
Detecting Phase Transitions in Anisotropic Heisenberg Models with the Stochastic Series Expansion

Johannes Hauschild



Master's Thesis
Theoretical and Mathematical Physics
LMU Munich

October 2014

Supervisor: Priv. Doz. Dr. Fabian Heidrich-Meisner

Abstract

We use the stochastic series expansion (SSE), a quantum Monte Carlo method, to study phase transitions in the two dimensional XXZ -model, i.e., the Heisenberg model with explicitly broken symmetry. Beside the uniform square lattice we also study the crossover to one dimension with different coupling in x and y direction and a pattern of strongly coupled two-site dimers embedded in a square lattice with weaker couplings. In the latter we find a quantum phase transition which we examine with a trial state. After a discussion of the usual methods to extract the critical temperature, we reproduce crossings near T_c in the Rényi mutual information caused by sub-leading terms [1]. We observe similar crossings in spatially anisotropic systems. Since the Rényi mutual information does not rely on a definition of an order parameter, this technique might be especially interesting to detect unconventional phase transitions. Motivated by recent experiments with ultracold atoms [2], we finally consider how the phase transition manifests in projections onto local singlet and triplet states on certain bonds. We illustrate a close relation between the projections and the energy, which shows that they share the same singular behavior.

Contents

1	Introduction	7
1.1	Experiments with Ultracold Quantum Gases	9
1.2	Outline	12
2	Heisenberg Model	13
2.1	Derivation from Hubbard model	13
2.2	Mermin-Wagner Theorem	16
2.3	Anisotropic Coupling	17
2.4	Single Heisenberg Dimer	18
2.5	Spatial Anisotropies	19
3	Method: Stochastic Series Expansion	21
3.1	Classical Monte Carlo	21
3.2	Stochastic Series Expansion: Mapping to Classical Configurations	26
3.3	Sampling	31
3.3.1	Diagonal Updates	32
3.3.2	Deterministic Loop updates	33
3.3.3	Directed Loop Updates	35
3.4	Observables and Improved Estimators	40
3.5	Program Verification	42
4	Phase Transitions	45
4.1	Correlation Length	48
4.2	Scaling Hypothesis	51
4.3	Finite Size Scaling	52
4.3.1	Binder Cumulant	56
4.4	1D-2D-Crossover	58
4.5	Dimers: Quantum Phase Transition	63
5	Mutual Information	69
5.1	Some Basics of Information Theory	69
5.1.1	Mutual information	72
5.1.2	Rényi Entropies	75
5.2	Replica Trick	76
5.3	Ratio Trick	89
6	Singlet and Triplet Projectors	95
6.1	SSE Estimators for Singlet and Triplet Projectors	95
6.2	Critical Behavior of Projectors at Phase Transitions	99

6.3	1D-2D-Crossover	107
6.4	Coupled Dimers	111
7	Summary	115
A	SSE Estimators for parts of the Hamiltonian	117
B	Derivation of Formulas for Generalized Susceptibilities	118
C	Energy of the Trial State for Coupled Dimers	120
D	Perturbation Theory for large Δ	122
E	List of Tables	125
F	List of Figures	126
G	Bibliography	128

1 Introduction

The phenomenon of magnetism has been known for a very long time: the word stems from a region called $\mu\alpha\gamma\eta\sigma\acute{\iota}\alpha$ (Magnesia) in ancient Greece where natural magnet stones attracting and repelling each other were found and compasses were used for navigation since many centuries.

The classical Ising model [3] is the simplest model describing how a short-ranged interaction of spins can lead to long-range order, such that – in the ferromagnetic case – a macroscopic magnetization may arise due to an alignment of small magnetic moments. Due to its simplicity, the two dimensional (2D) Ising model serves as prototypical example for phase transitions in many text books and lectures; it is one of the very rare examples for which an exact, analytical solution is known [4].

However, the assumption of Ising that each spin may only have two possible states clearly is not valid in real solids. The interaction between different electron spins (relevant for the magnetic properties of a solid) is not even of a static nature: instead, an effective spin interaction stems from kinetic exchange processes (sometimes called ‘superexchange’) of electrons and the Pauli principle [5]. We will discuss this mechanism in section 2.1 in detail. The result is the Heisenberg Hamiltonian:

$$\mathcal{H} = J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j. \quad (1.1)$$

Here, $\langle i,j \rangle$ denotes nearest neighbors on a lattice, \vec{S}_i is the (quantum mechanical) spin operator on site i , and J is the coupling constant. Bethe gave the exact eigenvalues and -vectors of this Hamiltonian on a chain, i.e., in one dimension, in 1931 [6]. Depending on the sign of J , neighboring spins tend to align parallel ($J < 0$, ferromagnetic) or anti-parallel ($J > 0$, antiferromagnetic) at low temperatures. The Mermin-Wagner theorem [7], which we discuss in section 2.2, rules out long-range order at finite temperature in one and two dimensions: due to the continuous symmetry there exist Goldstone modes, i.e., massless excitations where the direction of spins changes slowly over large scales. Depending on the geometry of the lattice, the ground state of the two dimensional Heisenberg model however can possess long-range order, e.g., in the simplest case of a square lattice. [8]. In geometrically frustrated lattices where the classical ground state is highly degenerate, e.g., the Kagome lattice, no magnetic long-range order is found. A detailed discussion of the influence of the lattice can be found in Ref. [9].

The quantum nature of the interactions can not be neglected. In the antiferromagnetic case the classical Néel states are no longer eigenstates of the Hamiltonian; even for the simple case of a square lattice, quantum fluctuations reduce the value of the staggered magnetization to approximately 60 % of the maximum (classical) value (on a square lattice). In other cases, they can completely destroy the magnetic long-range order in the ground state, e.g., in so-called spin liquids or resonating valence bond states [10]. If some bonds are stronger than others, a so-called valence bond crystal can form, where

singlets are arranged in a regular pattern [11]. The different ground states are connected by quantum phase transitions [12], i.e., phase transitions which are caused by quantum fluctuations rather than thermal ones and driven by a change of a parameter in the Hamiltonian rather than the temperature. Some of these transitions are particularly interesting as they go beyond the Ginzburg-Landau theory; there are exotic phases without a local order parameter. Quantum spin models can be used to study such (exotic) quantum phase transitions, i.e., they take the same role of prototypical examples which the Ising model has for thermal phase transitions.

The interest in magnetism of strongly correlated electrons has been boosted by the discovery of high- T_c superconductors [13]. In contrast to the conventional superconductors, the formation of Cooper pairs in these materials is not caused by electron phonon interaction, but apparently spin fluctuations play an important role [14]. Common to these materials are two dimensional layers of copper oxide planes¹ (hence the name ‘cuprate’) where the super conductivity takes place [16]. The layers are doped to have holes. However, without doping they are well described by the Heisenberg model. The inter-planar coupling is weak; it is nevertheless strong enough to increase the Néel temperature, where magnetic ordering sets in, up to room temperatures [17]. A general review on the isotropic square lattice Heisenberg model, including different numerical and analytical approaches and a discussion how well it describes La_2CuO_4 , is given in Ref. [8]. Of course, one can find a agreement between theory and experiment in other materials (not related to high- T_c superconductors), too [18].

Since exact, analytical solutions are rare (especially in more than one dimension), numerical calculations and simulations of different models are an essential tool for the comparison with experiments and the verification (or falsification) of approximative solutions and predictions, e.g., from conformal field theory. Due to the exponential grow of the Hilbert space with the system size, an exact diagonalization of the Hamiltonian eq. (1.1) is only possible for small system sizes. The basic idea of Monte Carlo methods – namely to randomly sample only some of the states instead of all (in the ideal case the most relevant ones) – is a natural way to overcome this grow. Monte Carlo methods can be used to solve basically every model of classical statistical physics. Unfortunately, quantum Monte Carlo methods suffer from a so-called sign-problem in frustrated geometries [19]. Nevertheless, they clearly belong to the most powerful methods in more than one dimension². Thus, a goal of this work was the implementation of one of this methods from scratch, namely the stochastic series expansion (SSE) [21–23]. We discuss in detail how and why it works in chapter 3. In contrast to other numerical approaches, the SSE gives unbiased results; the errors have a purely statistical nature and can be reduced by longer simulation times. We restrict ourself to two dimension to keep the numerical costs on an appropriate level.

As mentioned above, the Mermin-Wagner theorem rules out a finite temperature phase transition in the 2D Heisenberg model. Thus, we explicitly break the continuous symmetry and consider the XXZ -model, i.e., we introduce (in section 2.3) an anisotropy

¹ More recently, another class of iron-based high- T_c superconductors has been found [15]. However, magnetism in two dimensional layers seems to play an important role in these materials, too.

² In one dimensions, the density matrix normalization group (DMRG) method [20] overcomes the grow of the Hilbert space by a restriction to (matrix product) states with low entanglement entropy, which turned out to be well suited to describe most physical ground states.

parameter Δ for the coupling of spins in z -direction such that a finite temperature transition (belonging to the Ising universality class) to an ground state with Néel order occurs. In above spirit, we use the XXZ -model (mostly at $\Delta = 4$) as prototype to examine phase transitions. Beside the uniform square lattice we consider the 1D-2D-crossover with weakly coupled chains and a geometry of strongly coupled two-site dimers embedded in the square lattice with weaker couplings. In the latter, we find a quantum phase transition from Néel order to a valence bond solid with singlets on the dimers, which we discuss with a trial state.

In the model which we consider, we have a well-defined local order parameter – the staggered magnetization. However, for more exotic phase transitions this is not the case; the most prominent example is the Kosterlitz-Thouless transition in the XY -model [24, 25]. One might wonder whether there are further, ‘hidden’ phase transitions in a specific model, which one simply does not detect in the usually considered observables. This motivated us to change our viewpoint in chapter 5 to the fundamental level of information theory and to consider entropies, i.e., measures of the information content in random distributions [26]. The von Neumann entanglement entropy [27] – or at finite temperature the mutual information – measures *all* correlations between a subregion and the rest of the system and should thus also signal exotic phase transitions. Unfortunately, the von Neumann entropy is not directly accessible in quantum Monte Carlo methods – but the Rényi entropies [28] are accessible with the help of the so-called ‘replica trick’. We reproduce crossings near the critical temperature caused by sub-leading scaling terms from Ref. [1] in the uniform square lattice and find them also in spatially anisotropic geometries with different coupling strength. The conventional methods to locate the phase transition are more precise; nevertheless it is interesting to see how the conventional phase transition can be detected with these methods in order to have a interpretation of the signals for hidden phase transitions. In addition, the scaling of the entanglement entropy is relevant for DMRG calculations [20] and related to other fields as well, e.g., the holographic principle and the black hole entropy [29].

The final part of this thesis (chapter 6) is motivated by a recent experiment of Greif *et al.* [2] with ultracold atoms, where they observed short-range quantum magnetism by projections onto singlet and triplet states on certain bonds³. We wondered whether and how these projectors signal the phase transition to Néel order i.e., whether they obey a singularity at the critical temperature. In the following, we briefly review the most important results regarding quantum magnetism in experiments of ultracold atoms.

1.1 Experiments with Ultracold Quantum Gases

Experiments with ultracold atoms provide a new route to enhance our understanding of quantum many body systems. Therefore, atoms are trapped by magnetic fields or lasers in a vacuum chamber, usually in a harmonic potential. The atoms are cooled – first by lasers and in a final step by evaporative cooling [30] – down to temperatures very close to absolute zero; Hart *et al.* [31] recently reached $T = 31nK$. The interaction between different atoms can be tuned with a magnetic field exploiting Feshbach resonances [32]. Standing waves of lasers can be used to create an optical lattice of various geometries,

³ We denote the coupling between neighboring sites as ‘bonds’.

e.g., simple cubic with three mutually perpendicular lasers. The depth of the lattice can be tuned via the intensity of the lasers, i.e., the tunneling of atoms between neighboring sites is controllable as well. Thus, the effective dimensionality of the system can even be reduced to 1D tubes or 2D planes if the tunneling is strongly suppressed in two or one direction, respectively. Some reviews on this active field of research can be found in Refs. [30, 33, 34].

In contrast to real solids, the system parameters are highly tunable. This allows to engineer theoretical models in experiments and to study them in a very clean environment, a concept which is often called quantum simulation. In that sense, one should rather classify these experiments as a theoretical tool enlarging our understanding of some simplified models. Many of the experiments naturally implement either the Bose or the Fermi Hubbard model [35], depending on the used atoms.

However, the realization of magnetism in these experiments still poses a challenge; it is very hard to cool the atoms down far enough to observe magnetic ordering. A precondition is the realization of a Mott insulator such that each lattice site is occupied by exactly one atom. In 2002, Greiner *et al.* [36] were able to observe the transition from a Bose-Einstein condensate to a Mott insulator. Their pioneering work was a proof of principle that one can consider strongly correlated systems with ultracold atoms and laid the foundations to a series of further experiments. The first realization of a Mott insulator with fermions was reported in 2008 by Jördens *et al.* [37] and shortly after by Schneider *et al.* [38].

In the same year, Trotzky *et al.* [39] directly observed superexchange interactions of two bosons in a double-well potential, i.e., spin dependent interactions stemming from the same mechanism as in a real solid. A magnetic field gradient causes an energy difference between the singlet and triplet state such that an oscillation between these two states can be observed.

As we will see in section 2.1, the kinetic exchange leads to a coupling $J = \frac{t^2}{U}$, where t is the hopping amplitude between neighboring sites and U is the on-site interaction. However, the description of the (low-energy) physics with the Heisenberg model is only valid deep in the Mott insulating regime, i.e., for $t \ll U$, and the maximum value of t is set by the requirement of a sufficiently deep lattice. Thus, the (exchange) coupling J generically is quite small and magnetic correlations can be observed only at very low temperatures. The atoms are isolated very well from the environment, the entropy per particle basically is constant and can hardly be lowered to cool the system. Thus, Greif *et al.* [2] used spatially anisotropic couplings to observe quantum magnetism. More precisely, they considered an array of two-site dimers or one dimensional chains embedded in a simple cubic lattice of weaker couplings (i.e., the lasers are tuned such that there is less tunneling). The entropy then ‘redistributes’ on the weaker bonds: while the accessible temperature is not low enough to reach magnetic long-range order, the correlations along the strong bonds obviously are stronger than along the weak bonds. Thus, Greif *et al.* indeed observed magnetic (short-range) correlations along the strong bonds, namely more singlets than triplets within the dimers and along the chains. As mentioned above, our work partly is motivated by their experiment: we consider similar spatially anisotropic couplings (see section 2.5) and discuss the observables of the experiment, projections onto singlets and triplets. However, we focus on the criticality of the projections and do not aim at a comparison of our results with this experiment; such a comparison can already be found in literature: Sciolla *et al.* [40] compared to DMRG in 1D and found that a com-

petition between charge and spin excitations takes place; Imriška *et al.* [41] performed a dynamical cluster expansion for the three dimensional model and found good agreement with the experiment in a direct comparison.

In a more direct approach, Hart *et al.* [31] have recently enhanced the technique of evaporative cooling by using additional lasers (partly) compensating the harmonic trapping potential, as proposed in [42]. With this method, they reach a temperature of just $1.4 T_c$, where T_c is the Néel temperature at which the phase transition occurs. Moreover, they directly observe antiferromagnetic order (averaged over the whole sample) with spin-sensitive Bragg scattering of light, similar as the neutron scattering used to classify solids in condensed matter. Since the amount of antiferromagnetic order depends sensitively on temperature in the proximity of the phase transition, they demonstrate that the Bragg scattering may even be used as a precise thermometer by comparison with numerical data.

Similar as the antiferromagnetic Heisenberg model describes the low energy physics of the Fermi Hubbard model, a two-species Bose Hubbard model can be mapped to the ferromagnetic Heisenberg model. This mapping has been exploited in the work of Fukuhara *et al.* [43]. They prepared chains in the ground state where all spins point in the same direction. Then they flipped a single spin in the middle of the chain with a radio frequency pulse, i.e., they created an excitation called magnon. With time, single-site, and spin resolved measurements, they observed the dynamics of the magnon, which basically propagates freely either to the left or to the right. In a subsequent work, Fukuhara *et al.* [44] flipped two neighboring spins at once. These two flipped spins may split into two single magnons and propagate as before. However, the two magnons can also form a bound state, where the two flipped spins remain neighboring and propagate with another (lower) velocity through the ‘bath’ of parallel, not flipped spins.

Simon *et al.* [45] considered (spin-less) bosons in one-dimensional chains. They realized a Mott insulator with one atom per site, which they tilted in the direction of the chains with an magnetic field gradient. If the tilt is strong enough, an atom may tunnel to its neighboring site creating a double occupancy and a hole – but only if the neighboring site has not yet tunneled away by itself. Whether an atom has tunneled or not may then be mapped to a pseudo spin. This results in an antiferromagnetic Ising model in a magnetic field (tuned by the tilt). They were able to ramp adiabatically through the quantum phase transition between a paramagnetic state with alignment along the field (in the experiment a Mott insulator for weak tilt) and an antiferromagnetic state (with a double occupancy on each second site for strong tilt). An advantage of the mapping is that the coupling is set by the tunneling t (rather than $\frac{t^2}{U}$ stemming from kinetic exchange) such that lower temperatures are easier accessible and faster dynamics could be observed with a single-site read out using a quantum microscope.

Struck *et al.* [46] exploited a mapping, too. They identified the phase of a bosonic superfluid with a classical spin (i.e., a classical XY model), arranged in a two dimensional triangular lattice. The interaction between the sites is caused by a fast, elliptic shaking of the lattice. This allows an independent tuning of the interactions along the x -axis and in other directions; even the sign of the interactions may be changed. Depending on the signs of the interactions, there may even be a frustration of the classical spins such that can not align completely anti-parallel along all bonds. Thus, a variety of different phases, with first and second order phase transitions, were observed in this experiment. This experiment is a nice demonstration how quantum simulation can be

used to understand simple models which nevertheless give rise to complex phenomena. Particularly interesting would also be the experimental realization of the analog quantum spin model: the geometric frustration gives rise to a sign problem in quantum Monte Carlo methods. Maybe, quantum simulation could soon be an alternative to these powerful numerical methods.

1.2 Outline

In the following chapter 2, we discuss the considered model in more detail. Therefore, we review the mechanism of the exchange interactions and the Mermin-Wagner theorem and discuss the explicit symmetry breaking of the coupling yielding the XXZ -model. In section 2.5, we introduce different coupling strengths for some of the bonds, i.e., spatial anisotropies in addition to the anisotropic coupling of the XXZ -model.

The numerical method of our choice, the stochastic series expansion, is described in chapter 3. As the name suggests, it is based on a series expansion around infinite temperature, which we explain in section 3.2. Subsequently, we discuss how the sampling of the configurations is possible, which is one of the key points for the efficiency of Monte Carlo simulations. Finally, we verify our program code by comparison with data obtained from an exact diagonalization of the Hamiltonian.

In chapter 4, we use the results of our SSE simulations for $\Delta = 4$ on the square lattice as an example for phase transitions in general. We discuss spontaneous symmetry breaking and observe the divergence in the correlation length. We briefly give the basic ideas of renormalization group theory to have the foundations for the finite size scaling discussed subsequently. We review different methods to extract the critical temperature T_c from the finite size data, as well as critical exponents. We consider the 1D-2D-crossover in section 4.4 and coupled dimers in section 4.5. In the latter we find a quantum phase transition between a phase with long-range Néel order (and thus a finite temperature transition as well) and a valence bond crystal with singlet states on the dimers. However, since SSE is a finite temperature method, we do not completely focus our simulations on this quantum phase transition, but rather use a trial state to understand it qualitatively.

Chapter 5 starts with a short introduction to the concepts of the entropy and mutual information. We explain how the replica trick can be used to measure the Rényi entropy of second order and show the signal of the phase transition in the corresponding mutual information. Further, we discuss the ratio trick which can be used to extract the scaling of the entanglement entropy in the ground state.

Finally, we consider the behavior of projection operators onto singlets and triplets in chapter 6. We illustrate a close relation between the energy and the projectors. We conclude and demonstrate numerically that the temperature derivative of the projections diverges in the same manner as the specific heat. We consider also spatial correlations of the projections. In the case of spatially anisotropic (non-equivalent) bonds, our analytic arguments do not give information whether the projections are singular on all bonds. Thus, we study the projections on weak and strong bonds for the 1D-2D-crossover in section 6.3 and for weakly coupled dimers in section 6.4.

A short summary of the most important results in chapter 7 finalizes the thesis.

2 Heisenberg Model

In this chapter, we describe the Heisenberg model and variants of it. To motivate the Heisenberg model we first show that it describes the low energy physics of the Fermi-Hubbard model at half filling. The Mermin-Wagner theorem discussed in section 2.2 proves that the isotropic Heisenberg model does not show long-range order. Therefore, we introduce anisotropy, on the one hand in the coupling itself in section 2.3, and later by varying the coupling strength between some of the sites. This leads to models with finite temperature and quantum phase transitions which we will discuss later in chapter 4 in detail.

2.1 Derivation from Hubbard model

The Hubbard model [35] is one of the most studied Hamiltonians in condensed matter theory. It describes electrons on a lattice. Despite its simplicity it shows a variety of phases including a Mott insulator and BCS superconductivity. It is also studied in many experiments with ultra cold gases [33, 34]. In second quantization, the Hamiltonian is given by

$$\mathcal{H} = -t \sum_{\langle i,j \rangle, \sigma} \left(c_{j,\sigma}^\dagger c_{i,\sigma} + \text{h.c.} \right) + U \sum_i n_{i,\uparrow} n_{i,\downarrow} = \mathcal{H}_t + \mathcal{H}_U. \quad (2.1)$$

Here $\langle i, j \rangle$ denotes nearest neighbors, $n_{i,\sigma} = c_{i,\sigma}^\dagger c_{i,\sigma}$ is the number operator and $c_{i,\sigma}^\dagger, c_{i,\sigma}$ are the fermionic creation and annihilation operators for electrons with spin $\sigma \in \{\uparrow, \downarrow\}$ at site i , satisfying the usual anticommutation relations

$$\{c_{i,\sigma}, c_{j,\sigma'}\} = \delta_{i,j} \delta_{\sigma,\sigma'}. \quad (2.2)$$

The first term of the Hamiltonian describes hopping between neighboring sites with amplitude t . In general, the hopping amplitude t can depend on the sites i, j , but for simplicity we will here assume an isotropic t between neighboring sites and generalize the result at the end. The second term produces an on-site interaction, i.e., electrons on the same site pay an energy of U . Due to the Pauli principle, there can be at most two electrons per site. In the following we will consider the case of half filling, where there are exactly as many electrons as sites¹.

