Implementations and applications of Rényi entanglement in Monte Carlo simulations of spin models

by

Stephen Inglis

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Although entanglement is a well studied property in the context of quantum systems, the ability to measure it in Monte Carlo methods is relatively new. Through measures of the Rényi entanglement entropy and mutual information one is able to examine and characterize criticality, pinpoint phase transitions, and probe universality. We describe the most basic algorithms for calculating these quantities in straightforward Monte Carlo methods and state of the art techniques used in high performance computing. This description emphasizes the core principal of these measurements and allows one to both build an intuition for these quantities and how they are useful in numerical studies.

Using the Rényi entanglement entropy we demonstrate the ability to detect thermal phase transitions in the Ising model and XY model without use of an order parameter. The scaling near the critical point also shows signatures identifying the universality class of the model. Improved methods are explored using extended ensemble techniques that can increase calculation efficiency, and show good agreement with the standard approach. We explore the "ratio trick" at finite temperature and use it to explore the quantum critical fan of the one dimensional transverse field Ising model, showing agreement with finite temperature and finite size scaling from field theory. This same technique is used at zero temperature to explore the geometric dependence of the entanglement entropy and examine the universal scaling functions in the two dimensional transverse field Ising model. All of this shows the multitude of ways in which the study of the Rényi entanglement entropy can be efficiently and practically used in conventional and exotic condensed matter systems, and should serve as a reference for those wishing to use it as a tool.

Acknowledgements

Dedication

I dedicate this work to my wife, Tiffany. That I have been so lucky to find someone who also loves the pursuit of the unknown is my greatest fortune, and to share this life with you is my greatest joy. I also dedicate this to my parents, who have been supporting in all facets of my life. Their help has given me the opportunity to pursue what I love without compromise.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Why are we interested in the study of entanglement in condensed matter systems? For those outside of this bubble of research, this question is not often anwsered in a satisfying way. Although explained better below, the short answer is that the Rényi entanglement entropy offers a measurement that allows one to positively classify phases that were previously only determined to exist through a process of elimination. In the design and study of these entanglement measures a large number of alternative novel applications have been found, equally interesting as their original purpose. Though we do not specifically address the type of measurements needed to identify these aforementioned elusive phases, the framework discussed in this thesis is sufficient for designing such numerical studies, and outlines the fundamental nature and characteristics of these new measurements in the context of classical and quantum Monte Carlo.

Although the Landau paradigm of phases and phase transitions has enjoyed decades of success [1, 2], there are quantum phases of matter that defy this classification. The earliest examples of such phases are the fractional quantum Hall states, each fractionally filling a well defined state but many of them sharing symmetries, and instead characterized by a topological order [3, 4]. Many exotic states of matter that break the Landau paradigm have been found since, a large class of them known as the quantum spin liquids. An idea has emerged in the study of spin liquids: although these systems cannot be classified by an order parameter, the characteristics of entanglement within the phase can be used to distinguish and identify them [5, 6]. With this idea in hand, theoretical calculations of the entanglement entropy followed, leading to algorithms for a broad range of numerical techniques, most recently applied to Monte Carlo approaches.

Fundamentally, the entanglement entropy should be viewed as a resource in the same
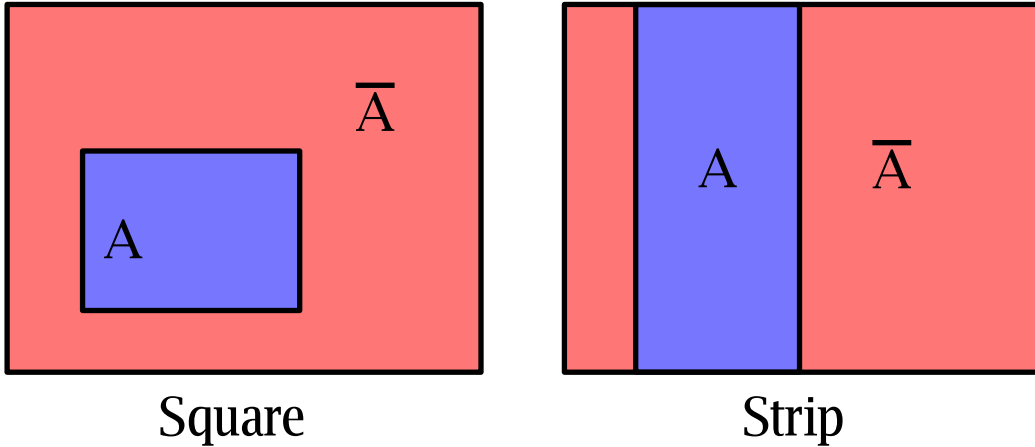
Figure 1.1: The basic region geometry studied in numerical calculations of the Rényi entanglement entropies. The strip region allows probing zero temperature shape functions (Ch. 6) and thermal phase transition (Ch. 3) while the square regions lets us access universal terms associated with corners (Ch. 6).

way that the two-point correlation function can be used to classify and quantify characteristics of a system. For instance, the exponential decay of the two-point correlation function generally signals that one is in a gapped or trivial phase, whereas a crossover to algebraic decay can herald criticality or scale invariance. In this way, the two-point correlation function can be used to understand both phases and phase transitions. On the other hand, the entanglement entropy is capable of encoding much more information than the correlation function. Since the entanglement entropy requires a region to be defined, it is possible that there are contributions to it that depend on the topology of the region [7]. It is in part the ability to isolate and measure these topological contributions that make the entanglement entropy far more versatile than simple correlation functions.

In addition to the topological dependence, the constant part of the Rényi entanglement entropy (REE) of trivial regions can detect the existence of a classical thermally ordered phase. Fig. 1.1 shows the types of region shapes we use in the examination of the REE for these transitions. In this case, the scaling of the constant part of the REE changes sign through the transition [8]. This behaviour suggests that the REE is useful for both exotic non-Landau phases of matter as well as conventional phases and phase transitions.

Finally, the nature of criticality itself can be probed by examining the REE. In critical systems, the scaling of the REE has universal properties. For example, in one dimension

2

the scaling takes the form [9, 10]

$$S_\alpha = \frac{c}{3\eta}\Big(1 + \frac{1}{\alpha}\Big) \log\Big[\frac{\eta L}{\pi a} \sin \frac{\pi x}{L}\Big], \tag{1.1}$$

where $\eta = 1$ for periodic boundary conditions and $\eta = 2$ for open boundary conditions, α is the Rényi index, L is the total system size, a the lattice spacing and x the size of the region. The only important part of the above is that the coefficient c, the central charge, can be extracted by examining the finite size scaling of the model. This coefficient c defines the central charge of the conformal field theory, and in one dimension this completely quantifies the universality class that the system belongs to. In higher dimensions work is just beginning to quantify and examine this type of scaling (see Ch. 6), and hopefully these methods will become part of the standard procedure for characterizing criticality. The difficulty is that analytical and numerical techniques for one dimensional system are either impossible in higher dimensions or scale in such a way that they are impractical. By using a Monte Carlo approach we are able to generate accurate results in two dimensions over a large range of system sizes and parameters.

The Rényi entanglement entropy represents a natural extension of correlation functions, with the ability to capture a much wider range of behaviour than two-point correlation functions can. With the rapid progress of numerical techniques there will soon be a library of known results to compare against, each of them capturing different subtle behaviour and characteristics of many-body systems. Improvements in technique mean calculating these quantities is becoming increasingly feasible, and conceptual advances improve our understanding of the sources of universal behaviour. The first hints of experimental realization are being worked on [11, 12, 13], and as these methods gain popularity further implementations will be realized. From this point of view, the REE represents a new fundamental paradigm in viewing collective behaviour, and through it our understanding of both standard and exotic phases will be unified in a common language.

1.1 Rényi and von Neumann, entanglement and entropy

We have discussed the exciting possibilities that measurement of the Rényi entanglement entropy might allow, but here we will discuss what this quantity is, both mathematically and conceptually. A more rigorous discussion can be found in Ch. 3, but we will repeat the salient concepts here while discussing some simple examples that might lend some insight to this quantity.

Mathematically, the REE is defined as

$$S(A)_\alpha = \frac{1}{1-\alpha} \ln \mathrm{Tr}\left[\rho_A^\alpha\right], \tag{1.2}$$

where α is the Rényi index, in the $\lim \alpha \to 1$ we recover the von Neumann entropy, and ρ_A is the reduced density matrix for some region A (see Fig. 1.1). If region A were the full system, this would simply be called the Rényi entropy (or von Neumann entropy), but when A is some subsystem we add the word "entanglement". Entanglement here refers to the sum of all correlations and mixing between region A and the remaining system.

To understand this naming scheme, and more so what this quantity actually is, we can examine some exactly solvable systems. First we will compare using the Rényi entropy and the REE on pure quantum state and a classical state. We will take the example in which these two differ, two sites coupled by an antiferromagnetic Heisenberg term

$$H = \vec{S}_1 \cdot \vec{S}_2 \tag{1.3}$$

where $\vec{S}$ can be interpreted as either quantum operators or classical vectors. Assuming spin-$1/2$ in the quantum case, these two interpretations give us two groundstates

$$|\psi_{\mathrm{cl}}\rangle = |\circ\bullet\rangle \tag{1.4}$$

$$|\psi_{\mathrm{q}}\rangle = \frac{1}{\sqrt{2}}\left(|\circ\bullet\rangle - |\bullet\circ\rangle\right), \tag{1.5}$$

where we use the filled and empty circle to represent spin down or spin up in the z-direction or z-basis. As each of these are pure states, their total entropy is zero (Rényi and von Neumann) but tracing out the first spin we get something different

$$\mathrm{Tr}_{S_1} |\psi_{\mathrm{cl}}\rangle \langle\psi_{\mathrm{cl}}| = |\bullet\rangle \langle\bullet| \tag{1.6}$$

$$\mathrm{Tr}_{S_1} |\psi_{\mathrm{q}}\rangle \langle\psi_{\mathrm{q}}| = \frac{1}{2}\left(|\bullet\rangle \langle\bullet| + |\circ\rangle \langle\circ|\right) \tag{1.7}$$

calculating the Rényi or von Neumann entropy of these reduced density matrices gives us zero in the classical case and $\ln(2)$ in the quantum case. The quantum wavefunction is a quantum entangled state (here using entanglement in the Einstein-Podolsky-Rosen or Schrödinger sense), and the REE is able to detect this, justifying the addition of the word "entanglement" in Rényi entanglement entropy.

The ambiguity comes in when, instead of considering simply a classical state, we consider the full classical thermal density matrix. We will take it in the limit of zero temper-

4

ature, and in that case (for the same model as above) it can be written as

$$\rho = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{1.8}$$

where the rows (or columns) of this matrix define the different states $|\circ\circ\rangle$, $|\circ\bullet\rangle$, $|\bullet\circ\rangle$, $|\bullet\bullet\rangle$, respectively. The difference between this and the above classical case is that this density matrix has a non-zero total Rényi entropy, equal to $\ln(2)$. By tracing out the states of S_1 we get the same reduced density matrix as the quantum case above, and when calculating the REE we again get $\ln(2)$.

This construction shows the fundamental potential for multiple interpretations. When looking at wavefunctions, entanglement refers to the quantity inherent to a state, but in numerical simulation studied in this thesis we sample the full density matrix, whether at finite or zero temperature. In this language, entanglement simply describes a more generalized quantification of correlations, whether they are quantum or classical in origin. From this perspective, purely classical systems often have a measurable REE, and it is through examining this quantity that we can learn about these systems as well as quantum systems.

Throughout this thesis we will refer to the REE in regards to both quantum and classical systems, but this is in light of the definition above rather than the more common use of the word entanglement in quantum systems. Fundamentally, this thermal entanglement cannot be easily distinguished from the quantum kind, and most attempts to separate the two in thermal quantum systems requires the full analytic solution. In addition, the REE can be interesting at finite temperature, displaying universal behaviour of its own that is as rich as the zero temperature behaviour (see Ch. 5) of quantum systems. For this reason we use this broad definition.

1.2   Spin liquids

The original motivation for studying entanglement in condensed matter was as a diagnostic tool for characterizing phases that defy the typical Landau paradigm [6]. Specifically, by looking at the Rényi entanglement entropy of regions with different genus, something requiring at least two dimensions, we can define a part of the Rényi entropy that probes the topological character of a groundstate called the topological entanglement entropy [14, 15].

More importantly, these topological probes can act as a positive indicator of certain types of quantum spin liquids [7]. Quantum spin liquids [16] are notoriously difficult to detect, characterized by a lack of all conventional order and unconventional correlations. For example, in some recent work studying the existence of quantum spin liquids [17] an exhaustive elimination of conventional order parameters was used as one of the arguments for a quantum spin liquid along with the scaling of the single-particle gap. Scaling these quantities was extremely difficult, as demonstrated by the fact that later work suggested differing results [18], but only by using significantly larger computational resources on one of the worlds top supercomputers.

The topological signature of quantum spin liquids provides an alternate method of characterizing these phases numerically. By examining this topological entanglement entropy as a function of temperature we can examine when the model enters the constrained phase, and numerical results match theoretical expectations [7]. To measure the topological entanglement entropy requires a careful subtraction of various Rényi entropies [7]. The finite size effects and subtraction requires high accuracy and efficient Monte Carlo methods, a further motivation for exploration of more efficient or alternate algorithmic approaches.

1.3   Phase transitions

Another major reason for studying the entanglement entropy and mutual information, a quantity constructed by differences of Rényi entropies, is for its ability to detect phase transitions. The mutual information, I_α, is defined as (more thoroughly in Ch. 3)

$$I_\alpha(A, B) = S_\alpha(A) - S_\alpha(A|B) \tag{1.9}$$
$$I_\alpha(A, B) = S_\alpha(A) + S_\alpha(B) - S_\alpha(A \cup B), \tag{1.10}$$

a symmetrized quantity made by taking a difference of Rényi entanglement entropies. $S_\alpha(A|B)$ is the conditional entropy of region A on region B, while the region $A \cup B$ is the region defined as the union between region A and region B. For this definition to work, A and B must be distinct regions, or degrees of freedom, with no overlap. By looking near phase transitions, we find that when scaling the mutual information by boundary length the values cross at the critical temperature [8, 19]. This is quite promising, as measuring the mutual information in classical or quantum Monte Carlo only requires two things: the ability to measure energy and the construction of a modified simulation. Energy is almost always available, and fundamental to typical Monte Carlo sampling. The modified

simulation is easy to implement, with the fundamental implementation details discussed in Ch. 3.

This is mutual information is novel as not only does it not require measuring the order parameter of the transition, but furthermore it does not even require knowledge of what the order parameter is! One benefit of such a measure is immediately obvious; when examining whether a new system undergoes a transition, this method can be used without consideration of the type of ordering. Of course such a claim is very tantalizing, and it becomes tempting to examine difficult systems such as the quantum spin liquid or even classical spin glass using this method. That being said, without a comprehensive theory and practical demonstration, any result on a controversial system will itself be controversial, so the development of base cases and theoretical background is necessary before these interesting problems can be analyzed.

An analytically complete understanding of this crossing behaviour would be a large step forward, but for the moment we are left with numerical evidence of the crossings over a variety of transitions. It is the hope that further numerical work might constrain any theories looking at these crossings. In addition, by examining many different well understood systems we hope that the common element of the physics can be discerned.

Along side the detection of phase transitions, there are suggestions that the REE and mutual information can be used to detect the universality class of phase transitions and critical phases.

1.4 Universality

One of the major concepts motivating this work is the study and characterization of universality. In condensed matter systems universality is usually discussed in the context of continuous phase transitions. These transitions are characterized by a diverging correlation length. More interestingly, continuous phase transitions can usually be characterized by a set of values known as the critical exponents. These exponents describe how the system diverges near criticality, and most importantly can be estimated in theory and measured in experiment. The reason these exponents are so important is that they are robust– they typically only depend on general qualities, like the dimension of the system and the symmetry of the Hamiltonian and order parameter. What this means is that there are a large set of irrelevant perturbations, operators that, when added to the Hamiltonian, might change details like the critical temperature or a prefactor in the susceptibility, but leave many properties intact.

Conversely, the existence of universal quantities means that we have powerful tools for detecting what class a model is within. Many of the critical exponents are related mathematically, meaning that measuring a small set of them can give us the rest of them. For simple thermal phase transitions we can usually find the critical temperature by examining an order parameter either in experiment or numerically. By measuring the scaling of specific heat, susceptibility, or the order parameter we can determine some of the exponents. Thorough reviews [20] of critical phenomenon and renormalization-group theory show how extensive and broad reaching concepts of universality apply.

If we define the reduced temperature

$$t = \frac{T - T_c}{T_c},$$ (1.11)

where T_c is the critical temperature, then the usual scaling relations define the critical exponents as

$$\chi \propto t^{-\gamma}$$ (1.12)
$$\xi \propto t^{-\nu}$$ (1.13)
$$C \propto t^{-\alpha}$$ (1.14)
$$M \propto (-t)^{\beta},$$ (1.15)

where the quantities are the susceptibility, the correlation length, the specific heat and the order parameter magnitude, respectively. With these defined, we can compare these coefficients for two prototypical models, the three dimensional Ising model and the three dimensional Heisenberg model and their respective critical points.

Table 1.1 shows the difference for some of the critical exponents between two universality classes using a variety of theoretical approaches as well as experimental systems. The consistency between the different methods is remarkable, yet using these critical exponents to differentiate between universality classes is still a non-trivial task, especially in experiment where not all measurements are easily possible or part of a complex system must be isolated and measured in a background of other effects. Nevertheless, critical behaviour remains an important tool for classifying systems and predicting behaviour.

1.5 Entanglement and universality

The concept of using quantum phenomenon to make a computing device is very old [30], but its popularity has only grown with age, and in the last few years we have seen the release and benchmarking of devices [31]. This interest has focused on condensed matter for

Model	γ	ν	α	β
3D Ising (HTE)[21]	1.2373(2)	0.63012(16)	0.1096(5)	0.32653(10)
3D Ising (MC)[22]	1.2366(15)	0.6297(5)	0.1109(15)	0.3262(4)
3D Ising (FT)[23]	1.2403(8)	0.6303(8)	0.1091(24)	0.3257(5)
3D Ising (Exp)	1.14(5)[24]	0.62(3)[24]	1.10(3)[25]	0.341(2)[26]
3D Heisenberg (HTE)[27]	1.406(3)	0.716(2)		0.3710(13)
3D Heisenberg (MC)[28]	1.3957(22)	0.7113(11)		0.3691(1)
3D Heisenberg (FT)[23]	1.3882(10)	0.7062(7)		0.3655(5)
3D Heisenberg (Exp)[29]	1.39(5)			0.36(7)

Table 1.1: A comparison of the critical exponents for the 3D Ising universality class and the 3D Heisenberg universality class measured in high temperature expansion (HTE), Monte Carlo (MC), field theory (FT), and experiment (Exp). In the case of the 3D Ising model the liquid-vapour transition is used for the experimental data, while for the 3D Heisenberg model the system $La_{0.95}Ca_{0.05}MaO_3$ was measured. Some of the scaling results are generated using the relationships $\gamma = (2 - \eta)\nu$, $2 - \alpha = 3\nu$, $\beta = \nu(1 + \eta)/2$, and $\beta\delta = \beta + \gamma$.

two reasons: condensed matter systems might hold the answer to realizing the fundamental building blocks of a quantum computer [32, 33], and understanding some condensed matter systems might be one of the interesting problems that a quantum computer can tackle [34]. The topic of computation in the context of condensed matter systems has lead to an overlap of the classical computational and algorithms community with the condensed matter community, and with that overlap many of the concepts of information and communication theory have been examined more widely as they apply to condensed matter systems. Although entanglement is a fundamental concept in any quantum system, the concept of entanglement as a resource or fundamental characterization of a system was likely influenced by the information theory community.

Examining entanglement, there are well known examples in which it can be used as a universal measure. Specifically, in one dimensional systems bipartitioned into two regions the entanglement entropy either saturates to a constant or scales in some non-trivial way with the region size. For most systems the entanglement saturates, embodying the concept that degrees of freedom in one region of a system only give a fixed amount of information about the complementary region, regardless of the system size. This behaviour is typical for gapped systems, but can be violated for gapless systems with conformal invariance. That the entanglement entropy of a groundstate typically scales with the size of the interface

between a region and the remaining system is known as the area law [35, 36], and in one dimension the violations carry interesting universal behaviour. The area law is a statement that for the groundstate wavefunction of a local Hamiltonian the leading order term in the entanglement entropy of a region scales with the size of the boundary of that region.

Given a boundary of size ℓ, the entanglement is said to satisfy the area law if it takes the form

$$S = a\ell + \cdots \tag{1.16}$$

where a is some constant and "$\cdots$" represents terms proportional to L^{D-2} and smaller, where D is the dimension of the system. If the leading order term is larger than a term proportional to the bundary length, such as $\ell \log(\ell)$, then the system is said to violate the area law. In one and two dimensional gapped and gapless systems, shown in Fig. 1.2, we expect the following scaling

$$\text{1D gapped} \quad S = a \tag{1.17}$$

$$\text{1D gapless} \quad S = a \log \ell + \gamma(x/L) \tag{1.18}$$

$$\text{2D gapped} \quad S = a\ell + \cdots \tag{1.19}$$

$$\text{2D gapless} \quad S = a\ell + \gamma(x/L_x, L_y/L_x) + \cdots \tag{1.20}$$

where a is a constant, ℓ is the size of the boundary and *gamma* is a shape function which in general may exist for gapless functions. These forms derive naturally from an assumed type of gapped or gapless system, specifically those shown in Fig. 1.2, and we expect that any subtle differences will only occur in smaller subleading terms.

In one dimensional systems the coefficient of the area law violating term can be related to the central charge of the associated conformal field theory. If a model is known to be described by a conformal field theory, there are powerful predictive tools to calculate universal properties of a subsystem. With this concept in mind, the measurement and scaling of the entanglement entropy becomes an interesting diagnostic tool in its own right in the same way that conventional order parameters tell us something about a system.

The success of universality in one dimension naturally suggests that we attempt to examine what universal quantities might exist in higher dimensions. Fig. 1.2 shows a categorization of states in one and two dimensions, gapped and gapless, and the scaling we expect of the entropy in each case. In the one and two dimensions, it is the γ function, the subleading term, that contains universal properties[1]. This γ is unrelated to the critical exponent γ used to describe phase transitions, and is merely an unfortunate case of

---

[1]In one dimensions this term can be lumped with the leading term.

Gapless (scale invariant)   Gapped

1D

$$S = a\log(x) + \gamma(x/L)$$ $$S = a$$

2D

$$S = a\ell + \gamma(x/L_x, L_y/L_x) + \cdots$$ $$S = a\ell + \cdots$$
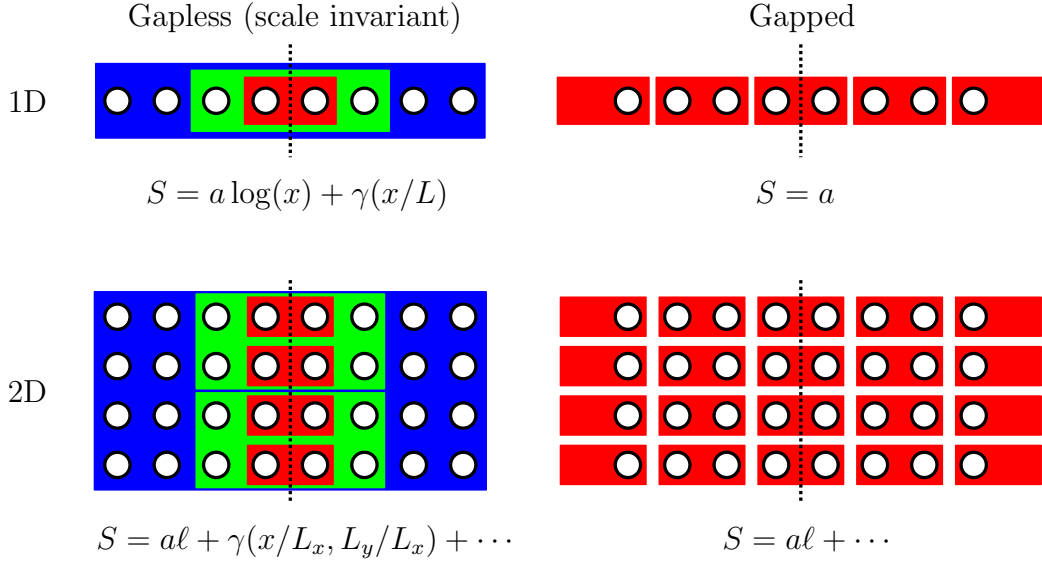
Figure 1.2: The general scaling form of the entanglement entropy in one and two dimensional gapped and gapless systems. Our prototypical gapped system is a valence bond solid, characterized by short range singlets. In this picture it is easy to see where the area law scaling comes from–the number of singlets cut is always proportional to the area between the two regions, in both one and two dimensions. The gapless systems assumes scale invariance, that is that contributions to the entanglement occur at every scale of the system, and each scale is related by some scaling factor, in this case by a factor of two. In one and two dimension the number of scales is proportional to $\log_2(x)$, where x is the size of the region. The one dimensional picture adds a contribution for every scale, leading to the logarithmic violation of the area law. In two dimensions the length of the interface also halves at every scale, and when summing all the contributions the sum does not violate the area law. In both gapless cases the subleading term γ may contain universal parts that help identify the universality class of the system.
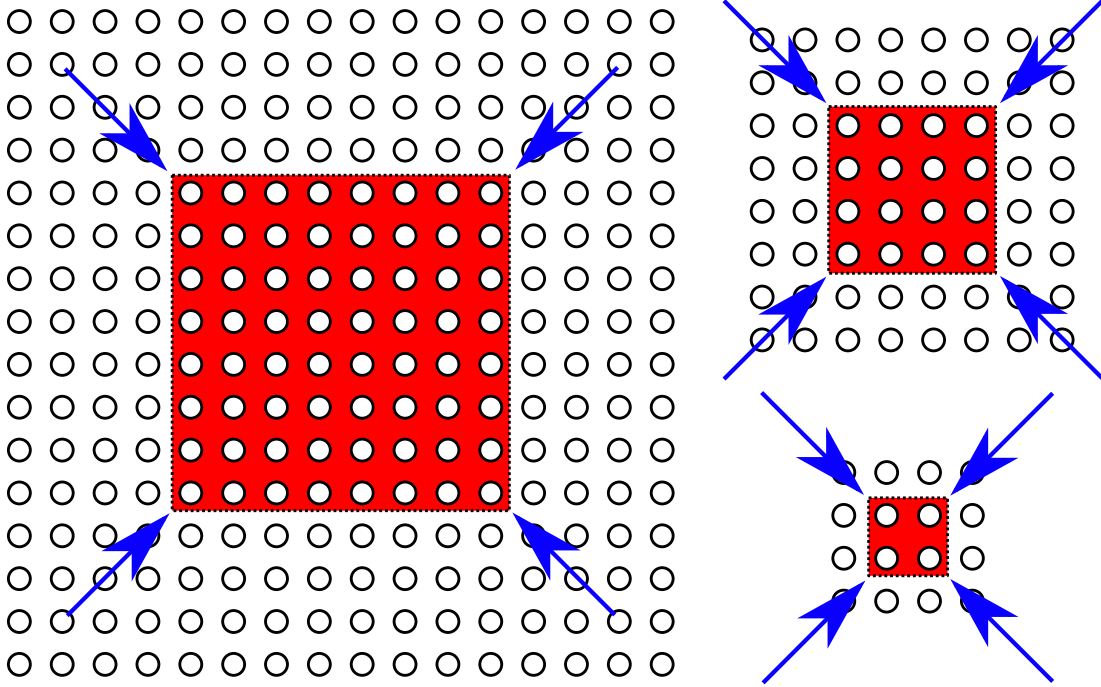
$$S = a\ell + \gamma(x/L_x, y/L_y) + n_c a(\pi/2)\log(\ell) + \cdots$$

Figure 1.3: The origin of the subleading logarithmic contribution to the area law in two dimensional systems with corners. Examining the square region, the contribution from the edges does not violate the area law, following the argument in Fig. 1.2. However, the number of corners at each scale is fixed (four in this case). If we assume a constant contribution from the corners at every scale, then the argument follows the one dimensional case, and a $\log_2(\ell)$ contribution can be realized from the corners. In this case it is not an area law violation, and this logarithmic correction is smaller than the leading area law term.

overloaded notation. Even though scale invariance leads to a log scaling of the entropy in one dimensional critical or gapless systems, the same argument does not lead to a log correction in two dimensions. The type of scale invariance we are assuming comes from a renormalization group perspective, and assumes that by scaling the system by some factor (typically denoted b) we return to a renormalized Hamiltonian with the same parameters. Thi renormalized Hamiltonian is calculated by integrating out some of the interactions to arrive at new effective degrees of freedom with new effective interactions, and if this new effective Hamiltonian looks the same as the original Hamiltonian, we are said to be at a fixed point. If we examine Fig. 1.2, in one dimensions this means that for a region of size x we can halve it $\log_2(x)$ times, where for $x = 8$ this means there are three scales at which we can count entanglement. Using the idea of scale invariance these scale contribute equally, and we immediately get that the entanglement scales as the logarithm of the region size. The difference in two dimensions is that the size of the boundary is also scaled at every step. Looking at the figure again, the first step contributes ℓ (the length of the boundary), while the second step contributes $\ell/2$, then $\ell/4$. Although there are $\log_2(\ell)$ steps in this procedure, the contribution from each step diminishes. Summing all the terms in the two dimensional case results in a contribution proportional to ℓ, and so even with scale invariance the area law remains unviolated in two dimensions.

The other geometry of interest is that of a square patch embedded in a two dimensional system, as shown in Fig. 1.3. The edges of the square contribute an area law whether the system is gapped or gapless, but the corners are special. With scale invariance, the number of corners at each scale is constant, and so this constant is multiplied by the number of steps in scaling procedure. This means that even though there is no log violation to the area law in two dimensions, corners can result in a subleading logarithmic term in the entanglement entropy. This examination of two dimensional systems is examined more closely in Ch. 6.

In two dimensional (periodic or open) systems we can examine the analogue of the one dimensional entanglement entropy by bipartitioning the system along a line, as shown in Fig. 1.2 and Fig. 1.1. For the Heisenberg model numerical results [37] show that such a geometry captures a subleading logarithmic term related to Goldstone modes [38], but finding a universal coefficient in two dimensions related to the entropy remains an open problem. Numerical schemes such as exact diagonalization always have access to the entanglement entropy, Rényi or von Neumann, and in methods such as density matrix renormalization group (DMRG) [39, 40, 41] the entanglement scaling is in some sense a condition for the procedure working efficiently. These methods both shine in one dimensional systems, with DMRG able to approach two dimensional system by studying long cylindrical systems and scaling the width, and make connections between theoretical results and numerical models

13

for one dimensional systems [42].

Methods for calculating the Rényi entropy and mutual information in quantum Mote Carlo [43] and classical Monte Carlo [8] are now available for a variety of systems. With the introduction of these methods studying larger two dimensional systems and approaching the thermodynamic limit in a more typical fashion is possible. For that reason, characterizations of these methods on known models is an important step to understand their accuracy and limitations in practice. These fundamental studies help push algorithmic development, such as the ratio trick (see Ch. 5), comparable to methods such as the Wolff algorithm [44] and the speedup it gives to systems like the Ising model. As these methods are still relatively new, adapting efficient techniques and examining their performance for this type of simulation is a current topic, and a large part of this thesis is devoted to exploring improvements of this type. It is our hope that these methods become ubiquitous as entropy and information measures become a fundamental measure in Monte Carlo methods.

1.6 Outline

This thesis is organized as follows. Ch. 2 discusses the quantum transverse field Ising model. It assumes no particular knowledge of Monte Carlo, but a working knowledge of classical Monte Carlo is helpful. This model will be our prototypical "quantum Monte Carlo" used to describe all of the algorithmic development. Many of the algorithmic developments are specialized for this model, but the concepts described are done so in a general way, so adaptation to other quantum Monte Carlo methods should be possible. We also discuss the interpretation of the phase transitions in the context of the $D + 1$ representation. Both the classical phase transition along the line of zero transverse field and the quantum phase transition are realized as percolation transitions in the $D + 1$ simulation cell, and each of these may realize universal physics.

Ch. 3 discusses the most fundamental methods for calculating the Rényi entanglement entropy and mutual information. It is discussed in the context of classical and quantum Monte Carlo, with examples to help keep concepts concrete. The concepts discussed here underpin all the more advanced algorithms. Here we demonstrate the ability of the mutual information to capture claisscal phase transitions without knowledge of an order parameter, as well as capturing the scaling relevant to the universality class of the phase transition.

Ch. 4 discussed an alternate method of calculating the Rényi entanglement entropy and mutual information without using thermodynamic integration. This method was

explored as a way of avoiding the temperature discretization necessary with the naive approach, and a minimal sample code is provided in the appendix. We are successful in reproducing the results from the integration approach in Ch. 3, but further advancements have made this approach suboptimal in some cases. Improvements are suggested should further improvements of this algorithm be pursued. The Wang-Landau approach is also used to detect thermal phase transitions, and motivates an alternate derivation of concepts like the ratio trick, crucial to the next two chapters.

Ch. 5 details the "ratio trick" for finite temperature classical and quantum Monte Carlo. This is probably the most fundamentally important algorithmic development for practically measuring the Rényi entanglement entropy. Several methods are detailed along with functional code for classical models in the appendix. Improving on the methods in this section has the most potential for improving the speed and accuracy of simulation. Using this method we explore the thermal Rényi entanglement entropy and show universal collapse for both finite size and finite temperature for the $c = 0.5$ universality class of the transverse field Ising model in one dimension. This matches predictions from analytical work, and shows how thermal results can be used to capture universal characteristics.

Ch. 6 discussed the modifications of the algorithm needed for the zero temperature projector quantum Monte Carlo. This is useful in the study of systems at a zero temperature quantum critical point, where many of the universal concepts discussed above take place. The ratio trick in Ch. 5 naturally extends to the zero temperature case with only minor modification. Accessing zero temperature results in two dimensions in an unbiased way allows us the first glimpse at universal scaling functions for two dimensional systems, and comparison to a quantum Lifshitz free field theory show remarkable agreement when compared to simply extended one dimensional results. We also extract the corner term associated with a subleading logarithmic contribution and compare this to results from non-interacting theories and other numerical approaches.

Finally, Ch. 7 concludes the thesis while discussing open problems that will hopefully be addressed soon. In addition, avenues for further algorithmic progress are discussed in a broader context.

1.7   Terminology

With any technical work there is necessary jargon, and although terms will be introduced as they are needed, this section will serve as an index of the most often used terms and acronyms. In addition, the overlap of different fields, condensed matter and quantum

information, for instance, mean that some terms may have slightly different standard definitions. Here we will outline what is meant when referring to each of these terms (in the context of this thesis).

Mutual information (MI) Described in Eq. (3.5), this is a quantity typically used when examining thermal systems, constructed by taking a sum and difference of different Rényi entanglement entropies. It is constructed in such a way that bulk contributions from each region are cancelled, and the remaining contribution only quantifies the information each region has about the other. This quantity also has an index α that specifies which Rényi entanglement entropy is being used to construct the mutual information.

Quantum Monte Carlo (QMC) When discussing techniques used in this thesis, it almost always refers to stochastic series expansion methods, both $T > 0$ methods and $T = 0$ methods. Outside of this document, there are a wide range of quantum Monte Carlo methods, including variational wave function, diffusion, world line, and many others.

Rényi entanglement entropy (REE) Described in detail from Eq. (3.1) and earlier in Sec. 1.1, the Rényi entanglement entropy describes the amount of (Rényi) entropy contained in some region, typically of a subsystem. Despite being called entanglement entropy, we use this term to describe the both the quantum and thermal contributions. This term can be calculated in classical systems in which there is no quantum contribution, as the definition only requires the existence of a density matrix. The index α is typically used to denote the Rényi index, and in the case of $\alpha = 1$ the Rényi entanglement entropy is equal to the von Neumann entanglement entropy.

Simulation cell The representation of the classical or quantum simulation, and more specifically the representation that is convenient for numerical simulation. In a classical simulation, the simulation cell consists of a particular spin configuration. This spin configuration is updated by Monte Carlo moves and evolves with time. In quantum Monte Carlo the simulation cell consists of a configuration as a set of operators, and this entire object is updated.

Stochastic series expansion (SSE) A type of quantum Monte Carlo, described from Eq. (2.7) onwards [45, 46, 47], in which one simulates a D-dimensional Hamiltonian as a $D + 1$-dimensional system. Able to simulate a wide range of systems, typically scaling with the size of the system and inverse temperature. Still vulnerable to the

sign problem, but has no problem with geometric frustration in the diagonal terms of the Hamiltonian.

Transverse field Ising model (TFIM) The most general form written in Eq. (2.4), the transverse field Ising model can be thought of as the quantum analogue to the classical Ising model. Despite it's simplicity, there are aspects of this model that are still unknown, such as the exact (geometry dependent) scaling of the entanglement entropy in two dimensions.

Chapter 2

Quantum Monte Carlo for the transverse field Ising model

In this section we will generally introduce quantum Monte Carlo and then delve into the algorithm used for most of the subsequent chapters–the stochastic series expansion method, or SSE. We will discuss the mathematical framework and practical implementation for the standard algorithm, with specific extensions or modifications being discussed in Ch. 5 and 6.

In addition we will discuss some insights into the thermal and quantum phase transitions. In the language of the simulation these points manifest as percolation transitions in the simulation cell, and there are suggestions that universality can be measured directly in this language as well.

2.1   Introduction to quantum Monte Carlo

Quantum Monte Carlo is a broadly used term describing a wide array of methods, but they share some common elements. They all make use of a random process (typically pseudo-random number generation [48]) to help them sample something that is like a sum or do some type of non-exact optimization. For instance, using a Monte Carlo method to do high dimensional integration is a common application [49]. For the sums, if we have an equation of the form

$$E = \frac{\sum_i W_i E_i}{\sum_i W_i}, \tag{2.1}$$

it is possible iterate over all terms i and do the direct sum to get the answer. Monte Carlo techniques take the approach that spend more time examining contributions to the sum where the value of W_i is large, and sample the terms i with probability proportional to W_i. In the case of optimization, Monte Carlo is a way of biasing a random walk through an exponentially large configuration space to try to find a minimum while satisfying two important Monte Carlo concepts, ergodicity and balance. This biasing is done in an identical way to the sampling of the sums mentioned previously. For a typical thermal system, these weights would be the Boltzmann weights, or $W_i = e^{-\beta E_i}$. A more detailed reading on the concepts of Monte Carlo can be found be looking into *Markov chain Monte Carlo* [50, 51] and the *Metropolis-Hastings algorithm* [52], both of which underpin much of current Monte Carlo philosophy.

The concepts of ergodicity and balance are fundamental to Monte Carlo methods, and any successful simulation implements these in some form or another. Ergodicity refers for the property that a simulation is able to reach all possible configurations. For an Ising model, this would mean the simulation should be able to reach all possible spin configuration via some set of updates. If we imagine any particular spin configuration, then we can explicitly calculate the probability of a set of moves that take us to any other state, and there is a non-zero probability that such a chain of moves will occur. In this way single spin updates are ergodic for the Ising model, although they may not necessarily be efficient. On the topic of ergodicity, it is sometimes said that a simulation loses ergodicity at low temperature. When this is said it means that there are states that are relevant to sample that typically will not be sampled, for instance the two polarized states of the Ising model at low temperature. For a large enough system single spin updates will choose a particular polarization and typically not sample the opposite polarization state, even for extremely long simulations. In this sense, ergodicity refers to the ability of a simulation to reach all important configurations in a reasonable amount of time, although what is reasonable changes depending on the application and resources available.

