid dissipation when a quantum system interacts with an extensive and uncontrollable macroscopic environment.

However, things have changed. While realizing experiments with one or few particles was not even imaginable some decades ago, nowadays it is not only possible but it is also a common tool to investigate the fundamental aspects of quantum mechanics. In a nutshell, we are now able to observe and manipulate systems consisting only of few "quanta" of matter or light [2]. A partial list of currently used quantum platforms include superconducting qubits (quantum bits) [6, 7], arrays of Rydberg atoms [8, 9], ultra-cold atoms in optical lattices [10], trapped ions [11], quantum dots [12] and cavity QED setups [13]. These experiments open the doors to a renovated and more profound exploration of the quantum world, commonly referred as *second quantum revolution* [14].

In this scenario, research on quantum physics and technologies has experienced a significant growth. A prominent field of investigation is quantum computing, namely the possible application of quantum systems to enhance or speed up classical computations. Richard Feynman's seminal speech at the Conference on the Physics of Computation in 1981 can be considered as one of the earliest conceptual advancements in this direction [15].

> *Now I explicitly go to the question of how we can simulate with a computer–a universal automaton or something–the quantum-mechanical effects. [...] the full description of quantum mechanics for a large system with* R *particles is given by a function* $\psi(x_1, x_2, ..., x_R, t)$ *which we call the amplitude to find the particles at* $x_1, x_2, ..., x_R$ *and therefore, because it has too many variables, it cannot be simulated with a normal computer with a number of elements proportional to* R *[...] Can you do it with a new kind of computer, a quantum computer? Now it turns out, as far as I can tell, that you can simulate this with a quantum system, with quantum computer elements.*

Feynman argued that simulating quantum mechanical systems inherently requires an exponential computational cost. Hence, he introduced the idea of employing quantum computers for simulating physical phenomena [3].

Nowadays, quantum computing has become a well established field of study, built upon the theoretical principles provided by quantum information theory. Several quantum algorithms have been developed, offering techniques to enhance classical computations. One notable example is Shor's

2 Unfortunately, we still cannot raise Ichthyosauria in the zoo...

3 A similar intuition was reached independently and apparently earlier, in 1980, by the soviet mathematician Yuri Manin.

Pictures from first and second quantum revolutions 1) Einstein receiving the Max Planck medal from Planck (1929). 2) Fermi, Heisenberg, Bohr at the "Convegno di fisica nucleare" (Rome, 1931). 3) Article headline of The New York Times regarding the EPR paper (1935). 4) The line spectrum of hydrogen (from G.Herzberg, Annalen der Physik, 1927). 5) A postcard from Walther Gerlach to Niels Bohr, showing first results of the celebrated Stern–Gerlach experiment (1922). 6) Charles Townes in front of first prototype of maser, forerunner of the laser (1953). 7) Richard Feynman (1980s). 8) Peter Shor, father of the quantum factorization algorithm (developed in 1994). 9) Ignacio Cirac, trailblazer of quantum computing and tensor networks techniques (1990s-2000s). 10) Alain Aspect, a pioneer in entangled photons experiments (in early 1980s). 11) Rainer Blatt with an experimental apparatus utilizing an ion trap (2000s). 12) The IBM superconducting quantum processor (2022).

factorization algorithm [16], which provides exponential speed up over the classical counterpart. Despite these successes, currently available quantum platforms suffer from strong noise issues, leading to short coherence times. As a result, achieving universal fault-tolerant quantum computation is impossible in the near future. We currently live in the era known as Noisy Intermediate-Scale Quantum (NISQ) [17], characterized by devices featuring a moderate quantity of qubits, usually ranging from tens to a few hundred. NISQ machines and methods are promising for tasks such as simulating quantum systems, optimization and machine learning. An example is given by Quantum Annealing (QA), an optimization method designed to find the minimum of a complex Hamiltonian by leveraging quantum fluctuations and tunneling effects generated by a time-dependent external field, similar to how simulated annealing uses thermal fluctuations. However, it remains uncertain whether methods like this will lead to a measurable quantum advantage.

In parallel with the outlined development of quantum information and computing, last decades have seen a substantial increase of our comprehension in quantum many-body physics. A crucial focus has revolved around understanding and examining the *entanglement properties* of quantum systems, both in equilibrium and out-of-equilibrium. In particular, entanglement has been used as a general and powerful tool for identifying and characterizing different phases, or phase transitions. This led to the development of new theoretical frameworks, such as the Cardy-Calabrese formula which relates the entanglement entropy of a one-dimensional critical quantum system to its conformal field theory description [18]. Besides, a series of works inspired by the physics of black holes and the holographic principle introduced the notion of *area law* for entanglement entropy [19–21]. In particular, it was shown that in 1D quantum systems, the ground-state entanglement entropy between a block of l spins and its complement is bounded by a constant in noncritical (gapped) regime [21], whereas it grows as $\log l$ in correspondence of a critical point. What makes the area law intriguing? One of the primary reasons of interest comes from a very practical situation: the numerical simulation of quantum many-body systems. Indeed, when a many-body state has limited entanglement, i.e. small quantum correlations between its constituents, the possibility of finding a suitable representation in terms of a reduced number of parameters becomes plausible. In this way one could circumvent the "curse of dimensionality" or "exponential wall" [22] that restricts exact treatments of quantum many-body systems to small system sizes. This observation is at the foundation of Tensor Networks methods. The origins of these techniques can be traced back to the early works of White in the 1990s, who introduced the density matrix renormalization group (DMRG) method [23, 24]. DMRG was initially designed as a numerical algorithm to determine the ground state of one-dimensional quantum spin chains by following a sort of renormalization flow of the density matrix of a block of spins. Over time, the concept behind DMRG was expanded and generalized, leading to the development of the *Matrix Product States* (MPS) [25–27], a powerful ansatz for 1D quantum systems. The value of the MPS many-body wave function for a particular system configuration

is obtained as a product of matrices (vectors at the boundaries) that depend only on a local degree of freedom, such for example a single spin. MPS today allow for the accurate simulation of strongly correlated quantum systems in 1D, both at thermal equilibrium and out-of-equilibrium, thanks to a variety of numerical methods and algorithms developed *ad hoc*. This has ensured undisputed success for MPS in various fields, such as quantum many-body physics, quantum computing, condensed matter physics and quantum chemistry [28]. The study of out-of-equilibrium dynamics is often carried out in numerical experiments of "quantum quench", by preparing the system in a certain state $|\psi_0\rangle$ and letting it evolve from $t = 0$ with the unitary dynamics generated by a specific Hamiltonian $\hat{H}$. In other words, it involves calculating the time-evolved state $|\psi(t)\rangle = e^{-i\hat{H}t}|\psi_0\rangle$. The simplest MPS method developed for this type of setup is Time Evolving Block Decimation, a straightforward algorithm in which the unitary dynamics is reduced to finite and discrete time steps, and the MPS matrices are updated using local unitary matrices. The intrinsic increase in the complexity of the time-evolved state, witnessed by the growth of its entanglement entropy, leads to an increase in the size of MPS matrices, which is usually exponential in time t. MPS thus encounter a sort of entanglement barrier, which, in this context, represents a manifestation of the exponential wall of quantum many-body systems. If a complete description of the many-body (micro)state at long times t is likely impossible to obtain, however, at least in principle, it should be possible to obtain a sort of effective MPS description (macrostate) where the values of some physically relevant local observables (for example, one or two-point correlation functions) are accurately reproduced even if one discard a certain irrelevant part of the information contained in $|\psi(t)\rangle$ [29]. Another constraint of MPS methods arises from the system's dimensionality, since the area law in dimensions greater than 1 still entails the need to encode entanglement entropy quantities that scale extensively with the linear dimensions of the system, requiring exponentially large MPS matrices. Despite these limitations, MPS remains one of our most robust and powerful computational tools in the NISQ era. They are likely going to help us, among other things, in understanding many-body quantum dynamics, analysing quantum circuits and protocols, exploring intrinsic resources of the quantum states.

This work presents some results proposed during the years of my doctoral studies that contribute to various aspects of the MPS numerical simulation techniques. As we will discuss, these advancements concern both aspects of quantum computing and aspects of quantum many-body systems.

PUBLICATIONS

This Thesis includes material published in the following articles.

1. "Matrix product states with backflow correlations", **Guglielmo Lami**, Mario Collura, Giuseppe Carleo (Phys. Rev. B 106, L081111)

2. "Quantum annealing for neural network optimization problems: A new approach via tensor network simulations", **Guglielmo Lami**, Pietro Torta, Giuseppe E. Santoro, Mario Collura, (SciPostPhys.14.5.117)

3. "Quantum Magic via Perfect Pauli Sampling of Matrix Product States", **Guglielmo Lami**, Mario Collura, (Phys. Rev. Lett. 131, 180401)

4. "Quantifying non-stabilizerness through entanglement spectrum flatness', Emanuele Tirrito, Poetri Sonya Tarabunga, **Guglielmo Lami**, Titas Chanda, Lorenzo Leone, Salvatore F. E. Oliviero, Marcello Dalmonte, Mario Collura, Alioscia Hamma, arXiv: (arXiv:2304.01175)

Instead the content of the paper "Continuously Monitored Quantum Systems beyond Lindblad Dynamics", **Guglielmo Lami**, Alessandro Santini, Mario Collura (arXiv:2305.04108) is not discussed here.

Other works are currently in preparation, focusing on possible extensions of articles 2) and 3).

SUMMARY

In Chapter 1, we provide an extensive Introduction to Matrix Product States and related numerical techniques.

In Chapter 2, we explore how Matrix Product States can be applied to simulate 2D quantum systems, utilizing a generalized ansatz referred to as Matrix Product Backflow States.

In Chapter 3, we discuss the application of Matrix Product States to the simulation of the Quantum Annealing process for a classical optimization problem of particular importance: the binary perceptron. To this end, we develop an ad hoc MPS method. We also draw general conclusions about the potential relevance and utility of MPS projection on states with bounded entanglement entropy in order to effectively address optimization problems.

In Chapter 4, we introduce a new method to measure the nonstabilizerness, also dubbed quantum magic, of a Matrix Product State. The method is based on a new sampling Algorithm in the Pauli basis.

In Appendix A, we present a concise overview of Variational Monte-Carlo techniques, which have been employed for optimizing the Matrix Product Backflow States discussed in Chapter 2. In Appendix B, we provide supplementary related to the study of Quantum Annealing with MPS in Chapter 3. Finally, Appendix C features an extensive introduction to the stabilizer formalism, while also elaborating on the concept of nonstabilizerness. Additionally, we present original results that shed light on the relationship between nonstabilizerness and entanglement.

LIST OF ABBREVIATIONS AND ACRONYMS USED IN THE TEXT

Acronym	Full Form
SVD	Singular Value Decomposition
TN	Tensor Network
MPS	Matrix Product States
MPO	Matrix Product Operator
DMRG	Density Matrix Renormalization Group
TEBD	Time Evolution Block Decimation
PEPS	Projected Entangled Pair States
EPS	Entangled Plaquette States
NN	Neural Network
RBM	Restricted Boltzmann Machine
MPBS	Matrix Product Backflow States
VMC	Variational Monte-Carlo
SGD	Stocastic Gradient Descent
OBC	Open Boundary Conditions
PBC	Periodic Boundary Conditions
QA	Quantum Annealing
dQA	digitized Quantum Annealing
VQA	Variational Quantum Algorithms
QAOA	Quantum Approximate Optimization Algorithm
ED	Exact Diagonalization
DFT	Discrete Fourier Transform
SRE	Stabilizer Rényi Entropies

1 INTRODUCTION

> As soon as the systems begin to influence each other, the combined function ceases to be a product [...] Thus one disposes provisionally (until the entanglement is resolved by an actual observation) of only a common description of the two in that space of higher dimension.
>
> E.Schrödinger,"The Present Situation in Quantum Mechanics" (1935)

In this introductory Chapter, we review some of the main aspects concerning Matrix Product States (MPS) and Matrix Product Operators (MPO), two tools that we will use extensively throughout the rest of the work. We will discuss their principles and practical applications as powerful numerical techniques studies to address the many-body quantum problem.

1.1 THE CURSE OF DIMENSIONALITY

We consider a quantum many-body system consisting of N qubits or spin-1/2 [1]. Each qubit, or particle, is described by a local Hilbert space $\mathcal{H}_i \sim \mathbb{C}^2$, which is spanned by a certain orthonormal basis $\{|0\rangle, |1\rangle\}$[2]. The total Hilbert space is obtained as the tensor product of the local Hilbert spaces

$$\mathcal{H} = \bigotimes_{i=1}^{N} \mathcal{H}_i \ . \tag{1.1.1}$$

A generic quantum state $|\psi\rangle \in \mathcal{H}$ can be decomposed as

$$|\psi\rangle = c_{s_1 \dots s_N} |s_1 \dots s_N\rangle \ , \tag{1.1.2}$$

where we adopted Einstein's convention on the repeated indices $s_i \in \{0, 1\}$. This representation offers a full description of the wave function $|\psi\rangle$. However, it requires to store all the components $c_{s_1 \dots s_N}$, that means 2^N complex numbers. Thus, the storage of a double-precision wave function requires approximately 17GB of memory for a system of size $N = 30$, whereas one

1 Sometimes, we will use the symbol d to denote the local Hilbert space dimension ($d = 2$ for the standard case of qubits).

2 A common choice is to identify the basis states $|0\rangle$ and $|1\rangle$ with the two eigenstates of the Pauli matrix $\hat{\sigma}^z$ up $|\uparrow\rangle$ and down $|\downarrow\rangle$ respectively.

needs $\approx 18\text{PB} = 18 \cdot 10^6\text{GB}$ for $N = 50$! Similarly, processing $|\psi\rangle$, for instance to compute a matrix element, requires an exponentially large number of operations. As a matter of fact, simulating quantum systems of size larger than $N \approx 30$ with exact numerical techniques is impossible [3]. This curse of dimensionality, or exponential wall [22], reflects the vastness of the many-body Hilbert space. Notice however that, if we had a product state, i.e. a state for which the components can be factorized as

$$c_{s_1 \dots s_N} = c^{[1]}(s_1) \dots c^{[N]}(s_N)\,, \tag{1.1.3}$$

the storage would not be challenging, since it would require to memorize only 2 complex numbers per site, i.e. $2N$ complex numbers in total. Hence, the computational problem of storing and manipulating quantum many-body states is somehow related to the intrinsic quantum correlations between the constituents (particles or spins), i.e. to the *entanglement*. For typical quantum many-body states, nearly all the information is encoded in the correlations among the constituents [30]. We aim to describe strongly correlated states in which the simple mean-field approximation of Eq. 1.1.3 is inadequate. This point will be discussed carefully in next Sections. In any case, it is clear that, to investigate larger systems and aim to approach the thermodynamic limit $N \to \infty$, it is necessary an alternative representation of the wave function $|\psi\rangle$ that allows for less data-intensive approximations.

Tensor Networks (TN) provide a possible solution. The goal is to represent the tensor $c_{s_1 \dots, s_N}$, which contains the coefficients of the many-body wave function, as a generic network of tensors, where each tensor corresponds to a local degree of freedom (such as the spin s_i). The entanglement between neighboring particles is captured by connecting the tensors, i.e. contracting them along auxiliary (unphysical) indices. This structure often impose a cut-off on the amount of entanglement between different regions of the system.

As an introductory example, let us consider a system of only $N = 2$ spins in the state $|\psi\rangle = \big(|00\rangle + |11\rangle\big)/\sqrt{2}$. Clearly, the 4 coefficients $c_{s_1 s_2}$ of the wave function cannot be expressed in the factorized form of Eq. 1.1.3, since $|\psi\rangle$ is not a product state (i.e. it is entangled). However, it should be noted that if we define the following two basis vectors

$$\boldsymbol{e}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \boldsymbol{e}(1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.1.4}$$

we will have

$$c_{s_1 s_2} = \frac{1}{\sqrt{2}} \big(\boldsymbol{e}^{[1]}(s_1)\big)^{\mathsf{T}} \cdot \boldsymbol{e}^{[2]}(s_2)\,. \tag{1.1.5}$$

It is therefore possible to rewrite c in a kind of generalization of Eq. 1.1.3 in which the $c(s_i)$ are replaced by $\boldsymbol{e}(s_i)$ $(i = 1, 2)$. In other words, we have explicitly factorized the wave function $|\psi\rangle$ using an additional local dimension that has transitioned from 1, in the case of scalar $c(s_i)$, to 2 for the vectors $\boldsymbol{e}(s_i)$. Eq. 1.1.5 is a simple example of a TN state. The local tensors $\boldsymbol{e}(s_1)$, $\boldsymbol{e}(s_2)$ are contracted by means of the scalar product to obtain the actual wave function value.

3 Except for the cases in which some symmetries can be exploited to reduce the dimension of the Hilbert space.

In the following Sections we will review the basics of the most used class of TN, namely the Matrix Product States. Before of that, we introduce some important tools commonly used in such a framework: the Singular Value Decomposition, the Schmidt decomposition and the graphical representation of tensors.

1.2 TENSOR NETWORKS TOOLS

1.2.1 Singular Value Decomposition (SVD)

Singular Value Decomposition is a crucial result in linear algebra stating that it is possible to transform any matrix into a diagonal form by employing the appropriate basis in the domain and in the image spaces [31]. To be more precise, let M be a complex matrix of dimension $m \times n$. Then M can be always decomposed as [31, 32]

$$M = U\Lambda V \,, \tag{1.2.1}$$

where

- U is a matrix of dimension $m \times \min(m,n)$ with orthonormal columns, $U^\dagger U = \mathbb{1}$;
- Λ is a diagonal matrix of dimension $\min(m,n) \times \min(m,n)$ with real positive entries that can be ordered as follows $\Lambda_1 \geqslant \Lambda_2 \geqslant ... \geqslant \Lambda_r > 0$;
- V is a matrix of dimension $\min(m,n) \times n$ with orthonormal rows, $VV^\dagger = \mathbb{1}$.

The real numbers $\Lambda_1 \geqslant \Lambda_2 \geqslant ... \geqslant \Lambda_r$ are called *singular values* of M. In this context, the *rank* r of M is defined as the number of non-zero singular values. The SVD decomposition is unique up to multiplication by complex phases of the columns of U (rows of V) or permutations of singular vectors having the same singular values [33]. The computational complexity of the numerical algorithms for SVD is $o(mn^2)$, if $m \geqslant n$.

The Eckart-Young-Mirsky theorem states that, fixed a matrix $M \in \mathbb{C}^{m\times n}$, for any unitarily invariant matrix norm $\|\cdot\|$, the minimum of $\|M - M'\|$ between all the matrices $M' \in \mathbb{C}^{m\times n}$ of rank r' is obtained by choosing

$$M' = U\Lambda' V \qquad \Lambda' = \mathrm{diag}\big(\Lambda_1, \Lambda_2, ..., \Lambda_{r'}, 0, ...0\big) \,, \tag{1.2.2}$$

i.e. by truncating the list of singular values to the first r' elements [34, 35]. An unitarily invariant norm is a matrix norm $\|\cdot\|$ such that $\|M\| = \|U^\dagger M V\|$ for all the unitary matrices U, V. An example is the Frobenius norm $\|M\| = \sqrt{\mathrm{Tr}(MM^\dagger)}$, which can be rewritten as $\|M\| = \sqrt{\mathrm{Tr}(\Lambda^2)}$ (i.e. the Frobenius norm square is equal to the sum of the squares of the singular values). This result is highly relevant and widely used in various fields to achieve dimensionality reduction, data approximation, etc. In Fig. 1.1, we show a possible application of the SVD truncation (Eq. 1.2.2) to image compression.

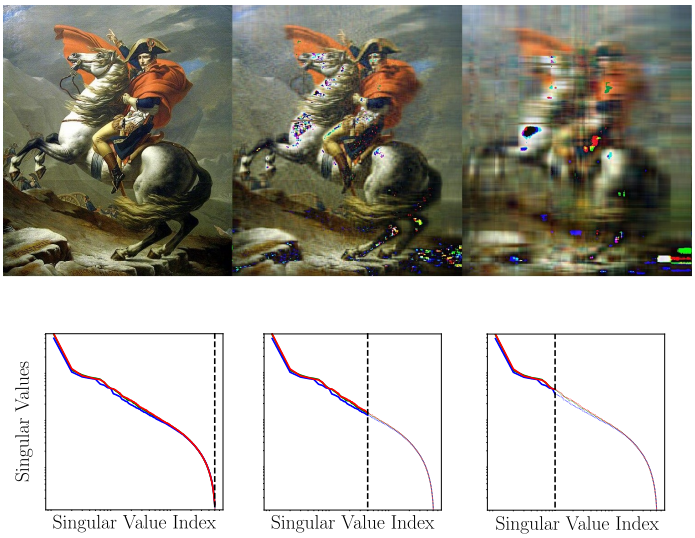

Figure 1.1: SVD can be used to compress images. Red, green, blue channels of the original picture are converted to matrices and SVD is applied discarding singular eigenvalues smaller than a given threshold (plots below). Afterwards, the image is reconstructed (figure above).

Finally, let us mention that there exists a closely related matrix factorization named QR decomposition. It states that we can factor any complex $m \times n$ matrix M as $M = QR$, where Q is a matrix of dimension $m \times \min(m, n)$ with orthonormal columns, $Q^\dagger Q = \mathbb{1}$, and R is upper-triangular.

1.2.2 Schmidt decomposition and Entanglement

One fundamental application of SVD in quantum mechanics is the Schmidt decomposition [36]. Let us consider a state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, where A, B are two subsystems forming a bipartition and $\mathcal{H}_A, \mathcal{H}_B$ are the corresponding Hilbert spaces. The state can be decomposed as

$$|\psi\rangle = c_{ij} \, |a_i\rangle \, |b_j\rangle \ , \tag{1.2.3}$$

where $|a_i\rangle \in \{|a_1\rangle, |a_2\rangle, ..., |a_m\rangle\}$, $m = \dim(\mathcal{H}_A)$ and $|b_j\rangle \in \{|b_1\rangle, |b_2\rangle, ..., |b_n\rangle\}$, $n = \dim(\mathcal{H}_B)$. c is a complex matrix of dimension $m \times n$, therefore we can use SVD to get $c = U\Lambda V$. We obtain

$$|\psi\rangle = U_{ik}\Lambda_k V_{kj} \, |a_i\rangle \, |b_j\rangle = \Lambda_k \, |a'_k\rangle \, |b'_k\rangle \ , \tag{1.2.4}$$

with $k = 1, 2... \ \min(m, n)$ and

$$|a'_k\rangle = U_{ik} \, |a_i\rangle \qquad |b'_k\rangle = V_{kj} \, |b_j\rangle \ . \tag{1.2.5}$$

These new states are orthonormal, in fact

$$\begin{aligned} \langle a'_l | a'_k \rangle &= U_{ik} U^*_{jl} \, \langle a_j | a_i \rangle = (U^\dagger U)_{lk} = \delta_{lk} \\ \langle b'_l | b'_k \rangle &= V^*_{li} V_{kj} \, \langle b_i | b_j \rangle = (VV^\dagger)_{kl} = \delta_{lk} \ . \end{aligned} \tag{1.2.6}$$

Eq. 1.2.4 is known as Schmidt decomposition. The normalization of the state $|\psi\rangle$ implies that

$$\langle \psi | \psi \rangle = \mathrm{Tr}\left[c^\dagger c\right] = \|c\|^2 = \sum_{k=1}^{r} \Lambda_k^2 = 1 \ , \tag{1.2.7}$$

where $r \leqslant \min(m, n)$ is the rank of c. Because of Eckart-Young-Mirsky theorem, the state

$$|\psi'\rangle = U_{ik}\Lambda'_k V_{kj} |a_i\rangle |b_j\rangle \qquad \Lambda' = \mathrm{diag}(\Lambda_1, \Lambda_2, ..., \Lambda_{r'}, 0, ...0) \qquad r' \leqslant r \tag{1.2.8}$$

is the best approximation of $|\psi\rangle$ between all the rank r' states[4]. The overlap between the truncated state and the original state is

$$\begin{aligned} \langle\psi|\psi'\rangle &= (U_{ik}\Lambda_k V_{kj})^* U_{il}\Lambda'_l V_{lj} = (U^\dagger U)_{kl}(VV^\dagger)_{lk}\Lambda_k\Lambda'_l = \\ &= \sum_{l=1}^{r'} (\Lambda'_l)^2 = 1 - \sum_{k=r'+1}^{r} (\Lambda_k)^2 \end{aligned} \tag{1.2.9}$$

and thus is determined by the sum square of all the discarded Schmidt eigenvalues. Furthermore, if we consider an operator $\hat{O}_A$ acting only on $\mathcal{H}_A$, we have

$$\langle\psi|\hat{O}_A|\psi\rangle = \sum_{k=1}^{r} \Lambda_k^2 \langle a'_k|\hat{O}_A|a'_k\rangle \qquad \langle\psi'|\hat{O}_A|\psi'\rangle = \sum_{k=1}^{r'} \Lambda_k^2 \langle a'_k|\hat{O}_A|a'_k\rangle$$

and

$$\begin{aligned} \delta &= |\langle\psi|\hat{O}_A|\psi\rangle - \langle\psi'|\hat{O}_A|\psi'\rangle| = \\ &= \sum_{k=r'+1}^{r} \Lambda_k^2 \langle a'_k|\hat{O}_A|a'_k\rangle \leqslant c \sum_{k=r'+1}^{r} \Lambda_k^2 \end{aligned} \tag{1.2.10}$$

where c is a constant representing an upper bound for the expectation value of the (bounded) operator $\hat{O}_A$. Hence, the error δ in the expectation value is on the order of the truncated weights. When these weights are sufficiently small, we obtain a truncated state $|\psi'\rangle$, which is a satisfactory approximation of $|\psi\rangle$ even with regard to the expectation values of physical observables. [5]

The Schmidt decomposition of a pure state $|\psi\rangle$ holds significance as it also provides immediate insight into the *entanglement* properties of the state. One crucial way to evaluate entanglement is the von Neumann entanglement entropy, which is defined as $S = -\mathrm{Tr}[\hat{\rho}_A \log \hat{\rho}_A]$, where $\hat{\rho}_A$ is the reduced density matrix of the subsystem A [36]. Such quantity is known to be an entanglement monotone, meaning that is not increasing under a large class of quantum transformation, namely the Local Operations and Classical Communication (LOCC) [36]. Given a Schimdt decomposition, one can easily obtain $\hat{\rho}_A$ as

$$\hat{\rho}_A = \mathrm{Tr}_B[|\psi\rangle\langle\psi|] = \Lambda_k^2 |a'_k\rangle\langle a_k| ,$$

and therefore the von Neumann entropy is the Shannon entropy associated to the squared Schmidt eigenvalues, i.e.

$$S = -\mathrm{Tr}[\hat{\rho}_A \log \hat{\rho}_A] = -\Lambda_k^2 \log(\Lambda_k^2) . \tag{1.2.11}$$

In general, the whole spectrum of the Λ_k^2, often called *entanglement spectrum* can be used to characterize the entanglement properties of the state.

4 After the truncation, $|\psi'\rangle$ should be normalized to 1 again.

5 However, it is important to note that linear functions in $|\psi\rangle$ ($|\psi'\rangle$) can potentially be affected by these approximations. Specifically, the error on $\langle\phi|\psi\rangle$ is: $\delta = |\langle\phi|\psi\rangle - \langle\phi|\psi'\rangle| \leqslant c\sum_{k=r'+1}^{r} |\Lambda_k|$, where c represents an upper bound on $\langle\phi|a_k\rangle|b_k\rangle$. However, the 1-norm $\sum_{k=r'+1}^{r} |\Lambda_k|$ is not bounded by the 2-norm $\sum_{k=r'+1}^{r} \Lambda_k^2$, and thus δ can in principle be significant.

1.2.3 Tensor Network graphical representation

In the context of Tensor Network theory is particularly useful to use a graphical notation [4, 37–39]. A generic tensor $\mathcal{T}_{i_1 i_2 \dots i_k}$ with k indices is represented as a shape with k outgoing lines (legs), each representing an index (see Fig. 1.2a)). The identity matrix $\mathbb{1}_{ij} = \delta_{ij}$ is depicted by a single continues line. The contraction of two tensors $\mathcal{T}, \mathcal{T}'$ is obtained by connecting the legs representing the indices that are summed over (Fig. 1.2a)).

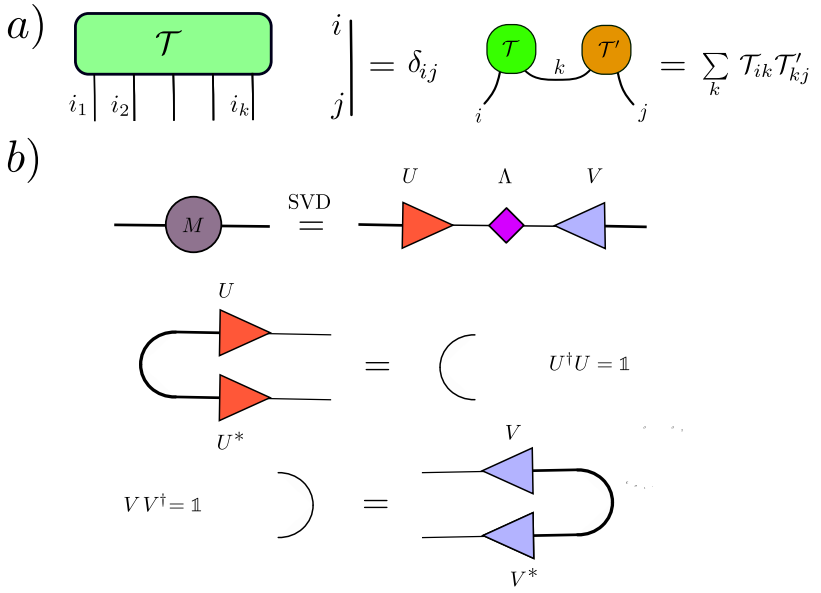

Figure 1.2: a) Graphical representations of a generic tensor $\mathcal{T}_{i_1 i_2 \dots i_k}$, of the identity matrix and of the contraction (product) of two matrices $\mathcal{T}, \mathcal{T}'$. b) The SVD decomposition of a generic complex matrix M. Different line thickness represent different dimensions of the matrices.

In such a framework, the SVD decomposition can be represented as in Fig. 1.2b). Notice that the triangular shape of U, V specify the direction of contraction giving rise to the identity when contracting with $U^\dagger, V^\dagger$.

1.3 MATRIX PRODUCT STATE (MPS) DECOMPOSITION

In this Section, we show that any many-body state $|\psi\rangle$ can be expressed in a specific form known as a Matrix Product State. The procedure is based on a series of SVDs. Let us consider the wave function tensor $c_{s_1 \dots s_N}$ (Eq. 1.1.2) and apply the following steps [27].

- We reshape c as a matrix $\psi^{[1]}$ of dimension $2 \times 2^{N-1}$, by grouping together the last $N-1$ indices:

$$\psi^{[1]}_{s_1,(s_2 \dots s_N)} = c_{s_1 \dots s_N} \ . \tag{1.3.1}$$

  Next, we apply the SVD obtaining

$$\psi^{[1]}_{s_1,(s_2 \dots s_N)} = U^{[1]}_{s_1 a_1} \Lambda^{[1]}_{a_1} V^{[1]}_{a_1,(s_2 \dots s_N)} = A^{[1]}(s_1)_{a_1} \psi^{[2]}_{(a_1 s_2),(s_3 \dots s_N)} \ . \tag{1.3.2}$$

For the second equality, we introduced the row vectors $A^{[1]}(s_1)$, that are labelled by the index s_1 and have entries $A^{[1]}_{a_1}(s_1) = U^{[1]}_{s_1 a_1}$. We also reshaped $\Lambda^{[1]}_{a_1} V^{[1]}_{a_1,(s_2 \ldots s_N)}$ into a new matrix $\psi^{[2]}_{(a_1 s_2),(s_3 \ldots s_N)}$ of dimension $2r_1 \times 2^{N-2}$, where r_1 is the rank resulting from the first decomposition.

- We apply again SVD to the new matrix $\psi^{[2]}$, getting

$$\psi^{[2]}_{(a_1 s_2),(s_3 \ldots s_N)} = U^{[2]}_{(a_1 s_2),a_2} \Lambda^{[2]}_{a_2} V^{[2]}_{a_2,(s_3 \ldots s_N)} = A^{[2]}_{a_1 a_2}(s_2)\, \psi^{[3]}_{(a_2 s_3),(s_4 \ldots s_N)} . \tag{1.3.3}$$

  The size of the new matrix $\psi^{[3]}$ is $2r_2 \times 2^{N-3}$.

- We iterate the first two steps over the next sites $i = 3, \ldots, N$. At the final step we get

$$c_{s_1 \ldots s_N} = A^{[1]}_{a_1}(s_1) A^{[2]}_{a_1 a_2}(s_2) \ldots A^{[N]}_{a_{N-1}}(s_N) \, . \tag{1.3.4}$$

The steps are represented with the graphical notation in Fig. 1.3 a), whereas in b) we show the final result.

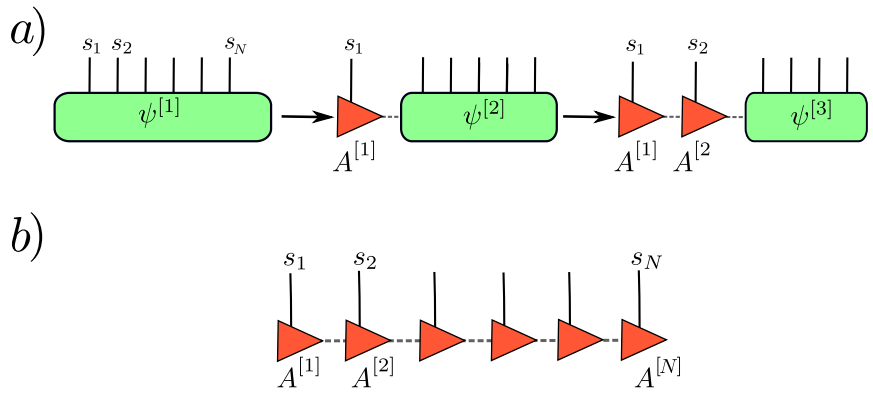

Figure 1.3: a) The outlined decomposition algorithm b) The final Matrix Product State. Dotted grey lines represent auxiliary unphysical indices, which are fully contracted.

The shape of matrices $A^{[1]} \ldots A^{[N]}$ is

$$1 \times r_1, \; r_1 \times r_2, \; r_2 \times r_3, \; \ldots, r_N \times 1 \, , \tag{1.3.5}$$

and therefore we can write the tensor c as a product of spin dependent matrices (vectors at the boundaries)

$$c_{s_1 \ldots s_N} = A^{[1]}(s_1) A^{[2]}(s_2) \ldots A^{[N]}(s_N) \, . \tag{1.3.6}$$

Frequently, the indices [1], [2]... [N] are omitted, yet it is important to be aware that the local tensors A depends on the site. By construction, the ranks r_j obey to the following inequalities

$$r_1 \leqslant 2, \qquad r_2 \leqslant 2r_1 \leqslant 2^2, \quad \ldots,$$
$$r_{N/2} \leqslant 2^{N/2}, \qquad r_{N/2+1} \leqslant 2^{N/2-1}, \qquad \ldots \, r_N \leqslant 2 \, .$$

The number of entries of the matrices A grows exponentially with N, since no information has been discarded in the construction. However, as we will discuss, in certain cases we can bound the dimension of the matrices A without significantly compromising the accuracy of the many-body state description.

Let us observe that, because of the unitarity of the matrices U, we have the following relationship

$$\begin{aligned}\delta_{\alpha_j\alpha'_j} &= (U^\dagger)_{\alpha_j,(\alpha_{j-1},s_j)} U_{(\alpha_{j-1},s_j),\alpha'_j} = \\ &= A^\dagger_{\alpha_j\alpha_{j-1}}(s_j) A_{\alpha_{j-1}\alpha'_j}(s_j) = \left(A^\dagger(s_j)A(s_j)\right)_{\alpha_j\alpha'_j} .\end{aligned} \tag{1.3.7}$$

Therefore, each set of matrices A obeys to the condition

$$A^\dagger(s_j)A(s_j) = \mathbb{1} . \tag{1.3.8}$$

Given our result (Eq. 1.3.6), we can attempt to reconstruct a Schmidt decomposition. We will divide the system into two parts: A, spanning from site 1 to site j, and B, spanning from site $j+1$ to site N. It is quite natural to introduce the following states

$$\begin{aligned}|a_{\alpha_j}\rangle &= \sum_{s_1,\dots,s_j} \left(A(s_1)\dots A(s_j)\right)_{1\alpha_j} |s_1\dots s_j\rangle \\ |b_{\alpha_j}\rangle &= \sum_{s_{j+1},\dots,s_N} \left(A(s_{j+1})\dots A(s_N)\right)_{\alpha_j 1} |s_{j+1}\dots s_N\rangle\end{aligned} \tag{1.3.9}$$

These states are labelled by the index α_j (ranging from 1 to r_j) and represent a set of auxiliary states of the Hilbert space of the first (last) j ($N-j$) spins. We can write our original state $|\psi\rangle$ as $|\psi\rangle = \sum_{\alpha_j} |a_{\alpha_j}\rangle |b_{\alpha_j}\rangle$. This decomposition looks very similar to the Schmidt's one. By using the property of Eq. 1.3.8, we can easily prove that the states $|a_{\alpha_j}\rangle$ are orthonormalized. Indeed:

$$\begin{aligned}\langle a_{\alpha'_j}|a_{\alpha_j}\rangle &= \sum_{s_1,\dots,s_j} \left(A(s_1)\dots A(s_j)\right)^*_{1\alpha'_j} \left(A(s_1)\dots A(s_j)\right)_{1\alpha_j} = \\ &= \sum_{s_1,\dots,s_j} \left(\left(A(s_1)\dots A(s_j)\right)^\dagger\right)_{\alpha'_j 1} \left(A(s_1)\dots A(s_j)\right)_{1\alpha_j} = \\ &= \sum_{s_1,\dots,s_j} \left(A^\dagger(s_j)\dots A^\dagger(s_1)A(s_1)\dots A(s_j)\right)_{\alpha'_j\alpha_j} = \delta_{\alpha'_j\alpha_j} ,\end{aligned} \tag{1.3.10}$$

where in the last line we used Eq. 1.3.8 for sites $1, 2\dots j$. However, the same calculation cannot be repeated for the states $|b_{\alpha_j}\rangle$. For this reason the matrices A are called *left-normalized*, and the result of our decomposition is a left-normalized Matrix Product States (MPS) [27].

Obviously, we can repeat the same steps starting from the last site of the chain. In this case, we would obtain a similar decomposition [27]

$$c_{s_1\dots s_N} = B^{[1]}(s_1)B^{[2]}(s_2)\dots B^{[N]}(s_N) , \tag{1.3.11}$$

in which each set of matrices B obey to the condition

$$B(s_j)B^\dagger(s_j) = \mathbb{1} \ . \tag{1.3.12}$$

This is a *right-normalized* MPS. In Fig. 1.4, we represent the left and right normalization conditions for MPS tensors A and B respectively.

Independently of the normalization of the local tensors, we can define an MPS as a many-body quantum state $|\psi[A]\rangle$ that can be expressed as

$$\psi[A](\mathbf{s}) = \langle \mathbf{s}|\psi[A]\rangle = A^{[1]}(s_1)A^{[2]}(s_2)...A^{[N]}(s_N) \ , \tag{1.3.13}$$

where

$$\begin{cases} A^{[1]} \text{ is a tensor of shape } (\chi_0, d, \chi_1) \ , \quad \chi_0 = 1 \\ A^{[2]} \text{ is a tensor of shape } (\chi_1, d, \chi_2) \\ ... \\ A^{[N]} \text{ is a tensor of shape } (\chi_{N-1}, d, \chi_N) \ , \quad \chi_N = 1 \end{cases} \tag{1.3.14}$$

and d is the dimension of local Hilbert space ($d = 2$ for qubits or spin-1/2). The integers $\chi_0, \chi_1 ..., \chi_{N-1}, \chi_N$ are named *bond dimensions*. More in general, one can also consider MPS with Periodic Boundary Conditions (PBC), defined as

$$\psi[A](\mathbf{s}) = \langle \mathbf{s}|\psi[A]\rangle = \mathrm{Tr}\left[A^{[1]}(s_1)A^{[2]}(s_2)...A^{[N]}(s_N)\right] \ . \tag{1.3.15}$$

Here, the boundary tensors $A^{[1]}, A^{[N]}$ are allowed to be matrices in the auxiliary space.

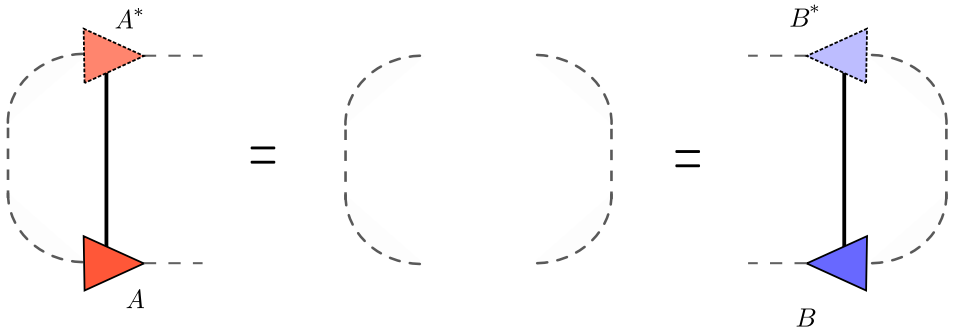

Figure 1.4: Left-normalization and right-normalization conditions of the MPS tensors. Conjugated tensors are represented as dotted lighter shapes.

1.4 BASIC PROPERTIES OF THE MPS

1.4.1 Entanglement

To examine one of the main properties of MPS, let us consider a left-normalized MPS and perform a series of SVD of the tensors A, starting from the site N

towards the site 1. At each site i we obtain $A_{a_{i-1}a_i}(s_i) = U_{a_{i-1}k}\Lambda_k B_{ka_i}(s_i)$ and we incorporate the matrices U, Λ into the next site $i-1$. Notice that in this way one can eventually pass from the initial left-normalized MPS to a right-normalized MPS (when $i = 1$) [27]. At a generic intermediate step, we obtain an object as the one depicted in Fig. 1.5, i.e. an MPS in a mixed form with a certain bond dimension χ. Importantly, the wave function's left and right sides are orthogonalized, ensuring a valid Schmidt decomposition and enabling access to the Schmidt spectrum Λ_k ($k = 1, 2...\chi$). As already discussed, this is strictly related to the entanglement properties of the state. For instance, one can easily evaluate the von Neumann entanglement entropy S as in Eq. 1.2.11. The entanglement entropy is maximized when all the eigenvalues takes the same value,i.e. $\Lambda_k^2 = 1/\chi$ so that

$$S \leqslant -\chi \cdot \frac{1}{\chi} \log\left(\frac{1}{\chi}\right) = \log \chi \,.$$

Thus, the MPS bond dimension χ gives a bound on the maximum entanglement content of the state: a state exhibiting an entanglement entropy value S for a given bipartition can be expressed as an MPS only using tensors with auxiliary dimension $\chi \sim \exp(S)$. This observation confirm our initial suspects: a profound link exists between entanglement and the intricacy involved in encoding and storing the information of a quantum state. MPS allow to write weakly entangled wave functions in a compressed form.

An example is provided by states that satisfy the *area law* of entanglement, namely those for which the entanglement entropy between a subsystem A and its complement B scales as $S \propto \partial A$. Random quantum states typically do not exhibit this behavior and conversely satisfy a *volume law*, i.e. $S \propto |A|$. Nevertheless, the ground states of one-dimensional gapped Hamiltonians for spin systems with local interactions are known to satisfy the area law [21]. In this case, since the boundary that separates two subsystems within a one-dimensional chain does not scale with the subsystem, the entanglement entropy remains bounded by a constant. Importantly, this allows to represent the state in an MPS form with a finite and reasonably small bond dimension χ, even in the thermodynamic limit. At a critical point however the energy gap typically vanishes and the entanglement entropy of a subsystem of size n acquires a logarithmic correction taking the form [18]

$$S(n) = \frac{c}{6} \log\left(\frac{2N}{\pi} \sin\left(\frac{\pi n}{N}\right)\right) + \mathcal{O}(1) \tag{1.4.1}$$

where c is a constant (dubbed central charge) that characterizes the underlying Conformal Field Theory. This equation translates into a logarithmic scaling of the half-chain entanglement entropy

$$S\left(n = \frac{N}{2}\right) = \frac{c}{6} \log N + \mathcal{O}(1) \,. \tag{1.4.2}$$

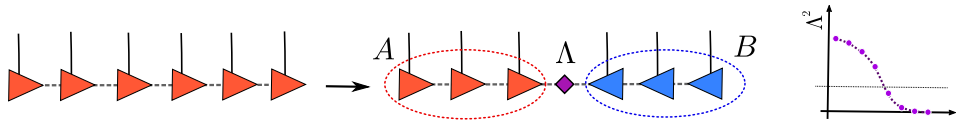

Figure 1.5: Converting a left-normalized MPS into a right-normalized MPS. At intermediate steps, a Schmidt decomposition of the state is obtained and entanglement spectrum is accessible.

1.4.2 Gauge freedom

The MPS representation is not unique. This can be easily verified by inserting an identity $X^{[i]}\left(X^{[i]}\right)^{-1} = \mathbb{1}$ in each auxiliary bond, i.e. replacing in Eq. 1.3.13 the original tensors $A^{[i]}$ with

$$\tilde{A}^{[i]} = \left(X^{[i-1]}\right)^{-1} A^{[i]} X^{[i]} \tag{1.4.3}$$

($X^{[0]} = X^{[N]} = 1$). The state is unchanged, namely

$$\psi[A](\mathbf{s}) = \psi[\tilde{A}](\mathbf{s})\,, \tag{1.4.4}$$

however our parameterization has been modified. In other words, the map $A \to |\psi[A]\rangle$ is not one-to-one. Thus, there is a *gauge freedom* in representing a quantum state with an MPS [27, 38, 40, 41]. This arbitrariness can be used proficiently to obtain computational advantages. For instance, we can remove part of the gauge freedom by imposing constraints, such as the left or right orthonormality conditions of the tensors.

1.4.3 Transfer matrix and correlation functions

Let us suppose to have an MPS $|\psi[A]\rangle$ and an operator $\hat{O}$ which acts locally on the Hilbert space, i.e. $\hat{O} = \hat{O}^{[1]} \otimes \ldots \otimes \hat{O}^{[N]}$. The expectation value of $\hat{O}$ over $|\psi[A]\rangle$ is

$$\begin{aligned}\langle\psi[A]|\hat{O}^{[1]}\ldots\hat{O}^{[N]}|\psi[A]\rangle =& A^*_{a_1}(s'_1)\ldots A^*_{a_{N-1}}(s'_N)\left(O^{[1]}\right)_{s'_1 s_1}\ldots \\ &\ldots\left(O^{[N]}\right)_{s'_N s_N} A_{a_1}(s_1)\ldots A_{a_{N-1}}(s_N)\end{aligned} \tag{1.4.5}$$

These tensor contractions can be represented graphically as in Fig. 1.6a). Often in this context it is useful to define the following tensor

$$\left(E_O^{[i]}\right)_{(a'_{i-1}a_{i-1}),(a'_i a_i)} = A^*_{a'_{i-1}a'_i}(s'_i)\left(O^{[i]}\right)_{s'_i s_i} A_{a_{i-1}a_i}(s_i) \tag{1.4.6}$$

which is usually dubbed *transfer matrix* (see Fig. 1.6b)). The transfer matrix acts as a super-operator in the auxiliary space, since it can map a matrix $M_{a'_{i-1}a_{i-1}}$ to a new matrix $M'_{a'_i a_i}$. In terms of the transfer matrix the expectation value of Eq. 1.4.5 becomes

$$\begin{aligned}\langle\psi[A]|\hat{O}^{[1]}\ldots\hat{O}^{[N]}|\psi[A]\rangle &= \left(E_O^{[1]}\right)_{(a'_1 a_1)}\left(E_O^{[2]}\right)_{(a'_1 a_1),(a'_2 a_2)}\ldots\left(E_O^{[N]}\right)_{(a'_{N-1}a_{N-1})} = \\ &= \mathrm{Tr}\left[E_O^{[1]}\ldots E_O^{[N]}\right].\end{aligned}$$

From a practical point of view, one can start by contracting the network starting from the left and creating an "environment matrix" $L_{a'_i a_i}$, which is updated site by site applying the transfer matrix $E_O^{[i]}$. If the update of L is performed contracting the tensors as in Fig. 1.6c), the computation cost is $\mathcal{O}(\chi^3)$, regardless of the system size N. After N updates, one reach the right boundary, obtaining the desired expectation value. In some cases the evaluation can become even simpler by using the MPS gauge freedom. For instance, the evaluation of a the expectation value of an operator $O^{[i]}$ acting only on the site i becomes trivial if one has the MPS in a mixed canonical form centered on the site i (see Fig. 1.6d)).