In the atomic limit $t = 0 \Rightarrow \mathcal{H} = \mathcal{H}_U$ at half filling, every site is occupied by exactly one electron in the ground state, and the system has a gap U to the excited states with double occupancies. Due to the spins the ground state has in this case a huge degeneracy of 2^N where N is the number of sites.

¹ At fillings below (on average) one electron per site, an analogous calculation as in this section leads to the $t - J$ model [5] which includes the motion of holes.

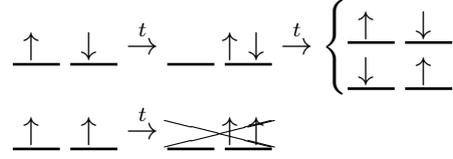


Figure 2.1: Possible hopping of electrons. If electrons on neighboring sites have different spins (upper row), a virtual hopping process is possible, lowering the energy $\propto \frac{t^2}{U}$ and eventually flipping the spins. If the electrons have the same spins (lower row), this process is forbidden due to the Pauli principle.

A nonzero hopping amplitude $t > 0$ leads to an effective spin dependent interaction by means of virtual hopping processes, as depicted in Figure 2.1. If electrons on neighboring sites have the same spin, the Pauli principle forbids hopping, since electrons with the same spin can not occupy the same site. In contrast, if the electrons are on neighboring sites, one electron can hop onto the other site in to a virtual state with higher energy U . Then either the same electron hops back, ending up with the previous state, or the other electron hops back. The second case leads to an effective spin flip of the two electrons. This process is called exchange interaction and leads to antiferromagnetism.

A formal correct calculation can be done by means of a Schrieffer-Wolff transformation [47], which can be seen as an elegant way to do degenerate perturbation theory [48]. The following derivation is along the lines of Ref. [5].

The Schrieffer-Wolff transformation aims at an effective Hamiltonian \mathcal{H}_{eff} in the limit of $t \ll U$ and low energies by treating \mathcal{H}_t as a perturbation. We know the eigenspace of \mathcal{H}_U with lowest energy: all states without any double occupancy. Let \mathcal{P}_0 be that subspace and \hat{P}_0 the projector onto that space. We can split $\mathcal{H}_t = \mathcal{H}_t^{\text{diag}} + \mathcal{H}_t^{\text{offd}}$ into a diagonal part preserving the eigenspaces of \mathcal{H}_U and an off-diagonal part, either increasing (\mathcal{H}_t^+) or decreasing (\mathcal{H}_t^-) the double occupancy:

$$\mathcal{H}_t^{\text{diag}} = -t \sum_{\langle i,j \rangle, \sigma} \left(n_{j,-\sigma} c_{j,\sigma}^\dagger c_{i,\sigma} n_{i,-\sigma} + (1 - n_{j,-\sigma}) c_{j,\sigma}^\dagger c_{i,\sigma} (1 - n_{i,-\sigma}) + (i \leftrightarrow j) \right) \quad (2.3)$$

$$\mathcal{H}_t^{\text{offd}} = -t \sum_{\langle i,j \rangle, \sigma} \left(\underbrace{n_{j,-\sigma} c_{j,\sigma}^\dagger c_{i,\sigma} (1 - n_{i,-\sigma})}_{\mathcal{H}_t^+} + \underbrace{(1 - n_{j,-\sigma}) c_{j,\sigma}^\dagger c_{i,\sigma} n_{i,-\sigma}}_{\mathcal{H}_t^-} + (i \leftrightarrow j) \right) \quad (2.4)$$

Since $\mathcal{H}_t^{\text{offd}}$ is not diagonal in the eigenspaces of \mathcal{H}_U , it changes the eigenstates and we can't directly project $\mathcal{H}_t + \mathcal{H}_U$ onto \mathcal{P}_0 . Instead, we first try to find a rotation e^{iS} with a hermitian generator S such that the rotated Hamiltonian $\tilde{\mathcal{H}}_{\text{eff}} = e^{iS} \mathcal{H} e^{-iS}$ is diagonal in the double occupancy, i.e., doesn't change the double occupancy. Then we can finally define the effective Hamiltonian as the projection onto \mathcal{P}_0 :

$$\mathcal{H}_{\text{eff}} = \hat{P}_0 \tilde{\mathcal{H}}_{\text{eff}} \hat{P}_0 = \hat{P}_0 e^{iS} \mathcal{H} e^{-iS} \hat{P}_0 \quad (2.5)$$

Since the Hamiltonian has a gap, it can be rigorously proven that this projection gives indeed the eigenstates with the lowest energies [48].

Expanding the rotation in powers of S gives

$$\begin{aligned}\tilde{\mathcal{H}}_{\text{eff}} &= \left(\mathbb{1} + iS + \frac{i^2}{2}S^2 \right) \mathcal{H} \left(\mathbb{1} - iS + \frac{i^2}{2}S^2 \right) + \mathcal{O}(S^3) \\ &= \mathcal{H} + i[S, \mathcal{H}] + \frac{i^2}{2}[S, [S, \mathcal{H}]] \\ &= \mathcal{H}_U + \mathcal{H}_t^{\text{diag}} + \mathcal{H}_t^{\text{offd}} + i \left[S, \mathcal{H}_U + \mathcal{H}_t^{\text{diag}} + \mathcal{H}_t^{\text{offd}} \right] - \frac{1}{2} [S, [S, \mathcal{H}]].\end{aligned}\quad (2.6)$$

In order to remove the off-diagonal terms in lowest order of t , we have to require

$$i[S, \mathcal{H}_U] = -\mathcal{H}_t^{\text{offd}} + \mathcal{O}(t^2). \quad (2.7)$$

Identifying the terms in $\mathcal{H}_t^{\text{offd}}$ as annihilation and creation of a double occupancy, it is straightforward to see

$$[\mathcal{H}_t^\pm, \mathcal{H}_U] = \mp U \mathcal{H}_t^\pm. \quad (2.8)$$

Thus, we find the solution of eq. (2.7) up to $\mathcal{O}(t^2)$ as

$$S = \frac{1}{iU} (\mathcal{H}_t^+ - \mathcal{H}_t^-). \quad (2.9)$$

It is thus clear that $S \in \mathcal{O}(t/U)$, as one can already expect from eq. (2.7). Since we want to project on the subspace without double occupancy in the end, the term $[S, \mathcal{H}_t^{\text{diag}}]$ will drop out of eq. (2.6). The remaining terms to second order in t evaluate as

$$\begin{aligned}[S, [S, \mathcal{H}]] &= i[S, \mathcal{H}_t^{\text{offd}}] \\ &= \frac{1}{U} [\mathcal{H}_t^+ - \mathcal{H}_t^-, \mathcal{H}_t^+ + \mathcal{H}_t^-] = \frac{2}{U} [\mathcal{H}_t^+, \mathcal{H}_t^-].\end{aligned}\quad (2.10)$$

Collecting the results eq. (2.5)-(2.10), we arrive at

$$\mathcal{H}_{\text{eff}} = \hat{\mathbb{P}}_0 \mathcal{H}_U \hat{\mathbb{P}}_0 + \hat{\mathbb{P}}_0 \mathcal{H}_t^{\text{diag}} \hat{\mathbb{P}}_0 + \frac{1}{U} \hat{\mathbb{P}}_0 [\mathcal{H}_t^+, \mathcal{H}_t^-] \hat{\mathbb{P}}_0. \quad (2.11)$$

In the case of half filling, we start from a state without empty sites and double occupancies and also have to end in such a state due to the projections. Thus, the second term drops and the first term gives a constant only. In the third term the part of the commutator with \mathcal{H}_t^- on the right drops also since we can not decrease the number of double occupancies if there are none. The remaining term $\mathcal{H}_t^- \mathcal{H}_t^+$ in general involves cross terms with three and four sites. But due to the projections these drop out again since they leave a double occupancy. The effective Hamiltonian thus simplifies to:

$$\mathcal{H}_{\text{eff}} = \text{const} - \frac{1}{U} \hat{\mathbb{P}}_0 \mathcal{H}_t^- \mathcal{H}_t^+ \hat{\mathbb{P}}_0, \quad (2.12)$$

$$\begin{aligned}\hat{\mathbb{P}}_0 \mathcal{H}_t^- \mathcal{H}_t^+ \hat{\mathbb{P}}_0 &= \hat{\mathbb{P}}_0 \sum_{\langle i,j \rangle, \sigma} \left\{ \left((1 - n_{i,-\sigma}) c_{i,\sigma}^\dagger c_{j,\sigma} n_{j,-\sigma} + (1 - n_{i,\sigma}) c_{i,-\sigma}^\dagger c_{j,-\sigma} n_{j,\sigma} \right) \right. \\ &\quad \left. \times n_{j,-\sigma} c_{j,\sigma}^\dagger c_{i,\sigma} (1 - n_{i,-\sigma}) + (i \leftrightarrow j) \right\} \hat{\mathbb{P}}_0 \\ &= t^2 \sum_{\langle i,j \rangle, \sigma} \left((1 - n_{i,-\sigma}) n_{i,\sigma} n_{j,-\sigma} (1 - n_{j,\sigma}) - c_{i,-\sigma}^\dagger c_{j,\sigma}^\dagger c_{j,-\sigma} c_{i,\sigma} + (i \leftrightarrow j) \right).\end{aligned}\quad (2.13)$$

Note that the terms in eq. (2.13) are exactly those depicted in Figure 2.1. From the second to the third line, we used the anti-commutation relations eq. (2.2) and removed redundant number operators. Due to the sum over the spins, switching i and j gives the same contribution, i.e., just a factor of two.

Using the definitions²

$$\begin{aligned} n_i &= n_{i,\uparrow} + n_{i,\downarrow}, & S_i^z &= \frac{1}{2} (n_{i,\uparrow} - n_{i,\downarrow}), \\ S_i^+ &= c_{i,\uparrow}^\dagger c_{i,\downarrow}, & S_i^- &= c_{i,\downarrow}^\dagger c_{i,\uparrow}, \end{aligned} \quad (2.15)$$

we can evaluate the sum over the spins to give

$$\hat{P}_0 \mathcal{H}_t^- \mathcal{H}_t^+ \hat{P}_0 = \text{const} - \frac{2t^2}{U} \hat{P}_0 \sum_{\langle i,j \rangle} \left[\left(\frac{n_i n_j}{2} - 2S_i^z S_j^z \right) - (S_i^- S_j^+ + S_i^+ S_j^-) \right]. \quad (2.16)$$

Dropping the constant we finally arrive at the Heisenberg Hamiltonian:

$$\begin{aligned} \mathcal{H}_{\text{eff}} &= J \sum_{\langle i,j \rangle} \frac{1}{2} (S_i^- S_j^+ + S_i^+ S_j^-) + S_i^z S_j^z \\ &= J \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j \quad \text{where } J = \frac{4t^2}{U}. \end{aligned} \quad (2.17)$$

The Heisenberg model thus describes the low-energy physics of the Hubbard model at half filling. Parallel spins on neighboring sites have a higher energy since $J = \frac{4t^2}{U} > 0$. Thus, the kinetic exchange in the Hubbard model gives rise to antiferromagnetism.

For the ferromagnetic Heisenberg model, $J < 0$, one can immediately write down a ground state: all spins are aligned in the same direction, e.g., $|\uparrow\uparrow\uparrow \dots\rangle$. This is referred to as *spontaneous symmetry breaking*: although the Hamiltonian is invariant under rotation, the system chooses a direction and breaks the $SU(2)$ symmetry of spin rotations, such that the ground state has less symmetry than the Hamiltonian. One might be tempted to assume that in the antiferromagnetic case $J > 0$ – to which the repulsive Hubbard model gives rise – the Néel-state with neighboring spins anti-aligned, $|\uparrow\downarrow\uparrow\downarrow \dots\rangle$, might be the ground state. However, that state is not even an eigenstate of the Hamiltonian: the action of the off-diagonal term $(S_i^- S_j^+ + S_i^+ S_j^-)$ does not vanish. It is thus a natural question to ask whether there exists long-range order in the antiferromagnetic Heisenberg model.

2.2 Mermin-Wagner Theorem

In 1966, Mermin and Wagner [7] gave a rigorous proof that there can not be a finite magnetization in the isotropic Heisenberg model eq. (2.17) in one or two dimensions at

² As one can check straightforwardly with the fermionic anticommutators (2.2), these definitions fulfill the usual commutation relations for spins:

$$[S_i^\pm, S_j^z] = \mp \delta_{i,j} S_i^\pm \quad [S_i^+, S_j^-] = 2\delta_{i,j} S_i^z \quad (2.14)$$

finite temperatures. The proof relies on an inequality from Bogoliubov [49] which they apply to show that the magnetization m (which may here refer to both the uniform and staggered magnetization for a ferromagnet and antiferromagnet, respectively) can be bounded for a sufficiently small external field h :

$$m < \begin{cases} \frac{\text{const}}{\sqrt{T|\ln|h||}} & (2 \text{ dimensions}), \\ \frac{\text{const}}{T^{\frac{2}{3}}} |h|^{\frac{1}{3}} & (1 \text{ dimension}). \end{cases} \quad (2.18)$$

Hohenberg [50] used the same inequality from Bogoliubov to rule out the existence of the order parameters associated with superfluidity and superconductivity³.

The physical reason for the absence of long-range order are Goldstone bosons [51]: Due to the continuous symmetry there exist massless excitations, which slowly vary the direction of the local order parameter. In one and two dimensions, those excitations are strong enough to completely destroy the ordering in the thermodynamic limit at any finite temperature. Thus, one often reads the following more general statement, although it was not proved in this form by Mermin, Wagner, and Hohenberg [52]:

Theorem (Mermin-Wagner-Hohenberg). *There is no spontaneous breaking of continuous symmetries in one and two dimensions at finite temperatures $T > 0$.*

2.3 Anisotropic Coupling

In section 2.1, we obtained the isotropic Heisenberg Hamiltonian eq. (2.17). Since already the Hubbard Hamiltonian has a $SU(2)$ spin-rotation symmetry, also the resulting Heisenberg Hamiltonian is invariant under a rotation of all spins. It is an interesting question what a breaking of that symmetry gives rise to. We introduce a different coupling in one direction, by convention usually the z -direction (thus also called XXZ -model):

$$\begin{aligned} \mathcal{H} &= J \sum_{\langle i,j \rangle} [(S_i^x S_j^x + S_i^y S_j^y) + \Delta S_i^z S_j^z] \\ &= J \sum_{\langle i,j \rangle} \left[\frac{1}{2} (S_i^- S_j^+ + S_i^+ S_j^-) + \Delta S_i^z S_j^z \right]. \end{aligned} \quad (2.19)$$

Clearly, the choice $\Delta = 1$ recovers the previous $SU(2)$ -invariant isotropic Hamiltonian eq. (2.17). As Mermin and Wagner mentioned in their paper, eq. (2.18) is still valid for $\Delta \neq 1$ and $h = (0, 1, 0)$ [7]. However, it just rules out a magnetization in the x - or y -direction, not in the z -direction.

$\Delta > 1$. For $\Delta \rightarrow \infty$, the x - and y -direction become unimportant compared to the z -direction, and the model approaches (after rescaling with $\frac{1}{\Delta}$) the antiferromagnetic Ising model $\mathcal{H} = J \sum_{\langle i,j \rangle} S_i^z S_j^z$. The Ising model only has the *discrete* \mathbb{Z}_2 symmetry corresponding to a sign-flip for all spins. Thus, the Mermin-Wagner-Hohenberg theorem does not rule out a finite temperature phase transition. Indeed, the two dimensional Ising model is even the prototypical example of phase transitions in statistical physics, see e.g., [52, 53]. It turns out that there exists a finite temperature transition at any $\Delta > 1$ with the critical temperature T_c approaching 0 when $\Delta \rightarrow 1$.

³ In fact, Mermin and Wagner were inspired by his paper.

		eigenstate	eigenenergy
singlet	S	$\frac{ \uparrow\downarrow - \downarrow\uparrow\rangle}{\sqrt{2}}$	$E_S = -\frac{J\Delta}{4} - \frac{J}{2}$
triplet	T_0	$\frac{ \uparrow\uparrow + \downarrow\downarrow\rangle}{\sqrt{2}}$	$E_{T_0} = -\frac{J\Delta}{4} + \frac{J}{2}$
	T_+	$ \uparrow\uparrow\rangle$	$E_{T_+} = \frac{J\Delta}{4}$
	T_-	$ \downarrow\downarrow\rangle$	$E_{T_-} = \frac{J\Delta}{4}$

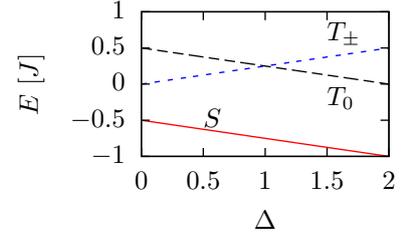


Table 2.1: Eigenstates and -energies of a single bond eq. (2.20). The energies are shown as a function of the anisotropy Δ on the right.

$|\Delta| < 1$. For $\Delta = 0$, one obtains the quantum XY -model, which obeys just the continuous $U(1)$ symmetry of rotations about the z -axis. Thus, the Mermin-Wagner-Hohenberg theorem indeed rules out a finite magnetization. However, the model still possesses a finite temperature transition in two dimensions, a so-called Kosterlitz-Thouless transition [24, 25]. There are ‘vortices’, which are point defects around which the direction of the spins makes a full rotation. Above some critical temperature T_{KT} , these vortices can move freely, whereas they are bound into vortex-antivortex pairs below T_{KT} . The characterization of this transition is quite hard since it has no classical order parameter.

$\Delta = -1$. This case is also $SU(2)$ symmetric and favors ferromagnetic alignment in the z -direction, but antiferromagnetic alignment in the x - and y -directions.

$\Delta < -1$. Similar to the case $\Delta > 1$, the Hamiltonian approaches the Ising model for large Δ , but in this case the ferromagnetic one. Again, there is a finite temperature transition for any $\Delta < -1$, from an disordered phase into ferromagnetic ordering.

The anisotropy $\Delta \neq 1$ thus serves as a way to introduce a finite temperature transition. Therefore, we will mainly study the case $\Delta = 4$ in chapter 4 about phase transitions.

2.4 Single Heisenberg Dimer

The simplest case of an XXZ -model one can study is a single dimer connecting just two sites

$$\mathcal{H} = \frac{J}{2} (S_1^- S_2^+ + S_1^+ S_2^-) + J\Delta S_1^z S_2^z \doteq J \begin{pmatrix} \frac{\Delta}{4} & & & \\ & -\frac{\Delta}{4} & \frac{1}{2} & \\ & \frac{1}{2} & -\frac{\Delta}{4} & \\ & & & \frac{\Delta}{4} \end{pmatrix}. \quad (2.20)$$

The matrix is written in the usual basis $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$ and easily diagonalized. The corresponding eigenstates and energies are collected in Table 2.1. It is worth noting that the eigenstates are independent of Δ . For any $\Delta \geq -1$ the ground state is the singlet state which is maximally entangled. The triplet states have the same energy $\frac{J}{4}$ right at the $SU(2)$ symmetric point $\Delta = 1$. For $\Delta > 1$ the T_0 triplet state is favored energetically over the other triplet states T_+ and T_- since the spins are anti-aligned in T_0 . On the other

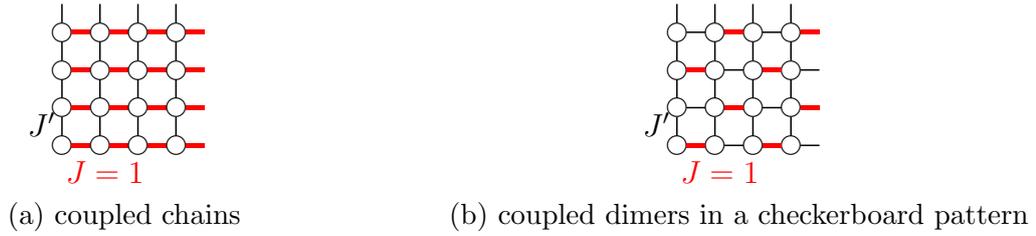


Figure 2.2: The spatially anisotropic couplings used in this thesis, based on a simple square lattice. We choose $J' \leq J$ and measure energies in units of J . *Note:* the experiment of Greif *et al.* [2] uses the same structure, but in a three dimensional cubic lattice with coupling J' between the depicted layers.

hand, the contribution of the kinetic term, which vanishes on T_+ and T_- , dominates for $\Delta < 1$; T_0 is then the state with the highest energy.

2.5 Spatial Anisotropies

As mentioned before (p. 13), it is possible that the hopping in the Hubbard model is stronger between some sites, $t = t_{i,j}$. The calculations and arguments of section 2.1 do not change, so the generalization of the resulting Hamiltonian eq. (2.17) is straightforward:

$$\mathcal{H}_{\text{eff}} = \sum_{\langle i,j \rangle} J_{i,j} \vec{S}_i \cdot \vec{S}_j \quad \text{where } J_{i,j} = \frac{4t_{i,j}^2}{U}. \quad (2.21)$$

In experiments with ultra cold atoms, the hopping can be tuned by the intensity of the lasers creating the optical lattice [34].

In this thesis, we study two possible setups of spatial anisotropies, depicted in Figure 2.2. The first setup, shown in panel (a), consists of one dimensional chains, which are connected to each other by weaker bonds with coupling strength J' . This configuration allows to study continuously the crossover between a one dimensional chain for $J' = 0$, and the simple two dimensional square lattice for $J' = J$.