The second necessary component of a simulation is that is satisfies balance. Mathematically, balance is the statement about the transition probabilities, requiring that

$$W_i = \sum_j W_j P_{j \to i}, \qquad (2.2)$$

where W_i is the weight of a state (such as the Boltzmann weight), and $P_{j \to i}$ is the probability of transitioning from state j to state i, and includes the case of $j = i$. Generally speaking, the weights of a system are inherent to it, while the transition probabilities are something the person designing the Monte Carlo can fine tune. To put it in other words,

the probability of transitioning to any state, for a long enough Markov chain, is proportional to the weight of that state. Very typically, the stronger constraint, detailed balance, is satisfied instead of just balance. Detailed balance is formulated as

$$W_i P_{i \to j} = W_j P_{j \to i}, \tag{2.3}$$

and automatically satisfies balance by construction. It can often be easier to implement due to the symmetry of the constraint, but there are cases where implementing balance can be more efficient than using detailed balance.

What distinguishes quantum Monte Carlo from classical Monte Carlo at a practical level, in a non-rigorous definition, is that quantum Monte Carlo addresses the class of Hamiltonians for which the diagonalizing basis is not known or not convenient to work in. In the transverse field Ising model, we have a Hamiltonian of the form

$$H = \sum_{\langle ij \rangle} J_{ij} \sigma_i^z \sigma_j^z + \sum_i h_i \sigma_i^x, \tag{2.4}$$

with $\langle \cdots \rangle$ representing the sum over all nearest neighbours of the lattice. For simplicity, we will generally examine the case of uniform coupling, $J_{ij} = J$, and uniform transverse field, $h_i = h$. In the limit of $h = 0$ this model reduces to the classical Ising model, and the above Hamiltonian is diagonal, that is

$$\langle \psi_i | H | \psi_j \rangle = 0 \text{ if } i \neq j, \tag{2.5}$$

in the spin basis states. For a diagonal Hamiltonian one can use the above mentioned methods easily since the weight of a state assuming a Boltzmann distribution, $W_i = e^{-\beta H}$, is well defined and easy to calculate. When the Hamiltonian has off-diagonal elements, formally the weights must be written as

$$\langle \Psi_i | e^{-\beta H} | \Psi_j \rangle = e^{-\beta \lambda_i} \delta_{ij}, \tag{2.6}$$

where Ψ_i are eigenvectors of the Hamiltonian with eigenvalue λ_i. Due to the fact that for an arbitrary Hamiltonian the eigenbasis is not typically known, and finding them is tantamount to exactly solving the problem, this description does not assist us in practice. Finding the correct eigenbasis is a hard problem, and typically requires techniques like exact diagonalization or density matrix renormalization group [39, 40]. The SSE allows us to sample a large class of such Hamiltonians, although there is a class of problems for which there is a sign problem and quantum Monte Carlo does not work. The sign problem is often encountered, but not unsurmountable, and there are cases where algorithmic

advanced allow one to access new problems that were previously thought to contain sign problems [53].

The description that follows can also be found in the seminal work by Sandvik [45, 46, 47], and is repeated here to focus on core aspects while introducing notation that is necessary for later chapters. Fundamentally, the SSE is based on the expansion of a general Hamiltonian in the form

$$e^{-\beta H} = \sum_n \frac{(-\beta)^n}{n!} H^n, \tag{2.7}$$

which is simply the Taylor expansion of the exponential. This alone is not enough to build a simulation on, to do that we have to analyze Eq. (2.1) and see that the proper quantum analogue takes the form

$$\langle E \rangle = \frac{\text{Tr}\left[E e^{-\beta H}\right]}{Z} \tag{2.8}$$

where $Z = \text{Tr}\left[e^{-\beta H}\right]$ is the partition function. Using this idea, the object that "represents" a contribution to the trace of the quantum partition function takes the form

$$Z = \text{Tr}\left[\sum_n \frac{(-\beta)^n}{n!} H^n\right] \tag{2.9}$$

$$= \sum_\alpha \sum_n \frac{(-\beta)^n}{n!} \langle \alpha | H^n | \alpha \rangle \tag{2.10}$$

$$= \sum_n \frac{(-\beta)^n}{n!} \sum_\alpha \prod_{i=0}^{n-1} \langle \alpha_i | H | \alpha_{i+1} \rangle, \tag{2.11}$$

where $|\alpha\rangle$ represents any complete set of states, and in the final line we insert $n-1$ resolutions of the identity, $\sum_\alpha |\alpha\rangle \langle \alpha|$, generating the states $|\alpha_i\rangle$. There is an important final restriction, that $|\alpha_0\rangle = |\alpha_n\rangle$, thus ensuring that the trace remains satisfied.

In contrast to classical Monte Carlo where the states sampled are simply states from a convenient basis, $|\alpha\rangle$[1], in the quantum Monte Carlo we can represent a "state" of the simulation by the ordered set of states from the same convenient basis $|\alpha_0\rangle, |\alpha_1\rangle, \ldots, |\alpha_{n-1}\rangle$. If one imagines each of these configurations as being stacked upon each other, we can see one way in which a D-dimensional quantum system maps to a $D+1$-dimensional classical

---

[1]Classical states are a subset of quantum states, and here by convenient basis we mean the classical one.

system, although the system it maps to is unusual in the new expansion direction (see Fig. 2.1).

The need for such a representation in the case of a quantum Hamiltonian can be seen by looking at a trivial case, a two-site Hamiltonian of the form

$$H = -\sigma_1^+ \sigma_2^- + \text{h.c.}, \tag{2.12}$$

with σ^+ and σ^- are the raising and lowering operators, and h.c. is the Hermitian conjugate of the first term (a shorthand used in the definition of many Hamiltonians), and 1 and 2 label the two sites of the system. This Hamiltonian can also be written as

$$H = \frac{1}{2}\left(-\sigma_1^x \sigma_2^x - \sigma_1^y \sigma_2^y\right), \tag{2.13}$$

where it can be interpreted as the ferromagnetic quantum XY model for two spins. The true groundstate for such a system is a wavefunction of the form

$$\frac{1}{\sqrt{2}}\left(|\circ\bullet\rangle + |\bullet\circ\rangle\right), \tag{2.14}$$

where $|\bullet\circ\rangle$ represents a spin up state (in the σ^z basis) on the first spin and a spin down state on the second spin. The groundstate of this Hamiltonian is clearly not a state in the classical basis. A well constructed SSE would find that for $|\alpha_0\rangle, |\alpha_1\rangle, \ldots$ we would get something of the form above, that is

$$|\alpha_0\rangle, |\alpha_1\rangle, \ldots = |\bullet\circ\rangle, |\circ\bullet\rangle, \ldots, \tag{2.15}$$

depending on the specific implementation. Different representations might also allow consecutive kets of the same state, but overall we expect any sufficiently low temperature state to have an equal fraction of the two possible Kets in the expansion

If the Hamiltonian had a positive leading coefficient rather than a negative one, naively it would have a *sign problem*, meaning that it would not be possible to simulate it using SSE. In this particular case the sign problem can be resolved since the system is bipartite, that is to say one can bipartition the system in to two species of sites such that the two species are only connected to species of the other type. Due to the bipartite nature, we are guaranteed to always require an even number of off-diagonal Hamiltonian elements in any expression of the trace, Eq. (2.11), implying that we will have an even number of negative terms, $(-H)$, ensuring that all non-zero terms contributing to that sum are always strictly positive.

22

One can see how this might go wrong by examining an SSE for the same Hamiltonian above, except for a three site triangular system and antiferromagnetic coupling.

$$H = \sigma_1^+ \sigma_2^- + \sigma_2^+ \sigma_3^- + \sigma_3^+ \sigma_1^- + \text{h.c.}, \tag{2.16}$$

In this case, a valid simulation cell could be one of the form

$$|\alpha_i\rangle = |\bullet \circ \circ\rangle, |\circ \bullet \circ\rangle, |\circ \circ \bullet\rangle, |\bullet \circ \circ\rangle. \tag{2.17}$$

In this case the weight of this configuration is $\beta^3(-1)^3/3!$ (including the prefactor), and is negative. The problem of negative weights is that the detailed balance transition probability does not make sense for negative weights. Although there are ways to account for this sign problem, in the naive implementation it is insurmountable. This also makes it slightly more clear why bipartite lattices do not have this type of sign problem, as in these systems such configurations do not occur.

2.2 Stochastic series expansion implementation of the transverse field Ising model

In this section we will describe the detailed algorithm for the transverse field Ising model. Most of the details for the base algorithm are discussed in a paper by Sandvik [47].

We will attempt to explain the stochastic series method in a contained way, relying only on basic concepts from classical Monte Carlo approaches. The first modification that is necessary in quantum Monte Carlo is that, for reasons evident later, we cannot directly simulate any Hamiltonian, but typically sample from a Hamiltonian that is modified by the addition of constants. For the transverse field Ising model, the modified Hamiltonian that we sample is of the form

$$H = -\sum_{\langle ij \rangle} J_{ij}(\sigma_i^z \sigma_j^z + \text{sgn}(J_{ij})) - \sum_i h_i(\sigma^+ + \sigma^- + \text{sgn}(h_i)), \tag{2.18}$$

where again, we typically take $J_{ij} = J$ and $h_i = h$, and in the case where $J > 0$ and $h > 0$, the sign of both terms is positive, and the additive constant to the Hamiltonian is equal to $-N_B J - N_s h$, where N_B is the number of bonds and N_s is the number of spins.

The reason for the modification of the Hamiltonian in this way is convenient for the algorithm, as we will see later, but an initial motivation can be seen as the following. In

this form each individual term of the Hamiltonian takes only one non-zero value. Looking at a single bond, the term $H_J = J_{ij}(\sigma_i^z \sigma_j^z + \mathrm{sgn}(J_{ij}))$ can only take two terms, $-2\,|J_{ij}|$ when the bond is satisfied–aligned if J_{ij} is negative, anti-aligned if J_{ij} is positive–and zero in the other case. If we draw this relationship (assuming a positive J) out it takes the form

$$\langle \circ\circ|\, H_J \,|\circ\circ\rangle = \langle \bullet\bullet|\, H_J \,|\bullet\bullet\rangle = -\,2J \tag{2.19}$$

$$\langle \circ\bullet|\, H_J \,|\circ\bullet\rangle = \langle \bullet\circ|\, H_J \,|\bullet\circ\rangle = \quad 0, \tag{2.20}$$

where we have omitted all off-diagonal terms, which have value zero for H_J.

Looking at a single spin term $H_h = h_i(\sigma^+ + \sigma^- + \mathrm{sgn}(h_i))$ is even more straightforward. Since the direction of the transverse field is irrelevant to the physics of the problem, the physics of the system is symmetric under $h \to -h$, we will assume h is positive[2]. In this case, in our convenient basis, the only value the term takes for any bra and ket is $-h$. Writing this in the same form as above we get

$$\langle \circ|\, H_h \,|\circ\rangle = \langle \bullet|\, H_h \,|\bullet\rangle = -\,h \tag{2.21}$$

$$\langle \circ|\, H_h \,|\bullet\rangle = \langle \bullet|\, H_h \,|\circ\rangle = -\,h. \tag{2.22}$$

With the matrix elements of the Hamiltonian well defined, we can discuss the representation of the simulation cell and how to construct the updates needed to simulate it effectively. In a classical system the simulation cell consists of a classical configuration, that is a specification of all the spins of the system. As mentioned earlier, the quantum simulation cell consists of a $D+1$ set of ordered classical configurations. In practice, rather than storing the set of all configurations $|\alpha_i\rangle$, we only store $|\alpha_0\rangle$ and the specific operators (shown below) that connect the $|\alpha_i\rangle$. In this way we can reconstruct the states if they are needed, but minimize the amount of information that needs to be stored to run the simulation. We label the Hamiltonian elements in the following way

$$\langle \bullet_i\bullet_j|\, H_{0,ij} \,|\bullet_i\bullet_j\rangle = -\,2J \tag{2.23}$$

$$\langle \circ_i\circ_j|\, H_{0,ij} \,|\circ_i\circ_j\rangle = -\,2J \tag{2.24}$$

$$\langle \bullet_i|\, H_{0,i} \,|\bullet_i\rangle = -\,h \tag{2.25}$$

$$\langle \circ_i|\, H_{0,i} \,|\circ_i\rangle = -\,h \tag{2.26}$$

$$\langle \bullet_i|\, H_{1,i} \,|\circ_i\rangle = -\,h \tag{2.27}$$

$$\langle \circ_i|\, H_{1,i} \,|\bullet_i\rangle = -\,h \tag{2.28}$$

$$H_{-1,0} \equiv \mathbb{1} \tag{2.29}$$

---

[2]Taking h negative in Eq. (2.18) natively gives us a sign problem, but it is related by symmetry to the opposite sign of h, and can be safely ignored.

which is essentially a site specific relabelling of the terms above. We add the operator $H_{-1,0}$ for convenience of the algorithm when discussing updates later. Using this above notation we can re-write the Hamiltonian as

$$H = \sum_{ij} H_{0,ij} + \sum_{i} \left(H_{0,i} + H_{1,i} \right). \tag{2.30}$$

This set of relations is an arbitrary labelling scheme for convenience, and defines all the non-zero weights that must be sampled in the simulation. These weights are necessary for the diagonal update, discussed below. With this specific labelling, if we specify a state and an ordered string of n operators, we know precisely the state and product of weights we have at the end, assuming we haven't applied any operators that annihilate the state. In construction of simple updates it will be convenient to treat the diagonal and off-diagonal terms in the Hamiltonian on equal footing. Between any two elements of the in the set of ordered states we will only ever insert a single term of the Hamiltonian. In addition, a configuration with multiple diagonal operators between two elements $\langle \alpha_i |$ and $| \alpha_{i+1} \rangle$ can be transformed in to a normal configuration by additional resolutions of the identity to restore exactly one operator per step in the expansion direction.

To give a trivial example of the above, let us consider the initial state $\langle \bullet_0 \, \bullet_1 \, \circ_2 \bullet_3 |$. Given the ordered set of operators $H_{0,0}, H_{0,01}, H_{1,2}, H_{1,0}$ (applying them from left to right on the state) we would have a weight of $h^3 J$ and a new state $\langle \circ_0 \, \bullet_1 \, \bullet_2 \bullet_3 |$. Such a string of operators would not represent a valid simulation cell, as the initial state is not identical to the final state, so in that language it would have weight zero due to the boundary condition being unsatisfied. The constraint on the states, that $| \alpha_0 \rangle = | \alpha_n \rangle$, can be viewed as enforcing periodic boundary conditions if we imagine the states being arranged in a circle that connects back to itself after n links. Keeping in mind that we only wish to sample simulation cells with non-zero weight in respect to the trace, we will discuss the updates that allow us to generate states and strings of operators in such a way that this boundary condition is automatically satisfied.

The two updates we need to be ergodic in the quantum system are referred to as the diagonal update and off-diagonal update. In the case of the transverse field Ising model the off-diagonal update is referred to as the deterministic loop algorithm, but other systems, such as the XXZ model, generally require different types of off-diagonal updates such as the directed loop algorithm [45, 46].

Let us imagine we already have a valid simulation cell (a set of identity operators with an arbitrary $| \alpha_0 \rangle$ is a trivial example) that we will assume has the ordered set of M Hamiltonian elements, S_M. If our original list of operators has less than M elements we

can imagine appending the operator list with $H_{-1,0}$ until we have a list of M operators. If we define the number of non-identity operators in S_M as n, we are allowed to choose an $M > n$ such that there are always some $H_{-1,0}$ elements in the list. Using this idea of a fixed list length, we can rewrite Eq. (2.11) as

$$Z = \sum_{\alpha} \sum_{S_M} \frac{(-\beta)^n (M-n)!}{M!} \prod_{i=0}^{M_1} \langle \alpha_i | H_i | \alpha_{i+1} \rangle, \tag{2.31}$$

where n is now the number of non-identity operators in S_M and H_i specifies the type of operator at each step in the expansion (see a list of operators types on page 24). The factor of $M!$ and $(M-n)!$ are a result of combinatorics, representing the number of ways in which the $(M-n)$ identity operators can be inserted into a list of n operators. The sum over S_M samples all valid choices[3] of H_i (see Eq. (2.23) and those below it) for a given choice of $|\alpha\rangle$. S_M is a somewhat complex "dummy index", including the sampling over initial states $|\alpha_0\rangle$ as well. Although the notation is a bit obtuse, in practice the implementation is straightforward. One more concept we must define before moving forward is the weights of the Hamiltonian, which is the sum of all *diagonal* terms in the full Hamiltonian matrix. We will define the total weight as $\sum_i W_i$ where i goes over all possible indices for the Hamiltonian elements $H_{0,i}$ and $H_{0,ij}$.

2.2.1 Diagonal update for the transverse field Ising Model

The diagonal update allows us to change all of the diagonal Hamiltonian elements in the list S_M, generating a new simulation cell. Given this ordered list S_M the diagonal update proceeds as follows, starting at the first operator in the list S_M and with knowledge of the state $|\alpha_0\rangle$

1. Examine the current operator.

   (a) If the operator is off-diagonal, $H_{1,i}$, update our knowledge of the current state $|\alpha_{i+1}\rangle = H_{1,i} |\alpha_i\rangle$.

   (b) If the operator is diagonal, $H_{0,ij}$ or $H_{0,i}$, replace the current operator with operator $H_{-1,0}$ with probability [47]

$$P_{\text{rem}} = \frac{M - n + 1}{\beta \sum_i W_i}. \tag{2.32}$$

---

[3]That satisfy the boundary conditions.

(c) If the operator is the identity, $H_{-1,0}$, choose the operator $H_{0,a}$ with probability $W_a / \sum_i W_i$, where $W_a = 2J$ for the two site operators and $W_a = h$ for the single site diagonal operators. One must keep in mind that there are typically a different number of two site and single site operators depending on the type and dimension of the lattice one is using. If a two site operator is chosen, ensure that the state on i and j (of the current $|\alpha_i\rangle$, assuming a ferromagnetic J) matches– if not, do not add the operator and go to the next step. If the operator is compatible with the configuration, replace the current operator with $H_{0,a}$ with probability [47]

$$P_{\text{add}} = \frac{\beta \sum_i W_i}{M - n}. \tag{2.33}$$

2. Update the number of non-identity operators in the list n.

3. Move to the next operator in the list until all operators have been visited.

By using this procedure we are able to modify all of the diagonal and identity operators in S_M. As a final step, the value M is typically[4] increased if $n/M > 0.75$ at the end of the diagonal update, ensuring the finite size of the list is not biasing the simulation. If M were too small, then this would restrict the growth of operators n as there would be no free room for inserting operators, and cause the simulation to give back incorrect results. As a result, the value M can change during a simulation, but this fact does not affect the collection of statistics or change anything fundamental when it occurs. This update alone is trivially non-ergodic, as the state $|\alpha_0\rangle$ (and in fact all $|\alpha_i\rangle$) is left unmodified, although the weight (see Eq. (2.11)) is potentially changed. To elaborate, changing the diagonal operators does not change what $|\alpha_i\rangle$ is compared to before the diagonal update, but $|\alpha_i\rangle$ may be different than $|\alpha_0\rangle$ due to off-diagonal operators present in S_M, and the diagonal update is not capable of modifying these operators. The deterministic loop update on the other hand is designed in such a way that it samples the states in S_M without changing the weights. To discuss this algorithm we have to examine the topological nature of the simulation cell, defined by the sites of the lattice that the operator list S_M interacts with.

2.2.2 Off-diagonal update for the transverse field Ising Model

In addition to the diagonal update, that is the update that modifies the diagonal operators, an off-diagonal update is needed. Fundamentally, the off-diagonal update transforms

---

[4]By convention, see Fig. 4.7 for the convergence as a function of "filling" of the list.

diagonal operators into off-diagonal ones (and vice-versa). This procedure, along with the diagonal update, allow for sampling of all possible contributing configurations of the operator list.

Fig. 2.1 shows a simulation cell, the operator expansion, and the modification of the state $|\alpha_i\rangle$ as a function of the operator expansion. In this $D + 1$ simulation cell, a cluster can be constructed following a simple set of rules. We must first define some terminology: a Hamiltonian element connecting n spins has $2n$ legs. These legs connect to the next (or previous) operator in the expansion that operators in the same spin. Fig. 2.2 shows how these legs are defined and how a cluster it built using the algorithm below. To enumerate all the clusters we use the following procedure.

1. Find a leg of any operator not currently in a cluster, this will define a new cluster.

2. If that operator is two spin operator, add all other legs of the operator to the cluster.

3. For every leg in the cluster, find the corresponding leg on another operator[5] it connects to, and add those legs to the cluster.

4. For every leg added to the cluster, check the type of operator it belongs to and add legs if there are any two spin operators.

5. Repeat the above two steps until no more legs are added to the cluster.

The above algorithm defines a cluster, and as such two properties emerge: all legs on a two spin operator are a part of the same cluster and the two legs of a single spin cluster may be a part of different clusters. The leg that an operator connects to is the next (or previous, depending if the leg is directed up or down the set of states $|\alpha_i\rangle$) leg that connects to the same spin; Fig. 2.1 shows this where the bottom left leg of $H_{0,23}$ connects to the upper leg of $H_{1,2}$ three layers below it. In this particular realization of the TFIM, a cluster also defines a set of spins of the same orientation in the simulation cell, but this is not a general characteristic, and changing to the antiferromagnetic TFIM will change this property.

The off-diagonal update is now trivial–for every cluster, flip the orientation of all spins in that cluster, and modify the attached operators appropriately, with probability $1/2$. The only operators that may be modified (or equivalently, the only operators a cluster can end on) are those of type $H_{0,i}$ and $H_{1,i}$, which can be transformed into each other by the process of flipping these clusters. If we look at how $H_{0,i}$ and $H_{1,i}$ interact with the $D + 1$

---

[5]Technically a leg can connect to another leg on the same operator in the case where the number of operators is small, i.e. at high temperature. See Eq. (2.32),(2.33).
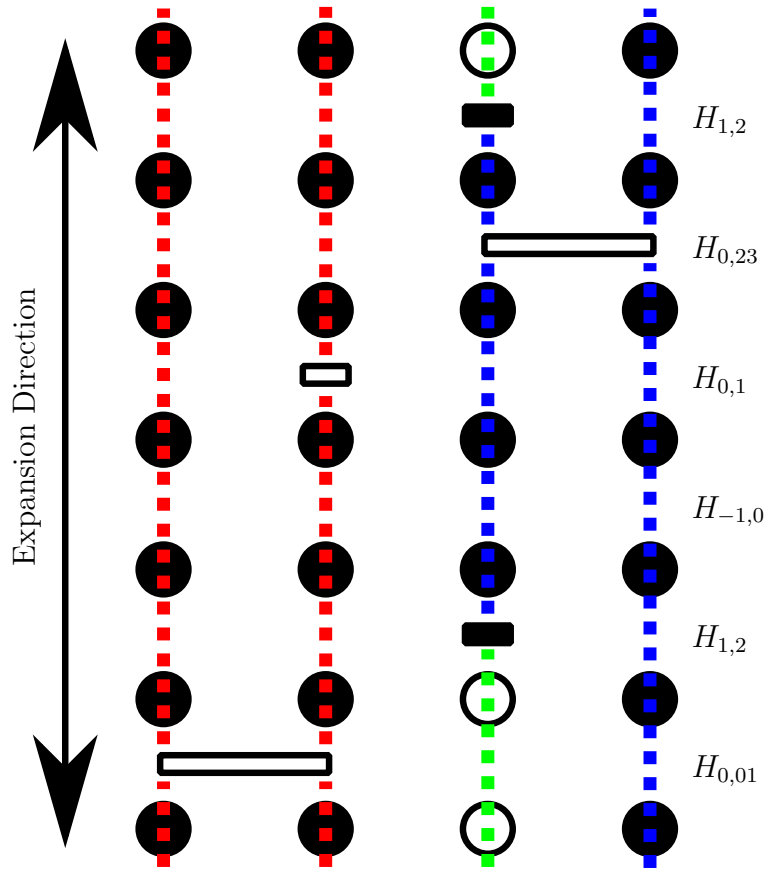
Figure 2.1: An example of a simulation cell for the transverse field Ising model. The state $|\alpha_i\rangle$ is represented by the spin up (filled) and spin down (empty) circles at each vertical layer. The block between layers represent operators, diagonal (empty) and off-diagonal (filled). Each unique cluster is marked with a different color, of which there are three for this particular configuration.
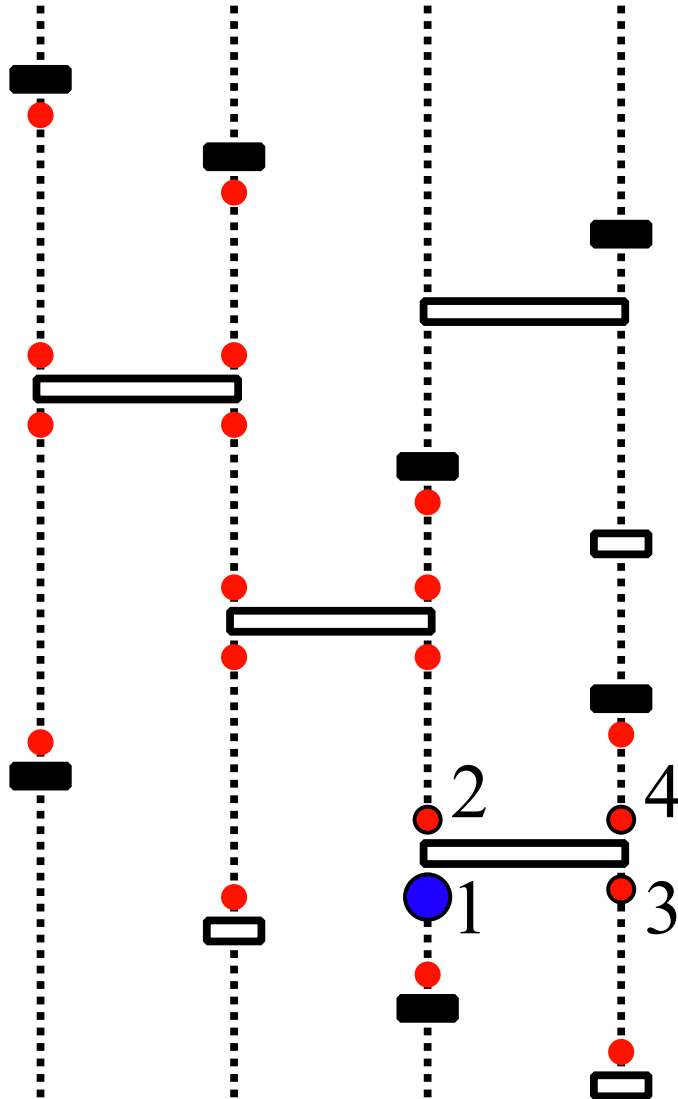
Figure 2.2: The construction of clusters for the off diagonal update. The four numbers show the four legs of a two site operator. Starting at the blue leg, we immediately add the other three legs because the current operator involves two spins. From there all the other red circles represent legs that are sequentially added to the cluster. Sometimes the new legs end the cluster on single site operators and when the legs are a part of two site operators they cause the cluster to branch and continue.

expansion, it becomes clear that if the spin above or below the operator in the expansion flips, we can change the operators into one another to locally satisfy the configuration. Flipping entire clusters ensures that the boundary conditions in the expansion direction remain satisfied, and the definition of the clusters is in such a way that both states have precisely equal weight, and it is for this reason that we can flip each cluster with probability one half. It should be noted that since we keep track of the spins at $|\alpha_0\rangle$, if any cluster that crosses that layer flips, we must record that change.

With these two updates we can sample states from Eq. (2.11) in an ergodic and efficient way. In this framework we will discuss the typical observables in SSE and how they are measured, although discussion of entropy measures will be left to Ch. 3.

The most fundamental observable in most Monte Carlo algorithms is the energy of the system. For observables that are composed of the Hamiltonian or diagonal within the working basis, measurement is easy. In the TFIM, this means that the energy, z-axis polarization and x-axis polarization are all trivial to measure. Starting with the energy, the operator form of this observable is simply the Hamiltonian used in simulation[6]. This means

$$\langle E \rangle = \left\langle \sum_{ij} H_{0,ij} + \sum_i H_{0,i} + H_{1,i} \right\rangle, \tag{2.34}$$

and so if we can find the expectation value of each operator independently, we can find the energy as the sum of all these results. Directly expanding Eq. (2.8) into something looking more like Eq. (2.11) we get

$$\langle H_a \rangle = \sum_n \frac{(-\beta)^n}{n!} \left(\sum_\alpha \prod_{i=0}^{n-1} \langle \alpha_i| H_i |\alpha_{i+1}\rangle \right) \langle \alpha_n| H_a |\alpha_{n+1}\rangle \tag{2.35}$$

$$= \sum_m \frac{(-\beta)^m}{m!} \frac{-m}{\beta} \sum_\alpha \prod_{i=0}^{m-1} \langle \alpha_i| H_i |\alpha_{i+1}\rangle \tag{2.36}$$

$$= \left\langle \begin{cases} \frac{-m}{\beta} & \text{if } H_{m-1} = H_a \\ 0 & \text{otherwise} \end{cases} \right\rangle. \tag{2.37}$$

In the second line we let $m = n + 1$ as a new dummy index to make the expression look like the normal partition function. The above naive measure is quite inefficient, as we add a large factor, $-m/\beta$ only when the operator is at the end of the list we sample. Due to

---

[6]Up to overall sign and a constant shift.

the symmetry in permuting the list, that is if we arbitrarily define $|\alpha_i\rangle = |\alpha_0\rangle$ and permute the operator list, we arrive at another state with equal weight. We can use this idea to average over all permutations of the list, arriving at equation

$$\langle H_a \rangle = \left\langle \frac{-m_a}{\beta} \right\rangle, \tag{2.38}$$

where m_a is the expectation of the number of that particular type occurring in the list, which can also be written as $(m_a/m)(m/\beta)$, where the first term is the density of the operator and the second is the prefactor from Eq. (2.37). This is actually quite remarkable: for an observable whose operator is in the Hamiltonian no modification is required to measure it, and we can simply count the occurrence of that operator in the normal simulation divided by the inverse temperature. For the expectation of the transverse field, this is equivalent the spatial average of the operator $H_{1,i}$ over all sites of the lattice.

Going back to the energy, which as we stated is merely the sum of all possible operators, we arrive at the simple expression

$$\langle E \rangle = \left\langle \frac{-n}{\beta} \right\rangle, \tag{2.39}$$

where n is just the number of non-identity operators in the list. This also implies that the energy is strictly negative, as the number of operators is strictly positive, but this is an artifact of simulating a modified Hamiltonian for which we have added sufficient negative constants to facilitate simulation. This reveals another interesting property, that the energy of the system is directly related to the average size of the operator expansion given a particular β. Conversely, for a system at a low enough temperature to be near its ground state energy the expected size of the operator expansion one must sample grows linearly with β.

For operators that are diagonal in the basis representation measurement is much easier. In this case the operator does not interfere the $|\alpha\rangle$ expansion, and it can be measured anywhere in this list without affecting the topology of the simulation. Formally, one has

$$\langle O_{\text{diag}} \rangle = \sum_n \frac{(-\beta)^n}{n!} \sum_\alpha \prod_{i=0}^{n-1} \langle \alpha_i | H | \alpha_{i+1} \rangle \sum_{k=0}^{n} \frac{O_k}{n}, \tag{2.40}$$

where $O_k = \langle \alpha_k | O | \alpha_k \rangle$, corresponding to one additional insertion of an identity operator and the measured operator averaged over all permutations of the list. In practice O_k may have very weak k dependence for a particular simulation cell. In this case, we may avoid

32
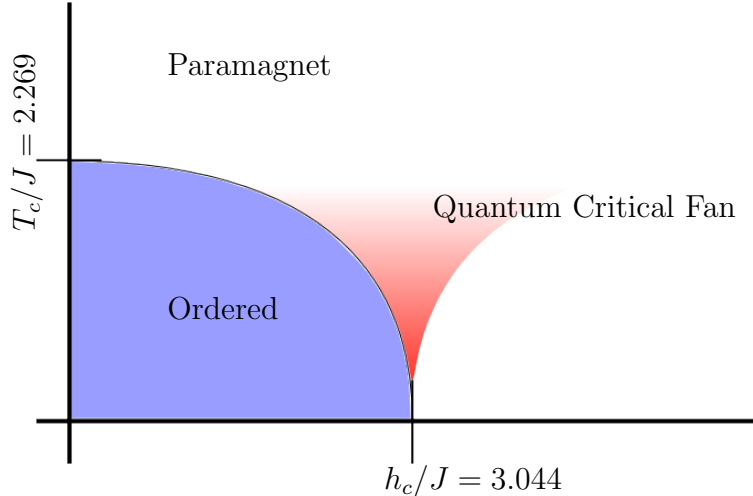
Figure 2.3: The phase diagram of the two dimensional transverse field Ising model on the square lattice. When the transverse field is small there is a continuous phase transition from a high temperature paramagnetic phase to an ordered phase (ferromagnet or antiferromagnet). Moving along the zero temperature axis there is a quantum phase transition from the ordered phase to a transverse polarized phase through a quantum phase transition. The effects of the critical point can be noticed in the finite temperature regime above it, as shown in Ch. 5.

measuring O at every $|\alpha_i\rangle$, and may instead just measure O at $|\alpha_0\rangle$. If the measurement of O is computationally expensive this can greatly increase the efficiency of the simulation.

Finally, although it is possible to measure off-diagonal operators that are not in the Hamiltonian, this is not typically done, and requires a non-trivial modification to the algorithm.

2.3 Phase transitions in simulations of the transverse field Ising model

As a final part of this chapter, we will discuss how the two phase transitions, the thermal and quantum phase transitions, manifest themselves when simulating the TFIM in two spatial dimensions. For the TFIM in 2D, there is a quantum critical point at $h_c/J \approx 3.044$

[54]; for $h > h_c$ there is no thermal phase transition and the phase is polarized in the x-axis at low temperature, while for $h < h_c$ the system goes through an Ising transition to a ferromagnetic (or antiferromagnetic, depending on the sign of J) ground state. This phase diagram is shown in Fig. 2.3. Since the TFIM SSE works for all choices of h and J, including $h = 0$, we can discuss the thermal phase transition in the pure Ising model limit.

In this picture, the simulation cell only has a single type of operator, $H_{0,ij}$, which connects two neighbouring spins on the lattice at some slice of the expansion. During the diagonal update these operators are removed and new operators are placed wherever the local conditions in $|\alpha\rangle$ allow. When temperature is very high (small β), there are very few operators, and in the extreme limit of infinite temperature there are none at all. If we view the off-diagonal update in this context, each spin is an independent cluster and can freely flip. Compare this to the high temperature phase of the classical Ising model, in which every spin can equally take both up or down configuration independent of its neighbours.

As we decrease the temperature we introduce a finite density of $H_{0,ij}$ operators, reducing the number of independent clusters. Each cluster must now flip together during the off-diagonal update. If we examine this in the classical picture, this represents the development of tiny clusters with some characteristic, but finite, correlation length. The ferromagnetic low temperature phase has a similarly simple picture, as it is the case where the number of operators is so large that every spin in the system is in a single cluster. In this case, this single cluster may be flipped with probability one-half, corresponding to a move that takes the system from one ferromagnetic ground state to the other.

Finally, the most interesting point is the thermal phase transition itself. At this point, classically, we know that the system becomes scale-invariant, that is clusters of spins of all sizes should exist. Since the operators in the $2 + 1$ expansion tie clusters together, this thermal phase transition in two spatial dimensions corresponds to the percolation transition of our three dimensional simulation cell. I use this term as, in an infinite system, this temperature is the point at which the first arbitrarily large clusters form. Prior to this temperature (when approaching from the high temperature phase) the clusters are finite in size even in an infinite system.

We can also look at the transition at zero temperature from the classical Ising model through the quantum phase transition. In this case, we start in the regime where the system is in a single large cluster and we increase the relative density of $H_{0,i}$ and $H_{1,i}$ operators compared to the $H_{0,ij}$ operators. These single spin operators define possible boundaries between clusters, and so we see that the ferromagnetic state is destroyed when the relative density of these operators is sufficient to break up the large single cluster. In

this way the quantum phase transition is also a percolation transition, although instead of connecting all the spins at some finite number of operators this argument holds at an infinite number of operators where only the relative density of the two types changes.

This possible alternative picture of the transitions in the simulation may be useful when examining dual models for which the statistics of the cluster structure itself is interesting [55].

With the framework for quantum Monte Carlo discussed (and the assumption of a basic understanding of classical Monte Carlo), we can now discuss the various modifications needed to examine the Rényi entanglement entropy in classical and quantum systems. These modifications are straightforward, and only require minor modification to the normal simulation to calculate these quantities. The following chapter discusses the implementation and procedure for calculating these quantities, as well as examples of the results that can be expected on a variety of standard Monte Carlo systems.

Chapter 3

Rényi entanglement entropy and mutual information in simulation

In this chapter we discuss the modifications and measurements needed to measure the Rényi entanglement entropy (REE) and the mutual information (MI) in both classical and quantum Monte Carlo methods. The essence of the numerical approach is the adaptation of the replica trick from field theory, a standard approach in that field for calculating the Rényi entropies through the use of a modified partition function (see [56] for a relevant example). The formulation of the replica trick in numerics is a direct implementation of these modified partition functions in Monte Carlo [43]. Since Monte Carlo methods can access some of the models that can be exactly calculated in field theory, this comparison sets precedent for making comparison in cases where the field theory solutions are approximate or where no field theory solutions are possible. In particular, this opens up the possibility of studying interacting models in higher dimensions, something that pure field theory techniques often have trouble even approximating.

In the field theory approach of calculating the REE, the relevant calculation requires taking the logarithm of a ratio of partition functions. Using the properties of the log, this is equivalent to the difference of the logarithm of these two partition functions. Recalling thermodynamic relations, the log of a partition function can be related to the free energy (up to a factor of $k_b T$), and so we can relate the calculation of the REE to a calculation of a free energy difference between two systems. Unfortunately, in Monte Carlo methods we rarely have access to the free energy. To calculate the free energy we would need to be able to calculate the partition function, and in Monte Carlo this is our normalization factor, and so it cannot be easily calculated directly. We can get around this by integrating

the energy from a point where the partition function is known (see Eq. (3.9)), typically at infinite temperature.

For this chapter we assume that thermodynamic integration will be used to calculate the log of the partition function, as this requires the least modification to a typical Monte Carlo. Using this technique we will describe the algorithm used to calculate the REE in Monte Carlo and the results for a set of basic models. Other techniques are discussed in Ch. 4,5,6, but all of them require knowledge of the framework discussed here.

Included in this chapter is the demonstration of the calculation of the REE for the Ising model and the classical XY model. In these systems we show that crossings of the mutual information per boundary length can be used to detect thermal phase transitions. In addition, for the Ising model we show that the scaling of the REE follow universal predictions for the universality class it belongs to.

3.1 Rényi entropy and mutual information

The REE is defined as [57]

$$S(A)_\alpha = \frac{1}{1-\alpha} \ln \text{Tr} \big[\rho_A^\alpha\big], \tag{3.1}$$

where ρ_A is the reduced density matrix of some subregion A, defined as

$$\rho_A = \text{Tr}_{\bar{A}}\rho, \tag{3.2}$$

where $\bar{A}$ represents all the degrees of freedom contained in the complement to region A and ρ is the full density matrix of a system. For a classical system, we can define the full density matrix as

$$\rho_{i,j} = \delta_{i,j} e^{-\beta E_i}/Z, \tag{3.3}$$

which has the property that it is diagonal by construction. To contrast, for a quantum system the density matrix is, in general, a Hermitian matrix that can have complex off-diagonal components.