Let us now consider the calculation of a correlator $\langle\psi|\hat{O}^{[i]}\hat{O}^{[j]}|\psi\rangle$ [27]. We can define the transfer matrices E_O as in Eq. 1.4.6, and

$$E_{(a'_{i-1}a_{i-1}),(a'_i a_i)} = A^*_{a'_{i-1}a'_i}(s_i)A_{a_{i-1}a_i}(s_i)\,. \tag{1.4.7}$$

Notice that E is obtained as a particular case of E_O, by replacing $\left(O^{[i]}\right)_{s'_i s_i}$ with the identity $\delta_{s'_i s_i}$. From its definition, it is straightforward to see that E is a completely positive map and the MPS matrices A play the role of the Kraus operators for quantum channels [36]. If the tensors A are left normalized, then the identity $\delta_{a'_{i-1}a_{i-1}}$ is a left eigenvector of E, with eigenvalues 1. Furthermore, one can prove that in this case 1 has to be the largest eigenvalue in modulus. In general, the transfer matrix E can be written in the basis of its left and right eigenvectors as

$$E = \sum_{k=1}^{\chi^2} \lambda_k \,|l_k\rangle\langle r_k|\,, \tag{1.4.8}$$

where the states $|l_k\rangle, \langle r_k|$ belong to the (doubled) auxiliary space. We can set $|l_1\rangle$ as the leading left eigenvector, with eigenvalue 1. We have

$$\begin{aligned}\langle\psi|\hat{O}^{[i]}\hat{O}^{[j]}|\psi\rangle &= E^{[1]}\dots E^{[i-1]}E_O^{[i]}E^{[i+1]}\dots E^{[j-1]}E_O^{[j]}E^{[j+1]}\dots E^{[N]} = \\ &= \mathrm{Tr}\left[E^{[1]}\dots E^{[i-1]}E_O^{[i]}E^{[i+1]}\dots E^{[j-1]}E_O^{[j]}E^{[j+1]}\dots E^{[N]}\right] = \\ &= \mathrm{Tr}\left[E_O^{[i]}E^{[i+1]}\dots E^{[j-1]}E_O^{[j]}E^{[j+1]}\dots E^{[N]}E^{[1]}\dots E^{[i-1]}\right]\end{aligned}$$

By supposing to have a translationally invariant system in the thermodynamic limit $N\to\infty$, we can assume that the A (and E) matrices are essentially site independents. Thus, we get

$$\begin{aligned}\langle\psi|\hat{O}^{[i]}\hat{O}^{[j]}|\psi\rangle &= \mathrm{Tr}\left[E_O^{[i]}(E)^{j-1-1}E_O^{[j]}(E)^{N-j+i-1}\right] = \\ &= \sum_{k,l}\langle l_l|E_O^{[i]}|r_k\rangle\,\lambda_k^{j-1-1}\,\langle l_k|E_O^{[j]}|r_l\rangle\,\lambda_l^{N-j+i-1}\end{aligned}$$

where we used the decomposition in Eq. 1.4.8. Since 1 is the dominant eigenvalue, we get

$$\langle\psi|\hat{O}^{[i]}\hat{O}^{[j]}|\psi\rangle \xrightarrow[N\to\infty]{} \sum_k \langle l_1|E_O^{[i]}|r_k\rangle\,\lambda_k^{j-1-1}\,\langle l_k|E_O^{[j]}|r_1\rangle = c_0 + \sum_{k>1} c_k e^{-\frac{|j-i-1|}{\xi_k}}\,.$$

Therefore, the generic form of a two-point correlator over an MPS is given by the superposition (sum) of decaying exponentials. Since correlators over typical states in gapped systems exhibit exponential decay over distance, MPS are exceptionally effective in characterizing such states. Furthermore, it turns out that a sum of exponential can provide a reasonably accurate approximation even for power-law decay, at least at short distances. This fact enables the use of MPS to describe critical systems with a good degree of accuracy.

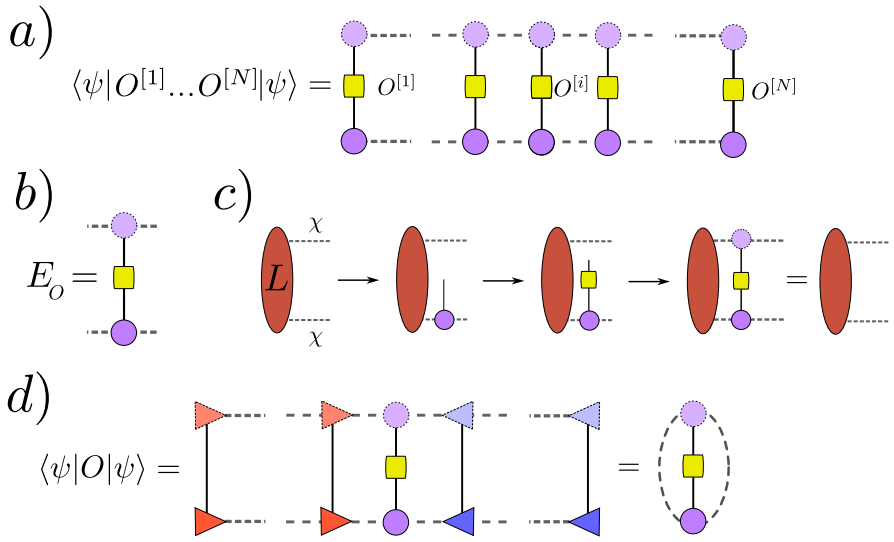

Figure 1.6: a) Evaluation of an expectation value of a tensored operator $\hat{O} = \hat{O}^{[1]}...\hat{O}^{[N]}$. over an MPS $|\psi[A]\rangle$. b) The transfer matrix E_O. c) Optimal contraction scheme. The environment matrix L of size $\chi \times \chi$ is updated by applying the transfer matrix and indices are contracted in the displayed order. d) Evaluation of a local expectation value over a mixed canonical MPS.

1.4.4 MPS compression

Compressing Matrix Product States from a high bond dimension D to a lower one χ is essential because it significantly reduces computational costs, making complex quantum simulations more practical and efficient. MPS compression can be achieved by the iterative algorithm reported in Ref. [27, 42]. This is designed to minimize site-by-site the Hilbert space distance between the two states with respect to the compressed local tensor $\tilde{A}^{[i]}_{a_{i-1},a_i}$, i.e.

$$\frac{\partial}{\partial(\tilde{A}^{[i]}_{a_{i-1},a_i})^*}\left(\| \, |\tilde{\psi}\rangle - |\psi\rangle \, \|^2\right) = 0 \, , \tag{1.4.9}$$

where $|\psi\rangle$, $|\tilde{\psi}\rangle$ are respectively the uncompressed and the compressed MPS. Eq. 1.4.9 can be rewritten as

$$\begin{aligned}\frac{\partial}{\partial \tilde{A}_{a_{i-1}a_i}(s_i)^*}\left(\| \, |\tilde{\psi}\rangle - |\psi\rangle \, \|^2\right) = \left(\frac{\partial}{\partial \tilde{A}_{a_{i-1}a_i}(s_i)^*} \langle\tilde{\psi}|\right)|\tilde{\psi}\rangle + \\ - \left(\frac{\partial}{\partial \tilde{A}_{a_{i-1}a_i}(s_i)^*} \langle\tilde{\psi}|\right)|\psi\rangle = 0 \, .\end{aligned} \tag{1.4.10}$$

which is linear in $\tilde{A}_{a_{i-1}a_i}(s_i)$ and can be represented graphically as in Fig. 1.7a). Interestingly, if one has the MPS in a mixed canonical form centered in i, the equation can be simplified becoming

$$\tilde{A}_{a_{i-1}a_i}(s_i) = L_{a_{i-1}\alpha_{i-1}} \, A_{\alpha_{i-1}\alpha_i}(s_i) R_{\alpha_i a_i} \, , \tag{1.4.11}$$

where L and R are left and a right environment tensors of shape $\chi \times D$ and $D \times \chi$ (see last line of Fig. 1.7a)). The key aspect here is that, after obtaining a solution for $\tilde{A}_{a_{i-1}a_i}(s_i)$, SVD is employed to maintain the normalization structure. For instance, one can left normalize $\tilde{A}^{[i]}$ and incorporate the remaining SVD matrices into the next site $\tilde{A}^{[i+1]}$ (Fig. 1.7b)). In this way, we obtain a mixed canonical form with respect to $i+1$. The left environment tensor L can be updated as in Fig. 1.7c), whereas the right environment tensors R must be computed and stored (starting from site N towards 1) before the procedure begins. At the beginning the compressed MPS is initialized in a right-canonical form. In practice, one iteratively performs a series of local minimizations, by sweeping along all the system sites a certain number of times N_{sweeps}. The computational cost of the MPS contractions involved in this procedure is $\mathcal{O}(N_{sweeps} N D^2 \chi)$ [27].

1.5 MATRIX PRODUCT OPERATORS (MPO)

An operator $\hat{O}$ acting on the Hilbert space $\mathcal{H}$ can always be decomposed as

$$\hat{O} = c_{(s_1,s_1')\ldots(s_N,s_N')} \, |s_1 \ldots s_N\rangle \langle s_1' \ldots s_N'| \, . \tag{1.5.1}$$

The similarity with Eq. 1.1.2 implies that we can decompose the tensor c in a way analogous to how we decomposed the state $|\psi\rangle$, whit the dual index (s_i, s_i') taking the role previously held by s_i [27]. We obtain the analogous of Eq. 1.3.6 in the world of operators, i.e. the following Matrix Product Operator (MPO)

$$c_{(s_1,s_1')\ldots(s_N,s_N')} = W^{[1]}(s_1, s_1') \ldots W^{[N]}(s_N, s_N') \, , \tag{1.5.2}$$

where $W^{[i]}$ are matrices (vectors at the boundaries) depending on the two spin variables s, s'. The size of auxiliary space defines the MPO bond dimension. A graphical representation of this entity is presented in Fig. 1.8a).

Importantly, the expectation value $\langle\psi|\hat{O}|\psi\rangle$ of an operator $\hat{O}$ represented in the form of an MPO over an MPS $|\psi\rangle$ can be computed efficiently (see Fig. 1.8b)) [27, 43]. In particular, one can proceed as illustrated in Fig. 1.6c), with the only difference being that now the environment tensors have an additional auxiliary index. If we denote the MPO bond dimension as D, the overall cost will be $\mathcal{O}(\chi^3 D^2)$. Another important task is the application of an

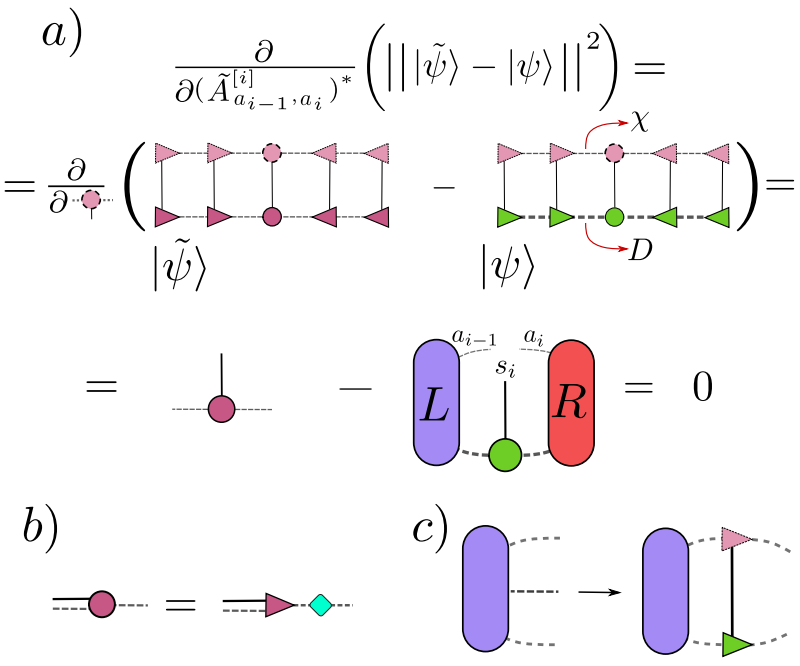

Figure 1.7: The iterative compression algorithm for MPS. The compressed (uncompressed) MPS tensors are represented by violet (green) shapes. a) The minimization of the states distance with respect to the uncompressed MPS tensor $(\tilde{A}^{[i]})^*$ (first line) gives a linear equation for $\tilde{A}^{[i]}$. Since the MPS tensors to the left (right) of $\tilde{A}^{[i]}$ are left (right) normalized, the linear system is greatly simplified (third line). b) The solution for $\tilde{A}^{[i]}$ is decomposed to preserve the mixed canonical form for the next step. c) The left environment tensor L is updated.

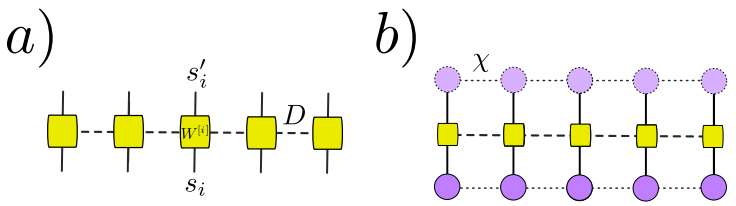

Figure 1.8: a) A generic Matrix Product Operator (MPO) with bond dimension D. b) The expectation value $\langle\psi|O|\psi\rangle$ of an MPO over an MPS.

MPO to an MPS. This operation can also be performed efficiently and preserves the MPS form, but increases its bond dimension from χ to $D\chi$ [27, 43].

In addition to their computational relevance, MPOs can be used also as an exact analytical tool in a variety of contexts, ranging from classical statistical mechanics to integrable systems (the Bethe ansatz can be rewritten in terms of an algebra of commuting MPOs) [44].

1.5.1 Examples of useful MPOs

In order to manipulate an MPO, it can be useful to define the following matrices of operators $\tilde{W}^{[i]}_{\alpha\beta} = W^{[i]}_{\alpha\beta}(s_i, s'_i)\,|s_i\rangle\langle s'_i|$, where α, β are indices belonging to the auxiliary space. We get

$$\hat{O} = \tilde{W}^{[1]} \dots \tilde{W}^{[N]} \,. \tag{1.5.3}$$

One can think at the above equation as the action of a finite state machine operating on D different virtual states, as in Fig.1.9.

Figure 1.9: An MPO having bond dimension D can be thought as a finite state machine acting on D virtual states.

To better understand the mapping, let us define the vectors

$$V^{[l]} = \tilde{W}^{[l]} \dots \tilde{W}^{[N]} \qquad l = 1, 2 \dots N \,,$$

whose entries are operators. We have

$$V^{[l-1]}_{\alpha} = \tilde{W}^{[l-1]}_{\alpha\beta} V^{[l]}_{\beta} \,,$$

meaning that the $\tilde{W}$ are transition matrices for the vectors V. V represent the internal states of the finite state machine. We will now show how the transition matrix (the MPO) can be adjusted to obtain a represantation of some useful operators $\hat{O}$.

Let us begin with a simple example. Suppose $\hat{O} = \hat{\Sigma}^z$ is the total $z-$magnetization operator

$$\hat{\Sigma}^z = \sum_{i=1}^{N} \hat{\sigma}^z_i = \sum_{i=1}^{N} \left(\hat{\mathbb{1}}_1 \otimes \dots \otimes \hat{\mathbb{1}}_{i-1} \otimes \hat{\sigma}^z_i \otimes \hat{\mathbb{1}}_{i-1} \otimes \dots \otimes \hat{\mathbb{1}}_N \right) \,.$$

Let us consider a finite state machine operating on $D = 2$ states and let us represent its action as in Fig. 1.10. The corresponding transition matrix is

$$\tilde{W}^{[l]} = \begin{pmatrix} \hat{\mathbb{1}} & 0 \\ \hat{\sigma}^z_l & \hat{\mathbb{1}} \end{pmatrix}$$

with $l = 2, \dots N$. The vector $\tilde{W}^{[N]}$ set the starting state of the machine. It can be both $\boxed{1}$ (and in this case we have to apply the identity operator $\hat{\mathbb{1}}_N$) or $\boxed{2}$ (and in this case we have to apply the operator $\hat{\sigma}^z_N$). Hence, we set

$$\tilde{W}^{[N]} = \begin{pmatrix} \hat{\mathbb{1}} \\ \hat{\sigma}^z_N \end{pmatrix}$$

It is easy to realize that, by applying the $\tilde{W}^{[l]}$ matrices to $\tilde{W}^{[N]}$, one get

$$V^{[2]} = \begin{pmatrix} \hat{\mathbb{1}}_2 \otimes \hat{\mathbb{1}}_3 \otimes \dots \otimes \hat{\mathbb{1}}_N \\ \sum_{i=2}^{N} \hat{\mathbb{1}}_2 \otimes \dots \otimes \hat{\mathbb{1}}_{i-1} \otimes \hat{\sigma}^z_i \otimes \hat{\mathbb{1}}_{i+1} \otimes \dots \otimes \hat{\mathbb{1}}_N \end{pmatrix} \,.$$

Now, it becomes obvious that to obtain the operator $\hat{\sigma}^z$ we have to set

$$\tilde{W}^{[1]} = \begin{pmatrix} \hat{\sigma}_1^z & \hat{\mathbb{1}}_1 \end{pmatrix} .$$

It is not difficult to recast standard 1D Hamiltonians as MPOs with such techniques [27]. For instance, let us consider the following Ising Hamiltonian, with exponentially decaying couplings

$$\hat{H} = -J \sum_{i=1}^{N} \sum_{j=1}^{i-1} \lambda^{i-j} \hat{\sigma}_i^z \hat{\sigma}_j^z - \sum_{i=1}^{N} \mathbf{h} \cdot \boldsymbol{\sigma} , \tag{1.5.4}$$

where $\mathbf{h}$ is the on-site magnetic field and $\boldsymbol{\sigma} \equiv (\hat{\sigma}^1, \hat{\sigma}^2, \hat{\sigma}^3)$. The finite state machine plotted in Fig. 1.11 can realize such operator. Indeed, at each loop on the intermediate state $\boxed{2}$ one acquire a factor λ, obtaining λ^{i-j} before of jumping on $\boxed{3}$. The corresponding transition matrix is

$$\tilde{W}^{[l]} = \begin{pmatrix} \hat{\mathbb{1}} & 0 & 0 \\ \hat{\sigma}_l^z & \lambda\hat{\mathbb{1}} & 0 \\ -\mathbf{h} \cdot \boldsymbol{\sigma}_l & -J\lambda\hat{\sigma}_l^z & \hat{\mathbb{1}} \end{pmatrix} ,$$

whereas the boundary vectors will be

$$\tilde{W}^{[1]} = \begin{pmatrix} -\mathbf{h} \cdot \boldsymbol{\sigma}_1 & -J\lambda\hat{\sigma}_1^z & \hat{\mathbb{1}}_1 \end{pmatrix} \qquad \tilde{W}^{[N]} = \begin{pmatrix} \hat{\mathbb{1}} \\ \hat{\sigma}_N^z \\ -\mathbf{h} \cdot \boldsymbol{\sigma}_N \end{pmatrix} .$$

Notice that if we set $J = J_0/\lambda$ and let λ approach 0, the Hamiltonian in Eq. 1.5.4 simplifies to the standard short-range Ising Hamiltonian. Consequently, in this scenario, we can employ the values of $\tilde{W}$ listed above with $\lambda = 0$. These tricks can be used also to get the power-law long-range Ising Hamiltonian with power-law decaying couplings, i.e.

$$\hat{H} = -J \sum_{i \neq j}^{N} \frac{\hat{\sigma}_i^z \hat{\sigma}_j^z}{(i-j)^\alpha} - \sum_{i=1}^{N} \mathbf{h} \cdot \boldsymbol{\sigma}$$

with $\alpha > 0$. To this purpose, one can fit the power law $1/r^\alpha$ as a sum of decaying exponentials, i.e. $1/r^\alpha \approx \sum_{k=1}^{n} c_k e^{-r/\xi_k} = \sum_{k=1}^{n} c_k \lambda_k^r$, with $\lambda_k = e^{-1/\xi_k}$. In this way, we can employ a finite state machine resembling the one represented in Fig. 1.11 but with more intermediate states, each corresponding to a fixed value of λ_k. In this case, the MPO bond dimension equals to $D = n + 2$. One can also easily design an MPO representing a 2D Hamiltonian on a square lattice of size $N_x \times N_y$, as for instance

$$\hat{H} = -J \sum_{\langle i,j \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z - \sum_{i=1}^{N} \mathbf{h} \cdot \boldsymbol{\sigma} .$$

The result is the finite state machine represented in Fig. 1.12. The price to pay is having a bond dimension D scaling with the system linear size N_y (or N_x).

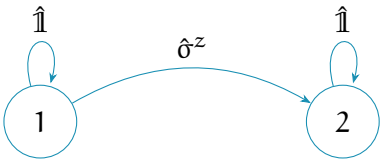

Figure 1.10: Finite state machine representation of the $\hat{\sigma}^z$ operator expressed as MPO.

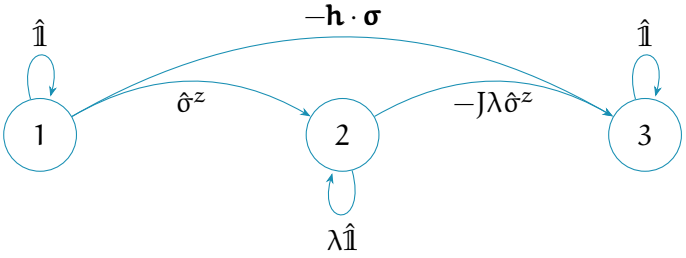

Figure 1.11: Finite state machine representation of the long-range Ising Hamiltonian expressed as MPO.

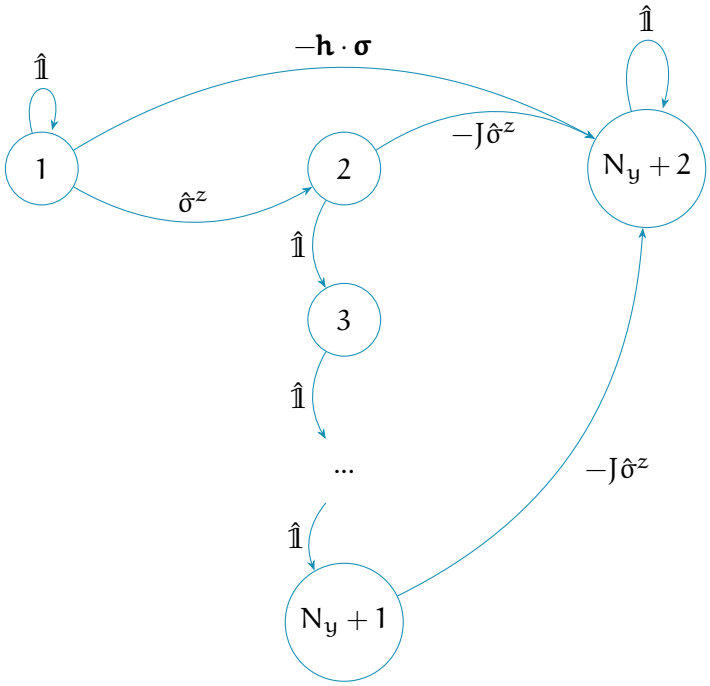

Figure 1.12: Finite state machine representation of the 2D nearest-neighbours Ising Hamiltonian expressed as MPO.

1.6 DENSITY MATRIX RENORMALIZATION GROUP (DMRG)

The ability of DMRG method to accurately describe ground-state quantum properties, capture entanglement and correlations has made it an indispensable tool for investigating low-energy physics in a wide range of systems [28]. DMRG was initially proposed in 1992 by Steven White in terms of a sort of renormalization flow for the density matrix of spins blocks [23, 24], and afterwards reformulated in terms of MPS optimization [27, 45, 46]. Nowadays nearly every cutting-edge DMRG implementation employs MPSs and

MPOs [27, 33, 47, 48]. Here, we provide a brief overview of this approach, specifically focusing on its simplest form, known as the one-site DMRG [27, 33].

In short, DMRG exploits MPS as a variational ansatz to approximate the many-body wave function describing the ground state. Following the principles of the variational method [49], one has to minimize the functional

$$\mathcal{E}[A,\epsilon] = \langle\psi[A]|\hat{H}|\psi[A]\rangle - \varepsilon\Big(\langle\psi[A]|\psi[A]\rangle - 1\Big) \tag{1.6.1}$$

with respect to the variational parameters $A_{a_{i-1}a_i}(s_i)$ ($i = 1, 2...N$). Notice that in Eq. 1.6.1 we introduced the Lagrange multiplier ε to enforce the normalization constraint $\frac{\partial}{\partial\varepsilon}\mathcal{E} = \langle\psi[A]|\psi[A]\rangle - 1 = 0$. Taking the partial derivative with respect to the conjugate of the tensor A, we get

$$\begin{aligned}\frac{\partial}{\partial A_{a_{i-1}a_i}(s_i)^*}\mathcal{E}[A,\epsilon] = \Big(\frac{\partial}{\partial A_{a_{i-1}a_i}(s_i)^*}\langle\psi[A]|\Big)\hat{H}|\psi[A]\rangle + \\ - \epsilon\Big(\frac{\partial}{\partial A_{a_{i-1}a_i}(s_i)^*}\langle\psi[A]|\Big)|\psi[A]\rangle = 0\,.\end{aligned} \tag{1.6.2}$$

This equation is linear in $A_{a_{i-1}a_i}(s_i)$ and can be represented graphically as in Fig. 1.13a) if one assume that the Hamiltonian $\hat{H}$ is expressed in as an MPO. Interestingly, if one has the MPS in a mixed canonical form centered in i, the equation can be simplified becoming

$$\mathcal{M}_{(a_{i-1}a_i),(a'_{i-1}a'_i)}(s_i,s'_i)\; A_{a'_{i-1}a'_i}(s'_i) = \epsilon A_{a_{i-1}a_i}(s_i)\,, \tag{1.6.3}$$

namely a simple eigenvalue problem. Furthermore, because of the MPS/MPO structure, the operator $\mathcal{M}$ can be expressed in the following form

$$\mathcal{M}_{(a_{i-1}a_i),(a_{i-1}a_i)}(s_i,s'_i) = L_{a_{i-1}\alpha_{i-1}a'_{i-1}} W_{\alpha_{i-1}\alpha_i}(s_i,s'_i)\, R_{a_i\alpha_i a'_i} \tag{1.6.4}$$

in terms of a left and a right environment tensor L and R (see last line of Fig. 1.13a)). Our (local) optimization problem can now be solved or by brute force diagonalization of the operator $\mathcal{M}$, or with an iterative solver, such as the Lanczos method. As in the MPS iterative compression (Sec. 1.4.4), it is important that, after obtaining a solution for $A_{a_{i-1}a_i}(s_i)$, SVD is employed to maintain the normalization structure. For instance, one can left normalize $A^{[i]}$ and incorporate the remaining SVD matrices into the next site $A^{[i+1]}$ (Fig. 1.13b)). In this way, we obtain a mixed canonical form with respect to $i+1$. The left environment tensor L can be updated as in Fig. 1.13c), whereas the right environment tensors R must be computed and stored (starting from site N towards 1) before the procedure begins. At the beginning the MPS is initialized in a right-canonical form. Typically, one repeats the outlined local optimization a number of times, each time sweeping all the sites of the chain. The convergence of the method can be assessed by tracking the variational energy ε or its variance $\delta^2 H = \langle\psi[A]|\hat{H}^2|\psi[A]\rangle - \langle\psi[A]|\hat{H}|\psi[A]\rangle^2$ (which can also be evaluated very easily).

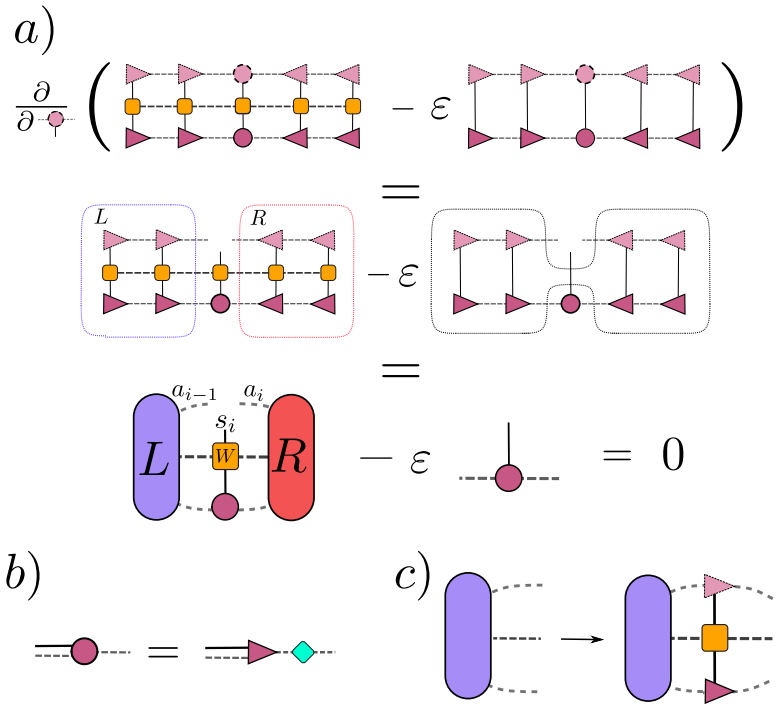

Figure 1.13: DMRG in a nutshell. a) The minimization of the energy functional with respect to the local MPS tensor $(A^{[i]})^*$ (first line) gives a linear equation for $A^{[i]}$ (second line). Since the MPS tensors to the left (right) of $A^{[i]}$ are left (right) normalized, the linear system is simplified to an ordinary eigenvalue problem. b) The solution for $A^{[i]}$ is decomposed to preserve the mixed canonical form for the next step. c) The left environment tensor L is updated.

1.7 TIME EVOLVING BLOCK DECIMATION (TEBD)

In addition to computing ground states, MPS enables efficient numerical simulations of the time evolution of a closed quantum system. The goal is to evaluate

$$|\psi(t)\rangle = \exp\left(-i\hat{H}t\right)|\psi(0)\rangle\,, \tag{1.7.1}$$

where H is the many-body Hamiltonian. Usually, the dynamics is discretized into time intervals of length δt. A direct calculation of the exponential $\exp(-iH\delta t)$ is possible only for very small N. If one has a Hamiltonian that includes exclusively nearest-neighbor interactions, i.e. $\hat{H} = \sum_i \hat{h}_{i,i+1}$, the first-order Trotter decomposition of the time evolution operator reads as follows

$$\begin{aligned} U^{\text{exact}}_{t,t+\delta t} &= \exp\left(-iH\delta t\right) = \exp\left(-i\hat{H}^{\text{even}}\delta t - i\hat{H}^{\text{odd}}\delta t\right) \simeq \\ &\simeq \exp\left(-i\hat{H}^{\text{even}}\delta t\right)\exp\left(-i\hat{H}^{\text{odd}}\delta t\right)\exp\left(-\frac{(\delta t)^2}{2}[\hat{H}^{\text{even}},\hat{H}^{\text{odd}}]\right) = \\ &= \underbrace{\exp\left(-i\hat{H}^{\text{even}}\delta t\right)\exp\left(-i\hat{H}^{\text{odd}}\delta t\right)}_{U^{\text{TEBD}}_{t,t+\delta t}} + O\left((\delta t)^2\right) \end{aligned} \tag{1.7.2}$$

where $\hat{H}^{\text{even}} = \sum_{i\in\text{even sites}} h_{i,i+1}$ and $\hat{H}^{\text{odd}} = \sum_{i\in\text{odd sites}} h_{i,i+1}$. Since all the summands $h_{i,i+1}$ in $\hat{H}^{\text{even}}$, $\hat{H}^{\text{odd}}$ commute with each other, we have

$$\begin{aligned} \exp\left(-i\hat{H}^{\text{even}}\delta t\right) &= \prod_{i\in\text{even sites}} \exp\left(-ih_{i,i+1}\delta t\right) \\ \exp\left(-i\hat{H}^{\text{odd}}\delta t\right) &= \prod_{i\in\text{odd sites}} \exp\left(-ih_{i,i+1}\delta t\right) \end{aligned} \tag{1.7.3}$$

Now, rather than exponentiating a single large matrix H, one has to compute $\exp\left(-ih_{i,i+1}\delta t\right)$. This can be done very easily since $h_{i,i+1}$ is a 4×4 matrix (or a $2\times 2\times 2\times 2$ tensor). The resulting object, usually dubbed "bond operator", is a unitary gate acting on two sites.

To complete a full time step $U^{\text{TEBD}}_{t,t+\delta t}$, the procedure involves applying the bond operators on odd sites first, followed by those on even sites. The full tensor network representing TEBD evolution is depicted in Fig. 1.14. One practical approach to apply the bond operators consists in starting from site 1 and moving to site N, sequentially applying $\exp\left(-ih_{i,i+1}\delta t\right)$ (i odd). Then, similarly, one evolves the bond operators on even sites $\exp\left(-ih_{i,i+1}\delta t\right)$ (i even) by moving from site N to 1. Importantly, using this procedure, at each step, one can apply the bond operators to an MPS in a mixed canonical form. The implementation of bond operators locally disrupts the MPS form, but this can be restored through the use of SVD (see Fig. 1.15). Notice that, since a mixed canonical form is used, the SVD provides access to the Schmidt eigenvalues (entanglement spectrum) of the full system with respect to the bipartition defined by the bond $i, i+1$. The SVD increase the local bond dimension by a factor of 2, but one can choose to truncate the Schmidt eigenvalues/eigenvectors in such a way as to keep the bond dimension within a fixed value χ set at the beginning. This technique can be applied also to simulate a quantum circuit with local unitary gates (indeed Fig. 1.14 represents a quantum circuit with a brick-wall structure).

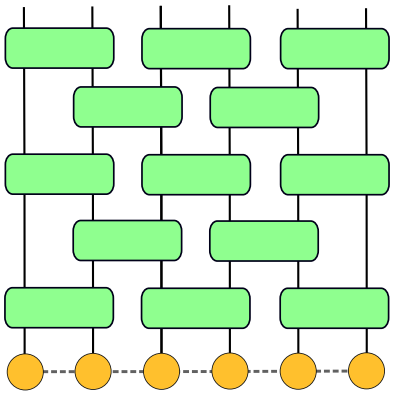

Figure 1.14: Sketch of TEBD applied to an initial MPS.

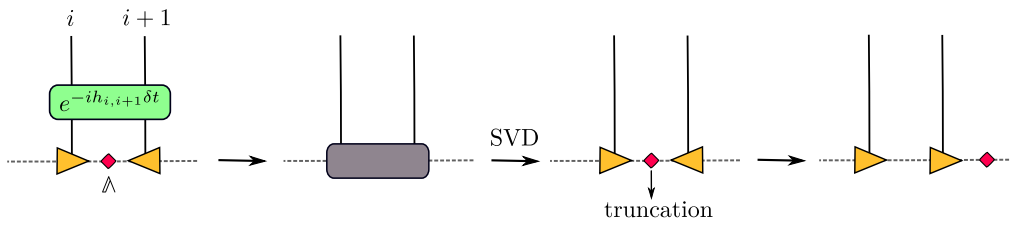

Figure 1.15: The bond operator $\exp\left(-ih_{i,i+1}\delta t\right)$ is applied to two MPS tensors (here in a mixed canonical form). Afterwards the MPS form is restored by means of an SVD. Schmidt eigenvalues/eigenvectors can be truncated to keep the bond dimension within a fixed value χ. Finally, the mixed canonical form is shifted for the next step of the sweep.

1.8 MPS AND QUANTUM CIRCUITS

The purpose of this Section is to explain how an MPS $|\psi\rangle$ can be obtained by applying a suitable unitary operator $\mathcal{U}^\dagger$ to the simple product state $|0\rangle^{\otimes N}$. This aspect is particularly relevant since it enables the implementation of an MPS on a quantum platform. We want to satisfy the equation $|\psi\rangle = \hat{\mathcal{U}}^\dagger|0\rangle$, or also $\hat{\mathcal{U}}|\psi\rangle = |0\rangle$. The latter shows that $\hat{\mathcal{U}}$ is a disentangler operator (because when applied to the entangled state $|\psi\rangle$ gives the disentangled state $|0\rangle$). Furthermore, in principle, we would like the unitary operator $\mathcal{U}$ to be expressible in terms of (local) one qubit or two qubits gates, making it more easily implementable on a quantum device. Our starting point is the MPS

$$\psi[A](\mathbf{s}) = A_{a_1}(s_1)A_{a_1a_2}(s_2)...A_{a_{N-1}}(s_N),$$

that we assume to be in the right-normal form. Let us begin by examining the case in which the local bond dimensions χ_i equal the physical dimension $d(=2)$, for all the sites $i = 1, 2...N$. We focus on a certain site i in the bulk ($1 < i < N$) and we reshape A as a rectangular matrix of shape $\chi \times d\chi$ by means of the mapping $A_{a_{i-1}a_i}(s_i) \to A_{a_{i-1},(a_i,s_i)}$. The right-normalization condition is now expressed by $AA^\dagger = \mathbf{1}_{\chi\times\chi}$. Let us notice that $A^\dagger A$ is not guaranteed to be the identity. Therefore A is an *isometry* matrix (not unitary!). The idea is to embed A into a proper unitary matrix U. To do this we have to add an extra (fake) index $\tilde{s}$, i.e.

$$A_{a_{i-1},(a_i,s_i)} \to U_{(a_{i-1},\tilde{s}_{i+1}),(a_i,s_i)} \tag{1.8.1}$$

so that U will be a square matrix of shape $\chi d \times d\chi$. The reason why we choose the label $\tilde{s}_{i+1}$ will be clear later. Let us now set

$$U_{(a_{i-1},0),(a_i,s_i)} = A_{a_{i-1},(a_i,s_i)}\,. \tag{1.8.2}$$

This condition is graphically represented in Fig. 1.16 and ensures that when U is applied on the state $|0\rangle$ the results is the original MPS tensor A.

This means that the first χ rows of the matrix U will be composed of the χ vectors of length $d\chi$ contained in A. The idea is now to fill the other $\chi d - \chi$ rows with other vectors in order to create an orthonormal basis of the complex vector space $\mathbf{C}^{\chi d}$. This can always done, for example by means of the Gram–Schmidt process. After that, U will necessarily be a unitary matrix,

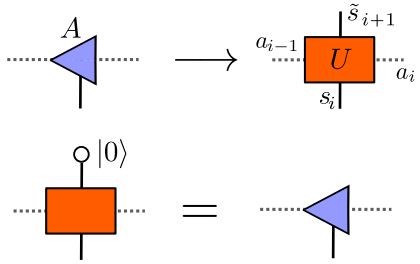

Figure 1.16: Embedding the local MPS tensors A into a unitary matrix U.

satisfying U

Something different occurs on the boundaries. If $i = 1$, A has shape $1 \times d\chi$ (i.e. it is just a single vector of length $d\chi$). Therefore, we have to add two extra (fake) indices in order to embed A into a unitary matrix of dimension $d\chi \times d\chi$. The mapping will be

$$A_{(a_1,s_1)} \to U_{(\tilde{s}_1,\tilde{s}_2),(a_1,s_1)} \ . \tag{1.8.3}$$

As usual, we have to set

$$U_{(0,0),(a_1,s_1)} = A_{(a_1,s_1)} \tag{1.8.4}$$

and fill the other $d\chi - 1$ rows of the matrix with new orthonormal vectors. On the other hand, when $i = N$ the tensor A has shape $\chi \times d$. Since $\chi = d$, this is already a square unitary matrix.

After these mappings, we found that the original MPS can be formally recast into a series of unitary operators acting, by the fake indices $\tilde{s}$, on N copies of the state $|0\rangle$ (see second line in Fig.1.17a)). Now, since $\chi = d$, we can think the auxiliary indices a to be themselves physical indices. We can therefore reshape the circuit and obtain the graphical representation in the third line in Fig. 1.17a). This is a proper staircase quantum circuit, composed only by one qubit and two qubits gates.

Let us now consider the general case. For a generic MPS the bond dimensions are

$$\chi_i = 2, 4, 8 ... \ \chi \ ... \ \chi \ ... \ 8, 4, 2 \quad \text{for} \quad i = 1, 2, 3 ... \ N-3, N-2, N-1 \, .$$

χ is the bond dimension in the bulk and we will assume to be a power of 2, i.e. $\chi = 2^n$. An MPS of this kind is represented in the first line of Fig. 1.17b), where each grey dotted line represents a binary index (in the bulk the auxiliary indices will have n of these). The exact mapping between the MPS and the quantum circuit is a straightforward generalization of the previous case, making use (in the bulk) of $\log_2 \chi + 1 = n + 1$ qubits gates. The mapping is schematically sketched in Fig. 1.17b). These gates can be in principle further decomposed into a series of sequential 2-site unitaries. As shown in [50], the number of CNOT gates necessary for the decomposition of a single $(n+1)$ qubits gate is $\mathcal{O}(3 \cdot 4^{n-1}) = \mathcal{O}(\chi^2)$. Thus, the optimized MPS can

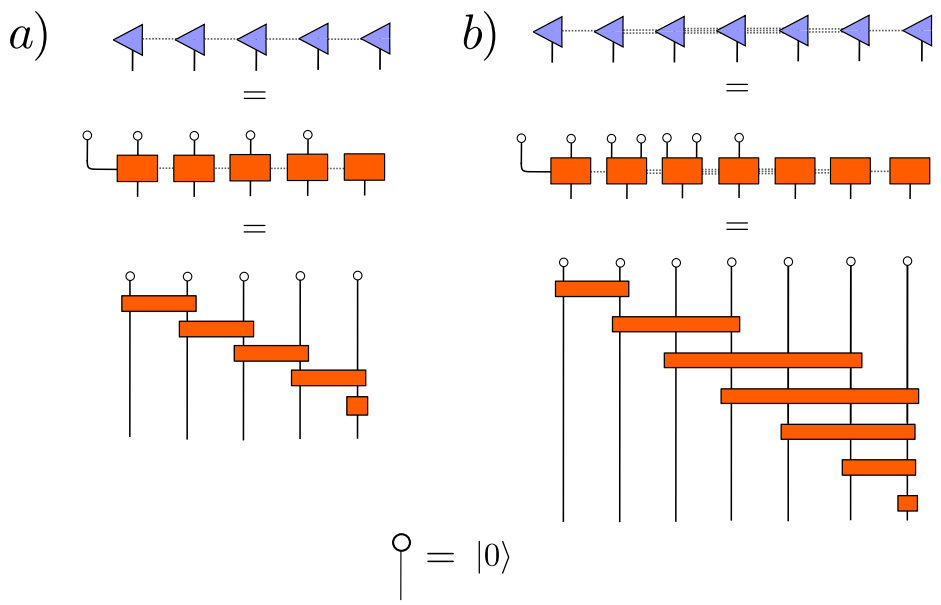

Figure 1.17: The MPS to quantum circuit mapping: a) for the simple case $\chi = d$; b) for the generic case $\chi > d$.

be recast into a quantum circuit consisting of $\mathcal{O}(N\chi^2)$ elementary two-qubits gates (such a CNOTS).

Finally, let us mention that one can achieve an approximate quantum circuit $\mathcal{U}$ by iteratively optimizing the fidelity between the target MPS and the quantum circuit, i.e.

$$\mathcal{F} = |\langle\psi|\hat{\mathcal{U}}|0\rangle|^2 .$$

$\mathcal{F}$ can be represented as in Fig. 1.18, where $\mathcal{U}$ is made of D "staircase" layers of two qubits gates. At a generic step of the iterative optimization, we fix all the two qubits gates except for one. We have

$$\mathcal{F} = |\langle\tilde{\psi}|\hat{U}|\tilde{\phi}\rangle|^2 ,$$

where $|\tilde{\psi}\rangle$ ($|\tilde{\phi}\rangle$) is the states obtained by applying the fixed gates to $|\psi\rangle$ ($|0\rangle$) and U is the two-qubits unitary gate we want to optimize. By defining the environment operator $E = |\tilde{\phi}\rangle\langle\tilde{\psi}|$, we have $\mathcal{F} = |\mathrm{Tr}(UE)|^2$. Let us notice that if U_{opt} is a possible unitary optimal solution of the problem also $U_{opt}e^{i\theta}$, for any phase θ, is a solution. Therefore we can assume $\mathrm{Tr}(UE)$ to be real. Now, the following inequality applies: $|\mathrm{Tr}(UE)| \leqslant ||E||_1$, where $||E||_1 = \mathrm{Tr}\left(\sqrt{E^\dagger E}\right) = \mathrm{Tr}\left(\sqrt{\tilde{V}\Lambda\tilde{U}^\dagger\tilde{U}\Lambda\tilde{V}^\dagger}\right) = \mathrm{Tr}(\Lambda)$ and $E = \tilde{U}\Lambda\tilde{V}^\dagger$ is the SVD of E. Let us notice that if we fix $U = \tilde{V}\tilde{U}^\dagger$ the inequality is saturated. Therefore this is the optimal choice of the local two-qubits gate. After fixing this, one can move to the next gate of the staircase, in an iterative fashion [51]. By performing an adequate number of iterations along all the gates of the circuit, this procedure is expected to reach convergence.

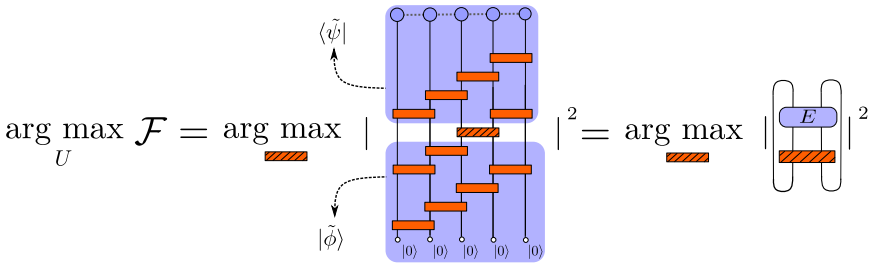

Figure 1.18: The fidelity maximization problem at a generic step of the iterative method.

2 MATRIX PRODUCT STATES WITH BACKFLOW CORRELATIONS

Understanding quantum many-body systems in and out of equilibrium is one of the most exciting open challenges in physics and chemistry. In recent years, significant progress has been made in the study of strongly correlated quantum systems, on many fronts. For example, several experimental approaches implementing Feynmans' quantum simulators [15] are allowing the controlled exploration of uncharted territory [52–58]. Recent experiments based on Rydberg atoms in optical tweezers allowed the simulation of quantum computational schemes to unprecedented system sizes. This has prompted a renewed interest in numerical techniques capable to provide accurate estimations for future quantum simulations [59–63].

On the theoretical level, the development of Tensor-Networks (TN) techniques has significantly expanded the scope of variational approaches to quantum many-body systems since the introduction of the Density-Matrix Renormalization Group (DMRG) algorithm [23]. The goal of TNs is to represent the quantum many-body wave functions by means of a set of local tensors, connected in a generic network via auxiliary bonds with finite dimension χ, thus overcoming computational limitations due to the exponentially large Hilbert space [64, 65]. The bond dimension χ can be adjusted to manipulate the information content of the TN, thus going from product states ($\chi = 1$), reproducing mean-field approximations, to the exact but inefficient wave function representation. In 1D, the Matrix Product State (MPS) geometry has demonstrated an unprecedented degree of accuracy for both equilibrium and out-of-equilibrium problems [27, 66]. We refer the reader to Chapter 1 for a detailed introduction to TNs and MPS. However, TN have some fundamental limitations, such as the intrinsic hardness of finding efficient contraction schemes [67] and unfavorable scaling of the required resources with the system size in higher dimension [64]. Most successful TN geometries, like Projected-Entangled Pair States (PEPS) [68] and Tree Tensor Networks [69], suffer from specific drawbacks: while the latter does not satisfy the Entanglement area law (although some effort has been spent to overcome this limitation in Ref. [61]), the former suffers from high algorithmic complexity, $O(\chi^{10})$, and lacks exact computation of expectation values.