Beside the chains, we study the case of weakly coupled dimers, depicted in Figure 2.2(b). In that pattern, every site belongs to exactly one strong bond J . As we have seen in section 2.4, the ground state of a single bond is a singlet. Thus, it is clear that in the case $J' = 0$ the ground state is just a product of singlets on the strong bonds. A nonzero value of J' produces a competition: a site can not be in a singlet state simultaneously with all its neighbors. Instead, the total state can be seen as a superposition of product states of singlets⁴. For J' below some critical value J_c the ground state is a so-called valence bond

⁴ Indeed one can expand any state with total spin 0 in the basis of so-called valence bond states [54, 55]

$$|(i_1, j_1)(i_2, j_2) \dots (i_{N/2}, j_{N/2})\rangle \quad \text{where } (i, j) = (\uparrow_i \downarrow_j - \downarrow_i \uparrow_j) / \sqrt{2}. \quad (2.22)$$

Here, each site needs to belong to exactly one singlet. These states form an overcomplete basis of the states with total spin zero. Further, they are nonorthogonal (indeed the overlap between any such states is nonzero), which comes in handy for an importance sampling scheme of the ground state, see Ref. [55] for details.

solid or valence bond crystal [9, 11, 52]. This state has no magnetic long-range order. On the other hand, for $J' = 1$ the ground state has long-range Néel order with some quantum fluctuations and spin wave excitations. Thus, one expects a so-called quantum phase transition at the critical coupling J_c . In contrast to a conventional phase transition, it is driven by a parameter in the Hamiltonian – in this case J' – and describes changes of the ground state, not a thermodynamic ensemble.

It is worth mentioning here that there has been some confusion about the universality class of the quantum Heisenberg model with the geometry of staggered coupled dimers as depicted in Figure 2.2(b). Wenzel *et al.* [56] estimated in 2008 for the exponent ν (for a definition see section 4.1) the value $\nu = 0.689(5)$ which was obtained from the scaling of the correlation length, Binder cumulant, and spin stiffness by SSE simulations. In contrast to expectations, this value of ν seemed to be inconsistent with the result $\nu = 0.7112(5)$ for the classical $O(3)$ Heisenberg model in 3D [57]. However, further, careful analysis showed that both models indeed belong to the same universality class [58, 59], but the former has anomalously large finite size corrections. Fritz *et al.* [58] explain these corrections by an additional term in the low energy quantum field theory which describes a two-particle decay of quantum critical fluctuations.

Finally, we can combine the spatial anisotropies with the anisotropic coupling of section 2.3 and obtain the most general form that will be studied in this thesis:

$$\mathcal{H} = \sum_{\langle i,j \rangle} J_{i,j} \left(\frac{1}{2} (S_i^- S_j^+ + S_i^+ S_j^-) + \Delta S_i^z S_j^z \right). \quad (2.23)$$

3 Method: Stochastic Series Expansion

The Stochastic Series Expansion (SSE) belongs to the class of quantum Monte Carlo (QMC) methods. The basic idea of Monte Carlo methods in general is to use a random process to determine a deterministic, non-random result. In the case of the SSE, one expands the Boltzmann weight in a Taylor series and samples stochastically over the resulting operator strings; details will be given in section 3.2. The basic idea of the expansion was given by Handscomb in 1962 [21]. However, he calculated the appearing trace analytically and was thus restricted to models where this was possible. This problem was avoided by a generalization of Sandvik and Kurkijärvi in 1991 [22] and pushed forward by Sandvik giving also a generalization to the Hubbard model in one dimension [23]. A more detailed introduction and review of the SSE with deterministic loop updates is given in Ref. [52]. For anisotropic couplings such as the ones given by eq. (2.19) and (2.23), a further generalization is necessary: directed loop updates. This generalization was given in 2002 by Syljuåsen and Sandvik [60] and is described in detail by Syljuåsen in Ref. [61].

3.1 Classical Monte Carlo

Although the Monte Carlo method is more general, we restrict ourselves here to the following, still very general, setup that we aim at the calculation of statistical averages for some observable A

$$\langle A \rangle = \sum_{\nu \in \mathcal{C}} w_{\nu} A_{\nu}. \quad (3.1)$$

Here we are given some configuration space \mathcal{C} with possible configurations $\nu \in \mathcal{C}$. For each configuration we can assign a value A_{ν} to the observable A . Further, each configuration has some weight¹ $w_{\nu} > 0$ which can be assumed to be normalized. Indeed, assuming that \tilde{w}_{ν} is not normalized, we just define

$$w_{\nu} := \frac{\tilde{w}_{\nu}}{\sum_{\nu \in \mathcal{C}} \tilde{w}_{\nu}}, \quad \text{then } \sum_{\nu \in \mathcal{C}} w_{\nu} = 1. \quad (3.2)$$

However, as we will see later, it may not be necessary to know the exact weight w_{ν} but only the relative weight $\tilde{w}_{\nu}/\tilde{w}_{\mu}$ between any two states $\mu, \nu \in \mathcal{C}$.

To fix ideas it might be helpful to think of the well known classical Ising model with Hamiltonian $\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i^z S_j^z$. In that case the configuration space is just the phase space consisting of all 2^N possible configurations of spins

$$\mathcal{C} = \left\{ \nu \equiv (\sigma_1, \sigma_2, \dots) \text{ with } \sigma_i \in \left\{ \uparrow \equiv +\frac{1}{2}, \downarrow \equiv -\frac{1}{2} \right\}, i = 1, \dots, N \right\}$$

¹ For frustrated spin systems and fermions one needs to evaluate an average in the form of (3.1), but with (partly) negative weights w_{ν} . This leads to exponential large errors in Monte Carlo methods and is known as the *negative sign problem* of Quantum Monte Carlo [19].

and the weight is given by the Boltzmann factor

$$w_\nu = \exp(-\beta E_\nu) / \mathcal{Z} \quad \text{with } E_{(\sigma_1, \sigma_2, \dots)} = -J \sum_{\langle i, j \rangle} \sigma_i \sigma_j. \quad (3.3)$$

Here $\beta = \frac{1}{T}$ is the inverse temperature² and the normalization is ensured by the partition function $\mathcal{Z} = \sum_{\nu \in \mathcal{C}} \exp(-\beta E_\nu)$. Examples for the observable A could be the energy E itself or the magnetization $M_{(\sigma_1, \sigma_2, \dots)} = \sum_i \sigma_i$.

Note that the value $\langle A \rangle$ for a given observable A is uniquely defined by (3.1) without any random error. However, often the number of possible configurations $|\mathcal{C}|$ is quite large – exponentially in the system size N even for the Ising model –, such that one can not carry out the complete summation in eq. (3.1). The basic idea of Monte Carlo methods is thus a restriction of the sum to some representative configurations μ_t , $t = 1, \dots, \mathcal{T}$, which are picked randomly according to some rules we will specify in the following. Usually the number of representatives \mathcal{T} is much smaller than the number of configurations $\mathcal{T} \ll |\mathcal{C}|$, but still very large; in this work we used $\mathcal{T} \approx 10^6 \dots 10^7$.

In 1953, Metropolis *et al.* suggested in their famous paper Ref. [62] to pick the configurations not uniformly in \mathcal{C} (as it was done before), but to choose them according to their weights w_ν :

$$w_\nu(t) := \text{Probability}(\mu_t = \nu) \stackrel{!}{=} w_\nu \quad \forall t \in \{1, \dots, \mathcal{T}\}. \quad (3.4)$$

The term Probability(...) here refers to a hypothetical average obtained when carrying out the whole Monte Carlo simulation multiple times, i.e., producing many such representative sets μ_t . But if this equation is fulfilled, since the right hand side is independent of t , an average over t gives the same result: say $w_\sigma = 0.25$ for some σ , then approximately a quarter of the μ_t , $t = 1, \dots, \mathcal{T}$, should be the configuration σ . One can write this as $w_\nu \approx \frac{1}{\mathcal{T}} \sum_{t=1}^{\mathcal{T}} \delta_{\mu_t, \nu}$, which leads to the following approximation of eq. (3.1):

$$\langle A \rangle = \sum_{\nu \in \mathcal{C}} w_\nu A_\nu \approx \frac{1}{\mathcal{T}} \sum_{t=1}^{\mathcal{T}} A_{\mu_t} =: \bar{A}. \quad (3.5)$$

A distribution according to eq. (3.4) can be generated with a Markov chain. Therefore one picks an initial μ_1 and chooses μ_{t+1} from μ_t according to the transition matrix³ (which needs to be determined)

$$P(\nu \rightarrow \sigma) := \text{Probability}(\mu_{t+1} = \sigma | \mu_t = \nu), \quad \text{such that} \quad (3.6)$$

$$w_\nu(t+1) = \sum_{\sigma \in \mathcal{C}} w_\sigma(t) P(\sigma \rightarrow \nu). \quad (3.7)$$

Such an update $\mu_t \rightarrow \mu_{t+1}$ is called Monte Carlo sweep. Clearly, $P(\nu \rightarrow \sigma)$ is a probability and thus has to fulfill the normalization

$$\sum_{\sigma \in \mathcal{C}} P(\nu \rightarrow \sigma) = 1. \quad (3.8)$$

² We use natural units and set $\hbar \equiv k_B \equiv 1$. Further, we measure all spatial lengths in units of the lattice spacing $a \equiv 1$.

³ One can view $P_{\sigma, \nu} := P(\nu \rightarrow \sigma)$ as a matrix. Equation (3.7) then combines with (3.4) to an eigenvector equation for w_ν corresponding to the eigenvalue 1. One can further show that other eigenvectors have an eigenvalue of magnitude less than one.

Equation (3.4) implies that $w_\nu(t)$ in fact is independent of t . Combining with eq. (3.7) and (3.8) we arrive at

$$\begin{aligned} 0 &\stackrel{!}{=} w_\nu(t+1) - w_\nu(t) \\ &= \sum_{\sigma \in \mathcal{C}} w_\sigma(t) P(\sigma \rightarrow \nu) - w_\nu(t) \sum_{\sigma \in \mathcal{C}} P(\sigma \rightarrow \nu) \\ &= \sum_{\sigma \in \mathcal{C}} [w_\sigma(t) P(\sigma \rightarrow \nu) - w_\nu(t) P(\sigma \rightarrow \nu)]. \end{aligned} \quad (3.9)$$

Although there might be other possible choices, the simplest solution to eq. (3.9) is to set the part in square brackets to zero individually for each σ :

$$\frac{P(\sigma \rightarrow \nu)}{P(\nu \rightarrow \sigma)} \stackrel{!}{=} \frac{w_\nu}{w_\sigma}. \quad (\text{Detailed Balance}) \quad (3.10)$$

This is one out of two conditions one has to ensure when developing a Monte Carlo algorithm. The second condition is ergodicity: it should be possible to reach any configuration $\sigma \in \mathcal{C}$ from any other $\nu \in \mathcal{C}$ in a finite number of sweeps with nonzero probability:

$$\forall \nu, \sigma \in \mathcal{C} \exists K \in \mathbb{N}, \mu_1, \mu_2, \dots, \mu_K \in \mathcal{C} \quad \text{such that} \\ P(\nu \rightarrow \mu_1) \cdot P(\mu_1 \rightarrow \mu_2) \cdots P(\mu_K \rightarrow \sigma) \neq 0. \quad (\text{Ergodicity}) \quad (3.11)$$

Thermalization

We derived the detailed balance condition from the requirement of invariance under a Monte Carlo sweep; thus, it ensures that $w_\nu(t) = w_\nu$ is an equilibrium distribution. However, we can not sample this distribution in the first step – that is the reason why we use a Markov chain in the first place. Instead, one starts from some configuration $\mu_0 \in \mathcal{C}$ with some other distribution $w_\nu(t=0) \neq w_\nu$, often just picked uniformly in \mathcal{C} . To circumvent this problem one first has to do a large number of Monte Carlo sweeps to reach the equilibrium distribution. This process is known as thermalization. Not till then one measures the observables A_{μ_t} and does further sweeps increasing t . Finally, one can approximate the observable averages according to eq. (3.5). Clearly, it is essential to estimate the error done in this approximation as well. This can be done after the simulation and is described in the following.

Error estimates

According to the central limit theorem, the average $\bar{A} = \frac{1}{\mathcal{T}} \sum_{t=1}^{\mathcal{T}} A_t$ itself is normally distributed for large \mathcal{T} if the A_t are independent and identically distributed. In a simple approach, one groups the measurements into B bins of length $M = \mathcal{T}/B$ and calculates the averages within these bins

$$\bar{A} = \frac{1}{B} \sum_{i=1}^B \bar{A}_i \quad \text{with } \bar{A}_i := \frac{1}{M} \sum_{t=(i-1)M+1}^{iM} A_t \quad \forall i = 1, \dots, B. \quad (3.12)$$

Then one can estimate the error from the variance of these bins

$$\sigma^2 = \frac{1}{B-1} \sum_{i=1}^B (\bar{A}_i - \bar{A})^2 \quad (3.13)$$

and quote the result as $\langle A \rangle = \bar{A} \pm \sigma/\sqrt{B}$ [52].

However, the $A_t \equiv A_{\mu_t}$ are not really independent random variables since, by definition, in a Markov chain the distribution of μ_{t+1} depends on the previous μ_t . This correlation is quantified by the integrated autocorrelation time [63, 64]

$$\tau_A^{int} = \frac{1}{2} + \sum_{t=1}^{\infty} \tau_A(t), \quad \text{where } \tau_A(t) = \frac{\sum_{t'} (A_{t'+t} - \bar{A})(A_{t'} - \bar{A})}{\sum_{t'} (A_{t'} - \bar{A})^2}. \quad (3.14)$$

Only for the bin length $M \gg \tau_A^{int}$ one can regard the bin averages as statistically independent such that the error bars are well defined from the central limit theorem. Due to correlations one effectively has only $\mathcal{T}/(2\tau^{int})$ independent measurements. Thus, one can expect the error to be $\sigma/\sqrt{B} \propto \sqrt{2\tau^{int}/\mathcal{T}}$ [52].

Jackknife

A slightly more sophisticated error estimate is obtained from the jackknife resampling [63, 64]. The advantage of this method is a correct error propagation also for nonlinear functions, e.g., for ratios of two observables or even fits. In this method, too, the measurements are first grouped into bins. But now one performs averages excluding one bin

$$\tilde{A}_i = \frac{1}{B-1} \sum_{\substack{j=1 \\ j \neq i}}^B \bar{A}_j. \quad (3.15)$$

Suppose we have a function $f(A, \dots)$ depending on one or more observable from the Monte Carlo simulation. With the definitions

$$f^{(0)} := f(\bar{A}, \dots), \quad f^{(i)} := f(\tilde{A}_i, \dots) \quad \forall i = 1, \dots, M, \quad f^{avg} := \frac{1}{B} \sum_{i=1}^B f^{(i)}, \quad (3.16)$$

we have two estimates using all measurements, namely $f^{(0)}$ and f^{avg} . For linear $f(A, \dots)$ these values coincide, but for nonlinear functions they may differ. We can then correct a possible bias of the estimate with the following formula⁴:

$$f^{bias-corrected} = f^{(0)} - (B-1)(f^{avg} - f^{(0)}). \quad (3.17)$$

⁴ Suppose we have an error $f^{(0)} = f + \frac{C}{B} + \mathcal{O}(B^{-2})$ for the correct value f and some constant C . Then $f^{avg} = f + \frac{C}{B-1} + \mathcal{O}(B^{-2})$ which in turn implies that $f^{(0)} - (B-1)(f^{avg} - f^{(0)}) = f + \mathcal{O}(B^{-2})$.

1. Initialization: random initial ν .
2. Thermalization: reach equilibrium with a large number of sweeps.
3. Measurement Loop:
 - Monte Carlo sweep: update $\nu_t \rightarrow \nu_{t+1}$ according to $P(\nu_t \rightarrow \nu_{t+1})$.
 - Measurement of Observables A_{ν_t} .
4. Calculation of average values and error estimates from bin averages.

Figure 3.1: Outer algorithmic structure for Monte Carlo simulations in order to evaluate eq. (3.1).

Since the \tilde{A}_i average over $\mathcal{T} - M$ compared to just M measurements with the simple binning, they are less noisy:

$$\begin{aligned} \tilde{A}_i - \bar{A} &= \frac{1}{B-1} \left(\sum_j \bar{A}_j - \bar{A} \right) - \frac{1}{B} \sum_j \bar{A}_j \\ &= \frac{1}{B(B-1)} \left(\sum_j \bar{A}_j - BA_i \right) = \frac{1}{B-1} (\bar{A} - \bar{A}_j). \end{aligned} \quad (3.18)$$

Compared to eq. (3.13) one thus has to include an additional factor $(B-1)^2$ into the variance:

$$\sigma_{JK}^2 = \frac{B-1}{B} \sum_{i=1}^B (f^{(i)} - f^{avg})^2. \quad (3.19)$$

As for simple binning, the final result can be quoted as $f^{bias-corrected} \pm \sigma_{JK}/\sqrt{B}$.

We have now developed the outer framework of Monte Carlo simulations, which is suitable to evaluate the weighted sum in eq. (3.1) for a huge configuration space. The necessary steps are summarized in Figure 3.1. At this point, the framework is very general and finds a variety of applications which are not restricted to physics, e.g., it can be used for any high-dimensional integrations.

Since we worked on a general configuration space, we have not exactly specified a Monte Carlo sweep yet, but only derived the detailed balance and ergodicity conditions for this update. We will now discuss the simple Metropolis-Hastings algorithm as the basic ideas of this algorithm will also be used in the SSE.

Metropolis-Hastings Algorithm

So how exactly can one choose the update probabilities in the simple case of the Ising model? Suppose we are in a configuration $\vec{\sigma} := (\sigma_1, \dots, \sigma_N)$. The Metropolis-Hastings algorithm [62, 65] divides one sweep into two steps: First select a random spin $i \in$

$\{1, \dots, N\}$ with uniform probability $P^{\text{sel.}} = \frac{1}{N}$; then accept a flip of that spin to the new configuration $\vec{\sigma}_{\bar{i}} := (\sigma_1, \dots, -\sigma_i, \dots, \sigma_N)$ with probability

$$P^{\text{acc.}}(\vec{\sigma}, i) := \min \{1, e^{-\beta \Delta E(\vec{\sigma}, i)}\} = \begin{cases} 1 & \text{if } \Delta E(\vec{\sigma}, i) < 0, \\ \exp(-\beta \Delta E(\vec{\sigma}, i)) & \text{else.} \end{cases} \quad (3.20)$$

Here $\Delta E = E_{\vec{\sigma}_{\bar{i}}} - E_{\vec{\sigma}}$ is the energy difference obtained by the flip. The total update probability is then given by $P(\vec{\sigma} \rightarrow \vec{\sigma}_{\bar{i}}) = P^{\text{sel.}} \cdot P^{\text{acc.}}(i, \vec{\sigma})$ and $P(\vec{\sigma} \rightarrow \nu) = 0$ if $\vec{\sigma}$ and ν differ by more than one spin. It is highly remarkable that one does not need to know the partition function, i.e., the normalization for the actual weight. Instead, one only needs to know the relative factor of the weight between every two states.

Since $\Delta E(\vec{\sigma}, i) = -\Delta E(\vec{\sigma}_{\bar{i}}, i)$, the Metropolis-Hastings-algorithm clearly satisfies detailed balance:

$$\begin{aligned} \frac{P(\vec{\sigma} \rightarrow \vec{\sigma}_{\bar{i}})}{P(\vec{\sigma}_{\bar{i}} \rightarrow \vec{\sigma})} &= \frac{P^{\text{sel.}}}{P^{\text{sel.}}} \cdot \begin{cases} \frac{1}{\frac{\exp(-\beta \Delta E(\vec{\sigma}_{\bar{i}}, i))}{\exp(-\beta \Delta E(\vec{\sigma}, i))}} & \text{for } \Delta E(\vec{\sigma}, i) < 0, \\ \frac{\exp(-\beta \Delta E(\vec{\sigma}, i))}{1} & \text{else,} \end{cases} \\ &= \frac{\exp(-\beta E(\vec{\sigma}_{\bar{i}}))}{\exp(-\beta E(\vec{\sigma}))} = \frac{w_{\vec{\sigma}_{\bar{i}}}}{w_{\vec{\sigma}}}. \quad \checkmark \end{aligned} \quad (3.21)$$

Also, ergodicity is fulfilled at least theoretically: One can reach any other state by N spin flips at most, each of which has a nonzero probability.

However, there is a problem with the algorithm described above. At a phase transition clusters of correlated spins form. Right at the critical temperature the correlation length diverges. With local updates it is very improbable to flip such a cluster: Flipping a spin inside the cluster costs much energy. Instead, the algorithm tends to just move the domain walls, which takes much longer. This manifests itself in long autocorrelation times and thus large errors. To overcome that problem one has to use so-called cluster algorithms, which flip a complete cluster of spins at once [53, 66]. Describing these algorithms goes beyond the scope of this thesis. However, it should be noted that the loop updates of the SSE described in section 3.3.2 and 3.3.3 are nonlocal and thus avoid a critical slowing down.

3.2 Stochastic Series Expansion: Mapping to Classical Configurations

When one tries to use Monte Carlo for quantum mechanical problems, one faces the problem that the eigenstates and eigenenergies of the system are unknown. Instead, we work in the local S^z basis for each site⁵. Thus, one needs to find a way to express the expectation values of observables

$$\langle \hat{A} \rangle = \text{Tr} \left\{ \frac{\exp(-\beta \mathcal{H})}{\mathcal{Z}} \hat{A} \right\} \quad (3.22)$$

in the form of eq. (3.1). To achieve that we first rewrite the Hamiltonian in a more general form. We will here demonstrate the mapping for the anisotropic spin- $\frac{1}{2}$ Heisenberg model

⁵For ground state properties one may also work in the valence bond basis [55], cf. the footnote on p. 19.

$\mathcal{H}_b^{\text{offd}} \uparrow\uparrow\rangle = 0$	$\mathcal{H}_b^{\text{diag}} \uparrow\uparrow\rangle = (C_0 - \frac{\Delta}{4}) \uparrow\uparrow\rangle$	
$\mathcal{H}_b^{\text{offd}} \downarrow\downarrow\rangle = 0$	$\mathcal{H}_b^{\text{diag}} \downarrow\downarrow\rangle = (C_0 - \frac{\Delta}{4}) \downarrow\downarrow\rangle$	
$\mathcal{H}_b^{\text{offd}} \uparrow\downarrow\rangle = \frac{1}{2} \downarrow\uparrow\rangle$	$\mathcal{H}_b^{\text{diag}} \uparrow\downarrow\rangle = (C_0 + \frac{\Delta}{4}) \uparrow\downarrow\rangle$	
$\mathcal{H}_b^{\text{offd}} \downarrow\uparrow\rangle = \frac{1}{2} \uparrow\downarrow\rangle$	$\mathcal{H}_b^{\text{diag}} \downarrow\uparrow\rangle = (C_0 + \frac{\Delta}{4}) \downarrow\uparrow\rangle$	

Table 3.1: Action and pictorial representation of diagonal and off-diagonal operators in eq. (3.23) for $J_b = 1$. The first spin refers to site $i(b)$, the second to $j(b)$. Spin up \uparrow (spin down \downarrow) is depicted with solid circles \bullet (open circles \circ).