The mutual information is a symmetrized analogue of the REE, defined as

$$I_\alpha(A, B) = S_\alpha(A) - S_\alpha(A|B) \tag{3.4}$$

$$I_\alpha(A, B) = S_\alpha(A) + S_\alpha(B) - S_\alpha(A \cup B). \tag{3.5}$$

The entropy $S_\alpha(A|B)$ is the entropy of region A given knowledge of region B, and as a whole the mutual information quantifies how much the configuration space of region A and B are dependent on one another. Formally this conditional entropy is defined as [58]

$$S_\alpha(X|Y) = \sum_i P(Y = Y_i)S_\alpha(X|Y = Y_i). \tag{3.6}$$

In the case that $S_\alpha(A|B) = S_\alpha(A)$, that is to say that the entropy of a region A is independent of the configuration in region B, then the mutual information is trivially zero.

The reason these REEs are interesting from a numerical point of view is that for integer REEs of two or greater, they can be measured easily in a Monte Carlo simulation [8, 43] and in field theory [56]. Compare this to the calculation of the von Neumann entanglement entropy where calculation of the reduced density matrix is needed, and typically requires either exact diagonalization or density matrix renormalization group methods to calculate [59, 60]. To calculate the REE we use a mapping known as the *replica trick* [56, 43], which defines

$$\mathrm{Tr}\left[\rho_A^\alpha\right] = \frac{Z[A, \alpha, \beta]}{Z(\beta)^\alpha} \tag{3.7}$$

where $Z(\beta)^\alpha$ is the normal partition function at inverse temperature β raised to the power α and $Z[A, \alpha, \beta]$ is the partition function of α copies of the system with some special boundary conditions in region A. It should be noted that $Z[\emptyset, \alpha, \beta] = Z(\beta)^\alpha$, and correspondingly $Z(\beta) = Z[U, n, \beta/n]$, where U is the set containing the entire system. Since the natural log of the reduced density matrix is what is needed to calculate the REE, this corresponds to the log of this ratio of partition functions. To calculate the log of the partition function we can rewrite the thermodynamic relationship

$$\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta} \tag{3.8}$$

$$\ln Z(\beta) = \ln Z(\beta \to 0) - \int_{\beta=0}^{\beta} \langle E(\beta') \rangle \, \partial\beta', \tag{3.9}$$

where $\ln Z(\beta \to 0) = S(\beta \to 0)$ is just the infinite temperature entropy of the system, a quantity that is proportional to the number of degrees of freedom in the system. Eq. (3.9) is valid for both the regular partition function and for $Z[A, \alpha, \beta]$, and so the problem of calculating the REE is reduced to the problem of calculating the energy of the normal and modified simulation over a range of temperatures and performing numerical integration. In this way we are able to (effectively) calculate the free energy difference for these two partition functions and the logarithm of the ratio of partition functions.

3.2 Rényi entanglement entropy in classical simulations

The modification to the classical simulation for the partition function $Z[A, \alpha, \beta]$ is very straightforward [8]. This section discusses the formulation of the modified partition function and its realization in simulation, the necessary considerations needed to perform the simulation, and present results using the Ising model and the XY model. For simplicity we will discuss the second Rényi entropy and mutual information, although extensions to higher Rényi entropies in the classical language are very straightforward.

It is easiest to see the nature of the modified simulation by writing out the trace over ρ_A^2 from a classical perspective. We remind that since ρ for a classical system is purely diagonal (see Eq. (3.3)), we can formulate the modified simulation in term of its diagonal elements, p_i. Writing it out in full we have

$$p_i = e^{-\beta E_i}/Z \tag{3.10}$$

$$p_{i_A, i_{\bar{A}}} = e^{-\beta E_{i_A, i_{\bar{A}}}}/Z \tag{3.11}$$

$$p_{i_A} = \sum_{i_{\bar{A}}} e^{-\beta E_{i_A, i_{\bar{A}}}}/Z \tag{3.12}$$

$$p_{i_A}^2 = \sum_{i_{\bar{A}}} e^{-\beta E_{i_A, i_{\bar{A}}}}/Z \sum_{j_{\bar{A}}} e^{-\beta E_{i_A, j_{\bar{A}}}}/Z. \tag{3.13}$$

In the first step we split the index i summing over states to $i_A, i_{\bar{A}}$ which sums over the states with a particular configuration in region A and region $\bar{A}$. This assumes the system can always be written as the tensor product of some configuration in region A and some other configuration in the complementary region, but this is typically the case for the lattice systems we deal with[1]. The probability $p_{i_A}^2$ defines the set of states that we need to sample in order to calculate $Z[A, 2, \beta]$. This can be seen by writing out the trace of p_i and $p_{i_A}^2$, each of which define a partition function

$$Z(\beta) = \sum_{i_A, i_{\bar{A}}} e^{-\beta E_{i_A, i_{\bar{A}}}} \tag{3.14}$$

$$Z[A, 2, \beta] = \sum_{i_A, i_{\bar{A}}, j_{\bar{A}}} e^{-\beta E_{i_A, i_{\bar{A}}}} e^{-\beta E_{i_A, j_{\bar{A}}}}. \tag{3.15}$$

---

[1]For a discussion on systems which do not necessarily admit this description, see appendix A.
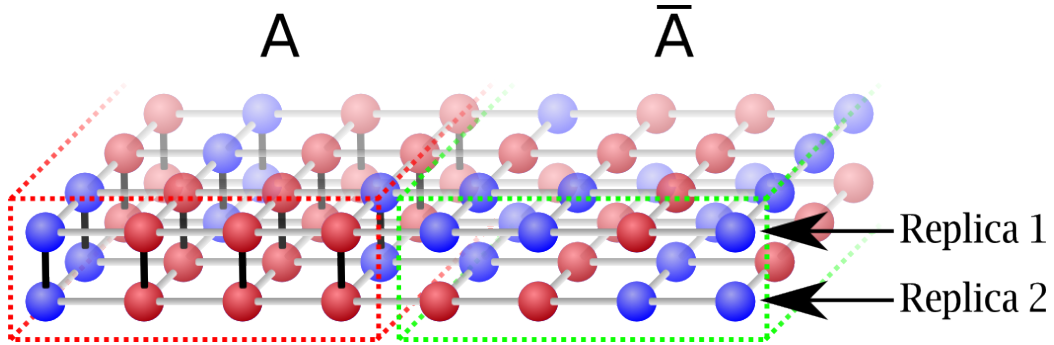
Figure 3.1: The modified classical simulation cell of a 2D square lattice Ising model, viewed edge on. The blue and red circles represent up and down spins, while the grey bonds represent the square lattice on which the Hamiltonian acts. The black bonds in the left area outlined in red, region A, represent the hard constraint between the replicas. In the right area outlines in green, region $\bar{A}$, there are no constraints and the two replicas fluctuate independently.

Fig. 3.1 shows an edge on view of the modified simulation cell for a 2D square lattice Ising model. Fundamentally the modified simulation consists of two copies of the original system. For the two dimensional classical lattice models, this means the modified simulation contains two $L \times L$ layers of spins. In region A, the degrees of freedom in the two layers are enforced to be the same; Eq. (3.15) shows this by the matching index i_A in the two energy terms while $i_{\bar{A}}$ and $j_{\bar{A}}$ are different in the two terms.

The hard constraint in region A that the spins must have the same configuration in both layers must be satisfied at every step of the simulation. For the Ising model, single spin updates are sufficient to satisfy balance and ergodicity, and so we can discuss how they must be modified in light of this constraint. In the complement of region A there is no constraint between the two replicas and updates proceed as they would in the unmodified system, that is to say that when updating a spin we only examine the neighbouring spins in the same layer and there is no interaction with the spins in the other replica. In region A a single spin in one layer cannot be updated in isolation and the constraint requires that we update spins in both replicas simultaneously. The need to update spins in both replicas necessarily means that the local configuration in both replicas will affect the acceptance probability of such an update, and it is in this detail that the modified simulation differs from the simulation of two completely non-interacting copies of the system.

The constraint on region A also has some generic effects on the dynamics of the system,

most of which are part of what make it useful as a system to study. Since spins in both replicas must be updated simultaneously, the number of spins that need to be considered for an update in region A is twice as many as an update in the complementary region. For a spin in region A surrounded by other spins in region A, this means that the typical update will have twice the energy cost when compared to the corresponding move without the constraint between replicas. To expand on this, let us examine the minimal non-zero energy update in the two dimensional Ising model on a square lattice. In this case, we will imagine a central spin up, surrounded by three spin up and one spin down. In a single replica, the local energy of this configuration is $-2J$, and flipping the central spin would take us to a configuration with $2J$ energy. In region A the energy in both replicas is $-4J$, while flipping the spin takes us to a state with a local energy of $4J$. Comparing these two scenarios, we calculate the acceptance probability of these moves as

$$P_{\bar{A}} = e^{-\beta 4J} \tag{3.16}$$

$$P_A = e^{-\beta 8J} = e^{-2\beta 4J}, \tag{3.17}$$

where we can see where the half temperature (2β) interpretation of region A can be seen.

This factor of two can just as easily be associated with β, in which the interpretation of region A is that of a single replica at half the temperature connected to two replicas at the normal temperature. The effective half temperature in region A is emergent in the design of the simulation cell, and does not require any fine tuning of the updates or the temperature. What this means is that when region A is large enough it will go through a thermal phase transition at $2T_c$ while the unconstrained part of the system will go through a phase transition at T_c. Using this view of the system we can try to estimate the behaviour of the mutual information.

3.2.1   Ising model results

First, we can write out the simplified form of the mutual information assuming that region A and B are related by some symmetry, such as two $L/2 \times L$ cylinders of an $L \times L$ torus representing a finite size approximation of the 2D Ising model with periodic boundary conditions. In this case we can write the mutual information as

$$I_2 = -2\ln(Z[A, 2, \beta]) + 2\ln(Z[\beta]) + \ln(Z[2\beta]), \tag{3.18}$$

where $Z[2\beta] = Z[A \cup B, 2, \beta]$, or rather the system in which all spins are a part of region A is identical to the unmodified system is at half the temperature.
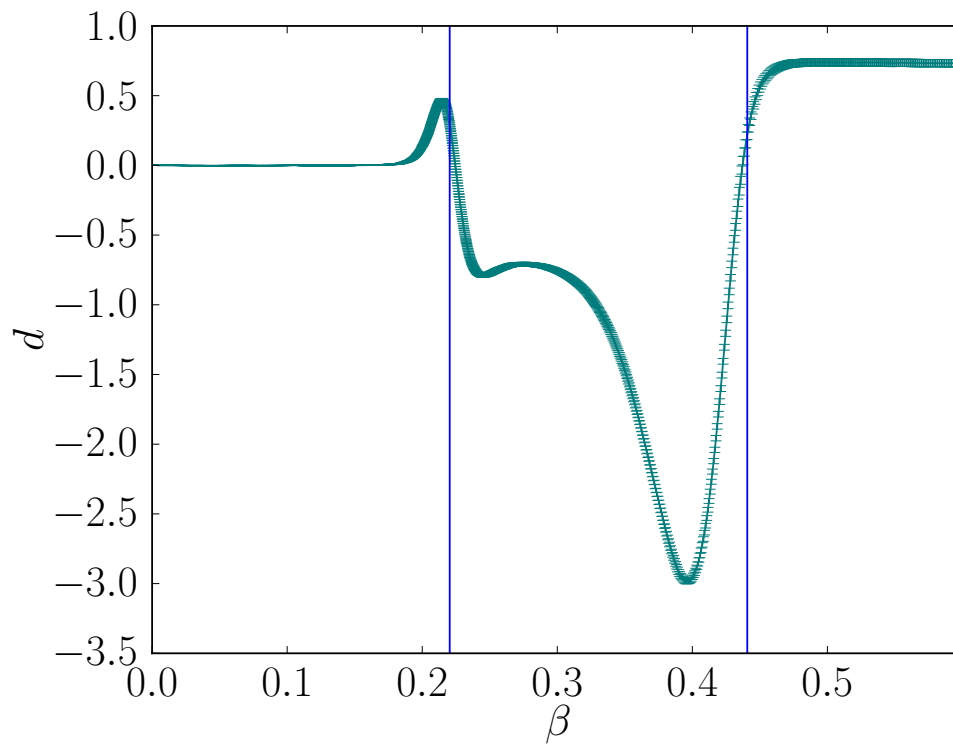
Figure 3.2: (Reproduced with permission from [61]) The constant part of the mutual information in the Ising model extracted using $L = 20$ and $L = 30$ system sizes. The two vertical lines represent (left) $2T_c$ and (right) T_c. The sign of the constant piece is consistent with predictions from theoretical calculations.

Above $2T_c$ both regions of the system are in the high temperature phase, and although region A is at a lower temperature than the unconstrained region, both are effectively in the paramagnetic regime and fluctuate randomly with some associated short correlation length. At these temperatures both region A and its complement are at high enough temperature that under the renormalization group procedure they flow to the infinite temperature fixed point. At infinite temperate the log of the partition function is simply $\ln(Z) = S = N \ln(2)$ where N is the number of spins in the system. For the modified simulation this number N is the number of independent degrees of freedom. This means that spins in region A are only counted once for all replicas, while spins in region $\bar{A}$ are counted once for every replica in the modified simulation. If we are above criticality it is reasonable to suggest that the log of the partition function takes the form

$$\ln(Z[\beta]) \approx (N/\xi_\beta^2) \ln(2) \tag{3.19}$$

$$\ln(Z[A, 2, \beta]) \approx (N/2\xi_{2\beta}^2 + N/\xi_\beta^2 - \mathcal{O}(\ell)) \ln(2). \tag{3.20}$$

These expressions are approximate as they certainly have not only finite size corrections, but there may be a lattice dependant geometric factor multiplying the correlation length. The term N/ξ_β^2 should be interpreted as the effective number of (independent) degrees of freedom, realized after some number of coarse graining steps in the renormalization group procedure. The $\mathcal{O}(\ell)$ term in the second line comes from the fact that at the interface between the constrained and unconstrained region we cannot use the bulk description to determine the number of degrees of freedom, and we can expect corrections proportional to the length of the boundary between the two region, ℓ. Plugging this form into Eq. (3.18) we get

$$I_2 = - 2(N/2\xi_{2\beta}^2 + N/\xi_\beta^2 - \mathcal{O}(\ell)) \ln(2) + 2(N/\xi_\beta^2) \ln(2) + (N/\xi_{2\beta}^2) \ln(2) \tag{3.21}$$

$$I_2 = \mathcal{O}(\ell) \ln(2). \tag{3.22}$$

Fig. 3.2 shows how this result is consistent with numerics, although near just above T_c there are $\mathcal{O}(1)$ corrections that we cannot explain in this simplified approximation [61].

In the region between $2T_c$ and T_c the picture of the simulation cell is very different. In region A the system is below the critical temperature, meaning that the system only effectively has two states it can access. The unconstrained region is identical to the previous picture, so the log of the partition function now takes the value of

$$\ln(Z[\beta]) \approx (N/\xi_\beta^2) \ln(2) \tag{3.23}$$

$$\ln(Z[A, 2, \beta]) \approx (N/\xi_\beta^2 - \mathcal{O}(\ell) + 1) \ln(2) \tag{3.24}$$

$$\ln(Z[2\beta]) \approx \ln(2). \tag{3.25}$$

The additional factor of one in $Z[A, 2, \beta]$ is from the inversion symmetry for the ordered region, and it is the same source of the entropy of the system at 2β. Plugging this in to Eq. (3.18) we now get

$$I_2 = -2(N/\xi_\beta^2 - \mathcal{O}(\ell) + 1)\ln(2) + 2(N/\xi_\beta^2)\ln(2) + \ln(2) \tag{3.26}$$

$$I_2 = (\mathcal{O}(\ell) - 1)\ln(2). \tag{3.27}$$

This is similar to the result at high temperature except that we have a constant negative contribution to I_2 in addition to the piece proportional to the boundary. A negative constant piece to the mutual information is seen in this temperature range, and indeed a brief plateau of the constant piece of the mutual information is seen in numerics (see Fig. 3.2).

Finally, the last region of interest is at low temperature below T_c. In this range all of the log of each of the partition functions is simply $\ln(2)$, reflecting the inversion symmetry of the ground state in each case. Calculating the limiting value of the mutual information in this last regime we get

$$I_2 = -2\ln(2) + 2\ln(2) + \ln(2) \tag{3.28}$$

$$I_2 = \ln(2), \tag{3.29}$$

which differs from the previous temperature range by the sign of the constant. In addition, near and below T_c we expect $\mathcal{O}(\ell)$ corrections due to the finite size of the fluctuation correlation length in the ordered phase for the constrained partition function. This saturation to a constant for the Ising model is the matched in numerical simulation below the critical temperature (Fig. 3.2).

Fig. 3.3 shows the mutual information per boundary length for the classical Ising model [8]. The finite size corrections give rise to a crossing of the curves near T_c and $2T_c$, and these crossings extrapolate with very high accuracy to the critical temperature. The sign of the constant part of the mutual information changing sign can also be seen by how the set of curves fan out in a different direction in the three relevant temperature regimes.

The divergence of the derivative of the total mutual information shown in Fig. 3.3 is a universal prediction [62, 63] which reinforces the idea that mutual information and the Rényi entanglement entropy can be used to extract universal quantities from simulation. The prediction for this divergence comes from the quantum spin-1/2 XXZ model [64, 65]. In that model there is a finite temperature phase transition in the 2D Ising universality class with a scaling prediction of the entropy of the form [19]

$$I_2 = [c_1(t) + t\log t]\ell + c_2(t) + \mathcal{O}(1/\ell). \tag{3.30}$$
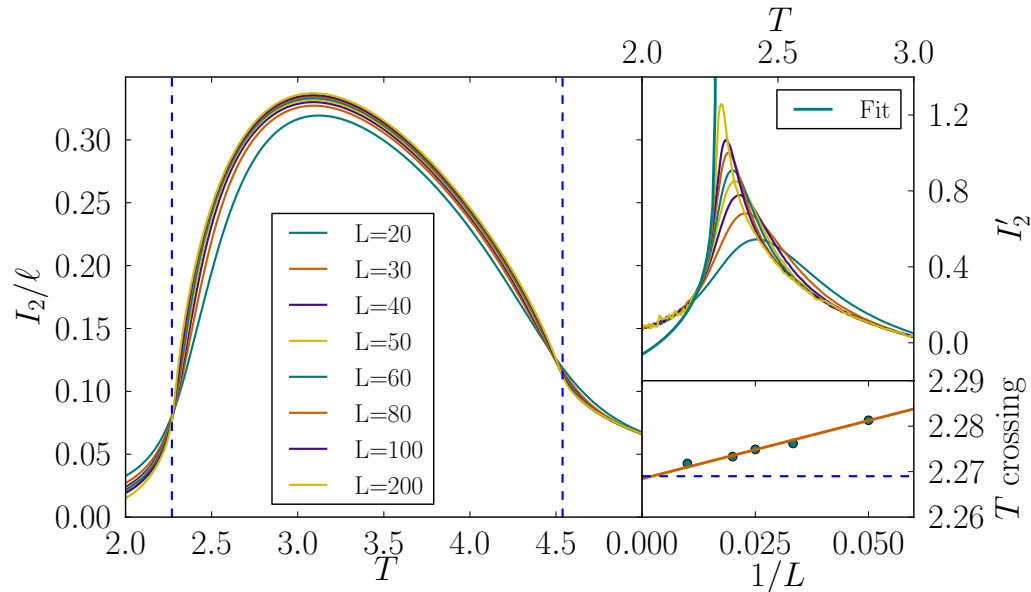
Figure 3.3: (Reproduced with permission from [8]) (left) The scaling of the mutual information per boundary length (ℓ) as a function of temperature for the classical Ising Model. Notice the convergent behaviour of the mutual information per boundary length, and crossings at T_c and $2T_c$ (vertical dashed lines). (top right) The derivative of the mutual information with respect to temperature, showing agreement with the expected log divergence at T_c [62, 63]. (bottom right) The finite size scaling of the crossing of systems of size L and $2L$, as a function of $1/L$. Since this is the leading order correction to the critical temperature for the Ising model, it gives very good scaling to the thermodynamic T_c (blue dashed line).
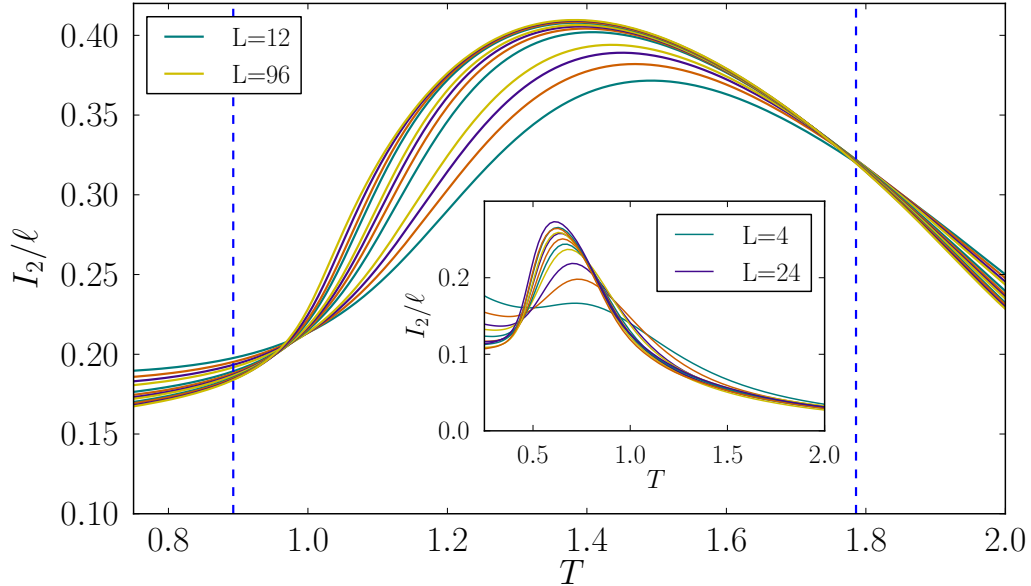
Figure 3.4: (Reproduced with permission from [8]) The scaling of the mutual information per boundary length (ℓ) as a function of temperature for the classical XY Model. Notice that in this case the crossings appear much further from the BKT-transition temperature, a result of the larger finite size corrections for this model. (inset) The crossings of the mutual information of the quantum XY model ($H = -\sum_{\langle ij \rangle} \sigma_i^+ \sigma_j^- + \text{h.c.}$) using a Wang-Landau technique [65, 8], for comparison. The critical temperature is not universal, and in the case of the quantum XY model it occurs at $T_c/J \approx 0.68$ [66].

Taking the derivative of the above expression one expects a divergence of the form $\log(t)$ which can be seen in Fig. 3.3. This suggests that this feature is universal to the 2D Ising universality class for the transition, and can be used as an strong indicator for classification.

3.2.2 XY model results

In addition to the Ising model, the classical XY model

$$H = -J_{\text{XY}} \sum_{\langle ij \rangle} \cos\left(\theta_i - \theta_j\right), \tag{3.31}$$

with $0 \leq \theta_i < 2\pi$ is also examined to see if the Berezinskii-Kosterlitz-Thouless vortex binding transition can be observed using the method of examining crossings in the mutual
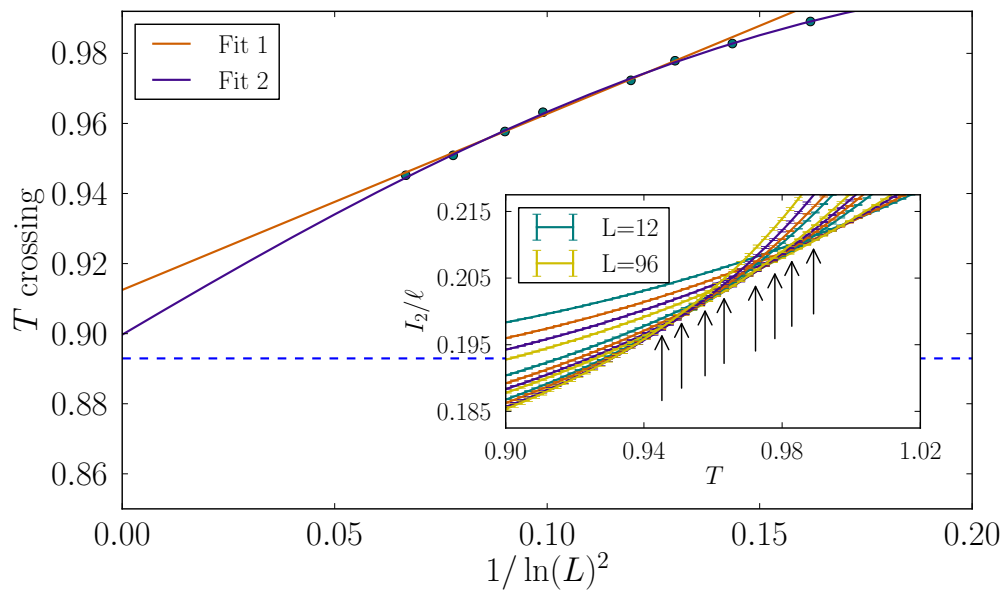
Figure 3.5: (Reproduced with permission from [8]) Finite size scaling of the crossings of system size L and $2L$ in the mutual information for the classical XY model near T_c (horizontal dashed line). Fit 1 assumes a leading order correction of $1/\log(L)^2$, while Fit 2 also assumed a term proportional to $1/\log(L)^5$ (see appendix C), each of these derived by examining the renormalization group flow for the XY model. (inset) A close-up of the crossings with arrows pointing out each crossing, and visible error bars on the data.

information. Fig. 3.4 and 3.5 show the mutual information per boundary length as a function of temperature and system size as well as the scaling of the crossings of system size L and $2L$ as a function of $1/\log(L)^2$, the leading order finite size correction. Although the finite size effects are much larger, there is a convincing fit–using the finite size scaling from KT renormalization group theory (see appendix C)–that shows the crossings scale to the thermodynamic transition temperature. This phase transition is typically very difficult to detect without measuring the spin stiffness [67], and this demonstrates precisely how useful the mutual information might be in more general applications.

3.3 Modifications to the quantum simulation cell

The modification to the quantum simulation is similar in spirit to the modification to the classical simulation, but the $D+1$ expansion in β changes the details of the implementation. Specifically, in the classical case the modification is realized as a hard constraint between the replicas, while in the quantum simulations the replicas are not flat objects. The special boundary conditions still apply to the quantum simulation cell, but instead affect how the expansions are connected together, modifying the $D+1$ topology of the quantum simulation as a whole. In the language of field theory these changes are usually discussed in terms of the modified boundary conditions of a n-sheeted Riemann surface [56], and that picture is very similar to the implementation for quantum systems.

These modifications are discussed in the context of the transverse field Ising model in Ch. 2, but familiarity with stochastic series expansion methods should be enough to understand the implementation presented here. These methods were adapted [43] from field theory approaches [68], but we will attempt to motivate them here using the same language as the classical approach.

Similar to the classical system, in the modified quantum simulation we take α copies of the system, two for the second Rényi entropy, and the geometry of region A determines how we join these replicas together. If we examine Eq. (2.31), for a normal simulation there is the restriction that $|\alpha_n\rangle = |\alpha_0\rangle$, or periodic boundary conditions in the expansion direction. This condition is derived from the property that while taking the trace we sum over the weights, in the stochastic series expansion of $e^{-\beta H}$, using some orthonormal basis. Without having the trace in mind, the stochastic series expansion of $e^{-\beta H}$ has effectively open boundary conditions in the expansion direction. Since $e^{-\beta H} = \rho$ (up to a normalization factor), if we wish to calculate ρ_A, then we must again trace over the degrees
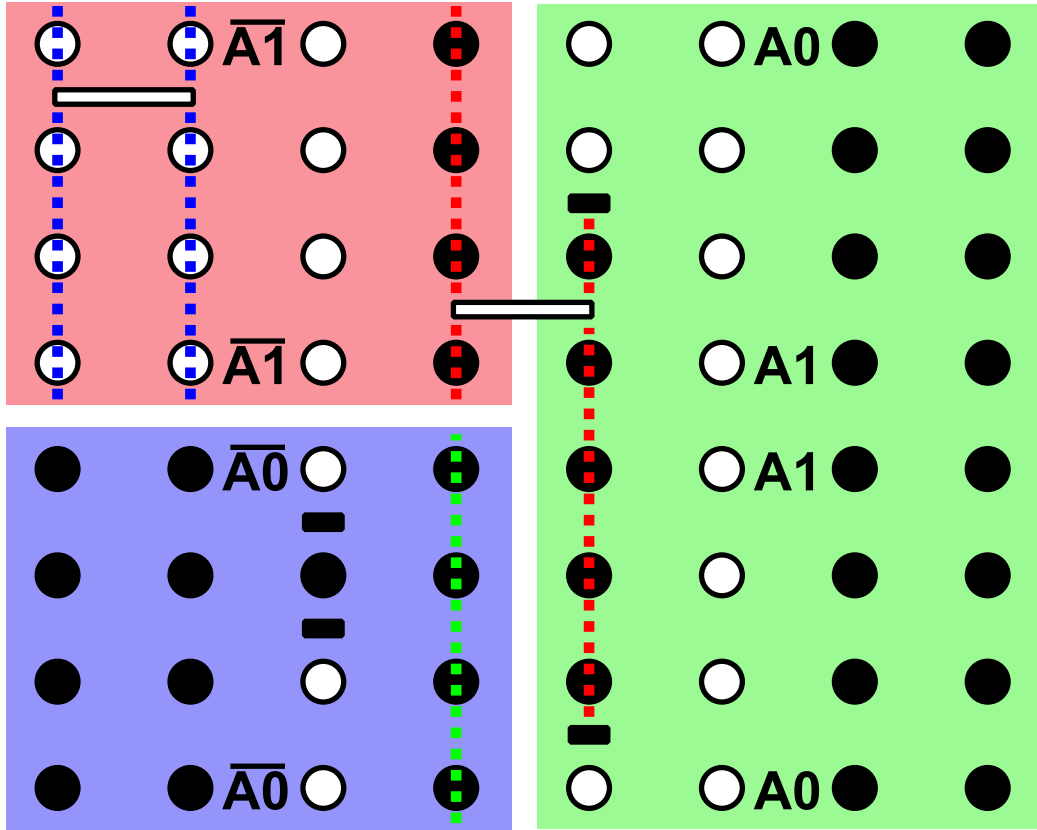
Figure 3.6: The modified quantum simulation cell, with regions and (some) clusters drawn explicitly. The labels, $A0$, $A1$, $\bar{A}0$, $\bar{A}1$, represent where the boundary conditions are enforced to match, either due to the partial trace ($\bar{A}0$, $\bar{A}1$), multiplication ($A1$) or the full trace ($A0$). Colors mirror this information, differentiating region A (green) and region $\bar{A}$ (blue and red) for the first and second replica. The number of operators in the quantum simulation is roughly proportional to the temperature, so region A is at effectively half the temperature as the complementary region. Boundary conditions affect the formation of clusters, especially when operators cross between region A and it's complement (red cluster).

of freedom in $\bar{A}$. This means,

$$\rho_A = \mathrm{Tr}_{\bar{A}}[e^{-\beta H}] \tag{3.32}$$

$$= \sum_\alpha \sum_{S_M} \frac{(-\beta)^n (M-n)!}{M!} \langle \alpha_{0_{\bar{A}}} | \left[\prod_{i=0}^{M_1} |\alpha_{i_A}\rangle \otimes |\alpha_{i_{\bar{A}}}\rangle \langle \alpha_{i_A}| \otimes \langle \alpha_{i_{\bar{A}}} | H_i \right] |\alpha_{0_{\bar{A}}}\rangle , \tag{3.33}$$

where the modification in notation is to emphasize the the that the boundary conditions are $|\alpha_{0_{\bar{A}}}\rangle = |\alpha_{M_{\bar{A}}}\rangle$ and there are open boundary conditions for $|\alpha_{i_A}\rangle$. The important concept is then how we represent ρ_A^2, which is for the most part simply writing the above expression twice. If we write out the square of the reduced density matrix as $\rho_{A0}\rho_{A1}$, where $A0$ and $A1$ are simply dummy indices, then the non-zero elements of this product will satisfy $|\alpha_{M_{A0}}\rangle = |\alpha_{0_{A1}}\rangle$. If we wish to take the trace of this product we then constrain the other indices, ensuring that non-zero contributions satisfy $|\alpha_{M_{A1}}\rangle = |\alpha_{0_{A0}}\rangle$. Fig. 3.6 shows how these boundary conditions would be satisfied for a $1+1$ quantum simulation cell, and how those boundaries change cluster formation.

With the modified simulation cell we can now calculate the energy of the modified system and the normal system, and plug this in to Eq. (3.7) and Eq. (3.9) to calculate the REE and MI. Updates in the modified simulation are performed in each replica without knowledge of the other replica, that is to say we still use the local number of operators in a replica to determine the acceptance rate of new operators rather than the total number of operators in the combined system. The cluster updates act on the system as a whole, as the modified boundary conditions determine the cluster topology, as shown in Fig. 3.6. There are some exceptions to this detail though, as Wang-Landau approaches (see Ch. 4) require using the total number of operators in all replicas to find the density of states. Observables other than energy are not as well defined in this strange geometry, and although they can be measured their interpretation should be considered carefully before comparing them to results from the normal simulation.

The Rényi entanglement entropy is a phenomenally powerful measure, used for a long time in field theory techniques and finally available to the broad set of problems in classical and quantum Monte Carlo. In this basic implementation, there are already incredible results that can be obtained, such as phase transition detection and identification of universal behaviour at phase transitions. The following chapters build up algorithmic improvements that let us further probe difficult problems, where the finite size scaling or very low temperature make this basic approach computationally expensive compared to improved techniques. The tradeoff is that the improved algorithms are more conceptually complex, and not as general as the approach in this chapter.

Chapter 4

Wang-Landau methods in quantum Monte Carlo

Typical Monte Carlo simulations sample states of a system from the weights of those configurations, such as the Boltzmann weights $e^{-\beta E}$. The advantage of such a straightforward approach is that most of the time sampling is spent in the configurations that probabilistically contribute most to any observable of the system. This is not a rigorous statement since it is always possible that a rare configuration of the system may have a large contribution for some observable, and so by this mechanism a small weight configuration may have a large contribution to the expectation value of an observable of the system.

Wang-Landau techniques fall into the broad category of extended ensemble Monte Carlo [69], in which we move beyond this usual approach, typically at the cost of having to account for this modification later when calculating expectation values. The advantage of Wang-Landau in particular is that it allows us to estimate the density of states, $g(E)$. If we know the density (or number in a discrete system) of states with a particular energy, we can explicitly sum over all energies using the Boltzmann distribution to find the energy of the system at any temperature. Recall from the previous chapter that knowledge of the energy over a large temperature range was precisely what was needed to calculate the Rényi entanglement entropy. In addition precise knowledge of the density of states allows direct calculation of the partition function, and by extension free energy, both of which also allow direct calculation of the Rényi entanglement entropies. This means, instead of creating a large set of simulations over a wide range of temperatures we could use a single Wang-Landau simulation and calculate the same information.

This chapter discusses the Wang-Landau implementation for classical and quantum

systems and demonstrates thermal phase transition detection in quantum models, similar to last chapter and classical models. This shows that the calculation of REE does not depend on the classical or quantum nature of the model, but rather on more universal properties. Finally, we shall discuss the efficiency and practicality of the scheme.

4.1 Classical Wang-Landau

To describe the Wang-Landau algorithm [70] we will use the example of a classical system, since extending to quantum systems is straightforward after the basic concepts of the method are understood. The fundamental idea is that we can rewrite the partition function as

$$Z = \sum_i \frac{e^{-\beta E_i}}{Z} \tag{4.1}$$

$$= \sum_E g(E)e^{-\beta E} \tag{4.2}$$

$$g(E) = \sum_i \frac{\delta_{E,E_i}}{Z}, \tag{4.3}$$

where $g(E)$ is the normalized density of states. The benefit of converting to this form is that if one has knowledge of $g(E)$, one can calculate the partition function and any observable based on the energy, such as specific heat, at any temperature. In principle there are ways to incorporate other observables, such as magnetization, so they can also be calculated over a range of temperature. Most crucially, the Rényi entanglement entropy, or more specifically the logarithm of the partition function, can be calculated with knowledge of the density of states. More amazing is that this calculation is valid at all temperature, avoiding the need for thermal integration. In the following sections we will discuss how one uses a modified Monte Carlo simulation to estimate $g(E)$.

The general outline of the Wang-Landau method is the following

1. Start with an assumed form of the density of states (often $g(E) = 1 \forall E$). Overall normalization is unimportant.

2. Perform typical Monte Carlo updates, except instead of using $\min(1, e^{-\beta(E'-E)})$ [52] as the acceptance condition, we use $\min(1, g(E)/g(E'))$, where E is the energy of the initial state, and E' is the energy of the proposed state.

E	$g(E)$	$\langle g(E) \rangle$
$-4J$	$2/16$	$0.168(2)$
0	$12/16$	$1.000(0)$
$4J$	$2/16$	$0.165(2)$

Table 4.1: The density of states, $g(E)$, for the one dimensional four site Ising spin chain with periodic boundary conditions, along with results from a simulation (see appendix B.1). The Wang-Landau method only gives $g(E)$ accurate up to some overall constant, and in light of this the above results are consistent between theory and simulation.

3. Whether the move is accepted or rejected, modify $g(E) \leftarrow kg(E)$, where k is some positive factor greater than one, typically starting at two or the natural log. Also, add one to a counter for the current energy, $C(E)$.

4. Continue the above two steps until $\left[\min(C(E)) - \max(C(E)) \right]/\mathrm{mean}(C(E)) < 0.1$.

5. When the above condition is met, reset $C(E)$ and modify $k \leftarrow \sqrt{k}$.

6. Stop when $g(E)$ is converged.

The above outline is the most straightforward implementation of the Wang-Landau approach. We will discuss optimizations and improvements in Section 4.3, but they do not change the core concept of the algorithm.

To use a concrete example we will discuss the four-site one dimensional ferromagnetic Ising model with periodic boundary conditions. In this case we know that the density of state is exactly, shown in Table 4.1. appendix B.1 has a brief code example that performs a Wang-Landau for this system. Averaging the results from a few runs of this code gives us this density of states presented in the above table, which are accurate to the theoretical results within error bounds.

To examine a bit more closely why this works at all, we can imagine the system when the internal representation of $g(E)$ is incorrect. If $g(E)$ is incorrect, say that the relative ratio of $g(E)$ is wrong for some pair of energies E_1 and E_2, then we can imagine what happens for the moves connecting these two energies. If the actual ratio of density of states, $g(E_1)/g(E_2)$ is lower than the current value assumed in the simulation, then the acceptance condition will cause the simulation to visit states with energy E_1 more often than it visits states with energy E_2. Visiting states causes the density of states to be

multiplied by the factor k, and visiting E_1 more often than E_2 will cause $g(E_1)$ to become larger relative to $g(E_2)$, increasing the ratio $g(E_1)/g(E_2)$. In this way all ratios are slowly and statistically corrected, and the reduction of k through the algorithm attempts to converge the results.

4.2 Wang-Landau in quantum Monte Carlo

In this section we will describe how one can use Wang-Landau concepts in the context of quantum Monte Carlo. Specifically, we will discuss how one would modify a code similar to the transverse field Ising model discussed in Sec. 2.2. This method was implemented prior [71] to the authors attempt [65], but the purpose of reimplementing is to see if it will allow for a more efficient method of calculating mutual information without thermodynamic integration [43].

Compared to the classical Wang-Landau algorithm in which we generate the density of states as a function of energy, in our quantum Monte Carlo we don't have direct access to this quantity. Instead, we take the expansion of the partition function in Eq. (2.11) and rewrite it as

$$Z = \sum_n \frac{(-\beta)^n}{n!} \sum_\alpha \prod_{i=0}^{n-1} \langle \alpha_i | H | \alpha_{i+1} \rangle \tag{4.4}$$

$$= \sum_n \beta^n g(n), \tag{4.5}$$

where $g(n)$ is the quantum analogue to the density of states for the classical case. In practice, quantum Monte Carlo only needs to sample up to a given order n for a particular value of β, and by using a hard cutoff in $g(n)$ we effectively introduce a low temperature cutoff below which we generate unreliable results.