In parallel to the progress of TN, artificial Neural Networks (NN) have been discovered and used in a plethora of different scientific fields, prov-

ing astonishing versatility in physics applications [70]. In recent years, they have been employed as a variational ansatz for quantum many-body problems [71]. In this context, a number of possible architectures have been tried, such as Restricted Boltzmann Machine (RBM) [71, 72], Feed-Forward NN [73, 74] and Recurrent NN [75]. These ansatze have been proven to have a great descriptive power [73, 76]. However, the number of parameters entering a NN wave function may be arbitrarily large and the appropriate network structure is usually not clear a priori. Understanding an optimal geometry encoding information from the specific dimensionality of the problem and taking advantage from both TN and NN structures could be the ultimate solution to the quantum many-body problem.

NN are usually optimized by means of variational Monte-Carlo (VMC) methods. Furthermore, a key tool in NN optimization is the so-called Automatic Differentiation [77], which allows to efficiently compute cost-function derivatives with machine precision. This paradigm have been recently applied also to the TNs optimization [78]. Combining such approaches with standard TN algorithms appears as a promising way to find new optimal strategies to solve open problems at the equilibrium and out-of-equilibrium. Efforts in this direction were made with the introduction of the Entangled Plaquette States (EPS) [79, 80], Monte-Carlo optimized MPS [81] and PEPS [82], and infinite PEPS optimized with automatic differentiation [83]. The space of possible hybrid wave functions is however still largely unexplored.

In this Chapter, we introduce a novel variational ansatz, generalizing the usual MPS. The ansatz is inspired by the so-called backflow techinque, commonly employed in electronic-structure theory [84–86]. These new Matrix Product Backflow States (MPBS) can overcome some limitations of MPS by encoding an extensive amount of entanglement and keeping the algorithmic complexity under control. We further introduce a simple optimization scheme mixing DMRG and VMC recipes which can be proficiently applied to MPBS in order to find quantum many-body ground-states. As a benchmark, we employ this approach against well-known 1D and 2D spin models. Finally, we simulate the $J_1 - J_2$ model, providing the ability to inspect some challenging highly non-trivial models.

2.1 MATRIX PRODUCT BACKFLOW STATES

A state $|\psi\rangle$ of a quantum many-body system consisting of N spin-1/2 variables is fully specified by the complex-valued function $\psi(\mathbf{s}) = \langle \mathbf{s}|\psi\rangle$, $\mathbf{s} \in \{\pm 1\}^N$ being the spin projections along the z direction. As discussed in Chapter 1, MPS [27] are defined by the functional form

$$\psi[A](\mathbf{s}) = A^{[1]}(s_1)A^{[2]}(s_2)...A^{[N]}(s_N)\ , \tag{2.1.1}$$

where local tensors $A^{[i]}(s_i)$ have one physical index s_i and two auxiliary indices. They can be graphically represented as three-legs shapes connected with lines, i.e. contracted along auxiliary indices (see Fig. 2.1) [27, 37]. These

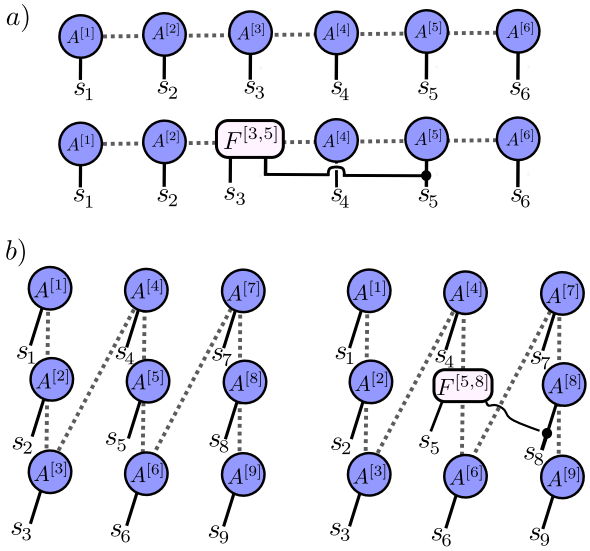

Figure 2.1: Graphical representations of MPS and MPBS applied to 1D (a) and 2D (b) quantum many-body systems. The F tensors encode correlations between different lattice sites. For illustrative purposes, the pictures represent only one of the terms in which F is involved.

indices run from 1 to a set of integers χ_i, called local bond dimensions, fixing the maximum amount of entanglement entropy which can be encoded by the state [27].

MPS can provide good approximations of low entangled states, as for instance ground states of local gapped Hamiltonians in 1D, for which an area law for entanglement entropy can be proven [87] (see Chapter 1). On the contrary, MPS cannot efficiently encode a volume law, since this would require an exponentially large value of χ.

In order to overcome these limitations, we introduce a new set of tensors $F^{[i,j]}_{\alpha,\alpha'}(s_i, s_j)$ with two physical indices s_i, s_j and two auxiliary indices α, α'. These tensors will encode correlations between different lattice sites i, j. We propose a new class of wave functions $\psi[A, F](\mathbf{s})$ obtained by formally replacing the MPS local tensors $A^{[l]}(s_l)$ with new tensors $\tilde{A}^{[l]}(\mathbf{s})$ as follows

$$A^{[l]}_{\alpha_{l-1}\alpha_l}(s_l) \rightarrow \tilde{A}^{[l]}_{\alpha_{l-1}\alpha_l}(\mathbf{s}) = A^{[l]}_{\alpha_{l-1}\alpha_l}(s_l) + \sum_{i_l \neq l} F^{[l,i_l]}_{\alpha_{l-1}\alpha_l}(s_l, s_{i_l}) \,. \tag{2.1.2}$$

Notice that the matrices $\tilde{A}^{[l]}$ now depend explicitly on all the global set of quantum numbers $\mathbf{s}$. The new wave function is

$$\psi[A, F](\mathbf{s}) = \tilde{A}^{[1]}(\mathbf{s})\tilde{A}^{[2]}(\mathbf{s})...\tilde{A}^{[N]}(\mathbf{s}) \,, \tag{2.1.3}$$

where we summed over all auxiliary indices α_l ($l = 1, 2...N-1$). This object can be considered conceptually similar to the well-known backflow wave function in electronic structure theory, which is commonly used to introduce correlations in the mean-field theory by taking the single-particle orbitals act on a configuration-dependent quasi-particle positions [84–86, 88]. In our case, the starting point is not a mean-field wave function, but rather an MPS,

which can be seen as a systematic and general improvement of the mean-field approximation. We thus name this class of variational states Matrix Product Backflow States (MPBS).

It is worth mentioning that MPBS wave functions admit a series expansion in increasing powers of F, where each term can be formally recast as an MPS with locally larger bond dimension, up to 2^n times the original bond dimension at the order n (see next Section). Examples of first-order ($n = 1$) terms are depicted in Fig. 2.1.

MPBS with F connected as in Fig. 2.1b) will be used in the next Sections to simulate 2D systems. The idea is that this type of connections provides a sort of "shortcut" for entanglement since they allow for the direct correlation of nearest-neighbor sites in the 2D geometry that are, however, placed far apart from each other in the effective 1D mapping (leading therefore to an undesired increase in the required bond dimension). Indeed, MPBS arranged as in Fig. 2.1b) satisfy an area law for the entanglement entropy, since any possible bipartition of the lattice cuts a number of auxiliary bonds and/or F tensors that grows linearly with the length of the subsystem perimeter. Remarkably, it can be easily proven that MPBS ability to encode entanglement can be even greater, since with a particular choice of the parameters one can encode a volume law for the entanglement entropy. Thus, MPBS can in principle provide good approximations not only of ground-states in 2D, but also of highly entangled quantum many-body states, as for instance time-evolved states after quantum quenches [89].

From an operative perspective, MPBS naturally suggest the adoption of the following two-step optimization algorithm:

1. first, the local A tensors are optimized with standard MPS techniques;

2. second, the non-local F tensors are optimized by means of VMC techniques.

This alternating optimization approach offers the advantage that the initial point for VMC stochastic optimization is not a random point in the parameter space, but instead it is already a reasonably good approximation of the quantum many-body wave function. Moreover, VMC optimization can further optimize the A tensors as well, thus providing an unrestricted variational search for our ansatz in the last optimization stage. Finally, the MPBS network can be exactly contracted during the Monte-Carlo steps (in contrast to other similar approaches where approximated contraction schemes are employed [82]), leading to a purely variational scheme. In the following Sections, we will focus on the ground-state search problem, benchmarking the MPBS ansatz on both 1D and 2D models. The numerical results are obtained by means of the two steps optimization algorithm just outlined. In Appendix A, we provide a short introduction to the VMC method.

2.1.1 Decomposing the MPBS as a sum of MPS

As mentioned, the MPBS wavefunction can be expanded in a series of terms at different orders in powers of F, i.e.

$$\psi[A,F](\mathbf{s}) = \sum_{n=0}^{N} \psi^{(n)}[A,F](\mathbf{s})\,.$$

The first terms are

$$\begin{aligned}\psi[A,F](\mathbf{s}) = &\underbrace{A^{[1]}_{a_1}(s_1)\dots A^{[N]}_{a_N}(s_N)}_{\psi^{(0)}[A](\mathbf{s})} + \\ &+ \underbrace{\sum_{i\neq j}\left(A^{[1]}_{a_1}(s_1)\dots F^{[i,j]}_{a_{i-1}a_i}(s_i,s_j)\dots A^{[N]}_{a_N}(s_N)\right)}_{\psi^{(1)}[A,F](\mathbf{s})} \\ &+ o\left(F^2\right)\,.\end{aligned} \tag{2.1.4}$$

The zero-order term $\psi^{(0)}$ corresponds exactly to the MPS wave function, while the first-order term $\psi^{(1)}$ is composed of N components. In a general term $\psi^{(n)}$ in the expansion, we would encounter n F tensors, each linking a pair of physical indices. The term $\psi^{(N)}$ can give rise to the well-known Jastrow wave function as a particular case (when setting $\chi = 1$, $A = 0$)

$$\psi_{\text{Jastrow}}(\mathbf{s}) = \prod_{i<j} F^{[i,j]}(s_i,s_j)\,. \tag{2.1.5}$$

Let us now focus on the linear term. We can reshape the tensor F into a square matrix of dimension 2χ

$$F^{[i,j]}_{a_{i-1}a_i}(s_i,s_j) = F_{(a_{i-1}s_i),(a_is_j)}\,.$$

By applying a QR-decomposition, we obtain

$$F_{(a_{i-1}s_i),(a_is_j)} = Q_{(a_{i-1}s_i),\alpha}R_{\alpha(a_is_j)} = Q_{a_{i-1},\alpha}(s_i)R_{\alpha a_i}(s_j)\,, \tag{2.1.6}$$

where the index α runs from 1 to 2χ. For simplicity, let us consider the case $j = i+2$. By combining the tensors R and $A^{[i+1]}$, and applying another QR decomposition, one can obtain a new MPS with a bond dimension increased to 2χ (see Fig. 2.2). It is clear that by means of these tricks one can re-write $\psi^{(1)}[A,F](\mathbf{s})$ as a sum of MPS with maximum bond dimension equal to 2χ. By summing these N MPS, we will get an MPS with maximum bond dimension $2N\chi$, which can be eventually compressed. With similar tricks, higher order terms can be formally recast in an MPS with locally larger bond dimension. In particular, it is not difficult to realize that $\psi^{(n)}$ contains terms, as the one represented in Fig. 2.3, that give rise to a local bond dimension $2^n\chi$.

Finally, let us notice that our proposed method to optimize the MPBS wave function, allows to take into account all the orders of the F tensors. This is possible because the full MPBS wave function can be exactly contracted, meaning that given a spin configuration $\mathbf{s}$ one can easily compute

the $\psi[A,F](\mathbf{s})$ as a product of matrices (exploiting Eq. 2.1.3). Therefore our numerical results can be considered as fully "non-perturbative" and cannot be replicated by using only some of the first terms of the expansion in Eq. 2.1.4.

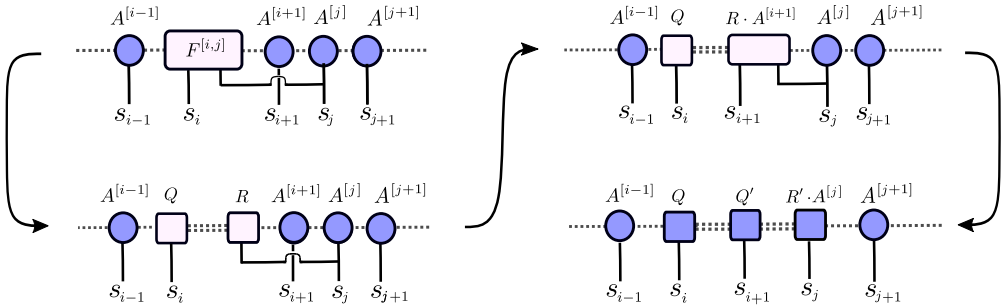

Figure 2.2: Decomposition of one term of $\psi^{(1)}[A,F]$ as MPS. We arbitrarily set $j = i+2$. Double dotted lines represent auxiliary bonds/indices with local bond dimension 2χ, whereas dotted lines have bond dimension χ. The first passage represents Eq. 2.1.6.

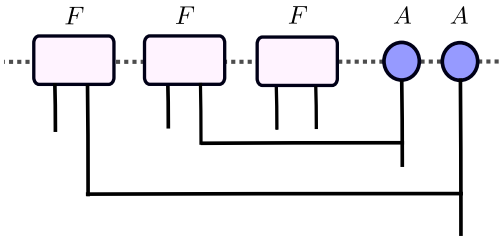

Figure 2.3: A term contained in $\psi^{(n)}$ ($n = 3$) with n overlapping long-range connections.

2.1.2 MPBS encoding a volume law for the entanglement

The aim of this section is to show, by means of an explicit example, that MPBS can efficiently encode a volume law for the entanglement entropy, thus representing an excellent generalization of MPS. In particular, we will focus on MPBS as readjusted to be applied on 2D systems (see Fig. 2.1b)). For this purpose, we will first consider an example of Restricted Boltzman Machine (RBM) state given in [90], showing that it can be rewritten as MPBS. Then, we will follow the proof given in [90] to demonstrate that this wave function, with a particular choice of the parameters, can encode a volume law for entanglement entropy in the two dimensional geometry. To begin, let us write the generic expression for the RBM representation of a quantum state ψ, i.e.

$$\psi_{\mathrm{RBM}}(\mathbf{s}) = \sum_{\mathbf{h}\in\{-1,+1\}^M} \exp\left[\sum_{i=1}^{N} a_i s_i + \sum_{m=1}^{M} h_m b_m + \sum_{i,m} W_{im} s_i h_m\right] . \quad (2.1.7)$$

Here, the wave function value $\psi_{\mathrm{RBM}}(\mathbf{s}) = \sum_{\mathbf{h}} \exp\left(-\tilde{H}(\mathbf{s},\mathbf{h})\right)$ for the spin configuration $\mathbf{s}$ is computed as a Boltzmann weight marginalized over a collection of 'hidden' fictitious spin variables $h_m \in -1,+1$ ($m = 1,2...M$). The parameters a,b,W, generally having complex entries, serve as free parameters in the ansatz. They correspond, respectively, to the external magnetic fields and the couplings of the fictitious Ising-like Hamiltonian $\tilde{H}(\mathbf{s},\mathbf{h})$. We will set $a_i = 0$, $\forall i$. If the couplings W are short-range, it can be demonstrated that the RBM wave function satisfies an area law for the entanglement entropy, as shown in Ref. [90]. However, when the couplings W are long-range, this is no longer true. For our purpose, let us now set $N = N_x N_y$ and $M = (N-1)+N_y(N_x-1) = 2N-N_y-1$. It is useful to split the hidden neurons in two sets, i.e. h_l ($l = 1,2...N-1$) and $\tilde{h}_n$ ($n = 1,2...N_y(N_x-1)$), and define the respective couplings as $W_{il} = W(\delta_{i,l}+\delta_{i,l+1})$ and $\tilde{W}_{in} = W(\delta_{i,n}+\delta_{i,n+N_y})$. The physical meaning under these choices is schematically represented in Fig. 2.4, where visible (hidden) variables are colored blue (yellow) and black lines represents W connections. The idea is to use $\tilde{h}_n$ variables to give rise to MPBS terms connecting spins on the same row but different columns, whereas h_l variables will contribute to the formation of the "snaking path" structure within the MPBS.

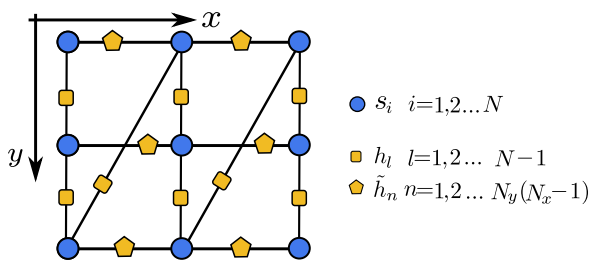

Figure 2.4: Graphical representation of the RBM state considered in the state. Visible layer variables (blue dots) are arranged in a two dimensional grid of shape N_x, N_y. Hidden variables (yellow shapes) give rise to the connections.

By explicitly calculating the sum over the hidden variables in Eq. 2.1.7, we obtain that the wave function is the product of two terms $\psi_{\mathrm{RBM}}(\mathbf{s}) = \phi_1(\mathbf{s})\phi_2(\mathbf{s})$, where

$$\phi_1(\mathbf{s}) = \prod_{l=1}^{N-1}\left(2\cosh\left[b_l + W(s_l + s_{l+1})\right]\right) = \prod_{l=1}^{N-1} T_1(s_l, s_{l+1})$$
$$\phi_2(\mathbf{s}) = \prod_{l=1}^{N_y(N_x-1)}\left(2\cosh\left[\tilde{b}_l + W(s_l + s_{l+N_y})\right]\right) = \prod_{l=1}^{N_y(N_x-1)} T_2(s_l, s_{l+N_y})\ .$$

Let us consider the first one. We can set $T_1(s_l, s_{l+1}) = \left(\mathbb{v}^{(l)}\right)^T \mathbb{w}^{(l+1)}$, where

$$\left(\mathbb{w}^{(l+1)}\right)^T = \left(\exp\left[b + W s_{l+1}\right],\ \exp\left[-b - W_l s_{l+1}\right]\right)$$
$$\left(\mathbb{v}^{(l)}\right)^T = \left(\exp\left[W s_l\right],\ \exp\left[-W s_l\right]\right)\ .$$

Thus, we get

$$\phi_1(\mathbf{s}) = \left(\mathbb{v}^{(1)}\right)^T \mathbb{w}^{(2)} \left(\mathbb{v}^{(2)}\right)^T \mathbb{w}^{(3)} ... \left(\mathbb{v}^{(N-1)}\right)^T \mathbb{w}^{(N)} = A^{[1]} A^{[2]} ... A^{[N]} \ ,$$

where we defined the following matrices

$$A^{[l]} = \begin{cases} \left(\mathbb{v}^{(1)}\right)^T & \text{if } l = 1 \\ \mathbb{w}^{(l)} \left(\mathbb{v}^{(l)}\right)^T & \text{if } l = 2 ... N-1 \\ \mathbb{w}^{(N)} & \text{if } l = N \end{cases} \ .$$

Let us observe that $A^{[l]}$ depends only on the local physical variable s_l, therefore $\phi_1(\mathbf{s})$ is in the form of an MPS, with bond dimension $\chi = 2$. The MPS follows the snaking-path, as displayed in Fig. 2.4. The wave function can now be written as

$$\begin{aligned} \psi_{RBM}(\mathbf{s}) = & \mathbb{A}(s_1) T_2(s_1, s_{1+N_y}) \mathbb{A}(s_2) T_2(s_2, s_{1+N_y}) ... \mathbb{A}(s_{N-N_y}) \cdot \\ & \cdot T_2(s_{N-N_y}, s_N) \, \mathbb{A}(s_{N-N_y+1}) ... \mathbb{A}(s_N) \ , \end{aligned} \tag{2.1.8}$$

where we used bold letters to distinguish the matrices, i.e. objects with two virtual indices, from scalars, i.e. objects with no virtual indices. We can reabsorb the scalars in the matrices $\mathbb{A}$, obtaining

$$\begin{aligned} \psi_{RBM}(\mathbf{s}) = & \left(\mathbb{A}(s_1) + \mathbb{F}^{[1]}(s_1, s_{1+N_y}) \right) ... \left(\mathbb{A}(s_{N-N_y}) + \mathbb{F}^{[N-N_y]}(s_{N-N_y}, s_N) \right) \cdot \\ & \cdot \mathbb{A}(s_{N-N_y+1}) ... \mathbb{A}(s_N) \ , \end{aligned} \tag{2.1.9}$$

where we defined the $\mathbb{F}$ matrices as follows

$$\mathbb{F}^{[l]}(s_l, s_{l+N_y}) = \mathbb{A}(s_l) \left(\sqrt{T_2(s_l, s_{l+N_y})} - 1 \right)$$

for $l = 1, 2 ... N - N_y$. Eq. 2.1.9 is just a particular case of our MPBS representation. Let us now set the RBM parameters as follows: $b_l = -\frac{i\pi}{4}$, $\tilde{b}_n = \frac{i\pi}{2}$ and $W = \frac{i\pi}{4}$ (see Ref. [90]). It is not difficult to realize that with this particular choice, one has $\phi_1(\mathbf{s}) = \pm c \, \forall \mathbf{s}$, and $T_2(s, s') = \pm c' \delta_{s,s'}$, where c and c' are numerical constants and the sign depends on the spins. Thus, the RMB state takes the form $|\psi_{RBM}\rangle = \sum_{\mathbf{s}_c} \pm C \, |\mathbf{s}_c ... \mathbf{s}_c\rangle$, where C is a constant and $\mathbf{s}_c$ is the spin configuration of the first column (i.e. $\mathbf{s}_c = (s_1, s_2 ... s_{N_y})$). This implies that the state $|\psi_{RBM}\rangle$ is a superposition with equal weights of all basis states corresponding to spin configurations where the N_x columns all have the same configuration $\mathbf{s}_c$. Let us now consider a bipartition of our 2D system A and B. The subsystem density matrix ρ_A can be computed as

$$\rho_A = \mathrm{tr}_B \left[|\psi_{RBM}(\mathbf{s})\rangle \langle \psi_{RBM}(\mathbf{s})| \right] = \sum_{\mathbf{s}_c, \mathbf{s}'_c} C^2 \sum_{\mathbf{s}_B} \langle \mathbf{s}_B | \mathbf{s}_c ... \mathbf{s}_c \rangle \langle \mathbf{s}'_c ... \mathbf{s}'_c | \mathbf{s}_B \rangle \ .$$

If subsystem B contains at least one whole column, then fixing $\mathbf{s}_c$ also fix $\mathbf{s}_B$ (and $\mathbf{s}'_c$). Therefore, we get $\rho_A \propto \sum_{\mathbf{s}_c} |\mathbf{s}_A\rangle \langle \mathbf{s}_A|$. Reintroducing the overall normalization constant, we get $\rho_A = \mathbb{1}/2^{|A|}$, where $|A|$ indicates the total number of lattice sites within the region A. This result shows that A and B are maximally entangled and therefore the Rényi entropies associated with

the bipartition are $S_\alpha(A|B) = |A| \ln 2$, for all Rényi indices α. These observations provide us a simple example of a wave function in the form of an MPBS efficiently encoding a volume law of the entanglement entropy in the 2D geometry (the number of non-zero parameters A, F scale polinomially with N).

2.2 NUMERICAL RESULTS

In this Section, we provide numerical results concerning the ground-state search problem for some 1D and 2D models. In all the simulations, DMRG was used as a first optimization step. The second stage involved the VMC optimization (see Appendix A) and was implemented using the NetKet library [91–93]. NetKet is an open-source project delivering cutting-edge methods for the study of many-body quantum systems with artificial NN and machine learning techniques. It is a Python library built on JAX [94], a package providing Automatic Differentiation routines. NetKet offers several methods to define custom models. The MPBS model was implemented by using the Flax Linen API framework [95]. Given a particular model, namely a parameter dependent ansatz $|\psi(\boldsymbol{\theta})\rangle$, the VMC approach use a stochastic Markov chain to generate batches of system configurations $\{\boldsymbol{s}^{(n)}\}_{n=1}^{\mathcal{N}}$ according to the probability distribution $p(\boldsymbol{s}) = |\langle \boldsymbol{s}|\psi(\boldsymbol{\theta})\rangle|^2 / ||\,|\psi(\boldsymbol{\theta})\rangle\,||^2$ [96]. The expectation value of the energy and its gradient with respect to the variational parameters $\boldsymbol{\theta}$, as well expectation values of other observables, are computed as statistical averages over a large-enough number of sampled system configurations. Having estimated $\mathcal{E}$ and the derivatives $\partial_{\boldsymbol{\theta}}\,\mathcal{E}$, the VMC approach uses some optimization algorithm to update the parameters $\boldsymbol{\theta}$ and iterate over a number of Monte-Carlo steps. In particular, NetKet makes use of the Stochastic Gradient Descent, with the Stochasitc-Reconfiguration preconditioner. The latter is also named Natural Gradient Descent in the machine learning literature [97] and allows to adjust the gradient descent trajectory by taking into account the geometry of the quantum state manifold. As sampler routines to generate the system configurations, we used or the standard local Metropolis algorithm [96] or, in the cases in which the total magnetization was keep fixed, a variant in which new system configurations are generated by flipping couples of spins, thus preserving the total $z-$magnetization. The number of samples $\mathcal{N}$ per Monte-Carlo step was chosen between 10^3 to 10^4, depending on the simulation. See Appendix A for a brief survey on VMC. At the beginning of the optimization, F tensors were initialized as follows

$$F^{[i,j]}(s_i, s_j) = \epsilon(\mathbb{1}_{\chi\times\chi} + \sigma\,\zeta_{\chi\times\chi})\,, \tag{2.2.1}$$

where $\mathbb{1}_{\chi\times\chi}$ is the identity matrix of dimension χ, $\zeta_{\chi\times\chi}$ is a random matrix with normally distributed entries and $0 < \sigma, \epsilon \ll 1$ are (small) real numbers (usually we set to $\epsilon = 0.01$ or 0.005 and $\sigma = 0.1$). After, the tensors F were considered as parameters of the ansatz, that is $\boldsymbol{\theta} \equiv F$, and optimized as sketched before. In the 2D simulations, we also included the A tensors as variational parameters, i.e. $\boldsymbol{\theta} \equiv (A, F)$, therefore allowing the algorithm to

further optimize the already optimized MPS tensors.

Concerning the computational cost of our proposed method, let us first mention that the computational cost of DMRG algorithm is proportional to χ^3, where χ is the bond-dimension. The computational cost of the VMC optimization is proportional to $N_{\text{iterations}}\mathcal{N}\chi^3$, where $N_{\text{iterations}}$ is the total number of Monte-Carlo iterations and $\mathcal{N}$ is the number of samples per Monte-Carlo step. The factor χ^3 comes from the (exact) evaluation of the MPBS wave-function. Indeed, given a spin configuration $\mathbf{s}$, one can easily evaluate the wave-function $\psi[A, F](\mathbf{s})$ as a product of N matrices each of size $\chi \times \chi$, N being the system size. In order to have the MPBS simulation competitive with the DMRG in terms of CPU time, it is therefore required to have $\chi_{\text{MPBS}} \ll \chi_{\text{MPS}}$. This is the case, since in 2D (or in higher dimensions), MPS are expected to require an exponentially large amount of resources to accurately encode physical states ($\chi_{\text{MPS}} \sim o(\exp(N))$) which can be instead compactly described by our ansatz. Furthermore, our simulations show that relatively small values of χ_{MPBS} allow to obtain accurate simulations of 1D systems.

2.2.1 Modified Haldane-Shastry model

First, we apply MPBS to a 1D quantum spin chain with Periodic Boundary Conditions (PBC). In particular, we consider the following modified Haldane-Shastry (HS) model

$$\hat{H}_{HS} = \sum_{j<i} \frac{1}{\widetilde{d}_{ij}^2}\left(-\hat{\sigma}_i^x\hat{\sigma}_j^x - \hat{\sigma}_i^y\hat{\sigma}_j^y + \hat{\sigma}_i^z\hat{\sigma}_j^z\right), \tag{2.2.2}$$

where $\widetilde{d}_{ij} = N/\pi \cdot \sin(\pi/N|i-j|)$. This model is known to be particularly challenging for standard DMRG, as it shows power-law scaling in the ground-state entanglement entropy [90]. To use our optimization scheme, we adapt the MPBS ansatz in order to explicitly realize translational invariance. This is achieved by adding an extra auxiliary index, connecting the first and the last site, as well as by taking the A tensors independent from the site i. Also, we set $F^{[i,j]}$ to be dependent only on the distance $d_{ij} = \min(|i-j|, N-|i-j|)$ between the two connected sites. We also introduced a cut-off r_c setting the maximum distance between sites for which the F tensors are non-zero (i.e. $F^{[d_{ij}]}(s_i,s_j) = 0$ if $d_{ij} > r_c$). Due to translational invariance and the imposed cutoff, the number of variational parameters of the ansatz is independent of the system size N, resulting in a reduced computational cost for the Monte-Carlo simulation. In the first optimization stage, we write the HS Hamiltonian as a Matrix Product Operator (MPO) and use standard two-sites DMRG [27] to get the optimized A tensors. In the second stage, a VMC optimization of the F tensors is realized, adopting the Stochastic Reconfiguration [97] *natural* gradient descent approach. Since H_{HS} commute with the total $z-$magnetization $\hat{\Sigma}^z = \sum_{i=1}^N \hat{\sigma}_i^z$ and the parity operator $\hat{P} = \hat{\sigma}_1^x\hat{\sigma}_2^x ... \hat{\sigma}_N^x$, we restrict the ground-state search to the $\hat{\Sigma}^z = 0$ sector of the Hilbert space. In Fig. 2.5, we show some selected results, obtained with a relatively small value of the MPBS bond dimension ($\chi = 5$) and $r_c = 3$. First subplot shows

the expectation value and the variance of the energy, tracked during the VMC optimization (red lines). Dotted lines represent DMRG energies/variances for increasing values of the bond dimension. After less than 10^2 VMC optimization steps the MPBS energy reach energy values smaller than the DMRG energy obtained with the larger value of χ ($\chi = 70$). Let us remark that number of parameters this MPS is much larger than the number of them parameters of our ansatz, meaning that MPBS provide good approximations of the true quantum many-body ground-state. Moreover, we also got substantially better results in terms of energy variance. Second subplot in Fig. 2.5 shows the two-points connected correlator $\left(c_{xx} + c_{yy}\right)_c$, computed by taking average of $\langle \hat{\sigma}^x_i \hat{\sigma}^x_{i+r} \rangle + \langle \hat{\sigma}^y_i \hat{\sigma}^y_{i+r} \rangle$ over i and then subtracting the square of the average x and y magnetizations. Red points represent estimations obtained at the end of the VMC optimization, whereas other points are DMRG results. These seem to converge to VMC values, when increasing the bond dimension χ. In the inset it is shown the correlator $c_{xx}(N/2) + c_{yy}(N/2)$ as estimated during the Monte-Carlo iterations. The convergence appears to be fast.

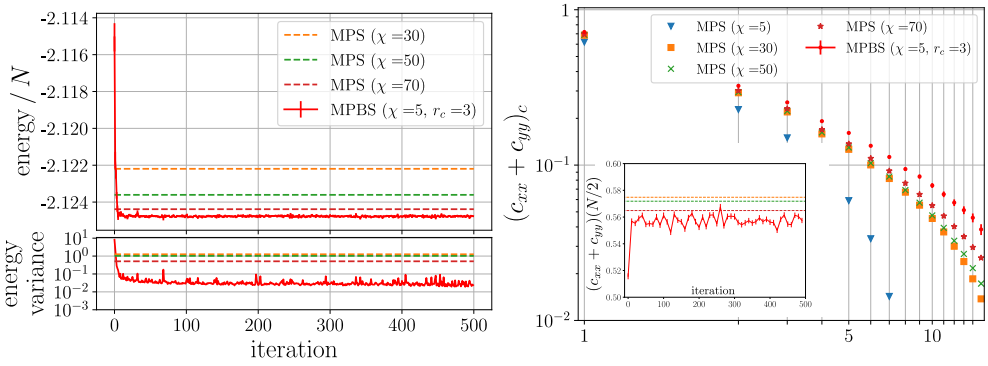

Figure 2.5: MPBS tested of bond dimension $\chi = 5$ the modified 1D HS model: energy density convergence (left) and $(c_{xx} + c_{yy})_c$ connected correlator (right). The system size is $N = 70$.

In Fig. 2.6, we show the two-points correlators $c_{zz}(r)$ as estimated at the end of the VMC optimization (red points). Since translational invariance is explicitly realized by our ansatz, a spatial average over was taken, i.e. we considered $c_{zz}(r) = 1/N \sum_i \langle \hat{\sigma}^z_i \hat{\sigma}^z_{i+r} \rangle$. Results are substantially in agreement with the ones displayed in Ref. [90]. The other points represent DMRG results for increasing values of the bond dimension χ. In the inset, we plot the absolute value of the difference between the DMRG correlators and the estimated VMC correlators. By increasing χ, the DMRG correlators goes monotonically to the VMC values.

2.2.2 Two dimensional Ising model

To corroborate the flexibility of MPBS in describing higher dimensional systems, we now start analyzing 2D quantum many-body models living on a square lattice of size $N_x \times N_y$ with Open Boundary Conditions (OBC). A simple way to adapt MPS to the description of such a system is to order the sites of the grid following a one-dimensional "snaking path" connecting all

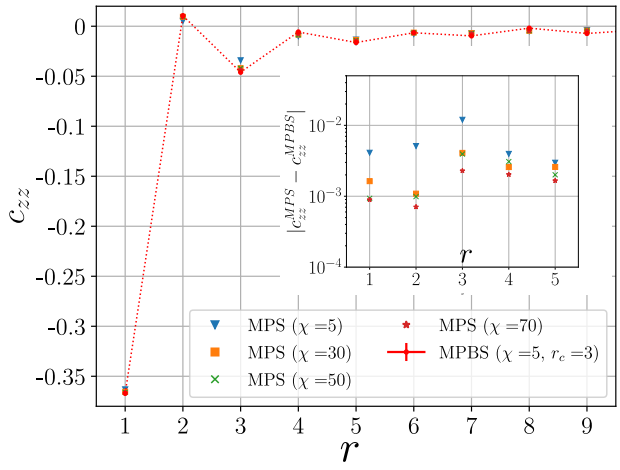

Figure 2.6: The spin-spin correlator $c_{zz}(r)$ of the 1D HS Model ($N = 70$). The inset represent $|c_{zz}^{MPS}(r) - c_{zz}^{MBPS}(r)|$.

the sites (see Fig. 2.1 b)) [98]. Others 2D to 1D mappings have been also studied [99], leading to increased numerical precision but not to a significant improvement in the codification of entanglement in 2D systems. The main issue is that, since area law in 2D implies that entanglement entropy grows linearly with the length of the subsystem perimeter, any MPS cannot describe efficiently typical ground-states of 2D Hamiltonians. As a possible improvement, we propose to arrange the MPBS ansatz in order to codify correlations between sites which are adjacent in the 2D geometry but which are placed at distance N_y along the 1D snaking path. This can be done by setting the $F^{[i]}(s_i, s_j)$ matrices different from zero in the cases in which $j = i \pm N_y$, where we label the lattice sites with a single integer $i = 1, 2, \dots N$. As already mentioned, an MPBS of this kind can encode the area law for the entanglement entropy and, at least for a particular choice of the parameters, the volume law. To benchmark the efficacy of MPBS in simulating 2D systems, we consider the following Ising Hamiltonian

$$\hat{H} = -\sum_{\langle \mathbf{i},\mathbf{j}\rangle} \hat{\sigma}^z_{\mathbf{i}} \hat{\sigma}^z_{\mathbf{j}} + h \sum_{\mathbf{i}} \hat{\sigma}^x_{\mathbf{i}} \,, \tag{2.2.3}$$

on a lattice of dimension $N_x = N_y = 11$. For the DMRG optimization, we used the MPO representation of 2D Hamiltonians discussed in Sec. 1.5.1. In Fig. 2.7, we show the results of an MPBS optimization with bond dimension $\chi = 5$ and transverse field $h = 3.0$, close to the quantum critical point of the system $h_c \simeq 3.044$ [100]. These results are compared with DMRG findings at different bond dimensions and with the energy value obtained by Lubasch and others by means of PEPS [101]. As in the previous case, MPBS with extremely small bond dimension leads, after ≈ 100 VMC optimization iterations, to results significantly better than DMRG, both in terms of energy density and energy variance. Since the system has rotational symmetry, during the last $\simeq 150$ Monte-Carlo iterations we explicitly symmetrize the MPBS with respect to the $\mathcal{C}_4$ group of fourfold rotations. To do this, we consider the following modified wave function $\psi'[A, F](\mathbf{s}) = \sum_{k=0}^{3} \psi[A, F](R^k \mathbf{s})$, where R

is a rotation of $\pi/2$ of the spin configuration. This results in a further improvement of the energy and energy variance. The value of energy density we find at the end of the optimization is $\langle \hat{H} \rangle / N = -3.17208(1)$. In the second subplot, we show the correlator $c_{zz}(r) = 1/N_r \cdot \sum_{\mathbf{r},|\mathbf{r}|=r} \langle \hat{\sigma}^z_{\mathbf{i}_c} \hat{\sigma}^z_{\mathbf{i}_c+\mathbf{r}} \rangle$, where $\mathbf{i}_c$ indicates the central site of the grid and N_r is the number of sites placed at distance r from this. MPBS points seem to be in good agreement with the trend of DMRG results for increasing bond dimension.

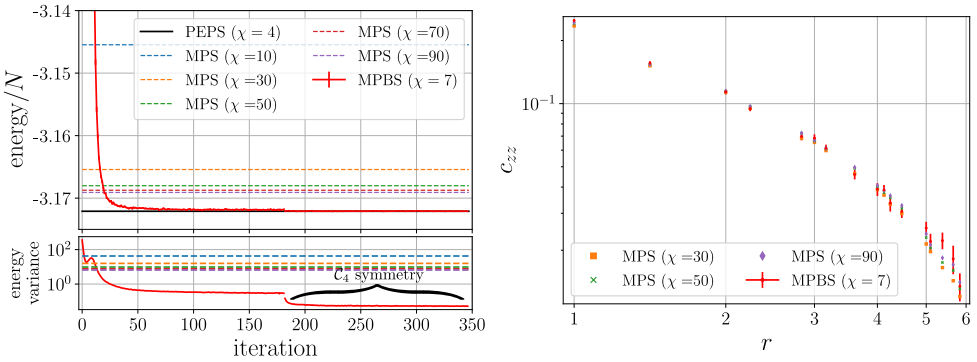

Figure 2.7: MPBS of bond dimension $\chi = 7$ tested with the 2D Ising model on a square lattice ($N_x = 11$, $N_y = 11$): energy density convergence (left) and c_{zz} correlator (right).PEPS result is taken from [101].

2.2.3 Two dimensional $J_1 - J_2$ model

Finally, we consider the anti-ferromagnetic $J_1 - J_2$ model, with Hamiltonian

$$\hat{H} = J_1 \sum_{\langle \mathbf{i},\mathbf{j} \rangle} \sigma_{\mathbf{i}} \cdot \sigma_{\mathbf{j}} + J_2 \sum_{\langle\langle \mathbf{i},\mathbf{j} \rangle\rangle} \sigma_{\mathbf{i}} \cdot \sigma_{\mathbf{j}} \,, \tag{2.2.4}$$

where the first (second) sum is on first (second) nearest neighbors couples of sites. This is a prototypical frustrated magnetic system. Despite active research in the past decades [102–105], the nature of the ground-state around the point of maximum frustration $J_2/J_1 = 0.5$ remains unclear. We address the problem Hamiltonian by means of MPBS arranged as in the previous paragraph and also adding F tensors connecting second nearest neighbors sites. As in the HS model, we reduce the simulation to the zero magnetization sector. In Fig. 2.8, we show some selected results obtained with a system of size $N_x = N_y = 8$, OBC and $J_1 = 1$, $J_2 = 0.5$. After $\simeq 350$ VMC optimization iterations, we apply $\mathcal{C}_4$ wave function symmetrization. We compare our results with the EPS and PEPS results reported in [79] and with Monte-Carlo optimized PEPS results reported in [105]. The final energy density of our simulation is $\langle \hat{H} \rangle / N = -1.9273(9)$ and is lower than both values reported in [79], whereas it is about $\approx 7 \cdot 10^{-3}$ greater than the value reported in [105]. It should be however remarked that the value in [105] is not strictly vari-

ational, because of the approximate contraction scheme adopted for PEPS. Finally, we measure some relevant observables as the correlators

$$\begin{aligned} c_{ver}(r) &= \frac{1}{N_x} \cdot \sum_j \langle \boldsymbol{\sigma}_{1,j} \cdot \boldsymbol{\sigma}_{1+r,j} \rangle \\ c_{hor}(r) &= \frac{1}{N_y} \cdot \sum_i \langle \boldsymbol{\sigma}_{i,1} \cdot \boldsymbol{\sigma}_{i,1+r} \rangle \end{aligned} \tag{2.2.5}$$

which are shown in the second half of Fig. 2.8. These are respectively the average spin-spin correlators along the columns and the rows of the grid. Since the wave function $\psi'[A, F](\mathbf{s})$ is symmetric under rotation of $\pi/2$, we always find values for these correlators compatible within the uncertainty bars. On the contrary, DMRG results show that MPS are unable to encode power-law decaying correlations along the horizontal direction. We also measure the structure factor

$$S^2(\mathbf{q}) = \frac{1}{(N(N+2)} \sum_{\mathbf{i},\mathbf{j}} \langle \boldsymbol{\sigma}_{\mathbf{i}} \cdot \boldsymbol{\sigma}_{\mathbf{j}} \rangle \, e^{-i\mathbf{q}\cdot(\mathbf{i}-\mathbf{j})} \tag{2.2.6}$$

or different pitch vectors $\mathbf{q}$. We find $S^2(0,\pi) \simeq 3.19(5) \cdot 10^{-2}$ and $S^2(\pi,\pi) \simeq 0.241(3)$. The latter corresponds to the Néel order parameter. Both values are compatible with similar findings in [104]. We also obtain $S^2(0,0) = 1.3(2) \cdot 10^{-4}$, which is consistent with the expectation that the $J_1 - J_2$ ground state is in a singlet under SU(2) global symmetry.

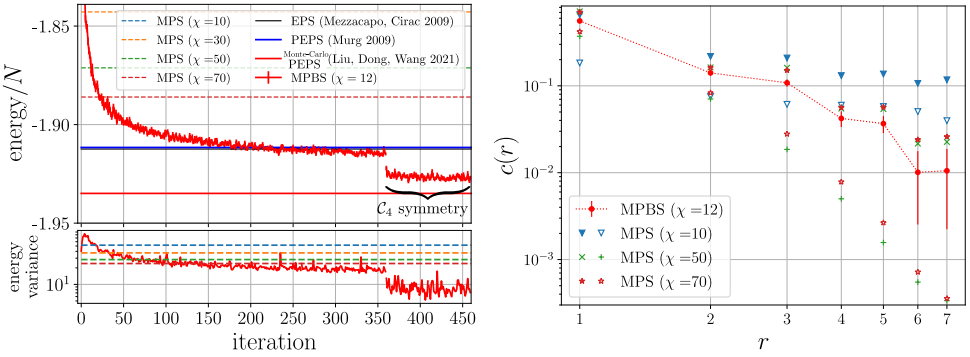

Figure 2.8: MPBS of bond dimension $\chi = 12$ tested with the $J_1 - J_2$ model on a square lattice ($N_x = 8$, $N_y = 8$). PEPS and EPS results are taken from Ref.[79]. In the right plot, filled (empty) markers represent DMRG results for $c_{ver}(r)$ ($c_{hor}(r)$), whereas red points are MPBS results.

3 QUANTUM ANNEALING FOR PERCEPTRON OPTIMIZATION: A NEW APPROACH VIA MPS SIMULATIONS

Numerous computational and physical problems can be reformulated as non-convex optimization problems involving a specific Ising-like cost function of the form $H(\mathbf{s}) = H(s_1, \ldots, s_N)$, which depends on N binary variables $s_i = \pm 1$ [106]. Finding an exact solution implies the daunting task of a discrete search in a space with exponentially-growing dimension in the system size N. Therefore, optimization is typically approached using heuristic minimization algorithms. Among these, a growing interest is devoted to quantum optimization, which aims at exploiting quantum effects, namely quantum superposition and entanglement, to obtain a wave function with a large overlap with classical solutions. Alongside with well-established schemes such as Quantum Annealing (QA) [107–111] and Adiabatic Quantum Computation (AQC) [112, 113], implemented in analogue dedicated hardware [114], recent approaches encompass the design of parameterized quantum circuits, implemented on a digital quantum device, which are run in loop with a classical computer in Variational Quantum Algorithms (VQA) [115].

A conceptual preliminary step in quantum optimization is to map classical spins on quantum spin-1/2 Pauli operators $\hat{\sigma}_j^z$, hence regarding the initial cost function as a quantum Hamiltonian that is diagonal, by construction, in the standard computational basis of quantum computation [116]:

$$H(s_1, \ldots, s_N) \to \hat{H}_z(\hat{\sigma}_1^z, \ldots, \hat{\sigma}_N^z) \,. \tag{3.0.1}$$

Next, in QA a non-commuting driving term, often a transverse field, is introduced, and the quantum Hamiltonian is taken to be

$$\hat{H}(s) = s\,\hat{H}_z + (1-s)\,\hat{H}_x \qquad \hat{H}_x = -\sum_{i=1}^{N} \hat{\sigma}_i^x \,, \tag{3.0.2}$$

with an interpolation parameter $s = s(t) \in [0, 1]$ such that $s(0) = 0$ and $s(\tau) = 1$, τ being the total annealing time. If τ is large enough, compared to the inverse square of the minimal spectral gap of $\hat{H}(s)$, one can rely on adiabatic theorems [113] to prove that the system will be driven into the ground state of the target Hamiltonian $\hat{H}_z$. In its digitized version [117–119] (dQA), the QA Schrödinger dynamics is implemented step-wise, in P

discrete time steps of length $\delta t = \tau/P$, after Trotter splitting the two non-commuting terms $\hat{H}_z, \hat{H}_x$. This leads to a final state of the form

$$|\psi_P\rangle = e^{-i\beta_P \hat{H}_x} e^{-i\gamma_P \hat{H}_z} \cdots e^{-i\beta_1 \hat{H}_x} e^{-i\gamma_1 \hat{H}_z} |\psi_0\rangle \,, \tag{3.0.3}$$

where $|\psi_0\rangle = |\rightarrow\rangle^{\otimes N}$ is the ground state of $\hat{H}_x$, with $|\rightarrow\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$. Assuming a linear annealing schedule $s(t) = t/\tau$ the parameters $\boldsymbol{\beta} = (\beta_1 \cdots \beta_P)$ and $\boldsymbol{\gamma} = (\gamma_1 \cdots \gamma_P)$ are given by $\beta_p = (1 - p/P)\delta t$ and $\gamma_p = (p/P)\delta t$, with $p = 1 \cdots P$. Alternatively, one can regard $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ as 2P variational parameters and optimizing the expectation value of the target Hamiltonian $\hat{H}_z$ over the variational state in Ed. 3.0.3. This approach is dubbed Quantum Approximate Optimization Algorithm (QAOA) [120].

Despite promising results in problem-specific settings, the actual effectiveness and scalability of quantum optimization schemes for classical optimization is still debated. In fact, the quest for quantum speed-ups [121, 122] and the real effectiveness of quantum optimization algorithms should ultimately be tested on real scalable quantum devices, beyond the reach of classical simulations by means of Exact Diagonalization (ED) techniques. To implement this program, however, one encounters two main hurdles.

The first, concerns available experimental platforms for quantum devices: despite major progresses the number of available physical qubits and their connectivity are quite limited. Moreover, experimentally available qubits are very sensible to noise, thus limiting realistic applications to shallow circuits requiring short coherence times. These technical issues severely limit, in practice, the feasibility of quantum simulations beyond the classical limits.

Secondly, the actual implementation of QA [114] on analogue devices, as well as that of digitized Quantum Annealing (dQA) [117] or VQAs [115] on a digital circuit-based quantum computer, usually requires an actual implementation of the unitary time evolution generated by the quantum Hamiltonian in Eq. (3.0.1). This often constitutes a formidable technical challenge: while few problems such as Max-Cut on regular graphs [123] only involve two-body interactions, directly implementable in an analogue/digital device, general optimization tasks usually yield a Hamiltonian $\hat{H}_z$ with non-local k−bodies interactions, hence difficult to implement.