(2.23), but a similar mapping is also possible for larger spin $S = 1, \frac{3}{2}, \dots$ or bosons [9, 61]. First of all, we can relabel the sum over next neighbors $\langle i, j \rangle$ as a sum over bonds $b = 0, \dots, N_b - 1$, where $i(b)$ and $j(b)$ are the two sites involved in bond b . This formulation has the advantage that it is independent of the detailed geometry used.

$$\begin{aligned}
\mathcal{H} &= \sum_b J_b \left[\frac{1}{2} \left(S_{i(b)}^- S_{j(b)}^+ + S_{i(b)}^+ S_{j(b)}^- \right) + \Delta S_{i(b)}^z S_{j(b)}^z \right] \\
&= \sum_b \left[\underbrace{\frac{J_b}{2} \left(S_{i(b)}^- S_{j(b)}^+ + S_{i(b)}^+ S_{j(b)}^- \right)}_{=:\mathcal{H}_b^{\text{offd}}} + \underbrace{J_b \left(\Delta S_{i(b)}^z S_{j(b)}^z - C_0 \right)}_{=:-\mathcal{H}_b^{\text{diag}}} + J_b C_0 \right] \\
&= - \sum_b \left[-\mathcal{H}_b^{\text{offd}} + \mathcal{H}_b^{\text{diag}} \right] + \sum_b J_b C_0
\end{aligned} \tag{3.23}$$

Table 3.1 lists the action of $\mathcal{H}^{\text{offd}}$ and $\mathcal{H}^{\text{diag}}$ on all possible combinations of spins. We have split the Hamiltonian in so many parts, that each part acting on a basis state produces at most one (maybe different) basis state, not a superposition. In the second line (3.23) we have further included a constant C_0 into the diagonal part $\mathcal{H}_b^{\text{diag}}$. This constant has to be chosen such that $\mathcal{H}^{\text{diag}}$ is positive. As shown in Table 3.1 this is the case for $C_0 \geq \frac{\Delta}{4}$. We will see later that this is necessary to ensure positive weights w_ν in eq. (3.1).

The starting point for the stochastic series expansion is the partition function:

$$\begin{aligned}
\mathcal{Z} &= \text{Tr} \{ \exp(-\beta \mathcal{H}) \} \\
&= \text{Tr} \left\{ \exp \left(\beta \sum_b \left(\mathcal{H}_b^{\text{diag}} - \mathcal{H}_b^{\text{offd}} \right) \right) \right\} e^{-\beta \sum_b J_b C_0}.
\end{aligned} \tag{3.24}$$

For all expectation values except the energy, an overall constant prefactor does not matter. Thus, we drop it in the following and include it only when we actually calculate the energy. The trace can be evaluated in the usual basis of local spins, which we label with $\{|\alpha\rangle\}$.

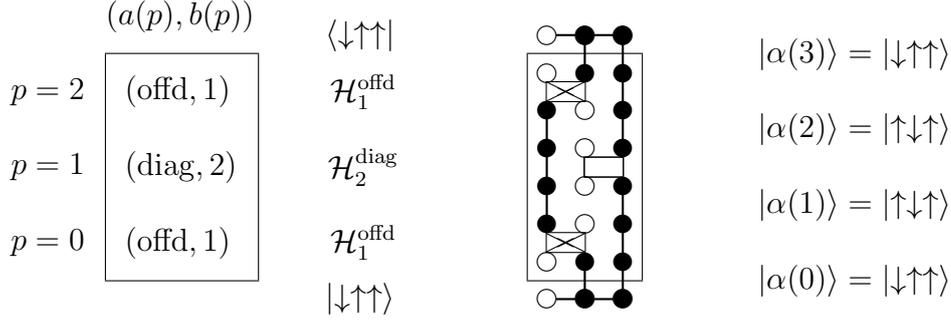


Figure 3.2: Simple example of an operator string of length 3. In the middle a pictorial representation is shown. Corresponding propagated states are shown on the right. For the sake of a better overview in the following figures the propagated states will not be shown but only the lines indicate the up-spins. The vertical direction will sometimes be referred to as imaginary time direction due to the correspondence with path integral Monte Carlo.

Further, we can expand the exponential in its series representation. We arrive at:

$$\begin{aligned}
\mathcal{Z} &\propto \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{|\alpha\rangle} \langle \alpha | \left[\sum_b \left(\mathcal{H}_b^{\text{diag}} - \mathcal{H}_b^{\text{offd}} \right) \right]^n | \alpha \rangle \\
&= \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{|\alpha\rangle} \langle \alpha | \prod_{p=0}^{n-1} \left[\sum_b \left(\mathcal{H}_b^{\text{diag}} - \mathcal{H}_b^{\text{offd}} \right) \right] | \alpha \rangle.
\end{aligned} \tag{3.25}$$

In the last line, we introduced the index $p \in \{0, \dots, n-1\}$ which directly corresponds to the imaginary time in path integral Monte Carlo [52]. Each term $-\mathcal{H}_b^{\text{offd}}$ and $\mathcal{H}_b^{\text{diag}}$ appears in the product exactly once at each p . We can collect the bond indices $b(p) \in \{0, \dots, N_b\}$ and indices for the operator type $a(p) \in \{\text{offd}, \text{diag}\}$ in an ordered list of length n , a so-called operator string S_n . Then we can write the product compactly as a sum over all possible operator strings of given length n :

$$\prod_{p=0}^{n-1} \left[\sum_b \left(\mathcal{H}_b^{\text{diag}} - \mathcal{H}_b^{\text{offd}} \right) \right] = \sum_{S_n} (-1)^{n^{\text{offd}}} \prod_{p=0}^{n-1} \mathcal{H}_{b(p)}^{a(p)}. \tag{3.26}$$

Here n^{offd} is the number of off-diagonal operators in the string. Due to the splitting of the Hamiltonian for a given operator string, there are unique states $|\alpha(p)\rangle$ after the propagation of the first p operators, which are referred to as propagated states. Figure 3.2 shows an example of an operator string and corresponding propagated states, which give a nonzero contribution. Clearly, we can exclude all operator strings where the action of on operator $\mathcal{H}_b^{a(p)}(b) |\alpha(p)\rangle$ vanishes, e.g., an off-diagonal operator can't act on parallel spins. Further, the trace structure requires a periodicity in p for an operator string of length n :

$$|\alpha(p = n)\rangle = |\alpha(p = 0)\rangle. \tag{3.27}$$

This periodicity helps us to get rid of the minus signs from the off-diagonal operators: We restrict ourselves to bipartite lattices, where we can assign each site to a sublattice, either

A or B , in such a way that each bond connects only sites between the sublattices A and B , but not within them. An off-diagonal operator "moves" an up-spin from sublattice A to B or back. It requires thus an even number of off-diagonal operators to end up with the same number of up-spins in sublattice A . This implies

$$(-1)^{n^{\text{offd}}} = 1 \quad \text{for bipartite lattices.} \quad (3.28)$$

Examples of bipartite lattices include simple cubic lattices in arbitrary dimension or the honeycomb lattice in two dimensions. When using periodic boundary conditions one is however restricted to even length.

Combining eq. (3.25), (3.26), and (3.28) we arrive at

$$\mathcal{Z} = \sum_{n=0}^{\infty} \sum_{|\alpha\rangle} \sum_{S_n} \frac{\beta^n}{n!} \langle \alpha | \prod_{p=0}^{n-1} \mathcal{H}_{b(p)}^{a(p)} | \alpha \rangle. \quad (3.29)$$

For any diagonal operator \hat{A} we can do an analogous calculation and arrive at an expression in the form of eq. (3.1):

$$\langle \hat{A} \rangle = \sum_{n=0}^{\infty} \sum_{|\alpha\rangle} \sum_{S_n} \underbrace{\frac{\beta^n}{n!} \langle \alpha | \prod_{p=0}^{n-1} \mathcal{H}_{b(p)}^{a(p)} | \alpha \rangle}_{=: w_{|\alpha\rangle, S_n}} \langle \alpha | \hat{A} | \alpha \rangle. \quad (3.30)$$

The configuration space \mathcal{C} consists here of all initial states $|\alpha\rangle$ and operators strings S_n of arbitrary length $n \in \mathbb{N}_0$.

A careful reader might worry here that the configuration space is infinite also for finite lattices due to the sum over n . But for finite temperatures $\beta < \infty$ the factorial in the denominator eventually wins such that the series is convergent. Indeed, one can relate the energy to the expansion order n , as we will see in the following. Up to the constant dropped in eq. (3.24), the energy is given by a derivative with respect to β :

$$\begin{aligned} E - \sum_b J_b C_0 &= -\frac{d}{d\beta} \ln(\mathcal{Z}) = -\frac{1}{\mathcal{Z}} \frac{d\mathcal{Z}}{d\beta} \\ &= -\frac{1}{\mathcal{Z}} \sum_{n=0}^{\infty} \sum_{|\alpha\rangle} \sum_{S_n} \frac{n\beta^{n-1}}{n!} \langle \alpha | \prod_{p=0}^{n-1} \mathcal{H}_{b(p)}^{a(p)} | \alpha \rangle = -\left\langle \frac{n}{\beta} \right\rangle. \end{aligned} \quad (3.31)$$

Here the average $\langle \cdot \rangle$ refers to the same weights as in eq. (3.30) and n is the expansion order, i.e., the number of operators in the operator string. A further derivative gives the specific heat:

$$\begin{aligned} C_V &= \frac{dE}{dT} = -\beta^2 \frac{dE}{d\beta} = \beta^2 \left[\frac{1}{\mathcal{Z}} \frac{d^2\mathcal{Z}}{d\beta^2} - \frac{1}{\mathcal{Z}^2} \left(\frac{d\mathcal{Z}}{d\beta} \right)^2 \right] \\ &= \frac{\beta^2}{\mathcal{Z}} \sum_{n=0}^{\infty} \sum_{|\alpha\rangle} \sum_{S_n} \left(\frac{n(n-1)\beta^{n-2}}{n!} \langle \alpha | \prod_{p=0}^{n-1} \mathcal{H}_{b(p)}^{a(p)} | \alpha \rangle \right) - \frac{\beta^2}{\mathcal{Z}^2} \left(\frac{d\mathcal{Z}}{d\beta} \right)^2 \\ &= \langle n^2 \rangle - \langle n \rangle - \langle n \rangle^2. \end{aligned} \quad (3.32)$$

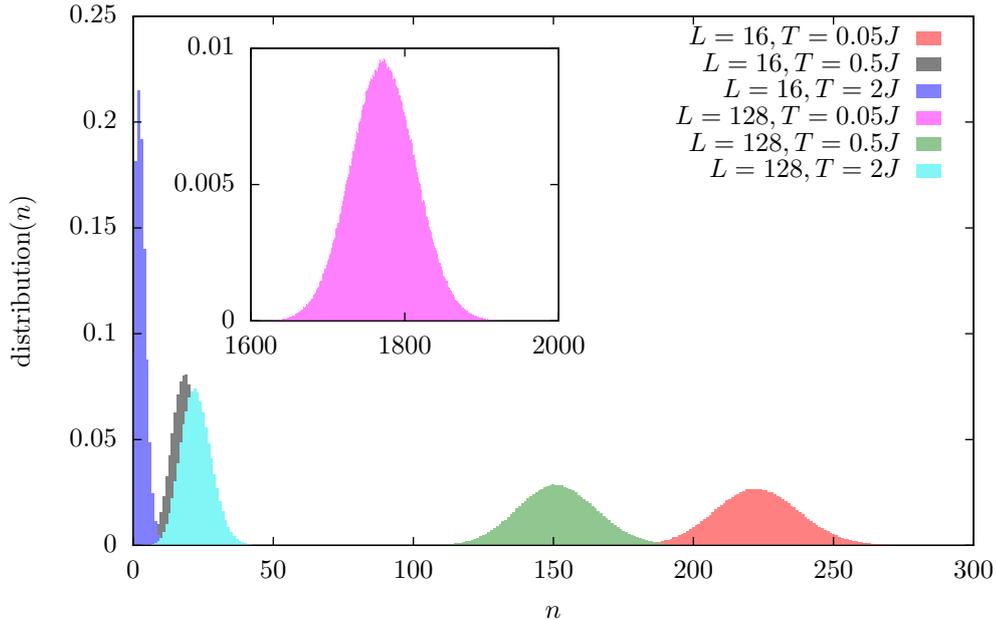


Figure 3.3: Distribution of the number of operators n in the operator string during a Monte Carlo simulation for a 1D isotropic chain ($\Delta = 1$) with periodic boundary conditions (PBC). For each curve 10^6 sweeps have been performed. The distribution is normalized such that the sum over all n gives 1.

Both the energy and the specific heat are finite, thus the expectation value $\langle n \rangle$ and the variance $\langle n^2 \rangle - \langle n \rangle^2$ are finite. This suggests that it should be possible to truncate the sum over n at some $L > \langle n \rangle$. The total weight of operator strings with larger length can then be safely neglected. In Figure 3.3 the distribution of 10^6 Monte Carlo sweeps shows that this is indeed the case. For high temperature $\beta \rightarrow 0$ only short operator strings give a relevant contribution. This is apparent also from a physical point of view: at high temperatures the spins are uncorrelated and we basically just have to sum over all basis states $|\alpha\rangle$. In contrast to that, at low temperatures the needed length scales with β . Since we expanded around $\beta = 0$, we can not expect that our method is well suited to determine the ground state. However, with some effort it is still possible to take the limit $T \rightarrow 0$ [52]. It is not clear a priori how large the cutoff has to be chosen. Thus, one starts with some relatively small cutoff, but allows to increase it in each step during the thermalization such that $L \approx 1.3n_{max}$ where n_{max} is the maximum number of operators encountered so far. One can also choose a higher factor than 1.3, but the computation time scales with L . Thus, it should not be chosen too high.

To avoid unnecessary memory copies in the program it is further a good idea to completely fix the length of the operator string to L . Therefore, one fills up the missing operators in the string with the identity operator $\mathbb{1}$. In other words one allows an additional element 'identity' for the index $a(p) \in \{\text{offd, diag, identity}\}$ and defines $\mathcal{H}_b^{\text{identity}} \equiv \mathbb{1}$ for any b . Let n furthermore denote the expansion order of the exponential, i.e., the number of operators which are not $\mathbb{1}$. This number n can still fluctuate. Since we allow to insert the identities at any of the L possible positions $p \in \{0, \dots, L-1\}$, there are $\binom{L}{L-n} = \frac{L!}{(L-n)!n!}$ operator strings with identities corresponding to the same operator string

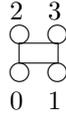


Figure 3.4: Convention for enumeration of the four legs of a vertex. Spins $\sigma_0, \dots, \sigma_3$ on legs 0 to 3 correspond to the matrix element $\langle \sigma_2 \sigma_3 | \mathcal{H}_{b(p)}^{a(p)} | \sigma_0 \sigma_1 \rangle$, i.e., leg 0 and 1 are incoming spins, leg 2 and 3 are outgoing.

without identities in eq. (3.30). To correct this combinatorial factor we substitute

$$\sum_{n=0}^L \sum_{S_n} \rightarrow \sum_{S_L} \frac{(L-n)! n!}{L!}, \quad (3.33)$$

where the sum on the right hand side includes additional identity operators, but not the one on the left hand. The left part of Figure 3.5 depicts an example for an operator string with $L = 5$ and $n = 3$. All in all, we arrive at the following expression for observable averages:

$$\begin{aligned} \langle \hat{A} \rangle &= \sum_{|\alpha\rangle} \sum_{S_L} w_{|\alpha\rangle, S_L} \langle \alpha | \hat{A} | \alpha \rangle \quad \text{with} \quad (3.34) \\ w_{|\alpha\rangle, S_L} &= \frac{\beta^n (L-n)!}{L!} \langle \alpha | \prod_{p=0}^{L-1} \mathcal{H}_{b(p)}^{a(p)} | \alpha \rangle \\ &= \frac{\beta^n (L-n)!}{L!} \prod_{p=0}^{L-1} \langle \alpha(p+1) | \mathcal{H}_{b(p)}^{a(p)} | \alpha(p) \rangle. \quad (3.35) \end{aligned}$$

In the last line, we used that the propagated states are determined uniquely for a given operator string and initial state. In fact, these matrix elements $\langle \alpha(p+1) | \mathcal{H}_{b(p)}^{a(p)} | \alpha(p) \rangle$ clearly only depend on the spin states on the two sites $i(b(p))$ and $j(b(p))$ connected by the bond $b(p)$ before and after the operator. We say that each operator has four legs – two incoming legs labeled 0 and 1, and two outgoing legs 2 and 3, as depicted in Figure 3.4. We further call the operator together with the four spin states at the legs a vertex.

3.3 Sampling

The result (3.34) now allows the evaluation along the lines described in section 3.1. The configuration space consists of all possible initial states α and the operator strings S_L . In the following we will give the details of the Monte Carlo sweeps used for the sampling.

A single sweep is split into two parts: First, diagonal operators are inserted or removed from the operator string which leads to a fluctuation of n . Then so-called loops are constructed and flipped which leads to changes from diagonal to off-diagonal operators (and back) and also includes updating the initial state α . If these updates both separately satisfy detailed balance (3.10), each of them leaves $w_\nu(t)$ invariant and so does the composition of both updates.

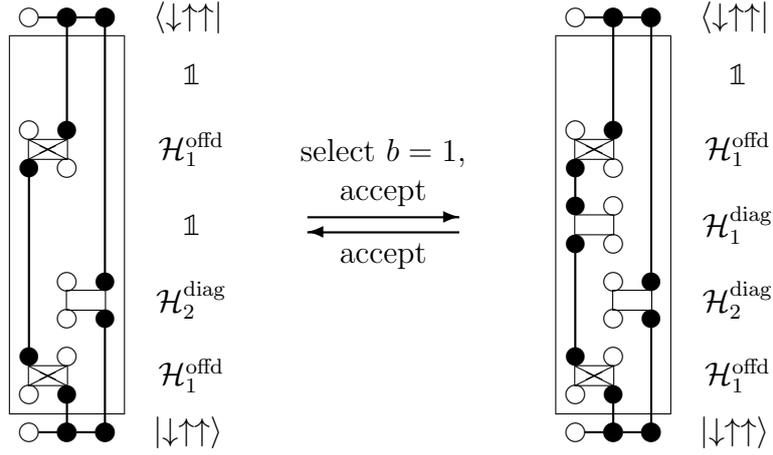


Figure 3.5: Simple example of a diagonal update for operator strings of length $L = 5$. From left to right a diagonal operator is inserted at the propagated index $p = 2$ such that the number of operators changes from $n = 3$ in the left to $n = 4$ on the right.

3.3.1 Diagonal Updates

As we saw in section 3.2, only an even number of off-diagonal operators is allowed in an operator string (i.e., leads to a nonzero weight). Thus, one can not insert or remove a single off-diagonal operator. Instead, one restricts this part of the update to switch off-diagonal operators with identities $\mathbb{1}$ to change the number of non-identity operators n . An example for such an update is shown in Figure 3.5.

To insert an operator at some $p \in \{0, \dots, L-1\}$ into an operator string S_L it is necessary to first select a bond b at which this operator is supposed to be inserted; let P_b^{sel} be the probability to select the bond b . In the spatially isotropic case one can select uniformly one of the N_B bonds, $P_b^{sel} = \frac{1}{N_B}$. We denote the resulting operator string with S_L^{ins} . We accept this insertion $S_L \rightarrow S_L^{ins}$ with some probability P_{ins}^{acc} . On the other hand the removal of an operator does not need a selection and we simply accept it with probability P_{rem}^{acc} . Thus, we have

$$P(S_L \rightarrow S_L^{ins}) = P_b^{sel} P_{ins}^{acc} \quad \text{and} \quad P(S_L^{ins} \rightarrow S_L) = P_{rem}^{acc}. \quad (3.36)$$

To satisfy detailed balance (3.10) we need to require

$$\begin{aligned} \frac{P_{ins}^{acc}}{P_{rem}^{acc}} &= \frac{P(S_L \rightarrow S_L^{ins})}{P_b^{sel} P(S_L^{ins} \rightarrow S_L)} \stackrel{!}{=} \frac{1}{P_b^{sel}} \frac{w_{S_L^{ins}}}{w_{S_L}} \\ &= \frac{1}{P_b^{sel}} \frac{\beta^{n+1} \frac{(L-(n+1))!}{L!} \prod_{p=0}^{L-1} \langle \alpha(p+1) | \mathcal{H}_{b(p)}^{\alpha(p, S_L^{ins})} | \alpha(p) \rangle}{\beta^n \frac{(L-n)!}{L!} \prod_{p=0}^{L-1} \langle \alpha(p+1) | \mathcal{H}_{b(p)}^{\alpha(p, S_L)} | \alpha(p) \rangle} \\ &= \frac{1}{P_b^{sel}} \frac{\beta}{L-n} \langle \alpha(p+1) | \mathcal{H}_b^{\text{diag}} | \alpha(p) \rangle. \end{aligned} \quad (3.37)$$

Given the relative factor of the acceptance probabilities eq. (3.37), we choose them in

analogy to the Metropolis-Hastings algorithm (3.20):

$$P_{ins}^{acc} = \min \left\{ 1, \frac{P_{ins}^{acc}}{P_{rem}^{acc}} \right\}, \quad P_{rem}^{acc} = \min \left\{ 1, \left(\frac{P_{ins}^{acc}}{P_{rem}^{acc}} \right)^{-1} \right\}. \quad (3.38)$$

In the case of spatial anisotropies J and J' , it is a good idea – although not strictly necessary – to balance the insertion of new operators. For $J' \rightarrow 0$ one would often need to reject an insertion. Thus, we first choose the type of the bond: we choose J with probability $\frac{JN_B(J)}{JN_B(J)+J'N_B(J')}$ and then one of the bonds of the corresponding type with $\frac{1}{N_B(J)}$, where $N_B(J)$ is the number of bonds with strength J . The total selection probability is

$$P_b^{sel} = \begin{cases} \frac{J}{JN_B(J)+J'N_B(J')} & \text{if } b \text{ has strength } J, \\ \frac{J'}{JN_B(J)+J'N_B(J')} & \text{if } b \text{ has strength } J'. \end{cases} \quad (3.39)$$

This choice reduces to the previous value $P_b^{sel} = \frac{1}{N_B}$ for $J = J'$, but does not select bonds of type J' for $J' \rightarrow 0$.

