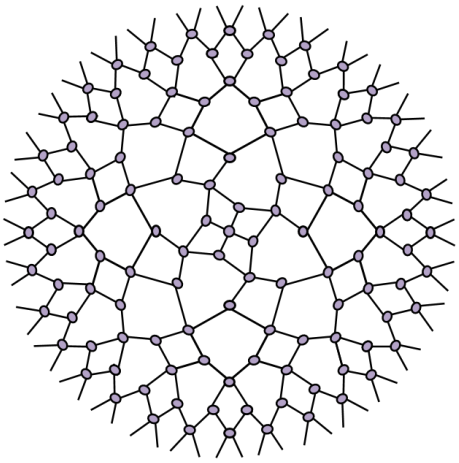
University of Groningen
Faculty of Mathematics and Natural Sciences
Van Swinderen Institute of Particle Physics and Gravity

Master Thesis
Physics
2016-2017

The Multiple Entanglement Renormalization Ansatz and the Anti-de-Sitter/Conformal Field Theory Correspondence



Author:
J.C. Borger
s2035022

Supervisor:
Prof. K. Papadodimas
Second corrector:
Prof. E.A. Bergshoeff

Abstract

The AdS/CFT correspondence offers a new way of studying gravity. In this correspondence it is not clear how the gravitational bulk physics emerges from the boundary CFT physics. Recently it was proposed that tensor networks could give more insight into this problem. A specific tensor network called a MERA resembles a discrete version of a spatial slice in the AdS/CFT correspondence. From a critical state it creates an extra spatial dimension resembling a spatial AdS slice. In this thesis this resemblance is further studied. Then two tensor network methods, TRG and TNR, are used to study thermal states. These states are of interest because a thermal CFT state has a black hole as its gravitational dual. It will turn out that TRG will break down at a certain scale for critical states while TNR is better able to study these systems, as expected.

Contents

Chapter 1

Introduction

Since the inception of quantum mechanics and general relativity at the start of last century, theoretical physics has struggled to understand gravity in a quantum mechanical framework. The Anti-de-Sitter/Conformal Field Theory correspondence offers a new way of studying this problem. It is a statement about the equivalence of two seemingly unrelated theories, one describing gravity in Anti-de-Sitter space (AdS) and one describing a Conformal Field Theory (CFT), without gravity. This equivalence makes it possible to understand gravity in a completely different setting namely on the CFT side of the equivalence. The AdS/CFT correspondence has been a hot topic ever since its proposition and although it is technically still a conjecture it is widely accepted and, globally at least, fairly well understood. It is however completely not understood how the gravitational AdS physics emerge (locally) from the CFT physics. Recently it has been proposed that tensor networks may play a role in this emergence. In this thesis tensor networks are examined and the similarity with AdS/CFT is studied. This research aims to answer the question: "How is AdS/CFT realized in the MERA network?" Later more general tensor network methods are used to calculate certain thermal properties of a CFT which are of interest in the context of AdS/CFT because a thermal CFT is dual to a black hole.

Thinking about AdS/CFT in terms of tensor networks can help us better understand gravity. A better understanding of gravity is of enormous academic interest. Moreover it is necessary to understand the Big Bang or black holes, thus it is not purely an academical question.

This thesis is structured in seven chapters. The first chapter is this introduction. In the second chapter the motivation behind this thesis is explained in more detail. The third chapter contains a basic introduction to CFT's, some specifics of the two-dimensional Ising CFT and a short treaty of the thermofield double state. The fourth chapter introduces tensor networks. In the fifth chapter the tensor network methods that were used during this research are treated. In chapter six the results are given and chapter seven contains the conclusion.

To fully understand this thesis some prior knowledge in quantum mechanics and quantum field theory is required. Knowledge about CFT's and AdS/CFT is desired. For the layman

there is a summary in Dutch in Appendix D.
Enjoy!

Chapter 2

Motivation

2.1 Gravity

In 1687 Isaac Newton published his seminal *Philosophiae Naturalis Principia Mathematica*. In this book he proposed the universal law of gravitation among other things. This law states that every massive object exerts a force on every other massive object proportional to the product of the masses divided by the distance between them squared. The motion of planets and the falling of apples suddenly could be explained by the same principle. Yet this law was phenomenological in nature as Newton himself pointed out:

> Thus far I have explained the phenomena of the heavens and of our sea by the force of gravity, but I have not yet assigned a cause to gravity. -Isaac Newton [1]

His law could describe a number of phenomena yet it did not become clear why the phenomena occured in this way.

In 1915 Albert Einstein published his theory of general relativity. In this theory space and time should be thought of as being part of spacetime. The dynamic relation between massive objects moving in a spacetime and the spacetime itself are given by the Einstein field equations. A laymans exlanation is often given by imagining a trampoline. If we would roll a small marble over the trampoline it would move in a straight line. Now suppose we put a heavy ball in the middle of the trampoline. If we now roll the marble over the trampoline, its trajectory will be diverted since the shape of the trampoline changed because of the heavy ball. Such an interplay also happens between heavy objects and spacetime and an apparent attraction is in fact just the moving of objects along geodesic lines in that spacetime.

Also in the early 1900's quantum mechanics came into play and it became apparent that Nature fundamentally works quantum mechanically. From then on people have tried to understand gravity in a quantum mechanical way, however this has not been succesfully done.

2.2 Hints from Black Holes

By combined efforts of Bekenstein and Hawking in the 1970's it became apparent that black holes have entropy proportional to the area of their event horizon. This is a strange property as can be demonstrated by thinking about a gas in some volume. The maximal entropy is proportional to the number of degrees of freedom needed to give a complete description of the system. The entropy is a measure of the amount of microscopic states that correspond to the same macroscopic state. Degrees of freedom are thought to be local and therefore it is clear then that in a larger volume the molecules have more possible configurations that correspond to the same macroscopic state. The entropy is thus proportional to the volume, which is very typical.

Now suppose we consider a volume with so many molecules that a large black hole will form due to gravitational collapse. Before the large black hole is formed the entropy is proportional to the volume. The moment the large black hole is formed the entropy suddenly becomes proportional to the area. For a large enough black hole the volume entropy is larger than the area entropy and thus the entropy will decrease significantly. However the second law of thermodynamics states that entropy cannot decrease.

The most natural explanation of this puzzle is that there must have been a theory the whole time in which a full description was proportional to the area of the system. This is known as the holographic principle. A gravitational theory, a theory that can make black holes, should admit a description that is proportional to the area of a system. This means that if we think of the theory in its fundamental degrees of freedom it is actually living in a dimension lower and that these degrees of freedom are not local.

Just as a hologram is a number of carefully tuned lights in a two-dimensional array that is observed as if being three-dimensional, the world is actually (2+1) dimensional yet we oberserve it as being (3+1) dimensional with gravity.

2.3 AdS/CFT

The AdS/CFT is a realization of this holographic principle. Is was first proposed by Maldacena in 1997 [2]. It is a statement about the equivalence of two theories that are seemingly unrelated.

On the one side of the duality is a string theory that describes gravity in D+1 dimensional anti-de-Sitter space. This is a maximally symmetric Lorentzian manifold with negative scalar curvature.

On the other side of the duality is a D dimensional conformal field theory which is a quantum field theory with an additional symmetry. For instance a certain type of Yang-Mills theory that describes strongly coupled elementary particles.

The duality makes it possible to study previously not understood phenomena by translating them to the other side of the duality and studying them there. One example of this is the calculation of the viscosity of quark-gluon plasma by translating this problem to the

gravitational side and solving it there [3]. It can also help to better understand gravity as it allows the translation of gravity to the CFT side, where it can be studied.

The duality is mostly pictured as a boundary that represents the D dimensional CFT and a bulk that represents the D+1 dimensional gravitational side.

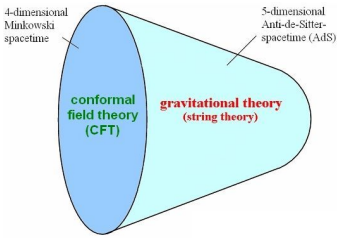


Figure 2.1: Depiction of the AdS/CFT conjecture with a CFT on the boundary and a gravitational theory in the bulk. Figure taken from [35].

This duality has given us a new way to study gravity and is still a very hot topic although it is not yet fully understood. One thing that is completely not understood is the way the gravitational physics in the bulk emerges from the CFT physics in the boundary. It is believed that the scale of phenomena in the boundary is related to depth of phenomena in the bulk. Meaning that things happening on a small scale in the boundary are dual to phenomena happening close to the boundary and phenomena happening at large scales in the boundary are dual to phenomena happening deep in the bulk, as shown in Figure 2.2.

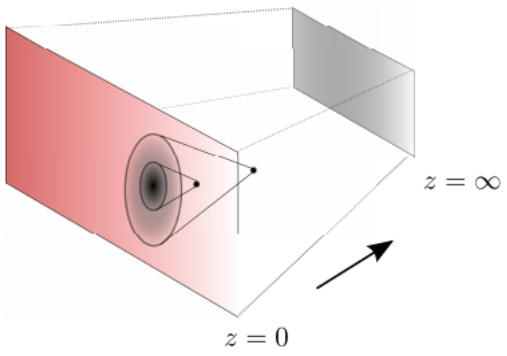


Figure 2.2: Depiction of the idea that scale is related to depth in the bulk. Figure taken from [36].

The original AdS/CFT article and some other relevant articles can be found here [2],[4],[5],[6].

2.3.1 Ryu-Takayanagi Proposal

Another insight into the workings of AdS/CFT is given by the recent Ryu-Takayanagi proposal [7]. This proposal states that the entanglement entropy of a subsystem in the

boundary with the rest of the boundary system, is proportional to the minimal area given by "dipping" into the bulk.

In classical physics the only reason to describe a system by a density matrix is lack of information. In quantum physics there is an additional reason. Suppose a system defined on area A and B is in a pure state with wavefunction $\psi(\alpha, \beta)$ where α and β are complete sets of commutating observables in subsystem α and β respectively. The state on subsystem A can be described by the reduced density matrix on A

$$\rho_A = \sum_{\beta} \psi^*(\alpha, \beta)\psi(\alpha', \beta)$$

.

For a factorizable state $\psi(\alpha, \beta) = \phi(\alpha)\chi(\beta)$ it is possible that the state on A is also pure. Thus for an entangled state there is a certain "entropy" due to the entanglement between subsystems.

$$S = -\mathrm{Tr}[\rho_A \log(\rho_A)],$$

this value is the same for ρ_B as long as the complete state is pure.

The Ryu-Takayanagi proposal states that the entanglement entropy of a certain subsystem in the boundary is given as

$$S_A = Area/4G_N \tag{2.1}$$

where G_N is Newtons gravitational constant and the area is given by the minimal area that can enclose subsystem A. For a graphical depiction see Figure 2.3. The area has a certain extension in the bulk because the bulk is hyperbolic and thus it can be area-efficient to extend into the bulk.

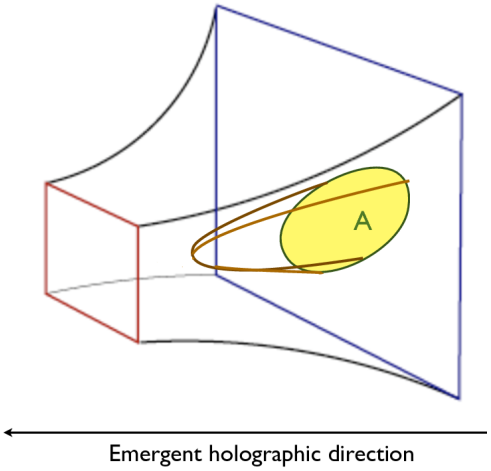


Figure 2.3: The entanglement entropy of the region in the boundary is given by the minimal area dipping into the bulk.

The Ryu-Takanayagi proposal is well established and tells us that entanglement plays an important role in going from the boundary to the bulk.

2.4 MERA as discrete AdS/CFT

In a paper from 2009, Swingle made a connection between tensor network methods, as were being used in condensed matter physics, and holography [8]. In particular he noticed that methods that used a real-space renormalization group flow could be thought of as representing a generalized notion of holography as it exists in efforts to describe quantum gravity. In other words he noticed a similarity between these tensor network methods and AdS/CFT.

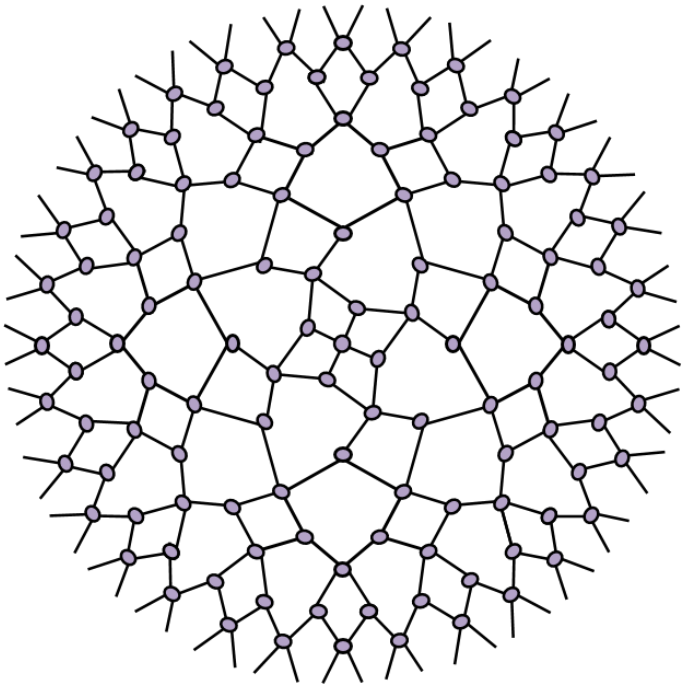


Figure 2.4: A MERA tensor network representing a quantum mechanical state defined on a circle. Figure taken from [34].

These methods can describe a quantum mechanical state by reorganizing it according to scale, the structure in which this information is reorganized can be thought of as representing a discrete extra spatial dimension. Representations of states with a finite correlation length give rise to a simple geometry. Representations of states with infinite correlation length give rise to an extra hyperbolic spatial dimension resembling empty Anti-de-Sitter space. An example of such a representation is given in Figure 2.4 where the similarity with a spatial AdS slice is apparent. States with an infinite correlation length are critical states which are the discrete equivalent of conformal field theories. Thus this particular procedure creates a discrete space resembling Anti-de-Sitter space starting from

a critical state, which is the discrete equivalent of a CFT.

These tensor network methods could explain how the Anti-de-Sitter physics in the bulk emerge from the CFT physics in the boundary. Or at least these methods can be studied as a toy model to increase our understanding of this emergence.

2.5 Entanglement growth due to time evolution of black hole interior

In a paper from 2013 Maldacena and Hartman [9] consider a thermofield double state which represents two spaces connected by a black hole. They then cut the complete system in two (through the black hole) and compute the entanglement entropy between the two cuts of the complete system. Then time evolution is turned on and the minimal surface going through the black hole keeps growing in size with time. For the calculation or further details see Appendix A or the original paper [9].

In the paper the authors propose that this growth can be understood from a tensor network point of view. If we represent the thermofield double state in tensor network form and then turn on time evolution such a growth is exactly what one would expect.

Tensor network methods are being used to study the inside of black holes and to better understand the emergence of spacetime in AdS/CFT. Hence they are interesting to study in the context of quantum gravity.