These experimental limitations and theoretical challenges call for efficient classical simulations of quantum optimization protocols, beyond the usual small-scale limits imposed by ED techniques. A prominent family of classical simulation techniques allowing for large-scale simulations of quantum systems is represented by Tensor Networks (see Chapter 1). Major results in this framework include winning strategies for 1D quantum many-body physics [27, 38, 40, 124–126] and, more recently, significant contributions in Machine Learning [127–129] and hybrid quantum-classical algorithms [130]. The goal of Tensor Networks (TN) is to provide an efficient representation of quantum many-body wave functions in the form of a generic network of tensors, connected by means of auxiliary indices [38, 131]. These indices are characterized by a fixed bond dimension, χ, which controls the information content of the network, characterizing its ability to encode entangled states. Matrix Product States (MPS) [27] are the simplest class of TN: an MPS

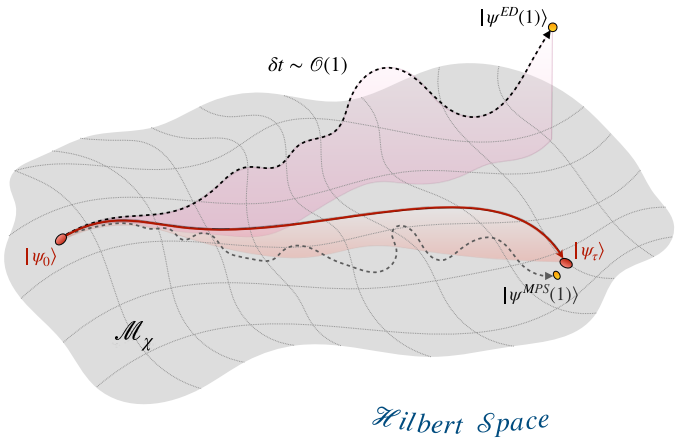

Figure 3.1: The MPS manifold $\mathcal{M}_\chi$, embedded in the exponentially larger Hilbert space. Different QA protocols are represented as trajectories in this space. The exact adiabatic dynamics (full red line) for a large annealing time $\tau \gg 1$ is not generally constrained into the MPS manifold. Nevertheless, it will lead from $|\psi_0\rangle$ (MPS with $\chi = 1$) to a final state $|\psi_\tau\rangle$, which also belongs to the MPS manifold. As soon as the adiabatic-theorem conditions are met, $|\psi_\tau\rangle$ has a big overlap with the ground-state eigenspace spanned by all classical solutions. However, when the dynamical protocol is dQA with a finite time step $\delta t \sim \mathcal{O}(1)$, the actual Trotterized dynamics (black dashed line) may largely deviate from the adiabatic one, due to an unwanted production of extra entanglement (see Figs. 3.10, 3.11 in Sec. 3.2.3 for a discussion of these aspects). Hence, the final state $|\psi^{ED}(1)\rangle$ may lay outside the MPS manifold, and it definitely differs from the final QA state $|\psi_\tau\rangle$, yielding very small overlap with classical solutions. Remarkably, when the Trotterized dynamics is performed within our novel framework (gray dotted line), i.e. constrained to $\mathcal{M}_\chi$, by alternating the Trotter steps with projections into the manifold, the MPS evolved state remains closer to the exact adiabatic dynamics, thus finally leading to a final state $|\psi^{MPS}(1)\rangle$ which is very close to the target state.

wave function is obtained by the multiplication of site-dependent $\chi \times \chi$ matrices $A^{[i]}(s_i)$, each depending on the spin variable $s_i \in \{+1, -1\}$. MPS can be manipulated efficiently in classical numerical simulations by means of well-established algorithms, as the Density Matrix Renormalization Group (DMRG) [27]. We refer the reader to Chapter 1 for a thorough introduction to MPS and the corresponding numerical techniques.

In this Chapter, we present a novel framework to efficiently simulate quantum optimization algorithms for a large class of hard classical optimization problems. We focus on a standard dQA approach by repeatedly applying the unitary operators $e^{-i\gamma_p \hat{H}_z}$, $e^{-i\beta_p \hat{H}_x}$ and iteratively projecting back, at each step, the resulting state on the MPS manifold $\mathcal{M}_\chi$ with a fixed bond dimension χ. A first main result of our work is a theoretical construction that yields an efficient MPS-based representation of dQA for a family of classical cost-function Hamiltonians $\hat{H}_z$, inspired by paradigmatic discrete neural net-

works and encompassing models with non-local multi-spin interactions, up to N-body terms. This results in an efficient algorithm, with computational cost scaling polynomially with the system size N, allowing for classical simulations well-beyond typical sizes analyzed by means of ED techniques.

Our numerical results are two-fold. First, in the regime of small time step $\delta t \ll 1$, where dQA closely approximates the continuous QA dynamics, our approach can systematically reproduce ED simulations of dQA, with a high degree of accuracy. Secondly, in the regime of large $\delta t \sim \mathcal{O}(1)$, characterized by large Trotter errors that spoil the dQA accuracy, we observe that our algorithm can significantly outperform ED simulations of dQA, surprisingly providing far better-quality solutions for the optimization problem. We provide the following interpretation of this unexpected effectiveness (see sketch in Fig. 3.1 and Sec. 3.2.3 for a comprehensive discussion). The initial state $|\psi_0\rangle$ is a trivial MPS of bond dimension $\chi = 1$. Moreover, the final annealed state $|\psi_\tau\rangle$, resulting from the exact QA time evolution with $\tau \gg 1$ and thus expected to yield a large overlap with classical solutions, is often a low-entanglement state, efficiently represented by a MPS of low bond dimension χ. This fact is certainly true for low-entangled many-body ground state preparation; nevertheless, it may also be verified in the context of classical optimization problems, whenever the number of classical solutions (spanning the ground-state eigenspace) is small enough, or when the exact QA converges to a cluster of solutions. These conditions are met for the models we examine, as detailed in Sec. 3.2.3 and B.3. Hence, in this case, both the initial and the final states of QA belong to the manifold $\mathcal{M}_\chi$, although the intermediate states may generally lay out of the manifold.

As stated above, in the regime of small δt, results based on ED and on MPS simulations essentially coincide; on the contrary, in the regime of large δt, dQA faces the blowing up of the Trotter errors, leading to an ED dynamics governed by an effective Hamiltonian that substantially differs from the original one. The resulting final state may thus deviate from the classical solutions of the target Hamiltonian. In addition, it could encode unwanted entanglement due to the spurious terms generated by the Trotter splitting. Our MPS approach relies instead on multiple projections of the time-evolved state on $\mathcal{M}_\chi$ for each time step δt: this may explain, for some class of optimization problems, the enhanced effectiveness of our MPS-based algorithm, since its final state may be closer to the optimal annealed state $|\psi_\tau\rangle$. These findings provide a novel promising application of TN techniques, which might be adapted to implement efficient classical simulations for other quantum optimization algorithms.

Finally, we show that a gate-decomposition of the final annealed MPS, yields efficient quantum circuits that effectively solve hard classical optimization problems, with a number of basis gates that grows only linearly in the system size N and quadratically with the MPS bond dimension χ. This result not only yields a numerical proof of principle on the effectiveness of quantum circuits in these regimes, but may also serve as a guideline to develop new classes of parameterized quantum circuits.

3.1 MODELS AND METHODS

3.1.1 Models

The classical optimization problems we analyze can be formulated as a ground-state search for a classical Hamiltonian $H(\boldsymbol{s})$, which can be mapped in a quantum setting as outlined in Eq. (3.0.1). We will focus on a broad class of classical Hamiltonians (or cost functions) that can be cast in the following form:

$$H(\boldsymbol{s}) = \sum_{\mu=1}^{N_\xi} h\left(\boldsymbol{\xi}^\mu \cdot \boldsymbol{s}\right), \tag{3.1.1}$$

where the $\xi_i^\mu \in \{-1, +1\}$ ($\mu = 1, 2 \ldots N_\xi$, $i = 1, 2 \ldots N$) are the components of N_ξ possibly random spin configurations, usually called patterns. h is any sufficiently regular function. As anticipated, our methods, detailed in the following Section, are quite general: they apply to any Hamiltonian that can be rewritten in this form.

As a preliminary benchmark for our strategy, we validate the results of our MPS-based technique against ED results for simple integrable p-spin models [132–134]:

$$H^{p-spin}(\boldsymbol{s}) = -N\left(\frac{1}{N}\sum_i s_i\right)^p, \tag{3.1.2}$$

which, for $p = 2$, is also known as the Lipkin-Meshkov-Glick (LMG) model (or infinite-range Ising model). Let us notice that the p-spin Hamiltonian can be rewritten as in Eq. 3.1.1 with a single pattern $\boldsymbol{\xi}^0 = (+1, +1, ..., +1)$ and $h(x) = -N^{1-p}x^p$. These benchmark models have trivial ground states: for even p, the classical ground-states are the two ferromagnetic states $\boldsymbol{s} = \pm\boldsymbol{\xi}^0$ (with energy $E_{gs} = -N$) whereas for odd p only $\boldsymbol{s} = +\boldsymbol{\xi}^0$ is a ground state (with energy $E_{gs} = -N$). Thanks to the integrability of p-spin models, we are able to verify the agreement of MPS and ED results up to large system sizes.

We then focus on two prototypical optimization problems coming from the realm of machine learning and artificial Neural Networks (NNs) [135, 136]. First, we consider the Hopfield model [135–142], a simple recurrent neural network studied in unsupervised learning

$$\begin{aligned} H^{\text{Hopfield}}(\boldsymbol{s}) = -\sum_{i,j} J_{ij} s_i s_j = -\sum_{\mu=1}^{N_\xi} \left(\frac{\boldsymbol{\xi}^\mu \cdot \boldsymbol{s}}{\sqrt{N}}\right)^2 \\ J_{ij} = \frac{1}{N}\sum_{\mu=1}^{N_\xi} \xi_i^\mu \xi_j^\mu \,. \end{aligned} \tag{3.1.3}$$

In this context, the patterns $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^{N_\xi}$ are i.i.d. random variables and the goal is to memorize them in the classical ground states. The Hopfield Hamiltonian entails infinite-range two-body interactions between any spin pair. Concerning the relation between the number of patterns N_ξ and the number of variables N, it is customary to set $N_\xi = \alpha N$ with $\alpha = \mathcal{O}(1)$. Different regimes/phases in the thermodynamic limit ($N \to \infty$) are distinguished by

different values of α. In particular, for the Hopfield model at zero temperature, $\alpha_c \simeq 0.138$ represents a critical value separating a retrieval phase in which the model works as a memory device ($\alpha < \alpha_c$), from a non-retrieval/spin-glass phase ($\alpha > \alpha_c$) [141].

Secondly, we examine the binary perceptron, the prototypical example of a single-layer binary classifier, which is a fundamental building block of NNs routinely used in supervised learning [135, 136, 143]. In this case, the spin variables $\mathbf{s}$ are identified with binary synaptic weights, classifying correctly a given pattern $\boldsymbol{\xi}^\mu$ into a prescribed binary label $\tau^\mu = \pm 1$ if $\mathrm{sgn}(\mathbf{s} \cdot \boldsymbol{\xi}^\mu) = \tau^\mu$ (see the sketch in Fig. 3.2). During the training phase, a given labeled data-set $\{\boldsymbol{\xi}^\mu, \tau^\mu\}_{\mu=1}^{N_\xi}$ is provided, and the objective consists in finding weight configurations $\mathbf{s}$ that classify correctly the whole training set. This is naturally formulated as a minimization problem of a suitable cost function, which assigns a positive energy cost for every pattern incorrectly classified, with the exact solutions to the classification problem being characterized as zero-energy configurations. A common choice for the such cost function is given by

$$H^{\text{perceptron}}(\mathbf{s}) = \sum_{\mu=1}^{N_\xi} \theta\left(-\boldsymbol{\xi}^\mu \cdot \mathbf{s}\right) \left(\frac{-\boldsymbol{\xi}^\mu \cdot \mathbf{s}}{\sqrt{N}}\right) , \tag{3.1.4}$$

where $\theta(x)$ is the Heaviside step function [1]. Let us observe that the Hamiltonian in Eq. 3.1.4 is again in the general form given by Eq. 3.1.1. Despite encouraging numerical and analytical evidence on the effectiveness of quantum optimization for this model [144, 145] and other closely related models [146], the perceptron Hamiltonian implies all possible interactions among spins, up to N-body terms, hence it is not efficiently implementable on a quantum device. Also for the perceptron model, it is customary to study the $N_\xi = \alpha N$ regime, with $\alpha = \mathcal{O}(1)$, since the critical capacity in the thermodynamic limit is $\alpha_c \simeq 0.83$ [147], separating a SAT region for $\alpha < \alpha_c$, admitting zero-energy solutions, from an UNSAT region $\alpha > \alpha_c$.

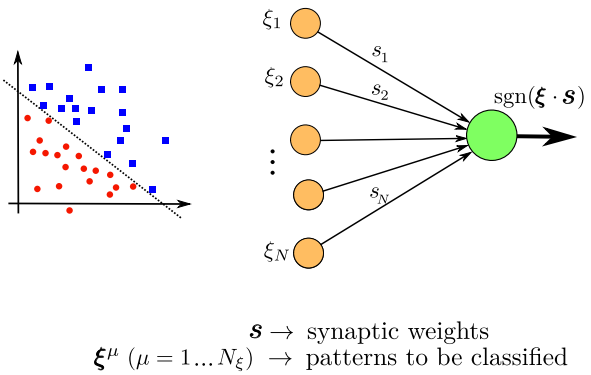

Figure 3.2: Schematic representation of a binary classifier consisting of a single perceptron with synaptic weights $\mathbf{s}$.

1 In the previous expression, the labels τ^μ are all set to 1, as it can be done without loss of generality for learning random patterns, i.e. if patterns and labels are both drawn from an unbiased Bernoulli distribution.

3.1.2 Digitized Quantum Annealing (dQA)

A standard procedure to implement Quantum Annealing on digital quantum simulators relies on a discretization of the continuous QA time evolution in $P \gg 1$ time steps of length $\delta t = \tau/P$, followed by a Trotter split-up of the two non-commuting terms. Albeit this scheme can be easily generalized to higher orders, here we stick to the lowest-order contributions

$$e^{-i\hat{H}(s_p)\,\delta t} \simeq e^{-i(1-s_p)\hat{H}_x\,\delta t}\, e^{-is_p\hat{H}_z\,\delta t} + \mathcal{O}\big(\delta t\big)^2\,, \tag{3.1.5}$$

where $p = 1, 2, \ldots, P$ and $s_p = t_p/\tau = p/P$. Introducing the shorthands $\beta_p = (1 - s_p)\delta t$, and $\gamma_p = s_p \delta t$, we can rewrite the previous expression more concisely as

$$e^{-i\hat{H}(s_p)\,\delta t} \simeq \hat{U}_x(\beta_p)\,\hat{U}_z(\gamma_p) + \mathcal{O}\big(\delta t\big)^2$$

with

$$\hat{U}_x(\beta_p) = e^{-i\beta_p\hat{H}_x}, \quad \hat{U}_z(\gamma_p) = e^{-i\gamma_p\hat{H}_z}\,, \tag{3.1.6}$$

thus recovering Eq. (3.0.3). The dQA framework can reproduce accurately the real QA dynamics for any value of the total annealing time τ, which is exactly recovered by simultaneously scaling $P \to \infty$ and $\delta t \to 0$, setting their product equal to τ. In practice, for a fixed value of P, the optimal value of the Trotter step δt depends on a trade-off between the Trotter errors and the annealing time τ [118]. Indeed, for small δt (small τ) the time evolution is not adiabatic, whereas for large δt (large τ) the Trotter split-up is expected to become a rough approximation and to introduce spurious quantum correlations. As discussed in Sec. 3.2.3, the discrete-time evolution obtained without any Trotter split-up is often unexpectedly accurate even for values of $\delta t \sim \mathcal{O}(1)$; however, the splitting in Eq. (3.1.5) is a necessary step to perform gate decomposition of the annealing dynamics on a quantum device, as well as to perform efficient classical simulations.

3.1.3 dQA with MPS

In this Section, we introduce our novel TN framework, which allows to efficiently simulate dQA for any classical Hamiltonian in the form of Eq. (3.1.1). First, let us notice that the initial state $|\psi_0\rangle$ can be trivially represented as an MPS of bond dimension $\chi = 1$, since it is a product state. Next, our goal is to rewrite the two unitaries $\hat{U}_z(\gamma_p)$ and $\hat{U}_x(\beta_p)$ as Matrix Product Operators (MPO, see Sec. 1.5). $\hat{U}_x$ admits an elementary decomposition into an MPO of bond dimension $\chi = 1$, since it is the exponential of an one-body Hamiltonian:

$$\langle \mathbf{s}'|\hat{U}_x(\beta_p)|\mathbf{s}\rangle = \prod_{i=1}^{N} \langle s_i'|e^{i\beta_p\hat{\sigma}_i^x}|s_i\rangle = \prod_{i=1}^{N} W^{[i]}(s_i', s_i)$$
$$W^{[i]}(s_i', s_i) = \delta_{s_i',s_i}\,\cos\beta_p + i\,(1 - \delta_{s_i',s_i})\,\sin\beta_p\,,$$

$|\mathbf{s}\rangle, |\mathbf{s}'\rangle$ being generic states of the computational basis. Since here $\chi = 1$, the tensors reduce to simple scalars $W^{[i]}(s_i', s_i)$. The MPO decomposition for $\hat{U}_z$

is much more challenging. First, one can factorize $\hat{U}_z$ into terms depending on a single pattern (see Eq. (3.1.1)), resulting in the following matrix element:

$$\langle \mathbf{s}'|\hat{U}_z(\gamma_p)|\mathbf{s}\rangle = \prod_{\mu=1}^{N_\xi} \langle \mathbf{s}'|\hat{U}_z^\mu(\gamma_p)|\mathbf{s}\rangle = \delta_{\mathbf{s}',\mathbf{s}} \prod_{\mu=1}^{N_\xi} e^{-i\gamma_p h(\xi^\mu\cdot\mathbf{s})} . \tag{3.1.7}$$

Let us now exploit the specific form of the Hamiltonian in Eq. (3.1.1). We notice that, by definition, any such Hamiltonian depends on the spin configuration $\mathbf{s}$ only via the following variables:

$$m^\mu(\mathbf{s}) = \boldsymbol{\xi}^\mu \cdot \mathbf{s} = \text{ overlap between } \mathbf{s} \text{ and } \boldsymbol{\xi}^\mu ,$$

or, equivalently:

$$x^\mu(\mathbf{s}) = \frac{N - \boldsymbol{\xi}^\mu \cdot \mathbf{s}}{2} = \text{ number of bits of } \mathbf{s} \text{ that are different from } \boldsymbol{\xi}^\mu. \tag{3.1.8}$$

The latter expression is the well-known Hamming distance between $\mathbf{s}$ and $\boldsymbol{\xi}^\mu$ and, accordingly, we observe that

$$m^\mu \in \{-N, -N+2, ..., N-2, N\} , \qquad x^\mu \in \{0, 1, ..., N\} .$$

This is a key point, as it represents the only hypothesis which our construction relies on. In fact, since x (for any pattern μ) is an integer variable taking values in the discrete set $\{0, 1, ..., N\}$, then any function $O(x)$ can be rewritten by means of the Discrete Fourier Transform (DFT) as follows

$$O(x) = \frac{1}{\sqrt{N+1}} \sum_{k=0}^{N} \tilde{O}_k \, e^{i\frac{2\pi}{N+1}kx} , \tag{3.1.9}$$

where the Fourier coefficients are computed as

$$\tilde{O}_k = \frac{1}{\sqrt{N+1}} \sum_{x=0}^{N} e^{-i\frac{2\pi}{N+1}kx} \, O(x) . \tag{3.1.10}$$

By setting $h(\xi^\mu \cdot \mathbf{s}) = f(x^\mu(\mathbf{s}))$, where $x^\mu(\mathbf{s})$ is defined by Eq. 3.1.8, and using the DFT expansion reported in Eqs. (3.1.9) and (3.1.10), we can further manipulate Eq. 3.1.7 as follows:

$$\begin{aligned} \langle \mathbf{s}'|\hat{U}_z^\mu(\gamma_p)|\mathbf{s}\rangle &= \delta_{\mathbf{s}',\mathbf{s}} \frac{1}{\sqrt{N+1}} \sum_{k=0}^{N} \tilde{U}_{k,p} \, e^{i\frac{2\pi}{N+1}kx^\mu(\mathbf{s})} \\ \tilde{U}_{k,p} &= \frac{1}{\sqrt{N+1}} \sum_{x=0}^{N} e^{-i\frac{2\pi}{N+1}kx} \, e^{-i\gamma_p f(x)} , \end{aligned} \tag{3.1.11}$$

where the Fourier components $\tilde{U}_{k,p}$ depend implicitly on the angle γ_p, so they can be regarded as a matrix of dimension $(N+1) \times P$. Remarkably, this Fourier decomposition allows us to find an *efficient* representation of $\hat{U}_z(\gamma_p)$ as an MPO. This is accomplished by using Eq. 3.1.8, which can be reformulated more explicitly as

$$x^\mu(\mathbf{s}) = \sum_{i=1}^{N} \left(\frac{1 - \xi_i^\mu s_i}{2} \right) .$$

Indeed, by identifying the wave-numbers $k = 0, 1, ...N$ as *auxiliary* indices in the MPO formalism, we can rewrite

$$\langle \mathbf{s}' | \hat{U}_z^{\mu}(\gamma_p) | \mathbf{s} \rangle = \prod_{i=1}^{N} W^{[i]}_{k_{i-1},k_i}(s'_i, s_i) \ , \tag{3.1.12}$$

where we defined the following tensors, diagonal by construction in the *physical* spin indices:

$$\begin{cases} W^{[i]}_{1,k}(s'_i, s_i) &= \delta_{\sigma'_i,s_i} \left(\frac{\tilde{u}_{k,p}}{\sqrt{N+1}} \right)^{\frac{1}{N}} e^{i \frac{\pi}{N+1} k (1 - \xi_i^{\mu} s_i)} & i = 1 \\ W^{[i]}_{k,k'}(s'_i, s_i) &= \delta_{\sigma'_i,s_i} \left(\frac{\tilde{u}_{k,p}}{\sqrt{N+1}} \right)^{\frac{1}{N}} e^{i \frac{\pi}{N+1} k (1 - \xi_i^{\mu} s_i)} \delta_{k,k'} & i = 2, \cdots, N-1 \\ W^{[i]}_{k,1}(s'_i, s_i) &= \delta_{\sigma'_i,s_i} \left(\frac{\tilde{u}_{k,p}}{\sqrt{N+1}} \right)^{\frac{1}{N}} e^{i \frac{\pi}{N+1} k (1 - \xi_i^{\mu} s_i)} & i = N \end{cases} \tag{3.1.13}$$

In Eq. (3.1.12) we set $k_0 = k_N = 1$, whereas tensors are implicitly summed over repeated auxiliary indices, each of them spanning $N + 1$ values. Therefore, we found that each unitary time evolution operator $\hat{U}_z^{\mu}(\gamma_p)$ associated to a given pattern μ can be written *efficiently* as an MPO of bond dimension $\chi = N + 1$. We remark that the four-indices tensors $W^{[i]}_{k,k'}(s'_i, s_i)$ are diagonal also in the auxiliary indices, effectively depending only on a single auxiliary index k as well as on a single physical one s_i. However, note that they also depend on the angle γ_p and on the pattern index μ, both of which are not explicitly indicated. In light of these results, the whole dQA time evolution can be represented exactly as the 2D Tensor Network in Fig. 3.3, corresponding to the application of a series of MPOs to the initial trivial MPS $|\psi_0\rangle$. This result is the starting point for our MPS-based algorithm for the classical simulation of dQA, which is summarized in the pseudo-code 1. Importantly, when contracting this 2D TN the MPS bond dimension would increase exponentially with the number of Trotter slices P. This calls for an effective compression procedure, which is detailed in the next section: indeed, a crucial input parameter of our algorithm is the fixed maximum bond dimension of the MPS.

Let us notice that the MPO representation shown in Eq. 3.1.13 can also be exploited to write the Hamiltonian itself as an MPO of bond dimension $\chi = N + 1$. This fact is remarkable, since it allows the *exact evaluation* of the classical cost function (i.e. the expectation value of $\hat{H}_z$) over *any* MPS: in particular, when applying the Algorithm 1, we can keep track of the exact value of the cost function along the time evolution. Moreover, the representation of $\hat{H}_z$ as MPO makes it possible to apply the DMRG algorithm, in order to find an approximate ground state of the system. However, as shown in Sec. B.1 of Appendix B, this method has some issues, since the quality of the final result is strongly dependent on the initial guess provided to the DMRG optimization. Moreover, even if DMRG reaches convergence, the optimized state turns out to have a sensible overlap with only one of the degenerate ground states configurations (classical solutions). The QA protocol, instead, always reaches delocalized quantum states, having large

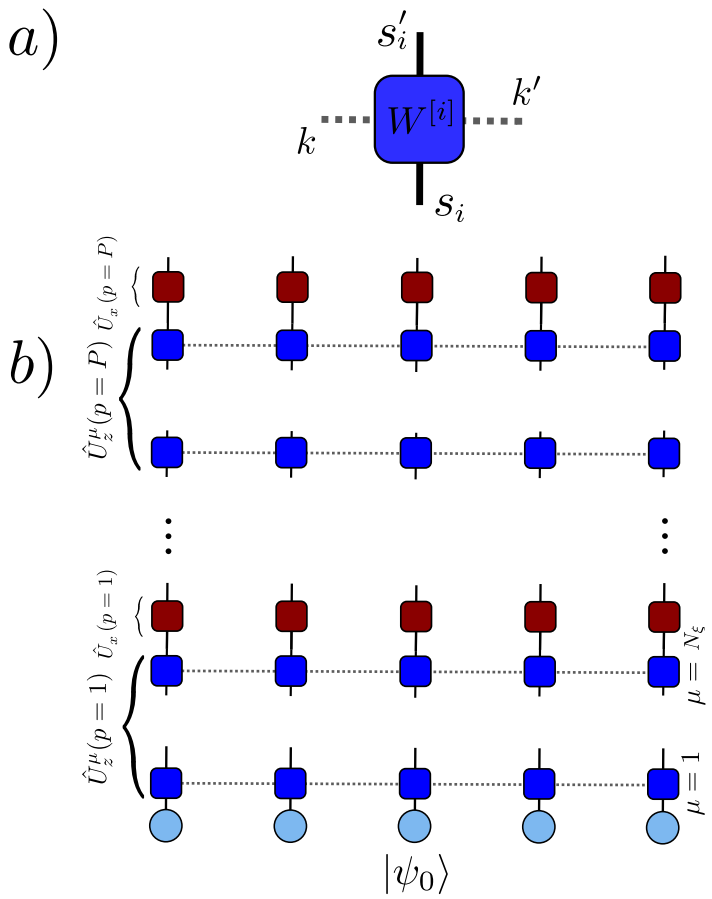

Figure 3.3: a) The four-indices MPO local tensor $W^{[i]}_{kk'}(s_i, s_i')$ of the $\hat{U}_z^{\mu}(\gamma_p)$ time evolution operator. Dotted grey lines represent auxiliary indices of dimension $N+1$ $(k, k' = 0, 1, ..., N)$. b) The trotterized time evolution of dQA as a 2D Tensor Network. Blue (red) squared shapes represent the MPO tensors for the decomposition of $\hat{U}_z^{\mu}(\gamma_p)$ $(\hat{U}_x(\beta_p))$ time evolution operators.

overlaps with many different solutions (see Sec. B.1 in Appendix B for details). This delocalization over clusters of classical solutions is an interesting confirmation of previous results[144, 145] in a new setting, and it may represent a winning feature for some classical optimization problems. In addition, whereas methods as DMRG are purely related to the classical simulation of many-body quantum systems, our MPS-scheme simulates dQA, providing a benchmark for Quantum Annealing experiments on real quantum devices, beyond the small system sizes reachable by means of ED methods.

Finally, it is worthwhile to mention that our TN representation can be easily adapted to study imaginary time evolution and the classical equilibrium properties of the corresponding classical models, i.e. without the transverse term $\hat{H}_x$.[2]

2 Indeed, in this case, the classical equilibrium partition function will assume the form of a 2D tensor network as in Fig. 3.3, without the single-qubit layers.

Algorithm 1 dQA with MPS

Input: the dQA parameters τ, P, $\delta t = \tau/P$, the MPS maximum bond dimension χ

1: Compute the Fourier matrix $\tilde{U}_{k,p}$ (Eq. 3.1.11) of dimension $(N+1) \times P$
2: Define the initial state $|\psi\rangle = |\psi_0\rangle$ as an MPS of bond dimension $\chi = 1$
3: **for** ($p = 1$, $p = P$, $p++$) **do**
4: **for** ($\mu = 1$, $\mu = N_\xi$, $\mu++$) **do**
5: apply $\hat{U}_z^\mu(\gamma_p)$ to the current MPS $|\psi\rangle$ and compress to a lower bond dimension χ
6: update $|\psi\rangle$ setting it equal to the resulting MPS
7: **end for**
8: apply $\hat{U}_x(\beta_p)$ to the current MPS $|\psi\rangle$
9: **end for**

Output: the final optimized MPS $|\psi\rangle$

The MPS compression

A key step in our proposed algorithm is the iterative compression of the MPS wave function, reported in line 5 of the pseudo-code 1. As anticipated, this is necessary to avoid an exponential increase of the MPS bond dimension with the number of Trotter steps P. In practice, the compression procedure projects the dynamics on the manifold of MPS with fixed maximum bond dimension χ. This projection can be achieved with the method in Section 1.4.4. In our particular case, $|\psi\rangle$ is given by the application of a unitary time evolution operator to the MPS at the previous step, i.e. $|\psi\rangle = \hat{U}_z^\mu(\gamma_p)|\phi\rangle$. In general, when an MPO is applied to an MPS, the corresponding bond dimensions are multiplied, thus we would get a bond dimension $(N+1)\chi$ and the overall cost of one compression would be $\mathcal{O}(N_{sweeps} N^3 \chi^3)$.

However, the compression efficiency can be greatly improved by exploiting the particular structure of the MPO in Eq. 3.1.13. To this scope, let us first observe that each operator $\hat{U}_z^\mu(\gamma_p)$ can be equivalently recast into a sum of $(N+1)$ MPOs, each of bond dimension $\chi = 1$. This fact can be seen directly from Eq. 3.1.13, by noticing that the tensors W are diagonal in the auxiliary indices k, k'. More formally, if we define the tensors

$$W^{k;[i]}(s_i', s_i) = \delta_{s_i', s_i} \left(\frac{\tilde{U}_{k,p}}{\sqrt{N+1}} \right)^{1/N} e^{i \frac{\pi}{N+1} k (1 - \xi_i^\mu s_i)} \quad k = 0, 1 \dots N \quad i = 1, 2 \dots N \,,$$

we have

$$\sum_{k=0}^{N} \prod_{i=1}^{N} \left(W^{k,[i]}(s_i', s_i) \right) = \delta_{\mathbf{s}', \mathbf{s}} \sum_k \frac{\tilde{U}_{k,p}}{\sqrt{N+1}} e^{i \frac{2\pi}{N+1} k x^\mu(\mathbf{s})} = \langle \mathbf{s}' | \hat{U}_z^\mu(\gamma) | \mathbf{s} \rangle \,, \tag{3.1.14}$$

namely the operator $\hat{U}_z^\mu(\gamma)$ is a sum of $N+1$ MPOs with bond dimension $\chi = 1$. Therefore, the uncompressed MPS $|\psi\rangle = \hat{U}_z^\mu(\gamma_p)|\phi\rangle$ can be written

as sum of $N+1$ MPS $|\psi_k\rangle$, each with bond dimension χ, so that Eq. 1.4.9 becomes

$$\frac{\partial}{\partial(\tilde{A}^{[i]}_{\alpha_{i-1},\alpha_i})^*}\left(\langle\tilde{\psi}|\tilde{\psi}\rangle - \sum_{k=0}^{N}\langle\tilde{\psi}|\psi_k\rangle\right) = 0\,. \tag{3.1.15}$$

Now, the second term involves $N+1$ contractions of MPS with bond dimension χ. The computational cost is therefore reduced by a factor N, i.e. to $\mathcal{O}(N_{sweeps}N^2\chi^3)$.

Estimate of the computational cost

The computational bottleneck of Algorithm 1 is the iterative compression of MPS sketched in the previous section. This is repeated PN_ξ times, and each repetition is expected to have a computational cost $\mathcal{O}(N_{sweeps}N^2\chi^3)$, where χ is the maximum bond dimension. If we set $N_{sweeps} = \mathcal{O}(1)$, the overall algorithmic cost is therefore estimated to be $\mathcal{O}(PN_\xi N^2\chi^3)$ and, in the regime $N_\xi \sim N$, we expect $\mathcal{O}(PN^3\chi^3)$.

The most computationally expensive steps are $3-9$, whereas the initial Fourier transforms can be performed at cost $\mathcal{O}(PN\log N)$ by means of the Fast Fourier Transform. In particular, the actual computational bottleneck is the application of each $\hat{U}_z^\mu(\gamma_p)$ to the current MPS, and the subsequent compression to the prescribed bond dimension χ.

Concerning the cost function evaluation, by exploiting the same MPO structure of Eq. 3.1.13 (with $\tilde{h}_k$ in place of $\tilde{U}_{k,p}$), once again one can rewrite $\hat{H}_z = \sum_{\mu=1}^{N_\xi}\hat{h}_\mu$ as a sum of $N_\xi(N+1)$ MPOs of bond dimension 1, instead of a single MPO of bond dimension $N_\xi(N+1)$. This allows to reduce the cost of the tensor contractions involved in the evaluation of $\langle\psi|\hat{H}_z|\psi\rangle$, that is $\mathcal{O}(N_\xi N^2\chi^3)$, instead of $\mathcal{O}(N_\xi^3 N^4\chi^3)$. If the cost function is evaluated at each step of dQA the overall cost is $\mathcal{O}(PN_\xi N^2\chi^3)$, that is of the same order of the algorithm itself.

Estimate of the computational cost

The ultimate result of simulating dQA through TN is an optimized MPS, ideally exhibiting a significant overlap with the subspace of ground states in the target Hamiltonian $\hat{H}_z$. Although our simulation is conducted using exclusively classical resources, it is possible to translate the final MPS into a quantum circuit, enabling real-world implementations on near-term quantum devices. This fact may set the stage to further manipulate the state with quantum resources, for example with additional hybrid quantum-classical optimizations [115]. The mapping onto a quantum circuit can be achieved by exploiting the MPS nature of the final state. As we discussed in Sec. 1.8, any MPS having maximum bond dimension $\chi = 2^n$ can be obtained from the trivial state $|\mathbf{0}\rangle = |0...0\rangle$ by applying sequentially N unitary gates, each acting at most on $\log_2\chi + 1 = n+1$ qubits. Although this proof of concept is remarkable, in order to practically compile the optimized MPS into a quantum circuit we focus on another method, reported in [51] and discussed in Sec. 1.8. This employs an algorithm that iteratively optimizes two-qubits unitaries in a fixed circuit architecture, in order to maximize the fidelity with

the target MPS. The geometry of the circuit is given by a fixed number D of staircase layers of two-qubits gates.

3.2 RESULTS

In this Section, we analyze and discuss our numerical results on MPS simulation of dQA. Here, we focus on the benchmark p−spin model and on the binary perceptron; the same qualitative results hold for the Hopfield model, as reported in Sec. 3.2.4. In the MPS simulations of the following sections, we fix a relatively small value of the bond dimension $\chi = 10$: in Sec. B.2 of Appendix B, we perform an analysis of the convergence for increasing values of χ. With such bond dimension, we are able to study systems up to size $N \simeq 100$ (see Sec. 3.2.4), whereas ED is necessarily limited to $N \simeq 20$.

It is useful to introduce the energy density

$$\varepsilon(s) = \frac{\langle\psi(s)|\hat{H}_z|\psi(s)\rangle - E_{gs}}{N} , \tag{3.2.1}$$

where $|\psi(s)\rangle$ is the instantaneous wave function for a certain value of the annealing parameter $s \in [0,1]$ and E_{gs} is the ground-state energy of $\hat{H}_z$. The residual energy density at the end of the annealing schedule, defined as $\varepsilon(1)$, can be regarded as a figure of merit of dQA effectiveness. The instantaneous standard deviation of the time-dependent Hamiltonian $\hat{H}(s)$ is written as

$$\sigma_{\hat{H}}(s) = \frac{1}{N}\left(\langle\psi(s)|\hat{H}^2(s)|\psi(s)\rangle - \langle\psi(s)|\hat{H}(s)|\psi(s)\rangle^2 \right)^{1/2} , \tag{3.2.2}$$

with an analogous definition for $\sigma_{\hat{H}_z}(s)$. Notice that the residual standard deviation of $\hat{H}_z$, i.e. $\sigma_{\hat{H}_z}(1) = \sigma_{\hat{H}}(1)$, is another possible figure of merit for dQA, since it is expected to vanish if the exact ground-state of $\hat{H}_z$ is reached. Moreover, $\sigma_{\hat{H}}(s)$ is expected to be constantly 0 in the case of a perfectly adiabatic dynamics.

3.2.1 Benchmark (p-spin models)

As a preliminary check, we focus on the integrable p−spin model, in order to benchmark our MPS-based simulations against ED results, up to large system sizes. Let us remark that the Hamiltonian considered in Eq. 3.0.2 (with $\hat{H}_z$ given by Eq. 3.1.2) has a phase transition as a function of the annealing parameter s. In the thermodynamic limit $N \to \infty$, one encounters a second order phase transition for $p = 2$, and a first order phase transition for $p > 2$ [134]. In particular, by means of mean field calculations, one can show that the critical values are $s_c = 1/3$ for $p = 2$ and $s_c \simeq 0.435$ for $p = 3$ [134]. These phase transitions represent a challenge for Quantum Annealing, since a system is expected to stay in the instantaneous ground state of $\hat{H}(s)$ only if the total evolution time τ is (at least) inversely proportional to the square of the minimum energy gap of $\hat{H}(s)$ [132], as stated by the adiabatic theorem. Besides, it is also known that the energy gap decreases exponentially as a function of the system size N at a first-order quantum

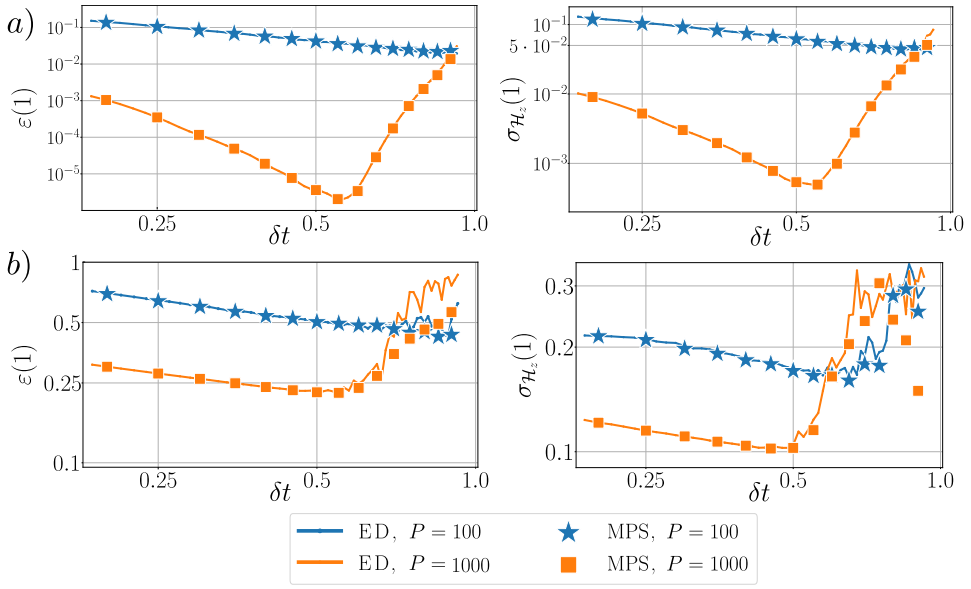

Figure 3.4: p−spin model. Residual energy density $\varepsilon(1)$ (left column) and residual standard deviation of the energy $\sigma_{\hat{H}_z}(1)$ (right column) as a function of the time step δt, for a) $p = 2$ and b) $p = 3$. We compare MPS ($\chi = 10$) and ED results. The system size is $N = 50$.

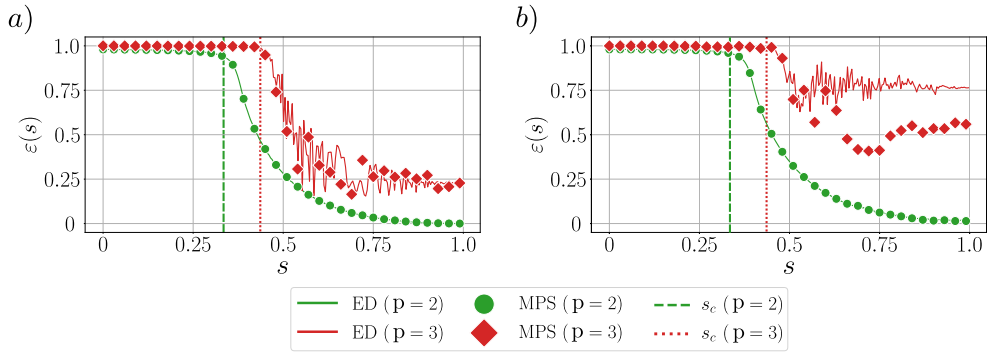

Figure 3.5: p−spin model. Energy density $\varepsilon(s)$ as a function of the annealing parameter s. Data refer to $P = 1000$ time steps and a) $\delta t = 0.5$ (close to the minimum) and b) $\delta t = 0.9$ (large Trotter errors regime). We consider both $p = 2$ (green) and $p = 3$ (red), comparing MPS ($\chi = 10$) with ED. The system size is $N = 50$.

phase transition, whereas the scaling is polynomial in N for a second-order transition [133]. Consequently, a QA implementation applied to finite-size systems is expected to perform worse for $p = 3$ than for $p = 2$. Let us notice that dQA can be easily simulated via ED up to large qubit numbers for the p−spin model. Indeed, the interpolating Hamiltonian $\hat{H}(s)$ in Eq. 3.0.2 commutes with the total spin operator

$$\hat{S}^2 = \frac{1}{4} \sum_{\alpha=x,y,z} \left(\sum_{i=1}^{N} \hat{\sigma}_i^{\alpha} \right)^2 ,$$

for any value of p. Since the initial state $|\psi_0\rangle$ of dQA belongs to the Hilbert space sector of maximum total magnetization, i.e.

$$\langle\psi_0|\hat{S}^2|\psi_0\rangle = \langle\rightarrow\cdots\rightarrow|\hat{S}^2|\rightarrow\cdots\rightarrow\rangle = \frac{N}{2}\left(\frac{N}{2}+1\right),$$

the time evolution of the system is constrained in this sector for any step $p = 1,\ldots,P$. This allows to write the time evolution operators $\hat{U}_z(\gamma_p)$ and $\hat{U}_x(\beta_p)$ as matrices of dimension $N+1\times N+1$, and to evaluate exactly the dynamics.

In Fig. 3.4, we benchmark our MPS results, obtained by setting a maximum bond dimension $\chi = 10$, with ED results. This is done for $p = 2,3$ models, with $N = 50$ qubits. We plot the residual energy density $\varepsilon(1)$ and the residual standard deviation $\sigma_{\hat{H}_z}(1)$ vs the time step $\delta t = \tau/P$, with a total number of steps fixed to $P = 100, 1000$. In all simulations we used a first-order Trotter approximation, as sketched in Sec. 3.1. As expected, ED data confirm that the final annealed state $|\psi(s=1)\rangle$ becomes a better approximation of the ground state by increasing the time step δt (at fixed P), until a minimum is reached, and then the protocol starts to become inaccurate due to large Trotter errors. For $p = 2$, the MPS results are in perfect agreement with ED for all values of δt, proving that our tensor network techniques can reproduce the exact (digitized) dynamics, despite fixing a finite bond dimension. For $p = 3$, the agreement is equally good, except for the regime of large δt, which is dominated by large Trotter errors. In this case, MPS results deviate from ED, surprisingly assuming lower values of the final energy density $\varepsilon(1)$.

To summarize, in the regime of sufficiently small δt and large values of P, i.e. where dQA approximates accurately the actual QA dynamics, our MPS techniques numerically coincide with ED simulation of dQA, thus proving an effective tool to simulate the annealing of large-size systems (at least for these benchmark models). In contrast, in the large δt regime, dominated by large Trotter errors, the MPS dynamics can sometimes depart from ED results. However, as sketched in Fig. 3.1 and in its discussion, the final MPS may turn out to be a better approximation (in terms of residual energy) of the ground state if compared to the final state obtained by exact dynamics. We will observe the same phenomenon in the next sections, for the other models considered. In Fig. 3.5, we also compare ED and MPS results along the annealing dynamics, for $P = 1000$. In particular, we consider two representatives values of $\delta t = 0.5$ (panel a) and $\delta t = 0.9$ (panel b), roughly corresponding to the minima of the residual energy landscape in Fig. 3.4 and to the regime dominated by Trotter errors, respectively. We plot the energy density $\varepsilon(s)$, as a function of $s \in [0,1]$. For $p = 2$, we observe a perfect match of MPS and ED values of $\varepsilon(s)$, during the whole time evolution. For $p = 3$, however, the agreement is good only for $\delta t = 0.5$ (corresponding to the minimum, i.e. low Trotter errors), whereas for $\delta t = 0.9$ our MPS dynamics does not reproduce the exact digitized dynamics (at least for $s > s_c$, where s_c corresponds to the phase transition point). Consistently with the discussion above, the MPS simulation yields substantially lower values of the energy density in the last part of the annealing (in particular for s approaching to 1).

3.2.2 Binary perceptron

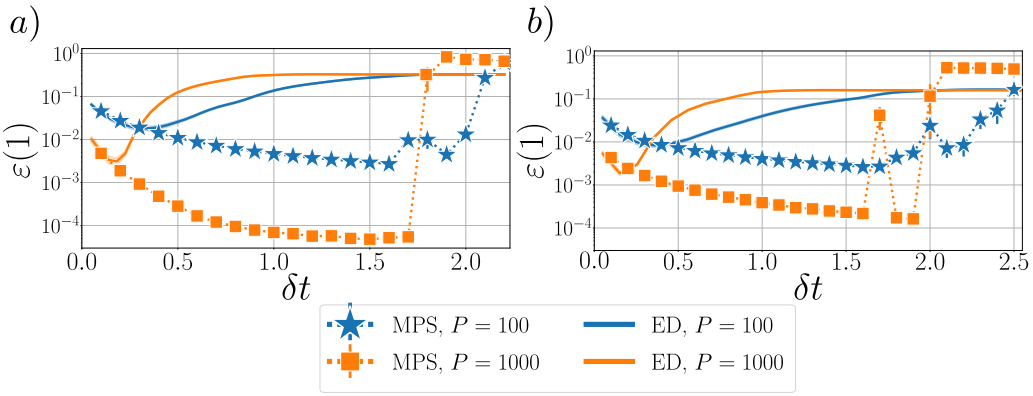

Figure 3.6: Binary perceptron. Residual energy density $\varepsilon(1)$ as a function of the time step length δt for $N = 21$ and $P = 100, 1000$. Data are averaged over five different training sets, for both values of N_ξ (or $\alpha = N_\xi/N$) considered: a) $N_\xi = 17$ ($\alpha \simeq 0.81$) close to the SAT/UNSAT transition, and b) $N_\xi = 8$ ($\alpha \simeq 0.38$) in the SAT phase. MPS results with $\chi = 10$ (full symbols) are compared with ED results (solid lines), showing good agreement for $\delta t \ll 1$, and remarkably better results for $\delta t = \mathcal{O}(1)$ (see discussion in the main text). Error bars for MPS data are given by the standard error of the means (seldom visible). Solid lines with lower opacity represent average $\pm$ the standard error of the mean for ED data.

