

# Area Law violations for bipartite Entanglement Entropy in Quantum Spin Chains

Menno van Laarhoven

to obtain the degree of Bachelor of Science in  
Applied Mathematics and Applied Physics  
at Delft University of Technology.

Student number: 4676750  
Project duration: February, 2020 – August, 2020  
Thesis committee: Prof. dr. B. Terhal, QuTech, Supervisor  
Dr. P. M. Visser, EEMCS, Supervisor  
Dr. J. M. Thijssen, TNW  
Prof. dr. D. Gijswijt, EEMCS

An electronic version of this thesis is available at <https://repository.tudelft.nl/>.

# Abstract

The concept of entanglement is one of the distinguishing features in quantum mechanics. Information about one particle can determine the state of another particle. This information and entanglement in subsystems is quantified by the entanglement entropy. The entanglement entropy has become an important research topic in recent years.

Entanglement entropy is expected to grow with the boundary size for systems described by Hamiltonians with local interactions, described by the area law. For one dimension there exists a bound quantifying this area law  $S = \mathcal{O}(2^{1/\Delta E})$  in terms of the energy gap  $\Delta E$  between the ground-state energy and the first excited-state energy. It is an open problem whether the area law holds in higher spatial dimensions. Therefore it is of great interest to study the entanglement entropy in systems that violate the area law.

In this thesis spin chain models with local interactions are studied for arbitrary spin. Hamiltonians with a unique ground state are constructed by mapping spin chains onto (coloured) Dyck and (coloured) Motzkin paths. We show that these models express logarithmic or power law violations of the area law for the bipartite entanglement entropy. These results are presented in table 1.

The known bound of the area law in one dimensional systems is dependent on the energy gap of these models. The energy gap of the Motzkin-path model is investigated by constructing the an orthogonal excited state with small energy. In doing so we associate the Hamiltonian with the Laplacian of a graph.

We present an alternative proof for the bound on the energy gap of the colourless Motzkin-path model than S. Bravyi et al and R. Movassagh et al. We show that the energy gap  $\Delta E$  scales in terms of the chain length  $n$  as  $\Delta E = \mathcal{O}(n^{-2})$  in the limit  $n \rightarrow \infty$ . Therefore  $\Delta E \rightarrow 0$  in this limit, resulting in area-law violations of the entanglement entropy while maintaining the known bound on the entanglement entropy.

Table 1: Leading order entanglement entropy scaling for the half-integer-spin Dyck-path and integer-spin Motzkin-path models.  $c$  is the number of colours.

Spin	Entanglement Entropy
$\frac{1}{2}$	$\log_2(n)$
1	$\log_2(n)$
$\frac{c}{2}$	$\sqrt{n} \log_2(c)$
$c$	$\sqrt{n} \log_2(c)$

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Notation</b>	<b>3</b>
<b>3</b>	<b>Background</b>	<b>4</b>
3.1	Quantum Mechanics . . . . .	4
3.2	Tensor Product . . . . .	5
3.3	Spin . . . . .	6
3.4	Product states and Entangled states . . . . .	7
3.5	Tensor Networks . . . . .	8
<b>4</b>	<b>Entropy</b>	<b>11</b>
4.1	Classical entropy . . . . .	11
4.2	Von Neumann Entropy . . . . .	11
4.3	Properties of the Von Neumann entropy . . . . .	12
4.4	Area Law for the entanglement entropy . . . . .	15
4.5	Tensor Networks and entanglement entropy . . . . .	16
<b>5</b>	<b>Motzkin path spin-1 chains</b>	<b>17</b>
5.1	Motzkin paths . . . . .	17
5.2	Motzkin path operations . . . . .	18
5.3	Motzkin Hamiltonian . . . . .	19
5.4	Entanglement entropy . . . . .	20
<b>6</b>	<b>Energy gap of the Motzkin-path Hamiltonian</b>	<b>24</b>
6.1	Hamiltonian as a graph . . . . .	24
6.2	Distance phase . . . . .	26
6.3	Estimate for the number of edges . . . . .	27
6.4	Distribution of the phase factor . . . . .	29
6.5	Energy gap scaling . . . . .	32
<b>7</b>	<b>Path models for arbitrary Spin</b>	<b>35</b>
7.1	Dyck paths . . . . .	35
7.2	Path Colourings . . . . .	39
7.3	Coloured Motzkin paths . . . . .	39
7.4	Coloured Dyck paths . . . . .	42
<b>8</b>	<b>Discussion and Outlook</b>	<b>46</b>
8.1	Spin-Operator interactions . . . . .	46
8.2	Area Law violations . . . . .	46
8.3	Tensor Network representation of Coloured Path models . . . . .	47
<b>9</b>	<b>Conclusion</b>	<b>50</b>
	<b>References</b>	<b>51</b>

<b>Appendices</b>	<b>53</b>
A Entanglement of the product state . . . . .	53
B Derivation of Schmidt decomposition coefficients . . . . .	53
C Normalization Constant for Phase-Shifted Motzkin-path states . . . . .	55

# 1 Introduction

Over the past century the field of physics has seen a revolution in the understanding of the world around us due to quantum mechanics. This theory distinguishes itself from classical physics through the concept of entanglement. Contrary to classical physics, the knowledge of the system as a whole does not necessarily include the knowledge about its parts. The concept of entanglement is therefore closely related to that of information, linking it to the study of entropy. With the rise of quantum computing, understanding these concepts are fundamental to understanding what can, and cannot be done using quantum computers.

While this relation in information and entanglement is interesting in its own right, a strong relation between gravity, holography and entanglement has emerged [1, 2]. As a result, the concept of entanglement and entropy has become a very active field of study in fundamental high-energy physics.

Entanglement occurs whenever the state of two particles is coupled in such a way that knowledge about the two particles separately is limited. Knowledge about one particle therefore gives you information about the second particle. Before a measurement on this first particle we would say that the second particle has a probability to occur in a certain state. The entanglement entropy is a measure of information that can be gained by performing a measurement on a subsystem. If the subsystem can occur in any state with equal probability we gain a lot of information, resulting in maximum entanglement entropy. If the state of the subsystem is known, the entanglement entropy is zero.

Under local interactions, one expects that entanglement should be a local phenomenon, expressed in a so-called area law. The area law states that the entanglement entropy grows with the boundary of the subsystem. Under the right conditions this can indeed be proven for one dimensional systems, drawing a close relation to the energy gap between the ground-state energy and the first excited-state energy. A proof is not known for higher dimensions, which makes it an interesting and active field of research to this date. This thesis will discuss systems that exhibit long-range entanglement while maintaining local interactions. Understanding these area law violations can lead to important insights for higher spatial dimensions. Computing the entanglement entropy for arbitrary states is often very difficult. We will discuss a set of one-dimensional models that can be associated with a (coloured) path representation. This representation converts the problem from a quantum-mechanical setting into that of combinatorics, resulting in a more manageable analytical derivation for the entanglement entropy. These models express logarithmic and power-law violations of the area law.

This thesis will discuss several examples of one-dimensional quantum spin systems that can be associated with paths and express violations of the area law for the entanglement entropy. Hamiltonians with local interactions and a unique ground state are given for each of these models. An introduction in quantum mechanics and the necessary background is presented in section 3. The concept of entanglement entropy and its properties is then presented in section 4. Associating a quantum spin system with Motzkin paths and computing its entanglement entropy is discussed in section 5. Due to the strong relation of the area law with the energy gap, the energy gap of the Motzkin-path Hamiltonian is investigated in section 6, where a new proof for the energy gap scaling for the Motzkin-path model is given. The construction for spin-1 Motzkin paths is extended for arbitrary spin in section 7. A discussion and conclusion follow in sections 8 and 9 respectively.

This thesis was written as part of the double Bachelor degree program in Applied Mathematics and Applied Physics at Delft University of Technology.

## 2 Notation

$\mathcal{H}$	A Hilbert space describing the state space of a quantum system.
$A^\dagger$	Hermitian conjugate of the matrix $A$ .
$S(A)$	Shorthand for the entanglement entropy of a system $A$ with density matrix $\rho_A$ .
$\bar{z}$	Complex conjugate of $z$ .
$ \psi\rangle$	The vector $ \psi\rangle \in \mathcal{H}$ describing a quantum state.
$\langle\psi $	The Hermitian transpose of the state vector $ \psi\rangle$ .
$\log x$	The logarithm with base $e$ , other bases are denoted using a subscript: $\log_2 x$ .
$\text{diag}(x_1, \dots, x_n)$	A diagonal matrix, where the elements $x_i$ appear on the diagonal.
$\text{tr}(A)$	Trace of the matrix $A$ , the sum over the diagonal elements.
$l, o, r$	Left, flat and right bracket notation for Motzkin and Dyck paths.
$m$	Index indicating the height of a Motzkin path or Dyck path.
$j$	Index indicating the position in a spin chain.
$n$	Length of a spin chain or string.
$N$	Size of a subsystem in a quantum spin chain.
$\hbar$	Reduced Planck's constant, $\hbar = 1.05 \times 10^{-34}$ Js/rad.
$I$	The identity matrix.
$A \otimes B$	The tensor product of $A$ and $B$ .
$S_x, S_y, S_z$	The angular momentum operators.
$S_+, S_-$	The raising and lowering operators.
$\mathcal{M}_n$	The set of Motzkin paths of length $n$ , $\mathcal{M}_n \subset \{l, o, r\}^n$ .
$ \mathcal{M}_n\rangle$	The superposition of Motzkin paths of length $n$ .
$M_n$	The $n$ 'th Motzkin number.
$p_m$	The Schmidt coefficients in a Schmidt decomposition.

## 3 Background

### 3.1 Quantum Mechanics

The theory of quantum mechanics concerns the world of the very small. The origins of the theory are rooted in describing the behaviour of absorption and emission of light in atoms, which could only occur at discrete energy levels, or quanta, which was originally proposed by Max Planck. The following decades the theory was developed rapidly and proved very successful. At the basis of quantum mechanics, is the understanding that particles have probabilistic behaviour. Any system, for example a single particle, is formally described by a unit vector  $|\psi\rangle$  in a Hilbert space  $\mathcal{H}$ . A Hilbert space is a vector space equipped with an inner product, such that the vector space is also complete under the induced norm. The inner product considered in quantum mechanics is often defined in terms of the conjugate operation. For a vector  $|\psi\rangle \in \mathcal{H}$ , its Hermitian conjugate is written  $\langle\psi|$ . If one then considers two vectors  $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}$ , the inner product is written  $\langle\psi_1|\psi_2\rangle$ . The conjugate operation is dependent on the choice of Hilbert space, but often takes the form of the Hermitian conjugate. The exact formalism associates the Hilbert space with a  $C^*$ -algebra, which is beyond the scope of this thesis.

Performing a measurement of an observable variable on a quantum system is described mathematically in terms of Hermitian operators on the Hilbert space of the system. For a Hilbert space identified with  $\mathbb{C}^2$ , an observable in a measurement is described by some  $2 \times 2$  matrix. The possible measurement results are then the eigenvalues of that operator. One expects different observations to yield different real results, which is ensured by the condition for them to be Hermitian operators. A Hermitian operator is always unitarily diagonalizable and has real eigenvalues. As a result the operator has orthonormal eigenvectors, corresponding to the different measurement results. Additionally, these eigenvectors constitute a basis for the Hilbert space. Suppose you perform a measurement on an observable with operator  $A$  on the system  $|\psi\rangle$ . The system  $|\psi\rangle$  can then be decomposed into eigenvectors  $|a_i\rangle$  of the measurement operator. The probability of measuring the result associated to the eigenvector  $|a_i\rangle$  is then given by the inner product  $|\langle a_i|\psi\rangle|^2$ . Decomposing a state  $|\psi\rangle$  into a sum of these eigenvectors yields a superposition of states for the system with their coefficients representing the probability amplitudes of being in that state. After performing the measurement, the state becomes the eigenstate associated to the measured result  $|a_i\rangle$ . In quantum mechanics this is called the collapse of the state into the eigenstate of the measurement.

In quantum mechanics the operator associated to energy is called the Hamiltonian, whose eigenvalues represent the possible energy-values of the system. Classically states tend to move towards states of lower energy due to loss of energy to the environment. Since the eigenvalues are bound, the state associated with the lowest energy-value is of particular interest. By reducing the energy of the environment through temperature and pressure we are able to obtain this state in a laboratory. Besides being interesting for research, most of the universe can be regarded as a low energy environment. This state with lowest energy-value is called the ground state.

The time evolution of a quantum state is characterized by the Schrödinger equation [3]

$$\hat{H}|\psi\rangle = i\hbar \frac{d}{dt}|\psi\rangle.$$

Where  $\hbar$  represents the reduced Planck constant and  $\hat{H}$  is the Hamiltonian operator acting on the state  $|\psi\rangle$ . For a single non-relativistic point particle the Hamiltonian can be defined like its

classical counterpart, as the sum of kinetic energy  $\hat{p}^2/2m$  and the potential energy  $\hat{V}$ , where  $\hat{p}$  is the momentum operator. However it can have different expressions, depending on the choice of basis and the interactions considered.

### 3.2 Tensor Product

Suppose you have two particles described by vectors  $|\psi\rangle_1$  and  $|\psi\rangle_2$  in Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  respectively. Bringing these particles together can then be described by the tensor product of the two vectors  $|\psi\rangle_1 \otimes |\psi\rangle_2$ , which is defined in the following definition [4].

**Definition 1** (Tensor Product). Let  $V_1, V_2$  be vector spaces with bases  $|i\rangle_1$  and  $|j\rangle_2$  respectively. Then the tensor product  $V_1 \otimes V_2$  is the vector space whose basis is given by  $|i\rangle_1 \otimes |j\rangle_2$ . Write  $|v\rangle_1 = \sum_i v_1^i |i\rangle_1 \in V_1$  and  $|v\rangle_2 = \sum_j v_2^j |j\rangle_2 \in V_2$ . Then we define

$$|v\rangle_1 \otimes |v\rangle_2 = \sum_{i,j} v_1^i v_2^j |i\rangle_1 \otimes |j\rangle_2.$$

This definition can be intuitively understood as the vector space describing all possible combinations of vectors. In the case of quantum mechanics, these vector spaces Hilbert spaces. This results in the ability to fully describe the two particles described by a single state vector  $|\psi\rangle_{12} \in \mathcal{H}_1 \otimes \mathcal{H}_2$ . The dimension of this new Hilbert space is given by  $\dim(\mathcal{H}_1) \times \dim(\mathcal{H}_2)$ .

In quantum mechanics, the operators  $A$  and  $B$  acting on state vectors of Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  respectively, a new linear operator on  $\mathcal{H}_1 \otimes \mathcal{H}_2$  can be defined. This results in a way to write  $A \otimes B$  acting on  $|\psi\rangle_1 \otimes |\psi\rangle_2$  according to

$$(A \otimes B) (|\psi\rangle_1 \otimes |\psi\rangle_2) = A |\psi\rangle_1 \otimes B |\psi\rangle_2.$$

The operators that act on a single state can therefore be used to define new operators, acting on the two particles at once. Or the operator can be considered acting on the single particle described by  $\mathcal{H}_1$ . The action of the operator  $A$  on the tensor product state is defined  $A \otimes I$ .

The inner product defined on two Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  can be used to obtain a inner product on the tensor product state,  $\mathcal{H}_1 \otimes \mathcal{H}_2$ , which for  $|\psi\rangle_1, |\phi\rangle_1 \in \mathcal{H}_1$  and  $|\psi\rangle_2, |\phi\rangle_2 \in \mathcal{H}_2$  becomes

$$(\langle\psi|_1 \otimes \langle\psi|_2) |\phi\rangle_1 \otimes |\phi\rangle_2 = \langle\psi|\phi\rangle_1 \langle\psi|\phi\rangle_2.$$

One can quickly check that this inner product is indeed conjugate symmetric, linear and positive definite by using distributivity of the tensor product. The state vector as a tensor product state therefore inherits the necessary properties for describing a quantum system.

The definition for the tensor product is rather abstract, therefore it is often useful to use the Kronecker product as matrix representation of the tensor product. For a  $m \times n$  matrix  $A$  and  $p \times q$  matrix  $B$  we have a  $mp \times nq$  matrix

$$A \otimes B = \begin{bmatrix} A_{1,1}B & A_{1,2}B & \cdots & A_{1,n}B \\ A_{2,1}B & A_{2,2}B & \cdots & A_{2,n}B \\ \vdots & & \ddots & \vdots \\ A_{m,1}B & \cdots & & A_{m,n}B \end{bmatrix}.$$

This is a linear operation and satisfies the definition of the Tensor Product. Using the Kronecker product yields several important and useful properties which are used without mention throughout this thesis. These properties are presented in the next theorem.

**Theorem 1** (Kronecker Product properties). Suppose  $A$  and  $B$  are arbitrary  $n \times m$  and  $p \times q$  matrices respectively. Then the following properties hold,

1. Mixed product property:  $(A \otimes B)(C \otimes D) = AC \otimes BD$ .
2. If  $A$  and  $B$  are invertible matrices,  $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ .
3.  $(A \otimes B)^* = A^* \otimes B^*$ .
4. If  $A$  and  $B$  are square matrices,  $\text{tr}(A \otimes B) = \text{tr}(A)\text{tr}(B)$ .

*Proof.* In order to prove the mixed product property, start by writing  $A = [a_{ij}]$ , and  $C = [c_{ij}]$  as the decomposition of their matrix elements. Then by definition one has  $A \otimes B = [a_{ij}B]$  and  $C \otimes D = [c_{ij}D]$ . Writing out the matrix multiplication, one obtains

$$(A \otimes B)(C \otimes D) = AC \otimes BD = \sum_{k=1}^m (a_{ik}B)(c_{kj}D) = \sum_{k=1}^m \underbrace{a_{ik}c_{kj}}_{[AC]_{ij}} BD.$$

Observing that the term on the right represents the matrix elements of  $AC$ , one can identify this with the Kronecker product  $AC \otimes BD$ .

The inverse property follows from applying the mixed product property and observing that

$$\begin{aligned} (A \otimes B)^{-1}(A \otimes B) &= (A^{-1}A) \otimes (B^{-1}B) = I, \\ (A \otimes B)(A \otimes B)^{-1} &= (AA^{-1}) \otimes (BB^{-1}) = I. \end{aligned}$$

For the hermitian transpose property of the Kronecker product, one can observe that the hermitian transpose acts elementwise on the matrix elements. From this it follows that  $(A \otimes B)^* = A^* \otimes B^*$ . The trace property follows from looking at the matrix elements, from which one obtains that the trace of  $A \otimes B$  becomes

$$\text{tr}(A \otimes B) = \sum_{i=1}^n a_{ii} \text{tr}(B) = \sum_{i=1}^n \sum_{j=1}^p a_{ii} b_{jj} = \left( \sum_{i=1}^n a_{ii} \right) \left( \sum_{j=1}^p b_{jj} \right) = \text{tr}(A)\text{tr}(B).$$

□

### 3.3 Spin

In this thesis we will consider Hilbert spaces with  $2s + 1$  dimensions. It is customary to relate these dimensions to the concept of spin, which is an (intrinsic) angular momentum of elementary particles. Contrary to what the name might suggest, the particles are not spinning about some axis. Instead spin is related to a deep symmetry result in the mathematical formulation of describing these particles with emerging properties similar to that of angular momentum. The spin can take values of  $s = n/2$ , with  $n$  a nonnegative integer, where half-integer spin states have different properties than the integer spin states. This leads to two separate names, with half-integer spin particles being called fermions, and integer spin particles being called bosons.

A spin state can be described by two quantum numbers  $s$  and  $m$ , with  $s$  the total spin as defined before, and  $m$  the projection of the (intrinsic) angular momentum along the  $z$  axis. The quantum numbers  $m$  are able to take  $2s + 1$  different values, ranging from  $-s$  to  $s$ , more specifically  $m \in \{-s, -s+1, \dots, s-1, s\}$ . Note that this results in a Hilbert space of dimension  $2s+1$ . The operators related to these quantum numbers are then defined by the fundamental commutation relations

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x, \quad [S_z, S_x] = i\hbar S_y.$$

Each of the operators  $S_i$  represent the projection of the intrinsic angular momentum along their respective axis. By writing the spin states defined by the quantum numbers  $s$  and  $m$  as  $|s\ m\rangle$ , one has that these operators satisfy

$$\mathbf{S}^2 |s\ m\rangle = \hbar^2 s(s+1) |s\ m\rangle, \quad S_z |s\ m\rangle = \hbar m |s\ m\rangle.$$

The operator  $\mathbf{S}^2$  is the total intrinsic angular momentum, defined as  $\mathbf{S}^2 = S_x^2 + S_y^2 + S_z^2$ . By deciding on a specific basis, the states  $|s\ m\rangle$  can be represented as vectors and the operators  $S_k$  as matrices. These matrices are then fully defined by the relations defined above.

The elementary particles of the standard model can only have a spin of 0,  $\frac{1}{2}$  or 1. Other orders of spins are obtained through superposition of multiple lower order spins. This way higher order spins for the nuclei of atoms are obtained.

### 3.4 Product states and Entangled states

Up to now we have mainly been considering states that could be described by a single state vector  $|\psi\rangle$ . These state vectors are called pure states. Let us consider the superposition of two qubits

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|0\rangle_1 \otimes |0\rangle_2 - |1\rangle_1 \otimes |1\rangle_2). \quad (1)$$

The two particles together have Hilbert space  $\mathcal{H}_1 \otimes \mathcal{H}_2$ . Isolating a single particle, one could expect its state to be described by a vector  $|\psi_1\rangle$  such that together with the other particle one has  $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$ . The state in expression 1 is can not be decomposed into a product. If we isolate particle 1, we say that is in state  $|0\rangle$  or in state  $|1\rangle$  with probability  $\frac{1}{2}$ . Two particles are called entangled when they cannot be written as a product. This can be generalized for arbitrary systems by saying that a state is in a statistical ensemble of pure states  $|\psi\rangle$ . Consider a system occurring in a pure state  $|\psi_i\rangle$  with probability  $p_i$ . Then we describe this system in terms of the density operator. The density operator is defined

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|, \quad (2)$$

This gives us a mathematical description of a system that cannot be written as a single vector. When the density operator of a state cannot be written as the product of a single vector  $|\psi\rangle \langle \psi|$  we call it a mixed state.