It is not necessary to store all the propagated states $|\alpha(p)\rangle$ in the program, but it suffices to store only the initial state $|\alpha\rangle$ and generate $|\alpha(p)\rangle$ on the fly. For the complete diagonal update, we iterate over the propagation index $p = 0, \dots, L-1$ in a loop. The action within the loop depends on $a(p)$: If $a(p) = \text{offd}$, we flip the spins on the corresponding sites of bond $b(p)$ to keep the propagated state up to date. Else, we try to insert or operate at the position p as described above in the case $a(p) = \text{identity}$; or try to remove it for $a(p) = \text{diag}$, respectively. In the end of the loop the original state $|\alpha\rangle$ is restored due to the periodicity eq. (3.27).

3.3.2 Deterministic Loop updates

We now need to find updates changing between off-diagonal and diagonal operators. We will first look at the easier example of a special case, namely the $SU(2)$ symmetric couplings with $\Delta = 1$. The matrix elements of the operators listed in Table 3.1 are then particularly simple for the choice $C_0 = \frac{1}{4}$:

$$\langle \alpha(p+1) | \mathcal{H}_{b(p)}^{a(p)} | \alpha(p) \rangle = \begin{cases} 0 & \text{for } |\alpha(p)\rangle \in \{|\uparrow\uparrow\rangle, |\downarrow\downarrow\rangle\}, \\ 0.5 & \text{for } |\alpha(p)\rangle \in \{|\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle\}. \end{cases} \quad (3.40)$$

In other words, we only need to include vertices into the operator string, where the spins on the two incoming legs 0 and 1 are different, as well as the spins at the outgoing legs 2 and 3. In that case, the matrix elements – and thus the weights of the vertices – are all the same. As mentioned before, it is not possible to change just one operator from diag to offd since the number of off-diagonal operators needs to be even. Indeed, by definition, any off-diagonal operator flips the spins on the sites $i(b)$ and $j(b)$ connected by the bond b , whereas diagonal operators keep them unchanged. In the simplest case the next operator in the operator string – or at least the next operator acting on any of the two sites – acts on the same bond. Then, we can update the type of both operators consistently by changing the propagated states in between. Such an update is indicated

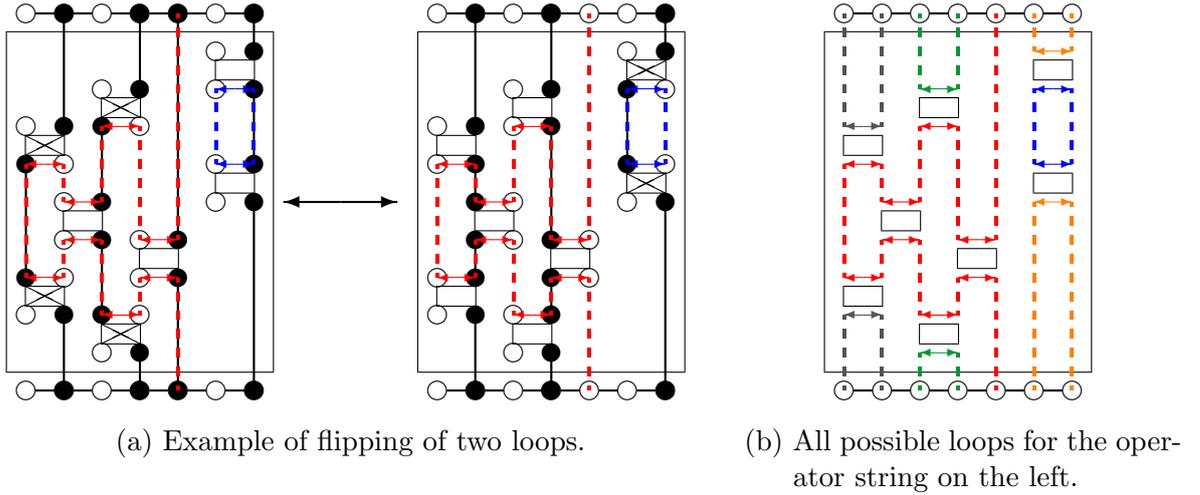


Figure 3.6: Example of deterministic loop updates.

in blue in the upper right corner of the operator string in Figure 3.6(a). All spins in the propagated states touching the blue loop need to be flipped at once; flipping just a part of them would result in an invalid operator string with zero total weight.

Let us now consider a more complicated example to see how we construct loops. Suppose we want to flip the operator in the lower left corner at the first bond. Therefore we start by flipping the spins at the legs 2 and 3 above the operator, which are at the sites 1 and 2. We have two open ends of the loop, one at leg 2 and one at leg 3. Now follow the spin from leg 3 on site 2 upwards: there is another operator acting on site 2 but not 1. Thus, we also have to flip the spins at the legs 0 and 1 of that operator – otherwise there would be an operator acting on parallel spins such that the whole operator string is invalid. To construct a loop we thus do two alternating moves:

- (i) We move horizontally between the incoming legs 0 and 1 or between the outgoing legs 2 and 3. In that way, we move along the bonds between different sites.
- (ii) During vertical movements we stay on one site and move away from the last operator – upwards from legs 2 or 3, and downwards from legs 0 or 1 – until we encounter another operator acting on that site. It may also happen that we do not encounter another operator, but we reach the initial state $|\alpha\rangle$. Then we need to flip the spin on the corresponding site both at $p = 0$ and $p = L$ due to eq. (3.27). That way we can go on moving on the open end of the loop: we just have a periodicity in p .

Eventually, we end up at the starting point connecting the two open loop ends such that the loop closes. Note also that in the given example in Figure 3.6(a), the red loop touches the spins both below and above the operator at the second bond. Thus, the flip of the red loop does not change the type of that operator; it remains diagonal here.

It is a good idea to construct links between the operator legs of operators for the vertical movement before actually performing the loop updates. These links have to be reconstructed after each diagonal update where operators are inserted or removed. Details on how to construct these links in an efficient way can be found in Ref. [52].

Once the position of all operators in the operator string is fixed, the loops are defined uniquely; hence the name deterministic loop update. Figure 3.6(b) shows all possible

loops for the same exemplary operator string as in Figure 3.6(a). Flipping any of these loops results in a configuration with the same total weight: n is not changed by loop updates in eq. (3.34), and the matrix elements of off-diagonal and diagonal operators are the same, cf. eq. (3.40). It is, however, not a good idea to flip all the loops: this would just flip all spins at any site and p and thus not give a new configuration. Instead, one flips each of the loops with some constant probability $P_0 \in (0, 1)$, a good choice seems to be $P_0 = \frac{1}{2}$ [52]. During the loop update also the initial state $|\alpha\rangle$ is updated when a loop goes through it. If there is a site on which no operator acts for any p , this can be seen as a loop, too. Accordingly, one also flips the spin on such empty sites with probability P_0 .

The loop updates clearly satisfy detailed balance (3.10). In addition with the diagonal updates described previously, this update scheme is also ergodic⁶; the combination of them both changes very effectively between possible configurations. In a Monte Carlo simulation one can start from an initially empty operator string consisting only of identities $\mathbb{1}$. The diagonal updates then successively inserts (and later removes) new diagonal operators, and the loop update changes some of them to off-diagonal ones and updates the initial state $|\alpha\rangle$.

In a first approach to get away from $\Delta = 1$, one can just use the same scheme of deterministic loop updates. Therefore one has to choose $C_0 = \frac{\Delta}{4}$ for the matrix elements in Table 3.1 and adjust the probability of flipping a loop according to the Metropolis-Hastings algorithm. In that case the weight depends on the number of diagonal operators $w_{\alpha, S_L} \propto (\Delta)^{n^{\text{diag}}}$. Thus, one needs to count how many off-diagonal operators would be changed to diagonal ones, and vice versa, during a flip of the loop. Since operators visited twice by a loop do not change their type, this requires some additional book keeping. The update probability can then be chosen as

$$P^{\text{acc}} = P_0 \min \left\{ 1, (\Delta)^{n_{\text{new}}^{\text{diag}} - n_{\text{old}}^{\text{diag}}} \right\} \quad \text{with } P_0 = \text{const} \in (0, 1). \quad (3.41)$$

Here P_0 should be chosen higher than before, e.g., $P_0 = \frac{3}{4}$, since some of the loops will not be flipped. However, this approach has limitations. For example, it is not possible to take the limit $\Delta \rightarrow 0$ towards the XY-model: the matrix elements of the diagonal operator $\langle \alpha(p+1) | \mathcal{H}^{\text{diag}} | \alpha(p) \rangle = \frac{\Delta}{2}$ vanish and thus any insertion of operators during the diagonal update gets extremely improbable. The way out is to choose another C_0 in that case. But then $\mathcal{H}^{\text{diag}}$ acting on parallel spins does not vanish, which was needed to determine the building rules of the loop above.

3.3.3 Directed Loop Updates

The particular case that only four out of the eight matrix elements in Table 3.1 do not vanish and have the same weight is restricted to only very few models. But the SSE is not restricted to these models. For example, it is possible to include a magnetic field or to simulate bosons [61]. In our case the Hamiltonian preserves the total magnetization –

⁶ It is also easy to construct configurations where a loop update changes winding numbers, e.g., around a plaquette in a 2D square lattice: the loop itself has to wind around the plaquette. That would not be possible if one would be restricted to local updates of only two operators following each other at the same bond; one needs to cut the world lines (say of up-spins) and reconnect them later.

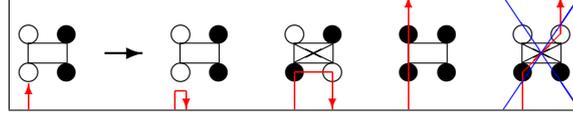


Figure 3.7: Example of the choice of an exit leg when the entry is at leg 0. Possible exits are from left to right at legs 0,1,2,3. An exit at the same leg as the entry is called bounce and does not lead to any changes. One of the possible exits leads to a violation of S^z conservation; in this case the exit at leg 3.

corresponding to particle conservation for bosons – such that the two matrix elements in the upper left corner of Table 3.1 always vanish leading to the six depicted vertices.

The general framework of the SSE still works; but we have to find another scheme for the off-diagonal updates: when we enter a vertex with an open end of a loop, we are not forced to a unique exit leg anymore. Instead, we will randomly choose one of the four legs as an exit. As was pointed out by Syljuåsen and Sandvik in 2002 [60], there are multiple possibilities how to choose the exit leg. The most efficient solution leads to the directed loop updates. In the following we will derive the directed loop equations and describe how to solve them. We follow Ref. [60] and [61], further details can be found therein.

Suppose we have an initial configuration $\nu \equiv \nu_0 = (|\alpha\rangle, S_L)$. To start the construction of a loop we randomly select a leg l_0 of any operator p_0 in the string; for simplicity we combine these indices in $e_0 = (p_0, l_0)$. Then we move one of the two open ends around from one operator leg to the next, $e_i \rightarrow e_{i+1}$. As before, we make two alternating movements:

- (i) The horizontal movement changes: we enter a vertex V_i of the configuration ν_i at position p_i in the operator string in a leg l_i . Now we are not forced to a unique exit leg as in the deterministic loop updates. Instead, we choose an exit leg \tilde{l}_i with some probability $P(V_i; l_i \rightarrow \tilde{l}_i)$ depending on the vertex and the entry leg. Then we flip the spins both at the entry and the exit leg l_i and \tilde{l}_i , resulting in a new vertex \tilde{V}_i . An example is depicted in Figure 3.7.
- (ii) The vertical movement $(p_i, \tilde{l}_i) \rightarrow (p_{i+1}, l_{i+1})$ is exactly the same as for the deterministic loop updates. As before, one should connect corresponding legs with links right after the diagonal updates. During the construction of the loops one can then simply follow these links.

Since the vertical movement is unique, the probability for a change $e_i \rightarrow e_{i+1}$ depends only on the vertex:

$$P(\nu_i, e_i \rightarrow \nu_{i+1}, e_{i+1}) = P(V_i; l_i \rightarrow \tilde{l}_i). \quad (3.42)$$

Finally, we close the loop after some number of steps m , reconnect the open ends $e_m = e_0$, and end up in a configuration $\nu' \equiv \nu_m$.

The probability for an update from the initial configuration ν to the final configuration ν' is a sum over all possible loops leading to that change:

$$\begin{aligned} P(\nu \rightarrow \nu') &= \sum_m \sum_{\{\nu_i, e_i\}} P^{sel}(e_0) P(\nu_0, e_0 \rightarrow \nu_1, e_1) \cdots P(\nu_{m-1}, e_{m-1} \rightarrow \nu_m e_m) \\ &= \sum_m \sum_{\{\nu_i, e_i\}} P^{sel}(e_0) P(V_0; l_0 \rightarrow \tilde{l}_0) P(V_{m-1}; l_{m-1} \rightarrow \tilde{l}_{m-1}). \end{aligned} \quad (3.43)$$

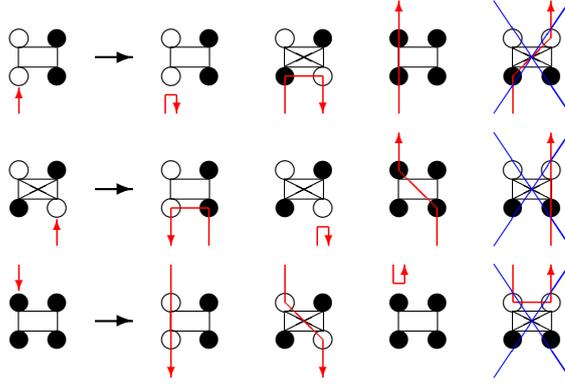


Figure 3.8: Example for a set of three vertices with entry legs which are coupled by the directed loop equations (3.47). The resulting vertices are the same in all three lines, see text.

Strictly speaking we have intermediate configurations ν_i with weight 0 since the open ends e_i correspond to discontinuities in the propagated state. However, the total weight eq. (3.35) consists of the product of matrix elements. We can then trace back the resulting total change of weight to the changes of the vertices in the single steps $i \rightarrow i + 1$. This motivates to call the matrix elements the weight of single vertices

$$w_{V_i} := \langle \alpha(p_i + 1) | \mathcal{H}_{b(p_i)}^{a(p_i)} | \alpha(p_i) \rangle. \quad (3.44)$$

To satisfy detailed balance (3.10) for the complete loop flip eq. (3.43) we thus require detailed balance in each single vertex update $V_i \rightarrow \tilde{V}_i$ for any $i \in 0, \dots, m^7$:

$$P(V; l \rightarrow \tilde{l}) w_V \stackrel{!}{=} P(\tilde{V}; \tilde{l} \rightarrow l) w_{\tilde{V}}. \quad (3.45)$$

Further, we have a normalization condition since $P(V; l \rightarrow \tilde{l})$ is supposed to be a probability; we have to choose one of the legs as exit:

$$\sum_{\tilde{l}=0}^3 P(V; l \rightarrow \tilde{l}) \stackrel{!}{=} 1. \quad (3.46)$$

The equations (3.45) and (3.46) form a set of coupled linear equations which are conditions for $P(V; l \rightarrow \tilde{l})$. With the definition $a(V, l, \tilde{l}) := P(V; l \rightarrow \tilde{l}) w_V$ they can be written as

$$\begin{aligned} a(V, l, \tilde{l}) &= a(\tilde{V}, \tilde{l}, l), \\ \sum_{\tilde{l}} a(V, l, \tilde{l}) &= w_V. \end{aligned} \quad \text{(Directed Loop Equations)} \quad (3.47)$$

It turns out that these equations decouple to smaller sets of equations containing just 3 vertices (or in the more general case without S^z conservation 4), as will be described in the following. Let us start with some vertex V_0 and some entry leg, say leg 0 at the vertex depicted in the upper left corner of Figure 3.8. There are only 3 possible resulting

⁷ For simplicity we drop the index i in the following.

vertices since one of the four possible exit legs violates spin conservation; in the given example it is exit leg 3. However, in one case we also choose the same exit leg as entry, we bounce back without changing the vertex. Thus, there are only two further resulting vertices V_1 and V_2 , in this case for the exit legs 1 and 2. But now consider entering the second vertex V_1 at the previous exit leg 1 as shown in the second line of Figure 3.8. If we exit at leg 0, we end up with V_2 since we go exactly the reverse path which led to that vertex in the first line. More generally, when we enter a vertex at leg l and exit at leg l' leading to a vertex V' and then enter V' at the same leg l' , we undo the flip at this leg l' . Thus, an exit at leg l'' leads to the *same* vertex V'' as if we would have gone from l to l'' immediately in V . Consequently, all resulting vertices in the second line are the same as in the first line. This is also the case in the third line where we enter V_2 at the exit leg of the first line, leg 2.

With the shorthand notation $a_{ij} := a(V_i, i, j)$ and $W_i := w_{V_i}$, $i, j \in \{0, 1, 2, 3\}$ we can immediately write down the corresponding directed loop equations

$$\begin{aligned} a_{00} + a_{01} + a_{02} &= W_0, \\ a_{10} + a_{11} + a_{12} &= W_1, \\ a_{20} + a_{21} + a_{22} &= W_2, \end{aligned} \qquad a_{ij} = a_{ji}. \quad (3.48)$$

Thus, we have three equations for six degrees of freedom a_{ij} , $i \leq j$, i.e., the solution is not unique. All in all, we have four possible entry legs to six vertices which makes eight groups of eq. (3.48). However, these are all equivalent due to three symmetries: (i) change the direction of the bond $i(b) \leftrightarrow j(b)$, i.e., left and right in the pictures, (ii) change in- and outgoing legs $(0, 1) \leftrightarrow (2, 3)$, i.e., up and down in the pictures, and (iii) flip all spins at once, i.e., exchange filled and empty circles. The last symmetry is not given for a nonzero magnetic field h ; then one actually has to solve two such groups of (3.48).

A simple and quite symmetric solution is given by

$$\begin{aligned} a_{ij} &= \frac{W_i W_j}{W_0 + W_1 + W_2}, \\ P(V_i; i \rightarrow j) &= \frac{a_{ij}}{W_i} = \frac{W_j}{W_0 + W_1 + W_2}. \end{aligned} \quad \text{(Heat Bath Solution)} \quad (3.49)$$

This was the first general solution proposed for directed loops [67]. A big advantage of the directed loops is, that it takes into account possible anisotropies or magnetic fields leading to different weights of the vertices from scratch. That way the loops can be flipped with unit probability already during the construction. In contrast to that, the naive approach described at the end of section 3.3.2 took the different weights into account a posteriori; there the loops can not always be flipped, which leads to unnecessary work without changes of the configuration.

However, there is also unnecessary work done during the construction of directed loops: in every bounce process, when we enter and exit a vertex at the same leg, the vertex does not change and even undoes the flip at the exit leg of the previous vertex since it enters there. This motivates to look for solutions minimizing the bounce probability a_{ii} . The heat bath solution (3.49) generically leads to high bounce probabilities; in particular the bounce probability even approaches $\frac{1}{2}$ when two weights are equal, $W_0 = W_1, W_2 = 0$; this is the case for the deterministic loop update at $\Delta = 1$.

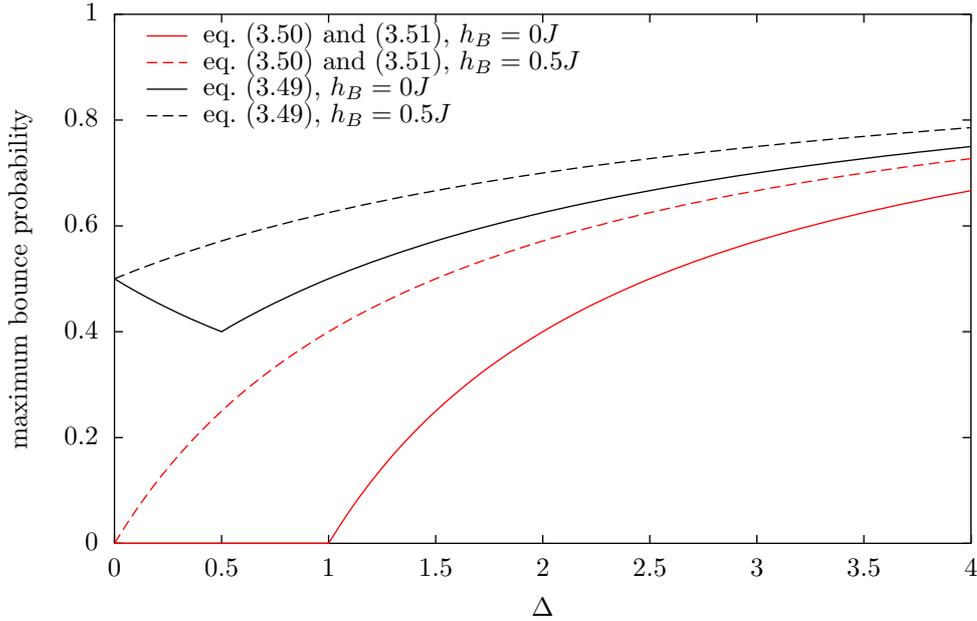


Figure 3.9: Comparison of the largest bounce probabilities between the heat bath solution, eq. (3.49), and the solution minimizing the bounce, eq. (3.50) or (3.51), respectively. Here $C_0 = \frac{\Delta}{4} + 0.25$ and h_B is the magnetic field per bond.

It is possible to find solutions with less bouncing [60, 61]. There is only one completely bounce free solution $a_{ii} = 0$ to eq. (3.48):

$$\begin{aligned} a_{01} &= (W_0 + W_1 - W_2)/2, \\ a_{02} &= (W_0 - W_1 + W_2)/2, \\ a_{12} &= (-W_0 + W_1 + W_2)/2, \end{aligned} \quad a_{00} = a_{11} = a_{22} = 0. \quad (3.50)$$

Since we need positive probabilities $a_{ij} \geq 0$, this bounce free solution can only be used if no term in the Hamiltonian dominates⁸, $W_0 \leq W_1 + W_2$. Otherwise, we can use the solution with a bouncing of the vertex with highest weight [61]

$$\begin{aligned} a_{00} &= W_0 - W_1 - W_2, \\ a_{01} &= W_1, \\ a_{02} &= W_2, \end{aligned} \quad a_{11} = a_{22} = 0 = a_{12}. \quad (3.51)$$

As shown in Figure 3.9, the bounce probability is indeed dramatically reduced by this solution compared to the heat bath (3.49). Note also that there is only one bounce in eq. (3.51), whereas in the heat bath solution there is a possible bounce for every entry.