The modification to the SSE algorithm from Ch. 2 comes in how the diagonal updates, those that change the number of operators n, are computed. Eq. (2.32) and Eq. (2.33) are modified from their original form to

$$P_{\text{add}} = \frac{g(n)/g(n+1) \sum_i W_i}{M - n} \tag{4.6}$$

$$P_{\text{rem}} = \frac{(M - n + 1)g(n)/g(n-1)}{\sum_i W_i}. \tag{4.7}$$

where the β dependence has been replaced with the density of states, but the weight and factor due to combinatorics (the $M - n$) still remain. In the quantum case the density of states $g(n)$ is updated after every successful or failed attempt to add or remove an operator, and exactly analogous to the classical case the refinement factor, k, is reduced after the histogram of sampling, $C(n)$ in this case, is found to be flat within some tolerance. Due to the fact that the Wang-Landau approach to sample the density of states will mean that we will necessarily sample all possible number of operators allowed by the system, this will include configurations where the system is completely empty and completely full. The approach in Sec. 2.2 using a bounded-length list is inefficient in these two limits, and for this reason when using Wang-Landau a method with a variable list size is preferred [72].

In addition to the modified diagonal update, the unmodified off-diagonal update–detailed in Sec. 2.2–is also needed to ensure that the simulation remains ergodic, but during these updates no modifications to the density of states $g(n)$ takes place. With just this modification it is possible to use the simulation to calculate the energy as a function of temperature (down to some cutoff), and hence use that to calculate the Rényi entropy and mutual information discussed in Ch. 3.

4.3 Convergence of Wang-Landau

Although the above algorithm calculates the density of states, the simple implementation has problems with convergence. In the example code in appendix B.1, we mitigate this problem by generating independent samples of the density of states and averaging them. Nevertheless, each individual simulation eventually stops converging, meaning that simulation time is being wasted beyond some point in every simulation.

To see why there is a convergence problem, we can construct a simple example. Let us call the approximation to the density of states that the simulation assumes $g'(n)$ while the true density of states is $g(n)$. In general we can write

$$g'(n) = g(n)(1 + \Delta(n)) \tag{4.8}$$

where $\Delta(n)$ is how far from convergence we are. The problem now arises if $\max(\Delta(n)) < 0.1$, since regardless of the value of k the system will seem to be flat within our ten percent tolerance in point 4 of Sec. 4.1. This means that the flatness condition for sampling all the energies will be satisfied when statistical noise is overcome by the number of sampling steps rather than when the distribution has become more converged to the accurate result. What this leads to in practice is that the refinement factor k quickly approaches unity when the

deviation of the density of states becomes smaller than our tolerance condition. Although decreasing the tolerance condition would help mitigate this problem, the tolerance of ϵ will always require $\mathcal{O}(1/\epsilon^2)$ number of steps to converge using a random walk, even in the case where $\Delta(n) = 0$.

The correction to this problem is to ensure that k does not decrease too quickly. One proposed algorithm to achieve this is use the above algorithm to reduce k until $k < e^{1/N_{\mathrm{MC}}}$, where N_{MC} is the number of Monte Carlo steps performed so far [73]. We will refer to this as the $1/t$ algorithm, as $\log(k) = 1/t$ after some time, where t represents the time of the simulation. After this threshold condition is met we simply define k to be

$$k = e^{1/N_{\mathrm{MC}}} \tag{4.9}$$

at every Monte Carlo step. This means that the flatness condition is left unused after this condition is met. Under this scheme for k, it reduces more slowly than it would under the previous scheme assuming the error is lower than the tolerance.

Within the simple Wang-Landau algorithm this improvement is the most straightforward, and it ensures that there is no wasted computational effort in the simulation. The next improvements requires a good starting estimate of the density of states as a starting point, and aims to improve the speed of convergence of the density of states. It is sometimes referred to as minimum round trip [72], as the objective of the algorithm is to minimize the number of Monte Carlo steps needed for the system to move from the largest energy state to the smallest energy state, or vice-versa. Minimum round trip works on the following idea: instead of keeping track of just the histogram of visits, we keep track of separate histograms depending on whether the largest or smallest energy state was visited most recently, calling them C_d and C_u, respectively. The old histogram, $C(n) = C_u(n) + C_d(n)$, is simply the sum of these two, while we also now have information about the ratio between the two histograms, $C_\Delta(n) = C_u(n)/C(n)$. A large jump in $C_\Delta(n)$ signifies a point where the random walk is getting stuck, and minimum round trip uses this condition as a handle for biasing the simulation.

The minimum round trip algorithm is similar to the Wang-Landau approach, but has slight modifications. In this algorithm we use a set of weights, $w(n)$, which in Wang-Landau were simply the inverse of the density of states, $1/g(E)$, but here may become different from them.

1. Starting with a good estimate for the weights, $w(n) \approx 1/g(n)$, sample the system using the Wang-Landau update probabilities.

2. After each accepted or rejected diagonal update, or every move in a classical Monte Carlo, update $C_u(n)$ or $C_d(n)$ depending on whether the full or empty operator string was most recently sampled. Unlike the Wang-Landau algorithm, the weights are not updated.

3. After N steps, update the weights as

$$w(n) \leftarrow w(n)\sqrt{\frac{|dC_\Delta(n)/dn|}{C(n)}} \tag{4.10}$$

4. Add to the running total of histograms, $C_T(E) \leftarrow C_T(E) + C(E)$, reset the histograms $C_u(E)$ and $C_d(E)$, double N, and return to step 1.

If we examine when this algorithm converges the weights, it is when

$$C(n) = \left|\frac{dC_\Delta(n)}{dn}\right|. \tag{4.11}$$

In this case the converged weights are not the same as the inverse density of states. To extract the density of states (again, up to an arbitrary multiplicative factor) we multiply by the total accumulated histogram $C_T(n)$. Specifically,

$$g(n) = C_T(n)/w(n), \tag{4.12}$$

and with the density of state we can calculate the energy and entropy measures over a range of temperatures.

A good estimate for the density of states is required in this algorithm, as for each set of simulation steps N, the updating procedure is only valid if multiple round trips occur. For a large range of n (in $g(n)$) this can be very computationally expensive. To improve this the range of n can be sliced up and simulated in many parallel simulations. To recover the full density of states we can ensure the ranges have enough overlap to recover normalization, and this method allows for optimization as well as increasing the rate of round trips by ensuring the range of n remains small. This division of the range has some overlap with the ratio trick discussed in Ch. 5, although there the range is divided up in to the most minimal range–one of size two.

4.4 Observables

After a thorough set of simulations, the Wang-Landau approach leaves one with an estimate for the density of states. With this quantity for the normal system we can calculate the Energy and total entropy, and if we also have this information for the modified simulation (see Ch. 3) we can calculate the Rényi entanglement entropy and the mutual information.

The most fundamental calculation in Wang-Landau is the partition function itself. In the classical and quantum simulation, the partition function can be calculated as

$$Z(\beta) = \sum_E g(E)e^{-\beta E} \tag{4.13}$$

$$Z(\beta) = \sum_n g(n)\beta^n, \tag{4.14}$$

where $g(E)$ is the classical density of states, and $g(n)$ is the quantum analogue for stochastic series expansion.

Typically our density of states is only known up to some multiplicative factor, but such a factor clearly enters the above expression. This is actually helpful, as at infinite temperature, $\beta = 0$, the value the partition function takes is known. Let us take the case of the 2D Ising model with N spins, and look at both the classical and quantum expression at infinite temperature. Looking at the above expressions, they become

$$Z(0) = \sum_E g(E) = 2^N \tag{4.15}$$

$$Z(0) = g(0) = 2^N \tag{4.16}$$

$$. \tag{4.17}$$

In the classical case this constrains the sum of all terms, but in the quantum case it is even more straightforward, constraining $g(0)$. The multiplicative factor that takes the representation of $g(0)$ to the true value must be applied to all elements of $g(n)$ to account for this unknown overall normalization factor.

An additional constraint arises from our knowledge of the energy at infinite temperature. Examining the Ising model again, or the transverse field Ising model in the classical limit (see Ch. 2), we can constrain one more part of the system. Notably, the general

expression of the energy is

$$E(\beta) = \sum_E E g(E) e^{-\beta E} / Z(\beta) \tag{4.18}$$

$$E(\beta) = -\sum_n n g(n) \beta^{n-1}, \tag{4.19}$$

and examining these in the limit of infinite temperature we get

$$E(0) = \sum_E E g(E) / Z(0) = 0 \tag{4.20}$$

$$E(0) = -g(1)/g(0) = -J N_b, \tag{4.21}$$

where we have assumed that the bonds of the classical Ising model take allowed values of $\pm J$, while for the transverse field Ising model they take values of 0 or $-2J$ on the N_b bonds of the system due to the constants added to the Hamiltonian needed for simulation. Since for both the classical and quantum system there is only one free parameter in the Wang-Landau approach, this second constraint can be used to check if the density of states is underconverged.

The last quantities of interest are the Rényi entanglement entropies and mutual information. Examining the Eq. (3.7), we see that the quantity of interest is the ratio of two partition functions, or more precisely the log of that ratio. Within the context of the modified simulations in Ch. 3, the Wang-Landau procedure requires no special modifications, and a new density of states can be easily extracted. The only point to be careful of is the normalization using the infinite temperature entropy, as the number of degrees of freedom in the modified simulation is affected by the hard constraints of the modified simulation cell. If we consider the normal Ising mode as have N degrees of freedom, one for each spin, then the modified simulation has less than $2N$ due to the constraint in region A. For the case of our Ising model, we have

$$Z[A, 2, \beta] = 2^{N_A + 2(N - N_A)}, \tag{4.22}$$

where N is the total number of spins in each replica, and N_A is the number of those spins in region A.

The limits of computational representation also require a small modification to the Wang-Landau in appendix B.1. As the partition function Z can become extremely large, for a scalable simulation one must store the log of the density of states, $G(n) = \ln(g(n))$, rather than the density of states directly [65]. In practice this is straightforward; the

59

multiplicative updates to the weights become additive and the acceptance probability becomes the exponentiation of the difference of the weights. Storing only $G(n)$ requires that the values be re-exponentiated before summing, and as far as we know there is no method for avoiding this calculation if one uses the log representation. The equation for the log of the partition function in this form is

$$\ln[Z(\beta)] = \ln\Big[\sum_n \beta^n e^{G(n)}\Big]. \tag{4.23}$$

If we define $G'(n) = G(n) + n\ln(\beta)$ and $K = \max_n\{G'(n)\}$ then we can rewrite the above as

$$\ln[Z(\beta)] = \ln[K] + \ln\Big[\sum_n e^{G'(n)-K}\Big]. \tag{4.24}$$

When calculating large quantities such as E or $\ln(Z)$ directly, this representation and subtraction prevents numerical overflow in a limited precision system.

Of final note is the calculation of the error for observables. In a typical simulation, we generate multiple estimates for the density of states, $G_i(n)$. For any observable that can be calculated from this density of states, $F[G(n)]$, one should calculate it in Wang-Landau as

$$\langle F[G(n)]\rangle = \frac{1}{N}\sum_{i=1}^{N} F[G_i(n)] \tag{4.25}$$

$$\Delta\langle F[G(n)]\rangle = \frac{1}{\sqrt{N}}\left(\Big[\frac{1}{N}\sum_{i=1}^{N}(F[G_i(n)])^2\Big] - \Big[\frac{1}{N}\sum_{i=1}^{N}F[G_i(n)]\Big]^2\right)^{1/2}. \tag{4.26}$$

The reason for this method in the calculation of error over attempting to estimate $\langle G(n)\rangle$ and $\Delta\langle G(n)\rangle$ is that by construction the estimate of $G(n)$ is highly correlated, that is to say if a particular $G_i(n)$ is larger than the theoretical value, then $G_i(n+1)$ will likely also be larger than the proper value. Another way to state this is that the errors in $G_i(n)$ are highly correlated in n, and attempting to calculate the error at each point, and then assuming such error is uncorrelated, will lead to an incorrect result. Fig. 4.7 in the next section shows this behaviour.

4.5   Results on quantum models

Using the above approach we use Wang-Landau to calculate the mutual information for variety of prototypical quantum models. In the previous sections, the mutual information

Figure 4.1: (Reproduced with permission from [65]) The mutual information for the XXZ model, Eq. (4.27), on the square lattice with $\Delta = 4$ using the quantum Wang-Landau method for sizes $L = 4$ to $L = 22$ in steps of two. In this limit the model is Ising like, and we see the mutual information cross near $2T_c$ and T_c (vertical blue lines). Scaling of the crossings (inset) of systems of size L and $2L$ shows that they converge to the thermodynamic transition temperature in the limit of large L with the $1/L$ corrections that are expected for the Ising model transition. See Eq. (3.19) for the theoretical discussion of these crossings.

is described as being calculated in quantum simulation by calculating the energy over a large temperature range and integrating the result to calculate the Rényi entanglement entropy and mutual information (see Eq. (3.9)). The difficulty with this approach is the discretization of energy as a function of temperature, which then itself has to be integrated. This naturally leads to a strictly increasing error as temperature is decreased. Since we can directly calculate the partition function at any temperature using Eq. (4.14), we can plug this in to Eq. (3.1) using the mapping in Eq. (3.7) to express the Rényi entanglement entropy in terms of the log of the ratio of two partition functions.
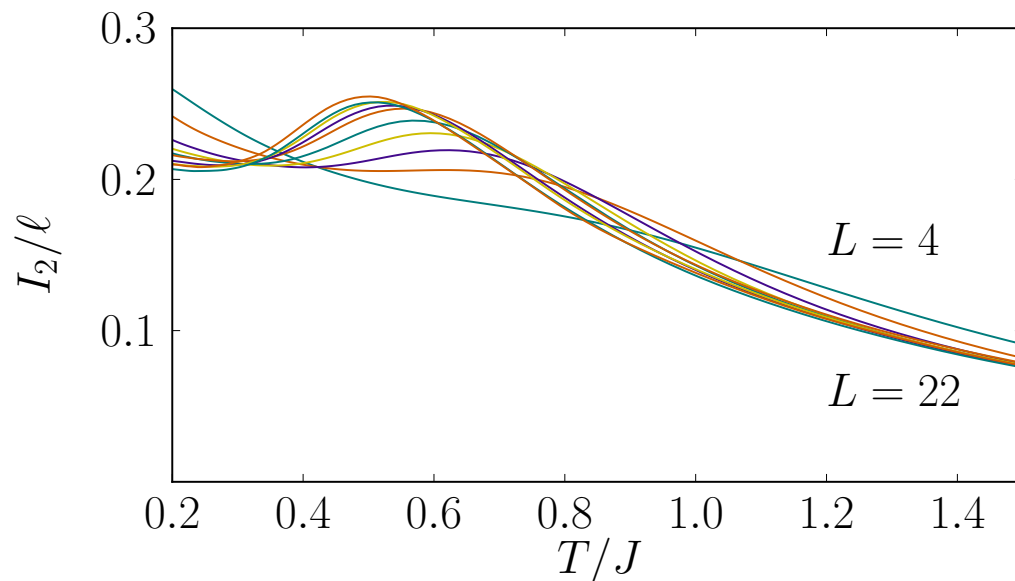
Figure 4.2: (Reproduced with permission from [65]) The mutual information for the XXZ model, Eq. (4.27), on the square lattice with $\Delta = 1$ using the quantum Wang-Landau method for sizes $L = 4$ to $L = 22$ in steps of size two. In this system there is no finite temperature phase transition, and the system only orders at zero temperature. The apparent crossings may be a result of the fact that for finite size systems the zero temperature transition is pushed up to a finite temperature. In this case we expect the crossings to converge to $T = 0$ in the limit of large system size.
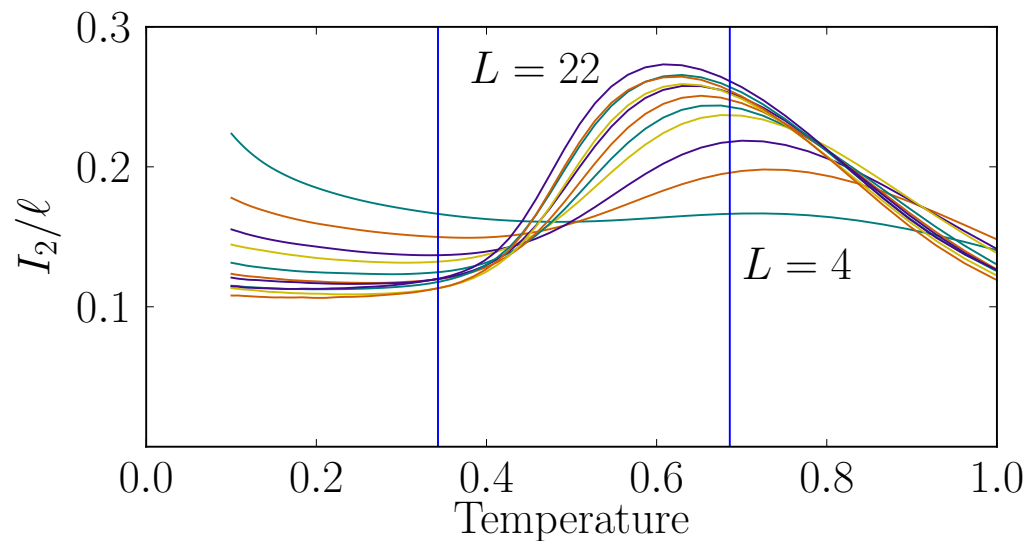
Figure 4.3: (Reproduced with permission from [65]) The mutual information for the XXZ model, Eq. (4.27), on the square lattice with $\Delta = 0$ using the quantum Wang-Landau method for sizes $L = 4$ to $L = 22$ in steps of size two. In the pure quantum XY model there is a finite temperature Kosterlitz-Thouless-Berezinskii transition in which vortices become bound. Although in the quantum simulation we do not have clear crossings between all systems due to our error resolution, it is plausible from this behaviour that crossings will occur at T_c and $2T_c$. Further work in classical systems confirmed that this is the case through a more thorough scaling analysis (see Fig. 3.4)
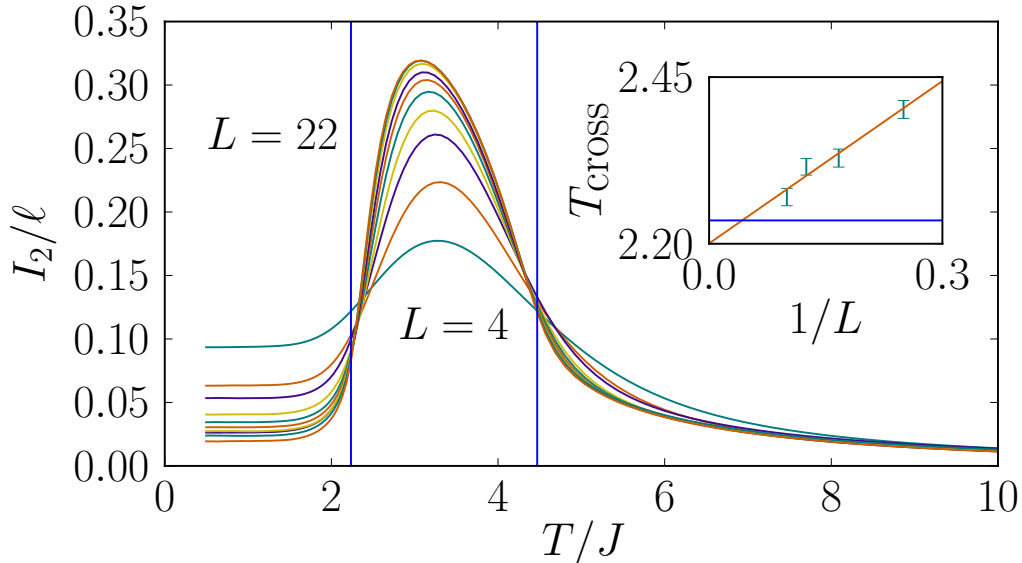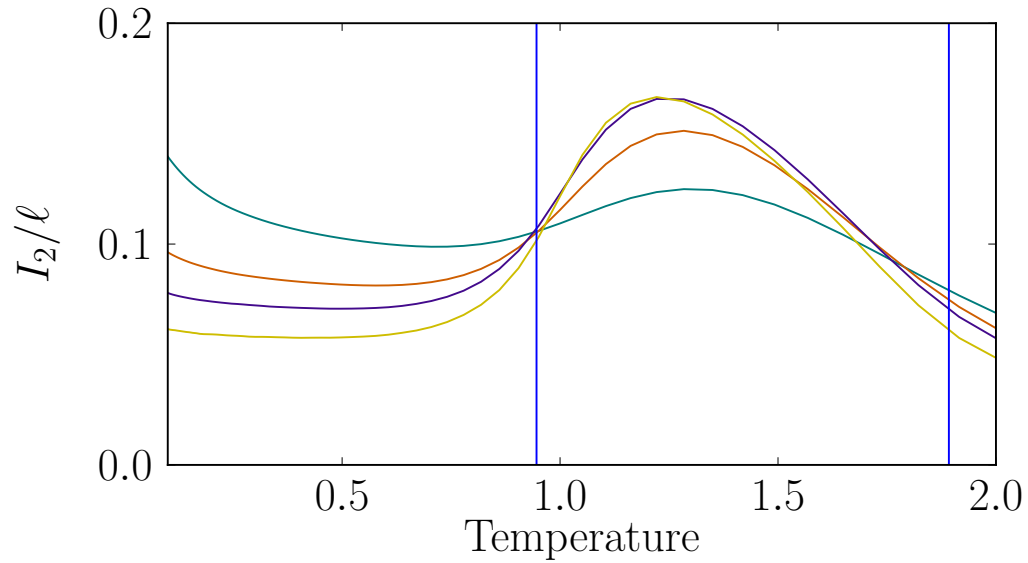
Figure 4.4: (Reproduced with permission from [65]) The mutual information for the XXZ model, Eq. (4.27), on the cubic lattice with $\Delta = 1$ using the quantum Wang-Landau method for system sizes $L = 4, 6, 8, 10$ (teal to yellow). Blue lines represent T_c and $2T_c$. Unlike the two dimensional version, the 3D Heisenberg model does have a finite temperature phase transition to a polarized state. Despite the limited set of sizes that the quantum Monte Carlo can explore in three spatial dimensions, the T_c crossings are quite clear and show small finite size effects.

The XXZ Hamiltonian can be written as

$$H = J \sum_{\langle ij \rangle} (\Delta S_i^z S_j^z + S_i^x S_j^x + S_i^y S_j^y). \tag{4.27}$$

The S term are spin-1/2 operators, typically described as the three Pauli matrices with eigenvalues $\pm 1/2$. At $\Delta = 4$ this model acts like an Ising model with a small quantum perturbation, while at $\Delta = 1$ we recover the $S = 1/2$ Heisenberg model and at $\Delta = 0$ we have the quantum XY model. We remind that for the $S = 1/2$ Heisenberg model there is no finite temperature phase transition, while for the quantum XY model there is a Berezinskii-Kosterlitz-Thouless transition. Fig. 4.1 shows the mutual information for the $\Delta = 4$ XXZ model on a 2D square lattice, while Fig. 4.2 and Fig. 4.3 show the same results for the Heisenberg model and the XY model. Moving to three dimensions Fig. 4.4 shows the results for the 3D Heisenberg model in which we have a finite temperature phase transition. All of these results used the simple implementation of the Wang-Landau algorithm with the $1/t$ algorithm for refining the density of states, then directly used this converged density of states, averaged over a few simulations, to calculate the mutual information. In all the cases where there is a finite temperature phase transition we find that the mutual information has crossings that converge in system size towards the critical temperature in the limit of large size.

To test the minimum round trip algorithm we implemented this method for the transverse field Ising model in two dimensions. Similar to the other systems with a finite temperature phase transition we see well defined crossings at T_c. For large values of J/h the system has an Ising transition similar to Fig. 4.1, and this can be seen clearly. As mentioned previously, all of these methods can be sped up through parallelizing the algorithm by ensuring each simulation only explores a small section of the function $g(n)$ and stitching it back together afterwards.

4.6   Discussion

Overall the implementation of Wang-Landau was successfully able to reproduce full mutual information curves, with the desired result of capturing crossings through the ability to generate data at arbitrary temperature in a large range [65]. The method has specific considerations when scaling to larger systems and lower temperatures, and compared to the original procedure there are implementation specific details that should be considered if one wishes to pursue this approach. As an aside, Wang-Landau describes a class of methods in which one projects a simulation along a particular parameter (energy, magnetization,
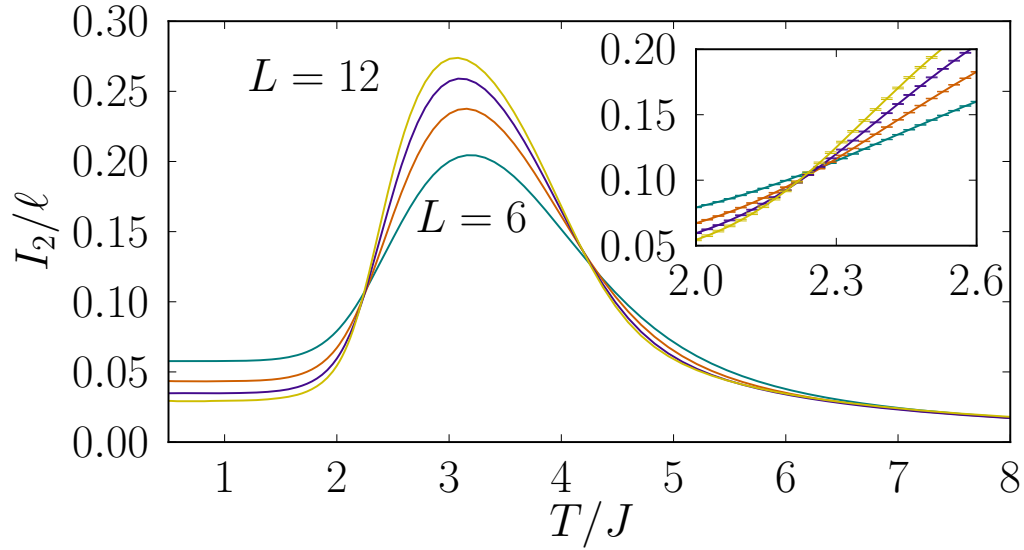
Figure 4.5: (Reproduced with permission from [65]) The mutual information for the transverse field Ising model with $J = h = 1$, Eq. (2.4), on the square lattice using the quantum Wang-Landau method and minimum round trip to refine the density of states, for sizes $L = 6, 8, 10, 12$. We again have clear definition of the crossings at $2T_c$ and T_c, with the inset showing the error at T_c using this approach. These crossings area near the pure Ising model crossings ($T_c/J = 2.269$), which is expected as the perturbation $h/J = 1$ is far from the quantum critical point at $h_c/J = 3.044$.
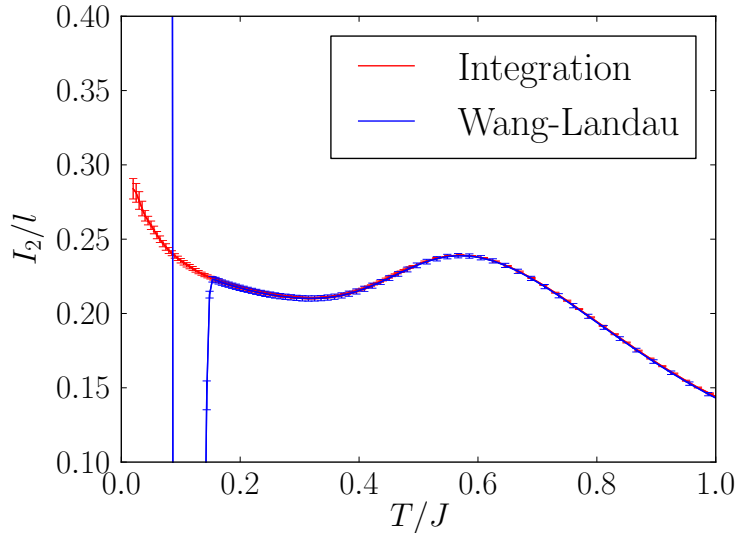
Figure 4.6: (Reproduced with permission from [65]) The mutual information of the 2D Heisenberg model for a 12×12 system calculated with Wang-Landau and by energy integration. Notice how the results track well down to $T/J \approx 0.15$ and then diverge wildly for the Wang-Landa method.

region size in the modified simulation) and attempts to find the ratio of the density of states with respect to that parameter. In this light it shares much in common with the "ratio trick", a common method for determining the ratio between two partition functions with different region sizes of a modified simulation (see Ch. 5). Improvements that work for the ratio trick should then also be efficient for improving Wang-Landau methods, and vice versa.

In the quantum language the density of states $G(n)$ is effectively the contribution from the n-th order of the Taylor expansion of $\text{Tr}[e^{-\beta H}]$. For any particular order or n there is a β at which it becomes relevant to calculating the energy. This means that if we cutoff at some value of n we put an upper limit on the β for which we can generate accurate results. Fig. 4.6 shows how this cutoff manifests if one attempts to use a truncated density of states to generate data beyond the intended range. To determine the cutoff in n needed to generate accurate results at a particular β a non Wang-Landau simulation is used to generate statistics at the cutoff. The normal simulation will sample a characteristic number of operators, and this number can be used as a guide for setting the cutoff for the
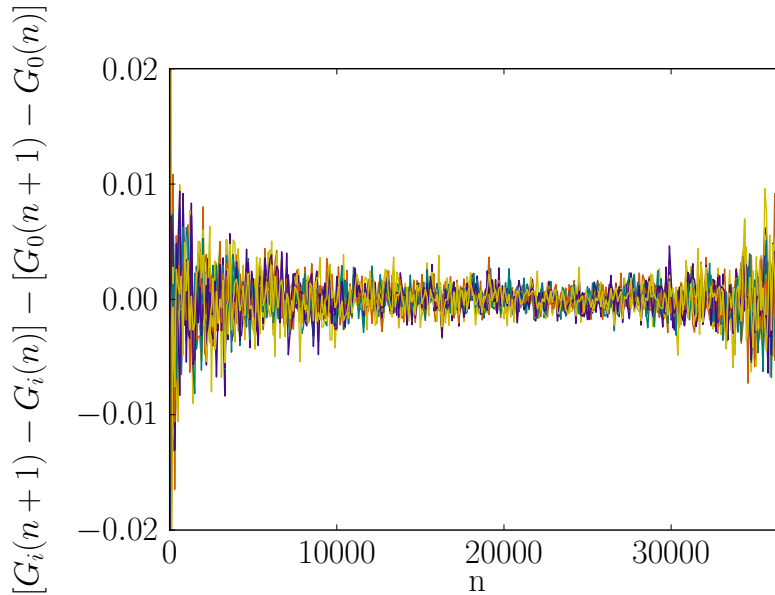
Figure 4.7: (Reproduced with permission from [65]) This graph shows the discrete derivative of the density of states from sixteen simulations of the 32×32 XXZ model. The value of the derivative is normalized to the value of one of the simulations. The error in the derivative is randomly distributed, but since we integrate this value to get $G(n)$, the value of $G(n)$ for large n can differ significantly between different simulations. In the fixed operator number representation for the quantum Wang-Landau there is also an enhancement of the error when the operator list is almost empty or almost full. Switching to a variable length representation has been shown to reduce this error [72].

Wang-Landau.

    As mentioned when discussing the convergence of the Wang-Landau algorithm, early errors can quickly become saturated and take a long time to recover from. Fig. 4.7 shows the derivative of the density of states for sixteen simulations normalized to the result of the one of the simulations. What one should gather from this figure is that the derivative of the density of states for all the simulations converge to a common value up to random statistical noise. This means that the sampled density of states $G_i(n)$ will have a larger spread for larger n, and results at lower temperature will tend to have larger uncertainty than results at higher temperature. Accurate results at small n are a result from the fact that we have an analytic way to pin $G(0)$ from the high temperature entropy, but if it were

possible to exactly know $G(n)$ for any other value of n, it would be possible to generate highly accurate results at a temperature range away from high temperature.

One difficulty with Wang-Landau type approaches is that they, by definition, have to sample all of the energies and equal number of times. The main difficulty with this is, ideally, the simulation acts as a random walk in one dimension, meaning that the number of steps it takes to make a round trip grows with the square of the length of the interval. For a fixed temperature cutoff, the interval size grows with the number of Hamiltonian elements in the system, or $\mathcal{O}(N)$ for short range Hamiltonians, meaning simulation time at least grows as $\mathcal{O}(N^2)$. Recent papers suggest cutting up the density of states in to small chunks and restricting the Wang-Landau to only operate in that range [72]. By restricting walkers to small ranges they make round trips more frequently and converge faster. The full density of states can be built up by having the ranges overlap slightly and recalling that the difference between adjacent $G(n)$ is converged, rather than the absolute value. Using this multiple ranges can be stitched together to form the full density of states, which can then be used to calculate any observables.

Finally, there is one potential extension to Wang-Landau that might allow the use of a truncated density of states for all temperatures. In finite size systems with discrete degrees of freedom there always remains a finite size gap from the ground state. This means that at very low temperature there is an analytic expression for the energy by considering only the ground state and the low energy excited states. Since this expression is valid precisely in the region where the Wang-Landau fails due to the cutoff, it could be used to analytically extend $G(n)$.

In the next chapter we will introduce the ratio trick. Fundamentally, the ratio trick allows for the direct calculation of the ratio of two partition functions, the quantity needed to calculate the Rényi entanglement entropy. This direct calculation does not require thermodynamic integration, but the implementation is not as general as the methods discussed in Ch. 3. Specifically, we derive general methods for discrete classical systems, but general quantum systems require a unique algorithm that depends on the particular implementation of the quantum Monte Carlo.

Chapter 5

Finite temperature "ratio trick"

Since the concept of measuring the Rényi entanglement entropy in numerics first emerged in 2010, there have been a variety of methods in which to measure it [74, 75, 59, 60]. The integration approaches in Ch. 3 and the Wang-Landau approaches in Ch. 4 represent methods for calculating the full partition function at any finite temperature. In this chapter we will discuss methods that allow direct calculation of the ratios of partition functions *without knowing the magnitude of either*. This shares some conceptual overlap with the Wang-Landau methods for the density of states, but it is applied in a very different direction.

Conceptually, it is not more useful to know the ratio of two partition functions over knowing the absolute value of each. The reason the finite temperature ratio trick is useful is that the full knowledge of the partition function is unnecessary, and by giving up this information the simulation is much more efficient. Specifically, the problem of thermodynamic integration is avoided, as is the need to sample all states equally. We benefit from the efficiency of a normal simulation without needing a large set of simulations to integrate over[1]. These algorithms represent the current state of the art for accurate calculations of the Rényi entanglement entropy, and are well suited for studying criticality and phase transitions where finite size scaling is typically more difficult.

Using the ratio trick outlined below, we demonstrate its effectiveness by studying the thermal Rényi entanglement entropy of the one dimensional transverse field Ising model. We show that above the critical point, in the quantum critical fan, we are able to detect universal collapse using the finite size and finite temperature form of the REE derived from field theory calculations. The agreement between the results is remarkable, and

---

[1]More accurately, we trade an integration in temperature for one in space.

shows how even with thermal results the REE can be used to probe the universality of zero temperature quantum critical points.

5.1 Concepts

At the core of all ratio trick methods is looking at ways to directly compute

$$\frac{Z[A, \alpha, \beta]}{Z^\alpha[\beta]} \tag{5.1}$$

where α is the Rényi index, we will restrict our discussion to $\alpha = 2$, and β is the inverse temperature. $Z[A, \alpha, \beta]$ here represents the modified partition function in Eq. (3.7), where A represents some choices of the region A. To calculate the Rényi entanglement entropy one needs to be able to calculate Eq. (5.1). This is usually done by decomposing the above ratio into

$$\frac{Z[A, \alpha, \beta]}{Z^\alpha[\beta]} = \prod_{i=0}^{N-1} \frac{Z[A_{i+1}, \alpha, \beta]}{Z[A_i, \alpha, \beta]} \tag{5.2}$$

where we define $A_N = A$ and $A_0 = \emptyset$ (the empty set, denoting a region containing no spins). We should also note that, by definition, $Z[\emptyset, \alpha, \beta] = Z^\alpha[\beta]$, since if A is an empty region then the α replicas are non-interacting. The reason for decomposing the full ratio into the above form is that the convergence properties of a ratio where A_{i+1} and A_i only differ slightly are significantly better than those for which they differ by a large amount (see Fig. 5.5 for a more complete discussion).

This ratio can be calculated in classical and in quantum systems, although the details of the implementation are not completely general. Often significant progress is made when specific implementation allows for the efficient calculation of this ratio in new models. Below we will discuss a variety of Monte Carlo implementations for both classical and quantum models, as well as potential considerations for further improvements.

5.2 Classical ratio trick

In general, classical ratio tricks work by generating a representative state of the partition function $Z[A_i, 2, \beta]$ and attempting to calculate

$$\frac{Z[A_{i+1}, 2, \beta]}{Z[A_i, 2, \beta]} \tag{5.3}$$

71

in some way. Let us take the most naive algorithm, which although slow, works. If we assume A_{i+1} is simply the same as region A_i except it contains a single extra spin S_{i+1}, then in the classical modified simulation we can move from $Z[A_i, 2, \beta]$ to $Z[A_{i+1}, 2, \beta]$ if that spin matches between the two replicas. To reiterate, if the spin S_{i+1} matches between the two replicas, the particular configuration is a valid state of both partition functions. In this way, a Monte Carlo that runs in the following way can estimate Eq. (5.3):

1. Generate a new random state using single spin flips, cluster moves, or any other efficient means.

2. If currently in the modified simulation defined by $Z[A_i, 2, \beta]$, attempt to add S_{i+1} to the cluster A_i and change partition functions; this move succeeds if the configuration of S_{i+1} is identical in both replicas, and fails if the configuration differs. After this skip to step 4.

3. If currently in the modified simulation defined by $Z[A_{i+1}, 2, \beta]$, remove S_{i+1} from the cluster A_{i+1} and change partition functions; this move is always accepted.

4. If the current state is in the modified partition function $Z[A_i, 2, \beta]$, add one to a counter N_i, otherwise add one to N_{i+1}, then go to step 1.

5. After many steps, the ratio in Eq. (5.3) is estimated by $\frac{N_{i+1}}{N_i}$

This algorithm almost seems too simplistic, but implementation shows that it does reproduce the correct REE. If one is attempting ratios in which one adds more than one spin, then *all* the spins have to match for the move to be accepted. In this light, it is clear why a large change in the region from A_i to A_{i+1} would result in poor statistics. Imagine the Ising model at high temperature, in which case the chance of N addition spins matching is roughly 2^{-N}. This means that our Monte Carlo algorithm is attempting to sample an event with a very rare frequency, and as such we should only expect good statistics when our sampling is much larger than the inverse of that frequency. There are two improvements in efficiency to this simple approach that develop additional conceptual tools for other attempts to calculate the ratio trick.