If the states  $|\psi_i\rangle$  are orthogonal, we obtain that these are the eigenvectors of the density matrix  $\rho$  with eigenvalues  $p_i$ , which follows from applying equation 2 to  $\rho|\psi_i\rangle$ . We observe that the eigenvalues satisfy  $0 \leq p_i \leq 1$  and  $\sum_i p_i = 1$ , due to them representing probabilities. Additionally,

if  $\rho$  represents a pure state, the eigenvalues will be 0 and 1 respectively. The eigenvector associated with eigenvalue 1 is the state that the particle is in. By distributivity we obtain that  $\rho$  is Hermitian

$$\rho^\dagger = \sum_i \overline{p_i} (|\psi_i\rangle \langle \psi_i|)^\dagger = \sum_i p_i (\langle \psi_i|)^\dagger (|\psi_i\rangle)^\dagger = \sum_i p_i |\psi_i\rangle \langle \psi_i|.$$

Since we are concerned with separating a state with density operator  $\rho$  into its subsystems, one would like to obtain the density matrix of a subsystem. This can be done by introducing the partial trace. This is defined as the unique linear map  $\text{tr}_B : \mathcal{H}_A \otimes \mathcal{H}_B \rightarrow \mathcal{H}_A$  such that

$$\text{tr}_B(\rho_A \otimes \rho_B) = \rho_A \text{tr}(\rho_B),$$

where  $\rho_A \in \mathcal{H}_A$  and  $\rho_B \in \mathcal{H}_B$ , denoting the subsystems  $A$  and  $B$ . It is not immediately clear that this definition covers the situation of mixed states, in which the combined state  $\rho_{AB}$  cannot be written as  $\rho_A \otimes \rho_B$ . In order to see this, one needs to consider the Schmidt decomposition. The Schmidt decomposition is additionally one of the important tools in analytically calculating the entanglement entropy and will be used extensively throughout this thesis.

**Theorem 2** (Schmidt Decomposition). Let  $\mathcal{H}_1$  and  $\mathcal{H}_2$  be two Hilbert spaces, then for  $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$  there exist orthonormal bases  $\{|u_1\rangle, |u_2\rangle, \dots, |u_m\rangle\} \subseteq \mathcal{H}_1$  and  $\{|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle\} \subseteq \mathcal{H}_2$  such that

$$|\psi\rangle = \sum_i \sqrt{p_i} |u_i\rangle \otimes |v_i\rangle.$$

Here  $p_i$  are real and non-negative.

A proof of this theorem can be found in texts covering linear algebra and quantum information theory [5]. The linearity of the trace operator can now be applied on this Schmidt decomposition of the vector  $|\psi\rangle$  describing the state  $\rho_{AB}$ . Identifying the two Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  with two subsystems  $A$  and  $B$  respectively. One obtains

$$\begin{aligned} \rho_A &= \text{tr}_B \left( \sum_{i,j} \sqrt{p_i p_j} |u_i\rangle \langle u_j| \otimes |v_i\rangle \langle v_j| \right) = \sum_{i,j} \sqrt{p_i p_j} \text{tr}_B (|u_i\rangle \langle u_j| \otimes |v_i\rangle \langle v_j|) \\ &= \sum_{i,j} \sqrt{p_i p_j} |u_i\rangle \langle u_j| \underbrace{\text{tr}(|v_i\rangle \langle v_j|)}_{\delta_{i,j}} = \sum_i p_i |u_i\rangle \langle u_i|. \end{aligned}$$

From the orthonormality of the bases it follows that  $\text{tr}(|v_i\rangle \langle v_j|) = 1$  if and only if  $i = j$ , being 0 otherwise. Hence one obtains that our definition of the partial trace  $\text{tr}_B$  satisfies the desired properties, reducing a density matrix into its components. Additionally another important property can be derived from the above operation of  $\text{tr}_B$  as well. The coefficients  $p_i$  from the Schmidt decomposition are the eigenvalues of the reduced density matrix  $\rho_A$ . Another interesting feature of this decomposition is that both the reduced density matrices  $\rho_A$  and  $\rho_B = \text{tr}_A(\rho)$  have the same eigenvalues, by symmetry of the argument above. This detail will result in an important property of the entanglement entropy in the next section.

### 3.5 Tensor Networks

Tensors are generalizations of vectors and matrices. We have already seen the tensor product, which were introduced as a generalization of products between vector spaces. The tensors are exactly the

generalizations of these vector spaces. Before introducing them formally, we consider the column vector  $|v\rangle \in V$  with dimension  $\dim V = d$ . This induces a dual space  $V^*$ , which in a particular choice of basis can be considered as a tangent space to the vector space  $V$ . The dual space  $V^*$  can also be considered as the space of linear operators on  $V$ . As an example, if one identifies  $V$  as a space of column vectors we have that  $V^*$  can be considered as row vectors acting on these column vectors. We identify the dual space by the Hermitian transpose of the elements,  $\langle v| \in V^*$ . A rank  $(p, q)$  tensor  $T$  is then an element of the space

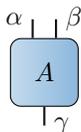
$$\underbrace{V \otimes V \otimes \cdots \otimes V}_p \otimes \underbrace{V^* \otimes V^* \otimes \cdots \otimes V^*}_q.$$

It follows that column vectors are  $(1, 0)$  tensors and matrices  $(1, 1)$  tensors. Additionally, the scalars are identified as rank  $(0, 0)$  tensors. The standard notation directly illustrates the rank of the tensor. Mostly we can consider tensors as arrays of numbers that require  $p+q$  indices to identify its elements. If we identify the basis elements of  $V$  by  $\mathbf{e}_i$  and of  $V^*$  by  $\mathbf{e}^j$ , a tensor  $T$  of  $(p, q)$  can be written

$$T = T_{j_1, \dots, j_q}^{i_1, \dots, i_p} \mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \cdots \otimes \mathbf{e}_{i_p} \otimes \mathbf{e}^{j_1} \otimes \mathbf{e}^{j_2} \otimes \cdots \otimes \mathbf{e}^{j_q}.$$

Note that a summation over all  $i$  and all  $j$  is implied. Whenever a  $i$  and  $j$  appear both as a raised index or a subscript such a summation is implied from now on, which is the standard Einstein summation convention. Additionally the basis notation is dropped so that  $T = T_{j_1, \dots, j_q}^{i_1, \dots, i_p}$ . Summation over an index that occurs both at the top and bottom is called tensor contraction.

The notation for tensors is rather cumbersome due to the many indices involved for higher rank tensors. A visual notation was therefore introduced by Penrose [6], which eventually evolved into the tensor network notation. Consider a rank- $(2, 1)$  tensor  $A_\gamma^{\alpha\beta}$ , which can then be represented by a node with 3 legs, 2 on top and 1 on the bottom. Which becomes



This notation is convenient because tensor contraction is indicated by the connection of the legs of the tensor nodes. Taking the transpose of a tensor, which is equivalent to swapping the indices from top to bottom and vice versa, is the same as flipping the tensor vertically in the network representation. Whenever two nodes are placed next to each other, they represent a tensor product. These are illustrated for three tensors  $A$ ,  $B$  and  $C$  of rank  $(2, 1)$ ,  $(2, 2)$  and  $(1, 2)$  respectively, as

follows

$$A_{\gamma}^{\alpha\beta} B_{\beta\mu}^{\epsilon\delta} C_{\zeta\chi}^{\mu} =$$

$$= [A \otimes C]_{\gamma\zeta\chi}^{\alpha\beta\mu} B_{\beta\mu}^{\epsilon\delta}.$$

The equation above illustrates the power of the visual notation. The structure in the clutter of indices is clearly illustrated by the connection between nodes.

We now have a visual way of writing down quantum states. A quantum product state described by a vector  $|\phi\rangle = |\psi\rangle_1 \otimes |\psi\rangle_2 \otimes \dots \otimes |\psi\rangle_n$ , consisting of  $n$  rank (1,0) tensors can now be written as a sequence of states,

$$|\phi\rangle = \begin{array}{c} | \\ \psi_1 \end{array} \begin{array}{c} | \\ \psi_2 \end{array} \dots \begin{array}{c} | \\ \psi_n \end{array},$$

which is not particularly interesting. For entangled states however, we are unable to write it down in such a sequence. Consider a state represented by a rank (2,0) tensor  $|\phi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ , which can be expanded using a Schmidt decomposition, which has a diagrammatic representation

$$|\phi\rangle = \begin{array}{c} | \\ \psi \end{array} = \begin{array}{c} | \\ \psi_A \end{array} \begin{array}{c} | \\ \psi_B \end{array} \begin{array}{c} \circ \\ D \end{array} = \sum_i \sqrt{p_i} |\psi_i\rangle_A \otimes |\psi_i\rangle_B.$$

The matrix  $D$  represents a diagonal matrix with the Schmidt decomposition values  $\sqrt{p_i}$  on the diagonal.

## 4 Entropy

### 4.1 Classical entropy

Classically the Boltzmann entropy of a thermodynamic system  $A$  is defined by

$$S_c(A) = k_B \log \Omega_A, \quad (3)$$

with  $k_B$  the Boltzmann constant and  $\Omega_A$  the number of possible states in a system  $A$  with fixed energy [7]. Entropy is often called a measure of disorder or chaos. A large number of possible states can be seen as disordered and can therefore be quantified by equation 3. An equivalent definition for entropy can be given in terms of probabilities

$$S = - \sum_i p_i \log_2 p_i. \quad (4)$$

Where  $p_i$  are the probabilities that the system is in a state  $i$ . Since probabilities must satisfy  $0 \leq p_i \leq 1$  and  $\sum_i p_i = 1$ , we have that  $S$  increases if there are more possible states with smaller probabilities compared to a few states with greater probabilities. This expression for entropy is used extensively in Information theory, and is often called the Shannon entropy or Gibbs entropy. The classical Boltzmann entropy in expression 3 is obtained whenever all states can occur with equal probability, including the Boltzmann constant.

Entropy is a fascinating topic in theoretical physics. This is mostly incorporated in the second law of thermodynamics. This states that, in an closed system, entropy does not decrease in time. This law has been tested extensively and is also intuitive, the milk added to your morning coffee will mix uniformly, as opposed to suddenly separating away from the coffee. In terms of entropy, there are more possible configurations of molecules in which the coffee and milk particles are completely mixed as opposed to a configuration with coffee particles and milk particles collectively together. Note that this is not a correct illustration for entropy, since we are talking about the system of milk and coffee as a whole already, but it does give some intuition for what entropy entails. For this reason, in which a system will move towards a system of higher entropy, it is often said that entropy is the only thing in physics that dictates the flow of time.

### 4.2 Von Neumann Entropy

A quantum-mechanical definition for the entropy is due to Von Neumann [8]. This quantity is consequently called the Von Neumann entropy. In this paper the information-theoretic version is used, which is defined as

$$S(\rho) = -\text{tr}(\rho \log_2 \rho), \quad (5)$$

where  $\rho$  denotes the density matrix of a system. Whenever one has a pure bipartite system and considers the reduced density matrix  $\rho_A$  of a subsystem  $A$ , the entropy of the subsystem is called the entanglement entropy. We will use a shorthand notation  $S(\rho_A) = S(A)$  when talking about the entanglement entropy of a system  $A$  with density matrix  $\rho_A$ . The above definition of the Von Neumann entropy is tedious to work due the logarithm of the density matrix, which, for matrices close to a multiple of the identity, can be calculated using its Taylor expansion [9].

The difficulties involved with calculating the matrix logarithm can be overcome by considering an equivalent definition of equation 5, given by

$$S(\rho) = - \sum_i p_i \log_2 p_i. \quad (6)$$

In this equation,  $p_i$  represent the eigenvalues of the density matrix  $\rho$ . This definition is closely related to the classical definition for the Shannon entropy in equation 4, where the probabilities are replaced with the eigenvalues of the density matrix  $\rho$ . Recall that those eigenvalues represented the probabilities of specific states occurring.

One can show that the definitions in 5 and 6 are equivalent by diagonalizing the density matrix  $\rho = UVU^\dagger$  with  $V = \text{diag}(p_1, p_2, \dots, p_n)$  and  $U$  unitary. Recall that  $\rho_A$  is Hermitian, and hence diagonalizable. Using that  $(ABA^{-1})^n = ABA^{-1}ABA^{-1} \dots BA^{-1} = AB^nA^{-1}$ , this yields

$$\begin{aligned} S(\rho) &= -\text{tr}(UVU^\dagger \log_2(UVU^\dagger)) = -\text{tr}\left(UVU^\dagger \left[ \frac{1}{\log 2} \sum_n \frac{(-1)^{n+1}}{n} (UVU^\dagger - I)^n \right]\right) \\ &= -\text{tr}\left(UVU^\dagger \left[ \frac{1}{\log 2} \sum_n \frac{(-1)^{n+1}}{n} (U(V - I)U^\dagger)^n \right]\right) \\ &= -\text{tr}\left(UVU^\dagger U \left[ \frac{1}{\log 2} \sum_n \frac{(-1)^{n+1}}{n} (V - I)^n \right] U^\dagger\right) \\ &= -\text{tr}(V \log_2 V) = - \sum_i p_i \log_2 p_i. \end{aligned}$$

Where the cyclic property of the trace  $\text{tr}(AB) = \text{tr}(BA)$  was used in the last step. The Taylor expansion is however only convergent for  $\|D - I\| \leq 1$ , with the matrix norm  $\|A\| = \lambda$  defined by  $\langle v|A|v\rangle^2 \leq \lambda \langle v|v\rangle^2$  for all  $|v\rangle$ . It follows that  $\|A\|$  is the eigenvalue  $\lambda'$ , such that  $|\lambda'|$  is largest. All eigenvalues  $p_i$  of the density matrix satisfy  $0 \leq p_i \leq 1$ , and are real. Hence the eigenvalues  $p'_i$  of  $\|\rho - I\| \leq 1$  satisfy  $0 \leq |p'_i| \leq 1$ . Since  $V$  has the same eigenvalues as  $\rho$ , the Taylor expansion converges in the above argument.

There is an issue with equation 6 for eigenvalues  $p_i = 0$ , in this case the limit to 0 is considered,  $\lim_{p_i \rightarrow 0} p_i \log_2 p_i = 0$ . This is in alignment with the physical aspect of entropy, in which states that cannot occur should not contribute to the entropy.

### 4.3 Properties of the Von Neumann entropy

The abstract definition of the Von Neumann entropy makes it difficult to compute and extract meaningful conclusions. Hence one can derive some important properties of the Von Neumann entropy. These properties are presented in theorem 3.

**Theorem 3** (Properties of the Von Neumann entropy). Let  $\rho$  be an arbitrary density matrix of a quantum state, then

1.  $S(\rho) \geq 0$  with equality if and only if  $\rho$  is the density matrix of a pure state.

2.  $S(\rho) \leq \log_2 d$ , where  $d$  is the dimension of the Hilbert Space with equality if and only if  $\rho$  is the density matrix of a completely mixed state  $\rho = I/d$ .
3. If systems  $A$  and  $B$ , with density matrices  $\rho_A$  and  $\rho_B$  respectively, form a pure state  $AB$ , then  $S(\rho_A) = S(\rho_B)$ .
4. For two density matrices  $\rho_A$  and  $\rho_B$  we have  $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$ .

The statement that the Von Neumann entropy is nonnegative follows from the fact that  $0 \leq p_i \leq 1$ , resulting in  $\log_2 p_i \leq 0$  and cancelling with the negative sign in the Von Neumann entropy definition. For a pure state, the eigenvalues of  $\rho$  are 0 and 1, and hence  $S(\rho) = 0$ . This is consistent with the notion of entropy, if a system is in a pure state, we know exactly what state it is in, and hence the Von Neumann entropy should be minimal.

The bound on the Von Neumann entropy given by the second property, is derived in the book by Nielsen and Chuang [5]. This is again consistent with what one expects for entropy, if a system is in the completely mixed state it represents the system in which you have a minimum of information. This is the system where any state can occur with equal probability and should have maximal entropy.

The third property follows directly from the Schmidt decomposition, given in theorem 2, and the succeeding paragraph, stating that the reduced density matrices  $\rho_A$  and  $\rho_B$  have the same eigenvalues. This makes sense, if the full state is exactly known, but there is some uncertainty in the two subsystems  $A$  and  $B$ , all the information that is necessary to know the state of  $A$  is contained in  $B$ , and the same information the other way around.

The fourth property in theorem 3 requires some calculations, presented in appendix A. This property is analogous to the classical image of entropy. Bringing two systems together, with a certain entropy will result in a sum of their components. The density matrix  $\rho_{AB}$  is often not of the form  $\rho_A \otimes \rho_B$ . We can however give a bounds on the entropy of a combined system  $AB$  given the entropy of its components  $A$  and  $B$ , quantified by

$$|S(A) - S(B)| \leq S(AB) \leq S(A) + S(B). \quad (7)$$

The subadditivity property of the entanglement entropy  $S(A) + S(B) \geq S(AB)$  is intuitive. Any correlation between the system  $A$  and system  $B$  will reduce the entropy of the combined system compared to the sum of their individual entropies, due to the overlap in information in systems  $A$  and  $B$ . Hence, equality follows when the systems are fully uncorrelated:  $\rho_{AB} = \rho_A \otimes \rho_B$ , as seen before. The inequality  $|S(A) - S(B)| \leq S(AB)$  is often called the triangle inequality, which is difficult to relate to anything physical.

Figure 1a can help to gain intuition in the behaviour of this entanglement entropy for small subsystems. A set of qubits are arranged in a circle, and entanglement between two qubits is denoted by the dotted lines. Consider the boundary of these subsystems  $A$  and  $B$  illustrated in figure 1a by the dashed lines. The entanglement entropy  $S(A)$  and  $S(B)$  will scale with the number of intersecting lines with this boundary. The inequalities from equation 7 can be verified in this graphical representation. This representation can also illustrate when  $|S(A) - S(B)| = S(AB)$  will hold, which is the state in which  $A$  is the completely mixed state with the rest of the system, but is not entangled

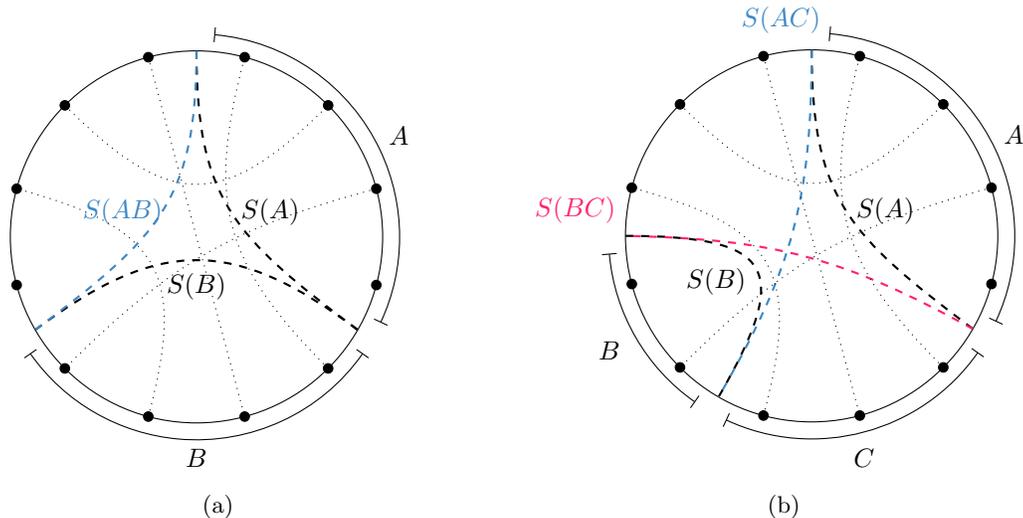


Figure 1: Visual representation of the entanglement entropy for two (a) or three (b) subsystems respectively. The black dots around the circle represent particles and the dotted lines connecting two particles represent entanglement. The subsystems  $A$ ,  $B$  and  $C$  are labeled. The entanglement entropy for these subsystems scales with the lines intersecting with the subsystem boundary, illustrated by the dashed lines.

with system  $B$ . The entanglement entropy inequalities given in 7 can be extended by introducing a third system  $C$ . This results in the strong subadditivity property, given by

$$\begin{aligned} S(ABC) + S(B) &\leq S(AB) + S(BC), \\ S(A) + S(B) &\leq S(AC) + S(BC). \end{aligned} \tag{8}$$

Both lines in equation 8 are equivalent, in which the first can be derived from the second by introducing an auxiliary system  $R$ , such that  $ABCR$  is a pure state. Applying the second inequality we obtain

$$S(R) + S(B) \leq S(RC) + S(BC).$$

Note that because  $ABCR$  is a pure state, that  $S(R) = S(ABC)$  and that  $S(RC) = S(AB)$ , from which we obtain the first inequality in equation 8. This strong Subadditivity property of the entanglement entropy was originally conjectured in 1968 by Langford and Robinson [10]. Proving this property deemed difficult and relied heavily on the study of convex functions. In 1973 a first proof was published by Lieb and Ruskai for the strong subadditivity, which is difficult and relies on some deep results in linear algebra [11]. Easier proofs have been developed since, and one can be found in Nielsens book [5].

The illustration that held for the inequalities involving two systems can be extended for these three systems  $A$ ,  $B$  and  $C$  as well, which is presented in figure 1b. Again the inequalities can be verified by considering the dotted lines that intersect the system boundaries. The second inequality from equation 8 effectively holds due to some of the information in the system  $AC$  being captured in

the system  $B$  and similarly for systems  $BC$  and  $A$ , while the information of  $A$  and  $B$  not captured by the other, is still present in the entropies for the combined systems  $AC$  and  $BC$ . The visual representation used here quickly becomes too complex for larger systems because combined systems must be adjacent to each other. The illustrations presented in figure 1 do not necessarily have to be presented in a circle.