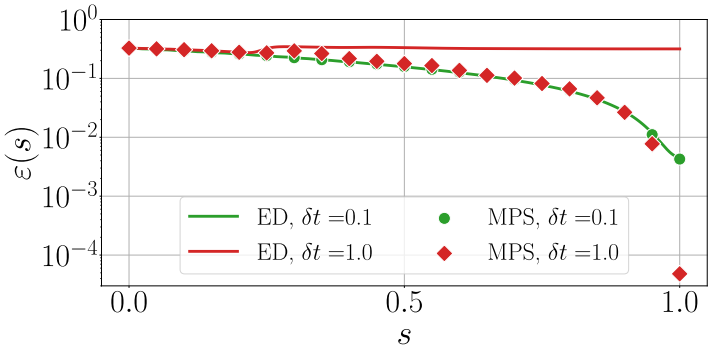

Figure 3.7: Binary perceptron. Energy density $\varepsilon(s)$ as a function of the annealing parameter s ($N = 21$, $N_\xi = 17$, $P = 1000$). We fixed $\delta t = 0.1$ (green) or $\delta t = 1.0$ (red).

We now focus on the binary perceptron model, as defined by Eq. 3.1.4. First, we consider a system of size $N = 21$, so that the classical solutions/-ground states can easily be found by enumeration, and MPS findings can be compared to ED results.

We set $N_\xi = 17$, corresponding to $\alpha = N_\xi/N \simeq 0.81$, close to the critical value $\alpha_c \simeq 0.83$, and we run our simulations for different training sets, labeled $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^{N_\xi}$. In particular, we considered the same training sets used in Ref. [144], also analyzed in Ref. [145]. These were originally selected to yield highly non-convex minimization tasks, proving particularly challenging for

standard Simulated Annealing techniques (i.e. Simulated Annealing suffers from an exponential slow-down, due to the trapping in meta-stable states). In the following, for conciseness, we show data averaged over five training sets in that list. However, the same results and considerations hold true for each single case in exam. In Fig. 3.6 (panel a)), we show the residual energy $\varepsilon(1)$ obtained by the MPS simulations and by ED, spanning different values of time step length $\delta t = \tau/P$. The total number of steps is fixed to $P = 100$ or $P = 1000$. Here and in the following, unless otherwise stated, MPS simulations are run by setting a maximum bond dimension $\chi = 10$. In both cases, we observe that MPS results replicate ED data for small values of δt ($\delta t \ll 1$). On the other hand, for higher values of δt (i.e. $\delta t \sim \mathcal{O}(1)$), ED data show an expected increase of the final energy density due to Trotter errors. On the contrary, MPS simulations surprisingly yield a further considerable decrease in the residual energy. This phenomenon supports the heuristic sketch in Fig. 3.1, replicating more distinctly what already observed in Fig. 3.4 b) for the $p = 3$ spin model at large values of δt. Strikingly, in this noise-dominated regime, our MPS framework still successfully performs the quantum optimization, and it yields non-trivial final quantum states that prove more efficient than those obtained by exact dQA. A detailed discussion of this phenomenon is given in Sec. 3.2.3. For completeness, we investigated whether these findings hold true also in the SAT phase, further away from the critical value α_c. The answer is positive, as shown in Fig. 3.6 (panel b)) for $N = 21$, $N_\xi = 8$ (i.e. $\alpha \simeq 0.38$). Here, data are also averaged over five different training sets, which are randomly generated.

The remarkable difference between the two regimes of $\delta t \ll 1$ and $\delta t = \mathcal{O}(1)$ is better elucidated in Fig. 3.7, where we compare ED and MPS results for the two representative cases of $\delta t = 0.1$ and $\delta t = 1.0$, by plotting the energy density $\varepsilon(s)$ during the annealing. In the first case, we observe an excellent agreement between the two methods, whereas in the second case the instantaneous MPS deviates from the ED state, finally reaching considerably lower energy values. In particular, the energy density of the MPS drops by almost two orders of magnitude in the very final annealing steps (i.e. for $s \to 1$).

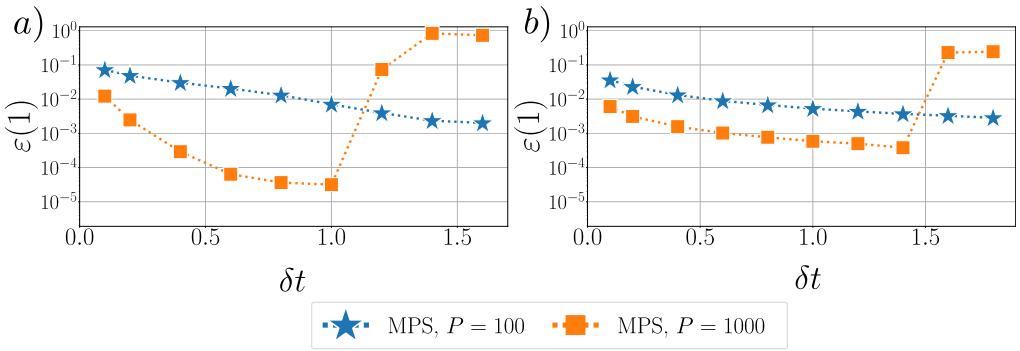

Figure 3.8: Binary perceptron. Residual energy density $\varepsilon(1)$ as a function of time step length δt, for a single randomly-generated training set ($N = 50$ and $P = 100, 1000$). Two values of N_ξ (α) are considered: a) $N_\xi = 40$ ($\alpha = 0.8$), b) $N_\xi = 20$ ($\alpha = 0.4$).

As already discussed, our MPS framework allows for large-scale simulations, beyond the reach of standard ED techniques. This may prove a useful benchmark for the actual implementation of quantum optimization schemes on near-term quantum devices. Here, to show the effectiveness of our methods, we address the example of a perceptron with $N = 50$ qubits, well beyond the reach of ED techniques, by studying a single randomly-generated training set. In Fig. 3.8 we plot the residual energy $\varepsilon(1)$ vs the time step length δt, as obtained by MPS simulations. The total number of steps is fixed again to $P = 100$ or $P = 1000$ and two similar values of α are considered: $\alpha = 0.8$ ($N_\xi = 40$, panel a) and $\alpha = 0.4$ ($N_\xi = 20$, panel b). Our MPS-based simulation yields a final state with low values of residual energy density, corresponding to a successful optimization. Furthermore, these results confirm the previously outlined scenario: the residual energy density keeps decreasing up to $\delta t \simeq \mathcal{O}(1)$. As a matter of fact, by increasing further the Trotter step beyond a model-dependent threshold, the algorithm enters into an unstable regime where the method fails to provide an optimal solution, and the residual energy $\varepsilon(1)$ suddenly jumps to very large values.

In the discussion above, we assumed the residual energy density $\varepsilon(1)$ to be a good figure of merit of dQA effectiveness. In the following, we verify this explicitly for the case of $N = 21$ and $N_\xi = 17$, by evaluating the overlap of the final annealed MPS $|\psi(s = 1)\rangle$ with the exact ground states $\{\mathbf{s}^*_\alpha\}_{\alpha=1}^{N_{sol}}$. The degeneracy of the perceptron model ranges in $N_{sol} \in [20, 160]$, depending on the particular training set in exam: here, for simplicity, we refer to the first training set with $N_{sol} = 80$. In Fig. 3.9 a) we plot (one minus) the total success probability, defined as $\sum_{\alpha=1}^{N_{sol}} p(\mathbf{s}^*_\alpha) = \sum_{\alpha=1}^{N_{sol}} |\langle \mathbf{s}^*_\alpha | \psi(s = 1)\rangle|^2$, vs δt for $P = 100, 1000$, confirming the same trend as previously shown in Fig. 3.6 (a). Additionally, in Fig. 3.9 b), we plot an histogram of probabilities $p(\mathbf{s}^*_\alpha)$ for a fixed value of δt (i.e. $\delta t = 1.4$, approximately corresponding to the best MPS performance) [3]. Noticeably, the final output of dQA has a non-vanishing overlap with a number of the classical solutions of the optimization problem, i.e. the final wave function is delocalized. As already mentioned, this fact represents a remarkable difference with the DMRG algorithm, which always converges to completely localized wave functions, overlapping with a single classical solution (see Sec. B.1 for details on DMRG results). Finally, let us mention that the final state has a non-vanishing overlap with the same set of classical solutions for both values of P, and the same is observed by comparing MPS results with ED (data not shown).

3.2.3 The MPS-projection mechanism to mitigate Trotter errors

In this section, we provide some theoretical insight to explain the numerical results presented above. Let us first summarize a few preliminary concepts. The set of MPS with any fixed (finite) value of bond dimension χ constitutes a smooth manifold $\mathcal{M}_\chi$ inside the many-body Hilbert space [40], as represented in Fig. 3.1. The dimension of $\mathcal{M}_\chi$ is $\simeq 2\chi^2 N$, since this is the number

3 We order the ground states $\mathbf{s}^*_\alpha$ such that $p(\mathbf{s}^*_\alpha)$ is in descending order for $P = 1000$, and we only select the 40 most probable solutions.

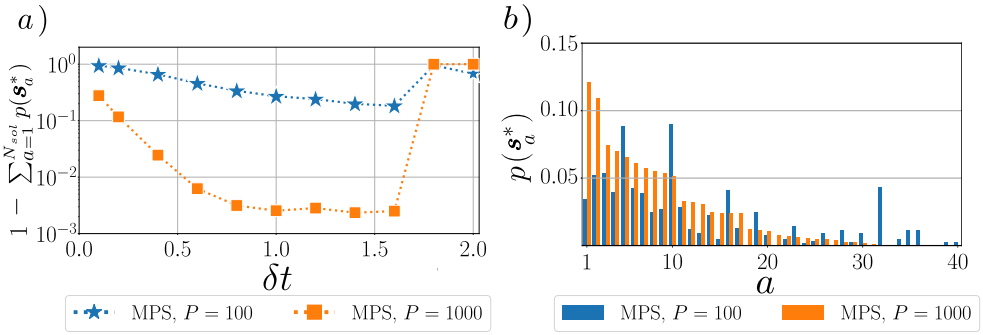

Figure 3.9: Binary perceptron. a) One minus the total success probability $\sum_{a=1}^{N_{sol}} p(s_a^*)$. The agreement with the corresponding plot for the residual energy density confirms that the latter is a good proxy for dQA effectiveness. b) Probabilities $p(s_a^*)$ of measuring the final state $|\psi(s=1)\rangle$ in a given solution s_a^*, with $\delta t = 1.4$. In both figures, data refer to a single instance with $N = 21$, $N_\xi = 17$. The same qualitative results hold for all the other training sets. Solutions are arbitrarily sorted such that $p(s_a^*)$ is in descending order for $P = 1000$.

of independent parameters in the MPS tensors, whereas the dimension of the many-body Hilbert space is exponentially greater, 2^N. Also, $\mathcal{M}_\chi$ can be equivalently characterized as the manifold of quantum states such that the entanglement entropy between any system bipartition is *upper bounded* by $\log\chi$ [27].

Let us now go back to our MPS-based Algorithm, introduced in Sec. 3.1 and summarized in the pseudo code 1, by examining it from a geometrical point of view (as depicted in Fig. 3.1). We remark that the initial state of dQA, namely the fully polarized state $|\psi_0\rangle = |\rightarrow\rangle^{\otimes N}$, is an MPS with bond dimension 1, thus it is a point belonging to the manifold $\mathcal{M}_\chi$ (for any $\chi \geqslant 1$). For every time step $p = 1 \cdots P$, our algorithm repeatedly projects the exact dQA dynamics on $\mathcal{M}_\chi$, after the application of each MPO $\hat{U}_z^\mu(\gamma_p)$ (corresponding to the patterns $\mu = 1, \ldots, N_\xi$). The projection is provided by the compression algorithm described in Sec. 3.1.3. This procedure results into an effective trajectory belonging to the manifold $\mathcal{M}_\chi$, sketched in Fig. 3.1 by a gray dashed line, and it affects the entanglement content of the quantum state, constrained to be $\leqslant \log\chi$.

The different qualitative results obtained in the two regimes of small $\delta t \ll 1$ and larger $\delta t \sim \mathcal{O}(1)$ can be understood in terms of entanglement entropy (see Sec. B.2 and B.3 of Appendix B for details). Indeed, in the regime $\delta t \ll 1$, the entanglement of the instantaneous ED state (during the whole annealing dynamics) is relatively low, allowing for an accurate and efficient simulation of dQA with our MPS implementation, with sufficiently low values of the bond dimension χ. In contrast, this is not possible for larger values of δt, since exact dQA generates large amounts of entanglement, hindering an efficient encoding with MPS techniques. Indeed, in this regime, MPS simulations largely deviate from exact dQA simulations, nevertheless surprisingly outperforming them. To clarify this peculiar aspect, we figure out that the high entanglement production and the large values of the final energy density $\varepsilon(1)$, observed in exact dQA for $\delta t \sim \mathcal{O}(1)$, are both

related to Trotter errors, which can be alleviated by a repeated projection on $\mathcal{M}_\chi$.

In previous studies (notably in Ref. [118] and [148]) it has been shown that Trotter errors largely dominate over time discretization errors in the regime of large δt. More explicitly, a comparison can be drawn between the two following protocols [4]

$$\begin{cases} |\phi_P\rangle = \prod_{p=1}^{P} e^{-i\,\hat{H}(s_p)\delta t}\,|\psi_0\rangle & \text{dQA without Trotterization} \\ |\psi_P\rangle = \prod_{p=1}^{P} e^{-i(1-s_p)\,\hat{H}_x\,\delta t}\; e^{-i s_p\,\hat{H}_z\,\delta t}\,|\psi_0\rangle & \text{dQA with Trotterization} \end{cases} \tag{3.2.3}$$

where $p = 1, 2, \ldots, P$ and $s_p = t_p/\tau = p/P$ (as in Eq. 3.1.5). Remarkably, approximating the *exact* continuous time-ordered evolution with P discrete time steps of length δt (dQA *without* Trotterization) turns out to be very accurate even for large time step values (i.e. $\delta t \sim \mathcal{O}(1)$), as long as P is large enough. This "robustness to time discretization" has been confirmed by several theoretical studies, such as Ref. [148], and we verify it numerically in Sec. B.3. On the contrary, in the same regime $\delta t \sim \mathcal{O}(1)$, dQA *with* Trotterization becomes highly inaccurate, leading to a sharp increase in the final energy density $\varepsilon(1)$. To better elucidate this fact, we can rewrite the Trotter split-up as

$$\begin{aligned} & e^{-i(1-s_p)\hat{H}_x\,\delta t}\; e^{-i s_p \hat{H}_z\,\delta t} = \\ & = \exp\left(-i\delta t\,\hat{H}(s_p) \; \overbrace{-\,\frac{1}{2}(\delta t)^2\, s_p\,(1-s_p)[\hat{H}_x, \hat{H}_z] + \cdots}^{\text{spurious Trotter terms}} \right), \end{aligned} \tag{3.2.4}$$

where use has been made of the lowest-order Baker-Campbell-Hausdorff expansion, neglecting $\mathcal{O}(\delta t)^3$ terms. In the regime $\delta t \sim \mathcal{O}(1)$, the large spurious Trotter terms in Eq. 3.2.4 induce non-adiabatic quantum transitions.

This scenario is confirmed by numerical simulations. In particular, we compare the two dQA protocols, *with* and *without* a first-order Trotter split-up, both of them simulated by means of ED, with our MPS scheme (we set $\chi = 10$). [5] This is done in Fig. 3.10 for a binary perceptron of increasing size N, with the parameter α fixed to 0.8 and $N_\xi = \alpha N$ scaling proportionally to N. Data are averaged over 10 random realizations of the patterns $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^{N_\xi}$. In panel a), we plot the final energy density $\varepsilon(1)$, for two representative time step values $\delta t = 0.1$ ($\delta t = 1.0$) for the small (large) time step regime, respectively. In the first case ($\delta t = 0.1$), the agreement betweeen the two dQA protocols and the MPS simulation is remarkable. Indeed, in this regime, Trotterization has negligible effects, and our MPS simulation closely resembles dQA without any Trotter split-up. In the second case ($\delta t = 1.0$), a few observations are in order. First, we notice that dQA *without* Trotterization performs even better than in the small time step regime, reaching

4 Notice that the authors of Ref. [118] adopt a different naming convention for these protocols. Indeed, they denote dQA without Trotterization as "linear-stepQA", while dQA with Trotterization as "linear-dQA".

5 We remark that our MPS methods allow for an efficient representation of dQA *with* Trotterization, as thoroughly explained in Sec. 3.1.

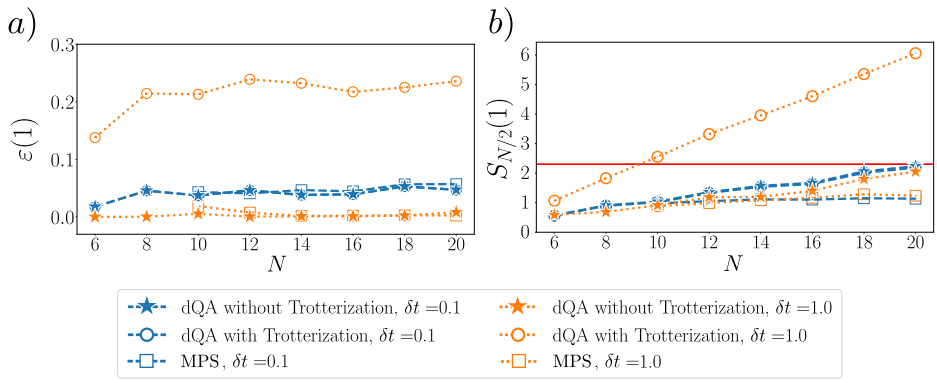

Figure 3.10: Binary perceptron. a) Final energy density $\varepsilon(1)$ and b) half-system entanglement entropy $S_{N/2}(1)$ of the final state, for systems of increasing size N. We set $\alpha = N_\xi/N = 0.8$ and $P = 100$. Data are averaged over 10 realizations of the random patterns, and the resulting standard deviations are smaller than the marker size. We fix two representative values of the time step, $\delta t = 0.1$ (blue) and $\delta t = 1.0$ (orange). We perform an exact simulation of dQA *with* and *without* Trotterization (see Eq. 3.2.3), and we compare these results with our MPS simulations ($\chi = 10$). The red line in panel b) coincides with the maximum entanglement that can be encoded into such MPS, namely $\log \chi$. In the small time step regime, Trotter errors are negligible, thus the two dQA protocols yield very similar results; moreover, they can be efficiently simulated with MPS, since the entanglement content of the final state is quite small. The qualitative picture is drastically different in the large time step regime: here, Trotter errors spoil the dQA effectiveness (large residual energy) and result in large entanglement values. Surprisingly, in this regime, MPS simulations still provide reliable low-entangled final states.

significantly lower values of $\varepsilon(1)$. Thus, dQA would benefit, in principle, of a larger time step. However, Trotter errors here become dominant, severely spoiling the final result and confirming our analysis above. Interestingly, as sketched in Fig. 3.1, our MPS simulation turns out to mitigate Trotter errors, providing good solutions even in this regime.

In panel b), we plot the von Neumann entanglement entropy for the final state.[6] In general, along the annealing protocol, this quantity is defined by

$$\begin{aligned} S_{N/2}(s) &= -\mathrm{Tr}_{s_1 \dots s_{N/2}} \left(\hat{\rho}(s) \log \hat{\rho}(s) \right) \\ \hat{\rho}(s) &= \mathrm{Tr}_{s_{N/2+1} \dots s_N} \left(|\psi(s)\rangle \langle \psi(s)| \right), \end{aligned} \tag{3.2.5}$$

where $|\psi(s)\rangle$ and $\hat{\rho}(s)$ are the whole-system state and the reduced density matrix of half-system, respectively. Here, we set $s = 1$, corresponding to the final annealed states reported in Eq. 3.2.3. For $\delta t = 0.1$, the entanglement entropy of the final state grows relatively slowly with N (for both dQA protocols), allowing for efficient MPS simulations. On the other hand, for $\delta t = 1.0$, we observe that Trotter errors give rise to large amounts of entanglement entropy in the final state. Now, since our algorithm iteratively reduces the entanglement of the quantum state at each time step, we argue

6 We consider the so-called entanglement entropy at half chain, i.e. between two subsystems of the same size.

that it acts by iteratively projecting away the contributions given by the spurious terms introduced by the Trotter split-up, since these, as shown, are the main source of entanglement. Since the same terms are also responsible of low-quality final states (large values of final energy density), our algorithm can quite re-establish the performance of dQA *without* the Trotter split-up, thus providing a significant improvement in the final result.

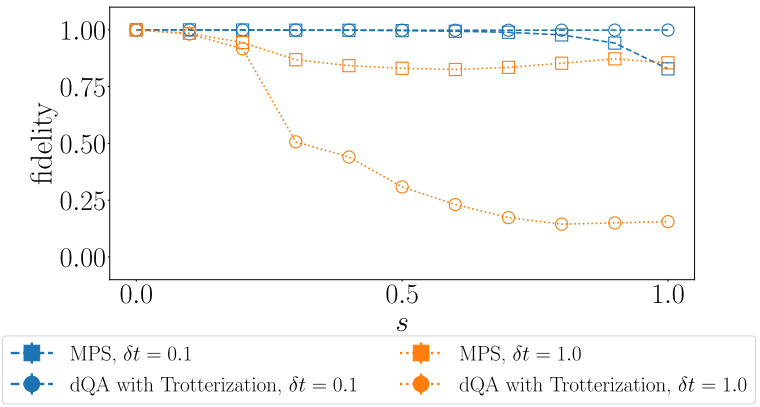

Figure 3.11: Binary perceptron. Fidelity between the reference state (instantaneous state of dQA *without* Trotterization) and the instantaneous state of a Trotterized dQA dynamics, either obtained with ED (circles) or with MPS simulations (squares). We fixed $P = 100$ and $\tau = 10$ (blue) or $\tau = 100$ (orange). The fidelity is plotted vs the rescaled annealing time $s = t/\tau \in [0, 1]$. We set $N = 18$, $N_\xi = 14$ ($\alpha \simeq 0.78$) and we averaged over 10 realizations of the random patterns.

In addition, a direct evidence on the MPS effectiveness in mitigating Trotter errors, and thus reproducing the actual QA dynamics, is provided in Fig. 3.11. Here, we fix the instantaneous state of dQA *without* Trotterization as a reference state, representing the true QA dynamics. [7] We plot the fidelity, during the annealing, between this reference state and two other states, namely the instantaneous dQA state *with* Trotter split-up obtained either with ED, or approximated by means of our MPS framework. Remarkably, in the small time step regime (once again $\delta t = 0.1$) both fidelity values are quite close to 1, whereas in the other regime ($\delta t = 1.0$) MPS simulations prove more faithful than exact dQA with Trotterization, especially in the last stages of the annealing.

Finally, let us remark that the outlined results provide some clear insight on the correctness of our initial sketch in Fig. 3.1. Indeed, as shown in Fig 3.10 (see also Sec. B.2), the final state from a dQA *without* Trotterization is generally low-entangled and close to the continuous time-ordered QA results, whereas the Trotterization represents the main source of entanglement in the regime of large δt. It may therefore be convenient to approximate this Trotterized evolution as an MPS with relatively low bond dimension. As shown numerically in Fig. 3.11, the MPS evolved state remains closer

7 Strictly speaking, this is certainly an approximation, since this state clearly does not satisfy the ideal continuous time-ordered dynamics. However, dQA *without* Trotterization often turns out to be an outstanding approximation for a wide range of time step values (even if $\delta t \sim \mathcal{O}(1)$), as anticipated above (see Ref. [148]), and proven numerically in Sec. B.3.

to the annealing dynamics, if compared to the exact Trotterized evolution, confirming the qualitative representation sketched in Fig. 3.1.

3.2.4 The Hopfield model

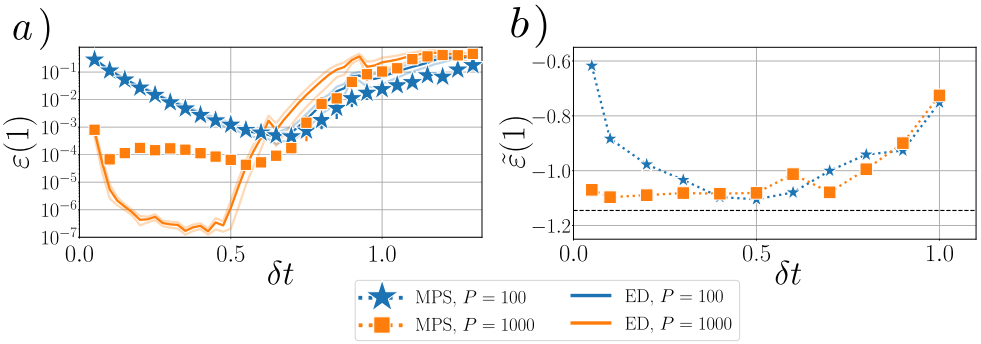

Figure 3.12: Hopfield model. Residual energy density as a function of time-step length δt. We set $N = 21$, $N_\xi = 2$ in panel a) and $N = 100$, $N_\xi = 13$ in panel b). In panel a), data are averaged over five different sets of random patterns, and MPS results (full symbols) are compared with ED results (solid lines). The (small) MPS error bars represent the standard error of the mean, often smaller than the marker size. Solid lines with lower opacity represent average $\pm$ the standard error of the mean for ED data. In panel b), we plot $\tilde{\varepsilon}(1)$ since the exact ground state energy is not known a priori and ED cannot be performed, due to the large system size. However, an estimate of the actual ground state energy is represented by a black horizontal dashed line (obtained by state-of-the-art classical solvers, see main text).

In this Section, we summarize results for the Hopfield model, defined by Eq. 3.1.3. As for the binary perceptron, we first focus on a relatively small size $N = 21$, such that exact solutions can be easily found by enumeration, and MPS results can be compared with ED. We set $N_\xi = 2$, so that $\alpha = N_\xi/N \simeq 0.095$ is close to the zero-temperature critical value $\alpha_c \simeq 0.138$, and we run our simulations for different randomly-generated training sets $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^{N_\xi}$. In Fig. 3.12 a), we plot the final energy density $\varepsilon(1)$ as a function of the time step δt. As for the other models examined in the main text, we observe a close agreement between MPS and ED for small δt (notice the log scale on the $y-$axis, therefore deviations for $P = 1000$ are actually small, of $\mathcal{O}(10^{-4})$). For δt values of $\mathcal{O}(1)$, larger deviations are observed, with MPS generally outperforming ED. Moreover, in Fig. 3.12 b), we consider a single instance of size $N = 100$, still close to the critical point (we set $N_\xi = 13$, so $\alpha = 0.13$). In this case, the exact ground state energy is not known *a priori* and therefore we plot $\tilde{\varepsilon}(1) = \langle\psi(1)|\hat{H}_z|\psi(1)\rangle/N$. An accurate estimate of the actual minimum energy can be obtained by means of an optimized classical solver (we used the online solver http://spinglass.uni-bonn.de/): this value is represented by a black dotted line. This solver employs state-of-the-art classical optimization methods, namely a version of the branch and bound algorithm [149]. In summary, these results show that our MPS methods are effective in finding a good approximation of the target ground state even for large system sizes, far beyond the reach of ED.

3.2.5 MPS compilation to quantum circuits

The scope of this section is to discuss the representation of an MPS as a quantum circuit, i.e. as a sequence of unitary quantum gates acting on a blank qubit register $|\mathbf{0}\rangle$. We perform several simulations by employing this iterative quantum circuit optimization described in Sec. 1.8. In our case, the target state $|\psi\rangle$ is the MPS obtained by contracting the Tensor Network structure representing the whole time evolution (see Fig. 3.3), by means of Algorithm 1. As an illustrative example, the following table includes results obtained for the binary perceptron ($N = 21$, $N_\xi = 17$, a single training set), by setting as target state the final state of dQA with $P = 500$. Different values of δt (and hence $\tau = P\delta t$) are considered. We report the values of fidelity $\mathcal{F}$ obtained after $N_{iter} = 3000$ optimization iterations of a circuit of depth $D = 4$. We also show the corresponding value of the cost function $\langle \hat{H}_z \rangle / N = \langle \mathbf{0}|\hat{U}^\dagger_{opt} \hat{H}_z \hat{U}_{opt}|\mathbf{0}\rangle / N$ and the total success probability $\sum_{a=1}^{N_{sol}} p(\boldsymbol{\sigma}^*_a) = \sum_{a=1}^{N_{sol}} |\langle \boldsymbol{\sigma}^*_a|\hat{U}_{opt}|\mathbf{0}\rangle|^2$, where $\hat{U}_{opt}$ represents the *optimized* quantum circuit.

δt	$\mathcal{F}$	$\langle \hat{H}_z \rangle / N$	$\sum_{a=1}^{N_{sol}} p(\boldsymbol{\sigma}^*_a)$
1.0	0.918	$1.61 \cdot 10^{-3}$	0.941
1.2	0.944	$1.63 \cdot 10^{-3}$	0.944
1.4	0.951	$1.10 \cdot 10^{-3}$	0.961
1.6	0.894	$1.08 \cdot 10^{-3}$	0.965

Data show that, even if the optimized quantum circuit does not reach particularly high fidelity $\mathcal{F}$ with the target state, anyway the final energy densities are $\mathcal{O}(10^{-3})$, and the overlap with classical solutions significantly large. This is indeed a remarkable result, for a shallow quantum circuit with $D = 4$. We mention that deeper circuits (larger values of D) are more challenging to optimize, possibly because the algorithm gets stuck into sub-optimal local minima, but we leave a more systematic numerical study to future work.

Let us notice that residual energy densities of $\mathcal{O}(10^{-3})$ are comparable with the best result attained by exact dQA with $P = 1000$ steps, and an optimal choice of δt (see Fig. 3.6). Thus, we argue that by applying the Algorithm 1, and then approximating the resulting MPS $|\psi\rangle$ with a quantum circuit, one can obtain a small set of two-qubits unitary gates with a simple architecture yielding a good quantum state, bearing large overlap with the classical solutions. This state may represent a valuable starting point, allowing for further circuit optimization on near-term quantum devices.

3.3 CONCLUSIONS

We tackled the quantum adiabatic optimization of a large class of hard classical minimization problems, encompassing p-spin models, and the paradigmatic binary perceptron and Hopfield models. We focused on the well-known Quantum Annealing protocol in its digitized version, and we developed a completely new, *ad hoc* approach, to simulate quantum adiabatic

time evolution for these systems. This approach relies on Tensor Network methods and Matrix Product States, exploiting efficiently the available classical resources, allowing for classical simulations well-beyond the usual size limits of Exact Diagonalization.

Our Tensor Network framework could be easily employed in future work to study different non-equilibrium time evolution or the imaginary time evolution for the same class of models. Moreover, our methods could be generalized to simulate other quantum optimization algorithms, e.g. hybrid quantum-classical variational schemes such as QAOA. In this case, the Tensor Network representing the time evolution would be parameter-dependent, with the parameters being iteratively optimized by some classical routine.

We stress that the broad class of models that can be simulated in our framework includes prototypical discrete Neural Networks of significant theoretical interest in Machine Learning, which do not admit, in general, a description in terms of few-body spin interactions. With the technological progress and the availability of real quantum devices, it is a promising route to assess the effectiveness of quantum optimization methods in this context.

Future perspectives concern the possibility to extend our methods to the optimization of more challenging discrete Neural Networks (essentially, multilayered perceptrons). This exciting chance could significantly expand our knowledge on the abilities of Quantum Computers to solve the hard optimization tasks involved in classical deep learning.

4 NONSTABILIZERNESS VIA PERFECT PAULI SAMPLING OF MATRIX PRODUCT STATES

Quantum advantage [30, 150] relies on harnessing the intrinsic complexity of quantum systems to surpass classical computing limitations, potentially enabling efficient solutions to NP problems [15, 151, 152]. Physicists agree to identify the entanglement as a fundamental feature accounting for this complexity, thus making necessary to exploit it proficiently in any quantum computation. Indeed, quantifying entanglement in many-body systems is a long-standing research focus, with various well-established measures, such as purity, entanglement entropy, negativity and mutual information [153–156].

Nevertheless, entanglement is not the sole *resource* which need to be quantified in order to discriminate between easy and hard to simulate quantum states. Indeed, it turns out that there exist several states encoding an extensive amount of entanglement which can still be simulated efficiently on a classical computer. These states are part of the *stabilizer states* [157], defined as quantum states exclusively achievable using Clifford unitaries from the computational basis state $|0\dots0\rangle$ [158–163]. Clifford group $\mathcal{C}_N$ represent a class of unitary transformations that maps strings of Pauli operators over N-qubits into other Pauli strings [36, 162], i.e.

$$\mathcal{C}_N = \{U \in \mathcal{U}_{2^N\times 2^N} \text{ s.t. } U\mathcal{P}_N U^\dagger = \mathcal{P}_N\},$$

where

$$\mathcal{P}_N = \{\pm\mathbb{1}, \pm i\mathbb{1}, \pm\sigma^x, \pm i\sigma^x, \pm\sigma^y, \pm i\sigma^y, \pm\sigma^z, \pm i\sigma^z\}^{\otimes N}$$

is the Pauli group [1]. Due to this underlying structure, stabilizer states can be compactly represented classically. Indeed, they are completely characterized by a set of N independents Pauli strings, which can be conveniently stored, allowing for a compact and efficient description. Clifford unitaries and measurement of Pauli operators can be efficiently executed using this representation by straightforwardly updating the stored Pauli strings [158, 159]. For a detailed introduction to the formalism of stabilizer states and the Clifford group, we refer the reader to Appendix C. Given the aforementioned considerations, assessing the difficulty of simulating a quantum state, irrespective of its entanglement content, necessitates the definition of a metric that accounts for the number of non-Clifford operations essential for state

1 In this Chapter, we will omit the symbol ˆ for operators.

preparation [164, 165]. This quantity has been dubbed *nonstabilizerness* or *quantum magic*. It turns out that nonstabilizerness is also related to the emergence of quantum chaos [166, 167].

Several measures of nonstabilizerness have been proposed so far in quantum information theory [168–170], as for instance the Robustness of Magic (see Appendix C and Ref. [169]). However, their computation is usually challenging [171], as it frequently involves challenging optimizations. As a matter of fact, quantifying nonstabilizerness beyond a few qubits remains a major challenge. Recently the Stabilizer Rényi Entropies (SREs) were introduced in Ref. [172] as a possible way of quantifiying the nonstabilizerness of a quantum state (see Appendix C for a detailed discussion). Interestingly, SREs allow the evaluation of the nonstabilizerness stored in the ground state of the paradigmatic transverse field Ising chain [173]. However, since they depend on expectation values of all possible Pauli strings, computing SREs of a generic state is exponentially costly with the number of qubits. Nevertheless when the N-qubits state admits a Matrix Product State (MPS) representation with finite bond dimension χ, the SREs can be computed as the norm of a "2n-replica" MPS with effective bond dimension χ^{2n}, where n (integer) represents the Rényi index [174]. Unfortunately, such norm can be computed at a cost $O(N\chi^{6n})$, thus having an unfavorable scaling with the bond dimension. For any practical purpose, this makes the approach unfeasible for $n > 2$ [2].

To overcome such limitations, we propose a new method which exploits the probabilistic nature of the SREs. The Algorithm relies on a novel and efficient MPS sampling in the Pauli basis, reminiscent of some well-established MPS techniques [175, 176]. The sampling is perfect since we directly obtain samples from the target probability distribution, without Markov chains. By sampling over $\mathcal{N}$ Pauli strings realizations, we are able to estimate the SREs with a computational cost scaling as $O(\mathcal{N}N\chi^3)$. We first benchmark our approach over a set of random realization of MPS states with large bond dimension. We then study the nonstabilizerness in the ground-state of the quantum Ising chain, showing a prefect agreement with the free-fermions calculation. Finally, we use our method to compute for the first time the non-equilibrium dynamics of the SREs after a quench. We consider the Ising model with or without a longitudinal field and show how the confinement of the excitations [177], hugely affecting the entanglement dynamics, may play a role also in the time-evolution of the SREs.

4.1 PRELIMINARIES

Let us consider a quantum system consisting of N qubits. We identify the Pauli matrices by $\{\sigma^\alpha\}_{\alpha=0}^3$, with $\sigma^0 = \mathbb{1}$, and with $\boldsymbol{\sigma} = \prod_{j=1}^N \sigma_j \in \tilde{\mathcal{P}}_N$ a

2 Only for $n = 2$, it is possible to exploit additional symmetries, further reducing the computational cost to $O(\chi^4)$.

generic N−qubits Pauli strings where $\tilde{\mathcal{P}}_N = \{\sigma^0, \sigma^1, \sigma^2, \sigma^3\}^{\otimes N}$ [3]. For a pure normalised state $\rho = |\psi\rangle\langle\psi|$, the SREs [172] are given by

$$M_n(\rho) = \frac{1}{1-n} \log \sum_{\boldsymbol{\sigma} \in \tilde{\mathcal{P}}_N} \frac{1}{2^N} \mathrm{Tr}[\rho\, \boldsymbol{\sigma}]^{2n}. \tag{4.1.1}$$

To understand the relation with usual Rényi entropies, one has to observe that the non-negative real-valued function $\Pi_\rho(\boldsymbol{\sigma}) = \frac{1}{2^N} \mathrm{Tr}[\rho\, \boldsymbol{\sigma}]^2$ sums to 1, since

$$\sum_{\boldsymbol{\sigma} \in \tilde{\mathcal{P}}_N} \Pi_\rho(\boldsymbol{\sigma}) = \mathrm{Tr}\left[\rho \sum_{\boldsymbol{\sigma} \in \tilde{\mathcal{P}}_N} \frac{\boldsymbol{\sigma}}{\sqrt{2^N}} \mathrm{Tr}[\rho \frac{\boldsymbol{\sigma}}{\sqrt{2^N}}]\right] = \mathrm{Tr}\left[\rho^2\right] = 1\,, \tag{4.1.2}$$

where we used the decomposition of ρ in terms of the orthonormal Pauli basis $\{\frac{\boldsymbol{\sigma}}{\sqrt{2^N}}\}$, and the fact that $\mathrm{Tr}[\rho^2] = 1$ for pure states. Thus, we can interpret $\Pi_\rho(\boldsymbol{\sigma})$ as a probability distribution on the set of Pauli strings. Thus

$$M_n(\rho) = \frac{1}{1-n} \log \sum_{\boldsymbol{\sigma} \in \tilde{\mathcal{P}}_N} \Pi_\rho(\boldsymbol{\sigma})^n - N \log 2\,, \tag{4.1.3}$$

apart from a constant, does coincides with the n-Rényi entropy of the distribution $\Pi_\rho(\boldsymbol{\sigma})$, and it reduces to the Shannon entropy

$$M_1(\rho) = -\sum_{\boldsymbol{\sigma} \in \tilde{\mathcal{P}}_N} \Pi_\rho(\boldsymbol{\sigma}) \log \Pi_\rho(\boldsymbol{\sigma}) - N \log(2)\,. \tag{4.1.4}$$

or $n \to 1$. It has been shown that SREs have the following properties [172], accordingly being a good measure of nonstabilizerness: i) M_n vanishes for stabilizer states whereas is positive for other states; ii) are invariant under Clifford unitaries; iii) are additive. Moreover, they grow extensively with the system size N, thus making possible to define nonstabilizerness density $m_n = M_n/N$ [174]. A violation of monotonicity for the SREs with $0 \leqslant n < 2$ has been reported for systems undergoing measurements in the computational basis [178].

Computing the SREs in Eq. (4.1.1) requires the evaluation of the expectation value of a generic power $\Pi_\rho(\boldsymbol{\sigma})^{n-1}$ (or $\log \Pi_\rho(\boldsymbol{\sigma})$ for $n = 1$) over the probability distribution $\Pi_\rho(\boldsymbol{\sigma})$ itself. This suggests a natural way to estimate the SREs, based on a sampling from $\Pi_\rho(\boldsymbol{\sigma})$.

4.2 CONDITIONAL SAMPLING

The task of sampling from the set of the Pauli strings $\boldsymbol{\sigma}$, which has size $D = 4^N$, may appear as exponentially hard. To overcome this difficulty, we rewrite the full probability in terms of conditional and prior (or marginal) probabilities as

$$\Pi_\rho(\boldsymbol{\sigma}) = \pi_\rho(\sigma_1)\pi_\rho(\sigma_2|\sigma_1)\cdots\pi_\rho(\sigma_N|\sigma_1\cdots\sigma_{N-1}) \tag{4.2.1}$$

3 Notice that, on the contrary of $\mathcal{P}_N$, $\tilde{\mathcal{P}}_N$ is not a group, since for instance $\sigma^1\sigma^2 = i\sigma^3$.

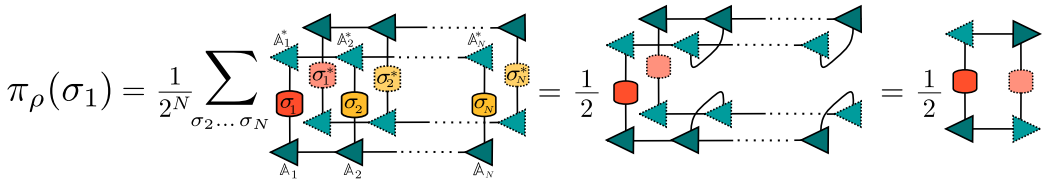

Figure 4.1: MPS evaluation of the marginal probability $\pi_\rho(\sigma_1)$. Dotted lighter shapes represent conjugate tensors. Contractions over the auxiliary indices can be easily carried out thanks to the property in Eq. (4.3.3), together with the right-normalization of the $\mathbb{A}_i$ tensors.

where $\pi_\rho(\sigma_j|\sigma_1\cdots\sigma_{j-1}) = \frac{\pi_\rho(\sigma_1\cdots\sigma_j)}{\pi_\rho(\sigma_1\cdots\sigma_{j-1})}$ is the probability that the Pauli matrix σ_j occurs at position j given that the string $\sigma_1\cdots\sigma_{j-1}$ has already occurred at positions $1\ldots j-1$, no matter the occurrences in the rest of the system (i.e. marginalising over all possible Pauli strings for the reaming qubits $j+1\ldots N$). Specifically, one has

$$\pi_\rho(\sigma_1\cdots\sigma_j) = \sum_{\boldsymbol{\sigma}\in\mathcal{P}_{N-j}} \frac{1}{2^N}\,\mathrm{Tr}[\rho\,\sigma_1\cdots\sigma_j\boldsymbol{\sigma}]^2\,. \tag{4.2.2}$$

In other terms, the conditional probability at the step j, i.e. $\pi_\rho(\sigma_j|\sigma_1\cdots\sigma_{j-1})$, can be thought as the probability $\pi_{\rho_{j-1}}(\sigma_j)$ of getting σ_j in the partially projected state

$$\rho_{j-1} \equiv \frac{\rho|_{\sigma_1\cdots\sigma_{j-1}}}{\pi_\rho(\sigma_1\cdots\sigma_{j-1})^{1/2}} \tag{4.2.3}$$

where we have defined the state

$$\rho|_{\sigma_1\cdots\sigma_{j-1}} \equiv 2^{-N} \sum_{\boldsymbol{\sigma}\in\mathcal{P}_{N-j+1}} \mathrm{Tr}[\rho\,\sigma_1\cdots\sigma_{j-1}\boldsymbol{\sigma}]\,\sigma_1\cdots\sigma_{j-1}\boldsymbol{\sigma} \tag{4.2.4}$$

where, in the Pauli matrices decomposition of ρ, we are only keeping the contribution with fixed $\sigma_1\cdots\sigma_{j-1}$. Notice that such state is not normalised, however $\mathrm{Tr}[\rho_{j-1}^2] = 1$, and the probability that the remaining string $\boldsymbol{\sigma} \in \mathcal{P}_{N-j+1}$ occurs is exactly given by $\pi_\rho(\boldsymbol{\sigma}|\sigma_1\cdots\sigma_{j-1})$. From the definition in Eq. (4.2.3), we can easily get the recursive relation $\rho_j = \pi_{\rho_{j-1}}(\sigma_j)^{-1/2}\rho_{j-1}|_{\sigma_j}$. Thanks to that, we can generate the outcomes (and the probabilities of that outcomes) by iterating over each single qubits, and sampling each local Pauli matrix according to the conditional probabilities. Once a local outcome occurs, the state is updated accordingly, and the iteration proceeds until all qubits are sampled. At the end of this procedure, as a result of Eq. (4.2.1), we generated configurations $\boldsymbol{\sigma}$ with probability $\Pi_\rho(\boldsymbol{\sigma})$. In order for this method to be computationally affordable, we need an efficient way of: *(i)* evaluating the conditional probabilities; *(ii)* updating the state according to the local outcome. In the following Section, we show that these conditions are met whenever the state admits an MPS representation.

Algorithm 2 Pauli sampling from MPS

Input: an MPS $|\psi\rangle$ of size N

1: Put the MPS in right-normalized form.
2: Initialize $\mathbb{L} = (1)$ and $\Pi = 1$ (see Fig.4.2 a))
3: **for** $(i = 1, i = N, i++)$ **do**
4: Compute the probabilities $\pi(\alpha) = \pi_\rho(\sigma^\alpha|\sigma_1 \cdots \sigma_{i-1})$ for $\alpha \in \{0,1,2,3\}$, as in Fig.4.2 b).
5: Generate a random value of α according to $\pi(\alpha)$
6: Set $\sigma_i = \sigma^\alpha$, update $\Pi \to \Pi \cdot \pi(\alpha)$
7: Update $\mathbb{L}$ as in Fig.4.2 c).
8: **end for**

Output: a Pauli string $\boldsymbol{\sigma}$ and the probability $\Pi(\boldsymbol{\sigma})$

4.3 MPS ITERATIVE ALGORITHM

We consider a pure state $|\psi\rangle$ represented in the MPS form [27, 38, 126]

$$|\psi\rangle = \sum_{s_1,s_2,\ldots,s_N} \mathbb{A}_1^{s_1}\mathbb{A}_2^{s_2}\cdots\mathbb{A}_N^{s_N}|s_1,s_2,\ldots,s_N\rangle\,, \tag{4.3.1}$$

with $\mathbb{A}_j^{s_j}$ being $\chi\times\chi$ matrices, except at the left (right) boundary where $\mathbb{A}_1^{s_1}$ ($\mathbb{A}_N^{s_N}$) is a $1\times\chi$ ($\chi\times 1$) row (column) vector. Here $|s_j\rangle \in \{|0\rangle,|1\rangle\}$ is a local computational basis. The state is assumed right-normalised, namely $\sum_{s_j}\mathbb{A}_j^{s_j}(\mathbb{A}_j^{s_j})^\dagger = \mathbb{1}$. We refer the reader to Chapter 1 for a thorough introduction to MPS. Following the conditional sampling prescription described in the previous section, we start from the first term of the expansion in Eq. (4.2.1). This can be written as

$$\pi_\rho(\sigma_1) = \frac{1}{2^N}\sum_{\boldsymbol{\sigma}\in\mathcal{P}_{N-1}} \langle\psi|\sigma_1\boldsymbol{\sigma}|\psi\rangle\,\langle\psi^*|\sigma_1^*\boldsymbol{\sigma}^*|\psi^*\rangle\,, \tag{4.3.2}$$

where we used the fact that the Pauli matrices are hermitian. In terms of the operators $\Lambda_{\sigma_i} = \frac{1}{2}\sigma_i\otimes\sigma_i^*$ and $\Lambda_i = \frac{1}{2}\sum_{\sigma_i}(\sigma_i\otimes\sigma_i^*)$, each acting on the local Hilbert space given by a spin and its replica, the previous equation reads $\pi_\rho(\sigma_1) = \big[\langle\psi|\otimes\langle\psi^*|\big]\Lambda_{\sigma_1}\Lambda_2\cdots\Lambda_N\big[|\psi\rangle\otimes|\psi^*\rangle\big]$. Now, the following property can easily be proven

$$\big[\langle s_i'|\otimes\langle r_i'|\big]\Lambda_i\big[|s_i\rangle\otimes|r_i\rangle\big] = \delta_{s_i',r_i'}\delta_{s_i,r_i}\,, \tag{4.3.3}$$

meaning that Λ_i is just two copies of the identity operator connecting the spin $|s_i\rangle$ and its replica (whose local computational basis is now indicated as $|r_i\rangle \in \{|0\rangle,|1\rangle\}$). Using Eq. (4.3.3) together with the right-normalization of the MPS, the computation of Eq. (4.3.2) reduces in the following local tensor contraction

$$\pi_\rho(\sigma_1) = \frac{1}{2}\sum_{s_1,s_1',r_1,r_1'} (\mathbb{A}_1^{s_1'})^*\mathbb{A}_1^{r_1'}(\sigma_1)_{s_1's_1}(\sigma_1^*)_{r_1'r_1}\mathbb{A}_1^{s_1}(\mathbb{A}_1^{r_1})^*\,, \tag{4.3.4}$$

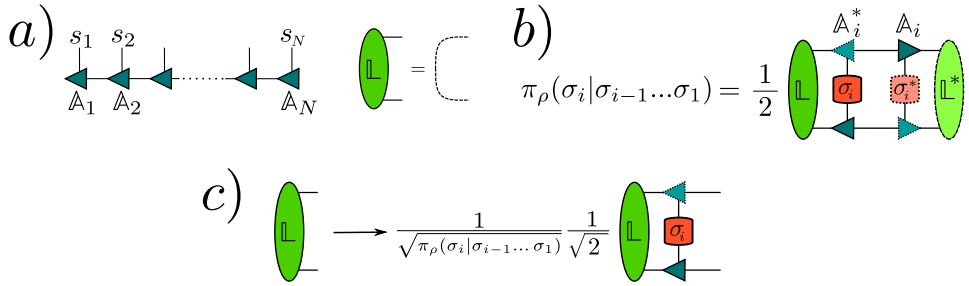

Figure 4.2: The iterative sampling Algorithm 2.

which is represented in Fig. 4.1 by means of the standard Tensor Network graphical notation [27, 38].