In this first example the simulation changes what the region size is as part of the simulation, but the interesting behaviour only occurs when increasing the region size while decreasing the region size is always a trivial move. In this sense all the information about the ratio in Eq. (5.3) is contained in how often we can increase the region size, and if we can directly measure this probability we can expect a large increase in speed. The most naive algorithm along this line of thinking would operate in the following way:
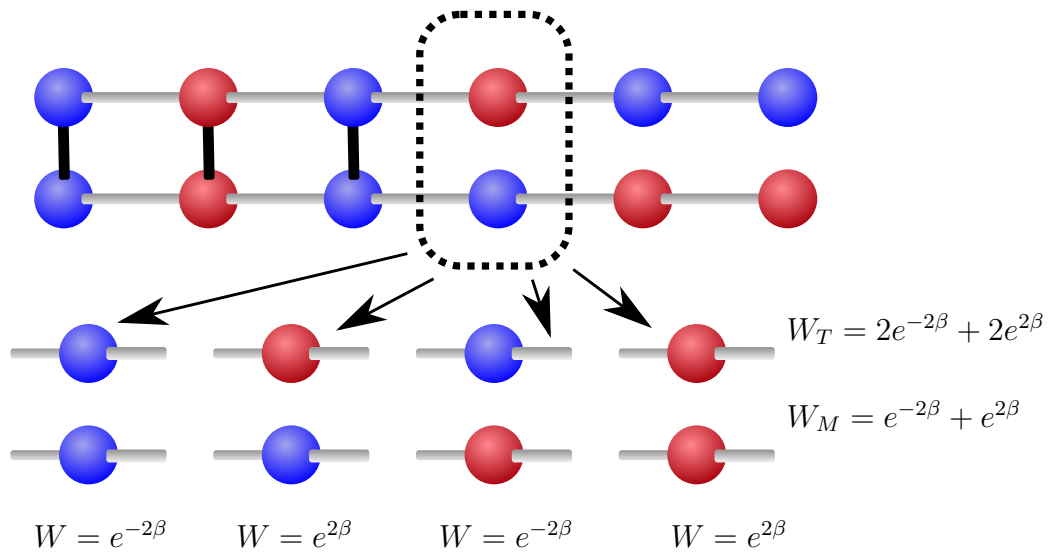
Figure 5.1: An improvement to classical calculation of the ratio trick by considering an ensemble of states closely related to the originally sampled state. Each of the weights W are calculated by examining the energy of the local configuration. Using this we can calculate the total weight, W_T, and the weight in cases where the spins match, W_M. The ratio gives us the (local) probability of accepting an update to increase the region size. The cost of such a calculation grows exponentially in the region size, but we can use many ratio tricks to ensure region growth is bounded in each step.

1. Generate a new random state using single spin flips, cluster moves, or any other efficient means.

2. Consider the four states corresponding to all the (weighted) possible configurations of spin S_{i+1} in both replicas, and calculate their weights up to an arbitrary constant. See Fig. 5.1 for an example.

3. Calculate the weight of all matching configurations, W_M, and divide by the weight of all configurations considered, W_T. Add W_M/W_T to R and go to step 1.

4. R/N, where N is the total number of weights sampled, gives an estimate of Eq. (5.3) after many steps.

Essentially this method differs from the first by sampling the possibility of increasing the region size given some fixed configuration of bath spins, and these bath spins are sampled over using the Monte Carlo. In principle, if we sampled over all possible configurations of the entire lattice, one wouldn't need to sample at all, and could exactly calculate the ratio in a single step. Enumerating over all the configurations of the entire lattice is an exponentially long process, and fundamentally the opposite of importance sampling in Monte Carlo, so in this procedure we use the exact procedure to target the type of states we are interested without biasing the simulation as a whole. There is one more improvement to the above train of thought in which we can eliminate the exponential growth of the exact calculation in some special cases.

If the region A_{i+1} only differs by A_i by the addition of some one dimensional strip of spins, then we can calculate the weights W_M and W_T using a transfer matrix approach. This approach comes from methods used in simulations of digital memory in computation, where the channel capacity, or ability of a medium to function as a storage device, is calculated using this method [76]. Let us refer to the strip of spins as $S_{\{i\}}$, the procedure works exactly the same as above, except since the row of spins is one dimensional we can calculate the weight over all configurations by taking the trace over the product of a set of matrices. There are two partition functions we calculate using the transfer matrices, one with the spins connected and one as the product of two separate strings of spins that are completely unrelated. Fig. 5.2 shows these two scenarios, and like the previous example the rest of the system acts as a static bath during this calculation. Appendix B.2 has an example code in which we demonstrate the use of transfer matrices to calculate the finite temperature Rényi entropy in the two dimensional Ising model. Fig. 5.3 shows the result of running the test code for the parameters given, reproducing the general behaviour and the proper low temperature limiting value of $\ln(2)$.
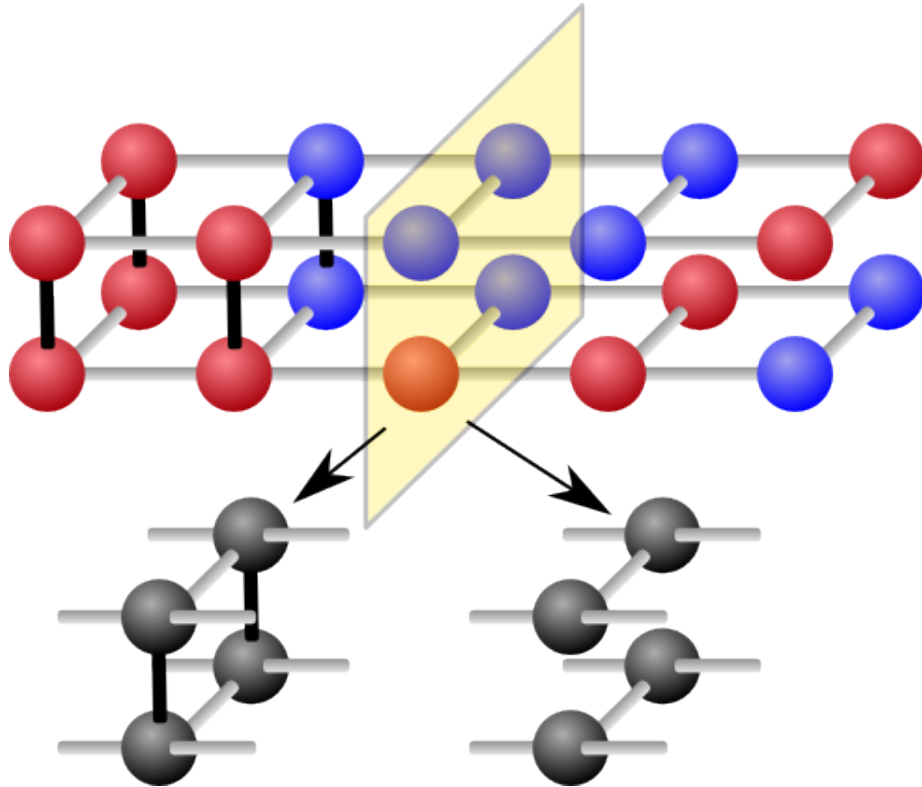
Figure 5.2: The modified simulation of A_i is shown, the with strip of spins $S_{\{i\}}$ highlighted in the box. We use a transfer matrix approach to calculate the two weights over this row of spins, one case in which they are (lower right) unconnected and one in which they are (lower left) connected. This transfer matrix is used to calculate the partition function assuming the rest of the system as a fixed bath. The ratio of these two partition functions approximates Eq. (5.3) after sampling over many configurations of the surrounding lattice.
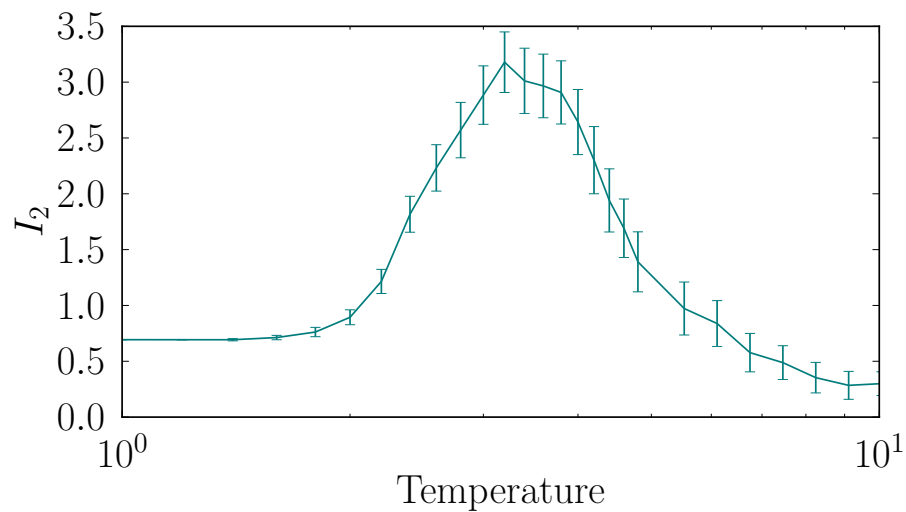
Figure 5.3: A short test using transfer matrices to calculate the total mutual information in the 2D Ising model for a 10×10 lattice. Notice that unlike the integration approach, error can decrease as we go to lower temperature, as each temperature point is generated separately and independently from every other point. When our primary interest is low temperature mutual information this method may be much more efficient than the integration approach.

The general concept of both of these improvements is that, instead of sampling a single state and examining if it satisfies the conditions for changing region size, we sample over an ensemble of states generated from a reference state, and measure the probability of changing region size within that ensemble. This concept of sampling over a large set of states at once rather than a single state is facilitated by the existence of methods such as transfer matrix that allow the calculation of the probability in a straightforward way. In the next section describing the same techniques for quantum Monte Carlo this concept of sampling over an ensemble of states will arise again, and it will be at the core of an efficient algorithm for calculating the Rényi entropy for the transverse field Ising model.

5.3 Quantum ratio trick

For the modified quantum simulation (see Sec. 3.3) we will start by explaining the trivial implementation of the ratio trick, Eq. (5.3), and then move on to a particular implementation for the transverse field Ising model.

The trivial implementation of the ratio trick will be explained in the same way as before, by examining the result of attempting to add a single spin, S_{i+1}, to the region A. Implementations of this have been shown to accurately reproduce the Rényi entanglement entropy for a variety of quantum models [77]. In the classical simulation, there are two copies of the spin S_{i+1}, one in each replica, and if they matched we are allowed to add the spin to region A, otherwise it is a forbidden move. In the quantum language, we have the stochastic series expansion representation, and the configuration of S_{i+1} changes depending where in that expansion we look. If we compare two modified simulations, $Z[A_i, 2, \beta]$ and $Z[A_{i+1}, 2, \beta]$, that differ by only a single spin in region A, then the only difference in the set of allowed states these two partition functions can represent is defined by the boundary conditions on the spin S_{i+1}. In the case of the two replica modified simulation, the condition for adding *or removing* a spin from region A is the same: the configuration at the bottom slice of each expansion must match between the replicas. Using Fig. 3.6 as a reference, in the case of spins already in region A, this means that the configuration at $A0$ and $A1$ match, while for spins not in region A this means that the configurations at $\bar{A}0$ and $\bar{A}1$ match.

Using this we can again construct a simulation using the following basic algorithm:

1. Generate a new random state using diagonal updates and off-diagonal updates.

2. If the current state is in the modified simulation defined by $Z[A_i, 2, \beta]$, attempt to add S_{i+1} to the cluster A_i and change partition functions.

3. If instead the current state is in the modified simulation defined by $Z[A_{i+1}, 2, \beta]$, remove S_{i+1} from the cluster A_{i+1} and change partition functions.

4. If currently the modified partition function $Z[A_i, 2, \beta]$, add one to a counter N_i, otherwise add one to N_{i+1}, then go to step 1.

5. After many steps, the ratio in Eq. (5.3) is estimated by $\frac{N_{i+1}}{N_i}$ [77].

The difference when compared to the classical analogue is the both the addition and removal of spins from region A is non-trivial, and in fact the condition is symmetric. Even though the condition for increasing or decreasing the region size is the same, the probability of acceptance is not identical since we are drawing states from different partition functions.

The conceptual hurdle comes when attempting to examine the modified quantum simulation cell and examining it in such a way that lets one imagine a particular configuration as a representation of some large ensemble of states in some tractable way. This concept is necessary for an efficient calculation of the ratio of the partition functions, similar to how the transfer matrix approach in classical systems let us consider a large set of related states. The transverse field Ising model is fortunate in this regard, as the off-diagonal updates precisely give us a route to this concept. If we examine the off-diagonal update, Fig. 2.1, it is designed so that every distinct cluster can be flipped independently, each one generating a new state with equal weight. This means that in a system with N_L total clusters, this state is easily connected to an ensemble of 2^{N_L} states, each generated by choosing whether or not to flip each of the clusters.

Even though we are connected to 2^{N_L} states, the difficulty comes in calculating the probability of adding a single or set of spins, $S_{\{i\}}$, to region A or removing that set from region A. Fig. 5.4 shows a representative example of the interface between two replicas in a modified quantum simulation cell. For a given configuration there is a well defined loop structure. Restricting ourselves to the boundary, and furthermore to the spins of interest we are attempting to add to region A, there are a specific number of total states possible given the loop configuration, 2^5 in the example. For each pair of spins, a connection is only allowed when spins on both side of the boundary match, which can be achieved by enforcing that the spin be a part of the same cluster[2]. Once the new clusters are formed the number of unique clusters is decreased or remains the same due to the fact that we assume reconnecting only combines clusters. The reason we assume reconnecting cannot

---

[2]Even in the case where the spins are in the opposite configuration initially, we can simply imagine a different state in which they match by flipping one of the loops, which is topologically identical to the representative state we are using.
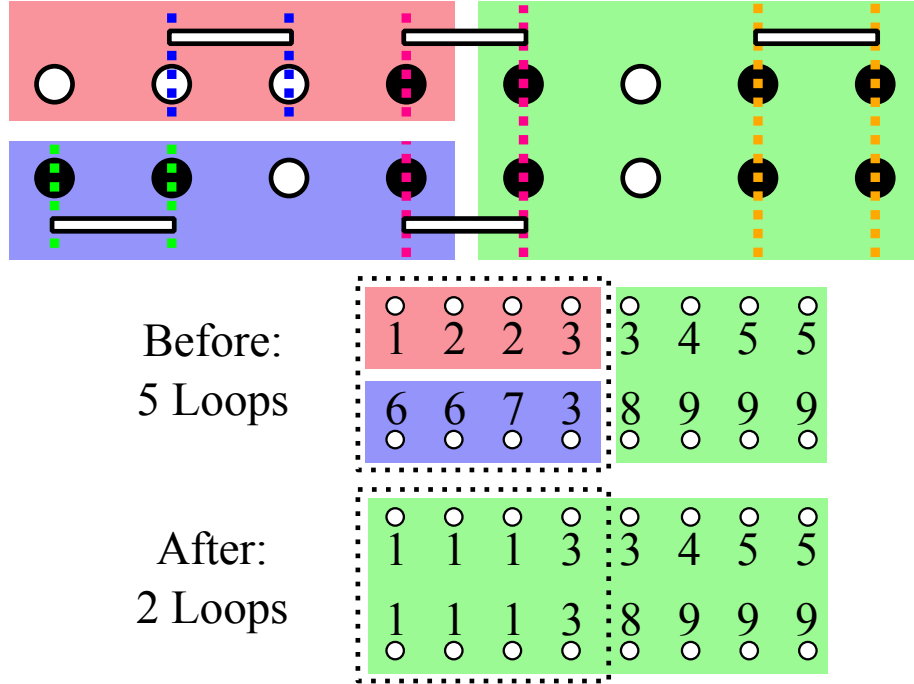
Figure 5.4: The probability of increasing the region size for a given configuration of the transverse field Ising model. For a given loop configuration we can examine what fraction of the total set of states generated by all possible configurations of all clusters will allow an increase in region A. The only relevant loops are those that intersect the boundaries where the difference between region A and the new region, in this case the region where all spins are in region A, is actually realized. In this case the number of distinct loops intersecting the boundary is 5. Matching boundary conditions is identical to enforcing that spins on each side be a part of the same cluster, and counting the number of clusters after we have two unique clusters (or loops) crossing the boundary of interest. Using this idea, we get that the fraction of allowable overlap states as 2^{2-5} for this particular configuration. Decreasing the size of region A is identical in practice, except we operate on the spins being removed from the region rather than those being added. Note that the cluster index is from $|\alpha_0\rangle$ in each replica, so the clusters won't necessarily match in region A even though the boundary conditions connect the replicas.

disconnect clusters is that we are examining the probability of changing the region size rather than the number of possible states in the reconfigured topology[3]. The reduced number of clusters gives us the fraction of available states after enforcing the connection, and since all states had the same weight we get an estimate of the probability of this move, and in the case of Fig. 5.4 this is 2^{-3}.

The algorithm for calculating Eq. (5.3) would proceed as follows

1. Generate a new random state using diagonal updates and off-diagonal updates.

2. During the last off-diagonal update, store the loop numbering at the boundary between the two replicas. The boundary is defined as the loop configuration at $|\alpha_0\rangle$ for both replicas. The number of unique loops at this point is L_0.

3. Connect the set of spins $S_{\{i\}}$. This means for each spin that is added (or removed) from the cluster identify the clusters attached to each spin as being the same cluster. Once completed, we have a reduced number of clusters L_s.

4. Measure $2^{L_s-L_0}$, and go to step 1.

5. Defining $Z_r = \langle 2^{L_s-L_0} \rangle$, after many steps, Eq. (5.3) is approximated by Z_r/Z_{r-1}

Again, conceptually this loop counting before and after changing the topology tells us what fraction of states allow for the reconnections to occur, and allow sampling of the ratio of partition functions at finite temperature. Here Z_{r-1} is the calculated expectation value of the reverse move of Z_r. That is, instead of starting at region A_i and adding the spins $S_{\{i\}}$ to get to region A_{i+1}, we start at region A_{i+1} and remove the spins $S_{\{i\}}$ from it to reach region A_i. In the classical case[4] $Z_{r-1} = 1$, but as mentioned earlier this move has a probability of acceptance in the quantum case. Sampling of Z_{r-1} is identical to sampling Z_r, except that the reference and target regions of consideration are reversed. One can relate this to measuring $\frac{N_{i+1}}{N_i}$ by viewing Z_r as the transition rate to state labelled by N_{i+1} from the state labelled by N_i, but since these transition rates are always fractional, we gain accuracy by measuring that fraction directly over attempting the high rejection move of actually changing the region size.

With this construction it seems reasonable to ask why a large region size cannot be done at once, as there is nothing specific to this algorithm that requires the region be one dimensional or even contain a large or small number of spins. Practically, the reason for

---

[3]This is an important distinction when examining zero temperature methods later.
[4]Assuming that Z_r is the move increasing region size and Z_{r-1} is decreasing region size.
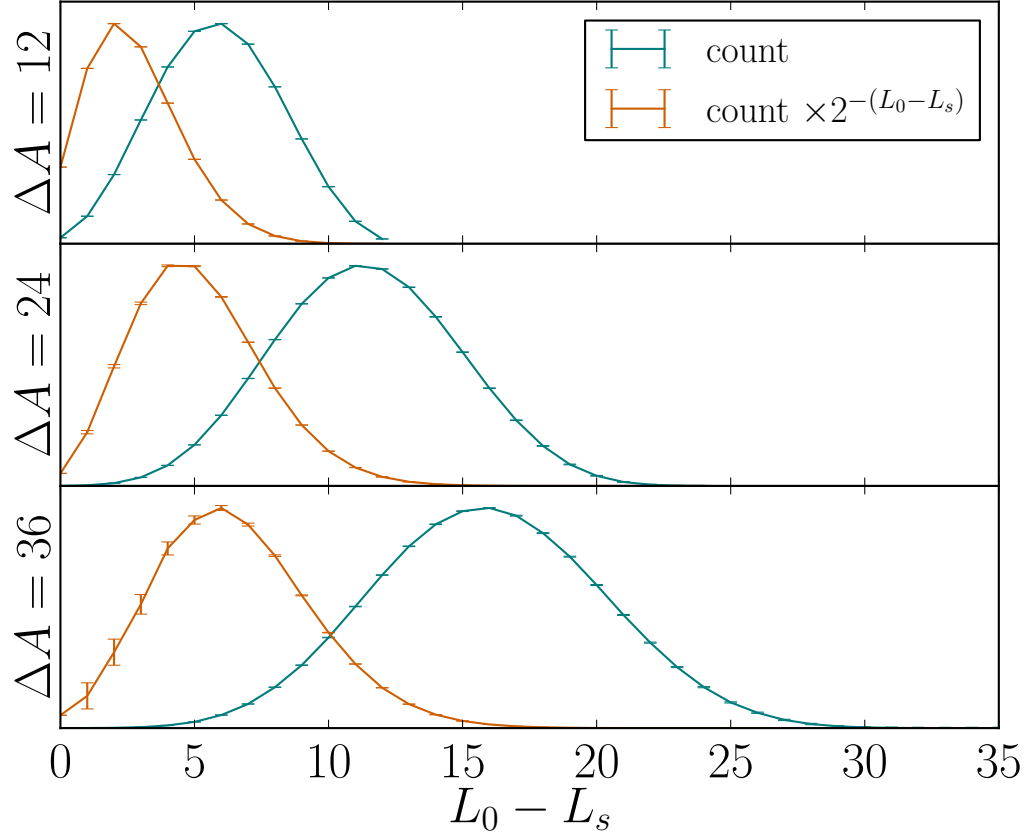
Figure 5.5:   The distribution of $L_0 - L_s$ for different sized ratio tricks for the one dimensional transverse field Ising model at $h = J = 1$, $L = 128$, and $\beta = 100$. The three graphs show the distribution of $L_0 - L_s$ and $2^{-(L_0 - L_s)}$ for ratio tricks where the change in region size, ΔA, is twelve, twenty four, and thirty six spins, normalized to their peak value. Notice for the larger changes in region size that, although the statistics of $L_0 - L_s$ are quite smooth, the factor of $2^{-(L_0 - L_s)}$ means that the distribution is dominated by rare events in the tail of the distribution where statistics are much worse. It is for this reason that we restrict ourselves to small changes in region size.

using small changes in region size is mostly empirical. From a technical standpoint, as long as both J and h are not zero, for any size region there is a non-zero probability of sampling states in which $2^{L_s - L_0}$ ranges from 2^0 to 2^{-N_s} where N_s is the number of spins being added or removed from region A. Even if $L_s - L_0$ is a smooth function with a well defined mean, $2^{L_s - L_0}$ may be very noisy when the domain of values is large. This is because if the mean of $L_s - L_0$ is large and negative, the function will be strongly affected by the tail in the positive direction. In this sense, the larger the domain and magnitude of the mean, the more difficult it is to converge statistics on the ratio. Fig. 5.5 shows the distribution of $L_s - L_0$ the 1D transverse field Ising model.

5.4   Finite temperature Ising model at criticality

As part of the zero temperature projector quantum Monte Carlo work (see Ch. 6), the finite temperature method was used to compare results for the one dimensional transverse field Ising model at criticality. Even though the finite temperature code is not ideally suited for zero temperature results, for small systems the zero temperature behaviour is reproduced when β is taken large enough. In this way, the comparison between the two methods is possible. This serves as one check to ensure that the results from simulation are reliable, as the codes were developed and implemented separately, and they were unlikely to share any errors.

As an additional check, we examined the Rényi entanglement entropy of the one dimensional transverse field Ising model at finite temperature. At zero temperature a conformal field theory assumes a scaling of the form [10]

$$S_\alpha(A) = \frac{c}{6}\left(1 + \frac{1}{\alpha}\right)\log\left(\frac{x}{a}\right) + c'_\alpha, \tag{5.4}$$

where α is the Rényi index, c is the central charge, A is a region of length x, a is the lattice spacing of the model and c'_α is a non-universal constant. The one dimensional transverse field Ising model corresponds to a conformal theory with a central charge of $c = 1/2$.

Fig. 5.6 shows the scaling of the second Rényi entropy for one dimensional systems at a low temperature. The form of Eq. (5.4) fits well, that is to say the data collapses well to the shape of $\log(x)$. Deviations occur when the region size approaches $L/2$, which is the right end of each data set. Considering the above equation is only valid in the thermodynamic limit at zero temperature, this deviation is expected.
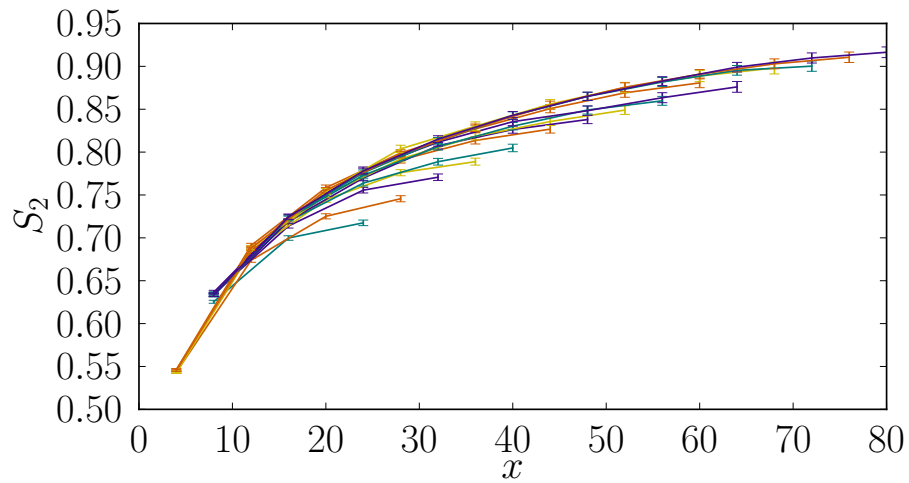
Figure 5.6: The second Rényi entropy of the one dimensional transverse field Ising model at $\beta = 100$ and $J/h = 1$. System sizes range from $L = 48$ to $L = 160$, and region A size from $x = 4$ to $x = L/2$ for each system size, in steps of four or eight, depending on system size. Notice how the results tend to collapse until $x/L \approx 0.5$ (the right end of each line). This makes sense, as we expect results to follow the form of Eq. (5.4) as long as the region size is small compared to the size of the entire system and the region size is smaller than the inverse temperature (up to some scaling factor).

Moving beyond this limit, we can examine the form of the entropy at finite temperature, but infinite size, get the form [10]

$$S_\alpha(A) = \frac{c}{6}\left(1 + \frac{1}{\alpha}\right)\log\left(\frac{\beta}{\pi a}\sinh\frac{\pi x}{\beta}\right) + c'. \qquad (5.5)$$

For small values of x/β this reduces to the form of Eq. (5.4), as one should expect.

Using Eq. (5.5) we can attempt to collapse the data by rewriting it as

$$S_2(A) - \frac{c}{4}\log\beta = \frac{c}{4}\log\left(\sinh\frac{x}{\beta}\right) + C', \qquad (5.6)$$

where C' has absorbed various constants. In this form, knowing the central charge c allows us to collapse all system sizes L, cut lengths x, and β as long as x/L is small enough. Fig. 5.7 shows this collapse, and using $c = 0.5$ for the transverse field Ising model gives a remarkable collapse over a large range of system size, temperature and cut size. We emphasize that this collapse has no free parameters and demonstrates that a finite temperature, finite size lattice model agrees remarkably well with predictions from continuum conformal field theory.

The effect of using the full data set is shown in Fig. 5.8. The collapse is still reasonable, but if data for strips larger than half were included a very large deviation would be apparent, due to the decreasing S_2 with increasing region size.

In the limit of zero temperature and finite size, there are also predictions of the entanglement entropy. The general form in that case can be written as

$$S_\alpha(A) = \frac{c}{6}\left(1 + \frac{1}{\alpha}\right)\log\left(\frac{L}{\pi a}\sin\frac{\pi x}{L}\right) + c'. \qquad (5.7)$$

This looks almost identical to Eq. (5.5) except that L replaces β and sin replaces sinh. This form also reduces to Eq. (5.4) when x/L is small.

We can attempt collapse for Eq. (5.7) by writing it as

$$S_2(A) - \frac{c}{4}\log L = \frac{c}{4}\log\left(\sin\frac{x}{\beta}\right) + C', \qquad (5.8)$$

where again C' has absorbed various constants. Fig. 5.9 shows this collapse using all the data for which $x/\beta \leq 1/8$, again using the same $c = 0.5$ for the transverse field Ising model critical point. Again, this collapse has no free parameters, and is simply fitting to the zero temperature finite size form of the scaling.

If we examine all of the temperature data in this manner, as opposed to only the low temperature data, then we get Fig. 5.10. In this case there is no collapse at all, with the high temperature data vastly deviating from the low temperature result.

Figure 5.7: The second Rényi entropy of the one dimensional transverse field Ising model at $h/J = 1$ when $x/L \leq 1/4$, with $L = 48$ to $L = 160$, and $\beta = 1$ to $\beta = 400$. Even though this represents systems quite far from the thermodynamic limit, the collapse of the data when plotted as a function of x/β is quite good. This suggests that corrections to Eq. (5.5) are not very large. The constant subtraction uses $c = 0.5$, the central charge for the one dimensional transverse field Ising model, and this collapse has no free parameters except for the cutoff in x/L.
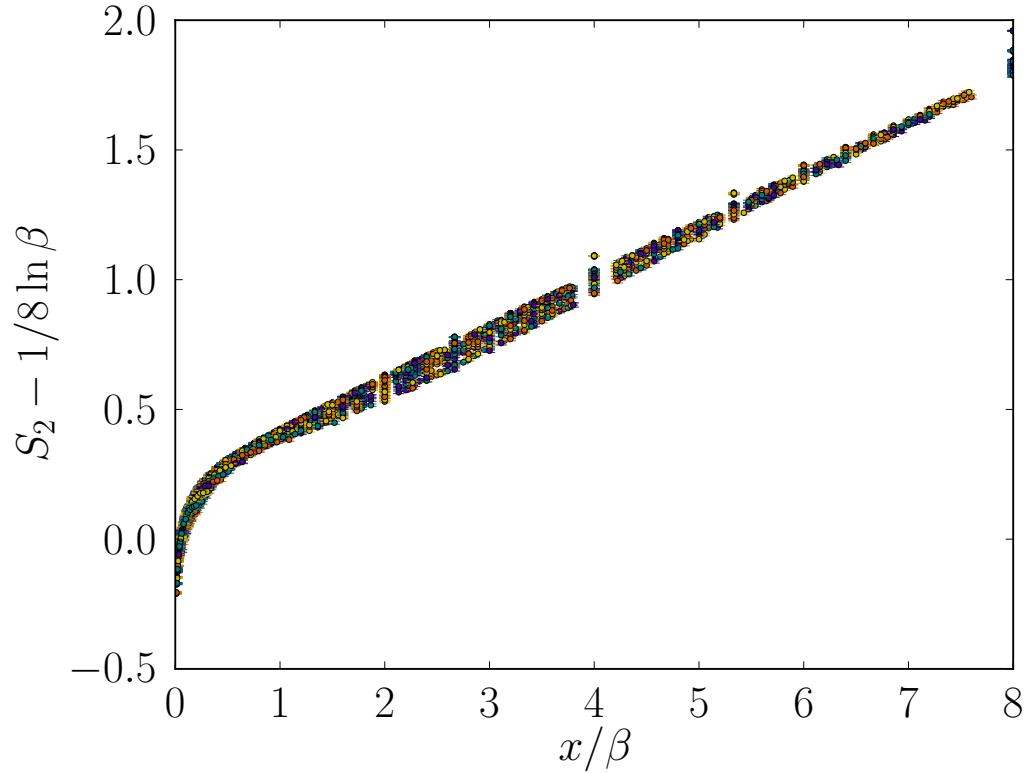
Figure 5.8: The second Rényi entropy of the one dimensional transverse field Ising model at $h/J = 1$ with $L = 48$ to $L = 160$, and $\beta = 1$ to $\beta = 400$. Using the full data up to $x = L/2$ the collapse only deviates slightly, suggesting that the cutoff in x/L for which the system is still in the infinite size limit is quite large. Including larger than half strips would show a strong deviation, as S_2 decreases when the region size grows larger than half the system, as $S(x) = S(L - x)$ due to symmetry.
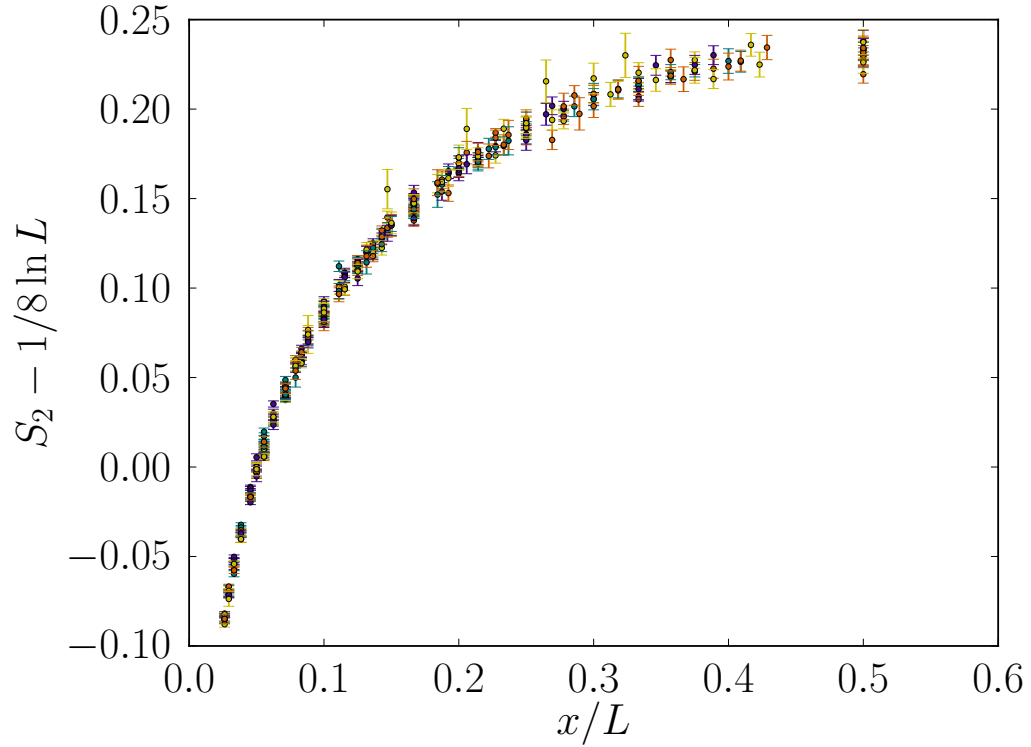
Figure 5.9: The second Rényi entropy of the one dimensional transverse field Ising model at $h/J = 1$ when $x/\beta \leq 1/8$, with $L = 48$ to $L = 160$, and $\beta = 1$ to $\beta = 400$. The collapse is good here as well, although not as impressive as Fig. 5.7. The corrections to Eq. (5.7) are likely larger than in Eq. (5.5), but there is likely a coefficient of conversion between lattice spacings x and inverse temperature β responsible for this. The constant subtraction uses $c = 0.5$, the central charge for the one dimensional transverse field Ising model, and this collapse has no free parameters except for the cutoff in x/β.
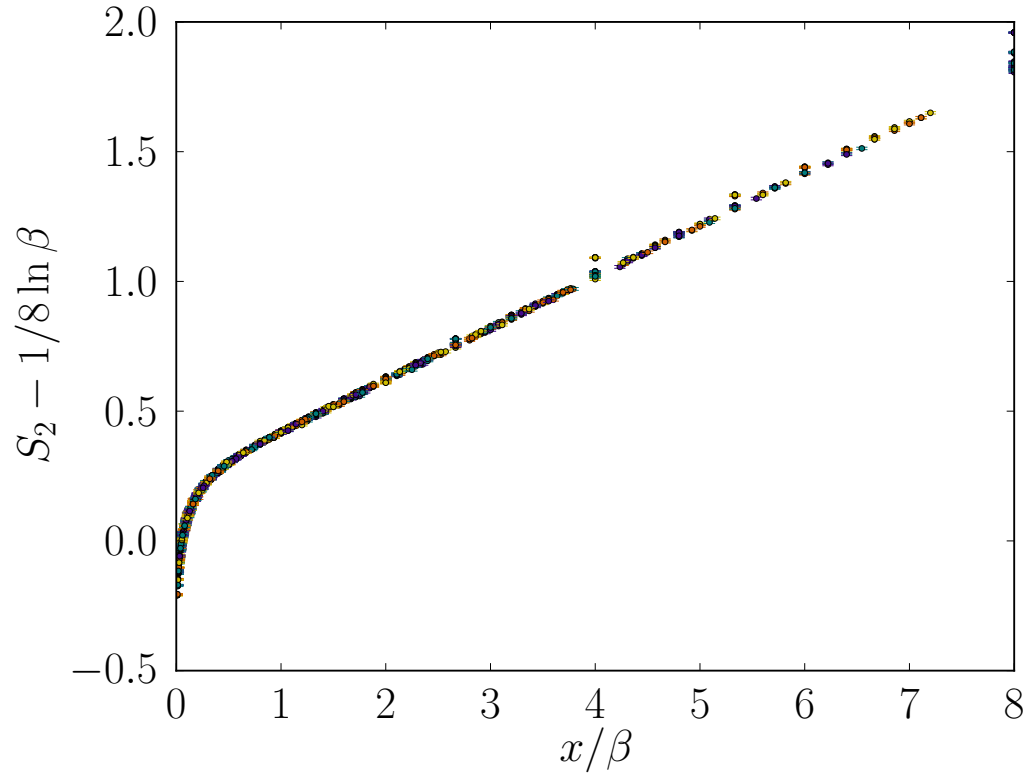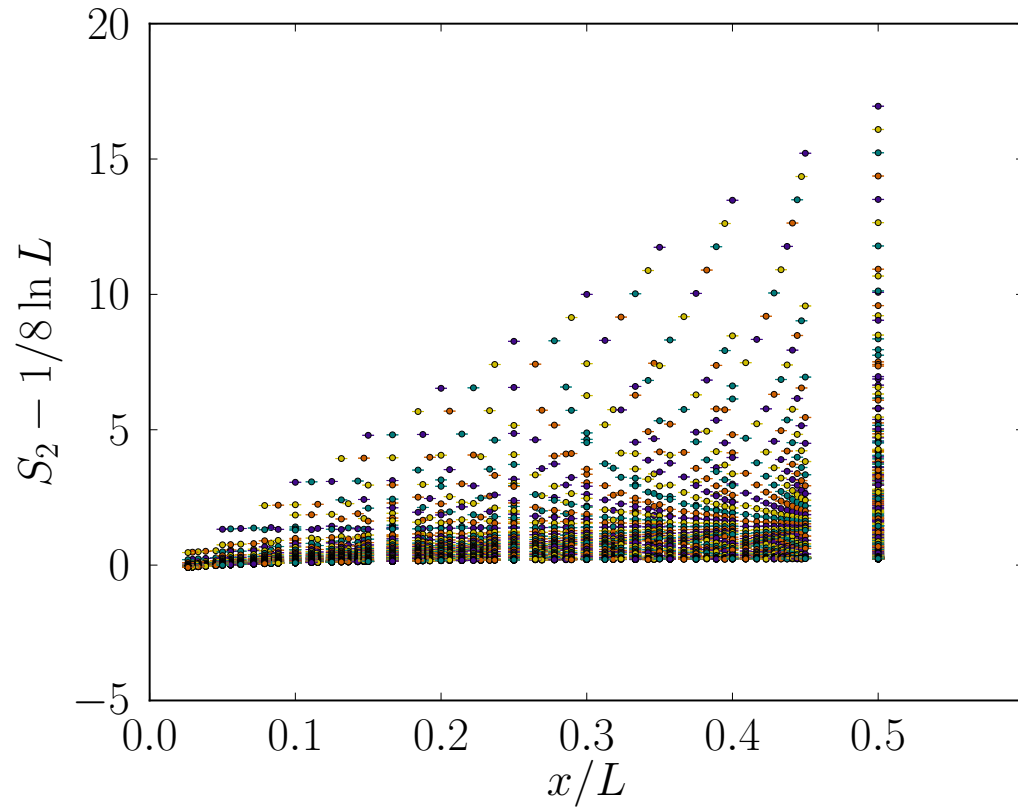
Figure 5.10: The second Rényi entropy of the one dimensional transverse field Ising model at $h/J = 1$ with $L = 48$ to $L = 160$, and $\beta = 1$ to $\beta = 400$. In this case, without restricting the data there is no collapse whatsoever; notice the scale of this plot is much different than that of Fig. 5.9.