#### 4.4 Area Law for the entanglement entropy

The study of the thermodynamic entropy of black holes in the 20th century by, among others, Bekenstein and Hawking led to the now famous black-hole entropy, given by  $S_{BH} = \frac{c^3}{4G\hbar} A$ . The entropy is related to the area  $A$  of the black-hole event horizon, through the gravitational constant  $G$ , the speed of light  $c$  and the reduced Planck constant  $\hbar$ . A series of arguments were put forward for this entropy scaling, in which the entanglement entropy took a central role [12]. Consider a

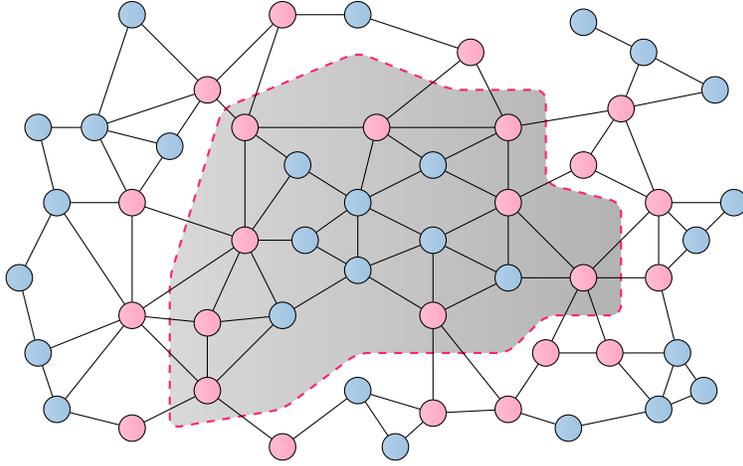


Figure 2: Near neighbour interactions between particles. The shaded area is the subsystem  $A$  with the surroundings, and particles that have interactions across its boundary  $\partial A$  are shaded red.

volume of spins in the ground state of a Hamiltonian with spatially local interactions, as illustrated in figure 2. Entanglement between particles is a result of these local interactions, and hence it is expected that the correlation between two regions of space will occur in the region near the boundary. This behaviour is illustrated in figure 2 for the subsystem  $A$  defined with the shaded area by the particles shaded in red. One can therefore expect the entropy to scale according to the size of the boundary  $|\partial A|$ . For a volume, the boundary is defined by the surface enclosing its volume, and its entropy therefore scales according to its surface area, clarifying the name area law. For two dimensional systems this boundary size is its perimeter. This area law is denoted

$$S(A) \sim K |\partial A|, \tag{9}$$

for a system  $A$  embedded in a space with spatial dimension greater than 1 [13]. The constant  $K$  depends on, among other things, the dimension and geometric properties of the space considered. Equation 9 determines the leading scaling term of the entropy, often however, there are additional

terms contributing to the entanglement entropy. This contribution often take the form of a logarithmic expression. For one-dimensional gapped quantum systems with short-range finite-strength interactions, there exists an analytical bounds on the entropy entailing the area law [14]. For Hamiltonians with a  $D$ -dimensional Hilbert space at each site, with nondegenerate ground state, acting only on adjacent particles with finite strength, this bound reads

$$S(\rho) \leq \frac{c_1}{\Delta E} \log\left(\frac{c_2}{\Delta E}\right) \log(D) 2^{c_3 \log(D)/\Delta E}, \quad (10)$$

for sufficiently small  $\Delta E$ . Where  $\Delta E$  denotes the energy gap between the ground state and the first excited state,  $c_1, c_2$  and  $c_3$  denote constants, dependent on certain correlative properties of the system [15]. Hence, violations of the area law in one dimension may occur if this energy gap closes [16]. There exist examples of violations of the area law for gapless systems [17, 18]. The gapless systems can be classified into three classes, those that obey the area law, have at most logarithmic violation or have more extreme violations [19]. This classification is based on a close relation to classical thermodynamics. It is unknown if the area law would hold for systems with a gapped ground state for dimensions larger than 1.

## 4.5 Tensor Networks and entanglement entropy

Isolating subsystems of tensor networks presents a very simple measure of entanglement entropy on these tensor networks. Suppose you can separate a tensor network  $|\psi\rangle$  in two systems  $A$  and  $B$  with indices  $\hat{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_j\}$ , where each  $\alpha_i$  can take up to  $D$  values. This is equivalent to the  $j$  bonds between the two subsystems. We perform a Schmidt decomposition on the system  $|\psi\rangle$  in terms of states on the subsystems  $|A(\hat{\alpha})\rangle$  and  $|B(\hat{\alpha})\rangle$ , giving

$$|\psi\rangle = \sum_{\hat{\alpha}} \sqrt{p_{\hat{\alpha}}} |A(\hat{\alpha})\rangle \otimes |B(\hat{\alpha})\rangle,$$

where  $p_{\hat{\alpha}}$  are the Schmidt coefficients. The summation is over all combinations of the different  $\alpha_i$ . Note there are  $j$  bonds which can take  $D$  values. The maximum entanglement entropy that can exist for the subsystem  $A$  occurs whenever  $A$  is completely mixed, represented by  $p_{\hat{\alpha}}$  being uniformly distributed. That is,  $p_{\hat{\alpha}} = 1/D^j$ , obtaining

$$S(A) \leq - \sum_{\hat{\alpha}} p_{\hat{\alpha}} \log_2 p_{\hat{\alpha}} = -D^j D^{-j} \log_2 D^{-j} = j \log_2 D.$$

Hence the maximum entanglement entropy that can occur in a tensor network is bound by the number of bonds that need to be cut in order to separate the subsystem from the rest of the tensor networks. For the matrix product state we have that  $j = 2$  or  $j = 1$  dependent on the choice of subsystem, only the bonds at the boundary need to be cut in order to separate the subsystem.

## 5 Motzkin path spin-1 chains

### 5.1 Motzkin paths

The emergence of long-range entanglement in the ground state of a quantum spin chain with short-range interactions is a fascinating subject of research. One model that can be constructed with a unique ground state and large entanglement property is a spin-1 chain model proposed by Bravyi et al [20]. The three basis states of the spin-1 particle in this model are denoted as the states  $\{l, o, r\}$ , such that the single spin-state is written  $|\psi\rangle = a|l\rangle + b|o\rangle + c|r\rangle$ . For a chain of multiple particles, this yields base states of the form  $|l\rangle \otimes |r\rangle \otimes \dots \otimes |o\rangle = |lr\dots o\rangle$ . Therefore the states of a chain of  $n$  spin-1 particles  $|s\rangle$  can be defined using strings  $s \in \{l, o, r\}^n$ , in which the elements  $l, o$  and  $r$  represent the basis of a single spin-1 particle. Bravyi et al impose a condition on the possible strings, identifying them with Motzkin paths. An example of a Motzkin path can be found in figure 3.

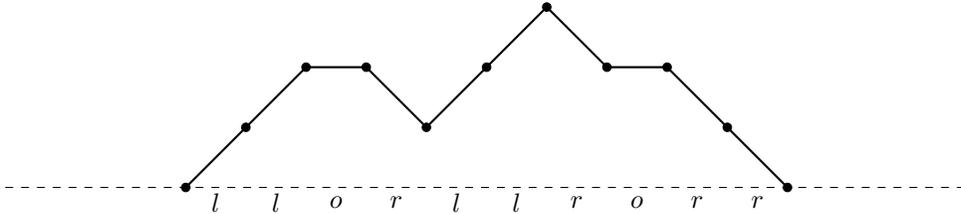


Figure 3: The Motzkin path that can be identified with the state  $|llorllrorr\rangle$ . The base line can be identified with the dashed line.

Effectively a Motzkin path is a path with three possible movement directions: up, horizontal or down, with the condition that it never drops below the dashed line, starts and ends at the base line. If one denotes  $L_j(s)$  and  $R_j(s)$  to be the number of  $l$ 's and  $r$ 's in the first  $j$  letters of a string  $s$  respectively, one can define a Motzkin path as in definition 2, using the notation  $\Sigma_n = \{l, o, r\}^n$  as the set of strings of length  $n$ .

**Definition 2.** A string  $s \in \Sigma_n$  is called a Motzkin path if

1. Any initial segment with length  $j \leq n$  of  $s$  has  $L_j(s) \geq R_j(s)$
2. The total number of  $l$  is equal to  $r$ , that is:  $L_n(s) = R_n(s)$

Whenever a string contains an  $l$ , there needs to exist an  $r$  later in the chain in order to make the path a Motzkin path. The difficulty lies in defining a local Hamiltonian, that does not directly act on 2 particles over the long distance, for example the left-most and right-most particle. We will denote the set of all Motzkin paths  $\mathcal{M}_n \subset \Sigma_n$ .

The number of different Motzkin paths of length  $n$  are given by the Motzkin numbers  $M_n$ . These have been widely studied in different contexts. We will give an expression for the  $n$ 'th Motzkin number shortly.

## 5.2 Motzkin path operations

The goal of Bravyi et al. was to construct a Hamiltonian such that the ground state is the normalized Motzkin state  $|\mathcal{M}_n\rangle = \sum_{s \in \mathcal{M}_n} |s\rangle$ , that is, the uniform superposition of all Motzkin paths of length  $n$ . To construct this Hamiltonian, one can consider three local exchange operations given by the exchange  $oo \leftrightarrow lr$ ,  $ol \leftrightarrow lo$  and  $or \leftrightarrow ro$ . Graphically these exchange operations can be identified with the operations in figure 4.

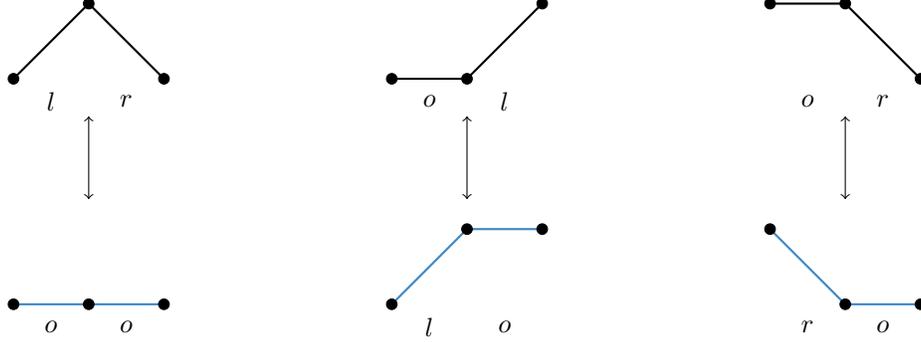


Figure 4: The three local string exchange operations  $oo \leftrightarrow lr$ ,  $ol \leftrightarrow lo$  and  $or \leftrightarrow ro$ .

Using these operations, one can say that an arbitrary string  $s$  is equivalent to  $t$  if  $t$  can be obtained from  $s$  by applying these operations repeatedly. This leads to a specific type of string, which can be defined with integers  $p, q \geq 0$  with  $p + q \leq n$ , yielding a string of length  $n$

$$c_{p,q} = \underbrace{rr \dots r}_p \underbrace{oo \dots o}_{n-p-q} \underbrace{ll \dots l}_q.$$

These type of strings lead to the following lemma.

**Lemma 4.** An arbitrary string  $s \in \Sigma_n$  is equivalent to a unique  $c_{p,q}$ .

*Proof.* Let  $s \in \Sigma_n$ . Then all substrings  $lr$ ,  $lor$  or  $lo \dots or$  can be converted to the  $oo \dots o$  string using the exchange operations. Doing this repeatedly comes down to flattening the peaks in the string. Eventually this will lead to a string  $\dots ro \dots ol \dots$ , with only  $r$ 's and  $o$ 's to the left of  $r$ , and,  $l$ 's and  $o$ 's to the right of  $l$ , which cannot be flattened any further. Exchange operations can then be used to position the  $r$ 's and  $l$ 's to the left and right respectively. Hence  $s$  is equivalent to some  $c_{p,q}$ .

To show that strings  $c_{p,q}$ , are not equivalent to each other, one can argue as follows. Suppose  $c_{p,q}$  and  $c_{p',q'}$  are equivalent, with  $p \geq p'$ . Note that  $R_p(s) - L_p(s) \leq p'$  for strings  $s$  equivalent to  $c_{p',q'}$ , which is the number of unmatched  $r$ 's in the first  $p$  letters of a string  $s$ . However, then  $p = R_p(c_{p,q}) - L_p(c_{p,q}) \leq p'$ , a contradiction unless  $p = p'$ . By symmetry, the same holds for  $q$  and  $q'$ , and thus  $c_{p,q} = c_{p',q'}$ .  $\square$

Note that the Motzkin paths are equivalent to the string  $c_{0,0}$  since for these strings there can not exist an  $r$  without preceding  $l$ . Lemma 4 leads to the equivalence classes  $C_{p,q}$  in which all  $s \in \Sigma_n$

equivalent to  $c_{p,q}$  are contained, which will be used to define normalized states  $|C_{p,q}\rangle$ , consisting of uniform superpositions of strings equivalent to  $c_{p,q}$ . That is,

$$|C_{p,q}\rangle = \frac{1}{K} \sum_{s \in C_{p,q}} |s\rangle,$$

where  $K$  is a normalization constant.

### 5.3 Motzkin Hamiltonian

In order to construct a Hamiltonian with the Motzkin paths as a ground state, one can approach it algebraically. The Hamiltonian describes the interactions in the system. A physical model might implement these interactions through a magnetic force. The Hamiltonian might be scaled by the magnetic moment in that case, but it does not affect the ground state of the model.

If one considers the Motzkin path operations illustrated in figure 4, one can come up with three orthogonal states, effectively implementing the aforementioned Motzkin path operations

$$|\phi\rangle = |oo\rangle - |lr\rangle, \quad |\psi_l\rangle = |ol\rangle - |lo\rangle, \quad |\psi_r\rangle = |or\rangle - |ro\rangle.$$

These states are orthogonal to the states  $|oo\rangle + |lr\rangle$ ,  $|ol\rangle + |lo\rangle$  and  $|or\rangle + |ro\rangle$ , for example,  $\langle\phi|(|oo\rangle + |lr\rangle) = \langle oo|oo\rangle + \langle oo|lr\rangle - \langle lr|oo\rangle - \langle lr|lr\rangle = 1 - 1 = 0$ . It follows that the operator defined as  $\Pi = |\phi\rangle\langle\phi| + |\psi_l\rangle\langle\psi_l| + |\psi_r\rangle\langle\psi_r|$  will annihilate any superposition of the three orthogonal states. Additionally it will annihilate the  $|rl\rangle$ ,  $|ll\rangle$  and  $|rr\rangle$  states. Considering only the two particles, annihilated states do not contribute to the energy in the system, while states that are not annihilated are penalized by some amount of energy. These penalized states are therefore not in the ground state of the system.

The idea is to implement the operator associated to these moves in the Hamiltonian, such that states that can be achieved with the different Motzkin operations have the same number of annihilated terms, and therefore the same energy. Since we are dealing with strings consisting of more than two spins, we will write  $\Pi_{j,j+1}$  for the operator  $\Pi$  acting on the pair of neighbouring spins located at positions  $j$  and  $j+1$ . In order to construct a state that is annihilated by all  $\Pi_{j,j+1}$  with component  $|s\rangle$ , consider the iterative projection  $\sum_j \Pi_{j,j+1} |s\rangle = m |s\rangle - \sum_{s'} |s'\rangle$ , where  $s'$  are the strings on which the string exchange operation is performed once at arbitrary location, and  $m$  is the number of different  $s'$ . Then for each  $|s'\rangle$ , one has  $\sum_j \Pi_{j,j+1} |s'\rangle = m' |s'\rangle - |s\rangle - \sum_{s''} |s''\rangle$ , with  $s''$  the strings on which the string exchange operation is performed twice at arbitrary location, and  $m'$  the number of different  $s''$ . Iterating these operations, the  $m$  and  $m'$  coefficients cancel out, and result in the annihilated state, yielding the state  $|C_{p,q}\rangle$ , the superposition of all strings in equivalence class  $C_{p,q}$ . For example, consider the effect of the operator  $\Pi$  on the Motzkin paths of length 3

$$\begin{aligned} (\Pi_{1,2} + \Pi_{2,3}) |ooo\rangle &= 2 |ooo\rangle - |lro\rangle - |olr\rangle, \\ (\Pi_{1,2} + \Pi_{2,3}) |lro\rangle &= 2 |lro\rangle - |lor\rangle - |ooo\rangle, \\ (\Pi_{1,2} + \Pi_{2,3}) |lor\rangle &= 2 |lor\rangle - |lro\rangle - |olr\rangle, \\ (\Pi_{1,2} + \Pi_{2,3}) |olr\rangle &= 2 |olr\rangle - |lor\rangle - |ooo\rangle. \end{aligned}$$

Note that the only annihilated state containing any Motzkin path of length 3, is the uniform superposition of all Motzkin paths of length 3. That is, the state in which the above equations

are summed together, yielding 0 under the operator  $\Pi$ . Thus the Hamiltonian implementing these operators, given by

$$H = \sum_{j=1}^{n-1} \Pi_{j,j+1},$$

has ground states within the space spanned by all  $|C_{p,q}\rangle$ . Now the superposition of all Motzkin paths  $|\mathcal{M}_n\rangle = |C_{0,0}\rangle$  is a possible ground state, as desired. In order to make this ground state unique, note that a Motzkin path never starts and ends with  $r$  and  $l$  respectively. This can be done by implementing operators acting on the  $r$  and  $l$  states at these locations respectively. This yields the Hamiltonian

$$H = |r\rangle_1 \langle r|_1 + |l\rangle_n \langle l|_n + \sum_{j=1}^{n-1} \Pi_{j,j+1}. \quad (11)$$

Here the notation  $|\cdot\rangle_i \langle \cdot|_i$  denotes the operator acting on the letter at position  $i$  in the string  $s$  of a state  $|s\rangle$ , that is  $I \otimes I \otimes \dots \otimes I \otimes |\cdot\rangle \langle \cdot| \otimes I \otimes \dots \otimes I$  with  $|\cdot\rangle \langle \cdot|$  at location  $i$ . The Hamiltonian in equation 11 only has nearest-neighbour interactions, but does have long-range properties in the ground state.

## 5.4 Entanglement entropy

In calculating the bipartite entanglement entropy of the ground state, one could try to naively calculate the possible coefficients of all possible strings. In order to make this computation more manageable one can consider a Schmidt decomposition for the Motzkin path state. Instead of considering arbitrary subsystems, we will partition the spin chain into subsystems described the index sets  $A = \{1, 2, \dots, N\}$  and  $B = \{N + 1, N + 2, \dots, n\}$ , with  $N < n$ . Denote  $s_A$  and  $s_B$  as the string  $s \in \Sigma_n$  partitioned into the two substrings. One can then obtain the original string again by concatenation of these strings  $s_A$  and  $s_B$ . Whenever  $s$  is a Motzkin path, one has that  $s_A$  and  $c_{0,m}$  are equivalent and that  $s_B$  and  $c_{m,0}$  are equivalent, since the substring  $s_A$  contains a certain number  $0 \leq m \leq N$  of  $l$ 's to be matched with a certain number of  $r$ 's in  $s_B$ . This detail is the principle on which we base the Schmidt decomposition performed here. We will use the equivalence classes  $|C_{0,m}(n)\rangle$ , describing the normalized uniform superposition of strings of length  $n$  equivalent to  $c_{0,m}$ . Considering all possible concatenations for all possible heights  $m$ , we find the decomposition

$$|\mathcal{M}_n\rangle = \sum_{m=0}^N \sqrt{p_m} |C_{0,m}\rangle_A \otimes |C_{m,0}\rangle_B. \quad (12)$$

This decomposition is illustrated in figure 5, giving an example for the Motzkin-path state of  $n = 5$  spins, in which the subsystem  $A$  has length  $N = 2$ .

We now turn to calculating the Schmidt coefficients  $p_m$  for the Schmidt decomposition in equation 12. In order to obtain the uniform superposition of Motzkin paths,  $p_m$  must be the fraction of strings in  $|C_{0,m}\rangle_A \otimes |C_{m,0}\rangle_B$ . Let  $|C_{p,q}(n)|$  be the number of elements in the equivalence class  $C_{p,q}$ , that is, the number of strings equivalent to  $c_{p,q}$  of length  $n$ . By definition of the tensor product  $|C_{p,q}\rangle \otimes |C_{u,v}\rangle$  is the uniform superposition of states consisting of the concatenation of strings equivalent to  $c_{p,q}$  and  $c_{u,v}$ . This yields that the total number of states in  $|C_{p,q}\rangle \otimes |C_{u,v}\rangle$  is given by the product  $|C_{p,q}| |C_{u,v}|$ . Note that due to symmetry  $|C_{m,0}| = |C_{0,m}|$ , from which one obtains

$$p_m = \frac{|C_{0,m}(N)| |C_{0,m}(n-N)|}{|C_{0,0}(n)|}. \quad (13)$$

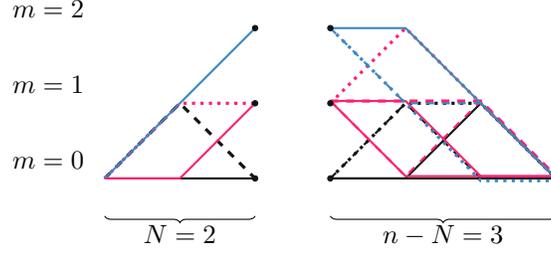


Figure 5: A graphical example of the Motzkin-path Schmidt decomposition for  $n = 5$  and  $N = 2$ . The blue paths represent the  $C_{0,2}$  and  $C_{2,0}$  states, the red paths represent the  $C_{0,1}$  and  $C_{1,0}$  states and the black paths the  $C_{0,0}$  states.

In order to calculate the entanglement entropy for the Motzkin state, the main difficulty lies in determining an expression for  $p_m$  in terms of  $m, n$  and  $N$ . Bravyi et al turn to another set of strings in order to solve this problem. Define the set of strings  $\mathcal{D}_{n,k} \subseteq \{l, r\}^{2n+k}$ , consisting of strings with  $k$  more  $l$ 's than  $r$ 's and any initial segment having more  $l$ 's than  $r$ 's. This results in lemma 5.