Chapter 3

Conformal Field Theories

3.1 Basics

A conformal field theory is a quantum field theory that is invariant under a conformal transformation. If $g_{\mu\nu}$ is the space-time metric then a conformal transformation is an invertible mapping $x \to x'$ that leaves the metric invariant up to a scale,

$$g'_{\mu\nu}(x') = \Lambda(x) g_{\mu\nu}(x). \tag{3.1}$$

These transformations form a group which has the Poincaré group as a subgroup, as these correspond to the special case $\Lambda(x) = 1$. These transformations are called conformal because they leave angles invariant, hence preserving 'form'. Apart from the regular translations and rotations, conformal transformations contain dilations and the special conformal transformation, respectively given as

$$x \to x'^{\mu} = \alpha x^{\mu}, \qquad x \to x'^{\mu} = \frac{x^{\mu} - b^{\mu} x^2}{1 - 2bx + b^2 x^2}. \tag{3.2}$$

The extra conditions due to the conformal symmetry constrain the theory more than a regular QFT. For instance, under a conformal transformation, a spinless field $\phi(x)$ transforms like

$$\phi(x) \to \phi'(x') = |\frac{\partial x'}{\partial x}|^{-\Delta/d} \phi(x), \tag{3.3}$$

where d is the space-time dimension and Δ is the conformal dimension of the field. This is an eigenvalue of the dilation operator working on the field, thus a scalar that corresponds to a field. P_μ and K_μ are the generators of the translations and of the special conformal transformations, these act as ladder operators for the eigenvalues of the dilation operator D_μ, on the irreducible representations of the transformations, the fields. Thus there are fields that are annihilated by the lowering ladder operator, these fields are called primary fields. From these primary fields other fields can be obtained by acting on the primary with the raising operator, these field are thus related to the primary fields and are known as descendant fields.

In two dimensions the conformal algebra can be extended to an infinite dimensional algebra known as the Virasoro algebra:

$$\begin{aligned} [L_n, L_m] &= (n-m)L_{n+m} + \frac{c}{12} n(n^2-1)\delta_{n+m,0} \\ [L_n, \bar{L_m}] &= 0 \\ [\bar{L_n}, \bar{L_m}] &= (n-m)\bar{L_{n+m}} + \frac{c}{12} n(n^2-1)\delta_{n+m,0} \end{aligned} \tag{3.4}$$

where L are the Virasoro generators and c is a number called the central charge. The number c is the only parameter in this definition hence a certain value of c uniquely defines a conformal field theory. The central charge can be thought of as an extensive measure of the number of degrees of freedom in a theory. A special class of two-dimensional CFT's is the one where c is smaller than one. The theories in this class are known as minimal models. The special property of this class is that they have a finite number of primary fields.

More information regarding CFT's can be found in [10].

In the next section a minimal model will be presented known as the two-dimensional Ising CFT.

3.2 Two-dimensional Ising CFT

The Ising model is a statistical model where a number of spins are put on a lattice, only nearest neighbour interactions are taken into account and there possibly is an external magnetic field.

The two-dimensional Hamiltonian without external magnetic coupling is given below,

$$H = -\sum_{i=1}^{N-1}\sum_{j=1}^{N} S_{i,j}S_{i+1,j} - \sum_{i=1}^{N}\sum_{j=1}^{N-1} S_{i,j}S_{i,j+1} \qquad \text{D} = 2. \tag{3.5}$$

For typical temperatures, two-point correlation functions decay exponentially fast with the distance between their two points. For a special temperature, the critical temperature, two-point correlations depend on the distance between their points like $|x_1 - x_2|^{-c}$ which is typical of a critical system. At criticality the Ising model is described by a CFT in the continuum limit. By performing a Jordan-Wigner transformation the action of this two-dimensional CFT can be written as

$$S = \int d^2z(\psi\bar{d}\psi + \bar{\psi}d\bar{\psi}), \tag{3.6}$$

where ψ is a fermion. This is a minimal model with a central charge of $\frac{1}{2}$. Since it is a minimal model it has only finite primary fields and the primary fields of the two-dimensional Ising CFT are I, σ, ϵ which are classified by the conformal dimensions $\Delta_I = 0, \Delta_\sigma = \frac{1}{8}, \Delta_\epsilon = 1$. For more details see di Francesco [10].

3.3 Eternal black hole and Thermofield double state

An eternal black hole consists of two space-time regions that are connected via a black hole with a future singularity and a white hole with a past singularity. It is a idealized because an actual black hole formed by gravitational collapse only has one external space-time region and no past singularity. The eternal black hole can be represented in a Penrose diagram as shown in Figure 3.1 .

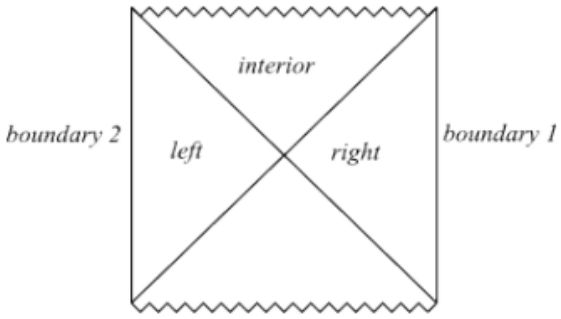


Figure 3.1: Penrose diagram of an eternal black hole.

In quantum mechanics a thermal state is described by the mixed state with density matrix $\rho = e^{-\beta H}$. A mixed state can always be thought of as being part of a larger system that in its totality is pure. The reduced density matrix of the actual system then gives us the original density matrix. This is known as purification. For the thermal state the whole system is doubled. For a QFT this means that for every field there exists another field in a different space. If the eigenstates of the Hamiltonian are $|m\rangle$ and $|n\rangle$, the states of the doubled system are $|m\rangle\,|n\rangle$. Now we consider the following state of the double system

$$|TFD\rangle = \frac{1}{\sqrt{Z(\beta)}} \sum_n e^{-\beta E_n/2}\,|n_1\rangle\,|n_2\rangle \tag{3.7}$$

also known as the thermofield double state. Here $Z(\beta)$ is the partition function. It is easy to check that the reduced density matrix on either one of the systems gives the thermal density matrix. This is a trick to study thermal behaviour of certain systems.

In [11] Maldacena proposed that the eternal black hole is the AdS equivalent of two CFT's that together describe the thermofield double state. Each CFT exists on the asymptotic boundary of a space-time region. The AdS/CFT equivalent of the special thermofield double state of these two CFT's is given by two AdS spaces that are connected via a black hole.

The eternal black hole is interesting to study in the context of the 'information paradox of black holes' which can be rephrased in the question of the smoothness of the horizon, also known as the firewall paradox. The horizon of the eternal black hole is smooth, as can be understood from for instance the calculation of Maldacena and Hartman [9] where the black hole has an interior. The eternal black hole is a very atypical black hole as

there is a very specific entanglement between the two sides. A more typical black hole can be made by acting on the eternal black hole with a time evolution operator. It turns out that it should be possible to do this in the tensor network picture. Thus in light of the information paradox it would be interesting to look at the smoothness of a black hole obtained from time evolving an eternal black hole, in the tensor network picture. For an introduction on the information paradox see [12], [13].

Chapter 4

Tensor Networks

4.1 Variational Principle

Suppose there is a Hamiltonian with an unknown ground state. The variational principle gives an upper bound on this ground state in the following way: An arbitrary normalized state $|\psi\rangle$ can be written as a linear combination of the orthonormal eigenvectors of the Hamiltonian.

$$|\psi\rangle = \sum_{n \in Spec\{H\}} c_n \psi_n \tag{4.1}$$

Computing the expectation value of the Hamiltonian then gives

$$\langle H \rangle = \langle \sum_m c_m \psi_m | H \sum_n c_n \psi_n \rangle = \sum_m \sum_n c_m^* E_n c_n \langle \psi_m | \psi_n \rangle = \sum_n E_n |c_n|^2 \tag{4.2}$$

The ground state is by definition the smallest eigenstate so that $E_{gs} \leq E_n$, and therefore

$$\langle H \rangle \geq E_{gs} \sum_n |c_n|^2 = E_{gs} \tag{4.3}$$

Normally the proposed state depends on a set of parameters $\{\alpha_i\}$ and minimizing the expectation value for the Hamiltonian with respect to these parameters gives the best bound on the ground state.
Varying over the whole Hilbert space will yield the exact ground state. However in most cases of interest this computation is extremely complicated. Therefore often a subspace is varied over. The choice for a specific subspace, and thus the set of proposed states, is called an "ansatz" which is german for 'approach'. Some ansätze approximate the ground state very well and the choice of ansatz is therefore very important.

A big problem in quantum many-body physics is that the Hilbert space of a system grows exponentially in its size. For example: the dimension of the Hilbert space of a one-dimensional chain of N spin 1/2 particles grows as 2^N. Exactly solving such a system

quickly becomes computationally hard and calculating the behavior of a macroscopical piece of matter with $N = \mathcal{O}(10^{23})$ becomes virtually impossible.

It is not straightforward to find an ansatz that reduces the Hilbert space tremendously yet gives a good approximation to the ground state. It is exactly in this context that tensor networks arise: as a variational ansatz that reduce the computational effort tremendously yet give a tight bound on the ground state for a certain class of physical systems.

4.2 Graphical Tensor Notation

In the context of Tensor Network States, tensors and tensor networks are often represented in a graphical way, originally proposed by Penrose[14]. A rank (n,m) tensor is represented by a shape with n legs going up and m legs going down, different shapes are used to distinguish between different tensors. Indices can be contracted by joining an upgoing leg with a downgoing leg. This graphical representation of tensors is very similar to the index notation of tensors ($R^d_{abc}, T_{\mu\nu}$) for instance: an expression with no external legs, no indices, is a number. An upgoing leg can be changed into a downgoing leg by contracting it with the metric, just as the metric raises or lowers indices in (pseudo)Riemannian geometry. This graphical notation is a tool to visualize multilinear mappings in an intuitive way. Since this can best be shown graphically Figure 4.1 hopefully explains a lot.

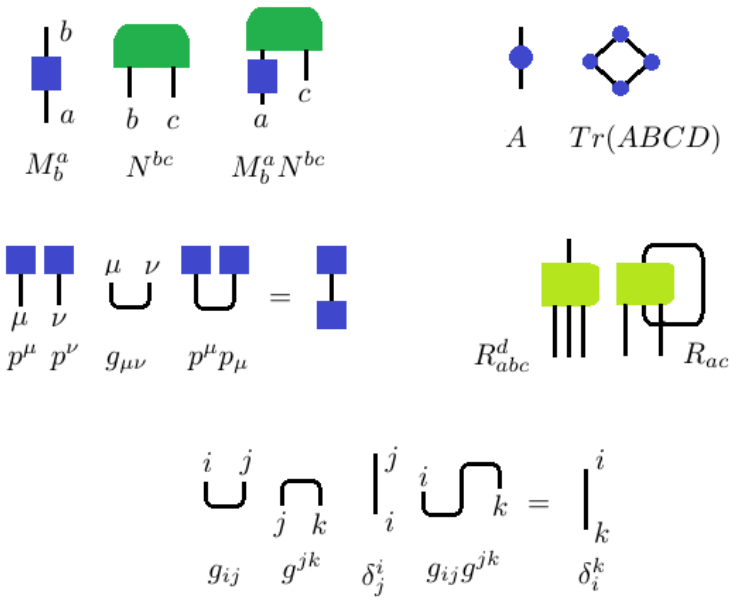


Figure 4.1: Some examples of the graphical tensor notation.

4.3 Tensor Network States

Using Dirac's 'braket' notation, a vector $|\phi\rangle$ is a (1,0) tensor and a dual vector $\langle\phi|$ is a (0,1) tensor. A N-particle state can then be respresented as a (N,0) vector consisting of N single particle states in the following way:

$$|\psi\rangle = \sum_{i_1=1}^{d} \sum_{i_2=1}^{d} \cdots \sum_{i_N=1}^{d} C_{i_1,i_2,\cdots,i_N} |\psi_{i_1}\rangle |\psi_{i_2}\rangle \cdots |\psi_{i_N}\rangle \tag{4.4}$$

If the Hilbert space of the single particle states is d the Hilbert space of the N-particle state has dimension d^N. In order to describe the complete N-particle state, d^N coefficients need to be specified.

The single particle states can also be connected via a tensor or a multilinear combination, a network, of tensors. For example in the following way:

$$|\psi\rangle = T^{1\cdots n} |\phi_1\rangle \cdots |\phi_n\rangle = \Xi^{1\cdots n a\cdots m} \Theta_{a\cdots m} |\phi_1\rangle \cdots |\phi_n\rangle . \tag{4.5}$$

The structure of the tensor network then describes the inner workings of the N-particle state. If a typical tensor in this network has p indices that can take χ values, $\mathcal{O}(N\chi^p)$ coefficients need to be specified in order to describe the complete state. Letting χ be exponential in N the whole Hilbert space of the state in equation (4.4) is covered. This would however give an equally complicated computation since again $\mathcal{O}(e^N)$ coefficients need to be specified. The power of tensor network states is that for some physical systems, a finite value of χ, (5-10) together with a certain network structure is enough to give a good approximation of a N-particle state.

Note that $\chi = 1$ corresponds to a N-particle state which is completely factorizable in terms of single particle states and thus corresponds to a state without entanglement. χ then can be thought of as the amount of entanglement that is permitted in the system.

A tensor network is useful if it can represent a N-particle wavefunction efficiently and if it is possible to efficiently extract information from it. This puts certain limitations on the possible network structures that can be formed.

The left tensor network in Figure 4.3 is exactly the state described in equation (4.4). Since d^N coefficients need to be specified, a computation of the N-particle state $|\psi\rangle$ scales as d^N, exponential in the scale of the system. The tensor network in the middle of Figure 4.3 consists of $N \times N$ tensors. The computation of the wavefunction then scales with the size N as $\mathcal{O}(N^2)$. The tensor network representation of the N-particle state on the right of Figure 4.3 consists of N tensors. The computation of this wavefunction scales with N as $\mathcal{O}(N)$.

In quantum mechanics physical quantities are calculated by expectation values of observables (Hermitian operators). A tensor network is useful if these expectation values are easily evaluated. This is not the case for the left state as given in Figure 4.3. This state is not efficiently described and this does not change when a $\langle bra|$ tensor network state is added. Expectation values for this state are therefore not efficiently evaluated. In Figure

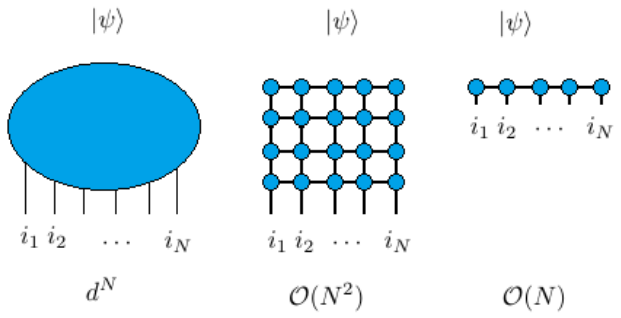


Figure 4.2: Three tensor networks representing a N-particle wavefunction $|\psi\rangle$ from N single site wavefunctions ϕ_{i_n} and associated computation cost of the wavefunction.

4.3 expectation values are shown for some local operator O at location x_i for the other two example states. To get the expectation value all tensors need to be contracted. For the

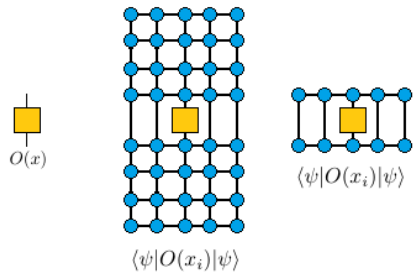


Figure 4.3: A local operator and its expectation value for two different states.

$N \times N$ state there is no efficient way to do this. For instance, naively starting to contract the tensors in the upper left corner, the number of legs of the upper left tensor quickly grow. It can be shown that the computation of this expectation value is exponential in the system size N and thus this tensor network state is not useful. The expectation value of the right tensor network state in Figure 4.3 can however be evaluated polynomial in N. With N contractions the upper tensors can be contracted with the lower tensors and then in again N-1 contractions all tensors can be contracted to a number. The computation is proportional to N^2 and thus is polynomial in N. Such a state can be efficiently evaluated and expectation values can be efficiently calculated using this state. It is thus a useful tensor network. This state, also shown in the right of Figure 4.3, with periodic boundary conditions can be written as:

$$M_{b_2}^{a_1 b_1} M_{b_3}^{a_2 b_2} \cdots M_{b_1}^{a_N b_N} |\psi_{a_1}\rangle |\psi_{a_2}\rangle \cdots |\psi_{a_N}\rangle \tag{4.6}$$

since this is just a product of matrices this state is known as the Matrix Product State which will be discussed in the next section.