The solutions eq. (3.50) and (3.51) (or the heat bath solution (3.49)) can be hard-coded in the program to determine the list of exit leg selection probabilities $P(V, l \rightarrow \tilde{l})$ for the given parameters during the initialization of the program. Further, it is necessary to store all four spin states for each operator in the operator string, such that one can look up the corresponding selection probabilities and flip the spins during the loop update. To ensure positive weights for all six vertices, C_0 has been chosen as $C_0 = J(\frac{\Delta}{4} + 0.25)$ in

⁸ We can assume without loss of generality that $W_0 \geq W_1 \geq W_2$.

the following. The constant 0.25 helps balancing the weights as it ensures that all six vertex weights are nonzero; the exact value is, however, irrelevant for the results. But the average number of operators $\langle n \rangle$ increases for larger C_0 ; thus, one should not choose it too large.

Of course, one should flip more than one loop during a Monte Carlo sweep such that a significant fraction of vertices changes. However, to avoid a bias, the number N_l of loop flips in each sweep should be constant during the measurement. The length of the single loops can vary a lot⁹. Thus, one can record the average loop length during the thermalization and periodically adjust N_l such that approximately $2L$ legs are visited during one sweep.

3.4 Observables and Improved Estimators

The SSE algorithm described above calculates thermodynamic weighted averages according to eq. (3.34) for any observable diagonal in the S^z basis. Simple examples for such diagonal operators are the magnetization, the order parameter in a ferromagnetic phase,

$$M = \left\langle \sum_i S_i^z \right\rangle \quad (3.52)$$

and the staggered magnetization, which serves as order parameter in the antiferromagnetic phase:

$$M_s = \left\langle \sum_i \phi_i S_i^z \right\rangle \quad \text{with } \phi_i = (-1)^{x_i+y_i}. \quad (3.53)$$

In some cases the statistical error can be reduced by a simple trick which will be explained in the following. Basically, there is nothing special about the initial state $|\alpha\rangle \equiv |\alpha(p=0)\rangle$ compared to the other propagated states $|\alpha(p)\rangle$. In particular one can permute the complete operator string $p+1 \rightarrow p \pmod L$ due to the periodicity in the imaginary time direction. This permutation leads to a new configuration with another initial state $|\alpha\rangle$, namely the previous first propagated state $|\alpha(p=1)\rangle$, but the same weight $w_{|\alpha(p=1)\rangle, S_L^{perm}} = w_{|\alpha\rangle, S_L}$. Clearly, one can iterate this procedure for any shift; thus, eq. (3.34) can be reexpressed with an average over all propagated states:

$$\langle \hat{A} \rangle = \sum_{|\alpha\rangle} \sum_{S_L} w_{|\alpha\rangle, S_L} \langle \alpha | \hat{A} | \alpha \rangle = \sum_{|\alpha\rangle} \sum_{S_L} w_{|\alpha\rangle, S_L} \frac{1}{L} \sum_{p=0}^{L-1} \langle \alpha(p) | \hat{A} | \alpha(p) \rangle. \quad (3.54)$$

Since two states $|\alpha(p+1)\rangle$ and $|\alpha(p)\rangle$ differ at most by the two sites at the bond $b(p)$ – if $\mathcal{H}_{b(p)}^{\alpha(p)}$ is off-diagonal – it is often easy to calculate the changes in the corresponding matrix elements $\langle \alpha(p) | \hat{A} | \alpha(p) \rangle$ with low computational effort. For example, the matrix

⁹ One might even worry that a loop never closes. To avoid that problem one can set an upper limit to the loop length, e.g., $100L$ for an operator string of fixed length L . If the loop did not close within so many steps, one can proceed immediately with the next Monte Carlo sweep discarding all changes during the loop update [60].

element of the staggered magnetization M_s has to be calculated only for the initial state from all N spins and can then be updated according to the local changes. The additional computational effort $\mathcal{O}(L)$ then easily pays off. Especially at low temperatures, where L is large, the additional average may dramatically decrease the statistical errors.

However, the trick does not work for the total magnetization M : M commutes with every local $\mathcal{H}_{b(p)}^{a(p)}$, the matrix element does not change at all in this case and the additional average does not help and should not be performed in that case.

To calculate the static structure factor

$$S(k) = \langle |S_k^z|^2 \rangle \quad \text{with } S_k^z = \frac{1}{\sqrt{N}} \sum_i e^{-ik \cdot i} S_i^z \quad (3.55)$$

for all momenta k , a fast Fourier transform has to be performed for each measurement of the matrix element. Thus, the author decided not to do the average over p ; otherwise this measurement would be $\mathcal{O}(N^2 \log(N))$ and completely dominate the calculations for large N . With $S(k)$ one can calculate a further Fourier transform post simulation to obtain the spatial correlations:

$$\langle S_i^z S_{i+r}^z \rangle = \frac{1}{N} \sum_i \langle S_i^z S_{i+r}^z \rangle = \frac{1}{N^2} \sum_{k,k'} \sum_i e^{ik \cdot i} e^{ik' \cdot (i+r)} \langle S_k^z S_{k'}^z \rangle = \frac{1}{N} \sum_k e^{-ik \cdot r} S(k). \quad (3.56)$$

For the deterministic loop updates at $\Delta = 1$, there exist so-called improved estimators reducing the noise further [52]. Since flipping a loop does not change the weight in that case, one can average over all possible loop flips which is often analytically possible. Although the improved estimators have been implemented for the deterministic loop updates, we will here refer the reader to the literature [52, 63] since we will mainly study anisotropic couplings $\Delta \neq 1$ in the following.

Susceptibilities

Often we want to calculate the linear response of an observable \hat{A} to a small perturbation $\mathcal{H} \rightarrow \mathcal{H} + b\hat{B}$.

$$\chi_{\hat{A}\hat{B}} = \left. \frac{\partial}{\partial b} \right|_{b=0} \langle \hat{A} \rangle_{\mathcal{H} + b\hat{B}}. \quad (3.57)$$

Such a generalized susceptibility is given by the Kubo formula [52]

$$\chi_{\hat{A}\hat{B}} = \int_0^\beta \langle \hat{A}(\tau) \hat{B} \rangle d\tau - \beta \langle \hat{A} \rangle \langle \hat{B} \rangle \quad \text{with } \hat{A}(\tau) = e^{\tau\mathcal{H}} \hat{A} e^{-\tau\mathcal{H}}. \quad (3.58)$$

A rigorous calculation given in appendix B shows that the average over the imaginary time corresponds to an average over the propagation index p , as one can expect from the correspondence between SSE and path integral Monte Carlo.

It is easy to see that this expression reduces to the connected correlator if A or B commute with the Hamiltonian. In practice this only happens for the susceptibility to the uniform magnetization

$$\chi = \frac{1}{N} \chi_{M,M} = \frac{1}{N} \frac{\partial \langle M \rangle}{\partial h} = \frac{\beta}{N} (\langle M^2 \rangle - \langle M \rangle^2) = \frac{\beta}{N} \langle M^2 \rangle. \quad (3.59)$$

Off-diagonal operators

The expression for the average eq. (3.34) (or equivalently (3.54)) is only valid for diagonal operators. For off-diagonal operators one needs to insert the identity operator $\mathbb{1} = \sum_{|\alpha'\rangle} |\alpha'\rangle \langle\alpha'|$ in the derivation of an analogous formula, such that the operator strings in general do not need to be periodic, $|\alpha(p=L)\rangle = |\alpha'\rangle \neq |\alpha\rangle = |\alpha(p=0)\rangle$. This leads to a sign problem as the bipartite lattice does not ensure an even number of off-diagonal operators $\mathcal{H}_{b(p)}^{\text{offd}}$ anymore. We will thus not discuss general off-diagonal operators here. However, there is a simple case that the off-diagonal operator is a part of the Hamiltonian itself. Indeed we have already derived eq. (3.31) for the total energy, which clearly is off-diagonal. Generalizing the derivation of eq. (3.31) (see appendix A), we obtain

$$\langle \mathcal{H}_b \rangle = \left\langle \frac{n_b}{\beta} \right\rangle \quad \text{for } \mathcal{H} = \text{const} - \sum_b \mathcal{H}_b, \quad (3.60)$$

where n_b is just the number of operators with index b in the operator string [23, 52]. We can go even further and distinguish between all vertices. For the calculation of the singlet and triplet projections in section 6.1 it will be necessary to split the diagonal part. The off-diagonal part can be split into $S_{i(b)}^+ S_{j(b)}^-$ and $S_{j(b)}^- S_{i(b)}^+$, which allows to calculate the current, e.g., in x -direction, from the number n^r of off-diagonal vertices transporting \uparrow to the right, or to the left n^l , respectively:

$$\langle \hat{j}_x \rangle = \left\langle \sum_{b \text{ x-dir.}} \left(S_{i(b)}^- S_{j(b)}^+ - S_{i(b)}^+ S_{j(b)}^- \right) \right\rangle = \sum_{b \text{ x-dir.}} (\langle n_b^r \rangle - \langle n_b^l \rangle). \quad (3.61)$$

Clearly, the current vanishes in thermodynamic equilibrium, but the current-current susceptibility used for the spin stiffness does not need to do so [52]. However, these quantities were not calculated in this thesis and will thus not be discussed further.

3.5 Program Verification

To spot possible logic errors in the written code it is essential to compare the results with known data. For this thesis the author checked the correctness mainly by comparison with data obtained from an exact diagonalization (ED) of the Hamiltonian as he has done for his bachelor's thesis [68]. The energy for different values of Δ in a one dimensional chain is exemplarily depicted in Figure 3.10. Figure 3.11 shows the difference between ED and SSE data in more detail such that the error bars are visible. For the energy the error bars tend to get smaller with decreasing temperature (i.e., increasing β) which can be related to the fact that we use an improved estimator (averaging over all propagated states) for the energy. In contrast to that the statistical errors for the specific heat increase dramatically with β and are generically larger. For this observable we need to take a difference between the variance $\langle n^2 \rangle - \langle n \rangle^2$ and average $\langle n \rangle$ of the number of operators n , see eq. (3.32). While C_V vanishes for $T \rightarrow 0$, both the variance and average of n increase with β , compare Figure 3.3. Thus, even if the relative error in $\langle n^2 \rangle$ and $\langle n \rangle$ can be kept constant, a large cancellation error appears in the final result due to the subtraction.

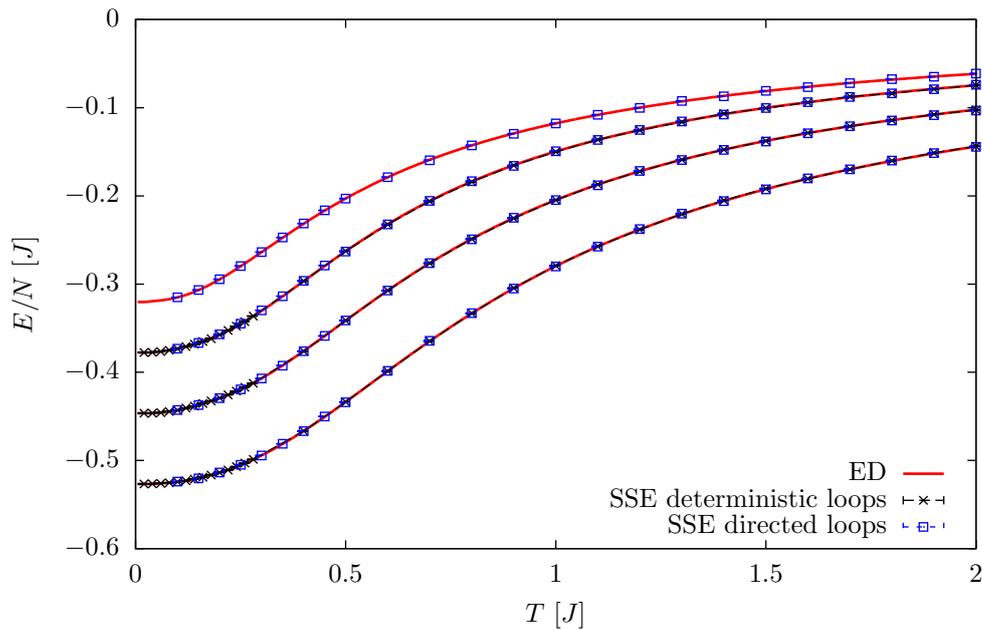


Figure 3.10: Energy per site of an 1D chain with PBC, $L = 16$, for $\Delta = 0, 0.5, 1.0, 1.5$ from top to bottom. The solid red line shows data obtained by an exact diagonalization (ED) of the Hamiltonian. The black crosses are obtained with deterministic loop updates as described in section 3.3.2, for $\Delta = 0.5$ and $\Delta = 1.5$ see the last paragraph therein. This method is not possible for $\Delta = 0$. The blue squares are obtained with directed loop updates as described in section 3.3.3. Error bars are included, but much smaller than the symbol size.

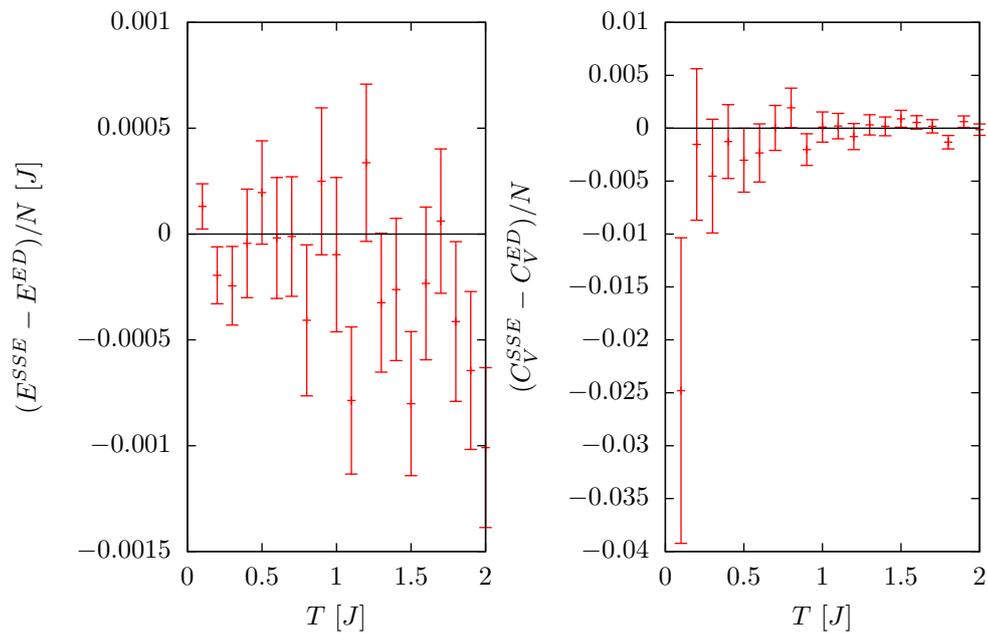


Figure 3.11: Difference between SSE data with directed loop updates and ED, showing the error bars for different temperatures, for a 1D chain with PBC, $L = 16, \Delta = 1$. For each point 10^6 sweeps have been performed. The left graph shows the energy, the right graph the specific heat.

4 Phase Transitions

We will now focus on anisotropic couplings with $\Delta = 4$ on a two-dimensional (2D) square lattice. To minimize boundary effects we use periodic boundary conditions. The lattice with $N = L^2$ sites is not frustrated for even side length L . As we have seen in section 2.3, this choice of Δ is deep in the Ising regime where we expect to see a finite temperature phase transition. We expect the ground state to be close to the classical Néel state $|\uparrow\downarrow\uparrow\downarrow\dots\rangle$ but with some quantum fluctuations. The correct order parameter is the average of the staggered magnetization

$$m_s = \frac{M_s}{N} = \frac{1}{N} \sum_i \phi_i S_i^z \quad \text{with } \phi_i = (-1)^{x_i+y_i}. \quad (4.1)$$

In any energy eigenstate of the system we can flip all spins and obtain another state with the same energy but negative m_s – this is the \mathbb{Z}_2 symmetry mentioned in section 2.3. Hence, $\langle m_s \rangle$, which is essentially $\langle S_i^z \rangle$, vanishes at any temperature in finite systems when the complete average is taken correctly.

Unfortunately SSE simulations do not give the correct result here, see Figure 4.1(a). However, that also happens in a real physical system: at low temperatures neighboring spins anti-align which leads to a nonzero value of m_s ; the sign of m_s is chosen spontaneously to point either up or down. Statistically we observe a second order phase transition where the distribution of m_s changes continuously from a peak centered around 0 to two peaks at $\pm |m_s|$. A real physical system will be in one of these two peaks; one says the system spontaneously breaks the \mathbb{Z}_2 symmetry. Alternatively, one might say here that the system breaks the translation invariance as it favors the spin correlations at a wave vector

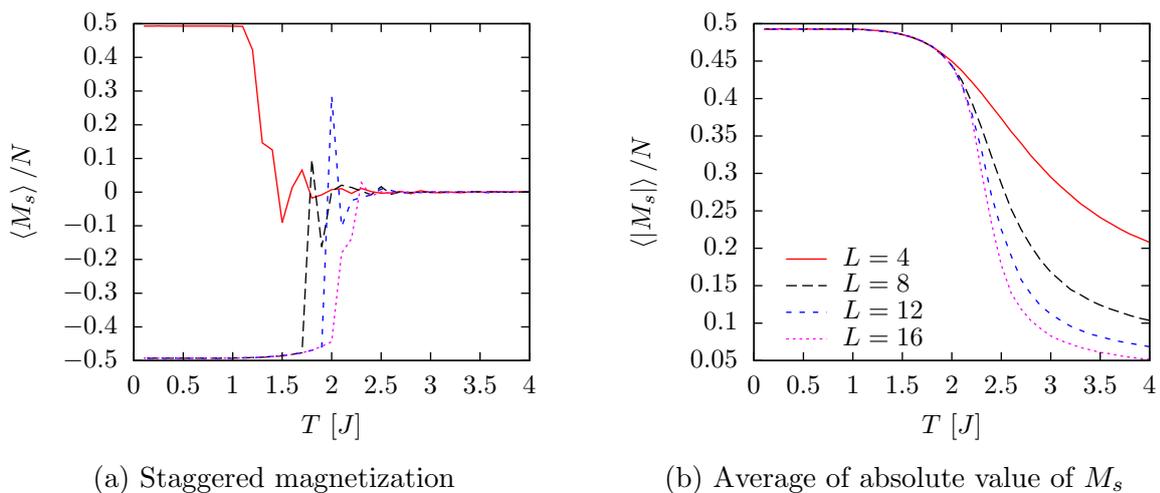


Figure 4.1: Staggered Magnetization per site on a 2D square lattice with $N = L \times L$ sites, PBC, $\Delta = 4$.

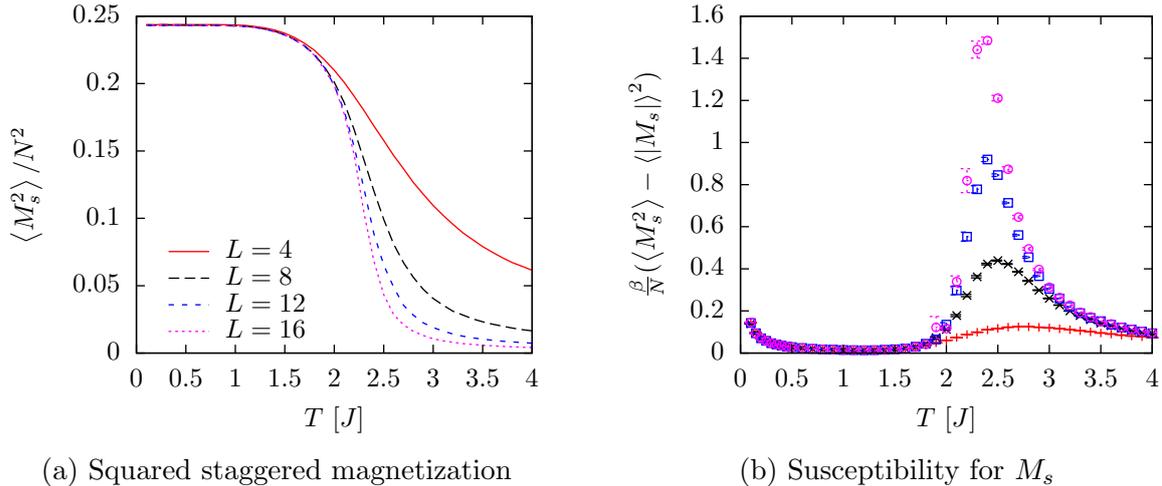


Figure 4.2: Squared staggered magnetization and corresponding susceptibility on a 2D square lattice, PBC, $\Delta = 4$.

$Q = (\pi, \pi)$ – the reason for the choice of ϕ_i in eq. (4.1); a shift of one lattice spacing a (which is set to unity here) will flip the sign of m_s . At low temperatures there are thus two regions with high weight in the phase space or the configuration space of SSE, respectively. At some point the system can no longer change the sign of m_s since this involves intermediate states with high energy. The necessary intermediate energy increases with L , thus the larger systems in Figure 4.1(a) get stuck in local energy minima earlier while the system is cooled down¹. However, it is obvious in the mechanism of SSE that flipping all spins will not change the weight of the configuration, thus it is not necessary to add such a mechanism in practice. Although our strictly speaking simulations do not fulfill ergodicity, we can thus expect correct results for quantities which are invariant under that symmetry, e.g., $\langle |m_s| \rangle$ shown in Figure 4.1(b) and $\langle m_s^2 \rangle$ in Figure 4.2(a).