**Lemma 5.** The number of strings  $D_{n,k}$  in  $\mathcal{D}_{n,k}$  is

$$D_{n,k} = \frac{k+1}{n+k+1} \binom{2n+k}{n}$$

A proof of lemma 5 can be found in [20]. Being mainly concerned with the scaling of  $p_m$  in terms of  $m$ , one can ignore the constant  $|C_{0,0}(n)|$ , representing the Motzkin numbers. Thus the problem is reduced to finding an expression for the coefficient  $|C_{0,m}(n)|$ . In light of lemma 5, one can think of constructing the Motzkin state as placing  $o$ 's at arbitrary location within strings  $s \in \mathcal{D}_{n,m}$ . E.g. suppose a string  $t$  of length  $n$  is desired, in which  $s \in \mathcal{D}_{i,m}$  is of length  $k$ , then there are  $n-k$  spots to be filled with  $o$ 's. The rest of the string is then constructed with the string  $s$ . Thus there are  $\binom{n}{n-k} = \binom{n}{k}$  ways to construct  $t$ . Since there are  $D_{i,m}$  strings of length  $2i+m$ , one obtains that there are  $\binom{n}{2i+m} D_{i,m}$  strings with  $i+m$  and  $i$   $l$ 's and  $r$ 's respectively. Hence, the total number possible strings becomes the sum over  $i$  such that the total length is at most  $n$ ,

$$|C_{0,m}(n)| = \sum_{\substack{i \geq 0 \\ 2i+m \leq n}} \binom{n}{2i+m} D_{i,m} = \sum_{\substack{i \geq 0 \\ 2i+m \leq n}} \underbrace{\frac{m+1}{i+m+1} \binom{n}{2i+m} \binom{2i+m}{i}}_{M_{n,m,i}}. \quad (14)$$

This expression can be used to calculate the number of possible strings for small  $n$  and  $m$ . The result for  $m = 0$  yields an expression for the  $n$ 'th Motzkin number,  $|C_{0,0}(n)| = M_n$ . It is very difficult to say anything about the scaling of  $|C_{0,m}(n)|$  as  $n$  or  $m$  increases. To solve this problem, call the summand in the above equation  $M_{n,m,i}$ , and let  $\alpha = m/\sqrt{n}$  and  $\beta = (i - \frac{n}{3})/\sqrt{n}$ . Then  $M_{n,m,i}$  can be approximated by

$$M_{n,m,i} \approx \frac{3\sqrt{3}}{2\pi n^{\frac{3}{2}}} 3^{n+1} \alpha \exp(-3\alpha^2 - 6\alpha\beta - 6\beta^2). \quad (15)$$

A derivation of approximation 15 can be found in appendix B. Note that this expression is only valid for  $m \neq 0$ , the case  $m = 0$  is also discussed in appendix B. The result of the sum of  $M_{n,m,i}$  over  $i$  is difficult to quantify, and hence Bravyi et al turn to approximating the sum by an integral. We will use the result of the Gaussian integral, given by

$$\int_{-\infty}^{\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \sigma\sqrt{2\pi}. \quad (16)$$

This follows from the normalization of a Gaussian distribution with variance  $\sigma$  and mean  $\mu$ . Equation 16 is a well known result. Note that the exponential in equation 15 can be rewritten  $\exp\left[-\left(\frac{\beta-\alpha/2}{1/3}\right)^2 - \frac{3}{4}\alpha^2\right]$ . This is a Gaussian in  $\beta$  centered around  $\alpha/2$  with a variance of  $\frac{1}{3\sqrt{2}}$ . Hence the integral over this Gaussian has most of its density around  $\alpha/2$ , thus the limits of the sum do not contribute significantly to the integral. The integral over  $i$  can be therefore be taken over  $\mathbb{R}$ . Making the change of variables  $di = \sqrt{n}d\beta$ , and applying it to equation 16, one obtains

$$\begin{aligned} M_{n,m} &\approx \frac{3\sqrt{3}}{2\pi n^{\frac{3}{2}}} 3^{n+1} \alpha \int_{-\infty}^{\infty} \exp\left(-\left(\frac{\beta-\alpha/2}{1/3}\right)^2 - \frac{3}{4}\alpha^2\right) \sqrt{n}d\beta \\ &\approx \frac{3\sqrt{3}}{2\pi n} 3^{n+1} \alpha \frac{\sqrt{\pi}}{3} \exp\left(-\frac{3}{4}\alpha^2\right) = \frac{3^n \sqrt{3}}{2\sqrt{\pi} n^{\frac{3}{2}}} m \exp\left(\frac{3m^2}{4n}\right). \end{aligned}$$

This confirms the result by Bravyi et al. We obtain the following unnormalized result for  $p_m$ ,

$$p_m \propto |C_{0,m}(N)| |C_{0,m}(n-N)| \propto m^2 \exp\left(\frac{3m^2}{4N} + \frac{3m^2}{4(n-N)}\right).$$

The state is normalized with  $\sum_m p_m = 1$ . To find the normalization constant we estimate the summation over a Gaussian exponential with an additional factor  $m^2$  by an integral over  $m$ . This modified Gaussian distribution has mean 0 and variance  $\sigma^2 = \frac{2N(n-N)}{3n}$  which is generally smaller than  $n$ , hence the limit of the sum can again be neglected for  $\frac{n}{2} \gg \sqrt{n}$ . Additionally one can make the substitution  $\gamma = m/\sigma$ , and  $dm = \sigma d\gamma$ . This yields for  $p_m$  the normalized expression

$$p_m \approx \frac{1}{\sigma} \frac{\gamma^2 e^{-\frac{1}{2}\gamma^2}}{K}, \quad K = \int_0^{\infty} \gamma^2 e^{-\frac{1}{2}\gamma^2} d\gamma = \sqrt{\frac{\pi}{2}}.$$

Hence, for the entanglement entropy, one obtains

$$\begin{aligned} S(A) &= -\sum_m p_m \log_2 p_m \approx \int_0^{\infty} \sigma p_m \log_2 \sigma K - \frac{1}{K} \gamma^2 e^{-\frac{1}{2}\gamma^2} \log_2 \left(\gamma^2 e^{-\frac{1}{2}\gamma^2}\right) d\gamma \\ &\approx \frac{1}{2} \log_2 \left(\frac{2N(n-N)}{3n}\right) + \log_2 K - \underbrace{\frac{1}{K} \int_0^{\infty} \gamma^2 e^{-\frac{1}{2}\gamma^2} \log_2 \left(\gamma^2 e^{-\frac{1}{2}\gamma^2}\right) d\gamma}_{\approx -1.1114}. \end{aligned} \quad (17)$$

The integral term in expression 17 can be numerically evaluated, obtaining  $-1.1114$ . By substituting  $N = \frac{n}{2}$  we can equate our result with the result of Bravyi et al,

$$S(A) \approx \frac{1}{2} \log_2 \frac{n}{6} + \frac{1}{2} \log_2 \frac{\pi}{2} + 1.1114 \approx \frac{1}{2} \log_2 n + 0.145.$$

This confirms the result found by Bravyi et al.

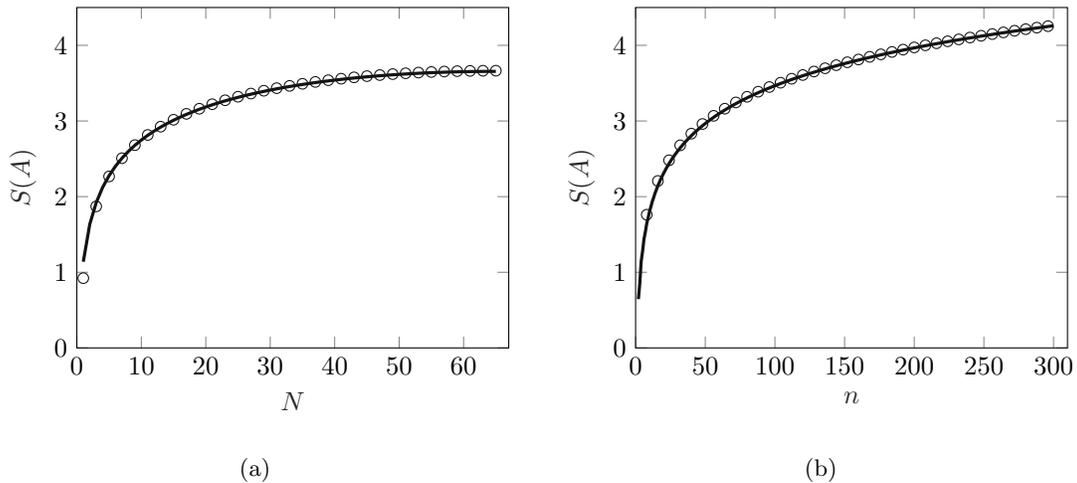


Figure 6: The entanglement entropy  $S(A)$  for the Motzkin path state. Figure a is the entanglement for a block  $A$  consisting of the first  $N$  spins in a chain of  $n = 130$  spins. Figure b is the entanglement entropy of the half chain, with  $N = \frac{n}{2}$  whenever  $n$  is a multiple of 8. The circles represent the exact entropy and the line is the approximation found in equation 17.

In figure 6a the entanglement entropy for the first  $N$  spins in a chain of  $n = 130$  spins is plotted. These values are computed with exact results from expressions 12 and 14 or the estimate from expression 17. The entanglement entropy is only shown up to  $n/2 = 65$ . For a pure system  $AB$  one has  $S(A) = S(B)$ , thus the entanglement entropy for the first  $N$  spins is equal to that of the rest of the chain, consisting of  $n - N$  spins. For small  $n$  our estimate is expected to be worse due to the assumptions made in the derivation. This is slightly visible in figure 6b.

In figure 6b the exact value of the entanglement entropy of the half chain  $N = n/2$  is again compared to the approximation in equation 17. Note that the error between the approximation and the exact value decreases as  $n$  increases, consistent with the assumptions made in the derivation.

## 6 Energy gap of the Motzkin-path Hamiltonian

As discussed in section 4.4, the entanglement entropy is related to the energy gap between the ground state and first excited state. Studying the behaviour of the energy gap is therefore essential in understanding the entanglement entropy. The ground state of the Motzkin-path Hamiltonian has 0 energy, hence the energy gap is given by the first eigenvalue  $\lambda_1 > 0$  of the Hamiltonian defined in equation 11. One can consider a minimization problem of finding a vector in the vector space orthogonal to the ground state such that the expectation of the energy is minimized. This idea is entailed in the Courant-Fischer formula, which states for the specific case of the second smallest eigenvalue of a Hermitian operator  $A$ , that

$$\lambda_1 = \min_{\substack{\langle \psi | \psi_0 \rangle = 0 \\ |\psi\rangle \neq \mathbf{0}}} \frac{\langle \psi | A | \psi \rangle}{|\langle \psi | \psi \rangle|^2}, \quad (18)$$

with  $|\psi_0\rangle$  the state vector associated to the smallest eigenvalue of the operator  $A$ . For us, the ground-state vector  $|\psi_0\rangle$  is the Motzkin state  $|\mathcal{M}_n\rangle$ . In solving this minimization problem we turn to an analogy of the Hamiltonian, considering it as a graph.

### 6.1 Hamiltonian as a graph

A graph is a collection of vertices with edges between them, as illustrated in figure 7. These vertices are described by the set  $V$ , and the edges between these vertices, by the set  $E$ . One often writes  $G = (V, E)$  for the graph. The number of edges connected to a vertex  $v \in V$  is called the degree of the vertex, denoted  $\deg(v)$ . The Motzkin-path Hamiltonian without boundary conditions can be considered as a Laplacian matrix  $L$  of a graph, which is defined for a graph  $G = (V, E)$  with vertices  $v_i \in V$  and edges  $(i, j) \in E$ , in terms of the degree matrix  $D$  and adjacency matrix  $A$

$$D_{i,j} = \begin{cases} \deg(v_i) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \quad A_{i,j} = \begin{cases} 1 & \text{if } (i, j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The Laplacian matrix is then  $L = D - A$ . For labeled nodes  $v_i \in V$ , we find that when there exists an edge  $(i, j) \in E$ , there is an entry of  $-1$  at position  $i, j$ . Furthermore, the degree  $\deg(v_i)$  of the vertex  $v_i$  is positioned on the diagonal. An example of a graph and its Laplacian matrix is illustrated in figure 7.

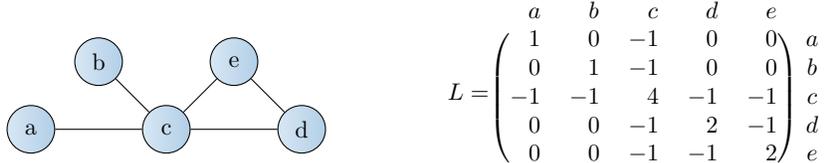


Figure 7: A graph and its Laplacian matrix  $L$ . The relation between the nodes in the graph and the entry positions in the Laplacian are indicated.

Note that the matrix rows and columns always sum to 0 for graphs without loops, since the degree  $\deg(v_i)$  of the vertex  $v_i$  and number of incidence edges are equivalent by definition. The Hamiltonian

from equation 11 can be thought of as such a Laplacian matrix if one removes the boundary terms. These boundary terms affect the diagonal of the Hamiltonian, but do not contribute to the incidence matrix. As a result the Laplacian of the entire Hamiltonian cannot be considered as a graph. Recall that the operators  $\Pi_{j,j+1}$  acting on a state  $|s\rangle$  result in  $\sum_i \Pi_{j,j+1} |s\rangle = m |s\rangle - \sum_{s'} |s'\rangle$ . Here  $|s'\rangle$  represents the string states a single operation away from  $|s\rangle$ , and  $m$  the number of different  $|s'\rangle$ . We find that the  $m$  get placed on the diagonal for each string  $s$  and the off-diagonal terms all have a  $-1$  entry. For the Motzkin-path Hamiltonian without boundary terms, the nodes represent the strings. They are connected by an edge when they are a single operation away. The example of Motzkin paths of length 3 in section 5.2 illustrates these vertices and edges. The  $|ooo\rangle$  state is a node with degree 2, connected to the states  $|lro\rangle$  and  $|olr\rangle$ . The Laplacian matrix of the operators  $\Pi$  acting on a chain of 3 spins yields the graph in figure 8.

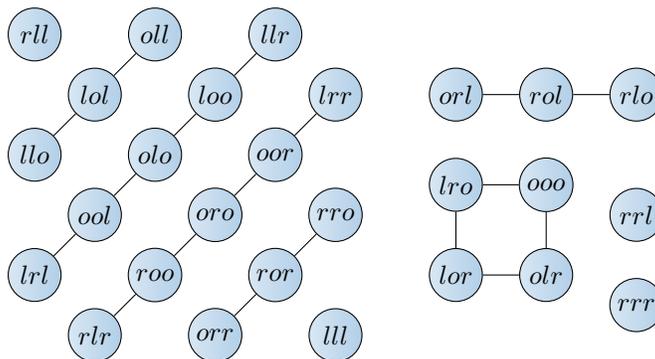


Figure 8: The graph associated to the Motzkin-path Hamiltonian on a chain of 3 spins. The nodes are labeled by the path they represent, and edges occur when two paths are a single operation away.

The different connected components in figure 8 give a good illustration of the different equivalence classes  $C_{p,q}$ . The connected component of 4 nodes represent the Motzkin paths of length 3. Now in order to quantify the behaviour of the Hamiltonian in light of the graph representation, consider the state  $|\psi\rangle = \sum_{s \in \Sigma_n} c_s |s\rangle$ , using again  $\Sigma_n = \{l, o, r\}^n$ . Then

$$\begin{aligned} \langle \psi | L | \psi \rangle &= \langle \psi | (D - A) | \psi \rangle = \sum_{s \in \Sigma_n} d_s \bar{c}_s c_s - \sum_{(s,t) \in E} (\bar{c}_s c_t + c_s \bar{c}_t) \\ &= \sum_{(s,t) \in E} (\bar{c}_s c_s + \bar{c}_t c_t - \bar{c}_s c_t - c_s \bar{c}_t) = \sum_{(s,t) \in E} |c_s - c_t|^2. \end{aligned}$$

The definition of the degree  $d_s$ , allowed us to replace the summation over the vertices  $s$  by a summation over the edges  $(s,t)$ . It is important to note that this equation for edges holds for undirected graphs, for which each edge should only be counted once. We find that equation 18 can be rewritten in terms of the edges in the graph

$$\lambda_1 = \min_{\substack{\langle \psi | \psi_0 \rangle = 0 \\ |\psi\rangle \neq \mathbf{0}}} \frac{\sum_{(s,t) \in E} |c_s - c_t|^2}{\sum_{s \in \Sigma_n} |c_s|^2}. \quad (19)$$

We find that the energy gap is mostly determined by the difference in coefficients of vertices connected by an edge. The goal is to find coefficients for the Motzkin state that are a small distance

apart.

In associating the Motzkin-path Hamiltonian with a Laplacian matrix we ignore the boundary terms. A string that is not a Motzkin path will increase the energy significantly, we will therefore restrict the optimization of equation 19 to the subgraph of Motzkin paths. In other words, we say  $c_s = 0$  if  $s$  is not a Motzkin path.

## 6.2 Distance phase

In light of equation 19, we introduce a phase factor on the Motzkin strings states  $|s\rangle$  in terms of its distance from the state  $oo\dots o$ . The Motzkin-path Hamiltonian is a real symmetric matrix. Therefore the complex conjugate of an eigenvector is also an eigenvector. A phase factor  $\exp i\theta$  can therefore be implemented by a cosine.

We define a function  $d(s)$  as the minimum number of moves required to obtain the string  $s$ , if one starts with the  $oo\dots o$  string. This can be defined more rigorously in terms of a string  $s \in \{l, o, r\}^n$

$$d(s) = \sum_{i=1}^n [L_i(s) - R_i(s)], \quad (20)$$

where  $L_i(s)$  and  $R_i(s)$  denote the number of  $l$ 's and  $r$ 's in the segment consisting of the first  $i$  initial letters in the string  $s$  respectively. This can be motivated by identifying  $d(s)$  with the area under the Motzkin path. For example,  $lor$  has area 2, equivalent to the distance,  $d(s) = 2$ . The definition for  $d(s)$  is effectively a summation over the height at each position in the string, which is equivalent to the definition of area. The string  $s$  of length  $n$  with largest distance  $d(s)$  is then the string  $ll\dots lrr\dots r$ , if  $n$  is even. The maximum distance  $d_{\max}$  is then  $d_{\max} = n \times n/2 \times 1/2 = n^2/4$ : the area under a triangle with side lengths  $n$  and  $n/2$ . If  $n$  is odd, this is  $d_{\max} = (n-1) \times (n-1)/2 \times 1/2 + (n-1)/2 = (n^2-1)/4$ . Implementing the phase factor for the Motzkin-path state  $|\mathcal{M}_n\rangle$ , one finds as coefficients for the original string states  $|s\rangle$ ,

$$|\hat{s}\rangle = \frac{1}{K} \cos\left(\pi \frac{d(s)}{d_{\max}}\right) |s\rangle, \quad K^2 = \sum_{s \in \mathcal{M}_n} \cos^2\left(\pi \frac{d(s)}{d_{\max}}\right).$$

Define the superposition of the phase-shifted strings  $|\psi\rangle = \sum_{s \in \mathcal{M}_n} |\hat{s}\rangle$ .  $K$  is then a normalization constant such that this state is normalized. Note however, that this state is not necessarily orthogonal to the ground state  $|\mathcal{M}_n\rangle$ , which was a condition on the Courant-Fischer formula in equation 18 and 19. Performing the Gram-Schmidt process, one constructs a orthogonal state from the to the ground state  $|\mathcal{M}_n\rangle$ . In performing this orthogonalization, we will ignore the normalization of  $|\psi\rangle$ , and normalize the state afterwards. The basis elements  $|s'\rangle$  of the orthogonal state in terms of the unmodified basis elements  $|s\rangle$  are

$$|s'\rangle = \cos\left(\pi \frac{d(s)}{d_{\max}}\right) |s\rangle - \frac{\langle \psi | \mathcal{M}_n \rangle}{\langle \mathcal{M}_n | \mathcal{M}_n \rangle} |s\rangle = \left[ \cos\left(\pi \frac{d(s)}{d_{\max}}\right) - \underbrace{\frac{1}{M_n} \sum_{t \in \mathcal{M}_n} \cos\left(\pi \frac{d(t)}{d_{\max}}\right)}_{\xi} \right] |s\rangle,$$

where the height shift  $\xi$  is a constant and  $M_n$  is the  $n$ 'th Motzkin number. Note that  $|\xi| \leq 1$  due to  $|\cos x| \leq 1$  and the summation over all Motzkin paths. Normalization of these strings leads us to

define a orthogonal state  $|\mathcal{M}_n^{\text{ex}}\rangle = \sum_{s \in \mathcal{M}_n} |s'\rangle$ , with the phase-shifted coefficients,

$$|\mathcal{M}_n^{\text{ex}}\rangle = \frac{1}{K} \sum_{s \in \mathcal{M}_n} \left[ \cos\left(\pi \frac{d(s)}{d_{\max}}\right) - \xi \right] |s\rangle, \quad K^2 = \sum_{t \in \mathcal{M}_n} \left( \cos\left(\pi \frac{d(t)}{d_{\max}}\right) - \xi \right)^2 \quad (21)$$

The excited state  $|\mathcal{M}_n^{\text{ex}}\rangle$  will by construction have a small energy in the Courant Fischer formula from equation 19, and will likely give a good approximation of the energy gap for large  $n$ , since the phase difference will become very small. This is due to the distance  $d(s)/d_{\max}$  introducing a phase shift of  $1/d_{\max}$  on the states connected by an edge. Applying equation 19 to the  $|\mathcal{M}'_n\rangle$  state yields

$$\begin{aligned} \lambda_1 &\leq \frac{1}{K^2} \sum_{(s,t) \in E} \left| \cos\left(\pi \frac{d(s)}{d_{\max}}\right) - \cos\left(\pi \frac{d(t)}{d_{\max}}\right) \right|^2 \\ &= \frac{1}{K^2} \sum_{(s,t) \in E} \left| \cos\left(\pi \frac{d(s)}{d_{\max}}\right) - \cos\left(\pi \frac{d(s) \pm 1}{d_{\max}}\right) \right|^2 \\ &\approx \frac{4}{K^2} \sum_{(s,t) \in E} \sin^2\left(\frac{\pi}{d_{\max}}\right) \sin^2\left(\pi \frac{2d(s)}{d_{\max}}\right) \end{aligned} \quad (22)$$

The relation  $\cos a - \cos b = 2 \sin(a+b) \sin(a-b)$  was used to bring the two cosine terms together. There are two terms in expression 22 requiring further inspection. We will firstly consider the summation over the number of edges  $(s,t) \in E$  and turn this into a summation over Motzkin paths  $s \in \mathcal{M}_n$ . Afterwards we will inspect the fraction  $d(s)/d_{\max}$  in further detail.

### 6.3 Estimate for the number of edges

The number of edges originating from a string  $s$  is equal to the number of operations that can be made on the string. In order to simplify the process, we note only the area-increasing moves. Therefore, each edge is only counted once, as desired. We will construct a function  $p : \{l, o, r\}^n \rightarrow \mathbb{N}$ , denoting the number of area-increasing moves possible on a string  $s$ . Writing down a closed form expression for this function is difficult, and requires further inspection of the moves that can be performed on a string. The exchange operations presented in section 5.2 can be defined as a unidirectional in order to become area-increasing. These unidirectional moves are  $oo \rightarrow lr$ ,  $ol \rightarrow lo$  and  $ro \rightarrow or$ , the moves presented in figure 4. Additionally we define an indicator function

$$\mathbf{1}_\chi(x) = \begin{cases} 1 & \text{if } x = \chi, \\ 0 & \text{otherwise.} \end{cases}$$

We will use the notation  $O(s)$  for the number of  $o$ 's in a string  $s$  and  $s_i$  for the  $i$ 'th letter in a string  $s$ .

**Proposition 1.** The number of area-increasing moves  $p(s)$  that can be performed on a string  $s \in \{l, o, r\}^n$  of length  $n$  can be expressed as

$$p(s) = O(s) - 1 + \sum_{i=1}^{n-1} [\mathbf{1}_l(s_i) \mathbf{1}_r(s_{i+1}) - \mathbf{1}_r(s_i) \mathbf{1}_l(s_{i+1})]. \quad (23)$$

The summand represents the number of peaks  $lr$  and valleys  $rl$  in the Motzkin path.

*Proof.* Let  $s \in \{l, o, r\}^n$  be an arbitrary string of length  $n$ .

We will first show that  $p(s) = 0$  if  $O(s) = 0$ . In this case  $p(s)$  is the difference in the number of pairs  $lr$  and  $rl$ , with an offset of 1. Suppose there exists a pair  $rl$ , then there must exist an additional  $lr$  later in the string, in order to satisfy the Motzkin path definition. Illustratively, the pair  $rl$  makes the path go up, thus in order to arrive back at the base height, the path needs to go back down again, which is only possible through a pair  $lr$  if there are no  $o$ 's. This holds for all pairs  $rl$  in the string  $s$ . Regardless of the rest of the string, the path needs to go from up to down at least once. Therefore there is an additional pair  $lr$ . Resulting in

$$p(s) = -1 + \sum_{i=1}^{n-1} \mathbf{1}_l(s_i)\mathbf{1}_r(s_{i+1}) - \mathbf{1}_r(s_i)\mathbf{1}_l(s_{i+1}) = 0,$$

provided  $s$  does not contain any  $o$ . Since a Motzkin path can be considered by placing  $o$ 's at arbitrary locations in this string, we will consider the different placements and count the contributions to the number of possible moves. We will cover the  $3 \times 3 = 9$  possibilities, which are illustrated in figure 9.