We have seen that not all possible tensor network states are useful. In the following sections we will treat two useful tensor network states.

4.4 Matrix Product State

Figure 4.4: A Matrix Product State in graphical tensor notation.

The Matrix Product State or MPS ansatz has a long and rich history. In retrospect it was first used in 1975 by Kenneth Wilson in his Numerical Renormalization Group[15] which is a technique to deal with quantum many body systems where impurities play a role. Wilson used this to solve the Kondo problem. It can be shown that the NRG procedure actually leads to a Matrix Product State.[16] In 1992 Steven White invented the Density Matrix Renormalization Group that also uses Matrix Product States[17]. This is a variational technique to estimate the low energy behavior of some quantum many body systems. For one-dimensional systems it is still the most efficient method[18].

In the previous section it was shown that the MPS is a useful tensor network. This is however no statement about its physical relevance. Since the MPS state lives in a drastically smaller Hilbert space then the actual state one might expect the MPS state to be oversimplified and thus to not be physically relevant. It turns out that this is not the case and that there are certain classes of systems that can be accurately described by a MPS state. There is no theorema that states which kind of systems can be accurately described by MPS states[18] but several years of numerical evidence lead to the class of one-dimensional gapped local Hamiltonians. This correspondence will be made more plausible by looking at two features of this class of Hamiltonians that the MPS states also turn out to possess. A gapped Hamiltonian is one that has an energy gap, $E_1 - E_0 > 0$. A typical correlator of a gapped Hamiltonian decreases exponentially with distance as

$$\langle\psi| A(x_1)B(x_2) |\psi\rangle = C(x_1, x_2) = e^{-|x_1-x_2|/\xi} \tag{4.7}$$

where ξ is the correlation length.

Another feature of one-dimensional gapped Hamiltonians is that the entropy as a function of the length gets saturated. This saturation happens from the correlation length and up.

$$S(L) \leq \text{constant}. \tag{4.8}$$

Correlations in Matrix Product States also depend exponentially on the distance between the operators. This can be understood by the following qualitative argument that is also shown in Figure 4.4. If operator A acts on x_1 and operator B on x_2 all tensors to the left of and including x_1 can be contracted into a dual vectorlike object $\langle L|$. This can also be done for the tensor at x_2 and all tensors to its right to form $|R\rangle$. In between these vectors we have $|x_1 - x_2|$ copies of M. The correlator can thus be thought of as the expectation

value of $M^{|x_1-x_2|}$ which will be proportional to $\lambda^{|x_1-x_2|}$ where λ is some eigenvalue of M that is smaller than 1 due to normalization. The correlator thus decreases exponentially with the distance between the two sites.

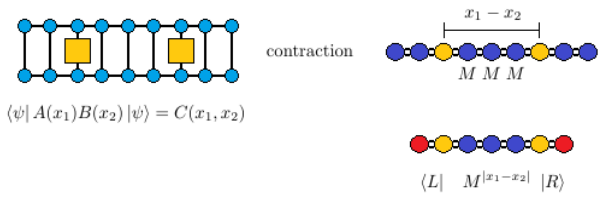


Figure 4.5: Consecutive contractions of a correlator.

If a Matrix Product State is arbitrarely divided in two parts, A and B, the state of only B is described by a density matrix where A is traced out, ρ_A. The bond index between parts A and B can take χ different values. The density matrix ρ_A can have at most rank χ since there are χ ways the bond can be contracted, all 'internal' states of A are traced out. If this density matrix is maximally mixed the entanglement entropy is $\log(\chi)$. The entanglement entropy is thus at most $\log(\chi)$. Just as a one-dimensional gapped Hamiltonians, Matrix Product States have exponentially decreasing correlators and a maximal entropy, which leads us to believe that this class of Hamiltonians can be described by Matrix Product States.

In the past two decades Matrix Product States have been extensively used as a variational ansatz in condensed matter physics.

Matrix Product States can be generalized to two dimensions, these are known as Projected Entangled Pair States.[19]

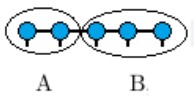


Figure 4.6: Division of a matrix product state.

4.5 MERA

The Multiscale Entanglement Renormalization Ansatz was proposed in 2006 by Guifre Vidal[20] as a class of quantum many-body states that can be efficiently simulated. The MERA is a tensor network that has a distinct causal structure and allows for efficient and exact computation of expectation values of local operators. The MERA represents a quantum many-body state on a D-dimensional lattice as a network of isometric tensors in

D+1 dimensions. This extra dimension can be thought of in two ways, as the time, where the tensor network represents a certain quantum circuit doing a computation or as a label representing scale in a coarse-graining procedure known as entanglement renormalization.

Entanglement Renormalization

Real-space renormalization methods, as originally proposed by Wilson[15], truncate Hilbert spaces of a neighbouring block of sites in a system with the aim of reducing its total Hilbert space. The total Hilbert space of a system on a lattice $\mathcal{L}$ in D spatial dimensions is given by

$$\mathcal{H}_{\mathcal{L}} = \bigotimes_{s \in \mathcal{L}} \mathcal{H}_s \tag{4.9}$$

were s is a lattice site and $\mathcal{H}_s$ the Hilbert space of a lattice site with finite dimension. A block of neighbouring sites $\mathcal{B}$ in $\mathcal{L}$ has Hilbert space

$$\mathcal{H}_{\mathcal{B}} = \bigotimes_{s \in \mathcal{B}} \mathcal{H}_s. \tag{4.10}$$

A renormalization step coarse-grains the lattice $\mathcal{L}$ to a new lattice $\mathcal{L}'$. Each lattice site s' is obtained from a block in $\mathcal{L}$. Since the aim of this procedure is to reduce the total Hilbert space, the Hilbert space of the new site s' should be smaller than the Hilbert space of the whole block. So a coarse-graining step is used where only a subspace of the block $\mathcal{S}_{\mathcal{B}}$ is mapped to the new site s'.

$$\mathcal{H}_{s'} = \mathcal{S}_{\mathcal{B}} \subset \mathcal{H}_{\mathcal{B}} \tag{4.11}$$

This coarse-graining is characterized by an isometric tensor $w^\dagger$,

$$w^\dagger : \mathcal{H}_{\mathcal{B}} \to \mathcal{H}_{s'}, \quad w^\dagger w = I. \tag{4.12}$$

Since information is lost while coarse-graining, generally $ww^\dagger \neq I$. The selection of a subspace $\mathcal{S}_{\mathcal{B}}$ is important, a large subspace will make computations hard yet for a subspace that is too small the coarse-graining will lead to a loss of important features of the original state. The optimal choice for this subspace was found by White as part of his DMRG algorithm[17]. This optimal choice of subspace is spanned by the largest m eigenvectors of the reduced density matrix on block $\mathcal{B}$ (where $\mathcal{L} - \mathcal{B}$ is traced out), where m depends on some prescribed truncation error $\epsilon > 1 - \sum_{i=1}^{m} p_i$ where p_i are the eigenvalues of the reduced density matrix. m may be thought of as a measure of entanglement between $\mathcal{B}$ and $\mathcal{L} - \mathcal{B}$ as it is approximately the rank of the Schmidt decomposition of the state

$$|\psi\rangle \approx \sum_{i=1}^{m} \sqrt{p_i} \, |\rho_i\rangle \otimes |\sigma_i\rangle \,, \quad |\sigma_i\rangle \in \mathcal{H}_{\mathcal{L}-\mathcal{B}}. \tag{4.13}$$

This means that for systems with a certain entanglement the effective Hilbert space per site grows with each coarse-graining. This makes the computation unfeasible after a

number of iterations of this coarse-graining procedure. It is also unsatisfactory from a conceptual point of view: a real-space renormalization procedure can be thought of as succesive rescalings and this procedure thus should have scale invariant systems as fixed points.

The MERA attacks this accumulation of entanglement by introducing 'disentanglers'. The aim of these disentanglers is to deform the boundaries of the block $\mathcal{B}$ using a unitary transformation so that the short range entanglement, the entanglement between the boundary sites of a block and their neighbours outside the block, is decreased. These disentanglers, u_i, are mappings of the following form:

$$u : \mathcal{H}_{s_1} \otimes \mathcal{H}_{s_2} \to \mathcal{H}_{s_1} \otimes \mathcal{H}_{s_2} \quad u^\dagger u = uu^\dagger = I. \tag{4.14}$$

Here the u's need to be chosen in such a way that they decrease the entanglement of the neighbouring pairs over the boundary of the block $\mathcal{B}$, where $\mathcal{B}$ is the block to be coarse-grained. In this way the effective rank of the density matrix of $\mathcal{B}$ will be lowered. For the one-dimensional case u needs to be chosen such that:

$$\text{rank}(\text{tr}_{\mathcal{L}-\mathcal{B}}\Big[(u_1 \otimes u_2)\rho^{\mathcal{B}}(u_1 \otimes u_2)^\dagger\Big]) < \text{rank}(\text{tr}_{\mathcal{L}-\mathcal{B}}\Big[\rho^{\mathcal{B}}\Big]). \tag{4.15}$$

Succesive applications of disentanglers u's and coarse-graining isometries w's is known as Entanglement Renormalization. The tensor network obtained by performing this procedure on a lattice is known as the Multiscale Entanglement Renormalization Ansatz. Part of a MERA is shown in Figure (4.5). A distinction is made between MERA's that have a different number of incoming legs in the isometry tensors, giving rise to a binary MERA, a ternary MERA and so on.

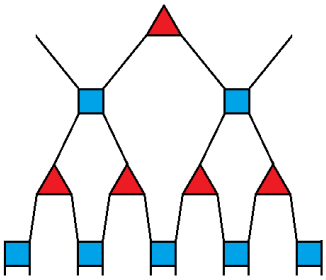


Figure 4.7: Graphical representation of a binary MERA. The triangles are isometries and the squares are disentanglers.

For a finite lattice after a finite amount of renormalization steps there are no tensors left to coarse-grain, the last tensors are then contracted with one last tensor. This tensor is called the top-tensor and can be thought of as a quantum state in the basis of the effective block sites, and thus is normalized: $T^{i_1,i_2}(T_{i_1,i_2})^* = 1$.

In [20] Vidal does computations involving the reduced density matrix for a lattice of 16,384 spins with a truncation error ϵ below 5×10^{-7} with an effective Hilbert space of dimension

8, thus keeping 8 singular values. Vidal then estimates that an equivalent computation without disentanglers would easily need a 500-1000 dimensional effective Hilbert space.

There is another point of view on MERA's. Here the tensor network is seen as a quantum circuit that transforms a product state on N sites into some entangled state on N sites. The MERA should then be considered in the opposite direction of the renormalization and the isometric tensors should be changed so that

$$w^\dagger : \mathcal{H}_\mathcal{B} \to \mathcal{H}_{s'} \otimes \mathcal{H}_{|0\rangle}, \quad w^\dagger w = I. \tag{4.16}$$

where $|0\rangle$ is one of the factorizable initial states going into the quantum circuit. Note that the only difference with the isometry of equation (4.12) is the $|0\rangle$. In this point of view the MERA takes some unentangled initial state on N sites and each level of tensors adds entanglement (on a certain scale) producing an entangled N site state.

Features of the MERA

One nice feature of the MERA is that it has some implicit notion of causality and that the causal cone of a site always has some bounded width. This causal cone of lattice site x_i is formed by all sites that can be influenced by site x_i through the engtanglement renormalization procedure as shown in Figure 4.8. The width of the causal cone is always

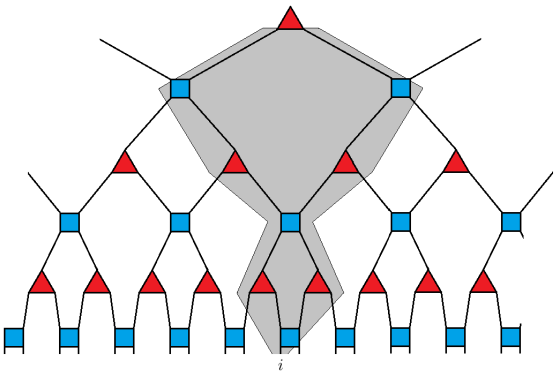


Figure 4.8: Causal cone of lattice site x_i.

smaller than some number independent of the number of lattice sites. For a lattice in D spatial dimensions the number of outgoing legs of a causal cone is at most $3^{D-1} \times 4$. In the point of view of the quantum circuit that entangles an unentangled state the causal cone consists of all legs and tensors that can influence the outcoming state on site x_i. Because of this the reduced density matrix on site x_i only depends on the legs and tensors in the causal cone, making the computation of a few reduced density matrices computationally affordable.

In section 4.3 two criteria were given for a tensor network state to be useful. It must be able to efficiently represent the wavefunction and it must be able to efficiently compute expectation values. The MERA is not able to efficiently represent wavefunctions. Lots

of contractions need to be done in order to evaluate the wavefunction and there is no algorithm to do this in a efficient way.

When computing expectation values with the MERA something nice happens. The tensors in the MERA satisfy

$$u^\dagger u = uu^\dagger = I, \quad w^\dagger w = I. \tag{4.17}$$

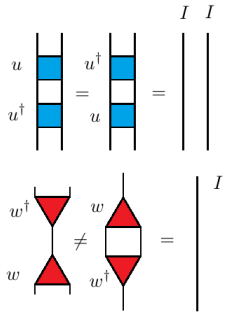


Figure 4.9: Properties of u and w in graphical notation.

These properties make sure that when we evaluate an operator acting on some sites, only the causal cone of these sites need to be calculated. The parts outside of the causal cone cancel each other out, as is illustrated in Figure 4.10.

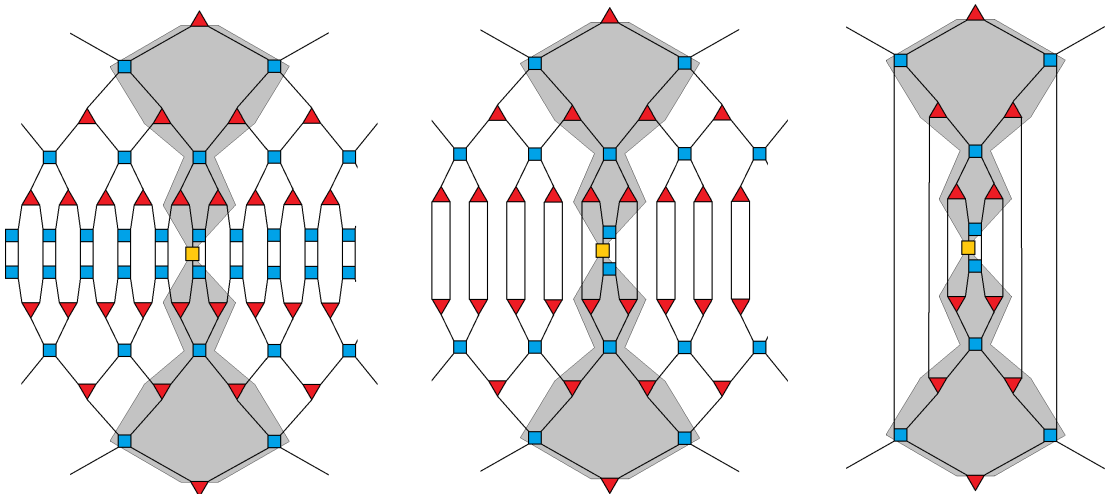


Figure 4.10: Illustration of cancellations happening when calculating an expectation value with a MERA.

Because of this property expectation values of MERA states can be calculated efficiently. This is normally done by introducing an ascending superoperator that raises the operator

to a new level of renormalization as illustrated in Figure 4.11. Then the operator is raised until the top tensor is reached and an expectation value is obtained.