After evaluating $\pi_\rho(\sigma_1)$ for $\sigma_1 \in \{\sigma^0, \sigma^1, \sigma^2, \sigma^3\}$, one can extract a sample from this distribution, obtaining the first element of the string. The information about the partially projected state Eq. (4.2.3) is encoded in an effective environment matrix $\mathbb{L} = \frac{1}{\sqrt{2\,\pi_\rho(\sigma_1)}} \sum_{s_1, s_1'} (\mathbb{A}_1^{s_1'})^*(\sigma_1)_{s_1' s_1} \mathbb{A}_1^{s_1}$. The calculation of the next terms of Eq. (4.2.1) and the extraction of the remaining σ_i proceeds following the same line. The full sampling recipe is summarized in the Algorithm 2, and graphically supported in Fig. 4.2. Extension of the Algorithm to mixed states is discussed in Supplementary Materials.

4.4 SAMPLING ERROR

We now discuss the statistical errors of the sampling algorithm, and their scaling with the system size N. We first consider the case of estimating the n−SRE, with $n > 1$. As we saw, the estimation of

$$q_n = \sum_{\boldsymbol{\sigma} \in \tilde{\mathcal{P}}_N} \Pi_\rho(\boldsymbol{\sigma})^n \tag{4.4.1}$$

is achieved by a statistical average over the samples $\{\boldsymbol{\sigma}_\mu\}_{\mu=1}^{\mathcal{N}}$, that means using the estimator

$$\tilde{q}_n = \frac{1}{\mathcal{N}} \sum_{\mu=1}^{\mathcal{N}} \Pi_\rho(\boldsymbol{\sigma}_\mu)^{n-1}\,. \tag{4.4.2}$$

Afterwards, we evaluate the density of nonstabilizerness as

$$\tilde{m}_n = \big(N(1-n)\big)^{-1} \log \tilde{q}_n - \log 2 \tag{4.4.3}$$

Notice that $\tilde{q}_n$ is an unbiased estimator of q_n, since $\overline{\tilde{q}_n} = q_n$ ($\bar{}$ indicating the average over the uncorrelated samples, each distributed according to $\Pi_\rho(\boldsymbol{\sigma})$). The fluctuations of $\tilde{q}_n$ are characterized by its variance, which can be easily evaluated as $\mathrm{Var}[\tilde{q}_n] = \mathrm{Var}[\Pi_\rho^{n-1}]/\mathcal{N}$. For every $n > 1$, one has $\mathrm{Var}[\Pi_\rho^{n-1}] < 1$ and thus we can upper bound the variance of the estimator obtaining $\mathrm{Var}[\tilde{q}_n] < \text{const.}/\mathcal{N}$, where const. is a constant of $o(1)$, whose value is independent of the size $D = 4^N$ of the support of $\Pi_\rho(\boldsymbol{\sigma})$. This means that the statistical error on $\tilde{q}_n$ can be reduced arbitrarily by increasing the

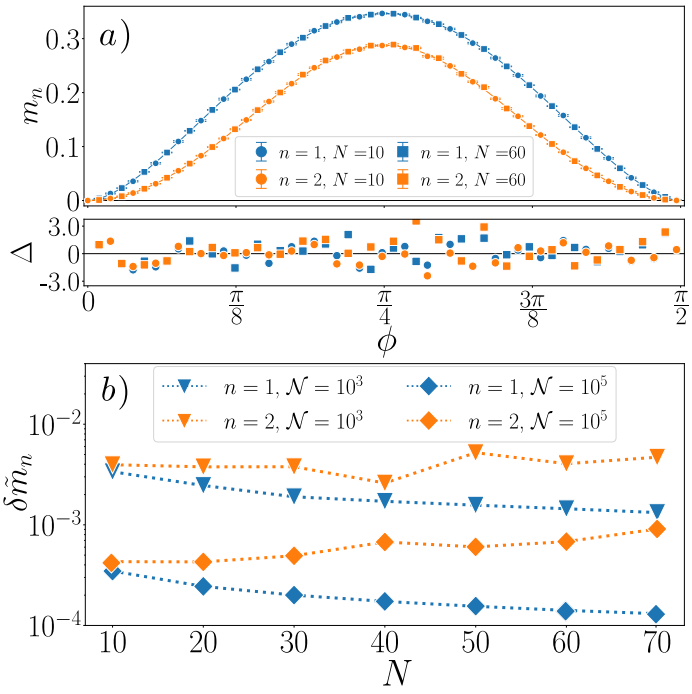

Figure 4.3: a) Density of nonstabilizerness of $|\psi\rangle = U_{\mathcal{C}}|T_\phi\rangle^{\otimes N}$ for $N = 10, 60$, $\mathcal{N} = 10^4$ and Rényi index $n = 1, 2$. In the lower strip we show the deviation from the analytical value $\Delta = (m_n - \tilde{m}_n)/\delta\tilde{m}_n$, $\tilde{m}_n$ being our estimation and $\delta\tilde{m}_n$ the propagated statistical error. b) The error $\delta\tilde{m}_n$ as a function of the system size N for fixed $\mathcal{N} = 10^3, 10^5$ and $\phi \simeq \pi/4$.

number of samples, no matter the system size N. However, since the uncertainty on $\tilde{m}_n$ propagates (at first order) as $\delta\tilde{m}_n \propto \delta\tilde{q}_n/\tilde{q}_n$ and both $\tilde{q}_n$, $\delta\tilde{q}_n$ are exponentially vanishing with N for typical probability distributions, $(\delta\tilde{m}_n)^2 \sim \frac{1}{\mathcal{N}}\mathrm{Var}[\Pi_\rho^{n-1}]/(\overline{\Pi_\rho^{n-1}})^2$ is generally exponentially increasing with N [4]. Nevertheless, for the physical states we have examined, the estimation error $\delta\tilde{m}_n$ is always under control for reasonable values of $\mathcal{N}$ (see next Section and Supplementary Materials for further details). For $n = 1$ we evaluate $q_1 = \sum_{\boldsymbol{\sigma}\in\tilde{\mathcal{P}}_N} \Pi_\rho(\boldsymbol{\sigma}) \log \Pi_\rho(\boldsymbol{\sigma})$ via the estimator $\tilde{q}_1 = 1/\mathcal{N} \cdot \sum_{\mu=1}^{\mathcal{N}} \log \Pi_\rho(\boldsymbol{\sigma}_\mu)$. We have $\mathrm{Var}[\tilde{q}_1] = \mathrm{Var}[\log \Pi_\rho]/\mathcal{N}$ and thus we are interested in giving an upper bound for $\mathrm{Var}[\log \Pi_\rho]$. Several works, e.g. Ref. [155], establish that $\mathrm{Var}[\log \Pi_\rho] \leqslant \frac{1}{4}\log^2(D) + 1$. Thus, in our case, $\mathrm{Var}[\tilde{q}_1] \lesssim N^2 \log^2(2)/\mathcal{N}$ meaning that *in the worst scenario* the number of samples has to scale as N^2 to reach a given accuracy in the estimation.

4 This because $\overline{\Pi_\rho^{2(n-1)}}/(\overline{\Pi_\rho^{n-1}})^2 \geqslant 1$

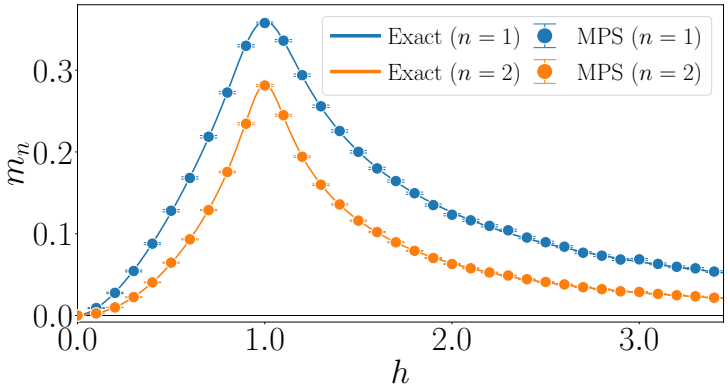

Figure 4.4: Nonstabilizerness density of the Ising ground state ($g = 0$) with periodic boundary conditions, for a system of size $N = 14$ and Rényi index $n = 1, 2$. Exact results obtained in the free fermions representation [173] are compared with MPS sampling ($\mathcal{N} = 10^4$).

4.5 SAMPLING OF MIXED STATES AND OPERATORS

The definition of the Stabilizer Rényi Entropies M_n can be easily extended to arbitrary non-pure states ρ by normalizing the probability distribution Π_ρ with the purity $\mathrm{Tr}[\rho^2] \leqslant 1$, thus redefining

$$\Pi_\rho(\boldsymbol{\sigma}) = \frac{1}{2^N} \frac{\mathrm{Tr}[\rho\, \boldsymbol{\sigma}]^2}{\mathrm{Tr}[\rho^2]} \,. \tag{4.5.1}$$

Our sampling approach can be generalized to estimate $M_n(\rho)$ in the case in which $\rho = \rho_{[L+1,N]}$ is the reduced density matrix describing the rightmost $N-L$ qubits embedded in a larger pure MPS state $|\psi\rangle$ of size N. Exploiting a mixed canonical form for $|\psi\rangle$, one can obtain $\rho_{[L+1,N]}$ as in Fig. 4.5a). After that, the only thing that changes in our algorithm is the initialization of the environment matrix $\mathbb{L}$. As we saw, this is set to (1) for a pure state, whereas in the general case $\mathbb{L} = {}^2/\sqrt{\mathrm{Tr}(^4)}$ in terms of the Schmidt eigenvalues associated to the bipartition between the first L qubits and the last $N-L$ qubits (see Fig. 4.5b)).

Moreover, it is straightforward to adapt the sampling technique to target Matrix Product Operators (MPO). Indeed, any MPO can be reshaped as an MPS whose local physical dimension is increased from 2 to 4 (see Fig. 4.6). In this way, by applying the proposed method, one could gain access to the SRE of an operator O written as MPO. In such scenario, the SRE serves as a form of participation entropy characterizing the operator spreading over the space of Pauli strings.

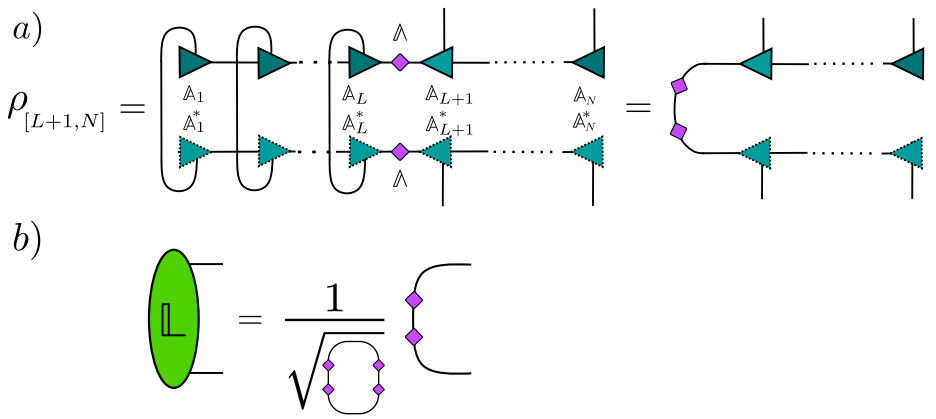

Figure 4.5: a) The reduced density matrix $\rho_{[L+1,N]}$ of the rightmost $N-L$ qubits embedded in a larger pure MPS state is obtained by tracing away the first L qubits. represent the Schmidt eigenvalue associated to the bipartition. b) The new appropriate initialization of the environment matrix $\mathbb{L}$, i.e. $\mathbb{L} = {}^2/\sqrt{\mathrm{Tr}(4)}$.

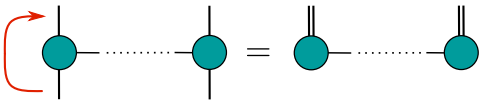

Figure 4.6: Any MPO can be reshaped as an MPS by merging the two physical indices of each local tensor.

4.6 NUMERICAL EXPERIMENTS

As a first benchmark of our algorithm, we consider the T-state $|T_\phi\rangle = (|0\rangle + e^{i\phi}|1\rangle)/\sqrt{2}$, with ϕ ranging in $[0, \pi/2]$. A straightforward calculation yields to

$$\begin{aligned} M_1(|T_\phi\rangle\langle T_\phi|) &= -\cos^2\phi\,\log(|\cos\phi|) - \sin^2\phi\,\log(|\sin\phi|) \\ M_2(|T_\phi\rangle\langle T_\phi|) &= -\log[(1+\cos^4\phi+\sin^4\phi)/2] \end{aligned} \tag{4.6.1}$$

Both quantities vanish for $\phi = 0, \pi/2$, while they have a maximum for $\phi = \pi/4$. We firstly initialize the system in the product state $|\psi_0\rangle = |T_\phi\rangle^{\otimes N}$, which is an MPS of bond dimension $\chi = 1$. Afterwards, we apply a random Clifford circuit $U_\mathcal{C}$ of depth N. In each layer, we randomly choose a sequence of one-qubits or two-qubits gates extracted from the generators [36]

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad S = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{pmatrix} \qquad \mathrm{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The final MPS $|\psi\rangle = U_\mathcal{C}|\psi_0\rangle$ has a larger bond dimension $\chi \gg 1$, whereas its nonstabilizerness is the same of $|\psi_0\rangle$, since this quantity is invariant under Clifford group. Thanks to the additivity of the SREs, the nonstabilizerness density $m_n(|\psi\rangle) = M_n(|\psi\rangle)/N$ is equivalent to the nonstabilizerness of a single T−state. We apply our sampling algorithm on $|\psi\rangle$, obtaining the estimation $\tilde{m}_n$. Results are shown in Fig. 4.3, for $n = 1, 2$ and size between

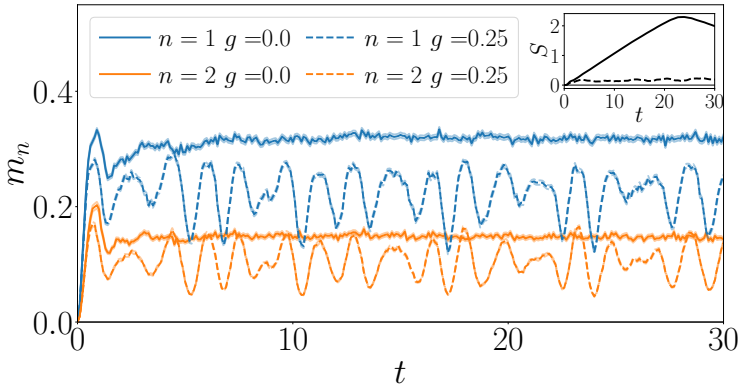

Figure 4.7: SRE density after a quantum quench in the transverse and longitudinal field Ising model ($N = 40$). The system is prepared in the ferromagnetic state $|+...+\rangle$ and quenched with parameters $h = 0.5, g = 0.0$ (solid line), $h = 0.5, g = 0.25$ (dotted line). The estimation is obtained with $\mathcal{N} = 10^3$ samples and pale lines represent the corresponding statistical uncertainty. Subplot: half-chain entanglement entropy.

$N = 10$ and $N = 70$. Notice that for $N = 70$, the bond dimension of $|\psi\rangle$ grows up to $\chi = 128$, depending on the particular arrangement of the Clifford layers. Values of χ of this order would be extremely challenging to target with previously known methods [174], whereas our approach takes only $\approx O(0.1)$ sec/sample on a single node simulation. Notice that the sampling can be easily parallelized, provided that the MPS is stored in multiple independent copies. All data points are in agreement with theoretical predictions within three error bars (see Fig. 4.3 a)). Moreover, a scaling of the statistical error $\delta\tilde{m}_n$ with N at fixed value of $\mathcal{N}$ suggests that the fluctuations do not grow significantly with the system size, even though in principle we might have expected them to increase exponentially with N for $n = 2$. Afterwards, we consider the quantum Ising model

$$H = -\sum_i \sigma_i^x \sigma_{i+1}^x - h \sum_i \sigma_i^z - g \sum_i \sigma_i^x \,. \tag{4.6.2}$$

For $g = 0$, this Hamiltonian can be mapped into a model of free fermions [179, 180], thus allowing the evaluation of the SREs in terms of $\sim 4^N$ determinants of matrices involving fermionic correlators [173]. In Fig. 4.4, we compare exact results for m_n ($n = 1, 2$) obtained in the fermionic representation with MPS estimations, for a system of size $N = 14$. For the MPS, we use Density Matrix Renormalization Group (DMRG) [27] ($\chi = 32$) to find the ground state. MPS data are in perfect agreement with the exact values, within small error bars.

We also use our algorithm to estimate the dynamics of the nonstabilizerness density during an out-of-equilibrium protocol. In particular, we prepare the system in the fully polarized state $|\psi(0)\rangle = |+\cdots+\rangle$, where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ is the eigenstate of σ^x with eigenvalue $+1$, and we consider the time-evolution generated by the Ising hamiltonian i.e. $|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle$. We set the transverse and longitudinal fields respectively to $h = 0.5$ and $g = 0, 0.25$. The latter value corresponds to a phase in which the system

is known to exhibit a dynamical confinement of the excitations [177, 181], whereas in the free case ($g = 0$) the quasiparticles give rise to a light cone spreading of correlations [182]. We use the Time Evolving Block Decimation (TEBD) to compute the time evolution of the post-quench MPS [27, 183], with bond-dimension up to $\chi = 128$. Results are shown in Fig. 4.7 for $N = 40$. For $g = 0$, the nonstabilizerness density seems to saturate rapidly to a stationary value (see Ref. [184]), although the half-chain entanglement entropy $S = -\mathrm{Tr}[\rho_{N/2} \log \rho_{N/2}]$, is still growing linearly with the time t as expected (see the subplot). In the confined phase $g = 0.25$, nonstabilizerness exhibits large and persistent oscillations around a slightly lower stationary value, whereas entanglement is strongly suppressed and approaches a low saturation value.

In Fig.4.8, we present the power spectra of the time-dependent order parameter $\langle \sigma^x(t) \rangle$ and of $m_n(t)$, which were obtained using Fourier transformation. The dominant frequencies of the two signals (i.e. peaks in the power spectrum) match with good accuracy. It is well-known that the time dependence of the order parameter in the longitudinal Ising model is strongly affected by the value of the longitudinal field g. As soon as $g > 0$, the model exhibit confinement, meaning that the longitudinal field confines a pair of domain wall excitations into a bound state dubbed meson. This physical phenomenon results in persistent oscillations of $\langle \sigma^x(t) \rangle$, whose characteristic frequencies are determined by the masses of the mesons (and their differences). Our plots in Fig. 4.8 provide evidences that also the nonstabilizerness of the time evolved state is strongly affected by the outlined physical picture.

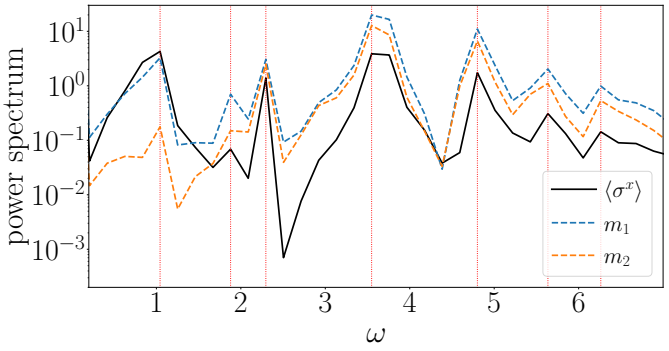

Figure 4.8: The power spectrum of the signals $m_n(t)$ and $\langle \sigma^x(t) \rangle$ computed as the squared modulus of the Fourier transform. The latter is evaluated by means of the Fast Fourier Transform (FFT) algorithm.

Finally, we consider the dynamics of the SRE for an operator O initially prepared as a localized Pauli matrix and undergoing temporal evolution in the Heisenberg picture, expressed as $O(t) = e^{iHt} O(0) e^{-iHt}$. In our case, we set $O(0)$ equal to the Pauli matrix σ^x localized in the central site of a chain of length $N = 51$. The time evolution was evaluated by means of TEBD applied to the MPO $O(t)$ reshaped as MPS (see Fig. 4.6). The maximum bond dimension used was $\chi = 256$. Results are shown in Fig.4.9 for $h = 0.5, g = 0.0$ (left) and $h = 0.5, g = 0.25$ (right). Note that we shifted m_n by $\log 2$ in order to eliminate the additive constant introduced in Eq. 4.1.3, which is unnecessary in this case. This results in $m_n + \log 2 = 0$ at $t =$

0, since the probability distribution associated to O has initially support on only a single Pauli string. Results presented in Fig.4.9 are preliminary and currently undergoing analysis. We are running multiple simulations to determine if assessing the operatorial SREs can characterize the dynamics in the Heisenberg picture. It would be very interesting if SRE behave differently for integrable or chaotic systems.

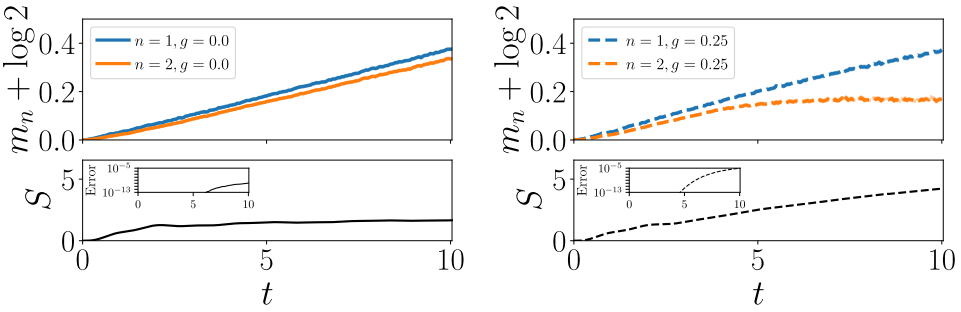

Figure 4.9: Operatorial SRE density after a quantum quench in the transverse and longitudinal field Ising model ($N = 51$). The operator O is prepared as a σ^x Pauli matrix localized in the central site of the chain, i.e. $O(0) = \sigma^x_0$, and quenched with parameters $h = 0.5, g = 0.0$ (left, solid line), $h = 0.5, g = 0.25$ (right, dotted line). Afterwards, we estimate the operatorial SRE density $m_n - \log 2$ of the time evolved operator $O(t)$. The estimation is obtained with $\mathcal{N} = 10^3$ samples and pale lines represent the corresponding statistical uncertainty. Lower plot: half-chain entanglement entropy. Subplot: the truncation error of the Schmidt coefficients.

4.7 CONCLUSIONS

We have shown that a relatively new measure of quantum nonstabilizerness, the Stabilizer Rényi Entropies [172], can be estimated efficiently in the MPS framework via a perfect sampling of Pauli strings operators. Our estimation neither suffers from the exponential growth of the size of the many-body Hilbert space, nor shows an unfavorable scaling with the MPS bond dimension. As a matter of fact, we are able to consider either equilibrium or non-equilibrium wave-functions with MPS bond-dimension up to values that were out of reach by any of the previously proposed methods for evaluating the nonstabilizerness. Specifically, we applied our method to evaluate the amount of nonstabilizerness generated after a quench in the quantum Ising chain, and its sensitivity to the presence of confinement of excitations. Although we mainly focused on pure MPS, our algorithm can be easily adapted to non-pure states obtained from an MPS tracing out a subsystem consisting of the first or last qubits.

Our approach pave the way to novel numerical studies of the nonstabilizerness, possibly providing new characterizations of the quantum phases of matter, in and out-of-equilibrium. In addition, our new Pauli sampling

technique for the MPS can be used to address crucial problems in quantum many-body theory, as for instance the operator scrambling.

Finally, we mention that an estimation of the SREs analogous to what we discussed is experimentally achievable in platforms enabling the preparation of duplicate states $|\psi\rangle \otimes |\psi\rangle$ and joint Bell basis measurements.

Appendices

A VARIATIONAL MONTE CARLO

A.1 THE VARIATIONAL PRINCIPLE

Let us suppose to have a family of quantum states parameterized by a set of real parameters $\boldsymbol{\theta}$

$$|\psi\rangle = |\psi(\boldsymbol{\theta})\rangle \ . \tag{A.1.1}$$

The aim of Variational MonteCarlo (VMC) is to minimize the energy functional

$$\mathcal{E} = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} \tag{A.1.2}$$

on this manifold of states, and finding a (hopefully good) approximation of the exact ground state of a many body Hamiltonian $\hat{H}$. Let us assume that $\{|\mathbf{s}\rangle\}$ is a basis of our Hilbert space. We can re-write $\mathcal{E}$ as follows

$$\mathcal{E} = \frac{\sum_{\mathbf{s},\mathbf{s}'} \psi^*_{\mathbf{s}} H_{\mathbf{s}\mathbf{s}'} \psi_{\mathbf{s}'}}{\sum_{\mathbf{s}} |\psi_{\mathbf{s}}|^2} = \frac{\sum_{\mathbf{s}} |\psi_{\mathbf{s}}|^2 \sum_{\mathbf{s}'} H_{\mathbf{s}\mathbf{s}'} \frac{\psi_{\mathbf{s}'}}{\psi_{\mathbf{s}}}}{\sum_{\mathbf{s}} |\psi_{\mathbf{s}}|^2} \ . \tag{A.1.3}$$

Let us now define the so-called *local energy*

$$E_{\text{loc}}(\mathbf{s}) = \sum_{\mathbf{s}'} H_{\mathbf{s}\mathbf{s}'} \frac{\psi_{\mathbf{s}'}}{\psi_{\mathbf{s}}} \ . \tag{A.1.4}$$

which is a function of the system configuration $\mathbf{s}$. We get

$$\mathcal{E} = \sum_{\mathbf{s}} p(\mathbf{s}) E_{\text{loc}}(\mathbf{s}) = \langle E_{\text{loc}} \rangle \qquad p(\mathbf{s}) = \frac{|\psi_{\mathbf{s}}|^2}{\sum_{\mathbf{s}'} |\psi_{\mathbf{s}'}|^2} \ . \tag{A.1.5}$$

Therefore, $\mathcal{E}$ is nothing but the expectation value of E_{loc} over the distribution$p(\mathbf{s})$. As we will show in the following, it is possible to define a stochastic algorithm (i.e. a Markov chain), which generates a sequence of configurations $\{\mathbf{s}_k\}$ distributed according to our p.d.f. $p(\mathbf{s})$. In this case we can estimate $\mathcal{E}$ by means of the sample mean

$$\overline{E_{\text{loc}}} = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} E_{\text{loc}}(\mathbf{s}_k) \ . \tag{A.1.6}$$

Since typically H is a highly sparse matrix, the sum $\sum_{\mathbf{s}'}$ in Equation A.1.4 is extended only to a tiny portion of the Hilbert space, i.e. to the states for which $H_{\mathbf{s}\mathbf{s}'} \neq 0$. For a given $\mathbf{s}$, the number of $\mathbf{s}'$ such that $H_{\mathbf{s}\mathbf{s}'} \neq 0$ is typically

polynomial in the system size, so the computation of E_{loc} is straightforward (in contrast to the computations of $\mathcal{E}$ and $\sum_{\mathbf{s}} |\psi_{\mathbf{s}}|^2$ which requires to sum over an exponentially large number of terms). The same arguments can be used for the great majority of physical observables, i.e. of the hermitian operators O. The legitimacy of using the estimator A.1.6 is strengthened by observing that the energy variance

$$\delta\mathcal{E}^2 = \frac{\langle\psi|\left(\hat{H}-\mathcal{E}\right)^2|\psi\rangle}{\langle\psi|\psi\rangle} \tag{A.1.7}$$

can be computed as the variance of E_{loc}. Indeed

$$\begin{aligned}\delta\mathcal{E}^2 &= \sum_{\mathbf{s}} p(\mathbf{s}) \sum_{\mathbf{s}'} \left((H^2)_{\mathbf{s}\mathbf{s}'} - 2H_{\mathbf{s}\mathbf{s}'}\mathcal{E} + \mathcal{E}^2 D_{\mathbf{s}\mathbf{s}'}\right)\frac{\psi_{\mathbf{s}'}}{\psi_{\mathbf{s}}} = \\ &= \sum_{\mathbf{s}} p(\mathbf{s}) \sum_{\mathbf{s}'} \Big(\sum_{\mathbf{s}''} H_{\mathbf{s}\mathbf{s}''}H_{\mathbf{s}''\mathbf{s}'} - \mathcal{E}^2 D_{\mathbf{s}\mathbf{s}'}\Big)\frac{\psi_{\mathbf{s}'}}{\psi_{\mathbf{s}}} \ .\end{aligned} \tag{A.1.8}$$

The first term gives

$$\begin{aligned}\sum_{\mathbf{s},\mathbf{s}',\mathbf{s}''} H_{\mathbf{s}\mathbf{s}''}H_{\mathbf{s}''\mathbf{s}'}\psi_{\mathbf{s}'}\psi^*_{\mathbf{s}} &= \sum_{\mathbf{s},\mathbf{s}',\mathbf{s}''} p(\mathbf{s}'')(H_{\mathbf{s}''\mathbf{s}})^* H_{\mathbf{s}''\mathbf{s}'}\frac{\psi_{\mathbf{s}'}\psi^*_{\mathbf{s}}}{\psi_{\boldsymbol{\sigma}''}\psi^*_{\boldsymbol{\sigma}''}} = \\ &= \sum_{\mathbf{s}''} p(\mathbf{s}'')|E_{loc}(\mathbf{s}'')|^2 = \langle|E_{loc}|^2\rangle \ .\end{aligned} \tag{A.1.9}$$

Thus

$$\delta\mathcal{E}^2 = \langle|E_{loc}|^2\rangle - |\langle E_{loc}\rangle|^2 \ . \tag{A.1.10}$$

Since any eigenstates of $\hat{H}$ is such that $\delta\mathcal{E}^2 = 0$, the variance of the stocastic variable E_{loc} will decrease by getting closer to the ground state of H. Thus, the approximation in equation A.1.6 will progressively become better.

A.2 SAMPLING METHODS

We want to properly define a Markov chain of configurations $\{\mathbf{s}_k\}$, $k = 1, 2, 3...$ in order to sample our probability distribution $p(\mathbf{s})$. A Markov process is completely specified by the transition matrix $\mathcal{T}_{\mathbf{s}\to\mathbf{s}'} = \mathcal{T}_{\mathbf{s}\mathbf{s}'}$. Obviously the matrix $\mathcal{T}$ must obey to

$$1 = \sum_{\mathbf{s}'} \mathcal{T}_{\mathbf{s}\to\mathbf{s}'} = \sum_{\mathbf{s}'} \mathcal{T}_{\mathbf{s}\mathbf{s}'} \ . \tag{A.2.1}$$

We have to require the detailed balance relation

$$p(\mathbf{s})\mathcal{T}_{\mathbf{s}\to\mathbf{s}'} = p(\mathbf{s}')\mathcal{T}_{\mathbf{s}'\to\mathbf{s}} \tag{A.2.2}$$

to have p as a stationary distribution for our Markov chain. Indeed, this equation implies

$$\sum_{\mathbf{s}} p(\mathbf{s})\mathcal{T}_{\mathbf{s}\mathbf{s}'} = p(\mathbf{s}') \sum_{\mathbf{s}} \mathcal{T}_{\mathbf{s}'\to\mathbf{s}} = p(\mathbf{s}') \tag{A.2.3}$$

which is the stationarity condition. The idea of the Metropolis-Hastings algorithm is to split the elements of $\mathcal{T}$ as a product of two terms, i.e.

$$\mathcal{T}_{\mathbf{s}\mathbf{s}'} = T_{\mathbf{s}\mathbf{s}'} A_{\mathbf{s}\mathbf{s}'} \tag{A.2.4}$$

where T represents a *local transition* and A represents the *acceptance*. Now, we observe that the detailed balance relation can be rewritten as

$$\frac{A_{\mathbf{s}'\to\mathbf{s}}}{A_{\mathbf{s}\to\mathbf{s}'}} = \frac{p(\mathbf{s})}{p(\mathbf{s}')} \frac{T_{\mathbf{s}\to\mathbf{s}'}}{T_{\mathbf{s}'\to\mathbf{s}}} \tag{A.2.5}$$

Interestingly, this condition is automatically fulfilled by

$$A_{\mathbf{s}\to\mathbf{s}'} = \min\left(1, \frac{p(\mathbf{s}')}{p(\mathbf{s})} \frac{T_{\mathbf{s}'\to\mathbf{s}}}{T_{\mathbf{s}\to\mathbf{s}'}}\right). \tag{A.2.6}$$

The Metropolis–Hastings algorithm then involves generating new configurations $\mathbf{s}'$ according to a certain rule for updates and accepting or rejecting them with a probability given by $A_{\mathbf{s}\to\mathbf{s}'}$. Often, the update is performed in a deterministic way, meaning that $T_{\mathbf{s}\mathbf{s}'} = 1$ for a certain configuration $\mathbf{s}'$. In this case

$$A_{\mathbf{s}\to\mathbf{s}'} = \min\left(1, \frac{p(\mathbf{s}')}{p(\mathbf{s})}\right). \tag{A.2.7}$$

A.3 STOCHASTIC OPTIMIZATION

We want to optimize the variational energy $\mathcal{E}(\boldsymbol{\theta})$. The simplest method is the Stocastic Gradient Descent (SGD), in which we update the parameters according to

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \eta\, \partial_{\boldsymbol{\theta}} \mathcal{E}\ . \tag{A.3.1}$$

Let us compute the gradient of $\mathcal{E}$ with respect to the real parameters $\boldsymbol{\theta}$. We have

$$\partial_k \mathcal{E} = \frac{\langle\psi|\hat{H}(\partial_k|\psi\rangle) + (\partial_k \langle\psi|)\hat{H}|\psi\rangle)}{\langle\psi|\psi\rangle} - \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle} \frac{\partial_k(\langle\psi|\psi\rangle)}{\langle\psi|\psi\rangle} \qquad \partial_k = \frac{\partial}{\partial\theta_k}. \tag{A.3.2}$$

Thus, by defining the log-derivative of the wave-function

$$D_k(\mathbf{s}) = \frac{\partial_k(\langle\mathbf{s}|\psi\rangle)}{\langle\mathbf{s}|\psi\rangle} \tag{A.3.3}$$

we find

$$\begin{aligned} \partial_k \mathcal{E} &= \langle E_{\mathrm{loc}} D_k^* \rangle + \langle E_{\mathrm{loc}}^* D_k \rangle - \langle E_{\mathrm{loc}} \rangle \langle D_k^* \rangle - \langle E_{\mathrm{loc}} \rangle \langle D_k \rangle = \\ &= \langle\ 2\,\mathrm{Re}\left(D_k^*\left(E_{\mathrm{loc}} - \langle E_{\mathrm{loc}} \rangle\right)\right)\ \rangle = \langle G_k \rangle\ . \end{aligned} \tag{A.3.4}$$

As usual, we will find an approximation of $\partial_k \mathcal{E}$ by computing a sample mean of the gradient estimator G_k. This method add some noise to the gradient descent.

One major problem of SGD method is that it totally ignores the geometry of the states manifold. This fact can be seen from equation A.3.1, in which

we have a sum of two objects of different types: a contro-variant coordinate vector with a co-variant coordinate gradient. Clearly, adding two such different terms leads to an object which does not transforme properly under change of coordinates (indeed, the ordinary gradient descent is not invariant under coordinate transformations). We can make invariant the alghoritm by trasforming the gradient to a contro-variant object. If we have a local metric g_{ij}, this can be done by contracting with its inverse matrix $g^{ij} = (g^{-1})_{ij}$. Our new method is just

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \eta\, g^{-1} \partial_{\boldsymbol{\theta}} \mathcal{E} \ , \tag{A.3.5}$$

This is called Natural Gradient Descent or also Stochastic Reconfiguration. Now, we have to establish what is a good metric for our state manifold. Let us begin by introducing a notion of distance between the states of the Hilbert space, for example the Fubini-Study distance which is defined as follows

$$d\big(\,|\psi\rangle\,,|\phi\rangle\,\big) = \arccos \sqrt{\frac{\langle\psi|\phi\rangle\,\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\ \langle\phi|\phi\rangle}} \ . \tag{A.3.6}$$

It is easy to show that infinitesimal distances are given by

$$ds^2 = \Big(d\big(\,|\psi\rangle + |\delta\psi\rangle\,,|\psi\rangle\,\big)\Big)^2 \simeq \frac{\langle\delta\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle} - \frac{\langle\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle}\frac{\langle\delta\psi|\psi\rangle}{\langle\psi|\psi\rangle} \ , \tag{A.3.7}$$

which defines the following quadratic form

$$S_{ij} = \langle D_i D_j^* \rangle - \langle D_i D_j^* \rangle \qquad ds^2 = S_{ij}\, d\theta^i d\theta^j \ . \tag{A.3.8}$$

This is the local metric of our manifold. The SR method can also be viewed as an approximate imaginary-time evolution in the tangent space of the wave function.

S is computed in presence of statistical error, therefore it may be an ill-conditioned matrix with very small eigenvalues, which affect the accuracy of its inversion. This issue can be solved by regularizing the elements of S on the diagonal, with the following regularization $S_{ii} \to S_{ii}(1+\epsilon)$.

B TECHNICALITIES ON MPS SIMULATIONS FOR QUANTUM ANNEALING

In this Appendix, we present some technical results related to the MPS simulations of Quantum Annealing discussed in Chapter 3.

B.1 DMRG RESULTS

As outlined in Sec. 3.1, our MPS framework allows to represent any target Hamiltonian $\hat{H}_z$ in the form of Eq. 3.1.1 as an MPO. Thus, one might consider using DMRG to find an MPS representation of a ground state, which is essentially a superposition of classical solutions. To test this method, we consider again the perceptron model (Eq. 3.1.4) with $N = 21$, $N_\xi = 17$ and the same patterns $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^{N_\xi}$ considered in the main text. We employ the standard one-site DMRG algorithm [27], with different bond dimensions and starting from different randomly-generated MPS. For each bond dimension value $\chi = 1, 10, 20$ we run 5 different simulations. As an illustrative example, we report the DMRG data for the first training set in the following table .

χ	Run	$\langle \hat{H}_z \rangle$	χ	Run	$\langle \hat{H}_z \rangle$	χ	Run	$\langle \hat{H}_z \rangle$
1	1	0.21822	10	1	0.65465	20	1	$-6.43 \cdot 10^{-14}$
	2	0.65465		2	$-3.88 \cdot 10^{-14}$		3	$-6.43 \cdot 10^{-14}$
	3	0.21822		3	$1.01 \cdot 10^{-13}$		2	0.21822
	4	$-4.49 \cdot 10^{-14}$		4	$1.80 \cdot 10^{-13}$		4	0.21822
	5	$-4.48 \cdot 10^{-14}$		5	$-3.60 \cdot 10^{-14}$		5	$-6.43 \cdot 10^{-14}$

These results show that DMRG convergence strongly depends on the initial guess, in our case a random MPS. Indeed, since DMRG relies on a local optimization of the MPS (see Sec. 1.6), we expect it to correctly converge to a classical ground state only in some favorable cases where the initial MPS state is already quite close to the absolute minimum, within the glassy energy landscape of $\hat{H}_z$. In other cases, DMRG will converge to an excited classical state. We also measure the standard deviation of the target Hamiltonian $\sigma_{\hat{H}_z}$, as defined in Eq. 3.2.2, always finding $\sigma_{\hat{H}_z} \lesssim 10^{-8}$, confirming that the final MPS is very close to be an exact eigenstate of $\hat{H}_z$. Furthermore, we measured the overlap between the MPS resulting from DMRG with the enumerated classical ground states of $\hat{H}_z$. When the minimization problem is solved properly (i.e. $\langle \hat{H}_z \rangle \approx 0$) the resulting MPS has overlap 1 (at machine precision) *only with a single solution*. This fact represents a major difference in

comparison with the Quantum Annealing approach, where the final quantum state always has a finite overlap with many different classical solutions (delocalization).

B.2 MPS CONVERGENCE WITH BOND DIMENSION

In this section, we provide extra results and details about MPS simulations. We restrict our numerics to an instance of the perceptron model (Eq. 3.1.4) for $N = 21$ spins and $N_\xi = 17$ patterns, but the same qualitative results are observed in general.

Throughout the Chapter, we set the bond dimension of MPS simulations to $\chi = 10$, since the validity of our methods and the main results are expected to be robust by varying the bond dimension in a reasonable range. More precisely, the bond dimension should be large enough to encode the entanglement produced by dQA in the regime of small time-step ($\delta t \ll 1$), allowing our MPS simulation to closely approximate the exact digitized dynamics. On the contrary, the large time-step regime ($\delta t = \mathcal{O}(1)$) is dominated by Trotter errors leading to high entanglement production, and we showed that MPS simulations largely deviate from ED (see Fig. 3.6 and Sec. 3.2.3 for details). Indeed, we argued that MPS simulations closely mimic a digitized dynamics *without* Trotterization, which is an excellent approximation of the continuous time-ordered dynamics of an ideal QA. Nevertheless, even in this regime, one would expect that by increasing the bond dimension, our MPS approximation of the digitized QA dynamics *with* Trotterization would converge to its exact version, implying a degradation of performance.

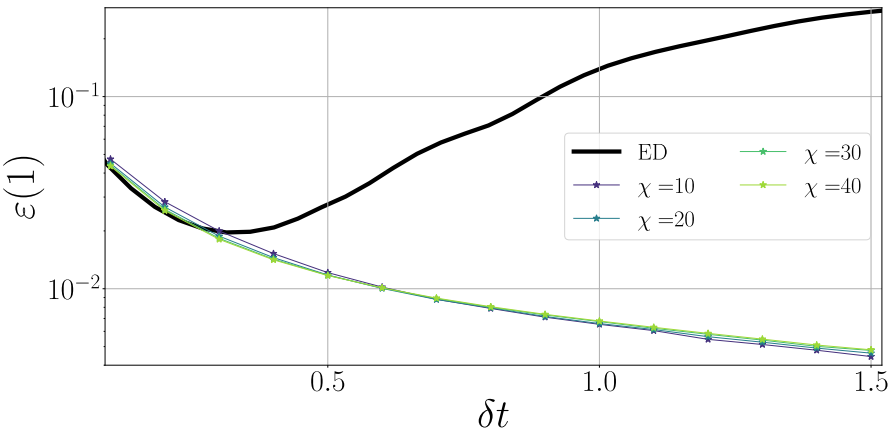

Figure B.1: Binary perceptron. Residual energy density $\varepsilon(1)$ as a function of time step length δt ($N = 21$, $N_\xi = 17$ and $P = 100$). MPS results for increasing values of the bond dimension χ are compared with ED results.

Here, we investigate this aspect, by performing a series of simulations for increasing values of bond dimension χ, having fixed the total number of annealing steps to $P = 100$. Fig. B.1 shows the final energy density $\varepsilon(1)$, for different values of δt, with MPS data compared with ED data (black line). In the small δt regime, as expected, MPS data converge to ED by increasing χ,

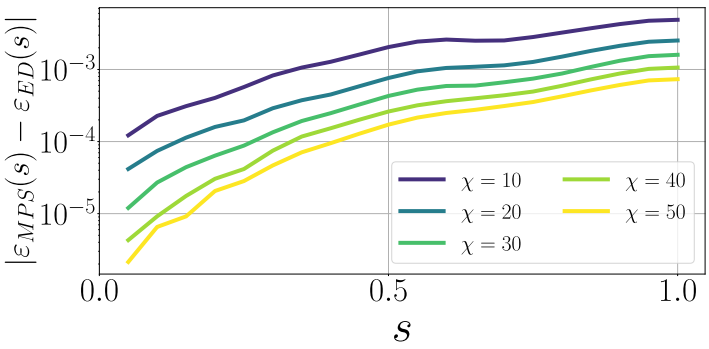

Figure B.2: Binary perceptron. The energy density difference $|\varepsilon_{MPS}(s) - \varepsilon_{ED}(s)|$ between MPS and ED, as a function of the annealing parameter s. We set $N = 21$, $N_\xi = 17$, $P = 100$, $\delta t = 0.1$ and we explored different bond dimensions χ.

confirming that our Algorithm 1 accurately simulates the exact Trotterized dQA dynamics in this regime. This is also visible in Fig. B.2, where we plot the energy density difference $|\varepsilon_{MPS}(s) - \varepsilon_{ED}(s)|$ between MPS data and ED data, in the small time-step regime ($\delta t = 0.1$) for $P = 100$; by increasing χ, this difference monotonically decreases to 0, for any value of s. On the contrary, in the large time-step regime $\delta t \sim \mathcal{O}(1)$ this interpretation seemingly breaks down, as in Fig. B.1 the values of $\varepsilon(1)$ are relatively stable with the increase of χ, very far from ED results. We argue that the reason for this apparent inconsistency is that, in order to correctly encode the large entanglement entropy due to (unwanted) spurious Trotter terms (see Sec. 3.2.3), one would actually need *much larger* values of χ. In the following, we show this fact more quantitatively.

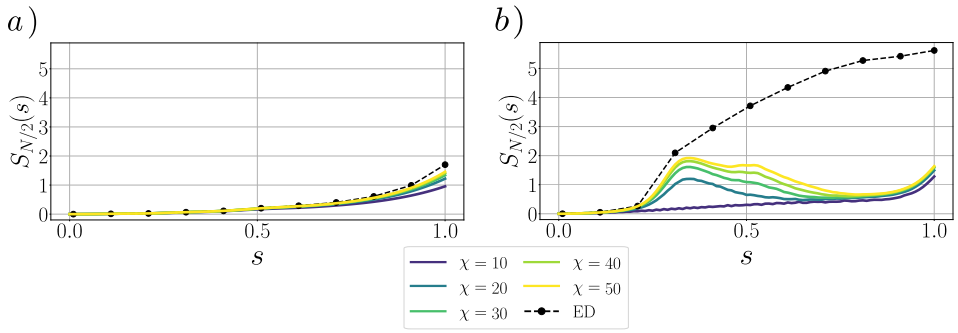

Figure B.3: Binary perceptron. Half system entanglement entropy $S_{N/2}$ as a function of the annealing parameter s during dQA. Here, δt is fixed to 0.1 in panel a) and 1.0 in panel b). ED data are compared with MPS data for increasing values of the bond dimension.

In Fig. B.3, we plot the entanglement entropy at half system $S_{N/2}(s)$, defined as in Eq. 3.2.5, which can be easily evaluated in the MPS framework [27]. Once again, we fix two reference values of δt in the two regimes: $\delta t = 0.1$ (left) and $\delta t = 1.0$ (right). In the first case, MPS data show a clear convergence towards ED data upon increasing the bond dimension χ. On the opposite, in the second case, MPS and ED largely deviate. In partic-

ular, the amount of entanglement produced by exact dQA reaches the value $S^{ED}_{N/2}(1) \simeq 5.62$: since the entanglement encoded by MPS is bounded by $\log\chi$, in order to encode the final steps of exact dQA in this regime, one would need (at least) $\chi \sim e^{S^{ED}_{N/2}(1)} \simeq e^{5.62} \approx 276$, far outside the range of values analyzed above.