5.5 Discussion

The important point from this analysis is that using a lattice model and Monte Carlo methods we are able to very accurately reproduce the thermal Rényi entropy in both the limit of small region size and low temperature. Despite the availability of a finite temperature and finite size form of the Rényi entropy [10], it is formally only valid for $c = 1$ systems. In the case of the infinite size form, this is reproduced well by only using data where the strip size is much smaller than the region size. When trying to collapse to the infinite β form, we again find good agreement when the ratio of the strip size and inverse temperature is small. Although there is no prediction for what the cutoff should be, this procedure can easily be repeated for other systems. In addition, the constant subtraction needed for the collapse is itself a measure of the universal central charge. With high enough accuracy this means thermal results well in to the quantum critical fan can be used to extract the central charge. This extraction is more accurate when the rate of change with S_2 with x/β is high, something that occurs at low temperature, but in cases where it is very difficult to extract low temperature results, using high accuracy thermal results is a viable alternative.

In the next chapter we discuss the scaling behaviour at the zero temperature critical point of the two dimensional square lattice transverse field Ising model. Zero temperature methods require a small conceptual step from finite temperature methods, but the gains in speed allow us to study the complex geometrical properties of the Rényi entanglement entropy. This geometric dependence at criticality shows insight into two dimensional universality, and demonstrates quantities of interest that should be considered in any further studies.

Chapter 6

Zero temperature 2D transverse field Ising model at criticality

In addition to finite temperature quantum Monte Carlo methods, we use a zero temperature Quantum Monte Carlo method [78] to examine the Rényi entanglement entropy in the two dimensional transverse field Ising model at its quantum critical point. The zero temperature method has the advantage of faster convergence when examining zero temperature properties compared to a finite temperature method pushed to very large β. The drawback is that, when a finite temperature method uses a smaller number of operators, the result can be related to a smaller β, where as when a zero temperature method uses too few operators it is simply underconverged, and the results cannot (or rather should not) be related to any physical quantity.

We will discuss the importance of the zero temperature entropy for this model, then describe the zero temperature algorithm and its application to examining the shape dependence and universal terms in the Rényi entanglement entropy for the two dimensional transverse field Ising model at criticality. Remarkably, we find that the one dimensional shape function, $\log(\sin(\pi x/L))$, does not seem to have a universal coefficient when examining strip geometries in two dimensions. This function has been used in various studies to examine the shape dependence of the subleading term in the REE in two dimensions, but never over the range of sizes in this study [79]. Instead, a shape function from the quantum Lifshitz free field theory (see Eq. (6.5)) shows a better fit along with a convergent coefficient when finite size scaling. This suggests this alternate shape function is more appropriate for the 2D TFIM, and more importantly this might represent a universal shape function in two dimensions.

6.1 Zero temperature entropy at criticality

In the case of two dimensional systems without a Fermi surface [80, 81], it is believed that the ground state entanglement obeys an area law [82, 83], that is the leading order term in the entanglement scales proportional to the boundary separating two regions. In gapped systems the entanglement is short ranged, and the area law is trivially satisfied since any entanglement is restricted to the volume near the boundary of the regions, as shown in Fig. 1.2. In the case of gapless systems, the assumption is that all length scales contribute equally to the entanglement as a result of the scale invariance typically seen at criticality. If we imagine a realspace renormalization procedure, the number of bonds crossing the boundary should scale like the logarithm of the region length in one dimension and the surface area of the region in two dimensions. Fig. 1.2 shows how this entanglement counting argument follows from an assumption of scale invariance in entanglement, and how the log correction to the area law in one dimension does not occur in two dimensions under the same assumption. Although there are a great deal of predictions for contributions to the entanglement, not all continuum geometries can be studies in lattice models [84, 85], but cornerless regions and those with right angles can be still be studied using square lattice geometries.

At two dimensional quantum critical points the entropy can generally be written as

$$S_\alpha = a\ell + \dots, \tag{6.1}$$

where the a coefficient is not universal, but parts of the subleading term can be. The term ℓ represents the boundary between region A and its complement, and in two dimensions this is the boundary length separating two regions.

At the two dimensional quantum critical point of the transverse field Ising model, there are no exact results for the scaling of the entanglement entropy or Rényi entanglement entropy from analytical work. There are theoretical results that we can compare to from related models [85], but it is unclear whether the universal results found in these cases should apply to the transverse field Ising model. The reason one might suspect the results hold is if the transverse field Ising model is connected to the related models by some irrelevant or marginal perturbation and the universal characteristics from these related models might be stable to these modifications. To put it in context, in one dimension all conformal field theories have the same functional form [9, 86], differing only in the leading coefficient, the central charge of the conformal field theory. The two universal quantities relevant are subleading terms in the Rényi entanglement entropy related to the geometry of the entangling region. One contribution is due to the presence of vertices
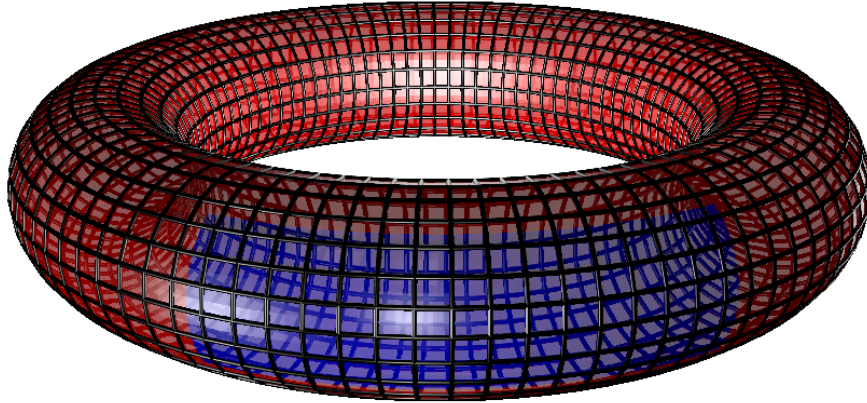
Figure 6.1: (Reproduced with permission from [79]) A square region embedded in a toroidal geometry. The four corners of the square give four times the contribution expected from an isolated corner, as long as they are far enough apart. This allows comparison to predictions for the entanglement entropy of sharp edges in theoretical models.

[87, 88, 89, 90, 91]. Fig. 6.1 shows an example of the square geometry we use to study the vertex contribution to the entanglement entropy. By examining the entanglement entropy of a square patch in a toroidal geometry we expect a contribution from each of the four corners and through careful subtraction schemes we can compare this to theoretical results. Fig. 1.3 shows how the same idea of scale invariance can lead to a subleading log contribution to the REE.

The other geometric contribution to the entanglement entropy that we can measure comes from examining a torus geometry split into two cylinders. In 1+1 models, the Rényi entanglement entropy takes the form [9, 86] (see also Sec. 1.5)

$$S_\alpha = \frac{c}{3\eta}\left(1 + \frac{1}{\alpha}\right)\log\left[\frac{\eta L}{\pi a}\sin\frac{\pi x}{L}\right], \tag{6.2}$$

where c is the central charge, L is the total size of the one dimensional system, a is the lattice spacing and x is the size of region A. The η term is one for open boundary conditions and two for periodic boundary conditions. Results in 1+1 have lead people to examine 2+1 data in a similar way, and attempt to fit to the form

$$S_\alpha = a\ell + c\log\left(\sin\frac{\pi x}{L}\right) + d, \tag{6.3}$$
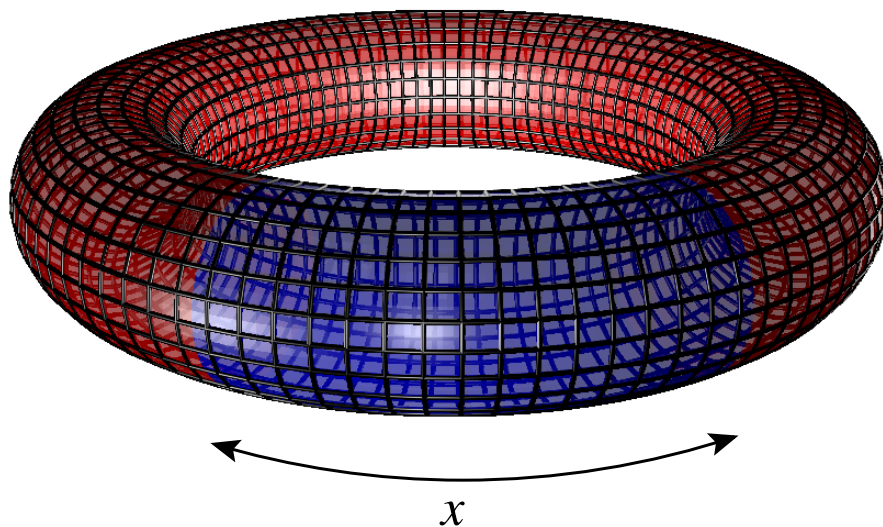
Figure 6.2: (Reproduced with permission from [79]) Splitting the torus into two cylinders. This choice of region A is closest to the geometry chosen in studying periodic one dimensional systems. The relevant parameter is the fraction of the torus of either region. Due to the symmetry of the entanglement entropy, one expects results for regions of size x/L to be identical for regions of size $(L-x)/L$.

where the assumption is that although the a term is not universal, the c term should be universal, and L now refers to the length of an $L \times L$ system. In general for this two dimensional geometry we expect a scaling of the form

$$S_\alpha = a\ell + \gamma\left(\frac{x}{L_x}, \frac{L_x}{L_y}\right) + d, \qquad (6.4)$$

where the common assumption is that this γ function was well approximated by the $\log(\sin)$ from one dimensional results. Fig. 6.2 shows how these two cylinders are mapped onto the torus. Using this assumption groups have tried to extract c numerically and show that it is a constant in the limit of large L [84], but in two dimensions the density matrix renormalization group methods used in one dimensional systems grow exponentially slow in the width of these systems, and so have difficulty scaling to large two dimensional systems. For this reason many of those results look at the long cylinder limit of this problem and attempt to draw results from there, but even in that case it is not clear if the coefficient c is saturating in the models of interest, something that would be necessary for it to be the analogue of the one dimensional case.

To extract the subleading terms in numerical studies requires the precise measurement of the entanglement entropy. We calculate the entanglement in a given finite size system and scale the region size and total system size up to accurately capture the part of the entropy that scales with the size of the interface between the regions. The difficulty of dealing with a finite size system is that there may be other subleading effects in the scaling not accounted for in the continuum theory. This means that although the smallest size systems are the most accurate in numerics, the largest size systems are the most useful for scaling. In addition, most smooth geometries discussed in continuum theories cannot be realized in lattice systems, dictating the use of the cylindrical and square patch geometries to test the predictions from these models. Along with the continuum results, we can also compare to other numerical results, such as high temperature expansion [75] and cluster expansion techniques [74], which approach the thermodynamic limit in different ways than the quantum Monte Carlo.

For the two cylinder geometry there are two effects that we can examine. One is the presence of an "even-odd" effect. In wavefunctions derived from the resonating valence bond (RVB) picture, there is an effect where there is an oscillation in the entanglement depending on whether the region size is even or odd [37, 92, 85]. This effect is smaller the larger the region is, and can be used to diagnose how RVB like the wavefunction is. The second effect is the shape dependence of the entanglement entropy [85, 37]. Eq. (6.3) shows one form of the entropy that depends on the size of the cylinder, but the $\log(\sin(\pi x/L))$ term could be replaced by other aspect ratio dependent $\gamma(x/L_x, L_x/L_y)$ functions as well.

94

One possibility is the shape function for the quantum Lifshitz free field theory [85], a function that has similar shape to the log(sin) but is still fundamentally different.

The shape function $J_\alpha(y)$ from the quantum Lifshitz model for an even region size takes the form

$$J_\alpha(y) = \frac{\alpha}{1-\alpha} \log\left[\frac{\eta(\tau)^2}{\theta_3(2\tau)\theta_3(\tau/2)} \frac{\theta_3(2y\tau)\theta_3(2(1-y)\tau)}{\eta(2y\tau)\eta(2(1-y)\tau)} \right], \tag{6.5}$$

where $y = x/L$, η is the Dedekind eta function and θ_ν is the Jacobi-Theta function. For an odd size there is an alternate form [85], but the important distinction is that there are large even-odd oscillations between the two functions at small y, and such behaviour does not describe the transverse field Ising model. When fitting two dimensional data the assumption is to fit to the form

$$S_\alpha = a\ell + cJ_\alpha(y) + d. \tag{6.6}$$

This is similar to Eq. (6.3) except that the dependence on the aspect ratio, y, is given by a slightly different function.

Finally, in systems such as the Heisenberg model at zero temperature, we know that the scaling of strip geometries is fit well to the form [92]

$$S_\alpha = a\ell + b\log(\ell) + c\log\left(\sin\frac{\pi x}{L}\right) + d, . \tag{6.7}$$

In the Heisenberg model the non-zero b term is argued by the presence of Goldstone modes in the groundstate of the model due to the continuous degrees of freedom [38]. There are unproven predictions that $b = 0$ in the case of the transverse field Ising model critical point, and we examine this possibility (see Fig. 6.5).

In the next section we describe the zero temperature algorithm. The projector quantum Monte Carlo is very similar to the finite temperature methods, with some unique modifications that allow for an increase in efficiency. Using this method we are able to test the predictions for the scaling of the REE in the two dimensional TFIM.

6.2   Zero temperature projector quantum Monte Carlo

At the core of the finite temperature stochastic series expansion is an attempt to expand the partition function in a tractable form so that one can measure observables of a system

in an efficient way. Eq. (2.11) is the result of that expansion, written in a way that facilitates simulation. Based on ideas from [78] and unpublished work by Sandvik, in the zero temperature projector methods the setup is different; we instead wish to look at the expectation of observables in the groundstate

$$\left\langle \hat{O} \right\rangle = \frac{\langle \psi_0 | \hat{O} | \psi_0 \rangle}{\langle \psi_0 | | \psi_0 \rangle}, \tag{6.8}$$

where $|\psi_0\rangle$ is the ground state wavefunction. There is an efficient way to sample this quantity that is conceptually very similar to the finite temperature methods described in Ch. 2.

In general we can write any wavefunction out as a superposition over the energy basis,

$$|\psi\rangle = \sum_i c_i |i\rangle, \tag{6.9}$$

where $|i\rangle$ is the energy eigenbasis of the Hamiltonian. If we add a constant to the Hamiltonian so that all energies are negative, then large application of the Hamiltonian to any trial state gives us

$$(-H)^n |\psi\rangle = \sum_i c_i (-E_i)^n |i\rangle \tag{6.10}$$

$$= c_0 (-E_0)^n \left[|0\rangle + \sum_{i=1} \frac{c_i}{c_0} \left(\frac{E_i}{E_0} \right)^n |i\rangle \right]. \tag{6.11}$$

The restriction that all energies are negative means that $E_i/E_0 < 1$ and when taken to a very high power these coefficients become very small. This means that sampling the groundstate wavefunction requires the ability to apply a high power of the Hamiltonian to a trial state, but this is precisely what the stochastic series expansion framework allows.

Formally, we sample the sum

$$\sum_{S_M} \frac{\langle \psi_L | \left(\prod_i H_i \right) \hat{O} \left(\prod_j H_j \right) | \psi_R \rangle}{\langle \psi_L | \left(\prod_i H_i \right) \left(\prod_j H_j \right) | \psi_R \rangle}, \tag{6.12}$$

where S_M is shorthand for sampling over all possible trial states $\langle \psi_L |$ and $|\psi_R\rangle$, as well as all possible operator expansions H_i and H_j. This is essentially identical to how S_M is used in Eq. (2.31). There are, however, a few differences between this formulation and the

finite temperature one. The first is that in the zero temperature method the far left and right states, $\langle\psi_L|$ and $|\psi_R\rangle$ do not need to be the same. As a shorthand, we will designate

$$\langle 0_L| = \langle\psi_L| \prod_i H_i \tag{6.13}$$

$$|0_R\rangle = \prod_j H_j |\psi_R\rangle \,, \tag{6.14}$$

representing the left and right projection of the ground state. In the finite temperature method we are sampling the trace, and so these states are enforced to be identical (see Sec. 2.2), but in the projector implementation there is no such restriction. This essentially modifies the boundary conditions in the expansion direction from periodic to open.

The second modification is that we only want to sample from wavefunctions that have a very high power of the Hamiltonian applied to them. This means we do not allow the identity operator, $H_{-1,0}$ (see Eq. (2.29)), in the expansion, and modify the diagonal update so that failed insertion steps are repeated until a successful operator insertion occurs, and operator removal is immediately followed by new operator insertion. Results must be converged in the number of operators, and the relevant scale that sets this cutoff is the inverse of the gap to the first excited state, as suggested by Eq. (6.11).

For describing the partition function of this simulation, the bottom part of Eq. (6.12), we proceed using the following steps for the diagonal update [79] (contrast this with the finite temperature algorithm in Sec. 2.2)

1. Starting at $\langle\psi_L|$, step to the next operator in the list.

2. If the operator is diagonal, remove it, if it is off-diagonal, update the state and go to the step 1.

3. If an operator was removed, insert a new diagonal operator H_a with probability $W_a/\sum_i W_i$, where these are the weights of all the diagonal operators over all possible locations they can be inserted. These weights are $2J$ for two site operators and h for single site operators.

4. If the operator cannot be placed due to local constraints, choose a new operator using the probability above until one is placed, and then proceed to step 1 until $|\psi_R\rangle$ is reached.

In this algorithm all diagonal operators are forced to change, although replacing an operator by itself is permitted, and the total number of operators remains fixed.

The off-diagonal update proceeds identically as in the finite temperature case (see Sec. 2.2), but with the exception that cluster now also end when they reach the end of the expansion, $\langle\psi_L|$ or $|\psi_R\rangle$. The clusters connect $\langle 0_L|$ and $|0_R\rangle$, and we can designate the division between the two by counting the number of operators we pass starting from either side. Connecting the left and right part of the expansion will be important when discussing the modified simulations necessary for examining the Rényi entropy. Clusters ending on these layers allow the off-diagonal update to change the local configuration at the end of the left and right projected states, and are an important part of the sampling procedure. With the diagonal and off-diagonal update we can sample the partition function for this construction.

Although many observables can be measured with the above construction, we will again restrict ourselves to measuring the Rényi entanglement entropy. The zero temperature method still makes use of the replica trick and ratio trick, and the methods presented in Ch. 5 also work here, but independent development historically lead to a slightly modified algorithm that we will discuss here. In the zero temperature algorithm we directly measure the ratio in Eq. (5.3), similar to the transfer matrix method discussed for classical systems except that there is no restriction on the regions A_i and A_{i+1} like the classical case. The ratio of partition function is related to the expectation value of a quantum *swap* operator [43], an operator that, in the context of the modified simulation, can be viewed as reconnecting the system with a modified region A.

The swap operator defines a region A in the modified simulation in the same way that the special boundary conditions do for a finite temperature simulation (see Sec. 3.3). To use the swap operator, one must have multiple replicas of the system, similar to the finite temperature approach. Fig. 6.3 shows what the zero temperature modified simulation looks like and what role the swap operator plays. To calculate the expectation of the swap operator, we simulate the system assuming a topology (or equivalently, assuming a particular swap operator inserted) and then calculate swap

1. Generate a new state using the diagonal and off-diagonal update and a choice of SWAP_A.

2. During the off-diagonal update, keep all of the cluster numbering. By this we mean record what cluster each spin in layer $|\alpha_0\rangle$ belongs to, on each replica of the simulation (see Sec. 2.2 for the off-diagonal update and Fig. 2.2 for how clusters are built) Modify the numbering so that clusters in each expansion and each replica have a unique numbering, i.e. if a cluster connects $\langle 0_L|$ in replica one to $|0_R\rangle$ in replica two, we would give them unique labels despite belonging to the same cluster.
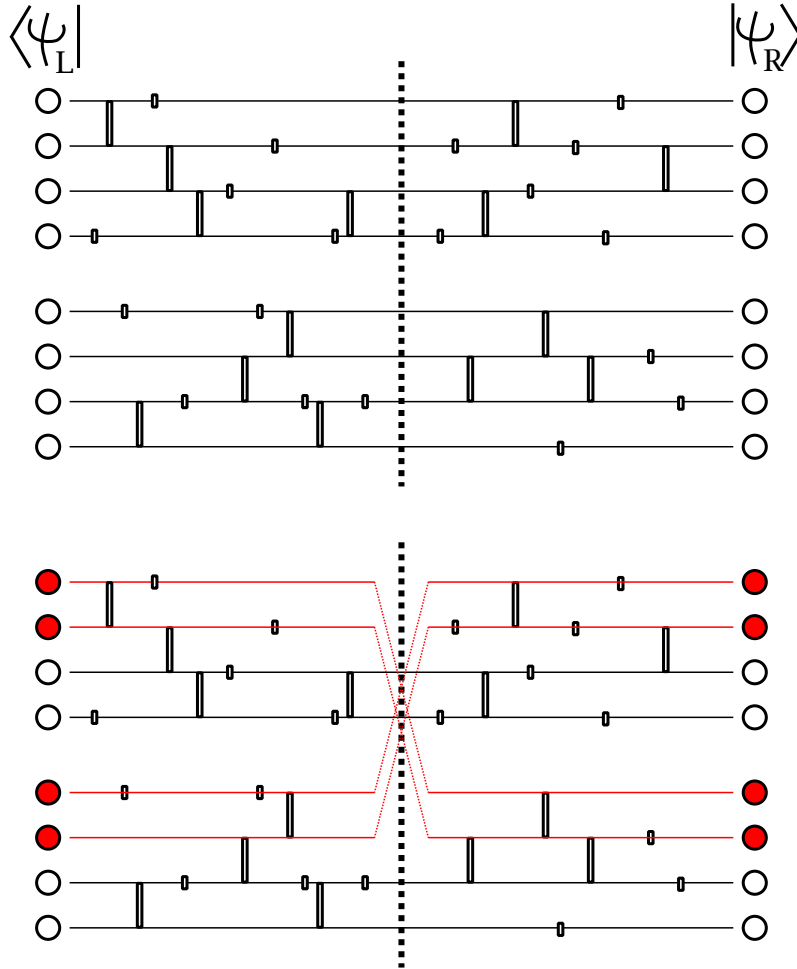
Figure 6.3: A modified zero temperature simulation cell of two replicas of four spins representing $\langle \psi_L | \left(\prod_i H_i \right) \text{SWAP}_A \left(\prod_j H_j \right) | \psi_R \rangle$ for different choices of region A. In this case, the region A differs between the two replicas by the spins (and lines) in red, and in the upper simulation the region A is empty. The figure shows the modified topology for the second Rényi entropy; for higher Rényi entropies the swap operator permutes the connection between he α copies. We generate the left and right half of the simulation using a particular topology, then calculate the number of loops crossing the boundary given different method of reconnecting. The weight of a configuration is to 2^{N_L}, where N_L is the number of loops crossing the center cut. Fig. 5.4 shows how this number of loops can be calculated using only local boundary information. The difference between that case and this one is that we label loops in each cluster as distinct, and then explicitly calculate reconnections for all different topologies (or choices of swap) rather than just the change in number of loops.

99

3. Using the new numbering, calculate the number of loops crossing the center using the original topology, $\langle 0_L | \text{SWAP}_A | 0_R \rangle$, and the modified topology, $\langle 0_L | \text{SWAP}_{A'} | 0_R \rangle$ (see Fig. 5.4).

4. The expectation of the ratio is given by

$$\frac{\langle \text{SWAP}_A \rangle}{\langle \text{SWAP}_{A'} \rangle} = 2^{N_A - N_{A'}}, \tag{6.15}$$

where N_A and $N_{A'}$ represent the number of loops crossing the center given different choices of region A, or equivalently different choice of swap.

Unlike the case used at finite temperature where we measured a transition probability that was strictly less than one, in the zero temperature approach we directly sample the ratio in one step. This does open the interesting question of what topology one should use the generate the states. In the current work, states are always generated using the modified simulation with a smaller region size, and then the ratio in Eq. (6.15) is sampled. In theory, one could generate the states using the modified simulation with the larger region size, or even an arbitrary region size, and then calculate the same ratio. The rationale for using one of the regions as a reference is because such a choice should bias the simulation towards regions where $\langle 0_L | 0_R \rangle$ is large, and those parts of the space contribute most significantly to the sum we are attempting to approximate. A rigorous study of whether this is accurate has not been performed to our knowledge. In principle, as long as for any choice of region A the configurations in the left and right projection are ergodic, the choice of reference region A should only bias the rate of convergence of the simulation, but not the result itself.

Developing this method allows us to study a fundamental but yet unexplored problem, the scaling of entropy at the critical point of the two dimensional transverse field Ising model at zero temperature, which we will discuss in the next section.

6.3   Results

We implement the zero temperature algorithms for the two dimensional transverse field Ising model and use it to examine the scaling of the entanglement entropy as a function of size and shape at the critical point.

The first necessary check is to ensure that we are working in the regime in which the results for the entanglement entropy are converged. The only parameter we have to tune
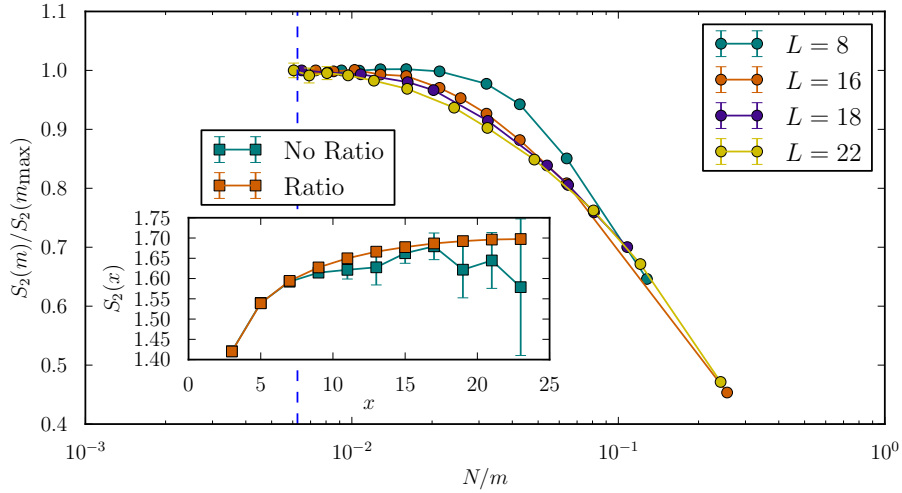
Figure 6.4: (Reproduced with permission from [79]) The normalized zero temperature entanglement entropy for different system sizes as a function of the number of operators applied to the trial state. For the two dimensional transverse field Ising model at criticality scaling the number of operators with N results in consistent saturation when $m/N > 160$ (dashed blue line). Inset shows the entanglement entropy calculated with and without the ratio trick, showing the large increase in error when not used, similar to the result in Fig. 5.5.
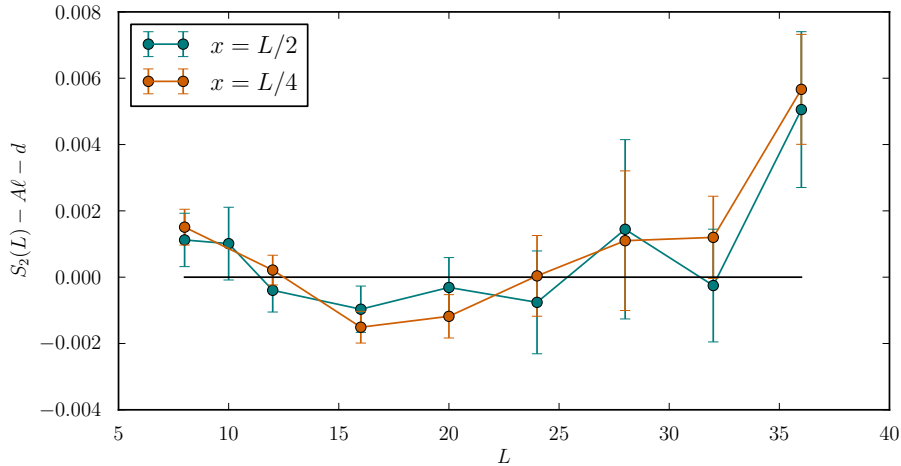
Figure 6.5: (Reproduced with permission from [79]) The residual of fitting all of the $x = L/2$ and $x = L/4$ cylinders to the form $S_2 = a\ell + d$. Since the aspect ratio, x/L, is fixed in both cases, and c term in Eq. (6.3) or 6.6 will be absorbed into the d term, as it will only contribute a constant. As this graph shows, the data is consistent within the form provided. If we also allow for a $\log(\ell)$ term in the fit, we would find that it would be very small and negative. Compared to the Heisenberg model where there is a known subleading log term [38, 92], the log term found here is smaller by two orders of magnitude. In addition, when examining the full strip data, the log is inconsistent. From this we conclude that no such term should be allowed in the fit.

is the number of operators, m, and the assumption is that for a large enough m the results will saturate. Eq. (6.11) shows how the error relates to the gap in the system, and for finite size systems we expect such a gap, although it may scale with system size. Fig. 6.4 shows how the entanglement entropy saturates to its low temperature value when the ratio of N/m (where N is the number of spins in the system) is below some threshold value. This result is important as it allows us to fix the number of operators and avoid an additional scaling in all of the results of the simulation.

The next characteristic to examine is whether the transverse field Ising model has $\log(\ell)$ terms in the entanglement entropy, similar to the Heisenberg model. Fig. 6.5 shows the residual, that is the data minus the fit, of fitting many system sizes using two fixed aspect ratios, $x = L/2$ and $x = L/4$. Using these cuts of the data, any shape dependence of the entanglement, that is terms that only depend on x/L, will be captured by the constant term in the fit. This allows us to determine the presence of a $\log(\ell)$ term while
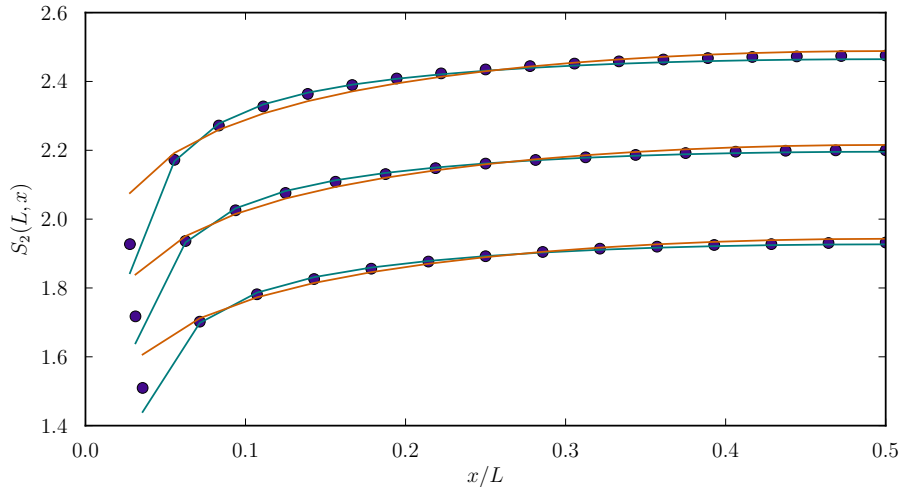
Figure 6.6: (Reproduced with permission from [79]) The entanglement entropy of systems of size $L = 28, 32, 36$ fit to (orange) Eq. (6.3) and (teal) Eq. (6.6). The important part of this fit is that in the raw data (dots) there is no even-odd effect, as seen in other systems such as pure resonating valence bond wavefunctions [37].

fitting a minimal number of other parameters. The data shows that the absence of such a term is consistent with the data, but allowing the presence of such a term finds a small negative coefficient for it. This coefficient is much smaller than in other systems where there is theoretical motivation for such a term, and this combined with the full fitting later suggest that the coefficient for the $\log(\ell)$ term should be set to zero. All other fits for the data assume that this is the case.

In the theoretical work examining the quantum Lifshitz free field theory, there is a prediction of strong even-odd effects in the region size [85]. Fig. 6.6 shows the entanglement entropy as a function of region size for both even and odd regions, and within the precision of the Monte Carlo there is no detectable even-odd branching behaviour. The even branch of the scaling used for the Lifshitz theory in Eq. (6.5) is shown to capture the data very well, capturing the small x/L and large x/L slightly better than the log(sin) function. With even-odd effects ruled out, we can quantitatively compare the two functions attempting to fit the two cylinder geometry shown in Fig. 6.2.

To compare the two fitting function, Eq. (6.3) and Eq. (6.6), requires a somewhat non-standard approach. We are looking for a function that captures the physics of all possible system sizes and bipartitions using only three coefficients. There are a few ways to make
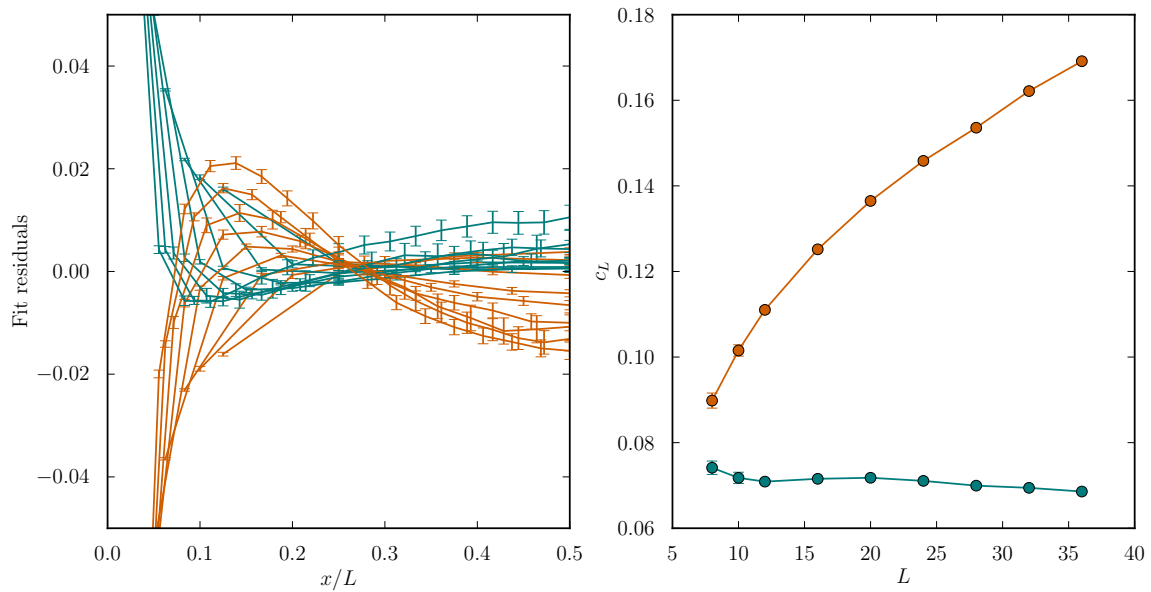
Figure 6.7: (Reproduced with permission from [79]) On the left we have the fit residual from fitting *all* system sizes to the form from (orange) Eq. (6.3) and (teal) Eq. (6.6). The right graph shows how the size-specific coefficient of the shape term depends on the system size. To reiterate, the area law term, a, and the constant term, d, are fit for all system sizes, while the system size specific coefficient, c, is fit for each system size individually. This is done for each shape function separately, allowing us to use the size dependence of the c-term to diagnose the quality of the fit.

104

such a comparison. One would be to fit the entire data set, all system sizes, using only three parameters, and then looking at the residuals over all system sizes. The fit with smaller residuals would better capture the physics of the system. Although this approach is possible, it does not tell us how well either shape function captures the nature of each system, but only gives us a comparison between the two shape functions and tells us which is better. A second approach would separately fit each system size using the functional form and examine the coefficient of the shape function from every fit. This is a fairly straightforward approach, but it has an unfortunate possibility for failure. The coefficient c that is chosen is highly sensitive on the coefficients a and d. That is to say, if we impose a constraint on either coefficient a or d it will have a strong effect on the c chosen for the fit. In fact, for a single system size the area of contact between the two regions, ℓ, is constant, and so we have no way of distinguishing a and d. Although a and d are not universal, they should be consistent for a particular choice of Hamiltonian, and should be extractable within a set of simulations.

The method we choose to extract the coefficient of the shape term is a combination of fitting all the systems together and fitting them all independently. We fit the function using a form

$$S_\alpha = a\ell + c_L f(x/L) + d, \tag{6.16}$$

where $f(x/L)$ are the specific functions from Eq. (6.3) and Eq. (6.6) and c_L is a coefficient that can be different for each size. Fig. 6.7 shows the residual for each of these fits on the left and c_L for each function on the right. Using this method we can examine the residual from each fit while also examining if the coefficient c_L is size independent. The fitting process shows that although both functions are able to capture the behaviour of the function fairly well, the residuals are comparable, the coefficient for the log(sin) shape function, the result in one dimension, varies almost linearly with system size, while the shape function from the quantum Lifshitz free field theory only varies marginally as a function of system size. This result strongly suggests that the Lifshitz shape function may be more appropriate for the critical transverse field Ising model. It should be noted that the result for the entanglement entropy is only valid for the Rényi entanglement entropies with $\alpha \geq 2$, and there is no analytic form for the von Neumann entanglement entropy. Many other results look directly at this quantity [84], and a direct comparison of these results to the von Neumann entanglement entropy is not possible.