Figure 9: The 9 possibilities an  $o$  can be placed in a Motzkin-path substring. The numbers refer to the listing in the proof.

Suppose the  $o$  is inserted between

1.  $ll$ , then the new substring is  $lol$ . This will add 1 to the number of possible moves, since it is now possible to perform  $ol \rightarrow lo$ .
2.  $lo$ , then the new substring is  $loo$ . This will also add 1 to the number of possible moves, since the  $oo \rightarrow lr$  can now be performed.
3.  $ol$ . Now the new substring is  $ool$ , in which two operations can be performed, namely  $ol \rightarrow lo$  and  $oo \rightarrow lr$ . One was already possible before, so the  $o$  adds 1 to the number of possible moves.
4.  $oo$ , yields the new substring  $ooo$ . The operation  $oo \rightarrow lr$  can now be performed at two locations, instead of the 1 location at that was initially possible. Again we find the  $o$  adds 1 to the number of possible strings.
5.  $or$ . The new substring is  $oor$ , similarly to one of the previous cases, we find there is 1 new possible operation, namely  $oo \rightarrow lr$ . Thus the placement of  $o$  adds 1 to the number of possible moves.
6.  $ro$ . Now the new substring becomes  $roo$ , in which two possible moves can be made. Similarly to a previous situation, one was already possible. We find that this placement adds 1 to the number of possible moves.

7.  $rr$ . The new substring  $ror$  has a contribution of 1 to the possible number of moves, since the operation  $ro \rightarrow or$  can now be performed.
8.  $lr$ , results in  $lor$ , where no additional moves are possible. Note however, that the number of peaks in the string is now reduced by 1. By counting the  $o$ , we compensate for this.
9.  $rl$ , yields  $rol$ , in which two possible operations are now able to be performed. More specifically  $ro \rightarrow or$  and  $ol \rightarrow lo$ . Note that the number of valleys is reduced by 1 and the additional  $o$  adds 1, so the number of possible moves is still increased by 2.

Performing this insertion of  $o$ 's iteratively on a string that originally had no  $o$ 's, any Motzkin path can be built. By keeping track of the additional number of moves, and the reduction of peaks in situation 8. and 9., we find the result in equation 23.  $\square$

The definition of  $p(s)$  in equation 23 allows us to exchange the summation over the edges. We replace this summation with  $p(s)$  for all strings  $s$ , turning the problem into a summation over Motzkin paths,

$$\lambda_1 \leq \frac{4p(s)}{K^2} \sum_{s \in \mathcal{M}_n} \sin^2 \left( \frac{\pi}{d_{\max}} \right) \sin^2 \left( \pi \frac{2d(s)}{d_{\max}} \right).$$

We are mainly interested in the scaling of the energy gap, therefore we can consider the average contribution of  $p(s)$  instead of considering  $p(s)$  for the individual strings. Consider a random position in the bulk in an arbitrary Motzkin path  $s$ . Then, assuming the boundary conditions do not affect the string significantly, the probability for finding a  $o$  at this location is approximately  $\frac{1}{3}$ . Hence, the average value of  $O(s)$  is approximately  $\frac{n}{3}$ . More formally,  $O(s)$  becomes a binomial random variable with success-probability parameter  $p = \frac{1}{3}$  and  $n$  trials. This binomial random variable has expectation value  $np = \frac{n}{3}$  [21].

The scaling of  $p(s)$  has an additional summation over the number of  $lr$ 's and  $rl$ 's. Both substrings should appear approximately equally often, and hence do not contribute significantly to  $p(s)$ . Hence  $p(s) \sim \text{Bin}(\frac{1}{3}, n)$ , with expectation value  $\frac{n}{3}$ . The normalized distribution of  $p(s)$  for strings of length 16 and the probability mass function of a binomial distribution in are plotted figure 10. The distribution of  $p(s)$  appears to be slightly more spread out than that of the binomial distribution. These are exactly the contributions of the pairs of  $lr$  and  $rl$ , increasing and decreasing the number of possible edges slightly, as illustrated in expression 23 for  $p(s)$ . The number of edges is correlated to the distance  $d(s)$ , since e.g.  $p(llrr) = 0$ . This is only the case at the boundaries, and is expected to not affect the bulk for large  $n$ .

## 6.4 Distribution of the phase factor

The main difficulty in deriving an estimate for the limit scaling of the energy gap lies in the behaviour of the function  $d(s)$  in terms of the number of Motzkin paths  $M_n$ . It is possible to compute the number of occurrences for each  $d(s)$  using a recurrence relation [22]. This recurrence relation generates a polynomial  $d_n(q)$ , in which the exponents of  $q$  represent the area and their coefficients the number of occurrences of that area. For example  $d_4(q) = q^4 + q^3 + 3q^2 + 3q + 1$ , representing 1 Motzkin path for areas 4 and 3, 3 paths with area 2 and 1, and only one path with

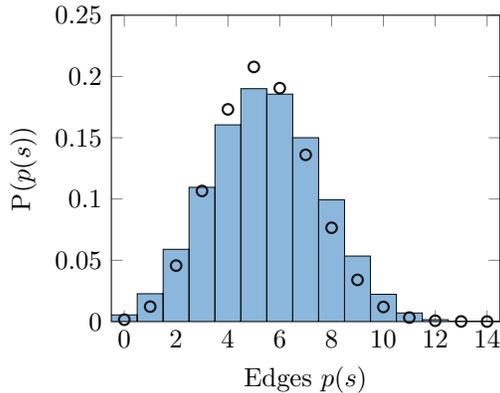


Figure 10: The normalized distribution of the number of edges  $p(s)$  for Motzkin paths of length 16 is presented by the histogram. The black circles are the probability mass function of a binomial distribution with success probability  $p = \frac{1}{3}$  and  $n = 16$  trials.

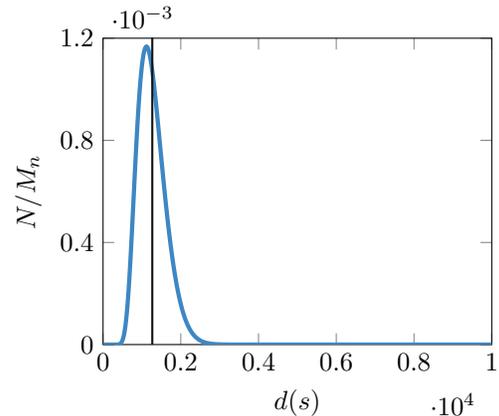


Figure 11: The normalized distribution of  $d(s)$  in terms for a chain of  $n = 200$  spins. The vertical black line is the mean of the distribution.

area 0, the flat path. The recurrence relation reads

$$d_n(q) = d_{n-1}(q) + \sum_{k=0}^{n-2} q^{k+1} d_k(q) d_{n-k-2}(q), \quad \text{with } d_0(q) = 1. \quad (24)$$

The normalized distribution of paths for a chain of  $n = 200$  spins is presented in figure 11. The goal will be to derive the distribution of  $d(s)$ , in terms of the chain length  $n$ . We turn to the first moment of the area distribution, which is defined  $\sum_k k \tau(k)$ . Here we write  $\tau(k)$  for the number of strings with distance and area  $k$ . Note that this is equivalent to the summation over the area of all Motzkin paths. The  $\tau(k)$  are equivalent to the coefficients of  $q^k$  in recurrence relation 24 for a chain of given length  $n$ . One defines a strict Motzkin path as a Motzkin path starting and ending with  $l$  and  $r$  respectively, and never returning to the base, equivalent to raising the Motzkin path up by a single trapezoid. For these paths there exist generating functions for arbitrary moments, which yields an expression for the summation over the area of all strict Motzkin paths of length  $n$  [23],

$$A'_n = \frac{1}{4} (3^{n-1} - (-1)^{n-1}).$$

We, however, are interested in the summation over all Motzkin paths of length  $n$ . Note that the area of the raised strip is  $n - 1$  for these strict Motzkin paths of length  $n$ . Hence, for normal Motzkin paths of length  $n$ , the expression  $A'_{n+2}$  overcounts  $M_n$  trapezoids of area  $n + 1$ . Correcting this results in the following result for the summation over the area of all normal Motzkin paths of length  $n$ ,

$$A_n = \frac{1}{4} (3^{n+1} - (-1)^{n+1}) - (n + 1) M_n. \quad (25)$$

This results in an expectation value for the distance factor  $d(s)$  of  $A_n/M_n$ . The scaling difference between the first term in equation 25 and  $M_n$  will need to be investigated further. Taking  $m = 0$  in expression 14 results in an exact expression for the Motzkin numbers. An estimate for this number was derived in appendix B. Analogously as for the arbitrary states in equation 15 we can integrate over  $\beta$ , by means of a change of variables  $di = \sqrt{n}d\beta$ , in order to obtain an estimate for  $M_n$ , yielding

$$M_n = \sum_{\substack{i \geq 0 \\ 2i \leq n}} M_{n,0,i} \approx \frac{3\sqrt{3}}{2\pi n^2} 3^{n+1} \int_{-\infty}^{\infty} \exp(-9\beta^2) \sqrt{n}d\beta = \frac{\sqrt{3}}{2\sqrt{\pi}n^{\frac{3}{2}}} 3^{n+1}.$$

Combining this expression with equation 25 yields the expectation value

$$\langle d(s) \rangle = \frac{A_n}{M_n} \approx \frac{\sqrt{\pi}n^{\frac{3}{2}}}{2\sqrt{3}} \left[ 1 - \left( -\frac{1}{3} \right)^{n+1} \right] - n - 1 \approx \frac{\sqrt{\pi}}{2\sqrt{3}} n^{\frac{3}{2}} - n.$$

The terms appearing in our derivation are normalized with a factor  $d_{\max} \approx n^2/4$ , yielding for the normalized expectation value

$$\left\langle \frac{d(s)}{d_{\max}} \right\rangle = 2\sqrt{\frac{\pi}{3n}} - \frac{4}{n}. \quad (26)$$

A random walk with integer steps converge to a Wiener process in the limit  $n \rightarrow \infty$  [24]. We can then define a Brownian excursion  $B(t)$ , which is a Wiener process on the unit interval  $B(t) : [0, 1] \rightarrow \mathbb{R}$  such that the path starts and ends at  $B(0) = 0 = B(1)$ , without going negative,  $B(t) \geq 0$ . Hence a random Motzkin path can be associated with that of a Brownian excursion in the limit  $n \rightarrow \infty$ . An example of a Brownian excursion is given in figure 12.

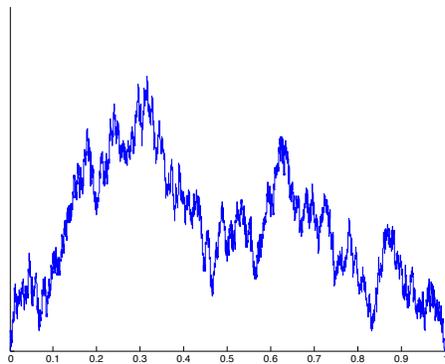


Figure 12: A Brownian excursion on the interval  $[0, 1]$ . Figure by I. Kortchemski [25].

The area under these Brownian Excursions, given by  $\int B(t)dt$ , then has a probability density function given by

$$f_A(x) = \frac{2\sqrt{6}}{x^2} \sum_{j=1}^{\infty} v_j^{\frac{2}{3}} e^{-v_j} \mathcal{U} \left( -\frac{5}{4}, \frac{3}{4}; v_j \right), \quad x \in [0, \infty) \quad (27)$$

with  $v_j = 2|a_j|^3/27x^2$ . In this case  $a_j$  are the zeroes of the Airy function and  $\mathcal{U}$  is the confluent hypergeometric function [26]. This distribution has expectation value  $\langle B \rangle = \sqrt{\pi/2}/2$  and variance

$\sigma_B^2 = \sqrt{5/12 - \pi/8}$  [27]. We now perform a change of variables  $d(s)/d_{\max} \sim C(n)B$ , projecting the Motkzin-path problem to a Brownian excursion. We do this by defining  $C(n)$  such that  $C(n)\langle B \rangle = \langle d(s)/d_{\max} \rangle$ . This yields of the distribution  $d(s)/d_{\max}$  the value

$$C(n) = \frac{\left\langle \frac{d(s)}{d_{\max}} \right\rangle}{\langle B \rangle} = 4\sqrt{\frac{2}{3n}} - \frac{8}{n}\sqrt{\frac{2}{\pi}}, \quad \sigma_d^2 = \left( 4\sqrt{\frac{2}{3n}} - \frac{8}{n}\sqrt{\frac{2}{\pi}} \right)^2 \sigma_B^2 = \mathcal{O}\left(\frac{1}{n}\right).$$

We can conclude that in the limit  $n \rightarrow \infty$  the variance becomes small. The estimation of terms consisting of  $\sin(d(s)/d_{\max})$  would involve integrals of terms similar to  $\int f_A(x) \sin(C(n)x) dx$ . The integration over the probability distribution of the Brownian excursion in equation 27 is extremely intricate, and still being researched actively. The extremely small variance in the limit  $n \rightarrow \infty$  leads to the estimate  $\langle \sin(d(s)/d_{\max}) \rangle \approx \sin\langle d(s)/d_{\max} \rangle$ , and analogously for the cosine. Combining this with the result from the previous section, we assume that the distributions for  $d(s)/d_{\max}$  and  $p(s)$  are independent. Therefore we will replace the product  $p(s) \sin^2(d(s)/d_{\max})$  with the product of expectation values,

$$\left\langle p(s) \sin^2\left(\frac{d(s)}{d_{\max}}\right) \right\rangle \approx \langle p(s) \rangle \sin^2\left\langle \frac{d(s)}{d_{\max}} \right\rangle.$$

These estimates will be used in the next section to estimate the energy gap scaling in the limit  $n \rightarrow \infty$

## 6.5 Energy gap scaling

In order to estimate the energy gap we can consider two separate contributions. These are the normalization constant  $K^2$  and the summation over the  $\sin^2$  terms. We will start with the normalization constant. The estimate in equation 26 will be used extensively. The scaling of the normalization constant in terms of the chain length  $n$  is derived in appendix C depending a Taylor expansion. This derivation yields the following expression for the normalization constant  $K^2$ .

$$K^2 \approx \frac{M_n \pi^4}{4} \left\langle \frac{d(s)}{d_{\max}} \right\rangle^4 \tag{28}$$

The relative error  $\epsilon$  of this estimate for the normalization constant, is presented in figure 13 together with the relative error in the estimate for the mean  $\langle d(s)/d_{\max} \rangle$ . The estimate improves as  $n \rightarrow \infty$  since the approximations made in the derivation have the limit  $\langle d(s)/d_{\max} \rangle \rightarrow 0$ . The other terms

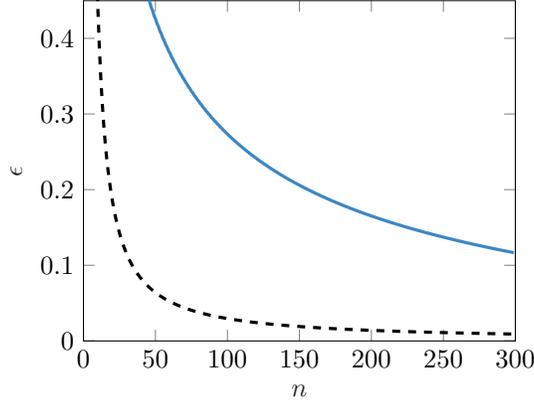


Figure 13: The relative error  $\epsilon$  between the normalization constant  $K^2$  and the estimate in expression 28, in terms of the chain length  $n$ , is the blue curve. The relative error between the estimate in expression 26 for the mean  $\langle d(s)/d_{\max} \rangle$  and the actual mean is presented in terms of the chain length  $n$  by the dashed curve.

in the bound on the energy gap will be estimated by the small angle approximation. We obtain

$$\begin{aligned}
\lambda_1 &\leq \frac{4n}{3K^2} \sin^2\left(\frac{\pi}{d_{\max}}\right) \sum_{s \in \mathcal{M}_n} \sin^2\left(\pi \frac{2d(s)}{d_{\max}}\right) \\
&\approx \frac{4n}{3} \frac{4}{M_n \pi^4} \left\langle \frac{d(s)}{d_{\max}} \right\rangle^{-4} \left(\frac{\pi}{d_{\max}}\right)^2 \sum_{s \in \mathcal{M}_n} \left(\pi \frac{2d(s)}{d_{\max}}\right)^2 \\
&\approx \frac{16n}{3M_n \pi^2} \frac{64\pi^2}{n^4} \left\langle \frac{d(s)}{d_{\max}} \right\rangle^{-4} \sum_{s \in \mathcal{M}_n} \left\langle \frac{d(s)}{d_{\max}} \right\rangle^2 \\
&= \frac{1024}{3n^3} \left\langle \frac{d(s)}{d_{\max}} \right\rangle^{-2} = \frac{1024}{3n^3} \left( \frac{4\pi}{3n} - \frac{8\sqrt{\pi}}{n\sqrt{3n}} + \frac{16}{n^2} \right)^{-1} \\
&= \frac{256}{3} \left( \frac{\pi n^2}{3} - 2\sqrt{\frac{n\pi}{3}} + 4n \right)^{-1} \tag{29}
\end{aligned}$$

We find that the leading term results in a  $n^{-2}$  scaling of the energy bound, similar to the results found by Movassagh et al. [16]. This estimate for the bound and the exact results is presented in figure 14. The exact smallest nonzero eigenvalues of the Motzkin-path Hamiltonian are presented by the black circles and compared to the phase-shifted values, represented by the blue crosses. The exact eigenvalues are only presented up to  $n = 13$  due to computational limits, which was already extended by the use of sparse matrices. Similarly, the exact phase shifted bounds are presented up to  $n = 16$ .

The presented bound appears to have different scaling than the exact eigenvalue  $\lambda_1$ . It is difficult to predict the behaviour for the exact smallest nonzero eigenvalue for larger  $n$ . The bound presented in equation 29 is presented by the black line, with its leading order term represented by the red

line. In addition to these bounds, the estimate for the distribution of strings with area  $d(s)$  by using the recurrence relation is presented by the green dots. The estimate for the number of edges  $p(s) \approx n/3$  appears to not have a large effect on the the estimated energy bound, due to the exact values and the estimate from the recurrence relation overlapping. The overestimate of the bound for small  $n$  is due to the underestimate in the normalization constant  $K^2$ , as indicated in figure 13.

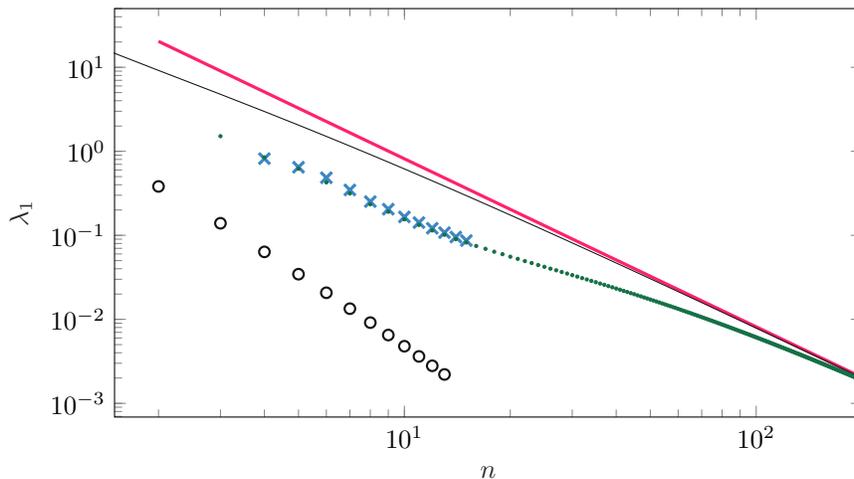


Figure 14: The eigenvalue gap  $\lambda_1$  of the Motzkin-path Hamiltonian as a function of the chain length  $n$  on logarithmic axes. The exact result obtained from the Hamiltonian is represented by the black circles. The approximation by a phase-distance is presented by the blue crosses. The black line is the approximation in the text and the red line represents the limit scaling. The scaling computed using the recurrence relation in the text is indicated by the green dots.

The closing of the energy gap in the limit  $n \rightarrow \infty$  allows the entanglement entropy to scale with the volume. This was illustrated in equation 10, which is therefore not violated for the Motzkin-path Hamiltonian. The logarithmic scaling is allowed by the bound presented. We obtain a similar result as Bravyi et al., with the energy gap being bound from above by an inverse polynomial,  $\lambda_1 \leq n^{-c}$  with  $c$  an arbitrary constant.

## 7 Path models for arbitrary Spin

The construction of the Motzkin path model for spin-1 systems leads to the question whether this model can be constructed for different order of spins. We will start by investigating a spin- $\frac{1}{2}$  model. Then one can construct generalizations of both the spin- $\frac{1}{2}$  and the spin-1 models, which leads to colourings of these models. The extensions presented in this section directly follow the work of R. Movassagh, P. Shor, O. Salberger and V. Korepin [16, 28].

### 7.1 Dyck paths

Analogous to the case of the spin-1 chain, we can consider strings consisting of  $l$ 's and  $r$ 's, identifying them with up and down movements. For a pair of neighbouring spins there is one operation that is closely related to that of the Motzkin path operations. The exchange operation  $lr \leftrightarrow rl$ , which is depicted graphically in figure 15a.

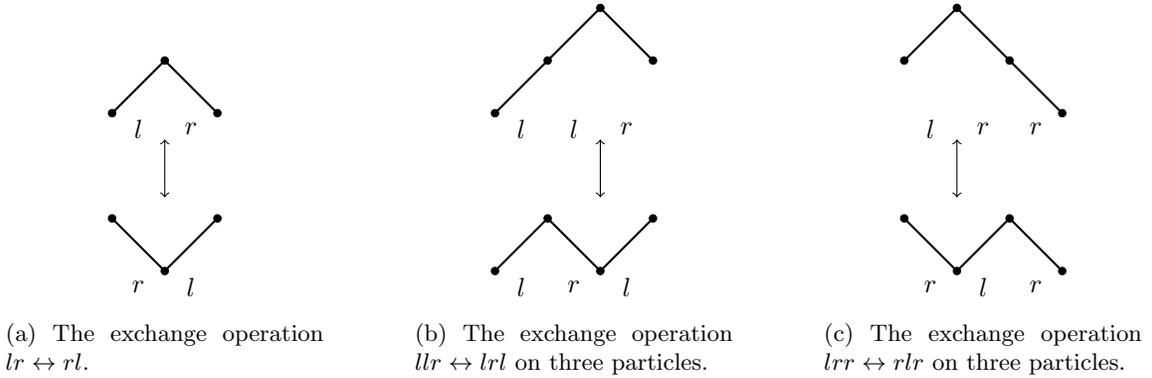


Figure 15: Candidates for the exchange operations on the spin- $\frac{1}{2}$  chain.