Finally, we provide extra results on the MPS compression accuracy (see Sec. 3.1.3). Let us address the Hilbert space distance between the compressed and uncompressed MPS at each step of the Algorithm 1, which is the same quantity being iteratively minimized in Eq. 1.4.9. This distance is written as

$$d^2_\mu(s) = \| \, |\psi^\mu(s)\rangle - |\tilde{\psi}^\mu(s)\rangle \, \|^2 \qquad \mu = 1, \dots N_\xi \, ,$$

$|\psi^\mu(s)\rangle$ being the uncompressed MPS after the application of the unitary operator $\hat{U}^\mu_z(\gamma_p)$ at a generic annealing step $s = s_p$ (with $p = 1 \cdots P$), and $|\tilde{\psi}^\mu(s)\rangle$ the MPS resulting from the compression. Thus, $d^2_\mu(s)$ is a measure of the compression accuracy. In Fig. B.4, we plot $d^2_\mu(s)$ vs the annealing parameter $s \in [0, 1]$, with a total number of annealing steps fixed to $P = 1000$. Different color shades refer to different patterns $\mu = 1 \dots N_\xi$ (the operators $\hat{U}^\mu_z(\gamma_p)$ are applied sequentially, as sketched in Fig. 3.3). Two values of δt are considered: $\delta t = 0.1$ (blue shades) and $\delta t = 1.0$ (purple shades), corresponding to the two usual regimes. Notice that $d^2_\mu(s)$ takes lower values, by some order of magnitudes, for the first case (with the exception of the last part of the annealing $s \simeq 1$). The reason can be traced back to Fig. 3.10: in the small time-step regime, exact dQA dynamics produces low-entangled states, thus the projection into the MPS manifold is easily performed by the compression algorithm (reaching high accuracy, i.e. low values of d^2_μ); for large time steps, on the contrary, the dynamics produces large amounts of entanglement, therefore the compression becomes rough (the projected state $|\tilde{\psi}^\mu(s)\rangle$ is located at larger distance from the uncompressed state $|\psi^\mu(s)\rangle$).

In the last stages of the annealing, however, the instantaneous state has projection close to one on the zero-energy ground state eigenspace, thus the unitary operators $\hat{U}^\mu_z(\gamma_p)$ act almost trivially (if the state is exactly in the ground state subspace, $\hat{U}^\mu_z(\gamma_p)$ equals the identity). This results into a sharp decrease in d^2_μ for s close to 1. In the case $\delta t = 1.0$, the final values of $d^2_\mu(s)$ are smaller than for $\delta t = 0.1$, since the final state has larger overlap on the ground state eigenspace of $\hat{H}_z$, as shown in Fig. B.2.

B.3 TIME DISCRETIZATION VERSUS TROTTERIZATION

The scope of this section is to provide supplementary results on the comparison between dQA *with* and *without* Trotterization (see Eq. 3.2.3 for a precise definition). In the first place, we test numerically a rather surprising fact, i.e. the robustness to time discretization discussed in Sec. 3.2.3: the error introduced by approximating the exact continuous time-ordered evolution with P discrete time steps of length δt is rather small, even for large $\delta t \sim \mathcal{O}(1)$. To show this, we simulate dQA *without* Trotterization (setting $P = 100$ and $\delta t = 0.1, 1.0$), and we evaluate the fidelity $\mathcal{F}$ between the evolved state $|\psi(s)\rangle$

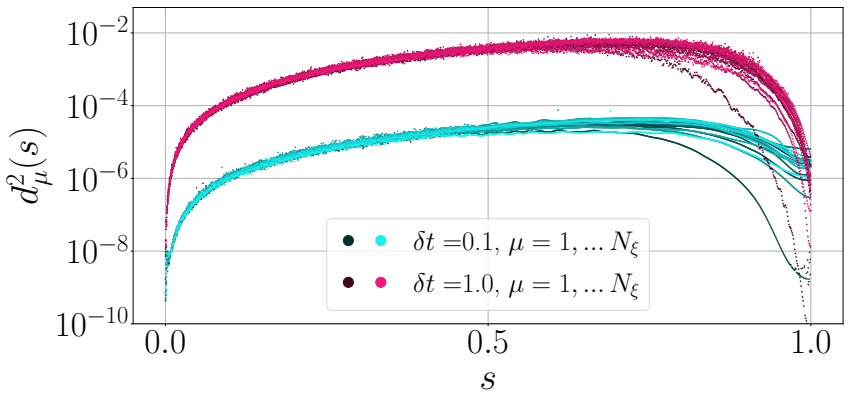

Figure B.4: Binary perceptron. Hilbert space distance $d^2_\mu(s)$ between the compressed and uncompressed MPS at each step of Algorithm 1. We set $N = 21$, $N_\xi = 17$, $P = 1000$. Two different time steps are considered: $\delta t = 0.1$ (blue shades) and $\delta t = 1.0$ (purple shades).

and a reference state, ideally representing the exact time-ordered evolution. This reference state, which we dub $|\psi'(s)\rangle$, is actually obtained by running a new approximate simulation with the same values of the total annealing time ($\tau = P\delta t = 10, 100$), now with a time discretization that is 50 times denser (i.e. $P' = 50\,P = 5000$, $\delta t' = \delta t/50 = 0.002, 0.02$). In practice, the simulation is performed with ED, for a perceptron model of size $N = 18$. Data are reported in Fig. B.5, averaged over 5 different realizations of the random patterns. Remarkably, fidelity values are very close to 1 along the whole dynamics ($1 - \mathcal{F} < 10^{-3}$), proving that time discretization injects negligible errors in the dynamics, even for time steps as large as $\delta t = 1.0$.

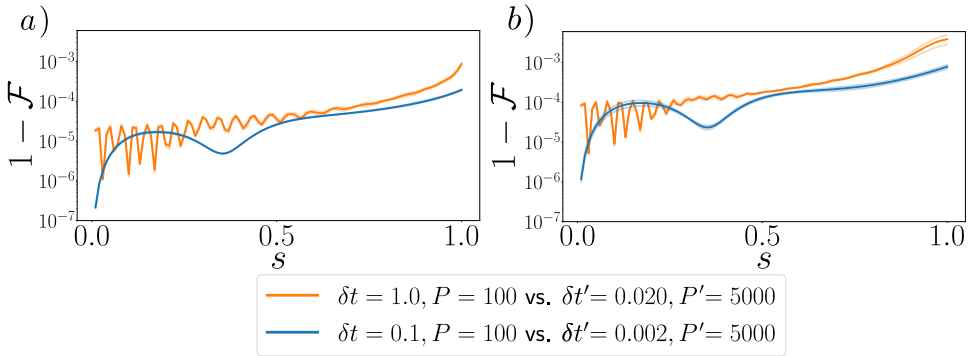

Figure B.5: Binary perceptron. One minus the fidelity $\mathcal{F}(s) = |\langle\psi(s)|\psi'(s)\rangle|^2$ between the states $|\psi(s)\rangle$ and $|\psi'(s)\rangle$ both obtained with dQA *without* Trotterization, setting respectively $P = 100$, $\delta t = 0.1, 1.0$ and $P' = 5000$, $\delta t' = 0.002, 0.02$. Notice that $\delta t P = \delta t' P' = 10, 100$, meaning that the simulations are different approximations of the same continuous time evolution (with total annealing time $\tau = 10, 100$). Data are averaged over 5 realizations of the random patterns $\{\boldsymbol{\xi}\}_{\mu=1}^{N_\xi}$. We set $N_\xi = 3\,(\alpha \simeq 0.17)$ in panel a) and $N_\xi = 14\,(\alpha \simeq 0.78)$ in panel b).

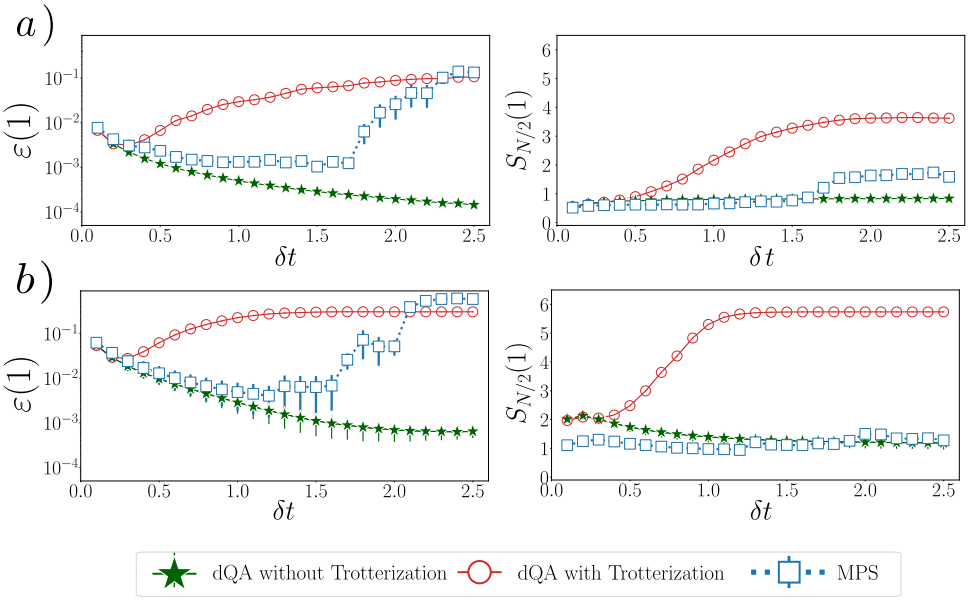

Figure B.6: Binary perceptron. Final energy density $\varepsilon(1)$ (left) and half-system entanglement entropy $S_{N/2}(1)$ of the final state (right), for $N = 18$ and $N_\xi = 3\,(\alpha \simeq 0.17)$ in panel a) and $N_\xi = 14\,(\alpha \simeq 0.78)$ in panel b). We set $P = 100$ and we perform dQA *with* and *without* Trotterization via ED, compared with MPS results. Data are averaged over 5 realizations of the random patterns, and the resulting standard deviations are plotted as error bars (for dQA *with* Trotterization these are smaller than the marker size).

An additional comparison of the two dQA methods with our MPS implementation is reported in Fig. B.6. Here, for the same model and data parameters specified above, we plot the final energy density $\varepsilon(1)$ and the half-system entanglement entropy $S_{N/2}(1)$ of the final state, for different values δt. These plots are similar to those shown in Sec. 3.2, but they also include dQA *without* Trotterization. Interestingly, residual energy data show that Trotterization spoils the effectiveness of dQA at large time steps, whereas the repeated projection on the MPS manifold can (partially) restore it. The entanglement plots also confirm that Trotterization results in a strongly enhanced entanglement production, if compared to dQA *without* Trotterization; the MPS simulation significantly reduces the entanglement of the final annealed state, as already shown in Sec. 3.2.3.

Concerning the entanglement entropy of the final annealed state, let us also notice a quite peculiar aspect of our results on QA. The (average) number of solutions N_{sol} is expected to decrease monotonically with the number of patterns N_ξ, interpolating the values $\sim 2^{N-1}$ for $N_\xi = 1$ and ~ 0 for $N_\xi = N\alpha_c$ (rigorously this holds for $N \to \infty$) [185]. As a consequence, one might expect the final entanglement entropy to be larger for a smaller value of N_ξ (i.e. smaller α, for fixed N), since the final QA wave function could acquire a non-vanishing overlap with many of these solutions. However,

Fig. B.6 shows that, for dQA *without* Trotterization[1], $S_{N/2}(1)$ is smaller for $\alpha = 0.17$ than for $\alpha = 0.78$ (for all values of δt).

Moreover, the half system entanglement entropy of a linear superposition of N_{sol} classical states is upper bounded by $\log(N_{sol})$. For $\alpha \ll \alpha_c$, we expect $N_{sol} \sim \mathcal{O}\left(2^{N-N_\xi}\right)$, and thus $\log(N_{sol}) \sim (N - N_\xi)\log 2$, which finally gives $\log(N_{sol}) \sim 15 \log 2$ for panel a). This is actually larger than the theoretical upper bound $N/2 \cdot \log 2 = 9 \log 2$; thus $S_{N/2}(1)$ could saturate the upper bound, but it is always observed to be much smaller. Consequently, we argue that the QA dynamics, for the problem in exam, yields a final wave function with non-vanishing overlap only with few of the many possible classical solutions. This fact is particularly relevant for our MPS approach: if QA resulted in highly entangled final states for small α, then MPS would not be able to accurately follow its dynamics in this regime (i.e., looking back at Fig. 3.1, $|\psi_\tau\rangle$ would not belong to the MPS manifold M_χ).

1 We assume again that dQA *without* Trotterization is a good approximation of the exact continuous QA. This is justified by the previous considerations.

C INTRODUCTION TO THE STABILIZER FORMALISM

Remark on the notation. Unlike the rest of the work, in this Appendix we will use the symbols $\{X, Y, Z\}$ to denote the Pauli matrices. Furthermore, we are going to use the symbol g (or h) to denote a generic Pauli string, whereas previously we used the symbol $\boldsymbol{\sigma}$. The reason for these choices is to adhere to the notation more commonly used in the literature. Finally, we will omit the symbol $\hat{}$ for operators.

C.1 PAULI GROUP AND STABILIZER GROUPS

Definition 1. The Pauli group $\mathcal{P}_1$ is the 16−elements group consisting of the following 2×2 matrices

$$\mathcal{P}_1 = \{\pm\mathbb{1}, \pm i\mathbb{1}, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}.$$

The Pauli group $\mathcal{P}_N$ is the group obtained by taking the tensor product of any N elements of $\mathcal{P}_1$, i.e.

$$\mathcal{P}_N = \{\pm\mathbb{1}, \pm i\mathbb{1}, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}^{\otimes N}.$$

Observation 1.

a) The elements of $\mathcal{P}_N$ are hermitian or anti-hermitian

b) The elements of $\mathcal{P}_N$ have eigenvalues ± 1 (if hermitian) or $\pm i$ (if anti-hermitian)

c) If $g \in \mathcal{P}_N$, then $g^2 = \mathbb{1}$ (if g is hermitian) or $g^2 = -\mathbb{1}$ (if g is anti-hermitian)

d) The elements of $\mathcal{P}_N$ are unitary

e) If $g, h \in \mathcal{P}_N$, then or they commute $[g, h] = 0$ or they anti-commute $\{g, h\} = 0$

f) With the exception of $\pm\mathbb{1}$, $\pm i\mathbb{1}$, the elements of $\mathcal{P}_N$ are traceless

Definition 2. We define *stabilizer group* any subgroup of $\mathcal{P}_N$ that does not contain the element $-\mathbb{1}$. Given a stabilizer group $\mathcal{S}$, we define $V_\mathcal{S}$ as the set of $N-$qubits states $|\psi\rangle$ such that $g|\psi\rangle = |\psi\rangle, \forall g \in \mathcal{S}$ and $\forall |\psi\rangle \in V_\mathcal{S}$. $V_\mathcal{S}$ is by construction a subspace. It is called *stabilizer subspace*. If it is one-dimensional, it defines a single state called *stabilizer state*.

Notice that in order to have a non trivial subspace V_S it is necessary to have $-\mathbb{1} \notin S$. Indeed otherwise we would have $(-\mathbb{1})|\psi\rangle = |\psi\rangle, \forall |\psi\rangle \in V_S \Rightarrow V_S = \{0\}$.

Observation 2. Any stabilizer group is an *abelian group*! Indeed, let us suppose that $\exists g, h \in S$ such that $[g, h] \neq 0$. Because of Observation 1 e), we have $\{g, h\} = 0 \Rightarrow gh = -hg$. Now, since $gh \in \mathcal{P}_N$, it can be either $(gh)^2 = \mathbb{1}$ or $(gh)^2 = -\mathbb{1}$. In the second case, we can conclude since we break the initial hyphotesis. In the first case, we have $(gh)(hg) = -(gh)(gh) = -(gh)^2 = -\mathbb{1}$, and therefore we find again the absurd.

Stabilizers are abelian groups, with all elements hermitian and squaring to identity.

Observation 3. Any stabilizer group S is generated by an independent set of k (commuting) elements of the Pauli group $g_1, ..., g_k \in \mathcal{P}_N$, meaning that

$$S = \left\{ \prod_{j=1}^{k} (g_j)^{\alpha_j}, \ \alpha_j \in \{0, 1\} \right\} = \langle g_1, ..., g_k \rangle$$

The size of S is therefore $|S| = 2^k$.

Definition 3. Given a stabilizer group S, we define the operator

$$P_S = \frac{1}{|S|} \sum_{g \in S} g \ .$$

It can be easily proven that P_S is nothing but the orthogonal projection on V_S [161]. Indeed, given $g \in S$, $P_g = (\mathbb{1} + g)/2$ is the orthogonal projector on its eigenspace with eigenvalue $+1$. The projector on S is therefore

$$P_S = \prod_{g \in \{g_1, ..., g_k\}} P_g = \frac{1}{2^k} \sum_{\{\alpha\}} \prod_{j=1}^{k} (g_j)^{\alpha_j} = \frac{1}{|S|} \sum_{g \in S} g \ .$$

Observation 4. Let $|\psi\rangle$ be a stabilizer state. Its density matrix operator $\rho = |\psi\rangle\langle\psi|$ is the orthogonal projection on $|\psi\rangle$. Thus

$$\rho = P_S = \frac{1}{|S|} \sum_{g \in S} g \ . \tag{C.1.1}$$

Lemma 1. The dimension of the stabilizer subspace V_S of a stabilizer group generated by k elements of $\mathcal{P}_N$ is 2^{N-k}. Indeed, since P_S is an orthogonal projection its eigenvalues are 0 or 1 and its trace is equal to the number of 1 eigenvalues, i.e. to the dimension of V_S. We have $\mathrm{Tr}[P_S] = 2^{-k} \sum_{g \in S} \mathrm{Tr}[g]$ and, since $\mathrm{Tr}[g]$ is non zero only if $g = \mathbb{1}$, $\mathrm{Tr}[P_S] = 2^{-k} 2^N$.

Observation 5. If $S = \langle g_1, ..., g_k \rangle$ is the stabilizer group of V_S, you cannot find any $g' \in (\mathcal{P}_N \setminus S)$ such that g' stabilizes entirely V_S. Indeed, otherwise $\langle g_1, ..., g_k, g' \rangle$ would be another possible stabilizer group for V_S, but this is impossible, since $2^{N-k} \neq 2^{N-(k+1)}$.

Observation 6. The stabilizer group of an $N-$qubits stabilizer state is generated by exactly N mutually commuting Pauli operators!

Do the stabilizer states cover the entire N−qubits Hilbert space? Obviously not! Indeed, since $\mathcal{P}_N$ has finite cardinality, the number of N mutually commuting Pauli operators is finite. For instance, the stabilizer states for $N = 1$ are just six, namely

$$|0\rangle\ , |1\rangle\ , \frac{|0\rangle + |1\rangle}{\sqrt{2}}\ , \frac{|0\rangle - |1\rangle}{\sqrt{2}}\ , \frac{|0\rangle + i\,|1\rangle}{\sqrt{2}}\ , \frac{|0\rangle - i\,|1\rangle}{\sqrt{2}}\ .$$

It can be shown that the cardinality of the set of pure stabilizer states is

$$|\text{Stab}| = 2^N \prod_{k=0}^{N-1} \left(2^{N-k} + 1\right).$$

C.2 THE CLIFFORD GROUP

Definition 4. The Clifford group $\mathcal{C}_N$ is the group of N−qubits unitaries that map elements of the Pauli group into elements of the Pauli group

$$\mathcal{C}_N = \{U \in \mathcal{U}_{2^N \times 2^N} \text{ s.t. } U\mathcal{P}_N U\dagger = \mathcal{P}_N\}.$$

One can prove that the Hadamard gate H, the phase gate S and the CNOT[1], i.e.

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad S = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{pmatrix} \qquad \text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

generate the entire Clifford group $\mathcal{C}_N$. Another gate is necessary to construct an universal set of gates, i.e. to span all the N−qubits unitaries. This is usually chosen to be

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix},$$

which is non Clifford, since for instance $TXT^\dagger = \frac{1}{\sqrt{2}}(X + Y)$.

Importantly, the uniform distribution on the Clifford group is a 3−design of random Haar unitaries, meaning that the average of any third order polynomial over $\mathcal{C}_N$ equals the average over the entire unitary group U_N [186]. Consequently, Clifford group can be used for the simulation of random unitary circuits, allowing the extraction of information up to the third order moments.

The following result gives a connection between the Clifford group and the set of stabilizer states.

1 CNOT can be replaced with controlled Z

Lemma 2. A state $|\psi\rangle$ is a stabilizer if and only if it is equal to $C|0\rangle^{\otimes N}$, for some $C \in \mathcal{C}_N$. The second implication is very easy to show. Indeed, $|0\rangle^{\otimes N}$ is clearly a stabilizer state having as generators

$$\begin{aligned} g_1 &= Z \otimes \mathbb{1} \otimes ... \otimes \mathbb{1} \\ g_2 &= \mathbb{1} \otimes Z \otimes ... \otimes \mathbb{1} \\ &... \\ g_N &= \mathbb{1} \otimes \mathbb{1} \otimes ... \otimes Z\,. \end{aligned}$$

Since $C\langle g_1, g_2 ... g_N\rangle C^\dagger$ stabilize $|\psi\rangle = C|0\rangle^{\otimes N}$ and since $\langle g_1, g_2 ... g_N\rangle$ has dimension 2^N, we conclude that $|\psi\rangle$ is a stabilizer (it has a stabilizer group of size 2^N). The converse can also be shown easily.

Consider now any stabilizer subspace V_S and an unitary operator $U \in \mathcal{C}_N$. The action of U maps V_S to UV_S. It is not difficult to realize that UV_S is the stabilizer subspace for the group $USU^\dagger$. Indeed, $\forall g' = UgU^\dagger$, $\forall |\psi'\rangle = U|\psi\rangle \in UV_S$ we have $g'|\psi'\rangle = g'U|\psi\rangle = UgU^\dagger U|\psi\rangle = Ug|\psi\rangle = U|\psi\rangle = |\psi'\rangle$. This means that we can keep track of the action of any Clifford unitary simply by applying it to the generator set of S.

C.3 THE TABLEAU REPRESENTATION

A simple but useful mapping exists between elements of $\mathcal{P}_1$ and the binary vector space $(\mathbb{Z}_2)^2$. The mapping is given by the following tableau [159, 187]

$$\begin{cases} r(\mathbb{1}) = (0\,|\,0) \\ r(X) = (1\,|\,0) \\ r(Y) = (1\,|\,1) \\ r(Z) = (0\,|\,1) \end{cases}$$

and neglects the overall phases ± 1 or $\pm i$ (thus, $r(-iX) = r(X)$). It is not difficult to show that $r(gh) = r(g) \oplus r(h)$, $\oplus$ denoting the sum modulo 2. The binary representation can be easily generalized to $\mathcal{P}_N$. Indeed, if we split $\mathbf{r}(g)$ as $(\mathbf{r}^{(1)}(g)\,|\,\mathbf{r}^{(2)}(g))$, where $\mathbf{r}^{(1)}$ and $\mathbf{r}^{(2)}$ are vectors of length N, we can assign the value 1 to the jth component of $\mathbf{r}^{(1)}(g)$ ($j = 1, 2 ... N$) iff the jth Pauli matrix in g is X or Y, and the value 1 to the jth component of $\mathbf{r}^{(2)}(g)$ iff the jth Pauli matrix in g is Y or Z. Furthermore, if we define the matrix

$$\Lambda_N = \begin{pmatrix} 0_{N\times N} & \mathbb{1}_{N\times N} \\ \mathbb{1}_{N\times N} & 0_{N\times N} \end{pmatrix}$$

of dimension $2N \times 2N$, we can realize that the symplectic inner product of two binary vectors evaluates whether the matrices commute or anticommute

$$\big(\mathbf{r}(g)\big)^T \Lambda_N \mathbf{r}(h) = \begin{cases} 0 & \text{if } g, h \text{ commute} \\ 1 & \text{if } g, h \text{ anticommute} \end{cases}\,.$$

Definition 5. Given a stabilizer state, fully specified by $S = \langle g_1, g_2 ... g_N \rangle$, its *generator matrix* is the $N \times 2N$ matrix

$$G = \begin{pmatrix} \mathbf{r}(g_1) \\ \mathbf{r}(g_2) \\ \dots \\ \mathbf{r}(g_N) \end{pmatrix} .$$

This matrix only specifies the generators up to an overall phase.

Observation 7. Notice that the stabilizer generators must be independent, meaning that no product of them can produce $\mathbb{1}$. Since $\mathbf{r}(g_i g_j) = \mathbf{r}(g_i) \oplus \mathbf{r}(g_j)$, no rows of G can sum to zero and therefore G must have full rank (over the field $\mathbb{Z}_2$). Furthermore, since generators must commute, then

$$G^T \Lambda_N G = 0 .$$

Let us notice that the Hadamard gate H and the phase gate S act as follows on the Pauli matrices

$$\begin{cases} HXH^\dagger = Z \\ HYH^\dagger = -Y \\ HZH^\dagger = X \end{cases} \qquad \begin{cases} SXS^\dagger = Y \\ SYS^\dagger = -X \\ SZS^\dagger = Z \end{cases}$$

In other words, H interchanges X and Z, giving Y a phase, and S interchanges X and Y, possibly adding a phase, and leaves Z unchanged. In the tableau representation, this means:

$$\left(\dots r_j^{(1)} \dots \middle| \dots r_j^{(2)} \dots \right) \overset{H_j}{\longleftrightarrow} \left(\dots r_j^{(2)} \dots \middle| \dots r_j^{(1)} \dots \right)$$
$$\left(\dots r_j^{(1)} \dots \middle| \dots r_j^{(2)} \dots \right) \overset{S_j}{\longleftrightarrow} \left(\dots r_j^{(1)} \dots \middle| \dots \left(r_j^{(1)} \oplus r_j^{(2)} \right) \dots \right) .$$

Regarding the CNOT gate, we have (CNOT $=$ CNOT †)

$$\begin{cases} \text{CNOT } (\mathbb{1} \otimes \mathbb{1}) \text{ CNOT } = \mathbb{1} \otimes \mathbb{1} \\ \text{CNOT } (\mathbb{1} \otimes X) \text{ CNOT } = \mathbb{1} \otimes X \\ \text{CNOT } (\mathbb{1} \otimes Y) \text{ CNOT } = Z \otimes Y \\ \text{CNOT } (\mathbb{1} \otimes Z) \text{ CNOT } = Z \otimes Z \\ \text{CNOT } (X \otimes \mathbb{1}) \text{ CNOT } = X \otimes X \\ \text{CNOT } (X \otimes X) \text{ CNOT } = X \otimes \mathbb{1} \\ \text{CNOT } (X \otimes Y) \text{ CNOT } = Y \otimes Z \\ \text{CNOT } (X \otimes Z) \text{ CNOT } = -Y \otimes Y \end{cases} \qquad \begin{cases} \text{CNOT } (Y \otimes \mathbb{1}) \text{ CNOT } = Y \otimes X \\ \text{CNOT } (Y \otimes X) \text{ CNOT } = Y \otimes \mathbb{1} \\ \text{CNOT } (Y \otimes Y) \text{ CNOT } = -X \otimes Z \\ \text{CNOT } (Y \otimes Z) \text{ CNOT } = X \otimes Y \\ \text{CNOT } (Z \otimes \mathbb{1}) \text{ CNOT } = Z \otimes \mathbb{1} \\ \text{CNOT } (Z \otimes X) \text{ CNOT } = Z \otimes X \\ \text{CNOT } (Z \otimes Y) \text{ CNOT } = \mathbb{1} \otimes Y \\ \text{CNOT } (Z \otimes Z) \text{ CNOT } = \mathbb{1} \otimes Z \end{cases}$$

This table can be summarized in following update rule for a CNOT where the qubit k control the target qubit j ($j, k \in \{1, 2, ...N\}$)

$$\left(\dots r_j^{(1)} \oplus r_k^{(1)} \dots r_k^{(1)} \dots \middle| \dots r_j^{(2)} \dots r_k^{(2)} \oplus r_j^{(2)} \dots \right) \overset{\text{CNOT}_{kj}}{\longleftrightarrow} \left(\dots r_j^{(1)} \dots r_k^{(1)} \dots \middle| \dots r_j^{(2)} \dots r_k^{(2)} \dots \right) .$$

As a simple example, let us consider a system with $N = 2$ qubits. The initial state is $|00\rangle$, which has $\{\mathbb{1}_1 Z_2, Z_1 \mathbb{1}_2\}$ as stabilizer generators. The stabilizer matrix is therefore

$$G = \left(\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

If we apply the gate H on the first qubit we can update the matrix G as follows

$$G' = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

Then, we apply a CNOT, that gives

$$G'' = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

Indeed, we have

$$\mathrm{CNOT}_{12}\big(H_1 |00\rangle\big) = \mathrm{CNOT}_{12} \frac{|00\rangle + |10\rangle}{\sqrt{2}} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

and the final state has $\{X_1 X_2, Z_1 Z_2\}$ as stabilizer generators, which are exactly the operators encoded in the tableau G''.

Other simple rules can be found to keep track of the overall sign ± 1 and to update the generator matrix after a Pauli measurements [159]. These observations lead to the following important Theorem.

Theorem 1 (Gottesman-Knill)**.** Suppose we perform a quantum computation which involves only state preparations in the computational basis, Hadamard gates H, phase gates S, CNOT gates, Pauli gates and measurements of observables in the Pauli group, together with the possibility of classical control conditioned on the outcome of such measurements. Such a computation may be *efficiently* simulated on a classical computer, meaning that there exists an Algorithm to do this classically in a polynomial time, specifically at cost $\mathcal{O}(N^2 M)$ operations, where N is the number of qubits and M the number of operations.

C.4 NONSTABILIZERNESS AS A QUANTUM RESOURCE

As we saw, a quantum computation involving only stabilizer states (or equivalently, Clifford circuits and measurements in the computational basis) can be simulated classically. Therefore, one has to consider nonstabilizer protocols to reach a true quantum advantage. In this sense, the "nonstabilizerness" can be considered as a key "resource" of the quantum world, making it more powerful of the classical one. This intuitive picture can be made mathematically rigorous by exploiting the meta-theory of quantum resources. In this framework, one usually introduce a set if "free-operations" that by construction do not generate resource. In the case of entanglement these are the

local operations and classical communication (LOCC). In the case of the nonstabilizerness, also dubbed *quantum magic*, they are the following stabilizer operations [2]:

- Clifford unitaries, i.e. $\rho \to U\rho U^\dagger$, with $U \in \mathcal{C}_N$
- Composition with stabilizer states, i.e. $\rho \to \rho \otimes \rho_S$, where ρ_S is a stabilizer state
- Measurements in the computational basis, i.e. $\rho \to \sum_k P_k \rho P_k$, where P_k are the projectors on the computational basis ($\sum_k P_k = \mathbb{1}$)
- Discarding some qubits, i.e. $\rho \to \mathrm{Tr}_A[\rho]$, where A is any subsystem
- The above operations conditioned on the outcomes of measurements

The goal is generally to evaluate the amount of resource stored into a state ρ by means of a quantity $\mathcal{M}(\rho)$, usually named *monotone*, that is required to be not increasing under any arbitrary free operations. For instance, $\mathcal{M}$ is a *magic monotone* if it respects $\mathcal{M}\big(\mathcal{E}(\rho)\big) \leqslant \mathcal{M}\big(\rho\big)$ for any combinations $\mathcal{E}$ of the stabilizer operations listed above. If one is able to identify and calculate a monotone measure $\mathcal{M}$, it can utilize it to discriminate between states that can or cannot be prepared from an initial state using free operations. The resource needed to prepare a state can be injected in the initial state, for instance having access to many copies of the magic T−state $|T\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + e^{i\frac{\pi}{4}}|1\rangle\big)$.

Examples of known genuine magic monotones are the following.

1. The robustness of magic, defined as

$$R(\rho) = \min\Big\{ \sum_{n=1}^{|\mathrm{Stab}|} |x_n| \text{ s.t. } \rho = \sum_n x_n \rho_S^{(n)} \Big\}, \tag{C.4.1}$$

   where $\rho_S^{(n)} = |\psi^{(n)}\rangle\langle\psi^{(n)}|$, $n = 1, 2 \ldots |\mathrm{Stab}|$, are all the pure stabilizer states over N qubits.

2. The min-relative entropy of magic, which for pure states reads

$$D(|\psi\rangle) = -\log\left(\max_n |\langle\psi^{(n)}|\psi\rangle|^2\right), \tag{C.4.2}$$

   with the same notation as above.

However, both these quantity are very difficult to handle in practice since their evaluation requires an optimization over exponentially large spaces. In the next Section, we will introduce the Stabilizer Rényi Entropies, a possible candidate for a simpler measure of quantum magic.

2 Sometimes an additional condition, known as *strong monotonicity*, is also imposed. It can be expressed as $\sum_k p_k \mathcal{M}(P_k \rho P_k) \leqslant \mathcal{M}(\rho)$ and it ensures that $\mathcal{M}$ does not increase, on average, when the experimenter has the ability to post-select multiple outcomes of a quantum measurement.

C.5 STABILIZER RÉNYI ENTROPIES

Definition 6. The projective Pauli group is the standard Pauli group $\mathcal{P}_N$ modulo global phases, i.e. $\tilde{\mathcal{P}}_N = \{\mathbb{1}, X, Y, Z\}^{\otimes N}$.

Lemma 3. Given a stabilizer state $|\psi\rangle$ and a Pauli string $g \in \tilde{\mathcal{P}}_N$, we have

$$\langle\psi|g|\psi\rangle = \begin{cases} +1 & \text{if } g \in \mathcal{S} \\ 0 & \text{if } g \notin \mathcal{S}, -g \notin \mathcal{S} \\ -1 & \text{if } -g \in \mathcal{S} \end{cases}$$

Indeed, the expectation value can be computed as

$$\langle\psi|g|\psi\rangle = \frac{1}{2^N} \sum_{g' \in \mathcal{S}} \mathrm{Tr}\left[g\, g'\right],$$

and we can conclude using the fact that the trace of a tensor product factorizes and that the matrices $\{\mathbb{1}, X, Y, Z\}$ are trace-orthogonal.

Definition 7. Given a N−qubits state $|\psi\rangle$, we define its *Stabilizer α-Rényi Entropies* (SRE) [188] as

$$M_\alpha(|\psi\rangle) = \frac{1}{1-\alpha} \log\left(\sum_{g \in \tilde{\mathcal{P}}_N} \frac{1}{2^N} \langle\psi|g|\psi\rangle^{2\alpha} \right). \tag{C.5.1}$$

To understand the relation with usual Rényi entropies it is useful to introduce the density matrix $\rho = |\psi\rangle\langle\psi|$ and the function $\Pi_\rho(g) = \frac{1}{2^N}(\mathrm{Tr}[\rho g])^2$. Let us observe that

$$\sum_{g \in \tilde{\mathcal{P}}_N} \Pi_\rho(g) = \frac{1}{2^N} \sum_{g \in \tilde{\mathcal{P}}_N} \mathrm{Tr}[\rho g] \cdot \mathrm{Tr}[\rho g] = \mathrm{Tr}\left[\rho \cdot \frac{1}{2^N} \sum_{g \in \tilde{\mathcal{P}}_N} \mathrm{Tr}[\rho g] g\right] = \mathrm{Tr}\left[\rho^2\right] = 1\,.$$

We used the decomposition of the matrix ρ in terms of $\{\frac{g}{\sqrt{2^N}}\}_{g \in \tilde{\mathcal{P}}_N}$, which is a complete base set for the hermitian matrices of size 2^N. We also use the fact that for pure states $\mathrm{Tr}\left[\rho^2\right] = 1$. Thus, we can consider $\Pi_\rho(g)$ as a probability distribution. We have

$$M_\alpha(|\psi\rangle) = \frac{1}{1-\alpha} \log\left(\sum_{g \in \tilde{\mathcal{P}}_N} \left(\Pi_\rho(g)\right)^\alpha \right) - N \log 2\,, \tag{C.5.2}$$

that shows that $M_\alpha(|\psi\rangle)$ is the α-Rényi entropy of $\Pi_\rho(g)$, apart from a constant. Notice that, to extend the definition of M_α to arbitrary density matrices ρ, we have to normalize with the purity, i.e.

$$\Pi_\rho(g) = \frac{1}{2^N} \frac{(\mathrm{Tr}[\rho g])^2}{\mathrm{Tr}[\rho^2]}\,. \tag{C.5.3}$$

Observation 8.

a) For a stabilizer state we have $M_\alpha(|\psi\rangle) = 0$. Indeed, by means of Lemma 3, we have

$$M_\alpha(|\psi\rangle) = \frac{1}{1-\alpha} \log\left(2^N \cdot \frac{1}{2^N} (\pm 1)^{2\alpha} \right) = 0\,. \tag{C.5.4}$$

b) In general, $M_\alpha(|\psi\rangle) = 0$ iff $|\psi\rangle$ is a stabilizer and $M_\alpha(|\psi\rangle) > 0$ otherwise.

c) The SRE are invariant under Clifford unitaries. Indeed, even if the Pauli strings are reshuffled by a Clifford unitary, the probability values $\Pi_\rho(g)$ remain unchanged and consequently also the values of the Rényi entropies.

d) The SRE are additive, i.e. $M_\alpha(|\psi\rangle \otimes |\phi\rangle) = M_\alpha(|\psi\rangle) + M_\alpha(|\phi\rangle)$. This is a trivial consequence of the fact that the sum over Pauli strings on a system can be factorized into the sum of the Pauli strings of two subsystems.

These properties have recently attracted a lot of attention to SREs, considered as a potential simpler monotone for magic. However, it was observed that

- for any $0 \leqslant \alpha \leqslant 2$, SRE are not genuine monotones since there exist protocols involving measurements and operations conditioned by their outcomes for which M_α increase;
- for any α, SRE are not strong monotones.

C.6 ENTANGLEMENT

C.6.1 Entanglement entropy

Let us suppose to have a pure stabilizer state

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2^N}\sum_{g\in\mathcal{S}} g \ .$$

We aim to assess the entanglement entropy $S(A|B)$ associated to the partitioning of the system into a subsystem A consisting of N_A qubits and a complementary subsystem B containing $N_B = N - N_A$ qubits. We start by computing the reduced density matrix $\rho_A = \mathrm{Tr}_B[\rho]$:

$$\rho_A = \frac{1}{2^N}\sum_{g\in\mathcal{S}} \mathrm{Tr}_B[g] \ .$$

Only some elements g of the group $\mathcal{S}$ give a non vanishing contribution, namely the Pauli strings that can be written as $g = g_A \otimes \mathbb{1}_B$. Notice that this subset $\mathcal{S}_A$ is also a stabilizer group, since $(g_A \otimes \mathbb{1}_B)(g'_A \otimes \mathbb{1}_B) = (g_A g'_A) \otimes \mathbb{1}_B$. Let us denote k_A the number of independent generators of $\mathcal{S}_A$. Now we have

$$\rho_A = \frac{1}{2^N} 2^{N_B} \sum_{g\in\mathcal{S}_A} g = \frac{1}{2^{N_A}} 2^{k_A} \frac{1}{2^{k_A}} \sum_{g\in\mathcal{S}_A} g \, .$$

The last part of the expression is the projector on the stabilizer subspace of $\mathcal{S}_A$, thus

$$\rho_A = \frac{1}{2^{N_A - k_A}} P_{\mathcal{S}_A} \ . \tag{C.6.1}$$

Now, since eigenvalues of P_{S_A} are only 1 (with degeneracy $2^{N_A-k_A}$) and 0 (with degeneracy $2^{N_A}-2^{N_A-k_A}$), we get

$$S(A|B) = -2^{N_A-k_A}\frac{1}{2^{N_A-k_A}}\log\frac{1}{2^{N_A-k_A}} = (N_A-k_A)\log 2\,.$$

How to compute k_A in practice given the generator matrix G of $|\psi\rangle$? It is not difficult to show that k_A is simply the rank of the matrix obtained by setting to 0 all the entries of G corresponding to the subsystem B. Indeed, this operation corresponds to projecting the original generators on the space of operators having the form $g_A \otimes \mathbb{1}_B$.

C.6.2 Entanglement fluctuations: new results

Let us notice that the entanglement spectrum of Eq. C.6.1 is flat, meaning that all non zero Schmidt eigenvalues take the same value. This is true for any stabilizer state $|\psi\rangle$ and any bipartition.

Let us now consider a possible measure of the flatness of the entanglement spectrum. We define

$$\mathcal{F}_A(\rho) = \mathrm{Tr}\left[\rho_A^3\right] - \mathrm{Tr}^2[\rho_A^2]\,. \tag{C.6.2}$$

One can easily check that $\mathcal{F}_A(\rho) = 0$ iff the entanglement spectrum is flat, whereas $\mathcal{F}_A(\rho) > 0$ in other cases. We have seen that being a stabilizer state implies $\mathcal{F}_A = 0$, while the converse is not true since there exist product states with a flat entanglement spectrum that have an extensive amount of magic (for instance, $\left(T|0\rangle\right)^{\otimes N}$). Nevertheless, it is possible to make the relationship between $\mathcal{F}_A$ and the nonstabilizerness more precise. In fact, the following result, obtained for the first time in Ref. [189], holds.

Theorem 2. For any pure state $|\psi\rangle$, the flatness of the entanglement spectrum averaged over the Clifford orbit is proportional to the linear SRE M_{lin}, which is defined as

$$\begin{aligned} M_{\mathrm{lin}}(|\psi\rangle) &= 1-\exp\left(-M_2(|\psi\rangle)\right) = \\ &= 1-2^N\sum_{g\in\tilde{\mathcal{P}}_N}\left(\Pi_\rho(g)\right)^2 \end{aligned} \tag{C.6.3}$$

Indeed

$$\langle\mathcal{F}_A(U_C|\psi\rangle)\rangle_{\mathcal{C}_N} = c(D,D_A)M_{\mathrm{lin}}(|\psi\rangle)\,, \tag{C.6.4}$$

where $\langle\cdot\rangle_{\mathcal{C}_N}$ is the average over the Clifford group $\{U_C\}_{U_C\in\mathcal{C}_N}$, $D = 2^N$ and $D_A = 2^{N_A}$ are the Hilbert space dimensions of the system and subsystem, and the proportionality constant is $c(D,D_A) \sim (D^2-D_A^2)D^{-3}$ for large D.

As a Corollary, we have that $|\psi\rangle$ is a stabilizer state if and only if the entanglement spectrum of any state in its Clifford orbit $\{U_C|\psi\rangle\}_{U_C\in\mathcal{C}_N}$ is flat for any bipartition.

SRE is a very expensive quantity to measure, typically demanding exponential resources (although it is more efficient than full state tomography).

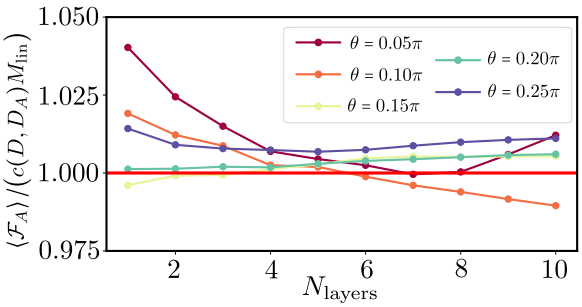

Figure C.1: Ratio $\langle \mathcal{F}_A \rangle / \left(c(D, D_A) M_{\text{lin}} \right)$ as a function of the number of layers N_{Layers} of a shallow Clifford circuit. The system is initially prepared in a volume law state. The ratio approaches 1 very fast verifying Eq. C.6.4.

The Theorem paves the way for a more efficient method of measuring SRE by averaging $\mathcal{F}_A$ over the Clifford orbit of $|\psi\rangle$. In principle, this might require averaging over a very large number of Clifford unitaries, which would be impractical. However, in Ref. [189], we numerically observed that this is not always the case.

In Fig. C.1, we show that one can accurately estimate M_{lin} even by shallow Clifford circuits provided one starts with volume law entanglement. We consider a system of $N = 14$ qubits in a volume law phase by subjecting the initial product state $|\psi(\theta)\rangle = 2^{-N/2}(|0\rangle + e^{i\theta}|1\rangle)^{\otimes N}$ to 1500 Clifford layers, for various values of θ. We then plot the ratio $\langle \mathcal{F}_A \rangle / \left(c(D, D_A) M_{\text{lin}} \right)$ as a function of the number of Clifford N_{Layers}. The theoretical line predicted by the theorem, corresponding to the value 1, is shown as a solid red line. Notably, we observe that even for circuits as short as $N_{\text{Layers}} = 7$ Clifford layers, the average flatness reaches the value predicted by the above Theorem.

In Ref. [189], we also propose a simple Algorithm to determine if a given state $|\psi\rangle$ is a stabilizer state or possesses some nonstabilizerness. The procedure works as follows: (1) Start with $|\psi\rangle$. (2) Draw a random Clifford gate U_C and apply it to the initial state: $U_C|\psi\rangle$. (3) Measure the entanglement spectrum flatness $\mathcal{F}_A(U_C|\psi\rangle)$. If the original state $|\psi\rangle$ is a stabilizer state, the output of the circuit is still a stabilizer state with zero flatness. On the contrary, if $|\psi\rangle$ has some magic, we expect that even a modest exploration of the Clifford orbit will result into a non flat entanglement spectrum. Therefore, if after a number of Clifford unitaries we measure $\mathcal{F}_A > 0$ we can establish that the initial state possesses nonstabilizerness. In a real-world scenario, one has to set both a number of iterations (which determines the maximum number of Clifford layers applied) and a threshold ϵ for measuring the flatness. In Fig. C.2, we present the probability of success of this Algorithm p_{succ} for $N = 12$ qubits and for a different number of Clifford layers N_{Layers}. We fix the threshold $\epsilon = 0.005$ and we compute the probability as a function of the SRE $M_2(|\psi\rangle)$ of the initial state. The plot shows that increasing the number of iterations of the Algorithm push the probability of success to 1 for any fixed values of nonstabilizerness.

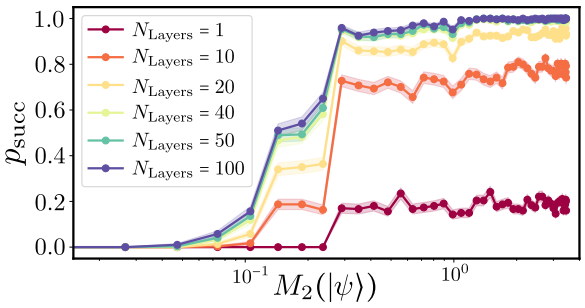

Figure C.2: Probability of success p_{succ} for a fixed threshold $\epsilon = 0.005$ and for different N_{Layers}. After collecting 10^3 realizations, we compute the probability of success p_{succ} as a function of the initial value of non-stabilizerness calculated using the second SRE $M_2(|\psi\rangle)$.

ACKNOWLEDGEMENTS

Firstly, I would like to thank Mario for the continuous support, patience, dedication. During the last four years, he left me free to explore new topics and spent lot of time discussing with me, whether in front of a blackboard or a computer.

I want to express my gratitude to my dear colleagues Nishan, Mandana, Alì, Mina, Gabriele, Pietro, with whom I have shared many days of work and fun at SISSA. A special place is reserved to Mr. Ranabhat, who taught me many things (especially regarding "The godfather").

Among all the friends, I want to thank Jess, Riccardino, Lucia, Davide and the new entry Gaia.

I warmly thank my love and partner in adventures Giorgia. Once again, we shared laughs, hopes, fears, disappointments, and joys.

Finally, a huge heartfelt thanks to my dad, my mum and Ludovico, who constantly supported me.

BIBLIOGRAPHY

[1] K. Konishi and G. Paffuti. *Quantum Mechanics: A New Introduction*. OUP Oxford, 2009. ISBN: 9780199560264. URL: https://books.google.it/books?id=YuQ-DgAAQBAJ.

[2] Max Born. *The Statistical Interpretations of Quantum Mechanics*. https://www.nobelprize.org/uploads/2018/06/born-lecture.pdf.

[3] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" In: *Phys. Rev.* 47 (10 May 1935), pp. 777–780. DOI: 10.1103/PhysRev.47.777. URL: https://link.aps.org/doi/10.1103/PhysRev.47.777.

[4] R. Penrose. *The Road to Reality: A Complete Guide to the Laws of the Universe*. Science: Astrophysics. Jonathan Cape, 2004. ISBN: 9780224044479. URL: https://books.google.it/books?id=csaaQgAACAAJ.

[5] E. Schrödinger. "Are There Quantum Jumps ?" In: *British Journal for the Philosophy of Science* 3.11 (1952), pp. 233–242. DOI: 10.1093/bjps/iii.11.233.

[6] G Wendin. "Quantum information processing with superconducting circuits: a review". In: *Reports on Progress in Physics* 80.10 (Sept. 2017), p. 106001. DOI: 10.1088/1361-6633/aa7e1a. URL: https://doi.org/10.1088%2F1361-6633%2Faa7e1a.

[7] Sergey Bravyi et al. "The future of quantum computing with superconducting qubits". In: *Journal of Applied Physics* 132.16 (Oct. 2022), p. 160902. ISSN: 0021-8979. DOI: 10.1063/5.0082975. URL: https://doi.org/10.1063/5.0082975.