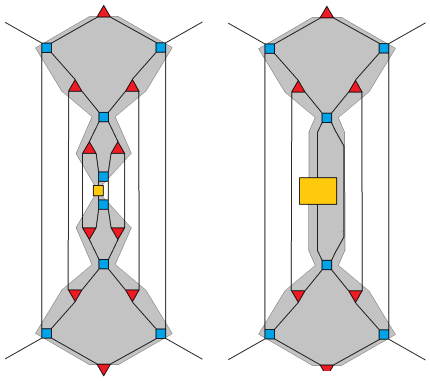


Figure 4.11: The ascending superoperator raises the local operator one level.

On top of the regular gapped Hamiltonians the MERA can also accurately describe gapless Hamiltonians. These systems do not have a gap between the ground state and the first excited state and often describe critical systems. Just as in the MPS case the correspondence between the MERA and the gapless Hamiltonians will be made more plausible by considering correlations and the behavior of entanglement entropy. Typically in systems with a gapless Hamiltonian, correlators depend on the distance between two sites as

$$\langle\psi| A(x_1)B(x_2) |\psi\rangle = C(x_1, x_2) = |x_1 - x_2|^{-c} \tag{4.18}$$

for some constant c. In one dimension the entanglement entropy of a block of L sites scales as[21]

$$S(L) \propto log(L). \tag{4.19}$$

The 'height' of a MERA with N sites is of the order of log(N). For a finite binary MERA it takes ${}_2\log N$ steps to be left with one tensor. A correlator on two sites x_1 and x_2 will depend on the distance $|x_1 - x_2|$ as $|x_1 - x_2|^{-c}$ for some constant c. After $\mathcal{O}(log|x_1 - x_2|)$ renormalization steps the sites x_1 and x_2 are mapped into one. After this, calculating the correlator is just a calculation of an expectation value, and thus not dependent on the distance $|x_1 - x_2|$. Just as in the MPS case we can argue that the correlator depends on the distance as some eigenvalue exponential in the distance. Because $x^{log(y)} = y^{log(x)}$ the correlator can be written as

$$C(x_1, x_2) = \lambda^{log|x_1 - x_2|} = |x_1 - x_2|^{log(\lambda)} = |x_1 - x_2|^{-c} \tag{4.20}$$

which is the same form as in equation 4.18 for the gapless Hamiltonian.

Gapped Hamiltonians typically have a finite correlation length. This can be represented by a MERA of a finite 'height'.

The entanglement entropy of a block of L sites scales with L as $log(L)$. It takes $\mathcal{O}(log(L))$ renormalization steps to map all L sites into one. After this the reduced density matrix

does not depend on L. The number of states that need to be traced out depends on the number of legs that need to be (minimally) cut in the graphical notation as seen in Figure 4.12. This scales with the number of sites as $log(L)$ and thus agrees with the result from the gapless Hamiltonian in equation 4.19.

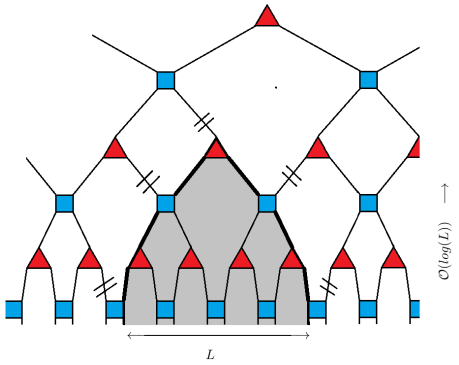


Figure 4.12: The number of bonds that need to be cut to get the reduced density matrix on L depend logarithmically on L.

When interpreting the MERA as a spatial AdS slice the behaviour of the entanglement entropy means that the MERA construction satisfies the Ryu-Takayanagi proposal.

4.6 Continuous MERA

As the MERA can be thought of as a real-space renormalization procedure it should be possible to define a similar procedure for continuous quantum states. Currently efforts are being made to make a continuous MERA or cMERA. For more information about this see [22],[23].

Chapter 5

Tensor Network Methods

5.1 MERA optimization

Earlier it was argued that a MERA can represent the lowest lying energy state of a given Hamiltonian with nearest neighbour interactions. In this section it will be explained how to obtain a MERA state representation, following [24].

5.1.1 Algorithm

The algorithm starts with a MERA state that consists of random unitary and isometric tensors. The lowest lying energy state is the state that minimizes the expectation value of the Hamiltonian, so the lowest lying energy state is the state that minimizes the tensor network representation of the expectation value of the Hamiltonian.

The algorithm optimizes tensors one by one. First the two-site Hamiltonians, $h_{r,r+1}$ are shifted by a constant value so all eigenvalues are negative. This can be done by redefining $h^*_{r,r+1} = h_{r,r+1} - \lambda_{max} I$. Suppose tensor w is optimized first. The expectation value of the Hamiltonian depends on w and $w^\dagger$,

$$E(w) = tTr[AwBw^\dagger] + c_w.$$

Here c_w is the contribution coming from the part of the MERA that is outside of the causal cone of w and tTr stands for tensortrace which means we have to contract the 'external' legs in the appropriate way. At this moment there is no known way of solving this equation, with the extra isometric constraint on w[24]. This is overcome by considering w and $w^\dagger$ to be independent. The tensor network representing the expectation value is divided in w and Γ_w, the environment of w.

$$E(w) = tTr[w\Gamma_w] + c_w,$$

where c_w is irrelevant in optimizing w. The environment Γ_w is decomposed in singular value decomposition $\Gamma_w = USV^\dagger$. The w that minimizes this is given by $-VU^\dagger$,

$$\min E(w) = \min tTr[w\Gamma_w] = tTr[-VU^\dagger USV^\dagger] = tTr[-S] = -\Sigma_i \lambda_i.$$

That this gives the minimal value might be intuitively clear and is proved in the appendix. After calculating the new w, it is updated giving rise to a new $w^\dagger$, a new environment is calculated $\Gamma_w(w^\dagger)$ and this is repeated a number of times. Because an isometry tensor lies in the causal cone of six two-site Hamiltonians the environment is made up of six contributions. A unitary tensor lies in the causal cone of three two-site Hamiltonians and therefore the environment is made up of three contributions.

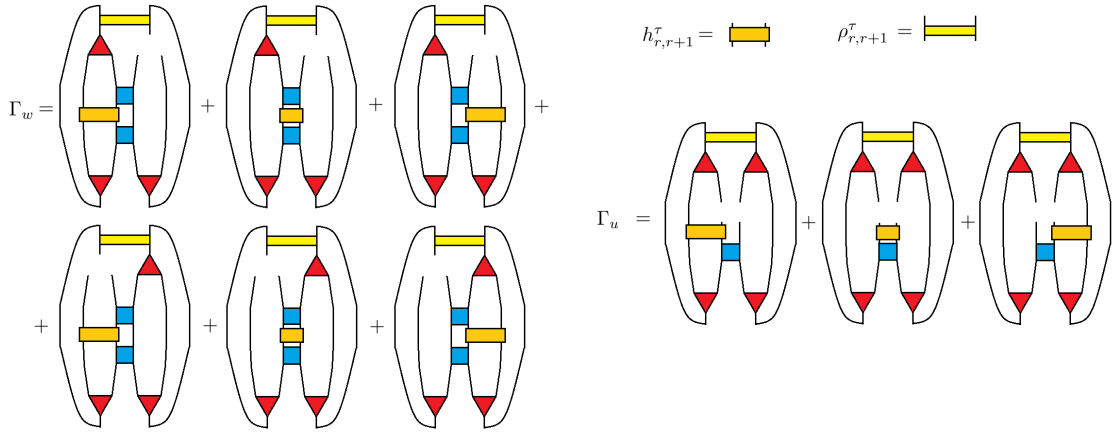


Figure 5.1: *left* Contributions to the environment of isometric tensor w. *top right* Definition of two-site Hamiltonian and density tensor for site $\{r, r+1\}$ and layer τ. *middle right* Contributions to the environment of unitary tensor u.

The environment of tensors at a certain layer can always be composed into unitaries and isometries from that layer, a specific density matrix for that site and layer $\rho_{r,r+1}$ and a two-site Hamiltonian, $h_{r,r+1}$. $\rho_{r,r+1}$ consists of all tensors in the causal cone in the layers above and $h_{r,r+1}$ consists of all tensors in the layers below the one being currently optimized.

The algorithm sequentially optimizes all tensors in the MERA. The whole MERA is swept over a couple of times until the algorithm converges. This can be done from the top down or from the bottom up. An example of an algorithm to optimize a complete MERA state is

Step 1: Calculate all density matrices for all sites and all levels.

Step 2: Starting from the lowest layer, optimize all isometric tensors w and all unitary tensors u of a certain layer τ.

Step 3: Compute the two-site Hamiltonians for layer τ.

Step 4: Repeat steps 2-3 for all layers.

Step 5: Update the top tensor.

Extra constraints can be added when dealing with a translationally invariant system or a scale invariant system. For a translationally invariant system all unitary tensors on a layer are the same and all isometries on a layer are the same. For a scale invariant system all isometries are the same and all unitaries are the same regardless of the layer.

5.1.2 Tensor network quotient takes vacuum to thermal state

In the paper by Czech, Evenbly, Vidal et al[26], it is shown that certain properties of Euclidean path integrals also hold in the MERA picture. More specifically the Euclidean path integral of a (1+1)D CFT on the upper half plane prepares the ground state on an infinite line. The conformal mapping $z \to w = (\beta/\pi)\mathrm{Log}(z)$ maps the whole upper plane to an infinite strip of width β. This prepares the thermal state with temperature $T = 1/\beta$ on an infinite line. Identifying w with $w + 2\pi L$ gives a thermal state on a ring of length L. In the z plane this mean identifying two semi-concentric circles of which the radii relatively differ $e^{2\pi^2 L/\beta}$. The thermal state on a ring is of interest to us because it makes it possible to study the thermal behaviour of the CFT, and in the AdS/CFT picture a finite temperature CFT is dual to a black hole.

In the paper it is shown this conformal mapping and this identification can also be done in the MERA picture.

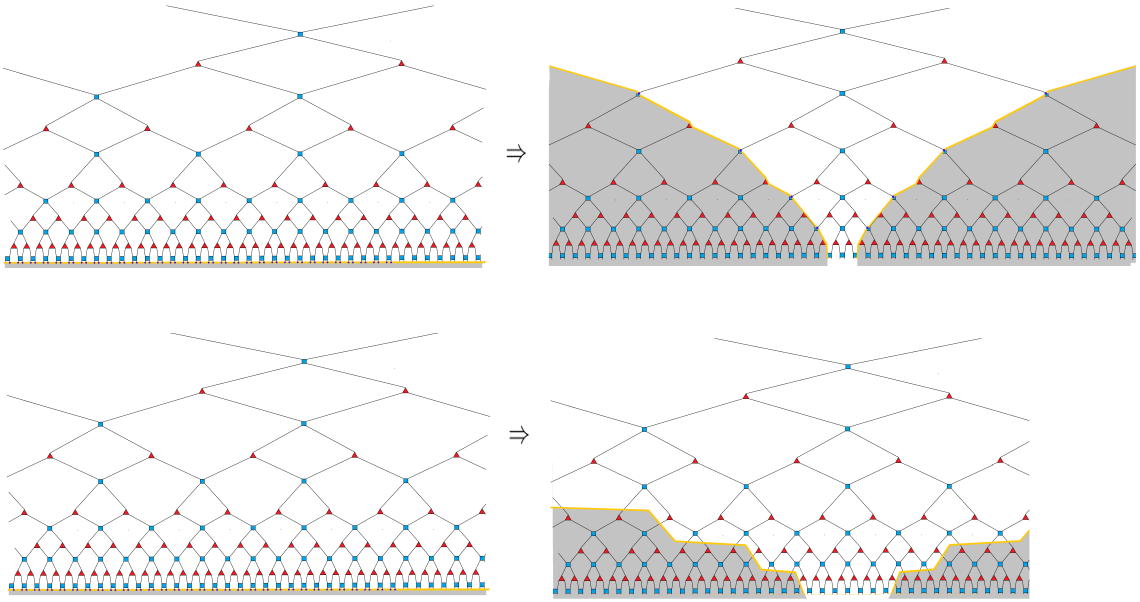


Figure 5.2: *top* Mapping $z \to w = Log_2(z) = Log(z)/Log(2)$. *bottom* Mapping $z \to \hat{w} = Log_{2^{\frac{1}{4}}} = 4Log(z)/Log(2)$

We start of with a MERA state covering the whole upper plane. This represents the ground state of the (1+1)D CFT on the infinite line just as the Euclidean path integral on the upper half plane. A discrete version of a logarithmic mapping can be made by considering

points that are exponentially discretized instead of uniformly, as seen in Figure 5.2. These new points represent the infinite line in the w-plane on which the thermal state is prepared. The form of the new line depends on the prefactor of the logarithm and in each case the best discretized mapping is chosen. Because of the discrete nature of the MERA there are not enough possible points near the origin to accurately represent the state there, effectively disconnecting the right and left part of the infinite line.

The identification of w and $w + 2\pi L$ is then done by contracting the appropriate legs as shown in Figure 5.3.

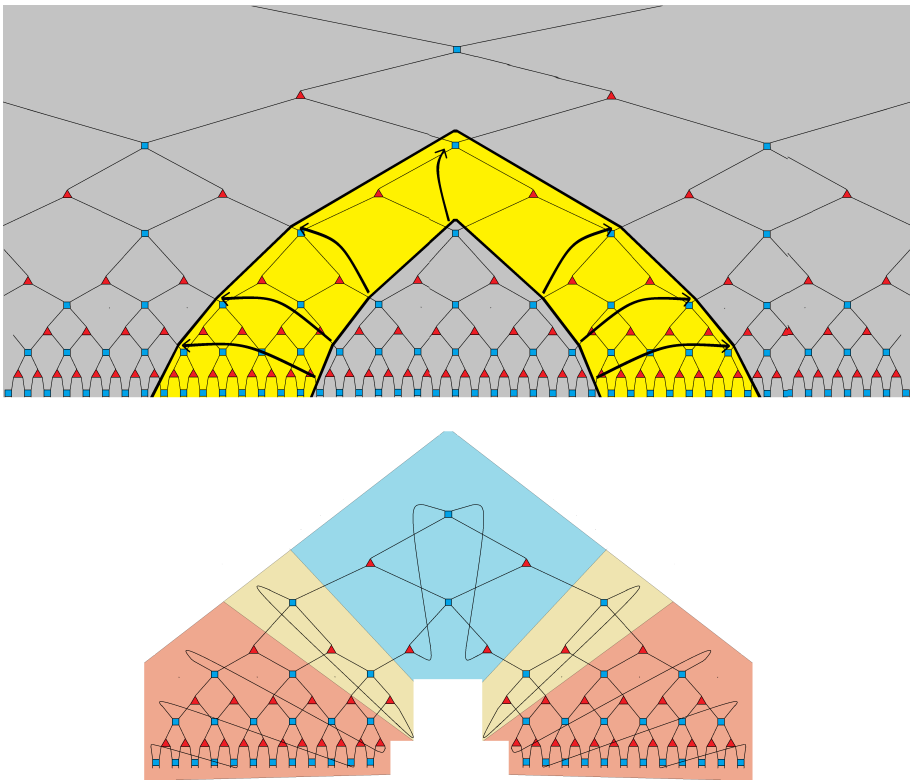


Figure 5.3: *(top)* Smallest possible identification of two constant values of w.*(bottom)* Because the MERA has an inherent causal structure it is possible to classify an interval between two tensors as spacelike, lightlike or timelike. With regard to this causal structure the blue part is connected in a timelike manner, the pink is connected in a lightlike manner and the red part is connected in a spacelike manner. The diagram is thus naturally divided in three sections.

In the article it is shown that this construction matches the theoretical predictions for the thermal spectrum of a CFT as $\beta/L \to 0$.

With some imagination this construction of the thermal state on a ring for a certain β already resembles the eternal black hole. The red parts in Figure 5.3 resemble a MERA, representing an AdS space, connected by the pink and blue part in the middle, representing the Einstein-Rosen bridge.