There is however an important difference to the Motzkin-path model when considering the exchange operation  $lr \leftrightarrow rl$ . Starting with the string  $llrr$  it is possible to construct any possible string with two  $l$ 's and  $r$ 's, including the  $rrll$  string, which is not a Motzkin path. There is a close relation to the Heisenberg model for this chain [29]. In order to see this, define analogous to the operators in the Motzkin-path model, a state implementing the exchange operation  $|\psi\rangle = |lr\rangle - |rl\rangle$ . Then the associated operator becomes  $|\psi\rangle\langle\psi| = |lr\rangle\langle lr| - |lr\rangle\langle rl| - |rl\rangle\langle lr| + |rl\rangle\langle rl|$ . Identifying the state  $l$  with a spin-up state, and  $r$  with a spin-down state, we can write this in terms of raising and lowering operators. That is,

$$|lr\rangle\langle rl| = S_+^1 S_-^2 \qquad |rl\rangle\langle lr| = S_-^1 S_+^2.$$

Indeed, the state  $rl$  in the exchange operation is projected to  $|lr\rangle$  by lowering the spin up state  $l$  to the  $r$  state, and the spin down state  $r$  to the spin up state  $l$ . A symmetric result holds for the  $|lr\rangle$  state. Note that in the spin basis, they are each other's Hermitian conjugate. Summation over the operators for all positions in the chain yields the Hamiltonian for the Heisenberg XX chain with  $\lambda = 0$ . The Heisenberg Hamiltonian, in terms of raising and lowering operators, for a chain of

length  $n$  is presented in equation 30.

$$H_{XX} = - \sum_{j=0}^{n-1} \left( S_-^j S_+^{j+1} + S_+^j S_-^{j+1} \right) + \lambda \sum_{j=0}^{n-1} \left( S_+^j S_-^j \right). \quad (30)$$

This chain expresses a logarithmic violation of the entanglement entropy for subsystems consisting of adjacent particles whenever  $\lambda \leq 1$ , under periodic boundary conditions [13]. It can again be shown that the energy gap closes in the limit of  $n \rightarrow \infty$  [30]. The boundary conditions for the model associated for paths is different, and can be expressed in relation to the number of kinks in the chain [31]. The expression above can be rewritten in terms of the operators  $S_x, S_y$  and  $S_z$ , and describe quantum-mechanical interactions related to the magnetic field. The factor  $\lambda$  describes a magnetic field  $B$  in the  $\hat{z}$ -direction perpendicular to the chain. When the magnetic field  $B$  is multiplied with the gyromagnetic ratio  $\gamma$  one obtains the coupling to the magnetic field  $B$ ,  $\lambda = \gamma B$ . This model is the quantum version of the Ising model, which is a classical model for ferromagnetic interactions.

Considering Motzkin paths without the  $o$  state leads to the definition of Dyck paths.

**Definition 3.** A string  $s \in \{lr\}^n$  is called a Dyck path if

1. Any initial segment with length  $j \leq n$  of  $s$  has  $L_j(s) \geq R_j(s)$
2. The total number of  $l$  is equal to  $r$ , that is:  $L_n(s) = R_n(s)$

Operations on the Dyck paths should maintain that any initial segment of the chain has more  $l$ 's than  $r$ 's. Consider the operations  $llr \leftrightarrow lrl$  and  $lrr \leftrightarrow rlr$ , as presented graphically in figures 15b and 15c. Using these operations on a Dyck path will return another Dyck path. For example, it is not possible to construct the  $rllr$  string from the  $llrr$  string using these operations. Again a class of strings can be defined

$$d_{p,q} = \underbrace{lr lr \dots lr}_{2k} \underbrace{rr \dots rr}_p \underbrace{r ll \dots ll}_q, \quad (31)$$

where  $2k + p + q = n$ . One can show that an arbitrary string  $s \in \{l, r\}^n$  of length  $n$  is equivalent to a unique  $d_{p,q}$  under the operations defined on three particles. The strings equivalent to  $d_{0,0}$  are then called Dyck paths, these are the spin- $\frac{1}{2}$  equivalent of the Motzkin paths without  $o$ . In order to construct a Hamiltonian implementing the aforementioned operations, one can write down two unnormalized states,

$$|\psi_l\rangle = |l\rangle \otimes (|lr\rangle - |rl\rangle), \quad |\psi_r\rangle = (|lr\rangle - |rl\rangle) \otimes |r\rangle.$$

This represents an interaction term for a singlet state  $|lr\rangle - |rl\rangle$ , with a spin  $l$  on the left, and a spin  $r$  on the right. The operator  $\Pi = |\psi_l\rangle \langle \psi_l| + |\psi_r\rangle \langle \psi_r|$  can then be used to construct a Hamiltonian, by writing  $\Pi_j$  for the operator acting on letters  $j, j+1$  and  $j+2$  in a string. For a chain of  $n$  spins, this Hamiltonian becomes

$$H = |r\rangle_1 \langle r|_1 + |l\rangle_n \langle l|_n + \sum_{j=1}^{n-2} \Pi_j. \quad (32)$$

Analogous to the Motzkin-path Hamiltonian, the summation over  $\Pi_j$  will annihilate any uniform superposition of the equivalence classes  $d_{p,q}$ . In order to isolate the Dyck-path state, boundary

conditions are necessary. This boundary condition is implemented by the two operators  $|r\rangle_1 \langle r|_1 + |l\rangle_n \langle l|_n$ , penalizing any state starting with  $r$  or ending with  $l$ , which is only impossible for the Dyck-path state. The Hamiltonian without boundary terms can be identified with Fredkin gate operations in quantum computing [28].

The energy gap of the Hamiltonian presented in equation 32 decays by  $n^{-c}$ , with  $c \geq 2$  [32]. Therefore it is possible that the entanglement entropy violates the area law. We will continue by investigating the entanglement entropy of this state further. The Schmidt decomposition for the Dyck-path ground state is a bit more straightforward than for the Motzkin-path state. Similar to the Motzkin-path state, one can consider the superposition of paths that reach a certain height  $m$  at a location within the string. These strings are then equivalent to the  $d_{0,m}$  strings, for which one can define a normalized state  $|D_{0,m}\rangle$ , consisting of the uniform superposition of all paths equivalent to  $d_{0,m}$ . Similarly, for  $d_{m,0}$  the  $|D_{m,0}\rangle$  state can be identified. One can now start to write down a Schmidt decomposition for the subsystems  $A$  and  $B$ , consisting of the first  $N$  particles and last  $n - N$  particles respectively. Because the aforementioned states are normalized, we can consider the Schmidt coefficients  $p_m$  as the fraction of states described by  $|D_{0,m}\rangle_A \otimes |D_{m,0}\rangle_B$ . This yields, for even  $n$ ,

$$|\mathcal{D}_n\rangle = \sum_{m=0}^N \sqrt{p_m} |D_{0,m}\rangle_A \otimes |D_{m,0}\rangle_B, \quad p_m = \frac{|D_{0,m}(N)| |D_{0,m}(n-N)|}{|D_{0,0}(n)|}. \quad (33)$$

Note that the number of paths equivalent to  $d_{0,m}$  of length  $2n+m$ , that is  $|D_{0,m}(2n+m)|$ , is given by the expression for  $D_{n,m}$  in lemma 5. We will use this lemma to derive expressions for  $|D_{0,m}(n)|$ . Further inspection of valid states yield that  $|D_{0,m}(N)| = 0$  whenever  $N$  and  $m$  have different parity, not both even nor both odd. Considering the subsystem of length  $N$ , such that the subsystem has  $m$  additional  $l$ 's yields that both  $N - m$  and  $n - N - m$  should be even, in order to accommodate the additional pairs of  $l$ 's and  $r$ 's. Because there are equal number of  $l$ 's and  $r$ 's for valid Dyck paths, one has that  $n$  is even, resulting in the condition on the parity of  $N$  and  $m$ . Writing for odd  $N$ ,  $N = 2l + 1$ , or  $N = 2l$  when  $N$  is even and  $m = 2h + 1$  or  $m = 2h$  when  $m$  is odd or even respectively. One obtains using the parity equivalence of  $N$  and  $m$  that

$$|D_{0,2h}(2l)| = \frac{2h+1}{l+h+1} \binom{2l}{l+h}, \quad |D_{0,2h+1}(2l+1)| = \frac{2h+2}{l+h+2} \binom{2l+1}{l+h+1}.$$

For large  $l$  and  $h$  one can approximate the binomial coefficients by a Gaussian in an analogous manner as for the Motzkin-path coefficients. This yields for the Schmidt coefficients

$$p_h = \frac{h^2}{Z} \exp\left(-h^2 \frac{n}{2l(\frac{n}{2} - l)}\right),$$

where  $Z$  is a normalization constant. Analogous to the entanglement entropy of the Motzkin-path state, we approximate the sum over  $m$  by an integral. Note that the dependence on  $m$  has changed into a dependence on  $h$ , which have a one-to-one relation. Depending on whether  $N$  is even or odd, only the even or odd  $m$  are counted, with  $h = \frac{m}{2}$  or  $h = \frac{m-1}{2}$  respectively. This means that a summation over  $m$  can be replaced by a summation over  $h$ . In addition, we again define variables

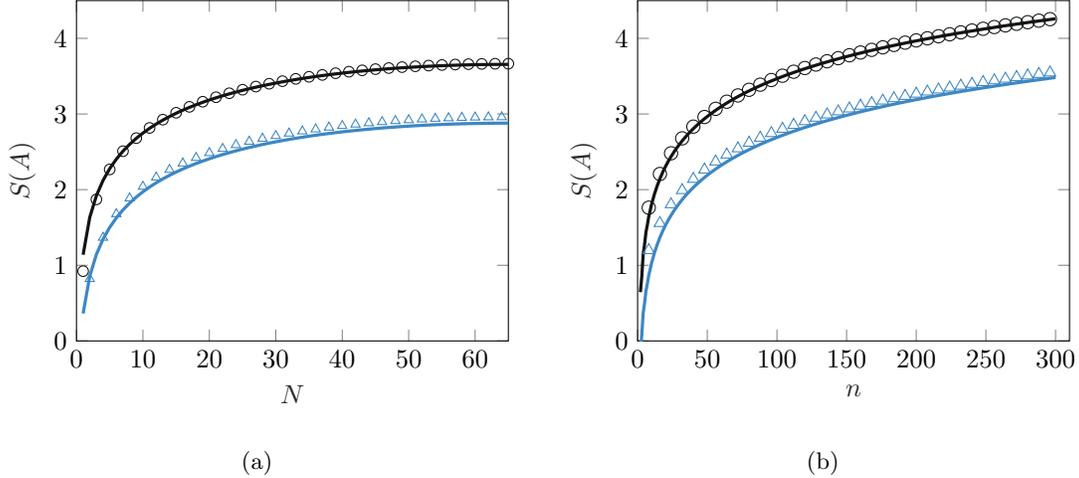


Figure 16: The entanglement entropy  $S(A)$  for the Dyck-path state and the Motzkin-path state. Figure a is the entanglement for a block  $A$  consisting of the first  $N$  spins in a chain of  $n = 130$  spins. Figure b is the entanglement entropy of the half chain, with  $N = \frac{n}{2}$  whenever  $n$  is even. The exact entanglement entropy of the Dyck-path state is represented by the blue triangles, the approximation from equation 34. For the Motzkin path entanglement entropy, the circles represent the exact entropy and the black line is the approximation found in equation 17.

$$\sigma = \sqrt{n/(N(n-N))} \text{ and } \alpha = \sigma h,$$

$$\begin{aligned} S(A) &= - \sum_{m=0}^N p_m \log_2 p_m \approx - \int_0^\infty p_h \log_2 p_h dh \\ &= - \int_0^\infty \rho_\alpha \log_2 \left( \frac{\rho_\alpha}{\sigma} \right) d\alpha = \log_2 \sigma - \int_0^\infty \rho_\alpha \log_2 \rho_\alpha d\alpha. \end{aligned}$$

We have  $\rho_\alpha = \alpha^2 \exp(-\alpha^2)/Z'$ , such that  $\int \rho_\alpha d\alpha = 1$ . The cancellation of the additional constants  $\sigma$  due to the change of variables happens in a manner that is mostly the same as for the Motzkin-path entanglement entropy. Evaluating the last expression numerically, yields

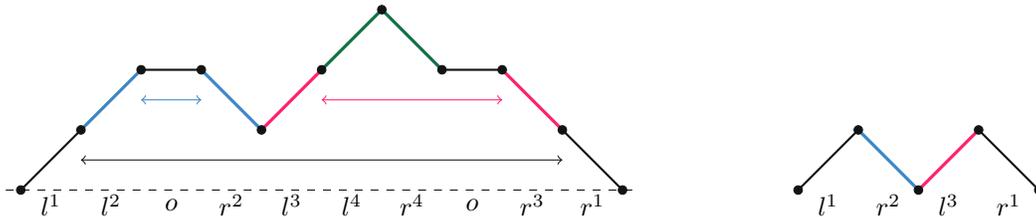
$$S(A) \approx \frac{1}{2} \log_2 \left( \frac{N(n-N)}{n} \right) + 0.3678. \quad (34)$$

The expression in equation 34 is the same as the entanglement entropy of the Motzkin paths in equation 17, up to a constant.

The scaling in terms of the initial chain length  $N$  and the total length  $n$  of the entanglement entropy for the Dyck-path states and of the Motzkin-path states is presented in figures 16a and 16b. The entanglement entropy of the half-chain, represented in blue, is slightly less than that of the Motzkin-path states, but expresses the same scaling behaviour as that of the Motzkin-path state. This identical scaling behaviour is expected by the constant difference in equations 17 and 34.

## 7.2 Path Colourings

One of the defining features of the Dyck and Motzkin-path states, is the fact that two particles are effectively paired. For each initial  $l$  in the chain, there exists a  $r$  later in the chain. So in order to generalize these models, the goal is to add an additional pair of spin states matching with each other. Since for a spin- $s$  system, there are  $2s + 1$  spin states, an integer increase in spin creates two additional spin states. In these systems, the letters will be denoted  $l^c$ ,  $o$  and  $r^c$ , where the index  $c$  can be identified as a colour [16, 28]. Integer spin chains can then be considered as colourings of Motzkin paths and half-integer spin chains as colourings of Dyck paths. An example of a coloured Motzkin path is presented in figure 17a.



(a) The coloured Motzkin path that can be identified with the state  $|l^1 l^2 o r^2 l^3 l^4 r^4 o r^3 r^1\rangle$ . In bracket notation  $([-]\{\langle\}-)$  (b) Invalid colouring of a Motzkin path.

Figure 17: Motzkin-path colourings. In figure a a valid Motzkin-path colouring is presented, while in figure b an invalid colouring is presented. In bracket notation  $(\{\})$

The notation used here makes it a bit more difficult to verify the validity of a Motzkin path. Instead of the  $l$  and  $r$  notation, one could use left brackets and right brackets  $()$ . For which the rules are familiar to any mathematician and physicist. Each opened bracket should be closed later. The  $o$  letter for the Motzkin strings, could be represented by an empty space, or a 0. Using this notation, the colourings are equivalent to using multiple types of brackets, for example  $\{\}$  and  $[\ ]$ . The rules that will define valid Motzkin paths are then equivalent to the standard bracket rules, in which all brackets opened after the opening  $[$  should be closed before it can be closed with a  $]$ . Consider for example

$$([\ ])\quad ([\ ]).$$

Then the first set of brackets is valid, but the second is not. In terms of colour such a violation is illustrated in figure 17b.

## 7.3 Coloured Motzkin paths

R. Movassagh and P. Shor start of constructing the Coloured Motzkin-path Hamiltonian analogously to the uncoloured Motzkin-path Hamiltonian. A difficulty arises due the importance of correlated pairs maintaining their height, required for correct colourings. For example, an exchange operation  $l^1 l^2 \leftrightarrow l^2 l^1$  is therefore not allowed. This leads to the same exchange operations presented in section 5.2, for each of the individual colours  $c$  respectively. These operations can be written  $l^c o \leftrightarrow o l^c$ ,  $o r^c \leftrightarrow r^c o$  and  $l^c r^c \leftrightarrow o o$ .

In the context of colourless strings, the equivalence classes  $c_{p,q}$  were key in understanding the different sets of strings that could be obtained. For a coloured string, such classes are difficult to write down. In addition to having the unmatched  $r^1 r^2 \dots l^4 l^2$ , there exists the possibility of crossed states, like  $l^1 r^2 l^2 r^1$  which cannot be reduced to zeroes using the discussed operations. The Hamiltonian that has the coloured Motzkin paths as ground state therefore requires additional terms, penalizing the crossed states.

The Hamiltonian for the coloured Motzkin paths can be deconstructed into three terms  $H = H_\partial + H_x + H_\sim$ . The Hamiltonian  $H_\sim$ , implementing the three local moves, is again given in terms of an operator  $\Pi'$ . This operator can be defined by writing the states

$$|\phi^c\rangle = |oo\rangle - |l^c r^c\rangle, \quad |\psi_l^c\rangle = |ol^c\rangle - |l^c o\rangle, \quad |\psi_r^c\rangle = |or^c\rangle - |r^c o\rangle,$$

and defining  $\Pi' = \sum_{c=1}^s [|\phi^c\rangle \langle\phi^c| + |\psi_l^c\rangle \langle\psi_l^c| + |\psi_r^c\rangle \langle\psi_r^c|]$ . This construction effectively inherits the structure found for Motzkin paths, but as a superposition for the distinct colourings. By writing  $\Pi'_{j,j+1}$  for the operator  $\Pi'$  acting at letters  $j$  and  $j+1$  in a coloured string state  $|s\rangle$ , we can define  $H_\sim = \sum_{j=1}^{n-1} \Pi'_{j,j+1}$ . The ground state of this Hamiltonian consists of all possible coloured strings. Since only the coloured Motzkin-path state is desired, additional terms need to be added. Analogous to the case of colourless Motzkin-path states, a Motzkin path will never start with  $r^c$  or end with  $l^c$ . This condition can be implemented by a boundary terms in the Hamiltonian, and can be defined  $H_\partial = \sum_{c=1}^s [|\phi^c\rangle_1 \langle\phi^c|_1 + |\psi_l^c\rangle_n \langle\psi_l^c|_n]$ , penalizing all states starting or ending with  $r^c$  or  $l^c$  respectively. Up to this point, the colourless and coloured Motzkin-path Hamiltonian are very similar, and consist simply of a summation over the individual colours. An additional constraint is necessary, due to a distinguishable feature between general strings and Motzkin-path strings: the possibility of crossed-states occurring. These occurrences are penalized by implementing

$$H_x = \sum_j^{n-1} \sum_{\substack{k,c \\ c \neq k}}^s |r^c l^k\rangle_{j,j+1} \langle r^c l^k|_{j,j+1}. \quad (35)$$

Any crossed-state will have some energy contribution, making the superposition of coloured Motzkin paths the ground state. The full Hamiltonian is then

$$H = \sum_{c=1}^s [|\phi^c\rangle_1 \langle\phi^c|_1 + |\psi_l^c\rangle_n \langle\psi_l^c|_n] + \sum_j^{n-1} \sum_{\substack{k,c \\ c \neq k}}^s |r^c l^k\rangle_{j,j+1} \langle r^c l^k|_{j,j+1} + \sum_{j=1}^{n-1} \Pi'_{j,j+1}$$

The energy gap between the ground state and first excited state satisfies the same upper and lower bounds as the colourless Motzkin-path Hamiltonian, given by  $\lambda_1 \leq \mathcal{O}(n^{-2})$ , and  $\lambda_1 \geq n^{-\mathcal{O}(1)}$  [16]. The vanishing of the gap in the thermodynamic limit of  $n \rightarrow \infty$  again allows for highly correlated states to occur.

Constructing a Schmidt decomposition is a bit more difficult than for the colourless Motzkin-path state, and only the half-chain is considered. The equivalence class states from the colourless construction are replaced by a modified state  $|C'_{p,q,x}\rangle$ , with  $p$  and  $q$  integers and  $x \in \{l^c, r^c | c = 1, \dots, s\}^{p+q}$  a coloured string with  $s$  different colours. The modified strings in  $|C'_{p,q,x}\rangle$  represent strings with  $p$  excess  $r$ 's and  $q$  excess  $l$ 's, similar to the equivalence classes for the Motzkin-path

states. The colouring of the unmatched letters is then defined by the coloured string  $x$ . Consider as an example  $l^2 r^2 r^1 l^1 r^2 \in C'_{2,1,s}$ , with  $s = r^1 l^1 r^2$ . Additionally we define the string  $\bar{s}$  as the complement string, which, for the example above becomes  $\bar{s} = l^2 r^2 l^1$ .

The Schmidt decomposition is similar to that of the colourless Motzkin-path state. One can consider all possible heights that the path reaches halfway in the string, corresponding to the letters existing in the first half of the chain with a matching letter in the second half. Using the states defined above, one can consider each of the colourings separately. For the first half of the string this then becomes  $|C'_{0,m,s}\rangle$  with  $s \in \{l^1, l^2, \dots, l^c\}^m = L_m^c$ . By defining the Schmidt decomposition again in terms of uniform superpositions, the Schmidt coefficients become the fraction of coloured strings reaching height  $m$ . Naming the first half of the chain  $A$  and the second half  $B$ , one can write for the Schmidt decomposition of the coloured Motzkin-path state of length  $n$  with  $c$  different colours

$$|\mathcal{M}_{n,c}\rangle = \sum_{m=0}^{n/2} \sqrt{p_{n/2,m,c}} \sum_{s \in L_m^c} |C'_{0,m,s}\rangle_A \otimes |C'_{m,0,\bar{s}}\rangle_B, \quad p_{n,m,c} = \frac{M_{n,m,c}^2}{\sum_{m=0}^n c^m M_{n,m,c}^2}. \quad (36)$$

$M_{n,m,c}$  represent the number of possible coloured Motzkin paths of length  $n$  reaching height  $m$  with  $c$  different colours. Starting from the case of uncoloured Motzkin paths, these coefficients were found by considering strings of  $i$  pairs of  $l$ 's and  $r$ 's, with  $m$  excess  $l$ 's. Then there are  $c^i$  ways to colour these pairs of  $l$ 's and  $r$ 's, and  $c^m$  ways to colour the excess  $l$ 's. This yields for the coefficients in equation 36, in which the colouring of the  $m$  excess  $l$ 's are already defined, that

$$M_{n,m,c} = \sum_{i=0}^{(n-m)/2} c^i M_{n,m,i} = (m+1) \sum_{i=0}^{(n-m)/2} \frac{n! c^i}{(i+m+1)! i! (n-2i-m)!},$$

where  $M_{n,m,i}$  are the coefficients from the uncoloured Motzkin-path Schmidt decomposition in equation 14. Normalizing the Schmidt decomposition yields the summation in the denominator of  $p_m$ , corresponding with coloured Motzkin paths reaching height  $m$ , with the additional  $c^m$  ways to colour the  $m$   $l$  and  $r$  pairs. Note that in this Schmidt decomposition, there are multiple terms with the same coefficient, due to the additional summation over the the different colourings  $s \in L_m^c$ . These are the number of possible ways one can colour the  $m$  excess  $l$ 's, of which there are  $c^m$ . Hence, under the additional counting of these Schmidt coefficients, the entanglement entropy of the half-chain becomes

$$S(A) = \sum_m^{n/2} c^m p_m \log_2 p_m. \quad (37)$$

A derivation similar to that of the colourless Motzkin-path states yields the following approximation for the entanglement entropy of the half-chain in terms of the number of different colours  $c$  and the chain length  $n$  [16],

$$S(A) \approx 2 \log_2(c) \sqrt{\frac{\sigma n}{\pi}} + \frac{1}{2} \log_2 \frac{n}{2} + \frac{1}{2} \log_2 \sigma + \underbrace{\left( \gamma - \frac{1}{2} \right) \log_2 e + \frac{1}{2} (1 + \log_2 \pi)}_{\approx 1.44}, \quad (38)$$

where  $\sigma = \sqrt{c}/(2\sqrt{c}+1)$  and  $\gamma \approx 0.577$  is the Euler-Mascheroni constant. Substituting  $c = 1$  yields the same result as in section 5.4. The coloured half-chain has unusually large entanglement entropy for the half-chain due to the  $\sqrt{n}$  scaling.