[8] Xiaoling Wu et al. "A concise review of Rydberg atom based quantum computation and quantum simulation". In: *Chinese Physics B* 30.2 (Feb. 2021), p. 020305. DOI: 10.1088/1674-1056/abd76f. URL: https://doi.org/10.1088%2F1674-1056%2Fabd76f.

[9] C S Adams, J D Pritchard, and J P Shaffer. "Rydberg atom quantum technologies". In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 53.1 (Dec. 2019), p. 012002. DOI: 10.1088/1361-6455/ab52ef. URL: https://doi.org/10.1088%2F1361-6455%2Fab52ef.

[10] Florian Schäfer et al. "Tools for quantum simulation with ultracold atoms in optical lattices". In: *Nature Reviews Physics* 2.9 (July 2020), pp. 411–425. DOI: 10.1038/s42254-020-0195-3. URL: https://doi.org/10.1038%2Fs42254-020-0195-3.

[11] Colin D. Bruzewicz et al. "Trapped-ion quantum computing: Progress and challenges". In: *Applied Physics Reviews* 6.2 (May 2019). DOI: 10.1063/1.5088164. URL: https://doi.org/10.1063%2F1.5088164.

[12] Guido Burkard et al. "Semiconductor spin qubits". In: *Reviews of Modern Physics* 95.2 (June 2023). DOI: 10.1103/revmodphys.95.025003. URL: https://doi.org/10.1103%2Frevmodphys.95.025003.

[13] Farokh Mivehvar et al. "Cavity QED with quantum gases: new paradigms in many-body physics". In: *Advances in Physics* 70.1 (Jan. 2021), pp. 1–153. DOI: 10.1080/00018732.2021.1969727. URL: https://doi.org/10.1080%2F00018732.2021.1969727.

[14] L. Jaeger. *The Second Quantum Revolution: From Entanglement to Quantum Computing and Other Super-Technologies*. Springer, 2018. ISBN: 9783319988252. URL: https://books.google.it/books?id=r8y0zQEACAAJ.

[15] R.P. Feynman. "Simulating physics with computers". In: *International Journal of Theoretical Physics* 21 (1982). DOI: /10.1007/BF02650179.

[16] Peter W. Shor. "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer". In: *SIAM Journal on Computing* 26.5 (Oct. 1997), pp. 1484–1509. DOI: 10.1137/s0097539795293172. URL: https://doi.org/10.1137%2Fs0097539795293172.

[17] John Preskill. "Quantum Computing in the NISQ era and beyond". In: *Quantum* 2 (Aug. 2018), p. 79. ISSN: 2521-327X. DOI: 10.22331/q-2018-08-06-79. URL: https://doi.org/10.22331/q-2018-08-06-79.

[18] Pasquale Calabrese and John Cardy. "Entanglement entropy and quantum field theory". In: *Journal of Statistical Mechanics: Theory and Experiment* 2004.06 (June 2004), P06002. DOI: 10.1088/1742-5468/2004/06/P06002. URL: https://dx.doi.org/10.1088/1742-5468/2004/06/P06002.

[19] G. Vidal et al. "Entanglement in Quantum Critical Phenomena". In: *Phys. Rev. Lett.* 90 (22 June 2003), p. 227902. DOI: 10.1103/PhysRevLett.90.227902. URL: https://link.aps.org/doi/10.1103/PhysRevLett.90.227902.

[20] Mark Srednicki. "Entropy and area". In: *Phys. Rev. Lett.* 71 (5 Aug. 1993), pp. 666–669. DOI: 10.1103/PhysRevLett.71.666. URL: https://link.aps.org/doi/10.1103/PhysRevLett.71.666.

[21] M B Hastings. "An area law for one-dimensional quantum systems". In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.08 (Aug. 2007), P08024. DOI: 10.1088/1742-5468/2007/08/P08024. URL: https://dx.doi.org/10.1088/1742-5468/2007/08/P08024.

[22] W. Kohn. "Nobel Lecture: Electronic structure of matter—wave functions and density functionals". In: *Rev. Mod. Phys.* 71 (5 Oct. 1999), pp. 1253–1266. DOI: 10.1103/RevModPhys.71.1253. URL: https://link.aps.org/doi/10.1103/RevModPhys.71.1253.

[23] Steven R. White. "Density matrix formulation for quantum renormalization groups". In: *Phys. Rev. Lett.* 69 (19 Nov. 1992), pp. 2863–2866. DOI: 10.1103/PhysRevLett.69.2863. URL: https://link.aps.org/doi/10.1103/PhysRevLett.69.2863.

[24] Steven R. White. "Density-matrix algorithms for quantum renormalization groups". In: *Phys. Rev. B* 48 (14 Oct. 1993), pp. 10345–10356. DOI: 10.1103/PhysRevB.48.10345. URL: https://link.aps.org/doi/10.1103/PhysRevB.48.10345.

[25] C. Lange, A. Klumper, and J. Zittartz. "Exact groundstates for antiferromagnetic spin-one chains with nearest and next-nearest neighbour interactions". In: *Zeitschrift for Physik B Condensed Matter* 96.2 (June 1994), pp. 267–270. DOI: 10.1007/bf01313293. URL: https://doi.org/10.1007%2Fbf01313293.

[26] Stellan Östlund and Stefan Rommer. "Thermodynamic Limit of Density Matrix Renormalization". In: *Physical Review Letters* 75.19 (Nov. 1995), pp. 3537–3540. DOI: 10.1103/physrevlett.75.3537. URL: https://doi.org/10.1103%2Fphysrevlett.75.3537.

[27] Ulrich Schollwöck. "The density-matrix renormalization group in the age of matrix product states". In: *Annals of Physics* 326.1 (2011). January 2011 Special Issue, pp. 96–192. ISSN: 0003-4916. DOI: https://doi.org/10.1016/j.aop.2010.09.012. URL: https://www.sciencedirect.com/science/article/pii/S0003491610001752.

[28] Alberto Baiardi and Markus Reiher. "The density matrix renormalization group in chemistry and molecular physics: Recent developments and new challenges". In: *The Journal of Chemical Physics* 152.4 (Jan. 2020), p. 040903. ISSN: 0021-9606. DOI: 10.1063/1.5129672. URL: https://doi.org/10.1063/1.5129672.

[29] Anupam Mitra et al. *Macrostates vs. Microstates in the Classical Simulation of Critical Phenomena in Quench Dynamics of 1D Ising Models*. 2023. arXiv: 2310.08567 [quant-ph].

[30] John Preskill. *Quantum computing and the entanglement frontier*. 2012. DOI: 10.48550/ARXIV.1203.5813. URL: https://arxiv.org/abs/1203.5813.

[31] Lloyd N Trefethen and David Bau. *Numerical linear algebra*. Vol. 181. Siam, 2022.

[32] Wikipedia contributors. *Singular value decomposition — Wikipedia, The Free Encyclopedia*. 2023.

[33] G. Catarina and Bruno Murta. *Density-matrix renormalization group: a pedagogical introduction*. 2023. arXiv: 2304.13395 [cond-mat.str-el].

[34] Carl Eckart and G. Marion Young. "The approximation of one matrix by another of lower rank". In: *Psychometrika* 1 (1936), pp. 211–218. URL: https://api.semanticscholar.org/CorpusID:10163399.

[35] L. MIRSKY. "SYMMETRIC GAUGE FUNCTIONS AND UNITARILY INVARIANT NORMS". In: *The Quarterly Journal of Mathematics* 11.1 (Jan. 1960), pp. 50–59. ISSN: 0033-5606. DOI: 10.1093/qmath/11.1.50. eprint: https://academic.oup.com/qjmath/article-pdf/11/1/50/7295335/11-1-50.pdf. URL: https://doi.org/10.1093/qmath/11.1.50.

[36] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511976667.

[37] Roger Penrose. "Applications of negative dimensional tensors". In: *Combinatorial Mathematics and its Applications* (1971), pp. 221–244.

[38] Pietro Silvi et al. "The Tensor Networks Anthology: Simulation techniques for many-body quantum lattice systems". In: *SciPost Physics Lecture Notes* (Mar. 2019). DOI: 10.21468/scipostphyslectnotes.8. URL: https://doi.org/10.21468%2Fscipostphyslectnotes.8.

[39] Shi-Ju Ran et al. *Tensor Network Contractions: Methods and Applications to Quantum Many-Body Systems*. Springer International Publishing, 2020.

[40] Laurens Vanderstraeten, Jutho Haegeman, and Frank Verstraete. "Tangent-space methods for uniform matrix product states". In: *SciPost Physics Lecture Notes* (Jan. 2019). DOI: 10.21468/scipostphyslectnotes.7. URL: https://doi.org/10.21468%2Fscipostphyslectnotes.7.

[41] Jutho Haegeman et al. "Unifying time evolution and optimization with matrix product states". In: *Physical Review B* 94.16 (Oct. 2016). DOI: 10.1103/physrevb.94.165116. URL: https://doi.org/10.1103%2Fphysrevb.94.165116.

[42] Hamed Saberi et al. "Constrained optimization of sequentially generated entangled multiqubit states". In: *Physical Review A* 80.2 (Aug. 2009). DOI: 10.1103/physreva.80.022334. URL: https://doi.org/10.1103%2Fphysreva.80.022334.

[43] Sebastian Paeckel et al. "Time-evolution methods for matrix-product states". In: *Annals of Physics* 411 (Dec. 2019), p. 167998. DOI: 10.1016/j.aop.2019.167998. URL: https://doi.org/10.1016%2Fj.aop.2019.167998.

[44] Jutho Haegeman and Frank Verstraete. "Diagonalizing Transfer Matrices and Matrix Product Operators: A Medley of Exact and Computational Methods". In: *Annual Review of Condensed Matter Physics* 9.1 (Mar. 2017), pp. 355–406. DOI: 10.1146/annurev-conmatphys-031016-025507. URL: https://doi.org/10.1146%2Fannurev-conmatphys-031016-025507.

[45] A. Weichselbaum et al. "Variational matrix-product-state approach to quantum impurity models". In: *Physical Review B* 80.16 (Oct. 2009). DOI: 10.1103/physrevb.80.165117. URL: https://doi.org/10.1103%2Fphysrevb.80.165117.

[46] Ian P McCulloch. "From density-matrix renormalization group to matrix product states". In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.10 (Oct. 2007), P10014. DOI: 10.1088/1742-5468/2007/10/P10014. URL: https://dx.doi.org/10.1088/1742-5468/2007/10/P10014.

[47] Matthew Fishman, Steven R. White, and E. Miles Stoudenmire. "The ITensor Software Library for Tensor Network Calculations". In: *SciPost Phys. Codebases* (2022), p. 4. DOI: 10.21468/SciPostPhysCodeb.4. URL: https://scipost.org/10.21468/SciPostPhysCodeb.4.

[48] Johannes Hauschild and Frank Pollmann. "Efficient numerical simulations with Tensor Networks: Tensor Network Python (TeNPy)". In: *SciPost Phys. Lect. Notes* (2018), p. 5. DOI: 10.21468/SciPostPhysLectNotes.5. URL: https://scipost.org/10.21468/SciPostPhysLectNotes.5.

[49] Wikipedia contributors. *Variational method (quantum mechanics) — Wikipedia, The Free Encyclopedia*. 2023.

[50] A. M. Krol et al. *Efficient decomposition of unitary matrices in quantum circuit compilers*. 2021. DOI: 10.48550/ARXIV.2101.02993. URL: https://arxiv.org/abs/2101.02993.

[51] Sheng-Hsuan Lin et al. "Real- and Imaginary-Time Evolution with Compressed Quantum Circuits". In: *PRX Quantum* 2.1 (Mar. 2021). DOI: 10.1103/prxquantum.2.010342. URL: https://doi.org/10.1103%2Fprxquantum.2.010342.

[52] Peter Schauss et al. "Observation of spatially ordered structures in a two-dimensional Rydberg gas". In: *Nature* 491.7422 (2012), pp. 87–91. DOI: 10.1038/nature11596. URL: https://doi.org/10.1038/nature11596.

[53] Henning Labuhn et al. "Tunable two-dimensional arrays of single Rydberg atoms for realizing quantum Ising models". In: *Nature* 534.7609 (2016), pp. 667–670. DOI: 10.1038/nature18274. URL: https://doi.org/10.1038/nature18274.

[54] Hannes Bernien et al. "Probing many-body dynamics on a 51-atom quantum simulator". In: *Nature* 551.7682 (2017), pp. 579–584. DOI: 10.1038/nature24622. URL: https://doi.org/10.1038/nature24622.

[55] Peter Schauss. "Quantum simulation of transverse Ising models with Rydberg atoms". In: *Quantum Science and Technology* 3.2 (Jan. 2018), p. 023001. DOI: 10.1088/2058-9565/aa9c59. URL: https://doi.org/10.1088/2058-9565/aa9c59.

[56] A. Omran et al. "Generation and manipulation of Schrödinger cat states in Rydberg atom arrays". In: *Science* 365.6453 (2019), pp. 570–574. DOI: 10.1126/science.aax9743. URL: https://www.science.org/doi/abs/10.1126/science.aax9743.

[57] Immanuel Bloch, Jean Dalibard, and Wilhelm Zwerger. "Many-body physics with ultracold gases". In: *Rev. Mod. Phys.* 80 (3 July 2008), pp. 885–964. DOI: 10.1103/RevModPhys.80.885. URL: https://link.aps.org/doi/10.1103/RevModPhys.80.885.

[58] Florian Schäfer et al. "Tools for quantum simulation with ultracold atoms in optical lattices". In: *Nature Reviews Physics* 2.8 (July 2020), pp. 411–425. ISSN: 2522-5820. DOI: 10.1038/s42254-020-0195-3. URL: http://dx.doi.org/10.1038/s42254-020-0195-3.

[59] Simone Notarnicola, Mario Collura, and Simone Montangero. "Real-time-dynamics quantum simulation of $(1+1)$-dimensional lattice QED with Rydberg atoms". In: *Phys. Rev. Research* 2 (1 Mar. 2020), p. 013288. DOI: 10.1103/PhysRevResearch.2.013288. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.013288.

[60] Federica M. Surace et al. "Lattice Gauge Theories and String Dynamics in Rydberg Atom Quantum Simulators". In: *Phys. Rev. X* 10 (2 May 2020), p. 021041. DOI: 10.1103/PhysRevX.10.021041. URL: https://link.aps.org/doi/10.1103/PhysRevX.10.021041.

[61] Timo Felser, Simone Notarnicola, and Simone Montangero. "Efficient Tensor Network Ansatz for High-Dimensional Quantum Many-Body Problems". In: *Phys. Rev. Lett.* 126 (17 Apr. 2021), p. 170603. DOI: 10.1103/PhysRevLett.126.170603. URL: https://link.aps.org/doi/10.1103/PhysRevLett.126.170603.

[62] Vincent Lienhard et al. "Observing the Space- and Time-Dependent Growth of Correlations in Dynamically Tuned Synthetic Ising Models with Antiferromagnetic Interactions". In: *Phys. Rev. X* 8 (2 June 2018), p. 021070. DOI: 10.1103/PhysRevX.8.021070. URL: https://link.aps.org/doi/10.1103/PhysRevX.8.021070.

[63] M. Morgado and S. Whitlock. "Quantum simulation and computing with Rydberg-interacting qubits". In: *AVS Quantum Science* 3.2 (June 2021), p. 023501. ISSN: 2639-0213. DOI: 10.1116/5.0036562. URL: http://dx.doi.org/10.1116/5.0036562.

[64] Pietro Silvi et al. "The Tensor Networks Anthology: Simulation techniques for many-body quantum lattice systems". In: *SciPost Phys. Lect. Notes* (2019), p. 8. DOI: 10.21468/SciPostPhysLectNotes.8. URL: https://scipost.org/10.21468/SciPostPhysLectNotes.8.

[65] Shi-Ju Ran et al. "Tensor Network Contractions". In: *Lecture Notes in Physics* (2020). ISSN: 1616-6361. DOI: 10.1007/978-3-030-34489-4. URL: http://dx.doi.org/10.1007/978-3-030-34489-4.

[66] Sebastian Paeckel et al. "Time-evolution methods for matrix-product states". In: *Annals of Physics* 411 (2019), p. 167998. ISSN: 0003-4916. DOI: https://doi.org/10.1016/j.aop.2019.167998. URL: https://www.sciencedirect.com/science/article/pii/S0003491619302532.

[67] M. Kliesch, D. Gross, and J. Eisert. "Matrix-Product Operators and States: NP-Hardness and Undecidability". In: *Physical Review Letters* 113.16 (Oct. 2014). ISSN: 1079-7114. DOI: 10.1103/physrevlett.113.160503. URL: http://dx.doi.org/10.1103/PhysRevLett.113.160503.

[68] Frank Verstraete and Juan Ignacio Cirac. "Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions". In: *arXiv: Strongly Correlated Electrons* (2004).

[69] Y.-Y. Shi, L.-M. Duan, and G. Vidal. "Classical simulation of quantum many-body systems with a tree tensor network". In: *Phys. Rev. A* 74 (2 Aug. 2006), p. 022320. DOI: 10.1103/PhysRevA.74.022320. URL: https://link.aps.org/doi/10.1103/PhysRevA.74.022320.

[70] Giuseppe Carleo et al. "Machine learning and the physical sciences". In: *Rev. Mod. Phys.* 91 (4 Dec. 2019), p. 045002. DOI: 10.1103/RevModPhys.91.045002. URL: https://link.aps.org/doi/10.1103/RevModPhys.91.045002.

[71] Giuseppe Carleo and Matthias Troyer. "Solving the quantum many-body problem with artificial neural networks". In: *Science* 355.6325 (2017), pp. 602–606. ISSN: 0036-8075. DOI: 10.1126/science.aag2302. URL: https://science.sciencemag.org/content/355/6325/602.

[72] Roger G. Melko et al. "Restricted Boltzmann machines in quantum physics". In: *Nature Physics* 15.9 (Sept. 2019), pp. 887–892. ISSN: 1745-2481. DOI: 10.1038/s41567-019-0545-1. URL: https://doi.org/10.1038/s41567-019-0545-1.

[73] Or Sharir, Amnon Shashua, and Giuseppe Carleo. *Neural tensor contractions and the expressive power of deep neural quantum states*. 2021. arXiv: 2103.10293 [quant-ph].

[74] Giuseppe Carleo, Yusuke Nomura, and Masatoshi Imada. "Constructing exact representations of quantum many-body systems with deep neural networks". In: *Nature Communications* 9.1 (Dec. 2018). ISSN: 2041-1723. DOI: 10.1038/s41467-018-07520-3. URL: http://dx.doi.org/10.1038/s41467-018-07520-3.

[75] Mohamed Hibat-Allah et al. "Recurrent neural network wave functions". In: *Physical Review Research* 2.2 (June 2020). ISSN: 2643-1564. DOI: 10.1103/physrevresearch.2.023358. URL: http://dx.doi.org/10.1103/PhysRevResearch.2.023358.

[76] Mario Collura et al. "On the descriptive power of Neural-Networks as constrained Tensor Networks with exponentially large bond dimension". In: *SciPost Physics Core* 4.1 (Feb. 2021). ISSN: 2666-9366. DOI: 10.21468/scipostphyscore.4.1.001. URL: http://dx.doi.org/10.21468/SciPostPhysCore.4.1.001.

[77] Michael Bartholomew-Biggs et al. "Automatic differentiation of algorithms". In: *Journal of Computational and Applied Mathematics* 124.1 (2000). Numerical Analysis 2000. Vol. IV: Optimization and Nonlinear Equations, pp. 171–190. ISSN: 0377-0427. DOI: https://doi.org/10.1016/S0377-0427(00)00422-2. URL: https://www.sciencedirect.com/science/article/pii/S0377042700004222.

[78] Hai-Jun Liao et al. "Differentiable Programming Tensor Networks". In: *Phys. Rev. X* 9 (3 Sept. 2019), p. 031041. DOI: 10.1103/PhysRevX.9.031041. URL: https://link.aps.org/doi/10.1103/PhysRevX.9.031041.

[79] F Mezzacapo et al. "Ground-state properties of quantum many-body systems: entangled-plaquette states and variational Monte Carlo". In: *New Journal of Physics* 11.8 (Aug. 2009), p. 083026. DOI: 10.1088/1367-2630/11/8/083026. URL: https://doi.org/10.1088/1367-2630/11/8/083026.

[80] F Mezzacapo and J I Cirac. "Ground-state properties of the spin-1/2 antiferromagnetic Heisenberg model on the triangular lattice: a variational study based on entangled-plaquette states". In: *New Journal of Physics* 12.10 (Oct. 2010), p. 103039. DOI: 10.1088/1367-2630/12/10/103039. URL: https://doi.org/10.1088/1367-2630/12/10/103039.

[81] A. W. Sandvik and G. Vidal. "Variational Quantum Monte Carlo Simulations with Tensor-Network States". In: *Phys. Rev. Lett.* 99 (22 Nov. 2007), p. 220602. DOI: 10.1103/PhysRevLett.99.220602. URL: https://link.aps.org/doi/10.1103/PhysRevLett.99.220602.

[82] Wen-Yuan Liu et al. "Gradient optimization of finite projected entangled pair states". In: *Phys. Rev. B* 95 (19 May 2017), p. 195154. DOI: 10.1103/PhysRevB.95.195154. URL: https://link.aps.org/doi/10.1103/PhysRevB.95.195154.

[83] Juraj Hasik, Didier Poilblanc, and Federico Becca. "Investigation of the Néel phase of the frustrated Heisenberg antiferromagnet by differentiable symmetric tensor networks". In: *SciPost Phys.* 10 (1 2021), p. 12. DOI: 10.21468/SciPostPhys.10.1.012. URL: https://scipost.org/10.21468/SciPostPhys.10.1.012.

[84] R. P. Feynman and Michael Cohen. "Energy Spectrum of the Excitations in Liquid Helium". In: *Phys. Rev.* 102 (5 June 1956), pp. 1189–1204. DOI: 10.1103/PhysRev.102.1189. URL: https://link.aps.org/doi/10.1103/PhysRev.102.1189.

[85] Luca F. Tocchio et al. "Role of backflow correlations for the nonmagnetic phase of the t–t′ Hubbard model". In: *Phys. Rev. B* 78 (4 July 2008), p. 041101. DOI: 10.1103/PhysRevB.78.041101. URL: https://link.aps.org/doi/10.1103/PhysRevB.78.041101.

[86] Di Luo and Bryan K. Clark. "Backflow Transformations via Neural Networks for Quantum Many-Body Wave Functions". In: *Phys. Rev. Lett.* 122 (22 June 2019), p. 226401. DOI: 10.1103/PhysRevLett.122.226401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.226401.

[87] M. B. Hastings. "An area law for one-dimensional quantum systems". In: *J. Stat. Mech. P08024* (2007).

[88] K. E. Schmidt et al. "Structure of the Ground State of a Fermion Fluid". In: *Physical Review Letters* 47.11 (1981). Publisher: American Physical Society, pp. 807–810. DOI: 10.1103/PhysRevLett.47.807. URL: https://link.aps.org/doi/10.1103/PhysRevLett.47.807.

[89] Vincenzo Alba and Pasquale Calabrese. "Entanglement and thermodynamics after a quantum quench in integrable systems". In: *Proceedings of the National Academy of Sciences* 114.30 (2017), pp. 7947–7951. ISSN: 0027-8424. DOI: 10.1073/pnas.1703516114. URL: https://www.pnas.org/content/114/30/7947.

[90] Xiaopeng Li Dong-Ling Deng and S. Das Sarma. "Quantum Entanglement in Neural Network States". In: *Phys. Rev. X 7, 021021* (2017).

[91] Giuseppe Carleo et al. "NetKet: A Machine Learning Toolkit for Many-Body Quantum Systems". In: *SoftwareX* (2019), p. 100311. DOI: 10.1016/j.softx.2019.100311. URL: http://www.sciencedirect.com/science/article/pii/S2352711019300974.

[92] Filippo Vicentini et al. *NetKet 3: Machine Learning Toolbox for Many-Body Quantum Systems*. 2021. arXiv: 2112.10526 [quant-ph].

[93] *NetKet*. https://github.com/netket/netket.

[94] James Bradbury et al. *JAX: composable transformations of Python+NumPy programs*. Version 0.2.5. 2018. URL: http://github.com/google/jax.

[95] Jonathan Heek et al. *Flax: A neural network library and ecosystem for JAX*. Version 0.3.0. 2020. URL: http://github.com/google/flax.

[96] Federico Becca and Sandro Sorella. "Monte Carlo Sampling and Markov Chains". In: *Quantum Monte Carlo Approaches for Correlated Systems*. Cambridge University Press, 2017. DOI: 10.1017/9781316417041.004.

[97] Sandro Sorella. "Green Function Monte Carlo with Stochastic Reconfiguration". In: *Physical Review Letters* 80.20 (May 1998), pp. 4558–4561. ISSN: 1079-7114. DOI: 10.1103/physrevlett.80.4558. URL: http://dx.doi.org/10.1103/PhysRevLett.80.4558.

[98] I. Glasser et al. "Neural-Network Quantum States, String-Bond States, and Chiral Topological States". In: *Phys. Rev. X 8, 011006* (2018).

[99] Giovanni Cataldi et al. "Hilbert curve vs Hilbert space: exploiting fractal 2D covering to increase tensor network efficiency". In: *Quantum* 5 (Sept. 2021), p. 556. ISSN: 2521-327X. DOI: 10.22331/q-2021-09-29-556. URL: http://dx.doi.org/10.22331/q-2021-09-29-556.

[100] Henk W. J. Blöte and Youjin Deng. "Cluster Monte Carlo simulation of the transverse Ising model". In: *Phys. Rev. E* 66 (6 Dec. 2002), p. 066110. DOI: 10.1103/PhysRevE.66.066110. URL: https://link.aps.org/doi/10.1103/PhysRevE.66.066110.

[101] Michael Lubasch, J. Ignacio Cirac, and Mari-Carmen Bañuls. "Algorithms for finite projected entangled pair states". In: *Physical Review B* 90.6 (Aug. 2014). ISSN: 1550-235X. DOI: 10.1103/physrevb.90.064425. URL: http://dx.doi.org/10.1103/PhysRevB.90.064425.

[102] Wen-Jun Hu et al. "Direct evidence for a gapless Z_2 spin liquid by frustrating Néel antiferromagnetism". In: *Phys. Rev. B* 88 (6 Aug. 2013), p. 060402. DOI: 10.1103/PhysRevB.88.060402. URL: https://link.aps.org/doi/10.1103/PhysRevB.88.060402.

[103] Luca Capriotti and Sandro Sorella. "Spontaneous Plaquette Dimerization in the $J_1 - -J_2$ Heisenberg Model". In: *Phys. Rev. Lett.* 84 (14 Apr. 2000), pp. 3173–3176. DOI: 10.1103/PhysRevLett.84.3173. URL: https://link.aps.org/doi/10.1103/PhysRevLett.84.3173.

[104] Kenny Choo, Titus Neupert, and Giuseppe Carleo. "Two-dimensional frustrated J1-J2 model studied with neural network quantum states". In: *Physical Review B* 100.12 (2019). DOI: 10.1103/physrevb.100.125124. URL: http://dx.doi.org/10.1103/PhysRevB.100.125124.

[105] Wen-Yuan Liu et al. "Gapless spin liquid ground state of the spin-$\frac{1}{2}$ $J_1 - J_2$ Heisenberg model on square lattices". In: *Phys. Rev. B* 98 (24 Dec. 2018), p. 241109. DOI: 10.1103/PhysRevB.98.241109. URL: https://link.aps.org/doi/10.1103/PhysRevB.98.241109.

[106] Andrew Lucas. "Ising formulations of many NP problems". In: *Frontiers in Physics* 2 (2014), p. 5.

[107] A. B. Finnila et al. "Quantum annealing: A new method for minimizing multidimensional functions". In: *Chem. Phys. Lett.* 219 (1994), p. 343.

[108] Tadashi Kadowaki and Hidetoshi Nishimori. "Quantum annealing in the transverse Ising model". In: *Phys. Rev. E* 58 (5 Nov. 1998), pp. 5355–5363. DOI: 10.1103/PhysRevE.58.5355.

[109] J. Brooke et al. "Quantum Annealing of a Disordered Magnet". In: *Science* 284 (1999), p. 779.

[110] Giuseppe E. Santoro et al. "Theory of quantum annealing of an Ising spin glass". In: *Science* 295 (2002), p. 2427.

[111] Giuseppe E. Santoro and Erio Tosatti. "Optimization using Quantum Mechanics: Quantum annealing through adiabatic evolution". In: *J. Phys. A: Math. Gen.* 39 (2006), R393–R431. DOI: doi:10.1088/0305-4470/39/36/R01.

[112] E. Farhi et al. "A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem". In: *Science* 292 (2001), p. 472.

[113] Tameem Albash and Daniel A. Lidar. "Adiabatic quantum computation". In: *Rev. Mod. Phys.* 90 (2018), p. 015002.

[114] M. W. Johnson et al. "Quantum annealing with manufactured spins". In: *Nature* 473 (2011), pp. 194–198.

[115] M. Cerezo et al. "Variational quantum algorithms". In: *Nature Reviews Physics* 3 (2021).

[116] M. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[117] R. Barends et al. "Digitized adiabatic quantum computing with a superconducting circuit". In: *Nature* 534 (June 2016), p. 222.

[118] Glen Bigan Mbeng, Luca Arceci, and Giuseppe E. Santoro. "Optimal working point in digitized quantum annealing". In: *Phys. Rev. B* 100 (22 2019), p. 224201. DOI: 10.1103/PhysRevB.100.224201.

[119] Glen Bigan Mbeng, Rosario Fazio, and Giuseppe Santoro. "Quantum Annealing: a journey through Digitalization, Control, and hybrid Quantum Variational schemes". In: *arXiv e-prints*, arXiv:1906.08948 (June 2019), arXiv:1906.08948. arXiv: 1906.08948 [quant-ph].

[120] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. "A Quantum Approximate Optimization Algorithm". In: *arXiv e-prints*, arXiv:1411.4028 (2014), arXiv:1411.4028. arXiv: 1411.4028 [quant-ph].

[121] Troels F. Rønnow et al. "Defining and detecting quantum speedup". In: *Science* 345.6195 (2014), pp. 420–424. DOI: 10.1126/science.1252319.

[122] G. G. Guerreschi and A. Y. Matsuura. "QAOA for Max-Cut requires hundreds of qubits for quantum speed-up". In: *Scientific Reports* 9.1 (2019). DOI: 10.1038/s41598-019-43176-9.

[123] Leo Zhou et al. "Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices". In: *Phys. Rev. X* 10 (2 June 2020), p. 021067. DOI: 10.1103/PhysRevX.10.021067. URL: https://link.aps.org/doi/10.1103/PhysRevX.10.021067.

[124] Jutho Haegeman et al. "Time-Dependent Variational Principle for Quantum Lattices". In: *Phys. Rev. Lett.* 107 (7 Aug. 2011), p. 070601. DOI: 10.1103/PhysRevLett.107.070601. URL: https://link.aps.org/doi/10.1103/PhysRevLett.107.070601.

[125] Nishan Ranabhat and Mario Collura. "Dynamics of the order parameter statistics in the long range Ising model". In: *SciPost Physics* 12.4 (Apr. 2022). DOI: 10.21468/scipostphys.12.4.126. URL: https://doi.org/10.21468%2Fscipostphys.12.4.126.

[126] Guifré Vidal. "Efficient Classical Simulation of Slightly Entangled Quantum Computations". In: *Phys. Rev. Lett.* 91 (14 Oct. 2003), p. 147902. DOI: 10.1103/PhysRevLett.91.147902. URL: https://link.aps.org/doi/10.1103/PhysRevLett.91.147902.

[127] James Dborin et al. "Matrix product state pre-training for quantum machine learning". In: *Quantum Science and Technology* 7.3 (May 2022), p. 035014. DOI: 10.1088/2058-9565/ac7073. URL: https://doi.org/10.1088/2058-9565/ac7073.

[128] E. Miles Stoudenmire and David J. Schwab. "Supervised Learning with Quantum-Inspired Tensor Networks". In: (2016). DOI: 10.48550/ARXIV.1605.05775. URL: https://arxiv.org/abs/1605.05775.

[129] Zhao-Yu Han et al. "Unsupervised Generative Modeling Using Matrix Product States". In: *Physical Review X* 8.3 (July 2018). DOI: 10.1103/physrevx.8.031012. URL: https://doi.org/10.1103%2Fphysrevx.8.031012.

[130] Rishi Sreedhar et al. *The Quantum Approximate Optimization Algorithm performance with low entanglement and high circuit depth*. 2022. DOI: 10.48550/ARXIV.2207.03404. URL: https://arxiv.org/abs/2207.03404.

[131] Shi-Ju Ran et al. *Tensor Network Contractions*. Springer International Publishing, 2020. DOI: 10.1007/978-3-030-34489-4. URL: https://doi.org/10.1007%2F978-3-030-34489-4.

[132] Beatriz Seoane and Hidetoshi Nishimori. "Many-body transverse interactions in the quantum annealing of the p-spin ferromagnet". In: *Journal of Physics A: Mathematical and Theoretical* 45.43 (Oct. 2012), p. 435301. DOI: 10.1088/1751-8113/45/43/435301. URL: https://doi.org/10.1088%2F1751-8113%2F45%2F43%2F435301.

[133] Yuki Susa et al. "Quantum annealing of the p-spin model under inhomogeneous transverse field driving". In: *Phys. Rev. A* 98 (4 Oct. 2018), p. 042326. DOI: 10.1103/PhysRevA.98.042326. URL: https://link.aps.org/doi/10.1103/PhysRevA.98.042326.

[134] Victor Bapst and Guilhem Semerjian. "On quantum mean-field models and their quantum annealing". In: *Journal of Statistical Mechanics: Theory and Experiment* 2012.06 (June 2012), P06007. DOI: 10.1088/1742-5468/2012/06/p06007. URL: https://doi.org/10.1088%2F1742-5468%2F2012%2F06%2Fp06007.

[135] B. Mehlig. *Machine Learning with Neural Networks: An Introduction for Scientists and Engineers*. Cambridge University Press, 2021. ISBN: 9781108849562. URL: https://books.google.it/books?id=DQlIEAAAQBAJ.

[136] J.A. Hertz et al. *Introduction To The Theory Of Neural Computation, Volume I*. Addison-Wesley Computation and Neural Systems Series. Avalon Publishing, 1991. ISBN: 9780201503951. URL: https://books.google.it/books?id=Z4BqAAAAMAAJ.

[137] J J Hopfield. "Neural networks and physical systems with emergent collective computational abilities." In: *Proceedings of the National Academy of Sciences* 79.8 (1982), pp. 2554–2558. DOI: 10.1073/pnas.79.8.2554. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.79.8.2554. URL: https://www.pnas.org/doi/abs/10.1073/pnas.79.8.2554.

[138] W.A. Little. "The existence of persistent states in the brain". In: *Mathematical Biosciences* 19.1 (1974), pp. 101–120. ISSN: 0025-5564. DOI: https://doi.org/10.1016/0025-5564(74)90031-5. URL: https://www.sciencedirect.com/science/article/pii/0025556474900315.

[139] D.O. Hebb. *The Organization of Behavior: A Neuropsychological Theory*. Taylor & Francis, 2005. ISBN: 9781135631901. URL: https://books.google.it/books?id=ddB4AgAAQBAJ.

[140] Daniel J. Amit, Hanoch Gutfreund, and H. Sompolinsky. "Spin-glass models of neural networks". In: *Phys. Rev. A* 32 (2 Aug. 1985), pp. 1007–1018. DOI: 10.1103/PhysRevA.32.1007. URL: https://link.aps.org/doi/10.1103/PhysRevA.32.1007.

[141] Daniel J Amit, Hanoch Gutfreund, and H Sompolinsky. "Statistical mechanics of neural networks near saturation". In: *Annals of Physics* 173.1 (1987), pp. 30–67. ISSN: 0003-4916. DOI: https://doi.org/10.1016/0003-4916(87)90092-3. URL: https://www.sciencedirect.com/science/article/pii/0003491687900923.

[142] Tamas Geszti. *Physical models of neural networks*. World Scientific, 1990.

[143] S. B. Kotsiantis. "Supervised Machine Learning: A Review of Classification Techniques". In: *Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in EHealth, HCI, Information Retrieval and Pervasive Technologies*. NLD: IOS Press, 2007, pp. 3–24. ISBN: 9781586037802.

[144] Carlo Baldassi and Riccardo Zecchina. "Efficiency of quantum vs. classical annealing in nonconvex learning problems". In: *Proceedings of the National Academy of Sciences* 115.7 (2018), pp. 1457–1462. DOI: 10.1073/pnas.1711456115.

[145] Pietro Torta et al. "Quantum Approximate Optimization Algorithm applied to the binary perceptron". In: (2021). DOI: 10.48550/ARXIV.2112.10219. URL: https://arxiv.org/abs/2112.10219.

[146] Steve Abel, Juan C. Criado, and Michael Spannowsky. "Completely quantum neural networks". In: *Physical Review A* 106.2 (Aug. 2022). DOI: 10.1103/physreva.106.022601. URL: https://doi.org/10.1103%2Fphysreva.106.022601.

[147] Krauth, Werner and Mézard, Marc. "Storage capacity of memory networks with binary couplings". In: *J. Phys. France* 50.20 (1989), pp. 3057–3066. DOI: 10.1051/jphys:0198900500200305700.

[148] Changhao Yi. "Success of digital adiabatic simulation with large Trotter step". In: *Phys. Rev. A* 104 (5 Nov. 2021), p. 052603. DOI: 10.1103/PhysRevA.104.052603. URL: https://link.aps.org/doi/10.1103/PhysRevA.104.052603.

[149] Franz Rendl, Giovanni Rinaldi, and Angelika Wiegele. "Solving Max-Cut to Optimality by Intersecting Semidefinite and Polyhedral Relaxations". In: *Math. Programming* 121.2 (2010), p. 307.

[150] John Preskill. "Quantum Computing in the NISQ era and beyond". In: *Quantum* 2 (Aug. 2018), p. 79. ISSN: 2521-327X. DOI: 10.22331/q-2018-08-06-79. URL: https://doi.org/10.22331/q-2018-08-06-79.

[151] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi. *Classical and Quantum Computation*. Graduate studies in mathematics. USA: American Mathematical Society, 2002. ISBN: 0821832298.

[152] P.W. Shor. "Algorithms for quantum computation: discrete logarithms and factoring". In: *Proceedings 35th Annual Symposium on Foundations of Computer Science*. 1994, pp. 124–134. DOI: 10.1109/SFCS.1994.365700.

[153] Luigi Amico et al. "Entanglement in many-body systems". In: *Rev. Mod. Phys.* 80 (2 May 2008), pp. 517–576. DOI: 10.1103/RevModPhys.80.517. URL: https://link.aps.org/doi/10.1103/RevModPhys.80.517.

[154] Pasquale Calabrese and John Cardy. "Entanglement entropy and quantum field theory". In: *Journal of Statistical Mechanics: Theory and Experiment* 2004.06 (June 2004), P06002. DOI: 10.1088/1742-5468/2004/06/P06002. URL: https://dx.doi.org/10.1088/1742-5468/2004/06/P06002.

[155] Jan de Boer, Jarkko Järvelä, and Esko Keski-Vakkuri. "Aspects of capacity of entanglement". In: *Physical Review D* 99.6 (Mar. 2019). DOI: 10.1103/physrevd.99.066012. URL: https://doi.org/10.1103%2Fphysrevd.99.066012.

[156] Marcello Dalmonte et al. "Entanglement Hamiltonians: From Field Theory to Lattice Models and Experiments". In: *Annalen der Physik* 534.11 (Aug. 2022), p. 2200064. DOI: 10.1002/andp.202200064. URL: https://doi.org/10.1002%2Fandp.202200064.

[157] Daniel Gottesman. *Stabilizer Codes and Quantum Error Correction*. 1997. DOI: 10.48550/ARXIV.QUANT-PH/9705052. URL: https://arxiv.org/abs/quant-ph/9705052.

[158] Daniel Gottesman. "Theory of fault-tolerant quantum computation". In: *Phys. Rev. A* 57 (1 Jan. 1998), pp. 127–137. DOI: 10.1103/PhysRevA.57.127. URL: https://link.aps.org/doi/10.1103/PhysRevA.57.127.

[159] Scott Aaronson and Daniel Gottesman. "Improved simulation of stabilizer circuits". In: *Physical Review A* 70.5 (Nov. 2004). DOI: 10.1103/physreva.70.052328. URL: https://doi.org/10.1103%2Fphysreva.70.052328.

[160] Héctor J. García, Igor L. Markov, and Andrew W. Cross. "On the Geometry of Stabilizer States". In: *arXiv preprint arXiv:1711.07848* (2017). DOI: 10.48550/ARXIV.1711.07848. URL: https://arxiv.org/abs/1711.07848.

[161] Simeon Ball, Aina Centelles, and Felix Huber. *Quantum error-correcting codes and their geometries*. 2020. DOI: 10.48550/ARXIV.2007.05992. URL: https://arxiv.org/abs/2007.05992.

[162] Daniel Gottesman. "The Heisenberg Representation of Quantum Computers". In: *arXiv preprint arXiv:9807006* (1998). DOI: 10.48550/ARXIV.QUANT-PH/9807006. URL: https://arxiv.org/abs/quant-ph/9807006.

[163] A.Yu. Kitaev. "Fault-tolerant quantum computation by anyons". In: *Annals of Physics* 303.1 (2003), pp. 2–30. ISSN: 0003-4916. DOI: https://doi.org/10.1016/S0003-4916(02)00018-0. URL: https://www.sciencedirect.com/science/article/pii/S0003491602000180.

[164] Mark Howard et al. "Contextuality supplies the 'magic'for quantum computation". In: *Nature* 510.7505 (2014), pp. 351–355. DOI: 10.1038/nature13460. URL: https://doi.org/10.1038/nature13460.

[165] James R. Seddon et al. "Quantifying Quantum Speedups: Improved Classical Simulation From Tighter Magic Monotones". In: *PRX Quantum* 2 (1 Mar. 2021), p. 010345. DOI: 10.1103/PRXQuantum.2.010345. URL: https://link.aps.org/doi/10.1103/PRXQuantum.2.010345.

[166] Lorenzo Leone et al. "Quantum Chaos is Quantum". In: *Quantum* 5 (May 2021), p. 453. ISSN: 2521-327X. DOI: 10.22331/q-2021-05-04-453. URL: https://doi.org/10.22331/q-2021-05-04-453.

[167] Kanato Goto, Tomoki Nosaka, and Masahiro Nozaki. *Chaos by Magic*. 2021. DOI: 10.48550/ARXIV.2112.14593. URL: https://arxiv.org/abs/2112.14593.

[168] Zi-Wen Liu and Andreas Winter. "Many-Body Quantum Magic". In: *PRX Quantum* 3 (2 May 2022), p. 020333. DOI: 10.1103/PRXQuantum.3.020333. URL: https://link.aps.org/doi/10.1103/PRXQuantum.3.020333.

[169] Mark Howard and Earl Campbell. "Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing". In: *Physical Review Letters* 118.9 (Mar. 2017). DOI: 10.1103/physrevlett.118.090501. URL: https://doi.org/10.1103%2Fphysrevlett.118.090501.

[170] Tobias Haug and M.S. Kim. "Scalable Measures of Magic Resource for Quantum Computers". In: *PRX Quantum* 4 (1 Jan. 2023), p. 010301. DOI: 10.1103/PRXQuantum.4.010301. URL: https://link.aps.org/doi/10.1103/PRXQuantum.4.010301.

[171] Markus Heinrich and David Gross. "Robustness of Magic and Symmetries of the Stabiliser Polytope". In: *Quantum* 3 (Apr. 2019), p. 132. DOI: 10.22331/q-2019-04-08-132. URL: https://doi.org/10.22331%2Fq-2019-04-08-132.

[172] Lorenzo Leone, Salvatore F. E. Oliviero, and Alioscia Hamma. "Stabilizer Rényi Entropy". In: *Phys. Rev. Lett.* 128 (5 Feb. 2022), p. 050402. DOI: 10.1103/PhysRevLett.128.050402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.128.050402.

[173] Salvatore F. E. Oliviero, Lorenzo Leone, and Alioscia Hamma. "Magic-state resource theory for the ground state of the transverse-field Ising model". In: *Phys. Rev. A* 106 (4 Oct. 2022), p. 042426. DOI: 10.1103/PhysRevA.106.042426. URL: https://link.aps.org/doi/10.1103/PhysRevA.106.042426.

[174] Tobias Haug and Lorenzo Piroli. "Quantifying nonstabilizerness of matrix product states". In: *Phys. Rev. B* 107 (3 Jan. 2023), p. 035148. DOI: 10.1103/PhysRevB.107.035148. URL: https://link.aps.org/doi/10.1103/PhysRevB.107.035148.

[175] E M Stoudenmire and Steven R White. "Minimally entangled typical thermal state algorithms". In: *New Journal of Physics* 12.5 (May 2010), p. 055026. DOI: 10.1088/1367-2630/12/5/055026.

[176] Andrew J. Ferris and Guifre Vidal. "Perfect sampling with unitary tensor networks". In: *Phys. Rev. B* 85 (16 Apr. 2012), p. 165146. DOI: 10.1103/PhysRevB.85.165146. URL: https://link.aps.org/doi/10.1103/PhysRevB.85.165146.

[177] Marton Kormos et al. "Real-time confinement following a quantum quench to a non-integrable model". In: *Nature Physics* 13.3 (Nov. 2016), pp. 246–249. DOI: 10.1038/nphys3934. URL: https://doi.org/10.1038%2Fnphys3934.

[178] Tobias Haug and Lorenzo Piroli. *Stabilizer entropies and nonstabilizerness monotones*. 2023. arXiv: 2303.10152 [quant-ph].

[179] Glen Bigan Mbeng, Angelo Russomanno, and Giuseppe E. Santoro. *The quantum Ising chain for beginners*. 2020. DOI: 10.48550/ARXIV.2009.09208. URL: https://arxiv.org/abs/2009.09208.

[180] Pasquale Calabrese, Fabian H L Essler, and Maurizio Fagotti. "Quantum quench in the transverse field Ising chain: I. Time evolution of order parameter correlators". In: *Journal of Statistical Mechanics: Theory and Experiment* 2012.07 (July 2012), P07016. DOI: 10.1088/1742-5468/2012/07/P07016. URL: https://dx.doi.org/10.1088/1742-5468/2012/07/P07016.

[181] Nishan Ranabhat and Mario Collura. *Thermalization of long range Ising model in different dynamical regimes: a full counting statistics approach*. 2022. DOI: 10.48550/ARXIV.2212.00533. URL: https://arxiv.org/abs/2212.00533.

[182] Pasquale Calabrese and John Cardy. "Time Dependence of Correlation Functions Following a Quantum Quench". In: *Phys. Rev. Lett.* 96 (13 Apr. 2006), p. 136801. DOI: 10.1103/PhysRevLett.96.136801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.96.136801.

[183] Guifré Vidal. "Efficient Simulation of One-Dimensional Quantum Many-Body Systems". In: *Phys. Rev. Lett.* 93 (4 July 2004), p. 040502. DOI: 10.1103/PhysRevLett.93.040502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.93.040502.

[184] Davide Rattacaso et al. *Stabilizer entropy dynamics after a quantum quench*. 2023. arXiv: 2304.13768 [quant-ph].

[185] Haiping Huang, K Y Michael Wong, and Yoshiyuki Kabashima. "Entropy landscape of solutions in the binary perceptron problem". In: *Journal of Physics A: Mathematical and Theoretical* 46.37 (Aug. 2013), p. 375002. DOI: 10.1088/1751-8113/46/37/375002. URL: https://doi.org/10.1088%2F1751-8113%2F46%2F37%2F375002.

[186] Richard Kueng and David Gross. *Qubit stabilizer states are complex projective 3-designs*. 2015. arXiv: 1510.02767 [quant-ph].

[187] Matthew B. Elliott. *Stabilizer states and local realism*. 2008. DOI: 10.48550/ARXIV.0807.2876. URL: https://arxiv.org/abs/0807.2876.

[188] Lorenzo Leone, Salvatore F. E. Oliviero, and Alioscia Hamma. "Stabilizer Renyi Entropy". In: *Phys. Rev. Lett.* 128 (5 Feb. 2022), p. 050402. DOI: 10.1103/PhysRevLett.128.050402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.128.050402.

[189] Emanuele Tirrito et al. *Quantifying non-stabilizerness through entanglement spectrum flatness*. 2023. arXiv: 2304.01175 [quant-ph].