5.1.3 Complexity issues

The relative difference between the radii that are identified is $e^{2\pi^2 L/\beta}$. Because of the discrete nature of the MERA only certain identifications are possible. The possible identifications are an integer times the smallest possible identification. To do an identification a MERA is needed that is large enough. The depth of a MERA depends logarithmically on the number of base legs. Increasing the range of the identification then exponentially increases the number of base legs of the MERA that is needed to make the thermal state. To get to a realistic value of β/L a MERA is needed with O(1024) legs.

As was mentioned earlier evaluating expectation values with a MERA can be done efficiently because only the tensors in the causal cone come into play and the complexity scales polynomial in L. This is also important for optimizing a MERA since then we repeatedly compute the expectation value of the energy of two-site Hamiltonians. So although the complexity of optimization scales polynomial in L and not exponentially, as with random networks, it still scales. The optimization of a MERA depends on the size of the MERA so at some point the optimization will become computationally inaffordable. This bound depends somewhat on the computer used and the extent to which the code is optimized but certain MERA sizes are just out of reach of even the best supercomputers.

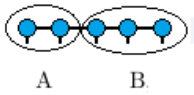


Figure 5.4: Division of tensor in two parts.

Decreasing complexity?

One might think that this complexity issue can be overcome by cutting the large tensor contractions into smaller parts. Consider the tensor given in Figure 5.4. If all legs have dimension χ, contracting the whole tensor in one time has a complexity of order $O(\chi^9)$. This because the final object has five external legs so χ^5 components where each components consists of four sums over χ values hence each component has χ^4 terms. The total complexity is then of order $O(\chi^9)$. Now suppose A and B are first contracted separately and then the results are contracted. Contracting A has complexity $O(\chi^4)$ and contracting B has complexity $O(\chi^5)$, contracting the two obtained tensors has complexity $O(\chi^7)$. So the total complexity of this procedure is $O(\chi^4 + \chi^5 + \chi^7)$ which is (for $\chi > 1$) less than $O(\chi^9)$ so in this case it is better to first do two separate contractions and then contract the results.

More generally consider a tensor network with a external legs and b internal legs. Contracting it in one time has a complexity of $O(\chi^{a+b})$. Now a cut is made through c legs dividing the network in two parts where the one system lies completely in the other. This system has d internal legs. Now evaluating the two networks separately has complexity

$O(\chi^{a+b-d}+\chi^{c+d}+\chi^{a+c})$. In the limit where there are a lot more internal legs than external and cut legs it is easily seen that the most optimal division in two parts is in parts with equal internal legs. Dividing the network in three parts leads to an optimal division in three equal parts, as long as $d > 1$. The optimal reduction of complexity that can thus be achieved is by cutting the network in equal blocks that have minimally one internal leg. This way the computational complexity can be reduced tremendously however there is one downside. Each block and each combination of blocks takes up memory, in the extreme case rendering a computer useless. Because of this the complexity problem cannot be reduced and it remains a limiting factor[25].

The sizes of the MERA states needed to build something that accurately represents the eternal black hole are simply too large to compute with the resources available. Therefore a different approach needs to be used, preferably one that does not depend on the system size but one that acts locally.

For an extensive treaty of the optimization of tensor contractions see [25].

5.2 Tensor network representations

5.2.1 Two-dimensional classical partition functions

Suppose we are interested in the thermal behaviour of a discrete classical system at some temperature $T = 1/\beta$ where the microscopic degrees of freedom are distributed according to the canonical ensemble. All information regarding such a system can be obtained from the partition function. The partition function sums over all possible configurations

$$Z(\beta) = \sum_{\{\sigma_i,\sigma_j\cdots\}} e^{-E(\{\sigma_i,\sigma_j\cdots\})\beta},$$

where E_i is the energy of a certain configuration. For concreteness we will consider the two-dimensional Ising model with periodic boundary conditions with Hamiltonian

$$H(\{s_i, s_j\cdots\}) = -\sum_{\{i,j\}} s_i s_j,$$

where $s_i = \{+1, -1\}$ and the sum only takes place over nearest neighbours(in two dimensions). The partition function can then be written as

$$Z(\beta) = \sum_{\{s_i,s_j\cdots\}} \prod_{\square ijkl} e^{-\beta(s_i s_j+s_j s_k+s_k s_l+s_l s_i)/2},$$

here the sum is again over all possible spin configurations, the product is over each spin, the factor of two is due to the double counting of the bonds. A change can be made to bond variables $\sigma_{ij} = s_i s_j$ so that

$$Z(\beta) = \sum_{\{\sigma_i,\sigma_j\cdots\}} \prod_{\square ijkl} \frac{1+\sigma_i\sigma_j\sigma_k\sigma_l}{2} e^{-\beta(\sigma_i\sigma_j+\sigma_j\sigma_k+\sigma_k\sigma_l+\sigma_l\sigma_i)/2},$$

where the factor in front of the exponential is introduced to get rid of configurations that are not physical. $\sigma_{ij}\sigma_{jk}\sigma_{kl}\sigma_{li} = s_i s_j s_j s_k s_k s_l s_l s_i = +1$ so all values that give $\sigma_{ij}\sigma_{jk}\sigma_{kl}\sigma_{li} = -1$ are dismissed. Now if we define a tensor

$$T_{ijkl} = \frac{1+\sigma_i\sigma_j\sigma_k\sigma_l}{2} e^{-\beta(\sigma_i\sigma_j+\sigma_j\sigma_k+\sigma_k\sigma_l+\sigma_l\sigma_i)/2}$$

the partition function is given by the contraction of a tensor network obtained by connecting all neighbouring sites, see Figure 5.5.

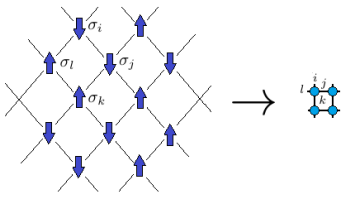


Figure 5.5: Tensor network representation of a 2d spin lattice.

$$Z(\beta) = \sum_{i,j,k,\cdots} T_{ijkl}T_{imno}T_{jpqr}\cdots = tTr(\otimes^N T),$$

5.2.2 One-dimensional quantum Euclidean time evolution operator

In quantum mechanics the partition function is given by

$$Z(\beta) = Tr[e^{-\beta H}].$$

It turns out that for a Hamiltonian with only nearest neighbour interactions the Euclidean time evolution operator $e^{-\beta H}$ can be represented in terms of tensor networks. For concreteness let us use the one-dimensional Ising model with an external magnetic field and a periodic boundary

$$H = \sum_i \sigma_i^x \sigma_{i+1}^x + \lambda \sum_i \sigma_i^z,$$

where the σ's are the Pauli matrices and λ is the coupling with the magnetic field.

For reasons that will become apparent later we introduce the integer $\frac{\beta}{\tau}$ so that the Euclidean time operator can be written as $\frac{\beta}{\tau}$ succesive operations of the operator $e^{-\tau H}$

$$e^{-\beta H} = (e^{-\tau H})^{\beta/\tau}.$$

The Hamiltonian can be split in terms that work on the even and on the odd numbered sites. The complete operator can be approximated as the succesive operation of the operator on the even sites and then of the operator on the odd sites

$$e^{-\tau H} \approx e^{-\tau H_{even}} e^{-\tau H_{odd}}$$

this is not an equality because of the Baker-Campbell-Hausdorff formula.

$$BCH : e^{A+B} = e^A e^B e^{\frac{1}{2}[A,B]} \cdots .$$

The error introduced in this approximation is of order e^{τ^2} and thus the error introduced in the full Euclidean time evolution operator is $O(e^{\beta\tau})$, the size of this error can thus be made arbitrarily small by tuning τ. This is the reason this representation of the Euclidean time evolution operator is called a semi-exact representation. Since the terms in the Hamiltonian acting on the even (and also on the odd) sites do not have any overlap they commute and

$$e^{-\tau H_{even}} = \prod_{i=even} e^{-\tau h_{i,i+1}}$$

$$e^{-\tau H_{odd}} = \prod_{i=odd} e^{-\tau h_{i,i+1}}.$$

The Euclidean time evolution operator can thus be represented as a tensor network in the following way:

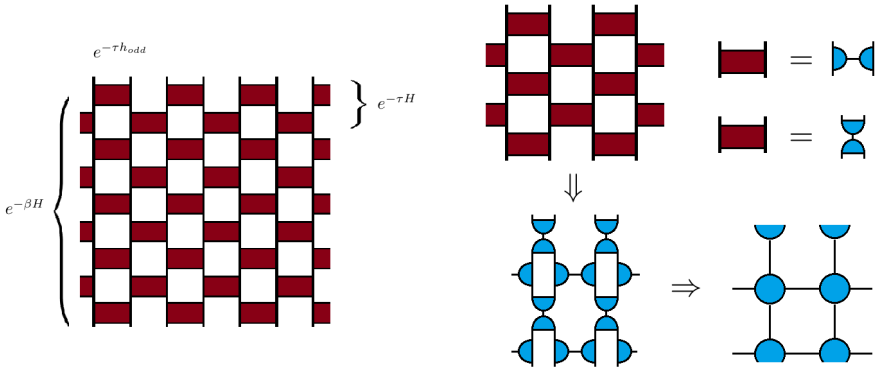


Figure 5.6: *(left)* Tensor network representation of a nearest neighbour Euclidean time evolution operator. *(right)* Reshaping the network into standard form.

For a typical β and system size both representations of the one- and two-dimensional partition functions are hard to compute. To use these representations of the partition function certain algorithms are used that will be discussed in the following sections. Both algorithms coarse grain the tensor network lattices in a certain way. In order for these algorithms to work on both representations of the one- and two-dimensional partition functions we will reshape the former so that both representations have the same form.

First the tensors are decomposed in a certain way. Suppose tensor T_{ijkl} will be decomposed. We group together indices i, j and k, l and reshape the resulting tensor in the form of a matrix. The dimension of every index was χ and so the resulting matrix indices will have dimension χ^2. This matrix is decomposed in singular value decomposition

$$M = U\Sigma V^\dagger,$$

here U, V are unitary matrices and Σ is a diagonal matrix with the singular values as it entries, selected in decreasing order. The matrix can thus be decomposed in $U\sqrt{\Sigma}$ and $\sqrt{\Sigma}V^\dagger$ both with dimensions (χ^2, χ^2). This matrices are reshaped into tensors of dimension (χ, χ, χ^2) or (χ^2, χ, χ) respectively. For a pair of indices i, j this gives

$$T_{[ij][kl]} = U^m_{[ij]} \Sigma^n_m V^{\dagger[kl]}_n.$$

In the representation of the one-dimensional partition function the two-site Hamiltonians are decomposed in this way, for each row alternating between a decomposition over a horizontal pair of indices and between a decomposition over a verical pair of indices. Then a combination of four different resulting tensors is combined into one single tensor with four indices to get a uniform tensor array. For a graphical representation see Figure 5.6.

5.3 Tensor Renormalization Group

The arrays created in the last section have a form comparable to Figure 4.2 *(middle)* and thus are computationally costly or even unaffordable. To overcome this difficulty a renormalization scheme is used based on local approximations. The first renormalization scheme we will inspect is the Tensor Renormalization Group (TRG) devised by Levin and Nahe[27].

The tensors in the array are alternatingly labelled A or B. The upper and left leg (index) of tensor A are grouped and $A_{[ul][rd]}$ is decomposed according to the singular value decomposition mentioned earlier. The coarse graining happens because of the singular values in Σ only a limited number χ_{end} is kept. This is a good approximation because typically a lot of information is carried by only a few singular values. Keeping the largest singular values then still gives you the coarse behaviour of the system. Keeping the largest singular values is realized by introducing a matrix d with dimensions (χ_{end}, χ^2) and ones on the diagonal. The decomposed tensors are then

$$S1 = U_{[ul]}\sqrt{\Sigma}d^\dagger, \qquad S3 = d\sqrt{\Sigma}V^\dagger_{[rd]}.$$

In the tensors B the indices are grouped as $[ur][ld]$ and a similar decomposition is done in S2 and S4.

The four new tensors are combined in a single tensor as shown in Figure 5.7 and the total number of tensors is halved. This proces is continued until the number of tensors in the array is of order $O(1)$. Now the partition function can be calculated by contracting the indices in the right way or a transfer matrix can be made. The largest eigenvalues of the transfer matrix are related to the energies of the system.

It is interesting to think about the physics behind the validity of this approximation and to consider when this approximation scheme is expected to fail. Suppose we start of with a large lattice of sites and after some big number of coarse graining iterations we have one tensor Ψ_{ulrd} representing a lot of other tensors. This new tensor should be thought of as a wavefunction of a one-dimensional system living on the boundary of the coarse

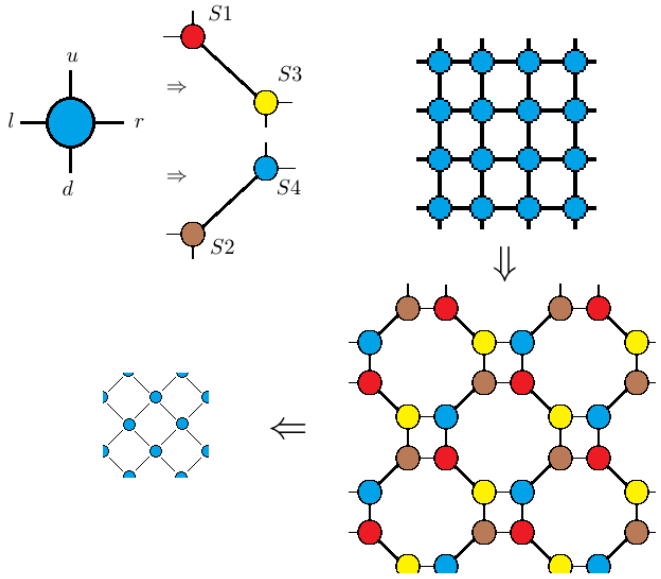


Figure 5.7: *(top left)* Tensor decomposition in S1 and S3 or S2 and S4. *(top right)* Starting array. *(bottom right)* Decomposing tensors alternating between S1,S3 and S2,S4. *(bottom left)* New tensor array with half the number of tensors.

grained square. This point of view is possible because of the deep similarity between a two-dimensional classical system and a quantum system with one spatial and one time dimension. The two-dimensional classical system is constructed by radially time-evolving the one-dimensionsal system outward with a one-dimensional quantum Hamiltonian. Because we are considering the system after a large number of coarse graining iterations this is equivalent to considering the system after a long evolution in Euclidean time hence Ψ can be thought of as the ground state. Now suppose we start of with a non-critical Ising model on the lattice. This means the ground state is gapped and typical correlators vanish exponentially fast with distance. This means that it is expected that the spins on each side of the square are minimally correlated and we can factorize the total wavefunction in wavefunctions that only depend on the particular spin states on each side

$$\Psi_{ulrd} = \sum_{ulrd=1}^{\chi} C_{ulrd}\Psi^u(\sigma_u)\Psi^l(\sigma_l)\Psi^r(\sigma_r)\Psi^d(\sigma_d).$$

This justifies the fact that we do not take into account any entanglement between spins on different sides when coarse graining.

If we start of with a critical Ising model on the lattice the ground state is gapless and spins are entangled over long distances. This makes the factorization of the wavefunction impossible, or a bad approximation. This is why Levin and Nahe[27] anticipated this coarse graining procedure to fail for critical systems.

5.4 Tensor Network Renormalization

Tensor Network Renormalization (TNR) is a coarsegraining procedure devised by Evenbly and Vidal[28] that aims to tackle the problem encountered in the last section. It aims to take into account entanglement properly and thus be able to coarsegrain critical systems. The procedure is based on removing entanglement as was done in the MERA approach. TNR again uses local replacements to replace blocks of tensors with a single tensor.

When naively repacing a block of tensors, the block of tensors should be replaced in such a way that it has a minimal effect on the values of quantities of our interest. This can be quantified by saying that the tensor norm should change as little as possible. The tensor norm is defined as

$$||A||^2 = tTr(A \otimes A^\dagger)$$

or in graphical notation as shown in Figure 5.8 (*top left*). If a block of tensors A will be replaced with a single tensor $\hat{A}$ the replacement error should be minimized

$$\epsilon^2 = ||A - \hat{A}||^2 = ||A||^2 - tTr(A \otimes \hat{A}^\dagger) - tTr(\hat{A} \otimes A^\dagger) + ||\hat{A}||^2.$$

or in graphical notation as shown in Figure 5.8 (*middle row*).