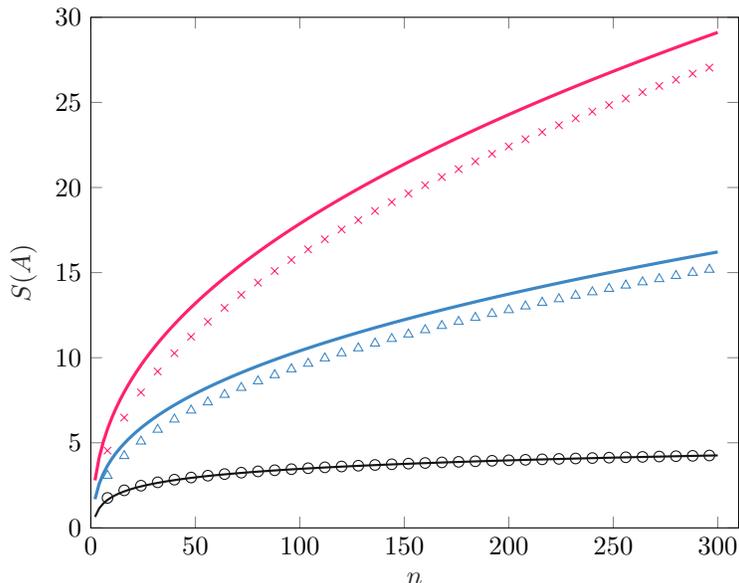


Figure 18: The entanglement entropy  $S(A)$  of the half-chain  $A$  of the coloured Motzkin-path state in terms of the chain-size  $n$ . The entanglement entropy is presented for 1, 2 and 4 colours, represented by black, blue and red respectively. The circles, triangles and crosses show the exact entanglement entropy for these respective colours, and the lines are the approximations given in equation 38.

The behaviour of the entanglement entropy of the half-chain in terms of  $n$  is presented in figure 18. The instances of  $s = 2$  colours is represented by blue and for  $s = 4$  colours in red. These are compared to the entanglement entropy for the uncoloured Motzkin half-chain represented in black. The number of different colours have a large effect on the entanglement entropy. The initial segment of the chain has greater effect on the rest of the chain as in the uncoloured situation. The approximation for the coloured Motzkin-path entanglement entropy appears to be considerably worse than that of the uncoloured Motzkin-path state. Figure 8 in the paper by Movassagh et al illustrate this slower convergence for higher order spin systems as well [16]. The colouring increases the variance in the Gaussian approximation of replacing the sum over  $p_m$  by an integral, decreasing the accuracy of the estimate.

## 7.4 Coloured Dyck paths

O. Salberger and V. Korepin extend the Dyck-path model from section 7.1 to a coloured model in a similar way to that of the coloured Motzkin-path model. Instead of having of three basis elements, the problem is reduced to two, represented by  $l^c$ 's and  $r^c$ 's, with  $c$  representing the color. The path operations presented in section 7.1 do not lend themselves easily into a coloured representation. The exchange operations are therefore slightly modified, becoming, for all colour sets  $c, j$  and  $k$ ,  $l^k l^j r^c \leftrightarrow l^j r^c l^k$  and  $l^k r^c r^j \leftrightarrow r^j l^k r^c$ . A graphical representation of these operations are presented in figure 19.



Figure 19: The exchange operations for coloured Dyck paths. On the left the exchange,  $l^1 l^2 r^2 \leftrightarrow l^2 r^2 l^1$  and on the right  $l^3 r^2 r^1 \leftrightarrow r^1 l^3 r^2$  is presented, with 1 representing blue, 2 representing red and 3 representing green.

The figure on the right an invalid coloured Dyck path, since it consists of a unmatched color pair. Note that for any valid Dyck path, particle pairing is not disturbed due to the unaffected colour being moved from front to back, or back to front, maintaining its initial height. Hence the operations above do not induce crossing or matching problems. The equivalence classes are defined by colourings of the strings  $d_{p,q}$  in expression 31.

The construction for the Hamiltonian is similar to that of the coloured Motzkin paths. Again we consider three components  $H = H_\partial + H_x + H_\sim$ . We will first discuss  $H_\sim$ . The operators implementing the two possible path operations for sets of different colors  $c, j, k$  becomes

$$|\psi_l^{c,k,j}\rangle = |l^k l^j r^c\rangle - |l^j r^c l^k\rangle, \quad |\psi_r^{c,k,j}\rangle = |l^k r^c r^j\rangle - |r^j l^k r^c\rangle.$$

Constructing an operator  $\Pi' = \sum_{c,j,k}^s [|\psi_l^{c,k,j}\rangle\langle\psi_l^{c,k,j}| + |\psi_r^{c,k,j}\rangle\langle\psi_r^{c,k,j}|]$  and again writing  $\Pi'_{j=1}$  for the operator acting on the particles at location  $j, j+1$  and  $j+2$  in the string  $s$ , one can define  $H_\sim = \sum_j^{n-2} \Pi'_j$  for the Hamiltonian with coloured string paths as ground state.  $n$  represents the length of the chain.

The crossed terms can again be prevented by an identical Hamiltonian as that of the coloured Motzkin-path Hamiltonian in equation 35, restricted to the basis of Dyck paths. The boundary conditions can be defined  $H_\partial = \sum_{c=1}^s [|r^c\rangle_1 \langle r^c|_1 + |l^c\rangle_n \langle l^c|_n]$  for a chain of  $n$  particles with  $s$  colors. The complete Hamiltonian is then

$$H = \sum_{c=1}^s [|r^c\rangle_1 \langle r^c|_1 + |l^c\rangle_n \langle l^c|_n] + \sum_{j=1}^{n-2} \Pi'_j + \sum_{j=1}^{n-1} \sum_{\substack{k,c \\ c \neq k}}^s |r^c l^k\rangle_{j,j+1} \langle r^c l^k|_{j,j+1}. \quad (39)$$

This Hamiltonian also expresses a decaying energy as the number of spins in the chain  $n \rightarrow \infty$ , scaling with  $n^{-c}$ ,  $c \geq 2$  [32].

The Schmidt decomposition for the coloured Dyck-path state can then be given using similar notation to the Motzkin-path states. We let  $|D'_{0,m,s}\rangle$  be the superposition of all strings with  $m$  excess

$l$ 's, with the colouring of the excess letters defined by the string  $s$ . Analogously  $|D'_{m,0,\bar{s}}\rangle$  represents the superposition of strings with  $m$  excess  $r$ 's, with the colouring of the excess letters defined by the conjugate of the string  $s$ , written  $\bar{s}$ . The concatenation of the two strings is then a Dyck path. For strings of length  $n$ , with  $c$  colours, and subsystem  $A$  representing the first  $N$  particles and  $B$  the last  $n - N$  particles, one finds for the Schmidt decomposition

$$|D_{n,c}\rangle = \sum_{m=0}^N \sqrt{p_{m,c}} \sum_{s \in L_m^c} |\hat{D}'_{0,m,s}\rangle_A \otimes |\hat{D}'_{m,0,\bar{s}}\rangle_B, \quad p_{m,c} = p_m \frac{c^{(N-m)/2} c^{(n-N-m)/2}}{c^{n/2}} = p_m c^{-m}.$$

The  $p_m$  are the coefficients for the Schmidt decomposition of the colourless Dyck-path state. To motivate this, consider the number of states represented in  $|\hat{D}'_{0,m,c}\rangle$ . For each of the  $(N - m)/2$  pairs  $l^i r^i$ , there are  $c$  possible colours. Combining this with the number of possible strings derived earlier in section 7.1, we find  $c^{(N-m)/2} |D_{N,m}|$  possible strings for this state. As a result we find the above expression for  $p_{m,c}$ .

The coefficients in the coloured Dyck-path state only differ by a factor  $s^{-m}$  from the colourless Dyck-path state. This is reflected in the expression for the entanglement entropy, which is given for the subsystem  $A$  in terms of the number of particles  $n$  and number of colors  $c$  by

$$S(A) = - \sum_{m=0}^N c^m p_{m,c} \log_2 p_{m,c} = - \sum_{m=0}^N p_m \log_2 (p_m c^{-m}) = S_0(A) + \log_2(c) \sum_{m=0}^N m p_m,$$

with  $S_0(A)$  representing the entanglement entropy of the uncoloured Dyck-path state, as defined in equation 34. This additional term can be computed by means of an integral, largely analogous to the construction on colourless Dyck paths. By successive approximation we obtain

$$S(A) \approx \frac{1}{2} \log_2 \left( \frac{N(n-N)}{n} \right) + 2 \log_2(c) \sqrt{\frac{2N(n-N)}{\pi n}} + 0.3678. \quad (40)$$

The relation between the entanglement entropy of the subsystem  $A$  with initial block size  $N$  is presented in figure 20a. An coloured initial segment of the chain is more closely related to the rest of the chain due to the extra limitations on the colouring of the chain.

In figure 20b the entanglement entropy of the half-chain subsystem  $A$  in terms of the chain size  $n$  is presented. Paying close attention to size of the entanglement entropy, comparing it to that of the coloured Motzkin-path state, the entanglement entropy for the coloured Dyck-path state becomes larger for  $c > 1$ . This is supported by the  $\sqrt{\sigma n}$  scaling for the Motzkin-path entanglement entropy, with  $\sigma < 0.5$  for  $c > 1$ . Resulting in a smaller  $\sqrt{n}$  coefficient for the coloured Motzkin-path state entanglement entropy. The additional pairs of  $l$ 's and  $r$ 's appear to have a greater effect on the uncertainty in the initial segment of the chain than the additional possibility of the  $o$ 's, contrary to the uncoloured situation.

Similarly to the Motzkin-path state, the estimate for the higher order spin systems appears to converge slower. This suggests that the error made in the integral over  $p_m$  is being amplified by the factor  $\log_2(c)$  instead of the increased variance on the approximation.

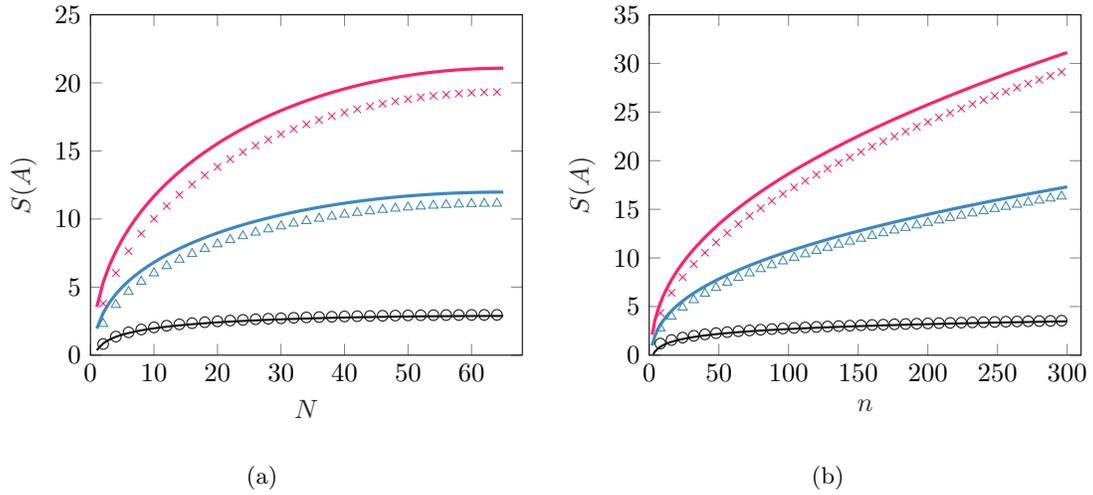


Figure 20: The entanglement entropy  $S(A)$  for the coloured Dyck-path state with 1, 2 and 3 colours, represented by black, blue and red respectively. Figure a is the entanglement for a block  $A$  consisting of the first  $N$  spins in a chain of  $n = 130$  spins. Figure b is the entanglement entropy of the half chain, with  $N = \frac{n}{2}$  whenever  $n$  is even. The circles, triangles and crosses represent the exact entanglement entropy for the respective colours, and the lines represent the approximation found in equation 40.

## 8 Discussion and Outlook

### 8.1 Spin-Operator interactions

The models discussed in this thesis are one-dimensional quantum chains with a dimension of  $2s + 1$ . Usually these states are identified with spin- $s$  states. Recall from section 3 that these higher order spin- $s$  systems are most likely to describe systems of interacting electrons or protons. Specific superpositions of these particles result in higher-order spin. Groupings of protons can exist as nuclear spins, implying interacting atoms.

In section 7.1, a clear relation between completely random spin- $\frac{1}{2}$  path model under path interactions and the Heisenberg spin chain was drawn. This relation might suggest investigation of the discussed operations in terms of spin operators. R. Movassagh presents the spin operator representation of the colourless Motzkin path model [33]. The Motzkin-path Hamiltonian has a rather involved description in terms of spin operators. The interactions described in these models are therefore not directly related to known spin interactions.

### 8.2 Area Law violations

The closing of the energy gap in the limit of  $n \rightarrow \infty$  allows the entanglement entropy of the discussed models to violate the area law. The models presented in this thesis have local interactions. The boundary conditions however, are often considered mathematically nonlocal, since they rely on an external property at the boundary of the chains. Boundary conditions are very common, so they do not necessarily affect the extend to which these models can be considered physically reasonable.

Swingle and McGreevy mention the violation of the area law for the colourless Motzkin-path model [19]. Denoting a Hamiltonian  $H$  restricted to the subsystem  $A$  by  $H_A$ , the entanglement entropy is closely linked to the ground-state degeneracy of  $H_A$ . Let  $\rho$  be the ground-state density operator of  $H$  and  $\rho_A$  be its reduced density operator. Then  $\text{tr}(H_A \rho_A) = 0$ . An upper bound is then given by the number of independent eigenstates of  $\rho_A$ . By definition, that is equivalent to the degeneracy of the restricted Hamiltonian  $H_A$ . We find

$$S(A) \leq - \sum_{j=1}^{G(H_A)} \frac{1}{G(H_A)} \log_2 \left( \frac{1}{G(H_A)} \right) = \log_2 G(H_A) \quad (41)$$

where  $G(H_A)$  denotes the degeneracy of the Hamiltonian  $H_A$ . Note that this means the ground state degeneracy will grow with the system size for the models in this thesis. These restricted Hamiltonians consist of the standard operators  $\Pi_{j,j+1}$  combined with the boundary term only at position 1 or at position  $n$ . We will first discuss the colourless Motzkin-path Hamiltonian. We will consider the restricted Hamiltonian  $H_A$  to a subsystem  $A$  of length  $N$ , for which only the boundary condition  $|r\rangle\langle r|$  remains. As a result the remaining ground states of  $H_A$  are exactly the states  $|C_{0,q}\rangle$ , of which there are  $N$ . Using relation 41, and that  $\log_2(x)$  is strictly increasing, we indeed find the  $S(A) \leq \log_2 G(H_A) = \log_2(N)$ , allowing the violation of the area law. Note that this argument holds for the colourless Dyck path model as well.

The ground state of a subsystem for the coloured path Hamiltonians also has its span as a subspace. The necessary boundary condition is inherited for the initial segment in the subsystem  $A$ , and any

violating colourizations are still penalized. We can similarly investigate the degeneracy of the restricted Hamiltonian  $H_A$  for the coloured Motzkin paths. Investigating the number of classes  $|C'_{0,q,x}\rangle$ , we again find that  $q$  grows linearly with  $N$ . But now each of the colourizations of these  $q$  excess  $l$ 's contribute a unique ground state to the restricted Hamiltonian  $H_A$ . These excess  $q$  elements can be coloured in  $c^q$  ways, hence we obtain that there are  $G(H_A) = \sum_{q=0}^N c^q = \frac{c}{c-1}(c^N - 1)$  ground states. Using that  $c^N \gg 1$ , one obtains  $S(A) \leq \log_2 \frac{c}{c-1} + N \log_2(c)$  bound for large  $N$ .

This argument illustrates why these highly entangled ground states are able to occur. The relation to the combinatorial path models allow a large number of equivalence classes to be constructed while isolating only a single equivalence class using the two boundary conditions. Breaking the chain into two segments removes one of the boundary terms and separates the original ground state into a large set of ground states on the restricted Hamiltonian. Combinatorial models that can similarly expressed in a large number of equivalence classes, and can be projected onto a spin system, will therefore likely lead to similar results. The area law violation will likely occur whenever the full Hamiltonian has a unique ground state that becomes extensively degenerate when considering the restricted Hamiltonian.

Whether similar models can be constructed in two dimensions is an important and interesting topic for further research. The spin lattice would be mapped to surfaces in a three-dimensional space. Similar interactions can be constructed and would effectively map to maximum height variations of this surface. The entanglement entropy would be determined by the possible configurations of the perimeter of the subsystem.

### 8.3 Tensor Network representation of Coloured Path models

Tensor Network states often allow for more efficient simulation of quantum-mechanical systems in comparison to considering the full Hilbert space, which grows exponentially with system size. Therefore it can be very practical to find a tensor network representation of the models in this thesis in order to study the behaviour of these systems more effectively. Alexander et al. introduced a tensor network representation for the ground state of the path models presented in this thesis [34]. In section 3.5 we briefly mentioned the the Schmidt decomposition in the tensor network notation. The entanglement entropy is therefore entirely contained in the boundaries of the subsystem, where the matrix  $D$  occurs. The exact details of the entanglement entropy estimate will be clarified later in this section. We will start by presenting the tensor network model by Alexander et al.

As illustrated in figure 17a, the pairs of colours are directly related in their height. By defining a set of tiles these relations can be constructed as a tiling of a pyramid. This process is illustrated in figure 21.

Let us shortly discuss the situation of the colourless model for the process sketched in figure 21a and 21b. Starting from the left end of the string, one can start placing the tiles presented for a single colour in figure 21a moving towards the right. Whenever an  $l$  is present, the tilings associated to that particle move up until reaching another path element or the base of the pyramid. Then the horizontal tilings are placed until the next time an  $r$  is found, connecting these. The coloured Motzkin paths are obtained by providing different colourizations of these connected components and Dyck paths are obtained by removing the tiles associated to  $o$ .

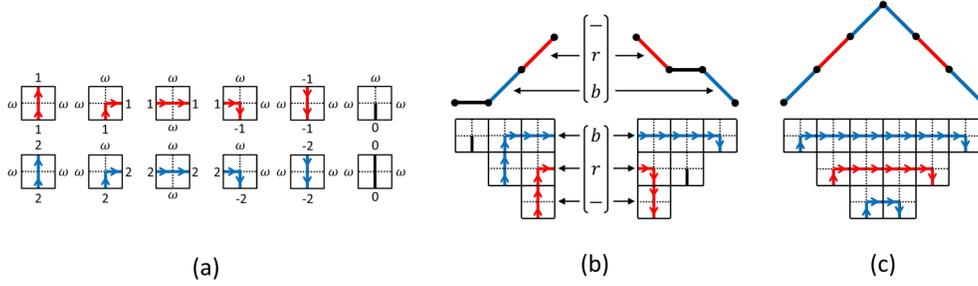


Figure 21: For the spin-2 models the associated 2-coloured Motzkin-path states are mapped onto a tiling of tensors. The 12 tiles are presented in (a). Two examples of coloured Motzkin paths of length 6 and their associated tensor network state are presented. Figures by Alexander et al. [34].

The different tiles are identified with tensors, which can be constructed using a set of Kronecker- $\delta$ 's. These tilings are for a colour  $r$  presented in figure 22a. Here  $t$  represents a deformation parameter, which for our models is 1,  $t = 1$ .  $\omega$  represents an index, which by construction is prohibited to occur. Now the tensor  $B$  is constructed in terms of the tiles  $A$ , by considering their sum,

$$B = \sum_{i=1}^5 \sum_{j=0}^c A_i(j) + A_6 + A_7.$$

Thus the tensor  $B$  represents the uniform summation over all tiles  $A_j$ . A tensor network can now be constructed using the tensors  $B$ , which is presented in figure 22b. The boundaries on top and on the left consist of the tiles  $v_j = \delta_{j,\omega}$ , prohibiting the invalid tilings of the Motzkin path state. The spins states can be identified with the open ends of tensors at the bottom, they are projected by the operator  $I - |\omega\rangle\langle\omega|$  to ensure they are orthogonal to the states associated with  $\omega$ .

The tensor  $B$  is the summation over all tiles  $A_j$ , the tensor network in figure 22b exactly represents the uniform superposition of all Motzkin paths. Contracting each column into a single tensor, the tensor network is equivalent with a matrix product state. For matrix product states we have that there are two bonds at the boundary. Recalling from section 4.5, we find  $S \leq 2 \log_2 D$ . Thus it is expected that the bond dimension grows exponentially with the system size in order to facilitate the power law violation of the area law for entanglement entropy.

We will consider the half-chain subsystem and analyse the bond dimension for the cut along the middle of the chain. For each vertical tile with a bond across this cut, there are  $c$  possible bonds, indicated by connection of tensors  $A_3$ ,  $A_4$  and  $A_5$  for each of the  $c$  colors across this gap. Since this number of vertical tiles grows with  $n/2$  for a chain of length  $n$ , we find that the bond dimension scales  $D = c^{\frac{n}{2}}$ , yielding a bound  $S \leq n \log_2 c$  for the entanglement entropy that can be encoded in this state.