The other interesting geometry that we examine is that of a square region embedded in the torus. In this geometry we expect a scaling of the form

$$S_2 = a\ell + \gamma(x/L, y/L) + n_c a_2(\pi/2) \log(\ell) + d, \tag{6.17}$$
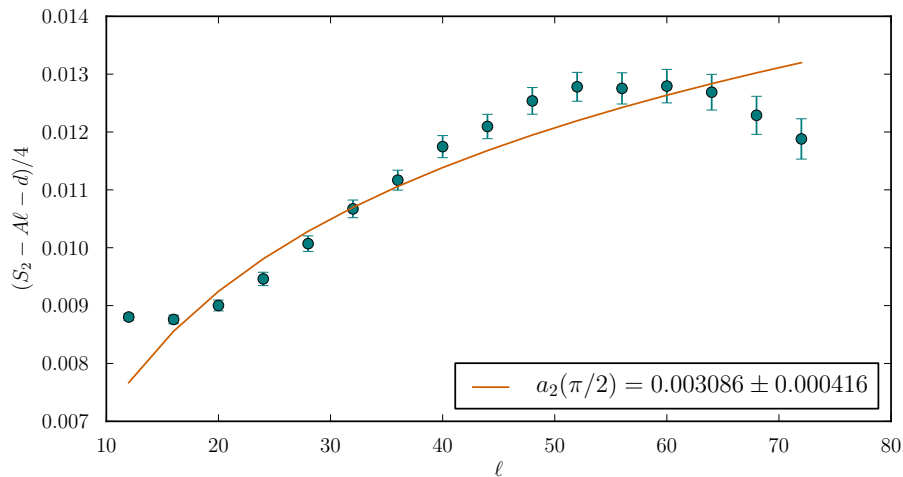
Figure 6.8: (Reproduced with permission from [79]) Extraction of the corner log term using a 36×36 system and scaling a square of size $\ell/2 \times \ell/2$ (total perimeter ℓ) inside to extract the corner term. Notice the shape of the raw data does not seem consistent with a log shape, suggesting this geometry of extraction is not well suited to capture the corner term. In addition, by fitting only the smaller squares we get a different sign of the log term, negative as opposed to positive. Formally the results for the corner term in theory are for an isolated corner far from other geometrical factors, so the more reliable results should be from larger squares rather than smaller ones.
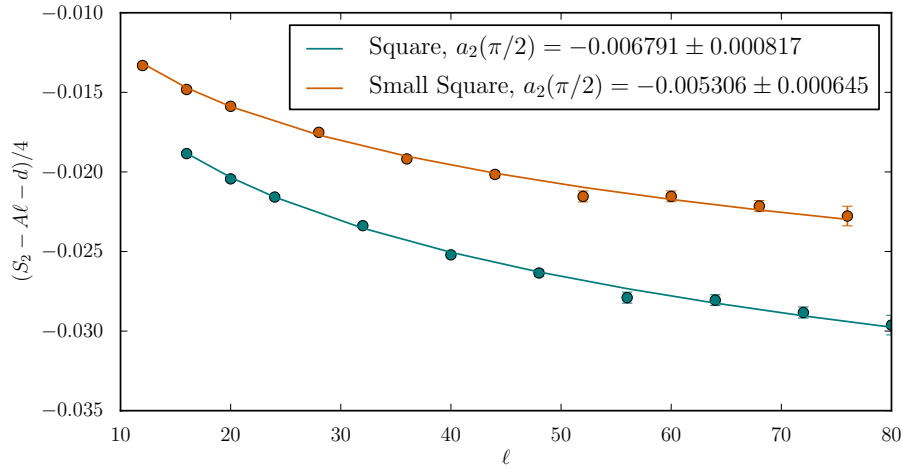
Figure 6.9: (Reproduced with permission from [79]) Extraction of the corner term using $L/2 \times L/2$ squares. The difference between "Square" and "Small Squares" is in the definition of the boundary length; for "Squares" it is the average of the number of sites on the outer and inner perimeter, while for "Small Squares" it is the inner, or smaller, perimeter which is used. By keeping the fractional size of the square fixed any contribution from the coefficient c in Eq. (6.17) will be absorbed into the constant part of the fit. Using this geometry we get a value for the corner term of $a_2(\pi/2) = -0.006(2)$ after accounting for size exclusion and different definitions of boundary length.

where $\gamma(x/L_x, y/L)$ is some unknown aspect ratio function contribution for a general rectangle (we assume a square $L \times L$ system in this case), n_c is the number of corners and $a_2(\pi/2)$ is the coefficient of a ninety degree corner. The aspect ratio function is necessary for the same reason the shape function for the cylinder geometry is necessary, and it is natural to assume that for a general rectangle both fractional widths matter in a finite size system. If one assumed that the function γ is zero, then the simplest scaling uses a fixed system size and increases the size of the square inside to extract the corner term. Fig. 6.8 shows this scaling using a 36×36 system and squares up to half that size. It is immediately obvious that this scaling does not capture a simple $\log(\ell)$ scaling, and that there are other factors influencing the scaling in a strong way. This may mean that the function γ in Eq. (6.17) is non-zero or there are further corrections that make this extraction difficult. Previous QMC work [77] used this approach to attempt extraction of the corner term, but the resolution of the REE in this study is higher, and able to resolve the shape dependence more clearly.

Fig. 6.9 shows one fit of the square region data using an $L/2 \times L/2$ square embedded in an $L \times L$ system. There are a few details that make this extraction of the corner term non-trivial. First is the definition of the boundary length. Although typically the boundary length has been defined as the average of the number of sites on either side of the boundary, there are arguments that it should be the minimum of the two [77]. One can see that for very small regions, such as a one dimensional strip of spins, it is unclear if one should count each side of the spins as part of the boundary, or simply the number of spins. Clearly in the case of very large regions the different between these two definitions is trivial, but typically we deal with small regions where these definitions have a noticeable effect. Our analysis shows that the extraction of the corner term coefficient is sensitive to this definition for the sizes of interest.

The second factor that adds difficulty is that the form of the entanglement entropy contribution from corners is only accurate for a system in which the corners are very far apart. Clearly a 4×4 square in an 8×8 system is not in this limit, but the question is how do these corrections affect the fitting procedure. We can eliminate such corrections by ignoring small system sizes in the fitting procedure, but by doing so we also eliminate the most highly accurate data. This is the primary reason that the result for the corner term is so large; the best estimate for the corner term from this Monte Carlo work is $a_2(\pi/2) = -0.006(2)$ [79].

This result for the corner term can be compared to other work. At the non-interacting Gaussian fixed point, the estimate for $a_2(\pi/2) = -0.0064$ [88], while numerical series expansion for the interacting quantum critical point of the transverse field Ising model give $a_2(\pi/2) = -0.0055(5)$ [75]. Numerical linked cluster expansion methods give $a_2(\pi/2) = -0.0053$ [74]. Previous Monte Carlo working using the fixed-systems scaling generate a result of $a_2(\pi/2) = -0.0075(25)$ [77], but reproducing this analysis we find that this geometry is ill-suited to extract the corner term.

6.4   Discussion

By using a zero temperature projector quantum Monte Carlo we have performed a detailed study of the Rényi entanglement entropy of the two dimensional transverse field Ising model at criticality. We examined two geometries amenable to large scale simulation on $L \times L$ lattices with periodic boundary conditions.

Using a fixed size of square region in the toroidal geometry, we confirm a scaling of the

108

form

$$S_2 = a\ell + 4a_2(\pi/2)\log(\ell) + d, \tag{6.18}$$

with a as a non-universal coefficient of the area law and

$$a_2(\pi/2) = -0.006(2) \tag{6.19}$$

is our best estimate for the universal coefficient of the subleading logarithmic term associated with corners. It is consistent with the result from series expansion and numerical linked cluster expansion, both of which approach the thermodynamic limit in a very different way. The difference between this result and that from the Gaussian fixed point is small. Since the Wilson-Fischer fixed point of the transverse field Ising model can be perturbatively reached in ϵ-expansion from the Gaussian fixed point, perhaps this result is unsurprising.

When bipartitioning the system into two cylinders of size $x \times L$ and $(L - x) \times L$ we get two systems with a smooth vertexless boundary between them. This geometry allowed us to show there is no resonating valence bond like character in the wavefunction, something that would cause noticeable even-odd oscillations. In addition, we were able to examine the shape function and quantitatively compare the log(sin) subleading term in Eq. (6.3) and compare it to the shape function from the quantum Lifshitz free field theory in Eq. (6.6). This analysis suggests that the quantum Lifshitz shape function is more appropriate, although it can only be applied to the Rényi entanglement entropies. The approach used to generate the shape function in the case of the Lifshitz free field theory also uses a replica trick, and so the results rely on the multi-sheeted Riemann surface for the calculation. For this reason they cannot be easily extended to the von Neumann entropies. Even if one has access to the von Neumann entropy, the Rényi entropies are often more amenable to numerical simulation, and may even have realization in experiment.

Finally, the accuracy of the shape function from the quantum Lifshitz free field theory suggests that this might be a more appropriate function for all fixed points in $2+1$, a point that could be examined in other numerical and theoretical studies. For one dimensional result, the log(sin), is typically used to examine the scaling of cylinders in two dimensional systems. In the case of gapless systems it fits the data fairly well, but the full scaling analysis shows that the leading coefficient is not universal in the case of the REE of the Wilson-Fischer fixed point. Since the shape function for the Lifshitz free field theory is derived from a fundamentally two dimensional system, it may be that the small deviations between the two fits (see Fig. 6.6–they are quite similar) contains the essential difference between one and two dimensional systems. If that is the case it should be unsurprising

if this shape function is applicable to a larger class of two dimensional systems, and with this study we demonstrate a comprehensive approach to examining the scaling of a two dimensional system. If this shape function is universal to $2+1$, there are a wide variety of systems for which it can be examined. Along the lines of the previous chapter, universal collapse can also be examined by adding a finite temperature to the model.

Chapter 7

Conclusions

Although entanglement measures are not in themselves new, their incorporation into Monte Carlo techniques is, and the development of algorithms to accurately measure entanglement is crucial. We have demonstrated a variety of methods for calculating the entanglement entropy and mutual information in classical and quantum Monte Carlo. Furthermore, the entanglement measures show potential for classifying critical phases and detecting phase transitions. These all point to further development, both numerically and theoretically, of entanglement as a resource in condensed matter physics.

The most promising aspect of modifications to Monte Carlo algorithms to allow for measuring the Rényi entanglement entropy is that they are simple. The modifications to a classical or quantum Monte Carlo are conceptually no more difficult than the construction of the Monte Carlo itself. In the simplest approach, the only quantity needed is the energy of the normal and modified simulation, and for most Monte Carlo techniques this is already available as it is needed to generate the acceptance rate of proposed updates.

The ability to detect phase transitions without knowledge of an order parameter, or for that matter without knowledge of the existence of a phase transition, is truly remarkable. Two point correlation functions and other traditional measures remain necessary, as much of our understanding of criticality and phase transitions comes from examining these quantities, but the existence of these new observables provides a invaluable window in systems where little else is understood or traditional measures fail. Along with the detection of phase transitions the finite size scaling of transitions shows universal behaviour that can be used to identify the universality class of the model. This is demonstrated by logarithmic divergence in the derivative of the mutual information in the Ising model phase transition.

Broad histogram techniques represent progress towards efficient calculation of the ther-

mal Rényi entanglement entropy, and future techniques may rely on a combination of this combined with conventional methods to quickly measure it in Monte Carlo. In principle, such broad histogram techniques could be used in other parameters, such as the region size. Doing so recovers the ratio trick from a different starting point, but given the highly efficient methods for directly calculating these ratios it is unlikely this direction will lead to a more efficient implementation. That being said, biasing the sampling of the simulation may lead to possible improvements if done in a more directed method than Wang-Landau, and is discussed in Sec. 7.1.

The finite temperature ratio trick methods are very new. This is partially because for most models of interest the naive approach works sufficiently well. For models with a first order transition or with difficult dynamics for a particular temperature range, integration approaches will necessarily have more trouble. It is for these cases that the finite temperature ratio trick might be far more useful. In addition, the integration approach is not in competition with direct calculation at a specific temperature, but rather complementary. As long as the Rényi entropy or mutual information is known at some temperature, the integration approach can use that point as a reference. This means the integration is only needed in a small range where we wish to examine crossing or scaling behaviour. The infinite temperature point is used because it is trivial to calculate both the mutual information and Rényi entropy at these points, but by using a numerical value near the critical point it is possible to use integration methods to examine crossings without simulating the entire temperature range of the model. The behaviour of the classical mutual information at critical points has only been studied in a very limited set of cases, and how it will manifest in the more general case may reveal entirely new applications for this quantity.

Using this finite temperature approach we have shown agreement with the analytical predictions of the finite temperature and finite size corrections to the transverse field Ising model at criticality in one dimension. This is one of the first attempt to match corrections to the thermodynamic limit for the scaling of the Rényi entanglement entropy, and demonstrates how this quantity can be used to directly examine the universal behaviour in critical systems. This suggests that in addition to the scaling of the REE across a phase transition that the scaling at a phase transition as a function of geometry can generally explore universal quantities. We have just begun to explore this with the zero temperature methods, but examining these quantities at classical thermal critical points should also probe universality in an interesting way.

Conceptually, the zero temperature methods are very similar to the finite temperature methods for quantum Monte Carlo systems. The faster convergence of this type of Monte Carlo is the main benefit, leading to the ability to generate results accurate enough to examine subleading terms in the Rényi entanglement entropy. Even so, this part of the

work, representing over 300 cpu-years of computational effort, highlights the amount of effort needed to accurately extract these universal parts of the entropy. Analysis of entropy in two dimensions is itself relatively new, with many of the possible geometric effects unexplored, along with nuances in the scaling and shape functions. The difference of two shape functions, as shown in Fig. 6.6, shows how similar different proposals can be. That small difference is sufficient to change the behaviour of the subleading constant from trending with system size to saturating. Hopefully this work prompts deeper exploration of interacting theories in $2 + 1$, and leads to a more comprehensive theory of universality in these cases.

We hope that these techniques continue to be developed and refined, and that the success thus far leads to a more complete theoretical understanding so that these methods may be applied to an even wider range of applications with confidence and understanding in the results found.

7.1 Future Directions

As with any project exploring new methods, there are often many avenues to explore, but we choose those we think will be the most interesting or necessary. In this section we discuss concepts for projects that have not, to our knowledge, been pursued by the scientific community. Any person with a working of the material in this manuscript should be able to make progress on these problems.

This ease of implementation of the classical mutual information makes the method available to a wide range of application. Models with hidden transitions are of immediate interest [93], as well as fundamental loop [94, 95] and dimer models [96, 97] that may hold universal properties of related critical systems [98]. Of course glassy and disordered systems [99, 100, 101, 102] represent an old problem where the existence of a transition itself is unclear, and the Rényi entanglement entropy and mutual information may be able to shed some light [103]. In addition, there are non-condensed matter systems such as financial markets [104] where the mutual information may act a resource to herald collective behaviour. Techniques from information theory [76] also inspire addition route to calculating these information quantities, and will likely be a rich resource for new algorithms and perspectives.

An interesting problem is examining two cylinder partitioned systems using a $2L \times L$ geometry as opposed to an $L \times L$ geometry, or in general examining different aspect ratio geometries. The specific quantity of interest is the constant part of the mutual information

through a phase transition, as the crossings in the mutual information are caused by this constant piece changing sign. Fig. 3.2 shows the numerical extraction of the constant part of the mutual information for the Ising model. If we recall the classical formulation of the replica trick, half of the systems is effectively at half the temperature of the other half of the system. Examining the behaviour of the constant piece we see it move away from zero before crossing through it at each transition as β is increased. Part of this behaviour may have to do with the aspect that each region of the system is an $L/2 \times L$ strip in these geometries. This squished geometry may come in to play when we imagine how it interplays with the growing correlation length at criticality. In the shorter direction the correlation length will see the boundary with the other system before the correlation length grows larger than L in the long direction of each region. Although the two temperatures at which these occur will likely be very close, as the correlation grows exponentially fast as one approaches criticality, it may contribute to the behaviour of the constant piece. The study of different aspect ratio systems might help examine if this theoretical approach to understanding the crossings of the mutual information is a useful direction to pursue.

The quality of the thermal results in one dimension outline a methodical way to approach two dimensional systems. The zero temperature work gives a good baseline of the limiting behaviour of any function also valid away from the thermodynamic limit, and suggest that the best way to examine thermal crossover functions as a function of strip size over the finite system size or strip size times temperature. By doing this we may learn if the quantum critical fan in two dimensions has similar behaviour to the one dimensional case. Conceptually the two dimensional code is no more challenging than the one dimensional code, but if the interesting behaviour is subleading the difficulty of accurate subtraction still precedes a universal collapse. As a possible benefit, if the quantum critical fan influences a large temperature range, modestly low temperatures may be able to access this physics, meaning that the thermal Monte Carlo algorithm might benefit from not necessarily needing to access the zero temperature physics. This is a benefit as the scaling of the finite temperature Monte Carlo is $\mathcal{O}(N\beta)$, where N is the number of Hamiltonian terms, proportional to the number of spins for short range models.

With the rise of numerical approaches to numerically measure the Rényi entropy, theoretical proposals for physical experiments are also following, in some cases looking philosophically similar to the numerical approach [12, 11, 13]. Pursuit along this direction, that is examining connections between numerical and experimental realization, should offer a rich variety of potential realizations. Along that direction, the ability of the classical Rényi entropy to detect the thermal phase transition may be related to the realization of a system where only some of the system is critical. The simulation of a system without an effectively uniform temperature implies that the temperature at which critical physics occurs may

114

be smeared out. If instead of the two replica system, we imagine a single system with a gradient of temperature from T to $2T$, then whenever the critical temperature lies in that range, part of the system will experience an effect from the critical phase. Whether such formulations lead to any interesting behaviour or are realizable in physical systems is uncertain, but they can be simulated without much difficulty.

Finally, Fig. 5.5 illustrates why the ratio trick is necessary, but perhaps it is possible to address the problem of sampling directly rather than relying on the ratio trick. The fundamental problem is that the primary contribution to the observable is very different than the natural weights of the system. To rephrase: if all the states, labelled by some index i, have weight W_i, then the states with large W_i are not the largest contributors to $O_i W_i$, which is the quantity we are measuring. Sampling with probability $O_i W_i$ could help this, but since we still need to know the weights W_i to normalize this simulation, just doing this is as equally problematic as the original formulation. What needs to be done is ensure that the sampling procedure spends equal time sampling the states that contribute strongly to $O_i W_i$ and to W_i. Looking at the bottom graph in Fig. 5.5, this would mean spending equal time sampling around $L_0 - L_s = 5$ and $L_0 - L_s = 16$. In principle this is not difficult, since only the diagonal updates can change the number of loops crossing the boundary, but there is a non-trivial part. Implementing such an approach requires that the change in the number of loops crossing the boundary for the insertion or removal of an operator is a fast calculation, otherwise the entire simulation becomes significantly slower. Although it is not clear to the author whether this is straightforward or not, solving this problem would potentially allow faster calculation of Rényi entropies than with ratio trick methods, by specifically targeting the distribution of interest during the sampling procedure.

APPENDICES

Appendix A

Entanglement in model systems

Here we present the calculation of the entanglement entropy of one of the spatial regions of a simple harmonic oscillator with the conjugate region of the same oscillator.

In the zero temperature limit, that is a single particle in the ground state of a simple harmonic oscillator, the density matrix describing the system takes the form

$$\rho(q, q') = \left(\frac{m\omega}{\pi\hbar e^{\beta\hbar\omega}}\right)^{1/2} \exp\left[-\frac{m\omega}{2\hbar}\left(q^2 + q'^2\right)\right] \tag{A.1}$$

$$\rho(q, q') = \left(\frac{1}{\pi}\right)^{1/2} \exp\left[-\frac{1}{2}\left(q^2 + q'^2\right)\right], \tag{A.2}$$

by taking $m = \omega = \hbar = 1$ for simplicity. This function is simply the product of the groundstate wavefunction and itself for two arguments q and q', and so equivalent to $|0\rangle\langle 0|$, except written in the real space basis. The full density matrix then would be the sum of all such terms, and could be written as

$$\rho = \iint_{-\infty}^{\infty} dq dq' \left(\frac{1}{\pi}\right)^{1/2} \exp\left[-\frac{1}{2}\left(q^2 + q'^2\right)\right] |q\rangle\langle q'|. \tag{A.3}$$

This above representation is correct, but does not lend itself to easy calculation of the entanglement entropy given a particular real space cut of the system. The most straightforward way to do this is to extend to the grand canonical ensemble.

Before we start, we should note that to fully describe an arbitrary one particle wavefunction formally requires an infinite set of coefficients, whether these are the amplitude and phases in the simple harmonic oscillator energy level basis, or complex wavefunction

in the real space basis. To describe an arbitrary two particle wavefunction would require either the full value of the wavefunction for all possible positions of the two particles or the set, but to describe the wavefunction of *zero* particles should require zero parameters. That is to say, there is only a single state in the zero dimensional space – let us call it $|\emptyset\rangle$, and define it such that

$$\langle\emptyset\,|\,\psi\rangle = 0 \tag{A.4}$$

$$\langle\emptyset\,|\,\emptyset\rangle = 1, \tag{A.5}$$

where here $|\psi\rangle$ represents an arbitrary one particle wavefunction. Using this notation we can now rewrite the density matrix of a system partitioned at x as

$$\rho = \begin{cases} \iint dq\,dq'\,\rho(q,q')\,|\emptyset\rangle_A\,|q\rangle_B\,\langle\emptyset|_A\,\langle q'|_B \text{ if } q > x \text{ and } q' > x \\ \iint dq\,dq'\,\rho(q,q')\,|\emptyset\rangle_A\,|q\rangle_B\,\langle q'|_A\,\langle\emptyset|_B \text{ if } q > x \text{ and } q' < x \\ \iint dq\,dq'\,\rho(q,q')\,|q\rangle_A\,|\emptyset\rangle_B\,\langle\emptyset|_A\,\langle q'|_B \text{ if } q < x \text{ and } q' > x \\ \iint dq\,dq'\,\rho(q,q')\,|q\rangle_A\,|\emptyset\rangle_B\,\langle q'|_A\,\langle\emptyset|_B \text{ if } q < x \text{ and } q' < x \end{cases}, \tag{A.6}$$

where the subscripts A and B define the region less than and greater than the dividing point x. These subscripts act in such a way that $\langle\psi|_i\,|\phi\rangle_j = \delta_{ij}\,\langle\psi\,|\,\phi\rangle$.

To calculate the entanglement entropy we have to trace out all the states in a subregion of the system. The addition of the state $|\emptyset\rangle$ into our Hilbert space means that it is also another vector we need to include in our trace. Tracing over the states in region B to give us the reduced density matrix in region A gives us

$$\mathrm{Tr}_B\rho = \int dz\,\langle z|_B\,\rho\,|z\rangle_B + \langle\emptyset|_B\,\rho\,|\emptyset\rangle_B \tag{A.7}$$

$$= c\,|\emptyset\rangle_A\,\langle\emptyset|_A + \iint_A dq\,dq'\,\rho(q,q')\,|q\rangle_A\,\langle q'|_A, \tag{A.8}$$

where $\rho(q,q')$ is the same one from Eq. (A.2) and the constant c is equal to $\int_B dq\,\rho(q,q)$. Note that in this density matrix is only defined for values of $q, q' \in A$, as if the particle is not in region A then we default to the function $|\emptyset\rangle$ which effectively states there is no particle in region A.

If we attempt to solve the eigenvalue equation using the vector $|\psi\rangle = |\emptyset\rangle$ we get

$$\rho_A\,|\psi\rangle = \lambda\,|\psi\rangle \tag{A.9}$$

$$c\,|\emptyset\rangle\,\langle\emptyset|\,|\psi\rangle + \iint_A dq\,dq'\,\rho(q,q')\,|q\rangle\,\langle q'|\,|\psi\rangle = \lambda\,|\psi\rangle \tag{A.10}$$

$$c\,|\emptyset\rangle\,\langle\emptyset|\,|\psi\rangle = \lambda\,|\psi\rangle \tag{A.11}$$

$$c = \lambda. \tag{A.12}$$

118

We can solve for the second eigenvalue assuming it is a vector from the single particle wavefunction basis

$$\rho_A |\psi\rangle = \lambda |\psi\rangle \tag{A.13}$$

$$c \langle q''| |\emptyset\rangle \langle \emptyset| |\psi\rangle + \iint_A dq dq' \rho(q, q') \langle q''| |q\rangle \langle q'| |\psi\rangle = \lambda \langle q''| |\psi\rangle \tag{A.14}$$

$$\int_A dq' \rho(q'', q') \langle q'| |\psi\rangle = \lambda \langle q''| |\psi\rangle \tag{A.15}$$

$$\pi^{-1/2} e^{-q''^2/2} \int_A dq' e^{-q'^2/2} \psi(q') = \lambda \psi(q'') \tag{A.16}$$

$$\pi^{-1/2} e^{-q''^2/2} \int_A dq' e^{-q'^2/2} e^{-q'^2/2} = \lambda e^{-q''^2/2} \tag{A.17}$$

$$\pi^{-1/2} \int_A dq' e^{-q'^2} = \lambda, \tag{A.18}$$

where at Eq. (A.17) we assume a form of the solution and plug it in to both sides, giving us a closed solution on the following lines. For any such infinite dimensional density matrix, there should be, correspondingly, an infinite number of eigenvalues. In the construction above, the two eigenvalues correspond to the weight of the wavefunction left and right of the point x. Due to normalization these two eigenvalues add to one, and since the sum of the eigenvalues of a density matrix add to one, we can say that we have found all of the non-zero eigenvalues.

One can repeat this analysis for single particle in any eigenstate of a simple harmonic oscillator or a particle in a box and the derivation holds. In fact, since any single particle pure state can be represented as some wavefunction $\psi(x)$, this derivation should also hold for any single particle pure state – that is to say, the reduced density matrix will only have two eigenvalues, where the eigenvalues are

$$c_1 = \int_A |\psi(x)|^2 \tag{A.19}$$

$$c_2 = \int_B |\psi(x)|^2. \tag{A.20}$$

With these quantities we can easily calculate arbitrary Rényi entropies, including the von

119

Neumann entropy

$$S_1 = -\sum_i c_i \log c_i \tag{A.21}$$

$$S_\alpha = \frac{1}{1-\alpha} \log \sum_i c_i^\alpha \tag{A.22}$$

In principle, one should be able to use the above formalism to write out a density matrix appropriate for bipartitioning given an arbitrary density matrix. The problem with a general density matrix is that if it does not correspond to a pure state then the original density matrix may have multiple non-zero eigenvalues, suggesting that the bipartitioned wave function will have more than the two eigenvalues shown above. If one is able to decompose the mixed state density matrix into the pure states it is composed of, then one may be able to repeat the above derivation on each pure component. It is plausible (but not proven or even attempted here) that the pair of eigenvalues generated for each pure component are simply multiplied by their weight in the density matrix, giving the result that starting from a density matrix of n pure states we end up with $2n$ eigenvalues in the new bipartitioned system.

The final complication comes in wavefunctions of many particles. In this case instead of each side of the system only containing a zero particle and one particle wavefunction, each side of the system will need to contain wavefunctions from zero to N particle number, where N is the total number of particles in the system. From here one has a framework to make progress on the entanglement, as tracing out over a region of the space is at least more well defined from this starting point than previously.

Appendix B

Code Examples

In this section we will present code snippets that help clarify some of the algorithms mentioned in this thesis using minimal examples. Most of the code is written in python, not for its efficiency, but rather because minimal functional examples are very short.

If you would like to test any of these codes yourself, they are freely available here https://github.com/dartonias/phd-samples

B.1   Wang-Landau

Below is a minimal Wang-Landau code, written for Sec. 4.1.

```python
"""
Short test code demonstrating Wang-Landau method on a 1D Ising chain
Stephen Inglis, 2013"""

import random

class SIM:
    def __init__(self,N=4):
        # Define the spins, 0 for spin up and 1 for spin down
        self.spins = [0 for i in range(N)]
        self.N = N
        # Define the modification to weights
        self.k = 2.
        # Define the weights -- we don't know the valid energies to
            start
```

```python
15              self.g = {}
16              # Define the counter for bins
17              self.C = {}
18      def calc_E(self):
19          E = 0
20          for i in range(self.N):
21              # Calculate the energy of all pairwise ferromagnetic
                    interactions
22              E += int(((self.spins[i]+self.spins[(i+1)%self.N])%2)*2 - 1)
23          return E
24      def spinflip(self):
25          currE = self.calc_E()
26          try:
27              currW = self.g[currE]
28          except:
29              self.g[currE] = 1.
30              currW = 1.
31          x = random.randint(0,self.N-1)
32          self.spins[x] = (self.spins[x] + 1)%2
33          newE = self.calc_E()
34          try:
35              newW = self.g[newE]
36          except:
37              self.g[newE] = 1.
38              newW = 1.
39          # Check the old and new weights to see whether to accept the
                proposed move
40          if random.random() < (currW / newW):
41              self.g[newE] *= self.k
42              try:
43                  self.C[newE] += 1
44              except:
45                  self.C[newE] = 1
46          else:
47              self.g[currE] *= self.k
48              self.spins[x] = (self.spins[x] + 1)%2
49              try:
50                  self.C[currE] += 1
51              except:
52                  self.C[currE] = 1
53      def renorm(self):
54          minW = self.g[0]
55          for i in self.g:
56              self.g[i] /= minW
57          minC = self.C[min(self.C,key=self.C.get)]
```

```python
58            maxC = self.C[max(self.C,key=self.C.get)]
59            avgC = 0
60            for i in self.C:
61                avgC += self.C[i]
62            avgC /= len(self.C)
63            if (maxC - minC)/avgC < 0.1:
64                self.k = self.k**0.5
65                for i in self.C:
66                    self.C[i] = 0
67
68  def main():
69      bins = 50
70      reconfig = 1000
71      numsteps = reconfig*20
72      stats  = {}
73      for j in range(bins):
74          s = SIM()
75          for i in range(numsteps):
76              s.spinflip()
77              if i%reconfig == reconfig-1:
78                  s.renorm()
79          if stats == {}:
80              stats = s.g
81              stats2 = {i:s.g[i]**2 for i in s.g}
82          else:
83              stats = {i:stats[i] + s.g[i] for i in s.g}
84              stats2 = {i:stats2[i] + s.g[i]**2 for i in s.g}
85      stats = {i:stats[i]/bins for i in stats}
86      stats2 = {i:stats2[i]/bins for i in stats2}
87      error = {i:(stats2[i] - stats[i]**2)**0.5/(bins**0.5) for i in stats
              }
88      minW = stats[0]
89      for i in stats:
90          stats[i] /= minW
91          error[i] /= minW
92      print sorted(stats.items(),key=lambda x:x[0])
93      print sorted(error.items(),key=lambda x:x[0])
94
95  if __name__ == "__main__":
96      main()
```

B.2 Classical ratio trick using transfer matrices

Below is a minimal code implementing the transfer matrix method of calculating the ratio of modified partition functions described in Sec. 5.2. Everything needed is provided, although only the first code represents the simulation.

```python
"""Simple ising model code to try to calculate S_2 / MI_2 without
    integration
in temperature, but instead using replica-like methods.
Author: Stephen Inglis
Date: July 19, 2012"""

import random
import math
import numpy as np
from xml.dom.minidom import parse

class PARAM:
    """Loads running controls from a parameter file 'param.xml'
Suggest modifying param.xml by script for batch jobs, to keep
track of running conditions.
Check param.xml for variable definitions."""
    def __init__(self):
        dom = parse('param.xml')
        node = dom.getElementsByTagName('vars')[0]
        self.mcs = int(node.getAttribute('mcs'))
        self.eqs = int(node.getAttribute('eqs'))
        self.beta = float(node.getAttribute('beta'))
        self.J = float(node.getAttribute('J'))
        self.L = int(node.getAttribute('L'))
        self.ANum = int(node.getAttribute('ANum'))
        self.random = int(node.getAttribute('random'))
        self.bins = int(node.getAttribute('bins'))

class MEASURE:
    """Observable tracking for statistics.
overlap - overlaps between the current topology (ANum)
and (ANum - 1 layer)
overlap_c - number of times overlap was measured"""
    def __init__(self):
        self.overlap = 0.0
        self.overlap_c = 0
        self.fout = open('data.txt','w')

    def Reset(self):
```

```python
39            self.overlap = 0.0
40            self.overlap_c = 0
41
42  class BONDS:
43      """Bond tracking for the lattice.
44  a - first spin on a bond
45  b - second spin on a bond
46  Spin a < Spin b (typically) in the construction."""
47      def __init__(self,i,j):
48          self.a = i
49          self.b = j
50
51      def __repr__(self):
52          return "(%d, %d)" % (self.a, self.b)
53
54  class SIM:
55      """Class containing everything needed to do the ferromagnetic Ising
              model."""
56      def __init__(self):
57          """Initialize the simulation with random bits and
58  using the data from the parameter file (notably, ANum).
59  Bits on both layers of the simulation are set to be the same
60  initially to make connecting arbitrary regions up trivial."""
61          self.p = PARAM()
62          self.m = MEASURE()
63          random.seed(self.p.random)
64          L = self.p.L
65          self.nSpins = L*L
66          self.Spins = []
67          self.Spins.append([random.randint(0,1) for i in range(self.
              nSpins)])
68          self.Spins.append(list(self.Spins[0]))
69          self.inA = [i<self.p.ANum for i in range(self.nSpins)]
70          self.Bonds = []
71          for i in range(self.nSpins):
72              if (i%L)==L-1:
73                  j = i+1-L
74              else:
75                  j = i+1
76              self.Bonds.append(BONDS(i,j))
77              if ((i/L)%L)==L-1:
78                  j = i+L-L*L
79              else:
80                  j = i+L
81              self.Bonds.append(BONDS(i,j))
```

```python
82              self.Energy = self.Calc_E()
83
84      def Calc_E(self):
85          """Calculate the energy by summing over all bonds."""
86          E = 0
87          for i in self.Bonds:
88              E += self.p.J*(((((self.Spins[0][i.a] + self.Spins[0][i.b])
                    %2)*2)-1)
89              E += self.p.J*(((((self.Spins[1][i.a] + self.Spins[1][i.b])
                    %2)*2)-1)
90          return E
91
92      def __repr__(self):
93          """Show the configuration of both layers of the lattice."""
94          r = "E = %0.2f\n" % self.Energy
95          r = "ANum = %d\n\n" % self.p.ANum
96          for i in range(self.p.L):
97              for j in range(self.p.L):
98                  r += "%d " % (self.Spins[0][i*self.p.L + j])
99              r += "\n"
100         r += "\n"
101         for i in range(self.p.L):
102             for j in range(self.p.L):
103                 r += "%d " % (self.Spins[1][i*self.p.L + j])
104             r += "\n"
105         return r
106
107     def Bonds_on_spin(self,i):
108         """Return a list of all the bonds connected to spin i."""
109         bond1 = self.Bonds[2*i]
110         bond2 = self.Bonds[2*i+1]
111         if (i%self.p.L)==0:
112             bond3 = self.Bonds[2*(i-1+self.p.L)]
113         else:
114             bond3 = self.Bonds[2*(i-1)]
115         if (i/self.p.L)==0:
116             bond4 = self.Bonds[2*(i-self.p.L+self.p.L*self.p.L)+1]
117         else:
118             bond4 = self.Bonds[2*(i-self.p.L)+1]
119         return [bond1,bond2,bond3,bond4]
120
121     def MCS(self):
122         """The usual Monte Carlo step, attempting to flip a single spin
                as an update."""
123         spin = random.randint(0,self.nSpins-1)
```

```python
124            dE = 0
125            if self.inA[spin]:
126                z = random.randint(0,1)
127                for i in self.Bonds_on_spin(spin):
128                    dE += -2*self.p.J*((((self.Spins[z][i.a] + self.Spins[z
                        ][i.b])%2)*2)-1)
129                if random.random() < math.exp(-self.p.beta*dE):
130                    self.Spins[z][spin] = (self.Spins[z][spin] + 1)%2
131                    self.Energy += dE
132            else:
133                for i in self.Bonds_on_spin(spin):
134                    dE += -2*self.p.J*((((self.Spins[0][i.a] + self.Spins
                        [0][i.b])%2)*2)-1)
135                    dE += -2*self.p.J*((((self.Spins[1][i.a] + self.Spins
                        [1][i.b])%2)*2)-1)
136                if random.random() < math.exp(-self.p.beta*dE):
137                    self.Spins[0][spin] = (self.Spins[0][spin] + 1)%2
138                    self.Spins[1][spin] = (self.Spins[1][spin] + 1)%2
139                    self.Energy += dE
140            if __debug__:
141                if self.Energy != self.Calc_E():
142                    raise Exception("Energy does not match after update")
143
144        def WolffAddNeighbor(self,bonds,spinlist,(spin,z)):
145            """Recursive helper function called by Wolff."""
146            if (spin,z) in spinlist:
147                return
148            spinlist.append((spin,z))
149            bondlist = self.Bonds_on_spin(spin)
150            nBond = -1
151            for i in bondlist:
152                nBond += 1
153                if ((i.a,z) in spinlist) & ((i.b,z) in spinlist):
154                    if (i,z) in bonds:
155                        bonds.remove((i,z))
156                    continue
157                if self.Spins[z][i.a] != self.Spins[z][i.b]:
158                    bonds.append((i,z))
159                else:
160                    if random.random() > (1-math.exp(-self.p.beta*2*self.p.J
                        )):
161                        bonds.append((i,z))
162                    else:
163                        if nBond/2  == 0:
164                            addme = i.b
```

```
165                        else:
166                            addme = i.a
167                        if self.inA[addme]:
168                            self.WolffAddNeighbor(bonds,spinlist,(addme,z))
169                        else:
170                            self.WolffAddNeighbor(bonds,spinlist,(addme,0))
171                            self.WolffAddNeighbor(bonds,spinlist,(addme,1))
172
173    def Wolff(self):
174        """Wolff algorithm: starting on a random spin (in a random layer
              ), build
175 a bluster of spins that can all be flipped.
176 This helps especially at very low temperatures and near criticality to
      speed
177 the equilibriation and reduce autocorrelation time."""
178        spin = random.randint(0,self.nSpins-1)
179        bonds = []
180        spinlist = []
181        if self.inA[spin]:
182            z = random.randint(0,1)
183            self.WolffAddNeighbor(bonds,spinlist,(spin,z))
184        else:
185            self.WolffAddNeighbor(bonds,spinlist,(spin,0))
186            self.WolffAddNeighbor(bonds,spinlist,(spin,1))
187        dE = 0
188        for (i,z) in bonds:
189            dE += -2*self.p.J*((((self.Spins[z][i.a] + self.Spins[z][i.b
                  ])%2)*2)-1)
190        self.Energy += dE
191        for (i,z) in spinlist:
192            self.Spins[z][i] = (self.Spins[z][i] + 1)%2
193        if __debug__:
194            if self.Energy != self.Calc_E():
195                raise Exception("Energy does not match after update")
196
197    def Try_change(self):
198        """Use a transfer matrix approach to look at the difference in
              partiton functions
199 when examining the difference between a layer being or not being in the
      simulation."""
200        # We'll look at the ration of Z_new / Z_current, where Z_new has
              one extra layer of connected spins in it
201        # starting from Anum+1
202        # Start by calculating Z_current with the free layer being the
              free variable
```

```
203            # Since the two layers don't interact, the total partition
                   function over those layers is simply the
204            # product of the upper and lower partiton function.
205            spins = [self.p.ANum-1-i for i in range(self.p.L)]
206            Z_new = 0.0
207            Z_curr = 0.0
208            temp_Mat = np.matrix([[1.,0.],[0.,1.]])
209            for i in range(2):
210                for s in range(len(spins)):
211                    # Build the matrix connecting this spin and the next
212                    bonds_a = self.Bonds_on_spin(spins[s])
213                    bonds_b = self.Bonds_on_spin(spins[(s+1)%len(spins)])
214                    tMat = np.matrix([[np.exp(-1.0*self.p.beta*self.p.J
                           *(-1.0 + ((self.Spins[i][bonds_a[1].b]+0)%2+(self.
                           Spins[i][bonds_a[3].a]+0)%2-1.0) + ((self.Spins[i][
                           bonds_b[1].b]+0)%2+(self.Spins[i][bonds_b[3].a]+0)
                           %2-1.0))),
215                    np.exp(-1.0*self.p.beta*self.p.J*(1.0 + ((self.Spins[i][
                           bonds_a[1].b]+0)%2+(self.Spins[i][bonds_a[3].a]+0)
                           %2-1.0) + ((self.Spins[i][bonds_b[1].b]+1)%2+(self.
                           Spins[i][bonds_b[3].a]+1)%2-1.0)))],
216                    [np.exp(-1.0*self.p.beta*self.p.J*(1.0 + ((self.Spins[i
                           ][bonds_a[1].b]+1)%2+(self.Spins[i][bonds_a[3].a]+1)
                           %2-1.0) + ((self.Spins[i][bonds_b[1].b]+0)%2+(self.
                           Spins[i][bonds_b[3].a]+0)%2-1.0))),
217                    np.exp(-1.0*self.p.beta*self.p.J*(-1.0 + ((self.Spins[i
                           ][bonds_a[1].b]+1)%2+(self.Spins[i][bonds_a[3].a]+1)
                           %2-1.0) + ((self.Spins[i][bonds_b[1].b]+1)%2+(self.
                           Spins[i][bonds_b[3].a]+1)%2-1.0)))]])
218                    temp_Mat = temp_Mat * tMat/tMat.max()
219                    Z_new += np.log(tMat.max())
220                Z_new += np.log(temp_Mat.trace()[0,0])
221            temp_Mat = np.matrix([[1.,0.],[0.,1.]])
222            for s in range(len(spins)):
223                bonds_a = self.Bonds_on_spin(spins[s])
224                bonds_b = self.Bonds_on_spin(spins[(s+1)%len(spins)])
225                tMat = np.matrix([[np.exp(-1.0*self.p.beta*self.p.J*(-2.0 +
                       ((self.Spins[0][bonds_a[1].b]+0)%2 + (self.Spins[0][
                       bonds_a[3].a]+0)%2 + (self.Spins[0][bonds_b[1].b]+0)%2 +
                       (self.Spins[0][bonds_b[3].a]+0)%2 + (self.Spins[1][
                       bonds_a[1].b]+0)%2 + (self.Spins[1][bonds_a[3].a]+0)%2 +
                       (self.Spins[1][bonds_b[1].b]+0)%2 + (self.Spins[1][
                       bonds_b[3].a]+0)%2-4.0))),
226                np.exp(-1.0*self.p.beta*self.p.J*(-2.0 + ((self.Spins[0][
                       bonds_a[1].b]+0)%2 + (self.Spins[0][bonds_a[3].a]+0)%2 +
```

```
                        (self.Spins[0][bonds_b[1].b]+1)%2 + (self.Spins[0][
                        bonds_b[3].a]+1)%2 + (self.Spins[1][bonds_a[1].b]+0)%2 +
                        (self.Spins[1][bonds_a[3].a]+0)%2 + (self.Spins[1][
                        bonds_b[1].b]+1)%2 + (self.Spins[1][bonds_b[3].a]+1)
                        %2-4.0)))],
227                     [np.exp(-1.0*self.p.beta*self.p.J*(-2.0 + ((self.Spins[0][
                        bonds_a[1].b]+1)%2 + (self.Spins[0][bonds_a[3].a]+1)%2 +
                        (self.Spins[0][bonds_b[1].b]+0)%2 + (self.Spins[0][
                        bonds_b[3].a]+0)%2 + (self.Spins[1][bonds_a[1].b]+1)%2 +
                        (self.Spins[1][bonds_a[3].a]+1)%2 + (self.Spins[1][
                        bonds_b[1].b]+0)%2 + (self.Spins[1][bonds_b[3].a]+0)
                        %2-4.0))),
228                     np.exp(-1.0*self.p.beta*self.p.J*(-2.0 + ((self.Spins[0][
                        bonds_a[1].b]+1)%2 + (self.Spins[0][bonds_a[3].a]+1)%2 +
                        (self.Spins[0][bonds_b[1].b]+1)%2 + (self.Spins[0][
                        bonds_b[3].a]+1)%2 + (self.Spins[1][bonds_a[1].b]+1)%2 +
                        (self.Spins[1][bonds_a[3].a]+1)%2 + (self.Spins[1][
                        bonds_b[1].b]+1)%2 + (self.Spins[1][bonds_b[3].a]+1)
                        %2-4.0)))]])
229                 temp_Mat = temp_Mat * tMat/tMat.max()
230                 Z_curr += np.log(tMat.max())
231             Z_curr += np.log(temp_Mat.trace()[0,0])
232             self.m.overlap += np.exp(Z_curr - Z_new)
233             self.m.overlap_c += 1
234
235 def main():
236     sim = SIM()
237     for i in range(sim.p.eqs):
238         sim.MCS()
239         if (i%sim.nSpins)==0:
240             sim.Wolff()
241     for j in range(sim.p.bins):
242         for i in range(sim.p.mcs):
243             sim.MCS()
244             if (i%sim.nSpins)==0:
245                 sim.Wolff()
246                 sim.Try_change()
247         sim.m.fout.write('%0.8e \n' % (sim.m.overlap / sim.m.overlap_c))
248         sim.m.fout.flush()
249         sim.m.Reset()
250
251 if __name__ == "__main__":
252     main()
```