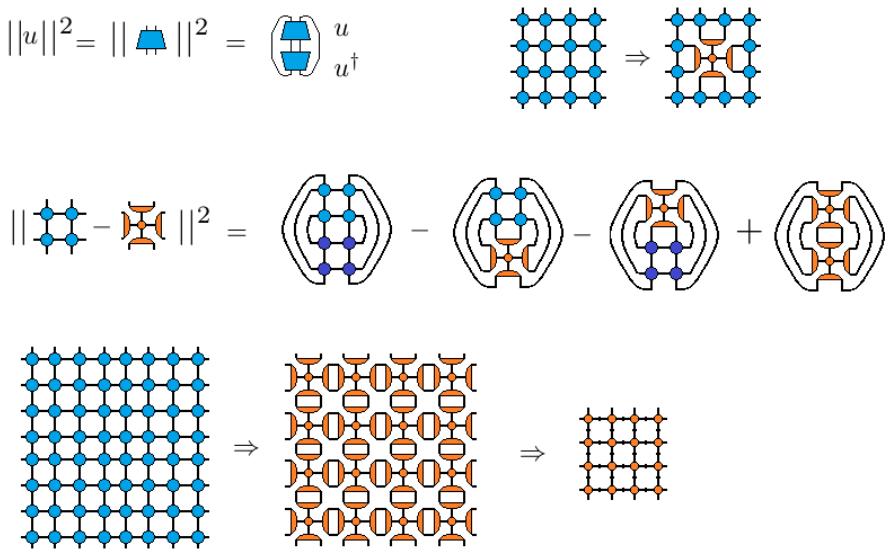


Figure 5.8: *(top left)* Definition of tensor norm. *(top right)* Local replacement of block of four tensors. *(middle row)* Error in replacing block of four tensors with new tensor. *(bottom row)* One coarse graining iteration.

Optimizing this highly nonlinear equation is hard and therefore the TNR procedure uses a slightly different replacement technique called a projective truncation.

5.4.1 Projective truncations

In a projective truncation the tensor $\hat{A}$ that will replace the block of tensors A has the specific form

$$\hat{A} = AP = Aww^{\dagger},$$

where P is a projector. Because of this restricion on the form of $\hat{A}$ the error can be simplified. P is a projector and thus is idempotent, $PP^{\dagger} = P$. This means

$$||\hat{A}||^2 = tTr(A \otimes \hat{A}^{\dagger}) = tTr(\hat{A} \otimes A^{\dagger})$$

the expression for the error can therefore be simplified to

$$\epsilon^2 = ||A||^2 - ||\hat{A}||^2 = ||A||^2 - ||AP||^2 = ||A||^2 - ||Aw||^2,$$

see Figure 5.9 for a graphical representation.

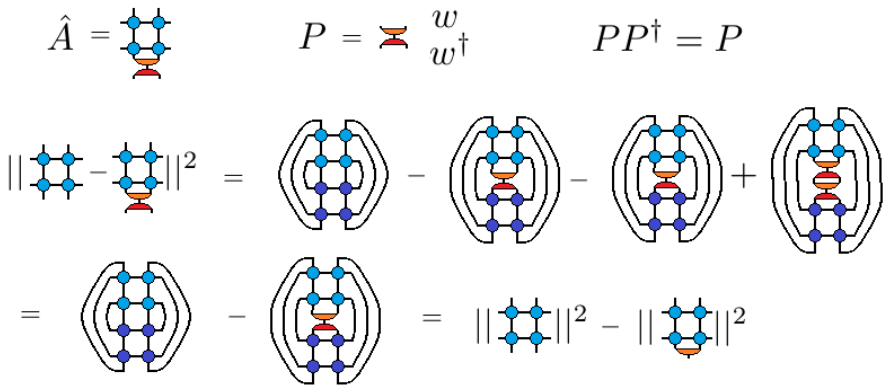


Figure 5.9: A projective truncation leads to a simpler expression for the replacement error.

Since $||Aw||^2 \leq ||A||^2$, this last expression is minimized for the maximal value of $||Aw||^2$. The problem is then to find a w that maximizes this expression. This again gives a nonlinear equation which is not easily solvable. The strategy to find this w is similar to the method used in the MERA optimization. First the expression is linearized, i.e. consider w and $w^{\dagger}$ as two separate objects. The complete expression $||Aw||^2$ is then separated in w and its environment Γ_w so that $||Aw||^2 = tTr(w \otimes \Gamma_w)$. The environment is decomposed in singular value decomposition and the w that maximizes the expression is $w = vu^{\dagger}$, which is proven in the appendix. Then $w^{\dagger}$ is updated and the whole proces is repeated until w converges.

5.4.2 Replacement steps

The TNR algorithm uses a number of projective truncations that combined lead to the coarsegrained tensor.

The TNR algorithm starts with a lattice of tensors with four indices A_{ijkl} which can represent a partition function of a classical two-dimensional system or it can represent the partition function of a one-dimensional quantum system. TNR starts with a large array of tensors and coarsegrains blocks of four tensors into one. In this way the one TNR step halves the linear size in the horizontal direction and $\frac{\beta}{\tau}$ in the vertical direction. After the array is reduced to $O(1)$ tensors the partition function is evaluated or the eigenvalues of the transfer matrix are extracted.

A block of four tensors has eight legs, two pointing towards each side. First a 'gauge' change is applied to every second row. That this can be done is shown in the appendix. This change effectively flips every row and conjugates the values in these tensors. This way a projective truncation can be done on the top two tensors of the block of four and the resulting projector will also be a projector for the bottom two conjugated tensors.

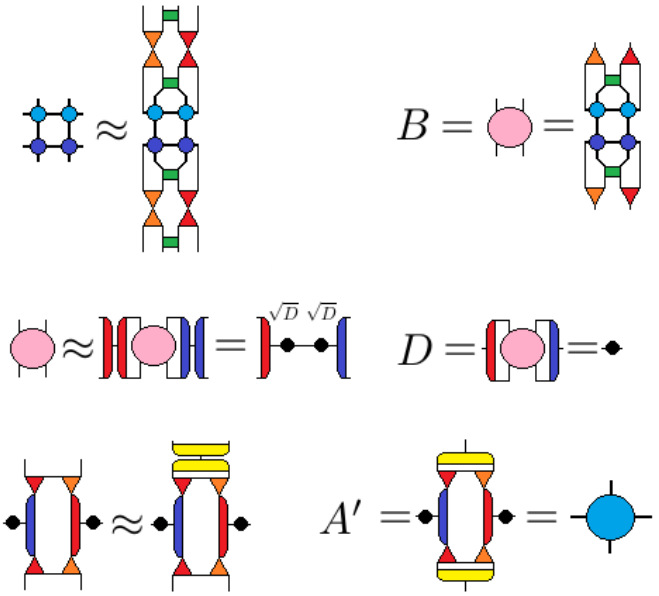


Figure 5.10: *(top left)* First projective truncation, defining w_l, u and w_r. *(top right)* Definition of tensor B. *(middle left)* Second projective truncation, defining y_l and y_r. *(middle right)* Definition of tensor D. *(bottom left)* Third projective truncation. *(bottom right)* Definition of new tensor A'.

The first step is to find a projector P for the top, right and left legs of a combination of two tensors. Because P is a projector the expression that needs to be maximized is relatively easy, as shown in the projective truncations section. The w is demanded to consist of two different isometries and one unitary. To optimize this combination of tensors first the environment of the left isometry w_l is computed, its singular value decomposition leads to a new w_l and $w_l^\dagger$, which are updated. Then the environment for the unitary u is computed, its singular value decomposition lead to a new u and $u^\dagger$, and then a new w_r is computed. These steps are repeated until the complete projector converges. Because of the particular shape of this projector the optimization can take some time.

The dimension onto which the isometries w_l, w_r project can be chosen. It should be large enough to keep the interesting physical features of the system but for too large values the optimization becomes computationally inaffordable. The key difference between TRG and similar methods compared to TNR is this unitary tensor u. Just as in the MERA its goal is to act as a basis transformation that minimizes the entanglement between the two sides of the system. After this transformation it is justified to coarse grain the two sides without regarding the entanglement with the other side, since this entanglement is minimized by the unitary.

The second step starts of with a block of four tensors, two normal tensors at the top that are acted upon by the half projector w made in step one, two flipped and conjugated tensors on the bottom acted upon by a flipped and conjugated half-projector $w^\dagger$. In this step two new projectors are made and optimized, P_l consisting of y_l and $y_l^\dagger$ and P_r consisting of y_r and y_l. For a graphical depiction see Figure 5.10.

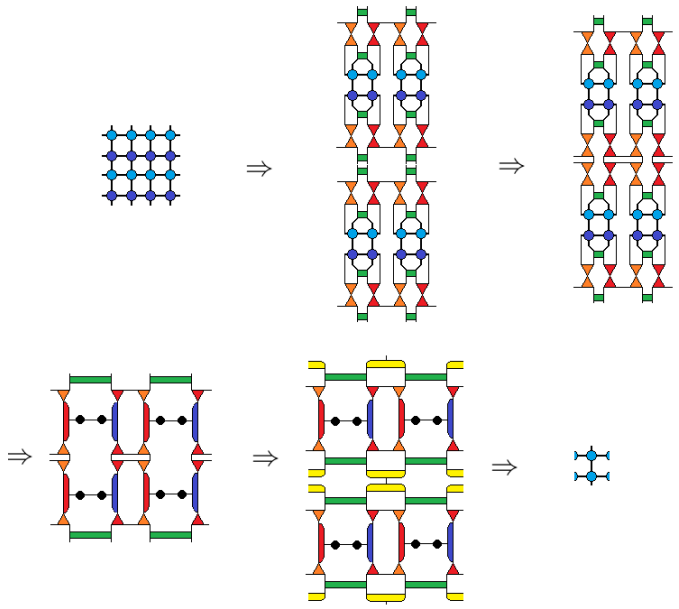


Figure 5.11: One TNR coarsegraining step.

In step three the lattice with the new tensors made in step one and two is reorganized into a lattice of tensors with two legs going up, one leg to each side and two legs going down. One last projective truncation is done on the top legs, the resulting tensor is used as a projector for the two top and two bottom legs resulting in a new tensor A' with four legs, as shown in Figure 5.11.

One iteration of TNR thus acts locally on four tensors and replaces them through a series of intricate replacements by one new tensor. Multiple coarse grainings define a renormalization group flow in the space of tensors

$$A^0 \to A^1 \to A^2 \cdots \to A^n.$$

5.4.3 TNR yields MERA etc

In the previous section tensor network methods were used to calculate partition functions. They can however also be used to calculate some other objects of our interest as was shown by Evenbly and Vidal[30].

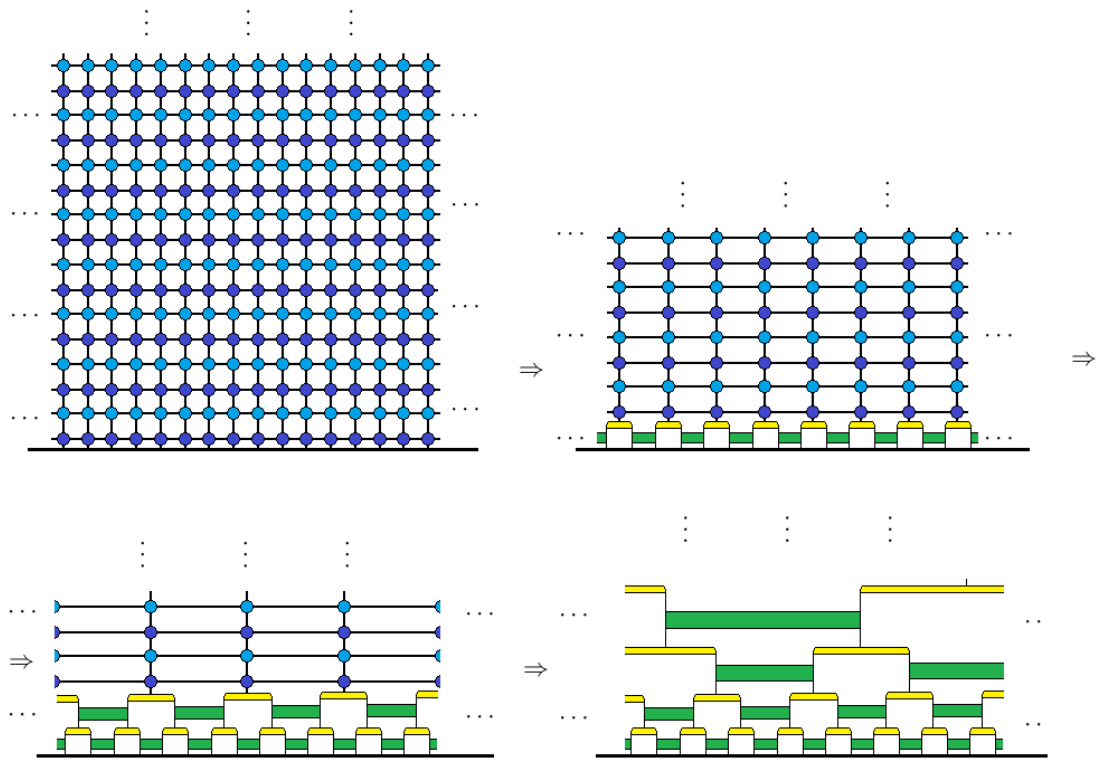


Figure 5.12: Applying TNR on an infinite lattice restricted to the upper half plane with a row of open indices at the bottom. The result is a MERA state.

If we consider the one-dimensional quantum partition function but do not perform the trace we are left with the Euclidean time evolution operator $e^{-\beta H}$. This operator takes a state and evolves it in Euclidean time. In the limit where β becomes large this operator suppresses all energy eigenstates except for the ground state. This means that acting with this operator on a state and evolving it for a large β gives you the ground state. Suppose we do this for the tensor network representing the Euclidean time evolution operator. We would get a network that has open legs at the bottom for some linear size and that is large vertically since β and thus $\frac{\beta}{\tau}$ is large. This network represents the ground state of the Hamiltonian H. Normally when applying the TNR coarsegraining procedure certain tensors are formed and are immediately cancelled against their conjugated versions acting on a different block. In this particular case this will not happen near the open legs, since there nothing is coarse grained. This means that at the open legs a certain structure will start to form composed of tensors that would have normally cancelled. It turns out

that this structure is exactly the MERA state as can be seen from Figure 5.12. This is a nice consistency check because the MERA was build by trying to represent the ground state and taking care of the entanglement in a certain way. Now that we take care of entanglement in the same way in a different setting the ground state is still represented by a MERA.

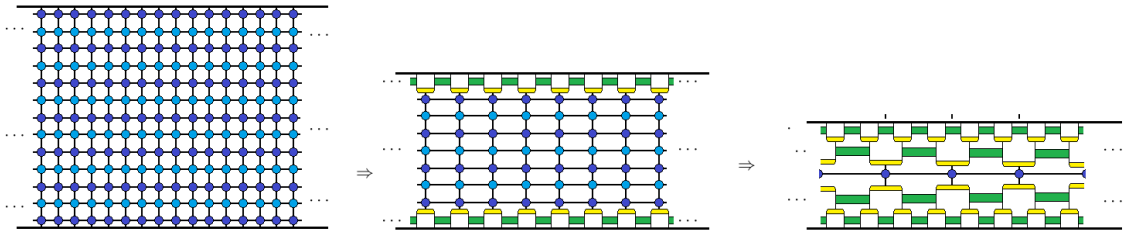


Figure 5.13: Applying TNR on a system on infinite sites for a finite β. The result resembles a thermofield double state.