At low temperatures $\langle |M_s| \rangle$ takes almost its maximum value of $\frac{1}{2}$ per spin, which corresponds to a complete anti-alignment of all spins as in the classical Ising model. This shows that we have only quite few quantum fluctuations in the ground state. With increasing system size a sharp decay develops around the critical temperature T_c . In the thermodynamic limit both $\langle m_s^2 \rangle$ and $\langle |m_s| \rangle$ will vanish in the disordered phase; fluctuations around the mean become negligible for $N \rightarrow \infty$. The order parameter then takes the form of a power law in the reduced temperature τ ,

$$\lim_{N \rightarrow \infty} \langle |m| \rangle \propto \begin{cases} |\tau|^\beta & \text{for } \tau < 0, \\ 0 & \text{for } \tau > 0, \end{cases} \quad \text{with } \tau \equiv \frac{T - T_c}{T_c}. \quad (4.2)$$

Unfortunately, the convention for the exponent is β , but it should be clear from the context whether β refers to that exponent or the inverse temperature. For $\tau < 0$ we are in the ordered phase, $\tau > 0$ corresponds to the disordered phase.

¹ It is possible to overcome this problem in Monte Carlo simulations with parallel tempering [69], yet this method has not been implemented for this thesis due to time restrictions. This is helpful, e.g., for first order phase transitions or spin glasses, where non-equivalent separated regions with high weight appear in phase space.

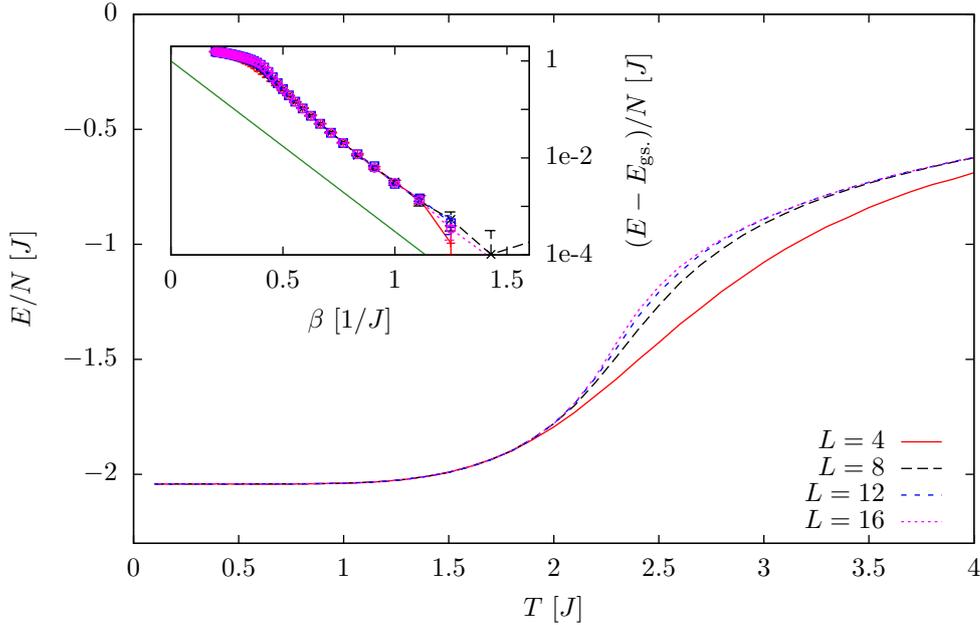


Figure 4.3: Energy per site on a 2D square lattice, PBC, $\Delta = 4$. The ground state energy is $E_{\text{gs.}} = -2.0417 \pm 0.0001$, independent of L within error bars. The inset shows $E - E_{\text{gs.}}$ versus β on a logarithmic y -axis revealing a gap. For $\beta \gtrsim 1.5$ the error bars grow larger than the values $E - E_{\text{gs.}}$. The green line is $\exp(-\beta 8.1J)$ for comparison.

The corresponding susceptibility $\chi = \frac{\beta}{N} (\langle M_s^2 \rangle - \langle |M_s| \rangle^2)$ for the staggered magnetization is shown in Figure 4.2(b). Note the different normalization with N here. The susceptibility χ diverges at the critical temperature. With increasing L the position of the maxima shifts, approaching T_c from above. The reason will be explained in section 4.3. However, away from T_c , χ takes finite values in the limit $N \rightarrow \infty$. Since $M_s \propto N$ this shows that the fluctuations around the mean value of M_s will indeed vanish in the thermodynamic limit.

The development of singularities – and thus a strong finite size dependence – is common for many quantities near a phase transition, although it is usually not as strong as for the order parameter. The energy E/N per site depicted in Figure 4.3 does not diverge, but shows an infinite slope at T_c . At high temperatures the energy approaches 0. This is clear from the following consideration: at infinite temperature each state has the same weight, thus the energy is approximately $\langle E \rangle \approx \sum_{\{s_1^z, s_2^z, \dots\}} \langle S_1^z, S_2^z, \dots | \mathcal{H} | S_1^z, S_2^z, \dots \rangle$, i.e., the sum over the *classical* Ising energies – the off-diagonal part drops. The spectrum of the Ising model is symmetric: changing the spin on each second site (on a bipartite lattice) flips the sign of the interaction. Thus, $E \rightarrow 0$ for $T \rightarrow \infty$.

Regarding the low temperature behavior, the plot further shows that we are in a gapped phase: At the lowest temperatures considered here, $T = 0.15J, 0.2J$ the ground state energy per bond is $E/N_b = -1.0208(2)J$, independent of both T and L within the error bars, which shows that we are already in the ground state. Note how close this value is to $E/N_b = -1$ for the classical Ising model neglecting S^x and S^y . The inset of Figure 4.3 shows the same data points after subtraction of the ground state energy and versus $\beta = \frac{1}{T}$.

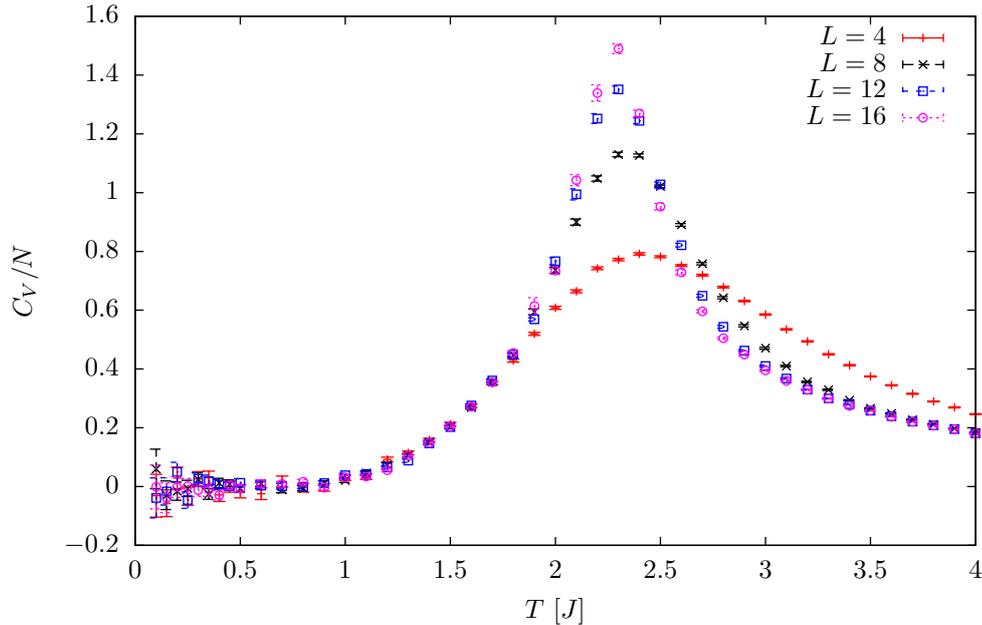


Figure 4.4: Specific heat on a 2D square lattice, PBC, $\Delta = 4$.

For large β the data points follow a straight line on the logarithmic y -axis until the relative error grows too large – the behavior one expects for an energy gap E_{gap} and $c \in \mathbb{N}$ excited states with energy $E_{\text{gs.}} + E_{\text{gap}}$:

$$\langle E - E_{\text{gs.}} \rangle = \frac{cE_{\text{gap}}e^{-\beta E_{\text{gap}}} + \dots}{2 + ce^{-\beta E_{\text{gap}}} + \dots} \approx \text{const} \cdot e^{-\beta E_{\text{gap}}}. \quad (4.3)$$

The 2 stems from the ground state which is two fold degenerate. From the slope we can estimate $E_{\text{gap}} \approx 8J$ for $\Delta = 4$. Indeed, the exact diagonalization for 4×4 sites gives two nearly degenerate ($\Delta E = 9.21 \times 10^{-5}$) ground states with energy -3.26685 and 16 excited states in the range $E = -25.085$ to $E = -24.584$ in the sector with total $S^z = 1$ (and another 16 such states with total $S^z = -1$ due to symmetry).

The derivative of the energy is given by the specific heat $C_V = \frac{\partial E}{\partial T}$ depicted in Figure 4.4. In general, the specific heat diverges as a power law $C_V/N \propto |\tau|^\alpha$, but in the universality class of the 2D Ising model the exponent for the specific heat is $\alpha = 0$ in the sense that C_V/N diverges only logarithmically² [70].

4.1 Correlation Length

We can easily relate $\langle m_s^2 \rangle$ to the spatial correlations of the spins:

$$\langle m_s^2 \rangle = \frac{1}{N^2} \sum_{i,j} \phi_i \phi_j \langle S_i^z S_j^z \rangle = \frac{1}{N^2} \sum_{i,r} \underbrace{\phi_r \langle S_i^z S_{i+r}^z \rangle}_{=: C(r)} = \frac{1}{N} \sum_r C(r). \quad (4.4)$$

² As a hand waving argument one can note $\frac{\partial x^\alpha}{\partial x} \propto x^{\alpha-1}$ for $\alpha \neq 0$ and $\frac{\partial \log(x)}{\partial x} = x^{-1}$ to motivate that a power law with $\alpha = 0$ corresponds to a logarithm.

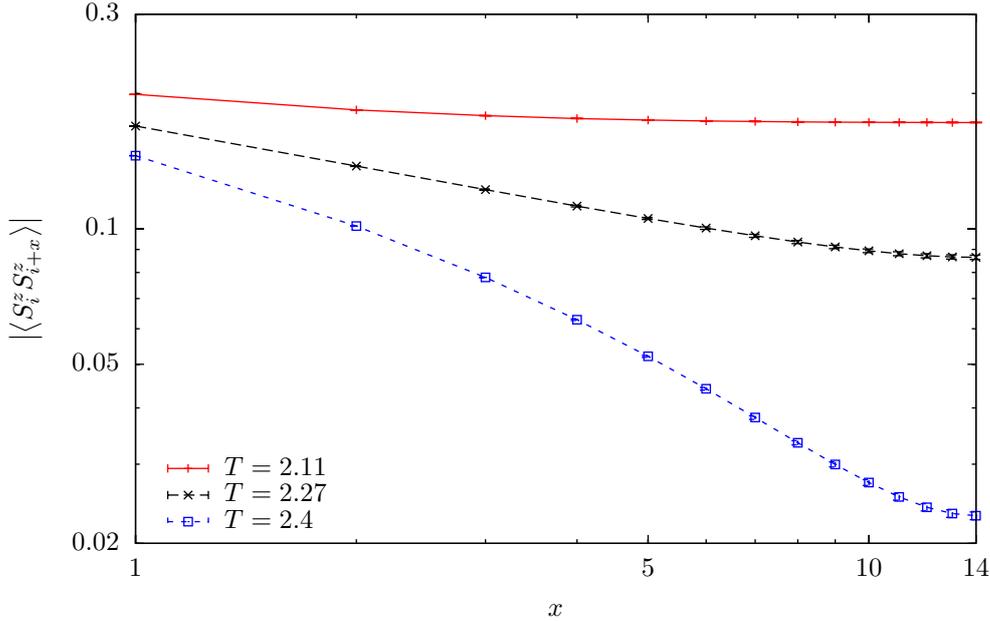


Figure 4.5: Decay of spin-spin-correlations $C(r = (x, 0))$ in x -direction for a 28×28 square lattice, calculated from eq. (3.56). Note the log-scale on both axes. Due to periodic boundary conditions the correlations increase again for $x > 14$, see text. $T = 2.27$ corresponds roughly to the critical temperature for this L . The lines serve as guides to the eye.

Here we have defined $r = j - i$ and used translation invariance for the definition of the correlation function $C(r)$. In the ordered phase we expect a decay of $C(r)$ from $C(0) = \langle (S_i^z)^2 \rangle = \frac{1}{4}$ to a nonzero long-range limit $C(r \rightarrow \infty)$. In the thermodynamic limit $N \rightarrow \infty$ only this value gives the relevant contribution in the average of eq. (4.4), and we have $C(r \rightarrow \infty) = \langle m_s^2 \rangle$.

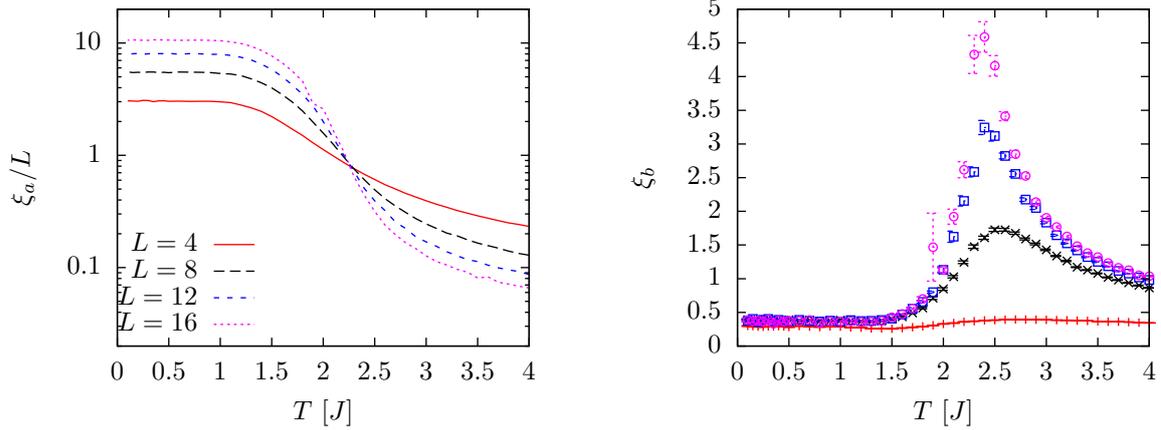
The correlations in x -direction are shown in Figure 4.5 for fixed system size $L = 28$. While $C(r)$ clearly decays to zero above the critical temperature, it reaches a finite value below T_c . The periodic boundary conditions impose the symmetry $C(r = (x, y)) = C((L + 1 - x, y)) = C((x, L + 1 - y))$ and lead to an increase of $C(r)$ when r approaches half the lattice. In general, one expects an exponential decay of the correlation function in the disordered phase, and of the connected correlations $C^*(r) := C(r) - \phi_i \langle S_i^z \rangle \langle S_{i+r}^z \rangle$ ³ in the ordered phase as well. The correlation length ξ is defined as the characteristic length scale of this decay:

$$C(r) \propto \exp\left(-\frac{|r|}{\xi}\right) \quad \text{for large } |r|. \quad (4.5)$$

Below the critical temperature we define ξ from an analogous relation for $C^*(r)$. When we approach the critical temperature $T \rightarrow T_c$ the correlation length diverges according to a power law [53]:

$$\xi \propto |\tau|^{-\nu}. \quad (4.6)$$

³ Here the average $\langle \cdot \rangle$ should be seen in the thermodynamic limit with a broken \mathbb{Z}_2 symmetry, such that $\langle S_i^z \rangle$ does not vanish.



(a) ξ_a rescaled with the lattice length L (b) Divergence of connected correlations at T_c

Figure 4.6: Correlation length from eq. (4.7) for a 2D square lattice, PBC, $\Delta = 4$.

Note that the exponent ν is the same for both sides $T > T_c$ and $T < T_c$. In the universality class of the 2D Ising model which we consider here $\nu = 1$ [70]. Right at the critical temperature the decay of $C(r)$ is no longer exponential, but also follows a power law which corresponds to a straight line in the double logarithmic scaling of Figure 4.5.

The calculation of all spatial correlations $\langle S_i^z S_{i+r}^z \rangle$ as the Fourier transform of the structure factor $S(k)$, see eq. (3.56), is naturally quite expensive. To get a rough estimate of the correlation length it is not necessary to calculate the structure factor $S(k)$ for all momenta k . With the anti-alignment of the spins $S(k)$ develops a peak at the wave vector $Q = (\pi, \pi)$ corresponding to the phase ϕ_i in eq. (4.1). We included that phase into the definition $C(r) = \phi_i \langle S_i^z S_{i+r}^z \rangle$ such that $C(r) > 0$. From (3.56) we have $S(k) = \sum_r e^{i(k-Q)\cdot r} C(r)$. The Fourier transform to $S(k)$ then encodes the long-range behavior of $C(r)$ in the momenta close to Q . For small q the peak $S(Q+q)$ should thus be well described by the Fourier transform of the exponential decay of $C(r)$, which is in 1D simply a Lorentzian function $\frac{1}{1+q^2\xi^2}$. This motivates the following two definitions for the correlation length [52]:

$$\xi_a := \frac{1}{q} \sqrt{\frac{S(Q)}{S(Q+q)} - 1}, \quad \xi_b := \frac{1}{q} \sqrt{\frac{\frac{S(Q+q)}{S(Q+2q)} - 1}{4 - \frac{S(Q+q)}{S(Q+2q)}}}. \quad (4.7)$$

Here we choose $q = (\frac{2\pi}{L}, 0)$ to be the smallest possible wave vector in a finite system.

Both quantities are shown in Figure 4.6. In the ordered phase with $C_\infty := C(r \rightarrow \infty) \neq 0$ only the connected correlations $C^*(r)$ decay exponentially. The additional constant in $C(r)$ transforms into a delta peak affecting only $S(Q)$. Thus, ξ_a diverges in the ordered phase, whereas ξ_b does not reflect the long-range order but describes the decay of $C^*(r)$ correctly. In a classical system $S(k \neq Q)$ vanishes in the ground state and ξ_a diverges when $T \rightarrow 0$. But here quantum fluctuations imply that $C(r=0) = \frac{1}{4} > C_\infty$ and thus $S(k)$ is nonzero for all temperatures. Indeed, suppose that $C(r) = C_\infty + (\frac{1}{4} - C_\infty)\delta_{r,0}$, then $S(k) = NC_\infty\delta_{k,Q} + (\frac{1}{4} - C_\infty)$. Thus, $\frac{S(Q)}{S(Q+q)} \propto N$ for low temperatures, which implies $\xi_a \propto L^2$. In contrast, at high temperatures the spins are uncorrelated, and $\xi \ll L$ is nearly independent of the system size. Thus, ξ_a/L shows a crossing as depicted in Figure 4.6(a).

The fact that all curves intersect at T_c is a hint that the exponent $\nu = 1$ since we are in the universality class of the classical 2D Ising model. Indeed, the maxima of the connected correlation length ξ_b shown in Figure 4.6(b) also diverge proportional to L .

4.2 Scaling Hypothesis

A critical point is characterized by scale invariance [53]: For the order parameter m_s in eq. (4.1) we average over all N spins of our finite system. Now consider such an average over a finite number of N sites in an infinite lattice. Scale invariance then means that the fluctuations in m_s will be independent of N – as long as N is large enough such that the finite lattice spacing does not matter. Clearly, this requires an infinite correlation length: On the one hand the correlations need to decay fast enough such that we have fluctuations on short scales, but on the other hand, slowly enough such that m_s takes nonzero values on large scales as well.

The description of the system at different length scales and scale invariance at critical points are the foundation of the renormalization group (RG) theory [71, 72]. In the following we will shortly explain the basic idea with the help of Kadanoff's spin block picture [73]. As the name suggests, one groups the spins into blocks of lateral length b in each dimension and describes the spins in each block collectively, combining the interactions of single spins at the boundaries of a block to effective interactions between the blocks. The resulting blocks then live on a lattice with lattice spacing $b \cdot a$, which is larger by a factor of b compared to the previous unit length a . After a final rescaling of all lengths with $\frac{1}{b}$, one obtains a description of the same model, but with renormalized, effective couplings. Such a procedure is called an RG transformation and is the mathematical formulation of how to view the system at a larger length scale.

We denote the free energy density with $f = -\frac{1}{\beta L^d} \ln \mathcal{Z}$. In the context of RG it is seen as a function of so-called scaling fields. In our model we have two relevant scaling fields: the reduced temperature τ and the magnetic field h which couples to the order parameter. Now consider an RG transformation: By rescaling $a \rightarrow a/b$, the correlation length is rescaled as well, $\xi \rightarrow \xi b$. We recall from eq. (4.6) that $|\tau| \xi^{\frac{1}{\nu}} = \text{const.}$ Thus, the RG transformation changes the reduced temperature $\tau \rightarrow \tau b^{\omega_\tau}$ with an exponent $\omega_\tau \equiv \frac{1}{\nu}$. Similar, the other scaling fields are changed as well, but with some other exponent; let us say $h \rightarrow h b^{\omega_h}$. Coupling fields with negative exponents ω are irrelevant: their values just renormalize to 0 after some number of RG transformations, and thus do not lead to singular behavior of f . However, from the non-rescaled point of view each spin block contains b^d spins, which arises as a prefactor to the *density* f . By definition, the value of the free energy density does not change under a rescaling for scale invariant systems, i.e., near a phase transition. Thus, we have the following highly remarkable result:

$$b^d f(\tau, h) = f(\tau b^{\omega_\tau}, h b^{\omega_h}). \quad (4.8)$$

With the choice $b = |\tau|^{-\nu} \Rightarrow \tau b^{\omega_\tau} = \pm 1$, we see that f can be written in terms of a scaling function Φ_f :

$$f(\tau, h) = |\tau|^{d\nu} \Phi_f \left(\frac{h}{|\tau|^\Delta} \right) \quad \text{with } \Delta := \frac{\omega_h}{\omega_\tau}. \quad (4.9)$$

This form was actually proposed as a hypothesis by Widom in 1965 [74] from phenomenological arguments. Many thermodynamic observables can be obtained from f through derivatives and thus obey very similar scaling laws, but with other scaling functions. For example the singular part of the specific heat is essentially given by two derivatives with respect to τ :

$$\begin{aligned} E(\tau, h) &\propto -\frac{\partial f(\tau, h)}{\partial \tau} = d\nu |\tau|^{d\nu-1} \Phi_f \left(\frac{h}{|\tau|^\Delta} \right) - h\Delta |\tau|^{d\nu-1-\Delta} \Phi_f \left(\frac{h}{|\tau|^\Delta} \right) \\ &\equiv |\tau|^{d\nu-1} \Phi_E \left(\frac{h}{|\tau|^\Delta} \right), \end{aligned} \quad (4.10)$$

$$\begin{aligned} C_V(\tau, h) &\propto -\frac{\partial^2 f(\tau, h)}{\partial \tau^2} = (d\nu - 1) |\tau|^{d\nu-2} \Phi_E \left(\frac{h}{|\tau|^\Delta} \right) - h\Delta \tau^{d\nu-2-\Delta} \Phi_E \left(\frac{h}{|\tau|^\Delta} \right) \\ &\equiv |\tau|^{d\nu-2} \Phi_{C_V} \left(\frac{h}{|\tau|^\Delta} \right). \end{aligned} \quad (4.11)$$

In the limit $h \rightarrow 0$ the argument of the scaling function will be 0 such that the scaling function gives a constant prefactor only. Above, we proposed a scaling with the exponent α for C_V (p. 48); thus, we see that not all exponents are independent. In fact, there are plenty of relations between them [57]; here, we arrive at the hyperscaling relation

$$d\nu = 2 - \alpha. \quad (4.12)$$

4.3 Finite Size Scaling

We will now review the influence of the finite lattice length L , an analysis which dates back to Fisher [75]. Formally, the thermodynamic average $\langle A \rangle = \text{Tr} \{ e^{-\beta \mathcal{H}} A \} / \mathcal{Z}$ of any observable A is an analytic function without any singularities for finite systems $N < \infty$. Thus, the thermodynamic limit $N \rightarrow \infty$ is essential to observe the singularities associated with a phase transition.