Using the parameter file below.

```
1  <params >
2  <!--
3  L - Linear size of the square system
4  ANum - the size of region A (in spins)
5  grow - how large we attempt to grow the system (scales as 2^N)
6  beta - the inverse temperature of the simulation
7  J - the coulping strength
8  mcs - Number of total monte carlo steps
9  eqs - Number of total equilibration steps
10  bins - Number of bins of data collection
11  random - random seed for the simulation
12  -->
13  <vars
14  L="20"
15  ANum="aaa"
16  beta="bbb"
17  J="1.0"
18  mcs="1000"
19  eqs="10000"
20  bins="200"
21  random="6546"
22  />
23  </params >
```

It can be run with this code, generating a script runme to be run locally for Unix or Mac systems supporting bash shell.

```
1  import numpy as np
2
3  fout = open('runme','w')
4  regions =
       [400,380,360,340,320,300,280,260,240,220,200,180,160,140,120,100,80,60,40,20]

5
6  numsteps = int((5.-1.)/0.2)
7  temps = [1.0+0.2*i for i in range(numsteps+1)]
8  temps[-1:] = [np.exp(np.log(5.) + 0.1*i) for i in range(1,10)]
9  temps = np.array(temps)
10  temps = 1./temps
11  temps.sort()
12
13  for b in temps:
14      fout.write('mkdir %4f\n' % b)
15      fout.write('cd %4f\n' % b)
16      for r in regions:
```

```
17        fout.write('mkdir %03d\n' % r)
18        fout.write('cd %03d\n' % r)
19        fout.write('cp ../../sim.py .\n')
20        fout.write("sed -e's/bbb/%4f/' -e's/aaa/%d/' <../../param.xml >
              param.xml\n" % (b,r))
21        fout.write('python sim.py\n')
22        fout.write('cd ..\n')
23     fout.write('cd ..\n')
```

Finally, this code will analyze and create a plot of the results

```
1  import numpy as np
2
3  from matplotlib import use
4  use('agg')
5
6  from matplotlib import rc, rcParams
7  rc('font',**{'family':'serif','serif':['Computer Modern'], 'size':24})
8  rc('text', usetex=True)
9  rc('xtick.major',pad=10)
10 rc('ytick.major',pad=10)
11 c1,c2,c3,c4 = '#007c7c','#ce5e00','#430a8c','#cebd00'
12 rcParams['axes.color_cycle'] = [c1,c2,c3,c4]
13
14 from matplotlib import pyplot as plt
15
16 regions = [100,90,80,70,60,50,40,30,20,10]
17 numsteps = int((5.-1.)/0.2)
18 temps = [1.0+0.2*i for i in range(numsteps+1)]
19 temps[-1:] = [np.exp(np.log(5.) + 0.1*i) for i in range(1,10)]
20 temps = np.array(temps)
21 temps = 1./temps
22 temps.sort()
23
24 fout_all = open('all_data.txt','w')
25 fout_f = open('plot_data.txt','w')
26 plot_me = []
27
28 plt.figure(figsize=(8,4))
29
30 for b in temps:
31     count = 0
32     p = 1.0
33     pe = 0.0
34     for r in regions:
35         filename = '%4f/%03d/data.txt' % (b,r)
```

132

```
36              try:
37                  temp_data = np.loadtxt(filename)
38              except IOError:
39                  break
40              avg = temp_data.mean()
41              err = temp_data.std()/np.sqrt(len(temp_data))
42              count += 1
43              p = p*avg
44              pe = pe + err/avg
45              fout_all.write('%4f %02d %8f %8f\n' % (b,count,-np.log(p),pe))
46              if r == 60:
47                  mid_data = (p,pe)
48              if r == 10:
49                  final_data = (p,pe)
50          fout_f.write('%4f %8f %8f\n' % (b,-2*np.log(mid_data[0])+np.log(
                final_data[0]),(mid_data[1]**2+final_data[1]**2)**0.5))
51          plot_me.append((b,-2*np.log(mid_data[0])+np.log(final_data[0]),(
                mid_data[1]**2+final_data[1]**2)**0.5))
52  fout_all.close()
53  fout_f.close()
54
55  plot_me = np.array(plot_me)
56
57  plt.gca().set_xscale('log')
58  plt.errorbar(1.0/plot_me[:,0],plot_me[:,1],yerr=plot_me[:,2])
59  plt.xlabel('Temperature')
60  plt.ylabel('$I_2$')
61  plt.xlim(xmax=10)
62  plt.ylim(ymin=0)
63  plt.savefig('transfer_test.pdf', bbox_inches='tight')
```

B.3   Numerical approaches to fitting

Although not directly related to the topic of information or entanglement, one of the minor developments was the adaptation of optimization schemes to provide good estimates for the error of a fitting procedure. Good methods for this exist, typically by using the covariant matrix and residuals to estimate the error, but these often rely on the assumption that the error is small. Given that we rely so much on Monte Carlo approaches, it is convenient to also use them for the calculation of error.

The concept is simple, given a set of data and some measure of errors, typically the standard deviation of the mean, determine the best fit parameters and a reasonable error

bar for the fit parameters. Reasonable here would suggest that, larger error bars on initial data should typically suggest larger error bars on fit parameters. A simple algorithm to achieve this uses the following algorithm

1. Generate a new data set using by adding Gaussian noise to the measured data. The width of the Gaussian distribution for each data point should be the "error bar" for each data point.

2. Fit this noisy data, extracting optimal fitting parameters

3. Repeat the above two steps until we have at least a hundred samples of the optimal fitting parameters.

4. Extract the mean and standard deviation of the optimal fitting parameters.

This is essentially the brute force procedure of re-sampling the data to get an estimate of error. One could even look at the distribution of the fitting parameters or the covariance of different parameters, but unless one has reason to expect these to be interesting they are not typically done. Below is a sample code showing the a general implementation of the above.

```
1  """A brute force approach to finding the error of a fitting procedure
2  by refitting the raw data with noise added.
3  Does not rely on the assumption of small errors, but does assume that
4  errors are uncorrelated.
5  Stephen Inglis, 2013"""
6
7  import numpy
8  from scipy.optimize import leastsq
9
10  def refit(x,y,dy,errf,c_guess,N_steps=100):
11      """Main module to fit the data, assuming one-dimensional data.
12  x - dependant variable of some data
13  y - independant variable
14  dy - standard deviation of the mean of the independant data
15  erry(C,x,y,dy) - error function of the fit, typically of the form (fitf
          (...) - y)/dy
16  c_guess - list of intial guess of the solution C for fitf(...), [c1,c2
          ,...]
17  N_steps - number of trials of the noise adding procedure, default 100"""
18      # Ensure the data is in numpy arrays, for easier manipulation
19      x = numpy.array(x)
20      y = numpy.array(y)
```

```python
21      dy = numpy.array(dy)
22      psum = None
23      for _i in range(N_steps):
24          # Generate data with noise, assuming each data point is
                 uncorrelated
25          ny = numpy.array([numpy.random.normal(loc=z[0],scale=z[1]) for z
                 in zip(y,dy)])
26          # Main workhorse from scipy, minimizing the sum of the squares
                 from 'errf' by modifying the first parameters of the function
27          # Insert your favourite minimization routine here
28          res = leastsq(errf, c_guess, args=(x,ny,dy), full_output=1)
29          (popt, pcov, infodict, errmsg, ier) = res
30          if psum == None:
31              psum = numpy.array(popt)
32              psum2 = numpy.array(popt)**2
33          else:
34              psum += numpy.array(popt)
35              psum2 += numpy.array(popt)**2
36      psum /= N_steps
37      psum2 /= N_steps
38      perr = (psum2 - psum**2)
39      # In the case of data without error, this prevents errors
40      perr = (perr*(perr>0))**0.5
41      return psum,perr
42
43  def main_test():
44      """A simple experiment.
45  Imagine we have 3 independent methods for calculating a value X.
46  We assume no systematic bias, but we assume they have different accuracy
47  Let us see how using a flat average and weighted average compare for
        finding the true mean."""
48      x = numpy.array([1.,2.,3.])
49      error = numpy.array([1.,5.,25.])
50      numtrials = 10
51      y = [0,0,0]
52      dy = [0,0,0]
53      for i in range(3):
54          temp = numpy.random.normal(loc=10.,scale=error[i],size=numtrials
                 )
55          y[i] = numpy.mean(temp)
56          dy[i] = numpy.std(temp)/(numtrials**0.5)
57      # 1-parameter fit, the average
58      fitf = lambda fc,fx: fc[0]
59      print main_test.__doc__
60      print 'Real mean = 10.00'
```

```
61    print 'Meaningful Error < %f' % dy[0]
62    # Using a weighted average
63    errf = lambda fc,fx,fy,fdy: (fitf(fc,fx) - fy)/fdy
64    popt,perr = refit(x,y,dy,errf,[10])
65    print 'Weighted Mean = %f +- %f' % (popt[0],perr[0])
66    # Using an unweighted (flat) average
67    errf2 = lambda fc,fx,fy,fdy: (fitf(fc,fx) - fy)
68    popt,perr = refit(x,y,dy,errf2,[10])
69    print 'Flat Mean = %f +- %f' % (popt[0],perr[0])
70
71 if __name__ == "__main__":
72    main_test()
```

By "Meaningful Error" we mean that the aggregate knowledge of multiple experiments, assuming no systematic error, should produce a result with less error than any of the individual input data. The "Meaningful Error" is the theoretical error of the most certain data point, and so by combining that with less accurate data it provides an upper bound on the error bar of the final average. Since the data is generated numerically, the actual error may be larger than this value in some fraction of cases. Note that the mean value of the fit and the optimal value of the fit may differ if the final distribution of values is not symmetric, and differentiating between the two can be important, as they do not always converge.

Appendix C

Finite size scaling of the XY model from renormalization group

Here we discuss the finite size scaling of the XY model using renormalization group (RG). This argument was originally laid out by Jason Iaconis, but was not published as part of any document. It is included here for completeness, as the finitesize scaling derived is used in matching the results from the classical XY model presented earlier.

For concreteness, we will discuss the classical XY model as an example of the XY universality class. This model is described by the Hamiltonian

$$H = -J \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j) \tag{C.1}$$

For the XY model we can define three parameters necessary in the RG procedure: the vortex fugacity $y = e^{-E_c/T}$, the temperature T, and the renormalization parameter $b = e^\ell$. Since the XY model permits arbitrarily small excitations in th form of Goldstone modes, the next non-trivial excitation corresponds to that of a vortex. A vortex corresponds to a non-zero lattice curl around some plaquette, and cannot be removed by smoothly deforming the spins locally, but can be annihilated by combination with another vortex of opposite vorticity. The low temperature phase of the XY model corresponds to the absence or vortices, or equivalently that they are bound. In the high temperature phase vortices unbind and proliferate.

The RG equations for the XY model using the above variables is

$$\frac{dy}{d\ell} = \left(2 - \frac{\pi}{T}\right)y \tag{C.2}$$

$$\frac{dT}{d\ell} = \frac{y^2}{2T}, \tag{C.3}$$

the folowing substitution is useful

$$z = y^2 \quad , \quad x = 2 - \frac{\pi}{T}, \tag{C.4}$$

giving us the new equations

$$\frac{dz}{d\ell} = 2xz \tag{C.5}$$

$$\frac{dx}{d\ell} = \frac{\pi}{T^2}\frac{dT}{d\ell} = \frac{\pi}{T^2}\frac{y^2}{2T} \tag{C.6}$$

$$\approx \frac{\pi}{2(\pi/2)^2}z = \left(\frac{2}{\pi}\right)^2 z, \tag{C.7}$$

where we use the approximation of assuming that T is near the critical temperature $T_{BKT} = \pi/2$

We can solve these last two equations to get

$$\frac{dz}{dx} = \frac{\pi^2}{2}x \tag{C.8}$$

$$z = \frac{\pi^2}{4}x^2 + \sigma. \tag{C.9}$$

Using equation, we wish to look at the behaviour of the flow near the critical point ($z = x = 0$). To do this, we will look at the equation $z(\ell) = \pi^2/4x(\ell)^2 + \sigma$ and integrate with respect to ℓ. Re-writing the solved equation we get

$$z = \frac{\pi^2}{4} + x^2 + \sigma \tag{C.10}$$

$$x = \frac{2}{\pi}\sqrt{z - \sigma} \tag{C.11}$$

$$\frac{dz}{d\ell} = 2xz \tag{C.12}$$

$$= \frac{4}{\pi}z\sqrt{z - \sigma}. \tag{C.13}$$

We wish to find the value of ℓ such that $z(\ell) = 1$

$$\int_0^\ell d\ell' = \frac{\pi}{4} \int_{z(0)}^{z(\ell)} \frac{dz'}{z'\sqrt{z' - \sigma}} \tag{C.14}$$

$$\ell = \frac{\pi}{4} \left[\frac{2\tan^{-1}\left(\sqrt{\frac{z-\sigma}{\sigma}}\right)}{\sqrt{\sigma}} \right]_\sigma^1 \tag{C.15}$$

$$= \frac{\pi}{4} \left[\frac{2\tan^{-1}\left(\sqrt{\frac{1-\sigma}{\sigma}}\right)}{\sqrt{\sigma}} - \frac{2\tan^{-1}(0)}{\sqrt{\sigma}} \right] \tag{C.16}$$

$$= \frac{\pi}{2} \frac{\tan^{-1}\left(\sqrt{\frac{1}{\sigma} - 1}\right)}{\sqrt{\sigma}}. \tag{C.17}$$

Since $\sigma = (T - T_c)/T_c$ is taken to be small, then $1/\sigma$ should be large, and the typical RG treatment is to approximate $\tan^{-1}\left(\sqrt{\frac{1}{\sigma} - 1}\right) \approx \tan^{-1}(\infty) = \frac{\pi}{2}$, set $\xi = b = e^\ell$ and solve for $t = \sigma$ as a function of ξ. We need to find the higher order contributions of ξ to t, and taylor expand around $\tan^{-1}(x)$ for large x.

$$\tan^{-1}(x) = \frac{\pi}{2} - \tan^{-1}\left(\frac{1}{x}\right) \tag{C.18}$$

$$\tan^{-1}(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 + \cdots \quad \text{for } |x| < 1. \tag{C.19}$$

Applying this approximation to Eq. (C.17) we get

$$\tan^{-1}\left(\sqrt{\frac{1-\sigma}{\sigma}}\right) \approx \frac{\pi}{2} - \sqrt{\frac{\sigma}{1-\sigma}} \tag{C.20}$$

$$\approx \frac{\pi}{2} - \sqrt{\sigma}\left(1 - \frac{1}{2}\sigma\right) \tag{C.21}$$

$$\approx \frac{\pi}{2} - \sqrt{\sigma} + \frac{1}{2}\sigma^{3/2}. \tag{C.22}$$

Plugging this in to the original form we find

$$\ell = \frac{\pi}{2} \left(\frac{\frac{\pi}{2} - \sqrt{\sigma} + \frac{1}{2}\sigma^{3/2}}{\sqrt{\sigma}} \right) \tag{C.23}$$

$$= \frac{\pi^2}{4\sqrt{\sigma}} - \frac{\pi}{2} + \frac{\pi\sigma}{4}. \tag{C.24}$$

From here, the typical procedure associates the correlation length ξ with the renormalization parameter b. Using this association and substitution the definition for $\sigma = t = (T - T_c)/T_c$ we find

$$\xi \approx e^\ell = \exp\left(\frac{c_1}{\sqrt{t}} - c_2 + c_3 t\right) \tag{C.25}$$

$$\xi = L/L_0, \quad \ln(L/L_0) = \frac{c_1}{\sqrt{t}} - c_2 + c_3 t \tag{C.26}$$

$$\ln(L/L_0') = \frac{c_1}{\sqrt{t}} + c_3 t, \tag{C.27}$$

where we have absorbed all the constants into the definition of $\ln(L_0')$ (which we will immediately rename $\ln(L_0)$). What we desire is a relationship relating L and t, so that we may make a statement about the finite size corrections to the critical temperature. We should keep in mind that our parameter t is small, and we will use this detail to ignore higher orders of t^n when convenient.

Solving the above equation for t we get

$$c_1 + c_3 t^{3/2} = \sqrt{t} \ln(L/L_0) \tag{C.28}$$

$$c_1^2 = (\sqrt{t} \ln(L/L_0) - c_3 t^{3/2})^2 \tag{C.29}$$

$$c_1^2 = \ln^2(L/L_0)t - 2c_3 \ln(L/L_0)t^2 + c_3^2 t^3 \tag{C.30}$$

If we drop the t^3 term under the assumption it is small, we get

$$2c_3 \ln(L/L_0)t^2 - \ln^2(L/L_0)t + c_1^2 = 0 \tag{C.31}$$

which we can then solve using the quadratic equation

$$t = \frac{\ln^2(L/L_0)}{4c_3 \ln(L/L_0)} \pm \frac{\sqrt{\ln^4(L/L_0) - 8c_3 c_1^2 \ln(L/L_0)}}{2c_3 \ln(L/L_0)} \tag{C.32}$$

$$4c_3 \ln(L/L_0)t = \ln^2(L/L_0) - \ln^2(L/L_0)\sqrt{1 - \frac{8c_3 c_1^2}{\ln^3(L/L_0)}} \tag{C.33}$$

$$4c_3 \ln(L/L_0)t \approx \frac{4c_3 c_1^2}{\ln^3(L/L_0)} - \frac{1}{8}\left(\frac{8c_3 c_1^2}{\ln^3(L/L_0)}\right)^2 \tag{C.34}$$

$$t \approx \frac{c_1'}{\ln^2(L/L_0)} + \frac{c_2'}{\ln^5(L/L_0)}. \tag{C.35}$$

140

This relationship gives us the finite size scaling by associating the finite size transition temperature with t given a particular value of L. Writing this out explicitly, we have

$$T_{BKT}(L) = T_{BKT}(\infty)\left[1 + \frac{c_1}{\ln^2(L/L_0)} + \frac{c_2}{\ln^5(L/L_0)}\right]. \tag{C.36}$$

This fitting form is shown in Fig. 3.5 as the two fitting curves for the finite temperature crossing, and shows good agreement with the critical temperature from other studies.

References

[1] Lev D Landau, EM Lifshitz, and LP Pitaevskii. Statistical physics (course of theo-retical physics, volume 5). *Butterworth-Heinemann,*, 3:4–1, 1984.

[2] LD Landau and VL Ginzburg. On the theory of superconductivity. *Journal of Experimental and Theoretical Physics (USSR)*, 20:1064, 1950.

[3] Xiao-Gang Wen. Vacuum degeneracy of chiral spin states in compactified space. *Physical Review B*, 40(10):7387, 1989.

[4] Xiao-Gang Wen. Chiral luttinger liquid and the edge excitations in the fractional quantum hall states. *Physical Review B*, 41(18):12838, 1990.

[5] Xiao-Gang Wen. Topological orders and edge excitations in fractional quantum hall states. *Advances in Physics*, 44(5):405–473, 1995.

[6] X.-G. Wen. Quantum orders and symmetric spin liquids. *Phys. Rev. B*, 65(16):165113, April 2002.

[7] S. V. Isakov, M. B. Hastings, and R. G. Melko. Topological entanglement entropy of a Bose-Hubbard spin liquid. *Nature Physics*, 7:772–775, October 2011.

[8] J. Iaconis, S. Inglis, A. B. Kallin, and R. G. Melko. Detecting classical phase tran-sitions with Renyi mutual information. *Phys. Rev. B*, 87(19):195134, May 2013.

[9] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev. Entanglement in quantum critical phenomena. *Phys. Rev. Lett.*, 90:227902, 2003.

[10] Pasquale Calabrese and John Cardy. Entanglement entropy and conformal field theory. *Journal of Physics A: Mathematical and Theoretical*, 42(50):504005, 2009.

[11] D. A. Abanin and E. Demler. Measuring Entanglement Entropy of a Generic Many-Body System with a Quantum Switch. *Physical Review Letters*, 109(2):020504, July 2012.

[12] John Cardy. Measuring entanglement using quantum quenches. *Phys. Rev. Lett.*, 106:150404, Apr 2011.

[13] H. Pichler, L. Bonnes, A. J. Daley, A. M. Läuchli, and P. Zoller. Thermal versus entanglement entropy: a measurement protocol for fermionic atoms with a quantum gas microscope. *New Journal of Physics*, 15(6):063003, June 2013.

[14] Michael Levin and Xiao-Gang Wen. Detecting topological order in a ground state wave function. *Phys. Rev. Lett.*, 96:110405, Mar 2006.

[15] Alexei Kitaev and John Preskill. Topological entanglement entropy. *Phys. Rev. Lett.*, 96:110404, Mar 2006.

[16] Leon Balents. Spin liquids in frustrated magnets. *Nature*, 464(7286):199–208, 2010.

[17] Z. Y. Meng, T. C. Lang, S. Wessel, F. F. Assaad, and A. Muramatsu. Quantum spin liquid emerging in two-dimensional correlated dirac fermions. *Nature*, 464(7290):847–851, 2010.

[18] Sandro Sorella, Yuichi Otsuka, and Seiji Yunoki. Absence of a spin liquid phase in the hubbard model on the honeycomb lattice. *Scientific reports*, 2, 2012.

[19] R. R. P. Singh, M. B. Hastings, A. B. Kallin, and R. G. Melko. Finite-Temperature Critical Behavior of Mutual Information. *Physical Review Letters*, 106(13):135701, April 2011.

[20] A. Pelissetto and E. Vicari. Critical phenomena and renormalization-group theory. *Physics Reports*, 368:549–727, October 2002.

[21] M. Campostrini, A. Pelissetto, P. Rossi, and E. Vicari. 25th-order high-temperature expansion results for three-dimensional Ising-like systems on the simple-cubic lattice. *Phys. Rev. E*, 65(6):066127, June 2002.

[22] Martin Hasenbusch. Monte carlo studies of the three-dimensional ising model in equilibrium. *International Journal of Modern Physics C*, 12(07):911–1009, 2001.

[23] F. Jasch and H. Kleinert. Fast-convergent resummation algorithm and critical exponents of phi^4-theory in three dimensions. *Journal of Mathematical Physics*, 42:52–73, January 2001.

[24] D. M. Sullivan, G. W. Neilson, H. E. Fischer, and A. R. Rennie. Small angle neutron scattering from D 2 O in the critical region. *Journal of Physics: Condensed Matter*, 12(15):3531, 2000.

[25] A. Haupt and J. Straub. Evaluation of the isochoric heat capacity measurements at the critical isochore of SF_6 performed during the German Spacelab Mission D-2. *Phys. Rev. E*, 59:1795–1802, Feb 1999.

[26] Shigeo Kuwabara, Hirokazu Aoyama, Haruki Sato, and Koichi Watanabe. Vapor-liquid coexistence curves in the critical region and the critical temperatures and densities of difluoromethane and pentafluoroethane. *Journal of Chemical & Engineering Data*, 40(1):112–116, 1995.

[27] P. Butera and M. Comi. N-vector spin models on the simple-cubic and the body-centered-cubic lattices: A study of the critical behavior of the susceptibility and of the correlation length by high-temperature series extended to order β^{21}. *Phys. Rev. B*, 56:8212–8240, October 1997.

[28] P. Calabrese, A. Pelissetto, and E. Vicari. Critical structure factors of bilinear fields in O(N) vector models. *Phys. Rev. E*, 65(4):046115, April 2002.

[29] J. H. Zhao, T. Song, H. P. Kunkel, X. Z. Zhou, R. M. Roshko, and Gwyn Williams. La 0.95 mg 0.05 mno 3 : an ideal ferromagnetic system? *Journal of Physics: Condensed Matter*, 12(30):6903, 2000.

[30] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982.

[31] S. Boixo, T. F. Rønnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer. Quantum annealing with more than one hundred qubits. *ArXiv e-prints*, April 2013.

[32] R. Raussendorf, J. Harrington, and K. Goyal. Topological fault-tolerance in cluster state quantum computation. *New Journal of Physics*, 9:199, June 2007.

[33] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma. Non-Abelian anyons and topological quantum computation. *Reviews of Modern Physics*, 80:1083–1159, July 2008.

[34] Christof Zalka. Simulating quantum systems on a quantum computer. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):313–322, 1998.

[35] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Reviews of Modern Physics*, 82:277–306, January 2010.

[36] P. Calabrese and J. Cardy. Entanglement Entropy and Quantum Field Theory: A Non-Technical Introduction. *International Journal of Quantum Information, vol. 4 (3), p. 429*, 4:429, June 2006.

[37] H. Ju, A. B. Kallin, P. Fendley, M. B. Hastings, and R. G. Melko. Entanglement scaling in two-dimensional gapless systems. *Phys. Rev. B*, 85(16):165121, April 2012.

[38] M. A. Metlitski and T. Grover. Entanglement Entropy of Systems with Spontaneously Broken Continuous Symmetry. *ArXiv e-prints*, December 2011.

[39] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, Nov 1992.

[40] Steven R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48:10345–10356, Oct 1993.

[41] Steven R. White. Strongly correlated electron systems and the density matrix renormalization group. *Physics Reports*, 301(13):187 – 204, 1998.

[42] G. Sierra and M. A. Martin-Delgado. The Density Matrix Renormalization Group, Quantum Groups and Conformal Field Theory. *eprint arXiv:cond-mat/9811170*, November 1998.

[43] M. B. Hastings, I. González, A. B. Kallin, and R. G. Melko. Measuring Renyi Entanglement Entropy in Quantum Monte Carlo Simulations. *Physical Review Letters*, 104(15):157201, April 2010.

[44] Ulli Wolff. Collective Monte Carlo updating for spin systems. *Phys. Rev. Lett.*, 62:361–364, Jan 1989.

[45] Anders W. Sandvik. Stochastic series expansion method with operator-loop update. *Phys. Rev. B*, 59:R14157–R14160, Jun 1999.

[46] Olav F. Syljuåsen and Anders W. Sandvik. Quantum Monte Carlo with directed loops. *Phys. Rev. E*, 66:046701, Oct 2002.

[47] Anders W. Sandvik. Stochastic series expansion method for quantum Ising models with arbitrary interactions. *Phys. Rev. E*, 68:056701, Nov 2003.

[48] Makoto Matsumoto and Takuji Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.*, 8(1):3–30, January 1998.

[49] Yosihiko Ogata. A monte carlo method for high dimensional integration. *Numerische Mathematik*, 55(2):137–157, 1989.

[50] William H Press. *Numerical recipes 3rd edition: The art of scientific computing.* Cambridge university press, 2007.

[51] Persi Diaconis. The markov chain monte carlo revolution. *Bulletin of the American Mathematical Society*, 46(2):179–205, 2009.

[52] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, 21:1087–1092, June 1953.

[53] Julia Wildeboer and Alexander Seidel. Correlation functions in su(2)-invariant resonating-valence-bond spin liquids on nonbipartite lattices. *Phys. Rev. Lett.*, 109:147208, Oct 2012.

[54] Heiko Rieger and Naoki Kawashima. Application of a continuous time cluster algorithm to the two-dimensional random quantum ising ferromagnet. *The European Physical Journal B-Condensed Matter and Complex Systems*, 9(2):233–236, 1999.

[55] V. S. Dotsenko, M. Picco, P. Windey, G. Harris, E. Martinec, and E. Marinari. Self-avoiding surfaces in the 3d Ising model. *Nuclear Physics B*, 448:577–620, February 1995.

[56] Max A. Metlitski, Carlos A. Fuertes, and Subir Sachdev. Entanglement entropy in the $o(n)$ model. *Phys. Rev. B*, 80:115122, Sep 2009.

[57] A. Renyi. *Proc. of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, 1960:547, 1961.

[58] Peter W. Shor. Capacities of quantum channels and how to find them. *Mathematical Programming*, 97(1-2):311–335, 2003.

[59] J. Wilms, J. Vidal, F. Verstraete, and S. Dusuel. Finite-temperature mutual information in a simple phase transition. *Journal of Statistical Mechanics: Theory and Experiment*, 1:23, January 2012.

[60] J. Wilms, M. Troyer, and F. Verstraete. Mutual information in classical spin models. *Journal of Statistical Mechanics: Theory and Experiment*, 10:11, October 2011.

[61] Jason Iaconis. Analytical and Numerical Studies of 2D XY Models with Ring Exchange. Master's thesis, University of Waterloo, 2012.

[62] Barry M. McCoy and Tai Tsun Wu. Theory of toeplitz determinants and the spin correlations of the two-dimensional ising model. iv. *Phys. Rev.*, 162:436–475, Oct 1967.

[63] Helen Au-Yang and Michael E. Fisher. Bounded and inhomogeneous ising models. ii. specific-heat scaling function for a strip. *Phys. Rev. B*, 11:3469–3487, May 1975.

[64] Roger G. Melko, Ann B. Kallin, and Matthew B. Hastings. Finite-size scaling of mutual information in monte carlo simulations: Application to the spin-$\frac{1}{2}$ xxz model. *Phys. Rev. B*, 82:100409, Sep 2010.

[65] S. Inglis and R. G. Melko. Wang-Landau method for calculating Rényi entropies in finite-temperature quantum Monte Carlo simulations. *Phys. Rev. E*, 87(1):013306, January 2013.

[66] R. G. Melko, A. W. Sandvik, and D. J. Scalapino. Aspect-ratio dependence of the spin stiffness of a two-dimensional XY model. *Phys. Rev. B*, 69(1):014509, January 2004.

[67] R. G. Melko, A. W. Sandvik, and D. J. Scalapino. Two-dimensional quantum XY model with ring exchange and external field. *Phys. Rev. B*, 69(10):100408, March 2004.

[68] P. Calabrese and J. Cardy. Entanglement entropy and quantum field theory. *Journal of Statistical Mechanics: Theory and Experiment*, 6:2, June 2004.

[69] Y. Iba. Extended Ensemble Monte Carlo. *International Journal of Modern Physics C*, 12:623–656, 2001.

[70] Fugao Wang and D. P. Landau. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.*, 86:2050–2053, 2001.

[71] Matthias Troyer, Stefan Wessel, and Fabien Alet. Flat histogram methods for quantum systems: Algorithms to overcome tunneling problems and calculate the free energy. *Phys. Rev. Lett.*, 90(12):120201, Mar 2003.

[72] Stefan Wessel, Norbert Stoop, Emanuel Gull, Simon Trebst, and Matthias Troyer. Optimized broad-histogram ensembles for the simulation of quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(12):P12005, 2007.

[73] R. E. Belardinelli, S. Manzi, and V. D. Pereyra. Analysis of the convergence of the 1?t and wang-landau algorithms in the calculation of multidimensional integrals. *Phys. Rev. E*, 78(6):067701, Dec 2008.

[74] A. B. Kallin, K. Hyatt, R. R. P. Singh, and R. G. Melko. Entanglement at a Two-Dimensional Quantum Critical Point: A Numerical Linked-Cluster Expansion Study. *Physical Review Letters*, 110(13):135702, March 2013.

[75] R. R. P. Singh, R. G. Melko, and J. Oitmaa. Thermodynamic singularities in the entanglement entropy at a two-dimensional quantum critical point. *Phys. Rev. B*, 86(7):075106, August 2012.

[76] M. Molkaraie and H.-A. Loeliger. Partition Function of the Ising Model via Factor Graph Duality. *ArXiv e-prints*, July 2013.

[77] S. Humeniuk and T. Roscilde. Quantum Monte Carlo calculation of entanglement Rényi entropies for generic quantum systems. *Phys. Rev. B*, 86(23):235116, December 2012.

[78] Anders W. Sandvik. Ground state projection of quantum spin systems in the valence-bond basis. *Phys. Rev. Lett.*, 95(20):207203, Nov 2005.

[79] S. Inglis and R. G. Melko. Entanglement at a Two-Dimensional Quantum Critical Point: a T=0 Projector Quantum Monte Carlo Study. *ArXiv e-prints*, May 2013.

[80] Michael M. Wolf. Violation of the entropic area law for fermions. *Phys. Rev. Lett.*, 96:010404, 2006.

[81] Dimitri Gioev and Israel Klich. Entanglement entropy of fermions in any dimension and the widom conjecture. *Phys. Rev. Lett.*, 96:100503, 2006.

[82] Luca Bombelli, Rabinder K. Koul, Joohan Lee, and Rafael D. Sorkin. Quantum source of entropy for black holes. *Phys. Rev. D*, 34:373–383, 1986.

[83] Mark Srednicki. Entropy and area. *Phys. Rev. Lett.*, 71(5):666–669, Aug 1993.

[84] A. J. A. James and R. M. Konik. Understanding the entanglement entropy and spectra of 2D quantum systems through arrays of coupled 1D chains. *Phys. Rev. B*, 87(24):241103, June 2013.

[85] Jean-Marie Stephan, Hyejin Ju, Paul Fendley, and Roger G Melko. Entanglement in gapless resonating-valence-bond states. *New Journal of Physics*, 15(1):015004, 2013.

[86] L. Taddia, J. C. Xavier, F. C. Alcaraz, and G. Sierra. Entanglement entropies in conformal systems with boundaries. 2013.

[87] Eduardo Fradkin and Joel E. Moore. Entanglement entropy of 2d conformal quantum critical points: Hearing the shape of a quantum drum. *Phys. Rev. Lett.*, 97:050404, 2006.

[88] H. Casini and M. Huerta. Universal terms for the entanglement entropy in dimensions. *Nuclear Physics B*, 764(3):183 – 201, 2007.

[89] Max A. Metlitski and Tarun Grover. Entanglement entropy of systems with spontaneously broken continuous symmetry. 2011.

[90] Michael P. Zaletel, Jens H. Bardarson, and Joel E. Moore. Logarithmic terms in entanglement entropies of 2d quantum critical points and shannon entropies of spin chains. *Phys. Rev. Lett.*, 107:020402, Jul 2011.

[91] Jean-Marie Stéphan, Grégoire Misguich, and Vincent Pasquier. Phase transition in the rényi-shannon entropy of luttinger liquids. *Phys. Rev. B*, 84:195128, Nov 2011.

[92] Ann B. Kallin, Matthew B. Hastings, Roger G. Melko, and Rajiv R. P. Singh. Anomalies in the entanglement properties of the square-lattice heisenberg model. *Phys. Rev. B*, 84:165134, 2011.

[93] Yifei Shi, Austen Lamacraft, and Paul Fendley. Boson pairing and unusual criticality in a generalized xy model. *Phys. Rev. Lett.*, 107:240601, Dec 2011.

[94] Adam Nahum, J. T. Chalker, P. Serna, M. Ortuño, and A. M. Somoza. 3d loop models and the cp^{n-1} sigma model. *Phys. Rev. Lett.*, 107:110601, Sep 2011.

[95] Scott D. Geraedts and Olexei I. Motrunich. Line of continuous phase transitions in a three-dimensional u(1) loop model with $1/r^2$ current-current interactions. *Phys. Rev. B*, 85:144303, Apr 2012.

[96] D. Charrier, F. Alet, and P. Pujol. Gauge Theory Picture of an Ordering Transition in a Dimer Model. *Physical Review Letters*, 101(16):167205, October 2008.

[97] G. Chen, J. Gukelberger, S. Trebst, F. Alet, and L. Balents. Coulomb gas transitions in three-dimensional classical dimer models. *Phys. Rev. B*, 80(4):045112, July 2009.

[98] B. Swingle and T. Senthil. Structure of entanglement at deconfined quantum critical points. *Phys. Rev. B*, 86(15):155131, October 2012.

[99] A. P. Young and Helmut G. Katzgraber. Absence of an Almeida-Thouless line in three-dimensional spin glasses. *Phys. Rev. Lett.*, 93:207203, Nov 2004.

[100] Helmut G. Katzgraber, Derek Larson, and A. P. Young. Study of the de Almeida-Thouless line using power-law diluted one-dimensional Ising spin glasses. *Phys. Rev. Lett.*, 102:177205, Apr 2009.

[101] L. Leuzzi, G. Parisi, F. Ricci-Tersenghi, and J. J. Ruiz-Lorenzo. Ising spin-glass transition in a magnetic field outside the limit of validity of mean-field theory. *Phys. Rev. Lett.*, 103:267201, Dec 2009.

[102] A. Coniglio and M. Nicodemi. The jamming transition of granular media. *Journal of Physics Condensed Matter*, 12:6601–6610, July 2000.

[103] C. Castelnovo, C. Chamon, and D. Sherrington. Quantum mechanical and information theoretic view on classical glass transitions. *Phys. Rev. B*, 81(18):184303, May 2010.

[104] M. Harré and T. Bossomaier. Phase-transition-like behaviour of information measures in financial markets. *Europhysics Letters*, 87(1):18009, 2009.