When we consider a tensor network representation for a strip of finite β we get a state that is proportional to the thermal state $e^{-\beta H}/Z$. When coarsegraining this network the result is a network that has two MERA's connected by an intermediate row of tensors. This is exactly the state that B. Swingle expected to represent the thermofield double state. The two MERA's represent two spaces that are connected via a Einstein-Rosen bridge or eternal black hole represented by the intermediate row of tensors.

Another construction that can be made is considering the Euclidean time evolution operator for large β on a periodic chain consisting of L sites. After $O(Log(L))$ coarsegraining steps the result is a small MERA with a semi-infinite tower of tensors on top of it. This semi-infinite tower can be thought of as the infinite product of a transfer matrix. The largest eigenvalue of such a matrix leads to the ground state, offering a possibility to study finite size systems.

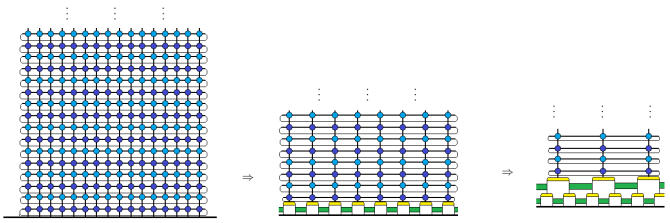


Figure 5.14: Applying TNR on a system on a finite system gives rise to a tower of transfer matrices.

Chapter 6

Results

6.1 Dimensions of two-dimensional Ising CFT

A tensor network representation of a classical two-dimensional partition function of the Ising model was made and coarse-grained using the Tensor Renormalization Group method. Coarsegraining iterations are performed until only one tensor is left. The horizontal legs of the tensor are connected to give a transfer matrix. The eigenvalues of the transfer matrix are exponentially proportional to their conformal dimensions. The ratios between the conformal dimension of our interest and the first conformal dimension are then extracted.

Figure 6.1 is a plot of the first conformal dimension as a function of the linear size. At criticality and in the continuum limit the two-dimensional Ising model is expected to behave like the two-dimensional Ising CFT. The value of the first conformal dimensions is 1/8 and Figure 6.1 clearly shows that for larger sizes the value of the first conformal dimension approaches the theoretical value.

The different lines are different values of χ which is the amount of singular values that are kept in a coarse graining iteration. Larger values of χ agree better with the theoretical result. From the considerations of the the original authors of the TRG method it became clear that this method is expected to fail after a certain size for critical systems. For the $\chi = 8$ case it is clear that the agreement with the theoretical result quickly drops as larger systems are examined. A similar effect is seen in the $\chi = 16$ line. Unfortunately it becomes hard to examine larger system sizes. This is because the largest component in one tensor is e^4. As larger systems are examined the largest components grow fast and for a system of $2^{10} = 1024$ spins the values are larger than Mathematica can accurately handle. This means that the largest examined system is of linear size $2^9 = 512$ spins. When we realize that the Hilbert space dimensions of this system is of order $O(e^{512})$ it is already impressive to be able to describe this system to some extent.

As the complexity of the algorithm is of order $O(\chi^4)$ larger values for χ are also hard to compute.

In Figure 6.2 the first 19 conformal dimensions of the $\chi = 48$ case are given. The y-axis

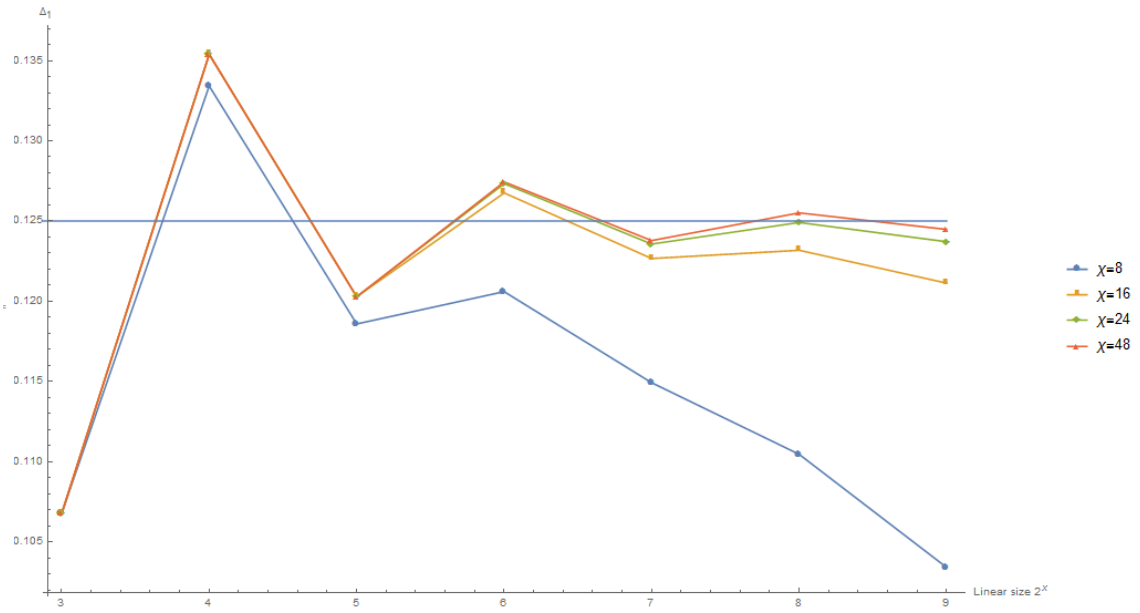


Figure 6.1: Second conformal dimension.

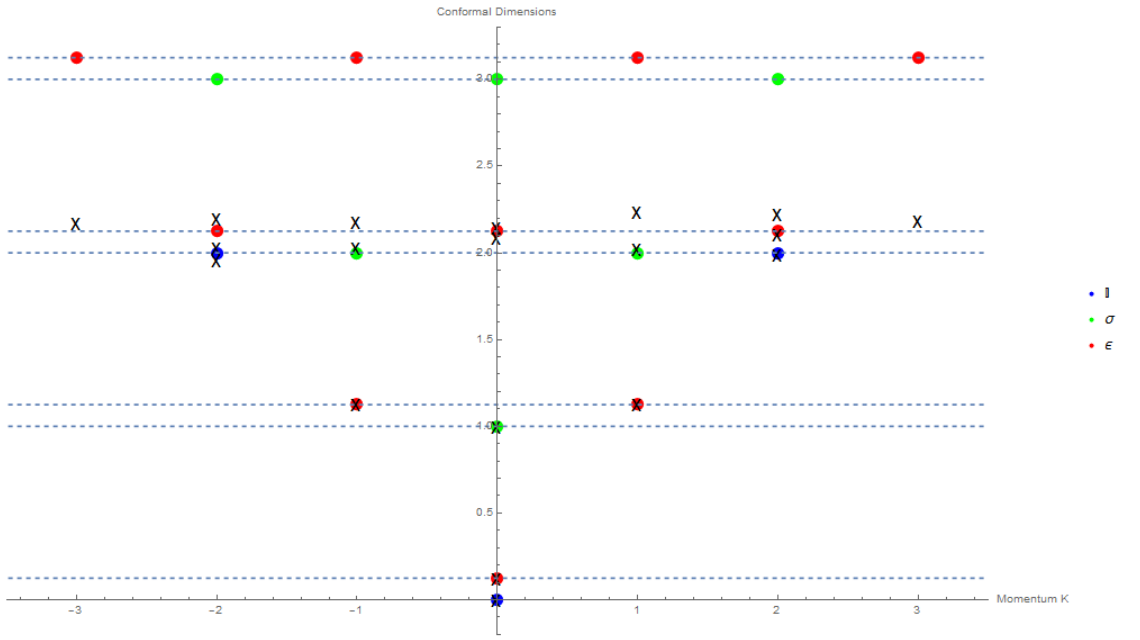


Figure 6.2: First 19 conformal dimensions for $\chi = 48$. Colors represent theoretical values. Crosses are values obtained by the TRG method.

represents the value of the conformal dimension and the x-axis represents the momentum of the conformal dimension. The three primary fields in the two-dimensional Ising CFT

I, σ, ϵ have conformal dimensions $0, 1/8, 1$ respectively. The structure of all conformal dimensions is then given by the primary fields and then the infinite towers corresponding to their descendant fields. In the plot the different colors represent the different towers. There are dashed lines at all theoretical values. Black x's are the values obtained from the TRG method. The first nine conformal dimensions closely agree with the theoretical value. The consecutive three dimensions deviate somewhat but after these the values are completely off.

This is not strange since it was expected that this method is not able to describe critical systems correctly. When keeping a large number of singular values (large χ) this method is able to describe the most important features (lowest lying conformal dimensions) of the system yet it quickly breaks down.

6.2 RG flow structure

TNR takes entanglement into account in the same way that the MERA takes entanglement into account. This leads to the expectation that TNR will be able to coarsegrain critical systems. Because critical systems are desribed by conformal field theories in the continuum limit, and these are scale invariant, the tensor should flow to a fixed point tensor after enough coarse grainings

In Figure 6.3 four graphical representations are shown of the transfer matrix at different points in the renormalization group flow. Also a diagram of the renormalization group flow is shown. The top matrix is a transfer matrix of a 8x8 grid of spins. Different colors represent different nonzero values and thus possible states. The top matrix thus has a lot of possible states. For different values of the coupling λ this matrix is nearly the same.

Now eight coursegrainings are performed, thus in a sense the continuum and low temperature limit are taken. For all three different ranges of the coupling λ a different fixed point tensor is obtained. In the range $\lambda < 1$ a trivial fixed point is obtained with two significant dots in the graphical representation representing two possible states, namely the state where all spins are aligned with the external magnetic field and one state where all spins are anti-aligned with the external magnetic field. In the range $\lambda > 1$ another trivial fixed point tensor is obtained with only one dot in the graphical notation representing only one state, namely the state where all spins are aligned with the external magnetic field. For the case where the coupling is exactly equal to its critical value $\lambda = 1$ a non-trivial fixed point tensor is obtained. This object has multiple possible states hence it has some structure and thus is non-trivial. The system reached a conformal point, since it does not change any more even after more coarse-grainings.

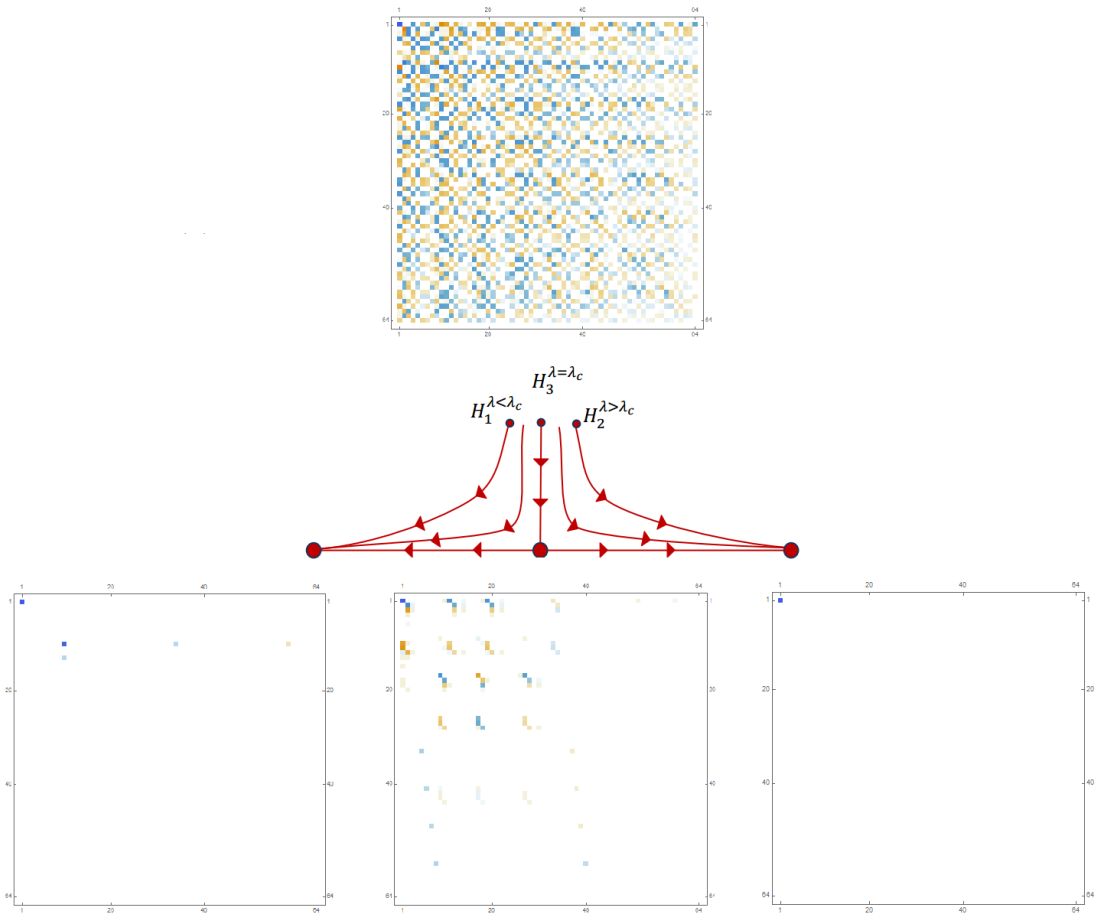


Figure 6.3: Graphical representation of transfer matrix showing structure of renormalization group flow.

Chapter 7

Conclusion

How is AdS/CFT realized in the MERA? The MERA captures some key features of the AdS/CFT correspondence. It naturally produces an extra hyperbolic spatial dimension starting from a critical state. Interpreted as a spatial AdS slice the MERA satisfies the Ryu-Takayangi proposal. Downsides of the MERA are that it is discrete, that it is not clear how excited states can be constructed. Efforts are made to make a continuous equivalent of the MERA, called CMERA[22][23]. Although there are no pure excited states it is possible to construct thermal states.

More recently it has become possible to study thermal behaviour by tensor networks. The TRG method is able to do this but fails for critical systems. The TNR is expected to be better at coarsegraining critical systems and this has been confirmed by reproducing the renormalization group flow structure of the two-dimensional Ising CFT. Moreover the TNR connects the ground state obtained from the Euclidean path integral to the original MERA ground state, giving a consistency check of sorts.

Because of these considerations tensor networks offer an interesting new way to study quantum gravity.

In the future it would be interesting to look at time dependence. The time evolution operator can be represented in a similar way as the Euclidean time evolution operator and thus it should be possible to study time dependence. This would be especially interesting for the thermofield double state since this would make it possible to study the typicallity of the eternal black hole in the tensor network picture.

It might also be interesting to put orthogonaliy constraints on the MERA so that it is able to represent the first excited state.