Alternatively one might wonder whether the path models presented in this thesis can be represented by a Multiscale Entanglement Renormalization Ansatz (MERA) network [35, 36]. The MERA network is a construction of a tensor network that allows for entanglement entropy scaling of

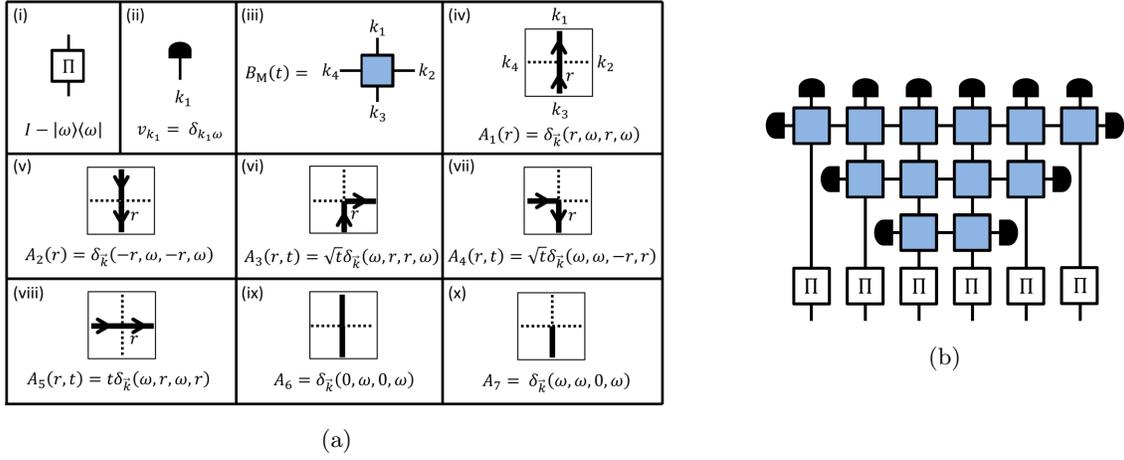


Figure 22: The tensor tiles that constitute the pyramid configuration are presented in figure (a). The associated tensor network to the pyramid configuration is presented in figure (b). Figures by Alexander et al. [34].

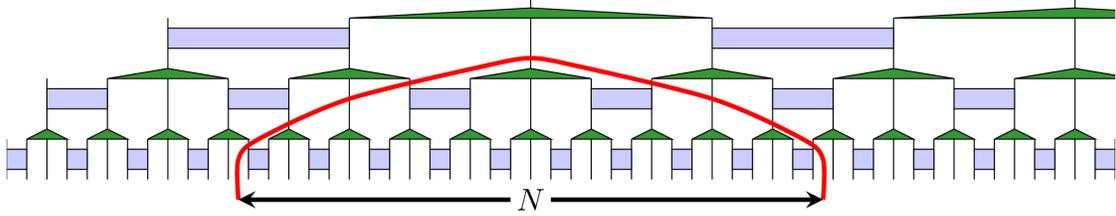


Figure 23: The Multiscale Entanglement Renormalization Ansatz network, where the cutting of bonds is indicated for a subsystem of  $N$  particles. Figure by Bridgeman and Chubb [35].

$S \leq \log_2 N \log_2 D$ , due to the number of bonds scaling as  $\log_2 N$  with the subsystem size. An example of a MERA network is presented in figure 23, in which the separation of a subsystem of length  $N$  is indicated.

## 9 Conclusion

We have shown that spin chains projected to Dyck and Motzkin paths under Hamiltonians with local interactions in the bulk express violations of the area law. More specifically, the subsystem consisting of the first  $N$  particles in the colourless Dyck-path and Motzkin-path states express an entanglement entropy  $S \sim \log_2 N$ . For the coloured Dyck-path and Motzkin-path states the entanglement entropy of the same subsystem has a leading  $S \sim \sqrt{N}$  scaling of the entanglement entropy. The results are summarized in table 2.

Due to the relations between the energy gap and the bound on the entanglement entropy, the energy gap of the Motzkin-path Hamiltonian is investigated further. We presented an alternative proof on the bound of the energy gap  $\Delta E$ , scaling in terms of the number of particles  $n$  as  $\Delta E = \mathcal{O}(n^{-2})$  in the limit  $n \rightarrow \infty$ . Hence we find  $\Delta E \rightarrow 0$ , indicating the vanishing of the energy gap and allowing highly entangled ground states to occur. The other models discussed in this thesis have similar scaling of the energy gap,  $\Delta E = \mathcal{O}(n^{-c}), c \geq 2$  [16, 28].

Table 2: Leading order entanglement entropy scaling for the half-integer-spin Dyck-path and integer-spin Motzkin-path models.  $c$  is the number of colours.

Spin	Entanglement Entropy
$\frac{1}{2}$	$\log_2(n)$
1	$\log_2(n)$
$\frac{c}{2}$	$\sqrt{n} \log_2(c)$
$c$	$\sqrt{n} \log_2(c)$

## References

- [1] Rangamani, M., & Takayanagi, T. (2017). *Holographic entanglement entropy* (Vol. 931). Springer.
- [2] Verlinde, E. P. (2017). Emergent Gravity and the Dark Universe. *SciPost Phys.*, 2.
- [3] Griffiths, D. J., & Schroeter, D. F. (2018). *Introduction to quantum mechanics*. Pearson.
- [4] Baez, J., & Munian, J. P. (1994). *Gauge fields, knots and gravity*. World Scientific Publishing Co. Pte. Ltd.
- [5] Nielsen, M. A., & Chuang, I. L. (2010). *Quantum computation and quantum information*. Cambridge University Press.
- [6] Pensrose, R. (1971). Applications of negative dimensional tensors. *Combinatorial Mathematics and its Applications*.
- [7] Schroeder, D. V. (2000). *An introduction to thermal physics*. Addison Wesley Longman.
- [8] von Neumann, J. (1927). Thermodynamik quantenmechanischer Gesamtheiten. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*.
- [9] Hall, B. C. (2003). *Lie groups, lie algebras, and representations: An elementary introduction*. Springer.
- [10] Lanford, O. E., & Robinson, D. W. (1968). Mean entropy of states in quantum-statistical mechanics. *Journal of Mathematical Physics*, 9.
- [11] Lieb, E. H., & Ruskai, M. B. (1973). Proof of the strong subadditivity of the quantum-mechanical entropy. *Journal of Mathematical Physics*, 14.
- [12] Bombelli, L., Koul, R. K., Lee, J., & Sorkin, R. D. (1986). Quantum source of entropy for black holes. *Phys. Rev. D*, 34.
- [13] Latorre, J. I., & Riera, A. (2009). A short review on entanglement in quantum spin systems. *Journal of Physics A*, 42(50).
- [14] Cho, J. (2018). Realistic area-law bound on entanglement from exponentially decaying correlations. *Phys. Rev. X*, 8.
- [15] Hastings, M. B. (2007). An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics*, (08).
- [16] Movassagh, R., & Shor, P. W. (2016). Supercritical entanglement in local systems: Counterexample to the area law for quantum matter. *Proceedings of the National Academy of Sciences*, 113(47).
- [17] Fradkin, E., & Moore, J. E. (2006). Entanglement entropy of 2d conformal quantum critical points: Hearing the shape of a quantum drum. *Phys. Rev. Lett.*, 97.
- [18] Gioev, D., & Klich, I. (2006). Entanglement entropy of fermions in any dimension and the Widom conjecture. *Phys. Rev. Lett.*, 96.
- [19] Swingle, B., & McGreevy, J. (2016). Area law for gapless states from local entanglement thermodynamics. *Phys. Rev. B*, 93.
- [20] Bravyi, S., Caha, L., Movassagh, R., Nagaj, D., & Shor, P. W. (2012). Criticality without frustration for quantum spin-1 chains. *Phys. Rev. Lett.*, 109.
- [21] Grimmett, G., & Welsh, D. (2014). *Probability: An introduction* (2nd ed.). Oxford University Press.
- [22] Drake, B. (2009). Limits of areas under lattice paths. *Discrete Mathematics*, 309(12).
- [23] Sulanke, R. A. (2000). Moments of generalized Motzkin paths. *Journal of Integer Sequences*, 3.

- [24] Prokhorov, Y. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theory of Probability and its Applications*, 1(2).
- [25] Kortchemski, I. (n.d.). Normalized Brownian excursion. Retrieved July 10, 2020, from <https://igor-kortchemski.perso.math.cnrs.fr/images.html>
- [26] Takács, L. (1991). A bernoulli excursion and its various applications. *Advances in Applied Probability*, 23(3).
- [27] Janson, S. (2007). Brownian excursion area, Wright’s constants in graph enumeration, and other Brownian areas. *Prob. Surveys*, 4.
- [28] Salberger, O., & Korepin, V. (2017). Fredkin Spin Chain. *Rev. Math. Phys.*, 29(10).
- [29] Gosset, D., Terhal, B. M., & Vershynina, A. (2015). Universal adiabatic quantum computation via the space-time circuit-to-Hamiltonian construction. *Phys. Rev. Lett.*, 114.
- [30] Cloizeaux, J. D., & Gaudin, M. (1966). Anisotropic linear magnetic chain. *Journal of Mathematical Physics*, 7(8).
- [31] Bravyi, S., Poulin, D., & Terhal, B. (2010). Tradeoffs for reliable quantum information storage in 2d systems. *Phys. Rev. Lett.*, 104.
- [32] Movassagh, R. (2018). The gap of Fredkin quantum spin chain is polynomially small. *Annals of Mathematical Sciences and Applications*, 3(2).
- [33] Movassagh, R. (2017). Entanglement and correlation functions of the quantum Motzkin spin-chain. *Journal of Mathematical Physics*, 58(3).
- [34] Alexander, R. N., Ahmadain, A., Zhang, Z., & Klich, I. (2019). Holographic rainbow networks for colorful Motzkin and Fredkin spin chains. *Phys. Rev. B*, 100.
- [35] Bridgeman, J. C., & Chubb, C. T. (n.d.). Hand-waving and interpretive dance: An introductory course on tensor networks. *Journal of Physics A*, 50(22).
- [36] Vidal, G. (2007). Entanglement renormalization. *Phys. Rev. Lett.*, 99.

## Appendices

### A Entanglement of the product state

We will consider the diagonalization of the density matrices  $\rho_A = U_A V_A U_A^\dagger$  and  $\rho_B = U_B V_B U_B^\dagger$ . Then the product state can be written as  $\rho_A \otimes \rho_B = (U_A \otimes U_B) (V_A \otimes V_B) (U_A \otimes U_B)^\dagger$ , giving the diagonalization of the product state. Consider the logarithm term

$$\begin{aligned} \log_2 V_A \otimes V_B &= \text{diag} (\log_2 (\lambda_1^A V_B), \log_2 (\lambda_2^A V_B), \dots, \log_2 (\lambda_n^A V_B)) \\ &= \text{diag} (I_n \log_2 \lambda_1^A + \log_2 V_B, I_n \log_2 \lambda_2^A + \log_2 V_B, \dots, I_n \log_2 \lambda_n^A + \log_2 V_B,) \\ &= \log_2 V_A \otimes I_m + I_n \otimes \log_2 V_B. \end{aligned}$$

Where it is important to note that these equalities hold by careful inspection of the elementwise interaction of the logarithm on the diagonal matrices  $V_A$  and  $V_B$ . Interestingly, this is an interesting similarity to the scalar logarithm, the logarithm of the tensor product of diagonal matrices becomes the direct sum of the components. The reason that we consider only the logarithm of the diagonal matrices, is due to the detail that  $\log_2(UVU^\dagger) = U \log_2(V)U^\dagger$  which was derived in the paragraph succeeding the definition of entropy in equation 6, section 4.2. Similarly, we can now apply the cyclic property, and only consider

$$(V_A \otimes V_B) \log_2(V_A \otimes V_B) = (V_A \log_2 V_A) \otimes V_B + V_A \otimes (V_B \log_2 V_B).$$

Now we can calculate the entanglement entropy directly, applying that  $\text{tr}(V_A) = 1 = \text{tr}(V_B)$ , by definition of the density matrix. That is,

$$\begin{aligned} S(\rho_A \otimes \rho_B) &= -\text{tr}((V_A \log_2 V_A) \otimes V_B) - \text{tr}(V_A \otimes (V_B \log_2 V_B)) \\ &= -\text{tr}(V_A \log_2 V_A) \times \text{tr}(V_B) - \text{tr}(V_A) \times \text{tr}(V_B \log_2 V_B) \\ &= -\text{tr}(V_A \log_2 V_A) - \text{tr}(V_B \log_2 V_B) \\ &= S(\rho_A) + S(\rho_B). \end{aligned}$$

### B Derivation of Schmidt decomposition coefficients

In order to approximate the factor  $M_{n,m,i}$  for large  $n$  and  $m$ , Bravyi et al. [20] start by making two substitutions  $\alpha = m/\sqrt{n}$  and  $\beta = (i - \frac{n}{3})/\sqrt{n}$ . This yields

$$M_{n,m,i} = \frac{n!(m+1)}{i!(m+i+1)!(n-2i-m)!} = \frac{n!(\sqrt{n}\alpha+1)}{(\frac{n}{3}+\sqrt{n}\beta)!(\frac{n}{3}+\sqrt{n}(\alpha+\beta)+1)!(\frac{n}{3}-\sqrt{n}(\alpha+2\beta))!}.$$

A common technique for approximating the ratio's of factorials is by considering the logarithm of  $M_{n,m,i}$  and applying Stirling's approximation, given by  $\log n! \approx n \log n - n + \frac{1}{2} \log 2\pi n$ . This

approximation can be used in order to obtain

$$\begin{aligned}
\log M_{n,m,i} &\approx \frac{1}{2} \log n + \log \alpha + n \log n - n + \frac{1}{2} \log 2\pi n \\
&\quad - \left( \frac{n}{3} + \sqrt{n}\beta + \frac{1}{2} \right) \log \left( \frac{n}{3} + \sqrt{n}\beta \right) + \left( \frac{n}{3} + \sqrt{n}\beta \right) - \frac{1}{2} \log 2\pi \\
&\quad - \left( \frac{n}{3} + \sqrt{n}(\alpha + \beta) + \frac{3}{2} \right) \log \left( \frac{n}{3} + \sqrt{n}(\alpha + \beta) + 1 \right) + \left( \frac{n}{3} + \sqrt{n}(\alpha + \beta) + 1 \right) - \frac{1}{2} \log 2\pi \\
&\quad - \left( \frac{n}{3} - \sqrt{n}(\alpha + 2\beta) + \frac{1}{2} \right) \log \left( \frac{n}{3} - \sqrt{n}(\alpha + 2\beta) \right) + \left( \frac{n}{3} - \sqrt{n}(\alpha + 2\beta) \right) - \frac{1}{2} \log 2\pi \\
&\approx \log n + \log \alpha + n \log n - \log 2\pi + 1 - \left( \frac{n}{3} + \sqrt{n}\beta + \frac{1}{2} \right) \left[ \log \frac{n}{3} + \log \left( 1 + \frac{3\beta}{\sqrt{n}} \right) \right] \\
&\quad - \left( \frac{n}{3} + \sqrt{n}(\alpha + \beta) + \frac{3}{2} \right) \left[ \log \frac{n}{3} + \log \left( 1 + \frac{3(\alpha + \beta)}{\sqrt{n}} + \frac{3}{n} \right) \right] \\
&\quad - \left( \frac{n}{3} - \sqrt{n}(\alpha + 2\beta) + \frac{1}{2} \right) \left[ \log \frac{n}{3} + \log \left( 1 - \frac{3(\alpha + 2\beta)}{\sqrt{n}} \right) \right].
\end{aligned}$$

Note that the approximation for  $\log(\sqrt{n}\alpha + 1) \approx \log(\sqrt{n}\alpha)$  was applied, assuming  $m > 0$ . This will have minor consequences for the coefficients in the Schmidt decomposition, due to the relatively small size of  $M_{n,0,i}$  compared to the sum over all  $M_{n,m,i}$  for large  $n$  and  $N$ . We will discuss the situation of  $\alpha = 0$  again later.

In this form, one can start to consider the scale of the terms within the logarithms. By definition,  $\alpha/\sqrt{n}$  and  $\beta/\sqrt{n}$  should always be smaller than 1, and larger than  $-1$ . Hence one can consider a second order Taylor expansion in terms of  $\alpha$  and  $\beta$ , by applying  $\log(1+x) = x - \frac{1}{2}x^2 + \mathcal{O}(x^3)$ . Additionally, since  $n \gg \alpha$  and  $n \gg \beta$ , one can neglect terms with  $n$  in the numerator, and, terms of  $\alpha^3$  or  $\beta^3$ . Hence

$$\begin{aligned}
\dots &\approx \log \alpha - \log 2\pi - \frac{3}{2} \log n + \left( n + \frac{5}{2} \right) \log 3 + 1 \\
&\quad - (3\beta^2 + 3(\alpha + \beta)^2 + 3(\alpha + 2\beta)^2) \\
&\quad - \left( \frac{n}{3} + \frac{1}{2} \right) \left[ \frac{3(\alpha + \beta)}{\sqrt{n}} + \frac{3\beta}{\sqrt{n}} - \frac{3(\alpha + 2\beta)}{\sqrt{n}} - \frac{1}{2} \left( \frac{9(\alpha + \beta)^2}{n} + \frac{9\beta^2}{n} + \frac{9(\alpha + 2\beta)^2}{n} \right) \right] \\
&\quad - \frac{n}{3} \frac{3}{n} \\
&\approx \log \alpha - \log 2\pi - \frac{3}{2} \log n + \left( n + \frac{5}{2} \right) \log 3 - 6\alpha^2 - 18\alpha\beta - 18\beta^2 + \frac{1}{6} (18\alpha^2 + 54\alpha\beta + 54\beta^2) \\
&\approx \log \left( \frac{3\sqrt{3}}{2\pi n^{\frac{3}{2}}} 3^{n+1} \alpha \right) - 3\alpha^2 - 9\alpha\beta - 9\beta^2.
\end{aligned}$$

Which yields equation 15, as desired.

The situation of  $\alpha \neq 0$  yielded two additional terms  $\frac{1}{2} \log n$  and  $\log \alpha$  in the expression for  $M_{n,m,i}$ . The approximations later in the derivation above are not affected for the case of  $\alpha = 0$ , hence these

additional terms can simply be subtracted. This results in

$$\log M_{n,0,i} \approx \log \left( \frac{3\sqrt{3}}{2\pi n^2} 3^{n+1} \right) - 9\beta^2.$$

### C Normalization Constant for Phase-Shifted Motzkin-path states

In order to derive the scaling of the energy gap, we will start by rewriting the definition of the normalization constant. Writing out the terms,

$$\begin{aligned} K^2 &= \sum_{s \in \mathcal{M}_n} \left[ \cos \left( \pi \frac{d(s)}{d_{\max}} \right) - \xi \right]^2 = \sum_{s \in \mathcal{M}_n} \left[ \cos^2 \left( \pi \frac{d(s)}{d_{\max}} \right) - 2\xi \cos \left( \pi \frac{d(s)}{d_{\max}} \right) + \xi^2 \right] \\ &= \sum_{s \in \mathcal{M}_n} \left[ \cos^2 \left( \pi \frac{d(s)}{d_{\max}} \right) \right] - \frac{2}{M_n} \sum_{s,t \in \mathcal{M}_n} \left[ \cos \left( \pi \frac{d(s)}{d_{\max}} \right) \cos \left( \pi \frac{d(t)}{d_{\max}} \right) \right] \\ &\quad + M_n \frac{1}{M_n^2} \sum_{s,t \in \mathcal{M}_n} \left[ \cos \left( \pi \frac{d(s)}{d_{\max}} \right) \cos \left( \pi \frac{d(t)}{d_{\max}} \right) \right] \\ &= \sum_{s \in \mathcal{M}_n} \left[ \cos^2 \left( \pi \frac{d(s)}{d_{\max}} \right) \right] - \frac{1}{M_n} \sum_{s,t \in \mathcal{M}_n} \left[ \cos \left( \pi \frac{d(s)}{d_{\max}} \right) \cos \left( \pi \frac{d(t)}{d_{\max}} \right) \right]. \end{aligned}$$

One can now investigate the two contributions in more detail separately. We will perform a second order Taylor expansion on the  $\cos^2$  term, which is  $\cos(x) = 1 - x^2/2! + \mathcal{O}(x^3)$ . Additionally we will use the mean for  $d(s)/d_{\max}$ , obtaining

$$\begin{aligned} \sum_{s \in \mathcal{M}_n} \left[ \cos^2 \left( \pi \frac{d(s)}{d_{\max}} \right) \right] &= \sum_{s \in \mathcal{M}_n} \left[ 1 - \frac{\pi^2}{2} \left( \frac{d(s)}{d_{\max}} \right)^2 \right]^2 \\ &= \sum_{s \in \mathcal{M}_n} \left[ 1 - \pi^2 \left( \frac{d(s)}{d_{\max}} \right)^2 + \frac{\pi^4}{4} \left( \frac{d(s)}{d_{\max}} \right)^4 \right] \\ &\approx \sum_{s \in \mathcal{M}_n} \left[ 1 - \pi^2 \left\langle \frac{d(s)}{d_{\max}} \right\rangle^2 + \frac{\pi^4}{4} \left\langle \frac{d(s)}{d_{\max}} \right\rangle^4 \right] \\ &= M_n \left[ 1 - \pi^2 \left\langle \frac{d(s)}{d_{\max}} \right\rangle^2 + \frac{\pi^4}{4} \left\langle \frac{d(s)}{d_{\max}} \right\rangle^4 \right]. \end{aligned}$$

Using the identity  $\cos(a) \cos(b) = \cos(a-b) - \sin(a) \sin(b)$ , we can rewrite the product of cosines. The difference  $a-b$  is generally small for  $d(s)/d_{\max}$ , hence the term  $\cos(a-b) \approx 1$ . Additionally

we will perform a small angle approximation on sine terms, considered  $\sin(x) \approx x$ . We obtain

$$\begin{aligned}
& \sum_{s,t \in \mathcal{M}_n} \left[ \cos\left(\pi \frac{d(s)}{d_{\max}}\right) \cos\left(\pi \frac{d(t)}{d_{\max}}\right) \right] \\
&= \sum_{s,t \in \mathcal{M}_n} \left[ \cos\left(\pi \frac{d(s) - d(t)}{d_{\max}}\right) - \sin\left(\pi \frac{d(s)}{d_{\max}}\right) \sin\left(\pi \frac{d(t)}{d_{\max}}\right) \right] \\
&\approx \sum_{s,t \in \mathcal{M}_n} \left[ 1 - \left(\pi \frac{d(s)}{d_{\max}}\right) \left(\pi \frac{d(t)}{d_{\max}}\right) \right] \\
&\approx \sum_{s,t \in \mathcal{M}_n} \left[ 1 - \pi^2 \left\langle \frac{d(s)}{d_{\max}} \right\rangle^2 \right] = M_n^2 \left( 1 - \pi^2 \left\langle \frac{d(s)}{d_{\max}} \right\rangle^2 \right)
\end{aligned}$$

Note that we obtain  $M_n^2$  due to the summation over pairs of  $s, t$ , of which there are  $M_n \times M_n$ . Bringing terms together, one obtains the result

$$K^2 \approx \frac{M_n \pi^4}{4} \left\langle \frac{d(s)}{d_{\max}} \right\rangle^4.$$