Bibliography

[1] I. Newton *General Schollium Book 3, Philosophiae Naturalis Principia Mathematica* (1687)

[2] J. Maldacena *The Large N Limit of Superconformal Field Theories and Supergravity*Adv.Theor.Math.Phys. Vol 2 p231-252 (1998)

[3] P. Kovtun, D.T. Son, O. Starinets *Viscosity in Strongly Interacting Quantum Field Theories from Black Hole Physics* Phys.Rev.Lett. 94 (2005) 111601

[4] S.S. Gubser, I.R. Klebanov, A.M. Polyakov *Gauge Theory Correlators from Non-Critical String Theory* Phys.Lett.B428 p105-114 (1998)

[5] E. Witten *Anti De Sitter Space And Holography*Adv.Theor.Math.Phys. Vol 2 p253-291 (1998)

[6] M. Ammon, J. Erdmenger *Gauge/Gravity Dualities* Cambridge University Press; 1 edition (2015)

[7] S. Ryu, T. Takayanagi *Holographic Derivation of Entanglement Entropy from AdS/CFT* Phys.Rev.Lett.96:181602,2006

[8] B. Swingle *Constructing holographic spacetimes using entanglement renormalization* arXiv:1209.3304

[9] T. Hartman, J. Maldacena *Time evolution of entanglement entropy from black hole interiors* J. High Energ. Phys. Vol 14.(2013)

[10] P. di Francesco, P. Mathieu, D. Sénéchal *Conformal Field Theory* Springer Science and Business Media New York (1997)

[11] J. Maldacena *Eternal Black Holes in AdS* JHEP 0304 (2003) 021

[12] D. Harlow *Jerusalem Lectures on Black Holes and Quantum Information* Rev. Mod. Phys. 88, 15002 (2016)

[13] L. Susskind, J. Lindesay *An introduction to Black Holes, Information and the String Theory Revolution* World Scientific Publising, Singapore (2005)

[14] R. Penrose *Applications of negative dimensional tensors, combinatorial mathematics and its applications* Academic Press, London (1971)

[15] K. Wilson *The renormalization group: Critical phenomena and the Kondo problem* Rev. Mod. Phys. 47, 773 (1975)

[16] A. Weichselbaum, F. Verstraete, U. Schollwck, J. I. Cirac, Jan von Delft *Variational matrix product state approach to quantum impurity models* Phys. Rev. B 80, 165117 (2009)

[17] R.G. White *Density matrix formulation for quantum renormalization groups* Vol. 69, Iss. 19 (1992)

[18] Talk in Mathematica Summer School on Theoretical physics 2015 by Guifre Vidal *Ground state entanglement and tensor networks*

[19] F. Verstraete, J. Cirac, V. Murg *Matrix Product States, Projected Entangled Pair States, and variational renormalization group methods for quantum spin systems* Adv. Phys. 57,143 (2008)

[20] G. Vidal *A class of quantum many-body states that can be efficiently simulated* Phys. Rev. Lett. 101, 110501 (2008)

[21] P. Calabrese, J. Cardy *Evolution of Entanglement Entropy in One-Dimensional Systems* J.Stat.Mech.0504:P04010,2005

[22] J. Haegeman, T. J. Osborne, H. Verschelde and F. Verstraete *Entanglement renormalization for quantum fields* Phys. Rev. Lett. 110, 100402 (2013)

[23] M. Nozaki, S. Ryu, T. Takayanagi *Holographic Geometry of Entanglement Renormalization in Quantum Field Theories* JHEP10(2012)193

[24] G. Evenbly, G. Vidal *Algorithms for entanglement renormalization* Phys. Rev. B 79, 144108 (2009)

[25] R.N.C. Pfeifer, J. Haegeman, F. Verstraete *Faster identification of optimal contraction sequences for tensor networks* Phys. Rev. E 90, 033315 (2014)

[26] B. Czech, G. Evenbly, L. Lamprou, S. McCandlish, X. Qi, J. Sully, G. Vidal *A tensor network quotient takes the vacuum to the thermal state* Phys. Rev. B 94, 085101 (2016)

[27] M. Levin, C.P. Nahe *Tensor renormalization group approach to 2D classical lattice models* Phys. Rev. Lett. 99, 120601 (2007)

[28] G. Evenbly, G. Vidal *Tensor Network Renormalization* Phys. Rev. Lett. 115, 180405 (2015)

[29] G. Evenbly *Algorithms for Tensor Network Renormalization* Phys. Rev. B 95, 045117 (2017)

[30] G. Evenbly, G. Vidal *Tensor network renormalization yields the multi-scale entanglement renormalization ansatz* Phys. Rev. Lett. 115, 200401 (2015)

[31] R. Orus *A Practical Introduction to Tensor Networks: Matrix Product States and Projected Entangled Pair States* Annals of Physics 349 (2014) 117-158

[32] R. Orus *Advances on Tensor Network Theory: Symmetries, Fermions, Entanglement, and Holography* Eur. Phys. J. B (2014) 87: 280

[33] R.N.C. Pfeifer, G. Evenbly, S. Sukhwinder, G. Vidal *NCON: A tensor network contractor for MATLAB* arXiv:1402.0939

[34] S. Singh, G. Vidal *Symmetry protected entanglement renormalization* Phys. Rev. B 88, 121108(R) (2013)

[35] www.desy.de

[36] K. Papadodimas Presentation: *Introduction to AdS/CFT* Corfu 2012

Appendix A

Calculation by Maldacena and Hartman

In the article by Maldacena and Hartman [9] the authors consider a thermofield double state. This is a state where two CFT states are entangled in the following way:

$$|\psi\rangle = \sum_n |E_{n_1}\rangle\, |E_{n_2}\rangle\, e^{-\frac{\beta}{2}E_n} \tag{A.1}$$

The gravitational dual of this state is the eternal black hole. They consider the situation where the two CFT's are simultaniously split in half at some time t_b and then look at the entanglement entropy of this half with the rest of the system. This entanglement entropy can be calculated via Conformal field theory but also via the gravitational side. This gravitational calculation is repeated here.

The metric of a planar black brane is considered:

$$ds^2 = -g^2(\rho)dt^2 + h^2(\rho)dx_{d-1}^2 + d\rho^2 \tag{A.2}$$

$$\mathrm{h} = (\frac{2}{d})^d \cosh(\frac{d\rho}{d})^{2/d}, \quad \mathrm{g} = \mathrm{h}\, \tanh(\frac{d\rho}{2}). \tag{A.3}$$

The interior region corresponds to $\rho = i\kappa$ and $t = t_I - i\pi/2$ so that ρe^t(the ingoing Kruskal Szekeres coordinate) is finite as we cross the horizon and t_I is real.

According to the Ryu-Takayanagi proposal the entanglement entropy in the boundary theory can be given by a minimal surface dipping into the gravitational bulk. The gravitational calculation of the entanglement entropy thus reduces to finding the minimal codimension two surface that connects the two entanglement boundaries in the CFT's. It reduces to finding a function $\rho(t)$ or $t(\rho)$ that minimizes the surface:

$$A = V_{d-2} \int [h(\rho)]^{d-2} \sqrt{-g^2(\rho)dt^2 + d\rho^2}. \tag{A.4}$$

Because of symmetry considerations we expect $\frac{d\rho}{dt} = 0$ at $t_I = 0$. The expression for the surface can be written as

$$A = V_{d-2} \int d\rho f(t, \frac{dt}{d\rho}; \rho). \tag{A.5}$$

Since this has the form of an action and the surface is minimized the equations of motion can be used, giving:

$$\frac{g^2 h^{d-2}}{\sqrt{-g^2 + (\frac{d\rho}{dt})^2}} = \text{constant} = -ig_0 h_0^{d-2} \tag{A.6}$$

where g_0, h_0 are just g, h evaluated at $t_I = 0$, $\rho_0 = i\kappa_0$. This last equation can be recast to

$$t(\rho) = -i\pi/2 + \int_{i\kappa_0}^{\rho} \frac{d\rho'}{g\sqrt{1 - \frac{g^2 h^{2d-4}}{g_0^2 h^{2d-4}}}} \tag{A.7}$$

and the formula for the surface can be rewritten as

$$A = V_{d-2} \int h^{d-2} \sqrt{-g^2 + (\frac{d\rho}{dt})^2} = 2V_{d-2} \int_{i\kappa_0}^{\infty} \frac{h^{d-2} d\rho}{\sqrt{1 - \frac{g_0^2 h_0^{2d-4}}{g^2 h^{2d-4}}}} \tag{A.8}$$

where the 2 comes from the symmetry in the system. It is useful to think about these integrals in terms of $a \equiv -igh^{d-2}$ because then

$$t_I = \int \frac{dp'}{g\sqrt{1 - \frac{a^2}{a_0^2}}}, \quad A = 2V_{d-2} \int \frac{h^{d-2} d\rho}{\sqrt{1 - \frac{a_0^2}{a^2}}}. \tag{A.9}$$

Generally $a = 0$ at the boundary $\rho = 0$ and then grows until it attains its maximum a_m and then decreases again. The contribution to the area from the interior can be rewritten using integration by parts as

$$\begin{aligned} &\int_0^{a_0} \frac{h^{d-2} a da}{\frac{da}{dk}\sqrt{a_0^2 - a^2}} = \\ &-\sqrt{a_0^2 - a^2} \frac{h^{d-2}}{\frac{da}{dk}}|_0^{a_0} - \int_0^{k_0} (\frac{(d-2)h^{d-1}\frac{dh}{da}}{\frac{da}{dk}} - \frac{h^{d-2}}{\frac{da}{dk}})(-\sqrt{a_0^2 - a^2}) da \end{aligned} \tag{A.10}$$

here it becomes apparant that the integral is wellbehaved everywhere in the interval as long as $a_0 \neq a_m$. For large t, $a_0 \to a_m$ and the surface lies along $k \approx k_m$ for a long time (this can be seen by noting that the largest contribution to the time integral comes from the a's near a_0) and then goes to the boundary. This contribution to the area is approximately $A = a_m t_b$ ($\frac{d\rho}{dt} = 0$) indicating a linear growth of the entanglement entropy due to the interior of a black hole.

Appendix B

Proof for maximal trace

In the space of all $m \times n$ matrices with complex entries an inner product can be defined as

$$\mathrm{Tr}(AB^\dagger) = \langle A, B \rangle \tag{B.1}$$

since this satisfies the three conditions for an inner product:

1. Positive definiteness
2. Linearity
3. Conjugate symmetry

1.

$$\langle A, A \rangle = \mathrm{Tr}(AA^\dagger) = \sum_{i=1}^{n} (AA^\dagger)_{ii} = \sum_{j=1}^{m}\sum_{i=1}^{n} A_{ij}A^\dagger_{ji} = \sum_{j=1}^{m}\sum_{i=1}^{n} |A|^2 \geq 0 \tag{B.2}$$

2.

$$\langle \lambda A + B, C \rangle = \mathrm{Tr}(\lambda AC^\dagger + BC^\dagger) = \lambda \mathrm{Tr}(AC^\dagger) + \mathrm{Tr}(BC^\dagger) = \lambda \langle A, C \rangle + \langle B, C \rangle \tag{B.3}$$

3.

$$\langle A, B \rangle = \mathrm{Tr}(AB^\dagger) = \mathrm{Tr}((AB^\dagger)^T) = \mathrm{Tr}(B^* A^T) = \langle B^*, A^* \rangle \tag{B.4}$$

The Cauchy-Schwarz inequality states that an inner product satisfies

$$| \langle A, B \rangle | \leq ||A||||B||. \tag{B.5}$$

Suppose we define the matrix $d = \sqrt{\Sigma}$, where Σ is the diagonal matrix from the singular value decomposition $M = U\Sigma V^\dagger$. The w that maximizes $\mathrm{Tr}(w\Gamma)$ is then given by $w = VU^\dagger$ as shown below.

$$\mathrm{Tr}(w\Gamma) = \mathrm{Tr}(wU\Sigma V^\dagger) = \mathrm{Tr}(wUdd^\dagger V^\dagger) = \langle wUd, Vd \rangle \leq ||wUd||||Vd|| = |d|^2 = \mathrm{Tr}(\Sigma).$$

(B.6)

In the second to last step the isometric and unitary properties of w, U, V are used. The equality is satisfied for $\mathrm{Tr}(w\Gamma) = \mathrm{Tr}(\Sigma)$ hence for $w = VU^\dagger$.

Appendix C

Gauge change in tensor network

When applying the TNR procedure on a one-dimensional Euclidean time evolution operator the Hermitian property of the Hamiltonian is used to perform a certain 'gauge' change on every second row, this reduces the complexity of the algorithm.

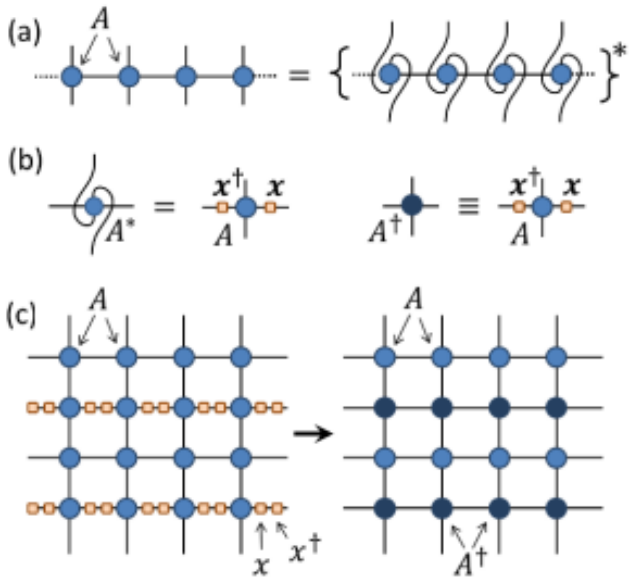


Figure C.1: *(a)* Hermitian property. *(b)* Conjugation and permutation of horizontal indices is equivalent to acting with unitary matrix x. *(c)* Gauge change as used in TNR.[29]

The unitary matrices x are defined as

$$A^\dagger = (A_{ilkj})^* = \sum_{j,l} (x^*_{j,j'} A_{ijkl} x_{l,l'}) \tag{C.1}$$

because the Euclidean path integral is taken for a system defined on a ring all $x's$ cancel and the associated change from A to $A^\dagger$ can be performed without introducing extra

tensors.

Because of this replacement the projectors created in the first TNR step for the two normal tensors, the four upper legs, can also be used on the conjugated tensors, the four lower legs. This reduces the algorithms' complexity.

In the case of a two-dimensional partition function such a symmetry can exist due to spatial symmetries present in the underlying lattice.

Appendix D

Nederlandse Samenvatting (Dutch Summary)

In 1628 publiceerde Isaac Newton de wet van de zwaartekracht. Deze wet stelt dat alle objecten met massa een kracht op elkaar uitoefenen. Hiermee kan de beweging van de planeten maar ook het vallen van een appel met dezelfde wet beschreven worden. Het probleem van deze wet is dat zij enkel beschrijft wat er gebeurd en niet zegt waarom een planeet en een appel een kracht op elkaar uitoefenen. In het begin van de twintigste eeuw kwamen er twee grote nieuwe theorien uit de natuurkunde voort. De algemene relativiteitstheorie en de kwantummechanica. De algemene relativiteitstheorie vertelt ons dat ruimte en tijd eigenlijk deel zijn van een groter geheel genaamd ruimte-tijd en dat zware objecten deze ruimte-tijd verstoren. Door deze verstoring bewegen objecten zich anders door de ruimte-tijd en lijkt het net alsof objecten met massa elkaar aantrekken. De andere grote nieuwe theorie genaamd kwantummechanica beschrijft hoe de natuur werkt op de allerkleinste schalen. Het probleem van deze twee theorien is dat ze niet consistent met elkaar zijn, specifieker: we weten niet hoe we algemene relativiteitstheorie moeten interpreteren in het raamwerk van de kwantummechanica. Dit wordt al sinds het begin van de twintigste eeuw geprobeerd en is nog steeds niet gelukt.

Sinds 1997 is er een nieuwe manier gevonden om dit probleem aan te vallen genaamd de Anti-de-Sitter/ Conformele Velden Theorie dualiteit. Dit is een dualiteit tussen twee theorien die in eerste instantie niets met elkaar te maken lijken te hebben. Aan de ene kant van de dualiteit staat een theorie die zwaartekracht beschrijft in een speciale Anti-de-Sitter ruimte op een manier die consistent is met kwantummechanica en aan de andere kant van de dualiteit staat een theorie (CFT) die deeltjes beschrijft die een sterke wisselwerking met elkaar hebben en waar geen zwaartekracht is. Door de dualiteit kan het probleem van de zwaartekracht 'vertaald' worden naar de andere kant van de dualiteit en daar bestudeerd worden. Hoewel de dualiteit in grote lijnen goed word begrepen is het nog helemaal niet duidelijk hoe de zwaartekrachtstheorie in de AdS ruimte uit de CFT voortkomt. Recentelijk is er een voorstel gedaan om dit te bestuderen met tensor netwerken. Dit zijn wiskundige objecten die gebruikt worden om grote hoeveelheden deeltjes te beschrijven. In deze scriptie heb ik onderzocht in hoeverre deze tensor netwerken kunnen verklaren

hoe de AdS zwaartekrachtstheorie voortkomt uit de CFT. Ook heb ik gekeken naar tensor netwerken die gebruikt worden om CFT staten met een eindige temperatuur te beschrijven want dit zou iets kunnen zeggen over zwarte gaten.