Handwavingly, we can argue that a finite system with $L < \infty$ is effectively ordered when the correlation length $\xi \propto |\tau|^{-\nu}$ is comparable to the length L . This happens already at some temperature $T_c(L)$ above the (real) infinite-length critical temperature T_c :

$$(T_c(L) - T_c) \propto \xi^{-\frac{1}{\nu}} \stackrel{!}{\approx} L^{-\frac{1}{\nu}} \quad \Rightarrow \quad T_c(L) \approx T_c + \text{const} \cdot L^{-\frac{1}{\nu}}, \quad (4.13)$$

which explains the shift of the maxima for divergent quantities, e.g., the susceptibility in Figure 4.2(b). The scaling theory of the previous section allows a more rigorous study which we will discuss in the following along the lines of Ref. [53].

For simplicity we will now set the magnetic field $h = 0$. When we consider a system of finite length L , we can simply see $\frac{1}{L}$ as a further, relevant scaling variable: clearly it will rescale as $\frac{1}{L} \rightarrow \frac{b}{L}$ under an RG transformation. Thus, the corresponding exponent is simply 1. Let us now consider a general thermodynamic variable $A(\tau, \frac{1}{L})$ generated from the free energy, e.g., the susceptibility. Similar to the scaling of the free energy in eq. (4.8), we expect the following behavior under an RG transformation [53]:

$$A_L(\tau) \equiv A \left(\tau, \frac{1}{L} \right) = b^{\omega_A} A \left(\tau b^{\frac{1}{\nu}}, \frac{b}{L} \right). \quad (4.14)$$

Here ω_A is the anomalous dimension of A and will later be related to the critical exponent of A . Again we can choose $b = |\tau|^{-\nu}$ such that we can write A_L in terms of just a single scaling function,

$$A_L(\tau) = |\tau|^{-\nu\omega_A} \tilde{\Phi}_A \left(\frac{|\tau|^{-\nu}}{L} \right) = L^{\omega_A} \left(\frac{|\tau|^{-\nu}}{L} \right)^{\omega_A} \tilde{\Phi}_A \left(\frac{|\tau|^{-\nu}}{L} \right) \equiv L^{\omega_A} \Phi_A \left(\tau L^{\frac{1}{\nu}} \right). \quad (4.15)$$

In the last step we absorbed some constants in the scaling function. Let us assume that A has a singularity with some exponent κ at the critical point, i.e., it behaves as $A_\infty(\tau) \approx c_\pm |\tau|^{-\kappa}$ in the limit $L \rightarrow \infty$, where we have possibly different prefactors c_+ and c_- corresponding to an approach to T_c from above ($\tau > 0$) or from below ($\tau < 0$), respectively. Such a behavior is only possible if the scaling function shows the asymptotic behavior $\Phi_A(x) \approx c_\pm |x|^{-\kappa}$ in the limit $x \rightarrow \infty$. Further, we have to identify $\omega_A = \frac{\kappa}{\nu}$ to compensate the scaling with L for a well-behaved thermodynamic limit. Thus, the finite size scaling of A will have the following form:

$$A_L(\tau) = L^{\frac{\kappa}{\nu}} \Phi_A \left(\tau L^{\frac{1}{\nu}} \right) \equiv L^{\frac{\kappa}{\nu}} \hat{\Phi}_A \left(\frac{L}{\xi(\tau)} \right). \quad (4.16)$$

This remarkable result has several consequences which we will give in the following.

First of all, eq. (4.16) shows that the finite size scaling exactly at the critical temperature, $T = T_c \Leftrightarrow \tau = 0$, is given by $A_L(\tau = 0) \propto L^{\frac{\kappa}{\nu}}$, and thus related to the singularity with respect to τ in the infinite size system.

Secondly, we understand the argument given in the beginning of this section more rigorously: For definiteness consider the divergent maxima of the susceptibility as shown in Figure 4.2(b). The scaling function $\Phi_\chi(x)$ may have a maximum at some $x_0 > 0$. If we denote the temperature of the maximum⁴ with $T_c(L)$, we can follow its movement with L which additionally gives the constant prefactor in eq. (4.13):

$$\frac{T_c(L) - T_c}{T_c} L^{\frac{1}{\nu}} = x_0 \quad \Rightarrow \quad T_c(L) = T_c + T_c x_0 L^{-\frac{1}{\nu}}. \quad (4.17)$$

The value of χ diverges at the maximum in exactly the same form as it does at $\tau = 0$, but with another prefactor $\Phi_a(x_0)$ (instead of $\Phi_a(0)$):

$$A_L \left(\tau = \frac{T_c(L) - T_c}{T_c} \right) = L^{\frac{\kappa}{\nu}} \Phi_a(x_0). \quad (4.18)$$

The result of eq. (4.17) allows a determination of T_c from the position of the maxima $T_c(L)$ in finite size data. Such an analysis is shown in Figure 4.7. Since Monte Carlo data have statistical noise, we did not just use the temperature of the data point with the maximal value for $T_c(L)$; this would also require a very fine grid of temperatures. Instead, a least squares fit of a polynomial of degree three has been performed, with the position of the maximum $T_c(L)$ as a fit parameter. For each L , only data points in the vicinity of the corresponding maximum have been used for the fit: As the width of the peaks decreases with L , we used a window size $\propto \frac{1}{L}$, based on the exact value $\nu = 1$ of the 2D

⁴ Note that the prefactor might depend on the variable x_0 in general, thus it is a bit misleading to talk about *the* critical temperature $T_c(L)$ at some length L .

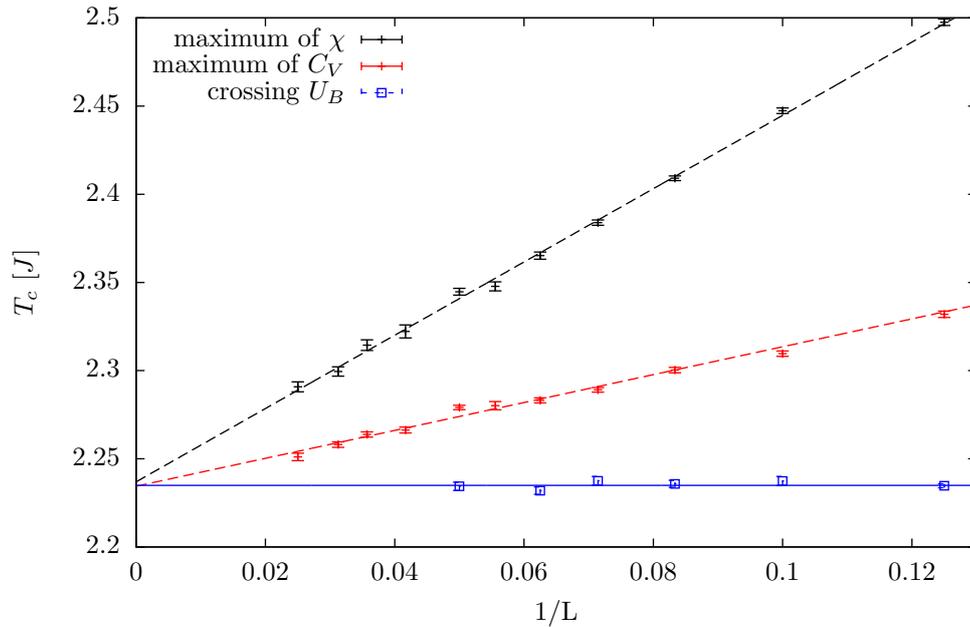


Figure 4.7: Finite size scaling (2D square lattice, $\Delta = 4$): position of maxima $T_c(L)$ for the susceptibility χ (black) and specific heat C_V (red) versus inverse length; see main text for an explanation how the data points were obtained. The data points follow a straight line which confirms the value of the exponent $\nu = 1$ expected for the 2D Ising universality class [57]. The dashed lines indicate least squares fits according to eq. (4.17), where the exact value for $\nu = 1$ has been used. From these fits, we obtain the estimates $T_c = 2.234(2)$ from C_V and $T_c = 2.237(2)$ from χ . In advance to section 4.3.1, the temperatures at which the curves of the Binder Cumulant U_B of sizes L and $2L$ cross are shown, too, see Figure 4.9. The solid blue line shows a fit of these crossings to a constant value, $T_c = 2.2349(7)$.

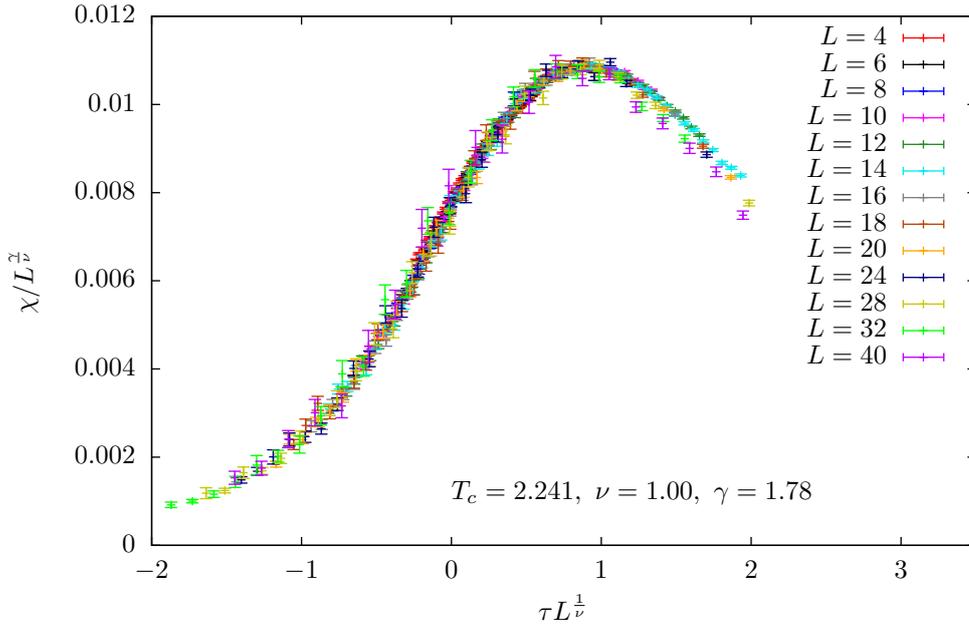


Figure 4.8: Data collapse of the staggered susceptibility χ on a 2D square lattice, $\Delta = 4$. The best collapse is obtained with the values $T_c = 2.2341(4)$, $\nu = 1.00(4)$ and $\gamma = 1.78(9)$, which compares well to $\gamma = 1.75$ for the 2D Ising model [57].

Ising universality class [57]. The obtained values for $T_c(L)$ are plotted in Figure 4.7. The data points $T_c(L)$ follow a straight line in the plot versus $\frac{1}{L}$, which confirms the exponent $\nu = 1$ (see eq. (4.17)) expected for the 2D Ising universality class. Although this method gives rough estimates for the critical temperature and the exponent ν , it is not the best possibility to determine these: especially if the exact exponent is not yet known, a fit as given in Figure 4.7 is quite hard and requires large system sizes for high accuracy.

As a final consequence of eq. (4.16), the finite size scaling gives rise to a so-called ‘data collapse’: With the definition $A_L^{\text{resc}}(\tau) := A_L(\tau)L^{-\frac{\kappa}{\nu}}$, equation (4.16) trivially implies

$$A_{L_1}^{\text{resc}}(\tau_1) = A_{L_1}(\tau_1)L_1^{-\frac{\kappa}{\nu}} = \Phi_A\left(\tau_1 L_1^{\frac{1}{\nu}}\right) = A_{L_2}^{\text{resc}}(\tau_2) \quad \text{if } \tau_1 L_1^{\frac{1}{\nu}} = \tau_2 L_2^{\frac{1}{\nu}}. \quad (4.19)$$

Given the critical temperature T_c and the exponents κ and ν , we can rescale the data points $A_L(T)$ obtained from finite size Monte Carlo simulations accordingly. Thus, in a plot versus $x := \tau L^{\frac{1}{\nu}}$ on the x -axis, the rescaled data points $A_L^{\text{resc}}(\tau)$ collapse to a single master curve; from eq. (4.19) we see that the master curve is just the scaling function $\Phi_A(x)$ with $x = \tau L^{\frac{1}{\nu}}$. Such a data collapse is shown in Figure 4.8 for the susceptibility χ , for which the above exponent κ is by convention denoted as γ .

Reversed, this scaling procedure provides another possibility to determine the exact values for T_c and the exponents: We can just guess some values and check whether the curves collapse. However, this requires a measure for the goodness of a collapse. The scaling function $\Phi_A(x)$ is a well behaved, analytic function. Thus, it should be possible to fit a polynomial of low degree (≈ 8) to the data points in the vicinity of $x = 0$. Throughout this work, we exclude data with length $L < 8$ from these fits since they have too strong finite size effects; yet we include them into the plots showing the collapse. One

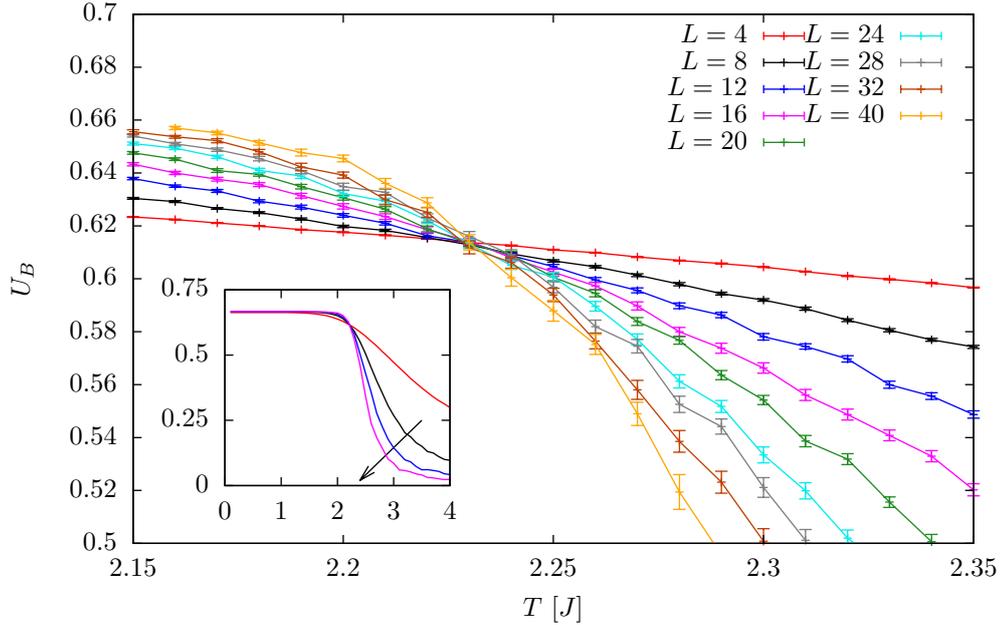


Figure 4.9: Binder cumulant U_B on a 2D square lattice, PBC, $\Delta = 4$. The inset gives an overview over a wide temperature range for $L = 4, 8, 12, 16$; the arrow indicates increasing L . The main figure shows the crossing at T_c in more detail.

can then use the fit quality – i.e., the χ^2 per degree of freedom for a least squares fit – as such a measure⁵ [52]. We have used the NLOpt library [77] to minimize this measure under variation of T_c and the critical exponents. In fact, we have used such a schema to obtain the exponent γ in Figure 4.8.

Unfortunately, most of our estimates are not in agreement with the exponents of the 2D universality class within the error bars. This might be related to the fact that we only can use comparatively small systems – for the classical Ising model, systems with $L \approx 10^3$ can be reached quite easily. For higher accuracy, it might help to include sub-leading corrections to the scaling ansatz eq. (4.16). Beach *et al.* [78] proposed – in consideration of first corrections from irrelevant scaling fields – the following scaling form:

$$A_L(\tau) = L^{\frac{\kappa}{\nu}} (1 + cL^{-\omega}) \Phi_A \left(\tau L^{\frac{1}{\nu}} - dL^{-\frac{\phi}{\nu}} \right). \quad (4.20)$$

Here, c, d, ϕ , and ω are further fit parameters and the exponents ϕ, ω are positive and usually at most 1. However, this has not been done for this thesis due to time restrictions.

4.3.1 Binder Cumulant

In 1981, Binder [79] studied the distribution of the order parameter on finite spin blocks for the Ising model. In particular, he studied a quantity which is nowadays known as Binder cumulant and reads, in our case with the order parameter given by M_s ,

$$U_B = 1 - \frac{\langle M_s^4 \rangle}{3 \langle M_s^2 \rangle^2}. \quad (4.21)$$

⁵ A similar measure which instead approximates Φ_A only in the local vicinity of data points is proposed in Ref. [76].

order k	moment	central moment ($\mu = 0$)
1	μ	0
2	$\mu^2 + \sigma^2$	σ^2
3	$\mu^3 + 3\mu\sigma^2$	0
4	$\mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$	$3\sigma^4$

Table 4.1: First four moments $\langle x^k \rangle_{\mu,\sigma}$ and central moments $\langle (x - \langle x \rangle)^k \rangle_{\mu,\sigma}$ for a normal distribution $\langle g(x) \rangle_{\mu,\sigma} = \int f(x, \mu, \sigma)g(x)dx$ with $f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right)$.

The results for different system sizes are shown in Figure 4.9. It turns out that this quantity is well suited to determine the critical temperature from finite size data: it approaches a step function with a jump at T_c in the thermodynamic limit and admits a clear crossing of finite size curves at T_c . In the following, we try to get some intuition for the behavior of the Binder cumulant.

In the paramagnetic phase at high temperatures, the spins are not correlated. M_s is then the sum of independent, identical random variables and is thus (according to the central limit theorem) normally distributed, centered around 0 as we have no magnetic field. From Table 4.1 we can read off that in this case $\langle M_s^4 \rangle = 3\langle M_s^2 \rangle^2$, thus $U_B \rightarrow 0$ at high temperatures. The central limit theorem requires a large number of independent variables; thus the value of U_B decreases with increasing L . In turn at fixed L , the correlation length increases when we approach T_c from above; we then effectively have less independent variables and U_B increases.

As we have discussed at the beginning of this chapter, the distribution of M_s shifts from a central peak to two peaks at $\pm M_s$ during the phase transition. Let us approximate these peaks with two normal distributions as well, centered at $\mu \approx \pm \langle |M_s| \rangle \propto N$. The variance is then given by

$$\sigma^2 = \langle M_s^2 \rangle - \langle |M_s| \rangle^2 = \frac{N}{\beta} \chi. \quad (4.22)$$

We have seen in Figure 4.2(b) that the susceptibility χ is well behaved in the thermodynamic limit if we are far enough from the critical temperature such that $\xi \ll L$. Thus, $\mu^2 \gg \sigma^2$ by a factor of N in the thermodynamic limit, which justifies the approximation with normal distributions and $\mu = \langle |M_s| \rangle$ a posteriori. From Table 4.1 we can estimate the leading order behavior of the Binder cumulant:

$$\begin{aligned} U_B &= \frac{3\langle M_s^2 \rangle^2 - \langle M_s^4 \rangle}{\langle M_s^2 \rangle^2} = \frac{3(\mu^2 + \sigma^2)^2 - (\mu^4 + 6\mu^2\sigma^2 + 3\sigma^4)}{3(\mu^2 + \sigma^2)^2} = \frac{2\mu^4}{3\mu^4 + 2\mu^2\sigma^2 + \sigma^4} \\ &= \frac{2}{3} \left(1 - \frac{2\chi}{N \langle |m_s| \rangle^2 \beta} + \mathcal{O}\left(\frac{1}{N^2}\right) \right). \end{aligned} \quad (4.23)$$

This result is valid in the ordered phase for $\xi \ll L$.

Binder noted in his paper [79] that the approximation with two normal distributions describes the peaks, but not the wings of the distribution: domain walls can form at which the order parameter changes its sign such that M_s is nearly 0 in the total considered block.

The energy cost of such domain walls grows only with the surface of the domains, not with their volume, and leads to a correction of the distribution function in M_s , increasing the weight at small M_s . Nevertheless, this approximation and eq. (4.23) give us the correct behavior of the Binder cumulant U_B in the thermodynamic limit: it approximates a step function with a jump from $\frac{2}{3}$ to 0 at the phase transition, i.e., at T_c . Assuming a scaling similar to eq. (4.16) for the distribution function itself, Binder further derived the following scaling form for U_B near the phase transition [79]:

$$U_B = 1 - C_0 \frac{f_4\left(\frac{\xi}{L}\right)}{3\left(f_2\left(\frac{\xi}{L}\right)\right)^2}. \quad (4.24)$$

Here C_0 is a universal constant, and $f_k\left(\frac{\xi}{L}\right)$ are the scaling functions of the k -th moment $\langle M_s^k \rangle$ (and are obviously related to the scaling of the distribution function). Exactly at T_c , the Binder cumulant is thus given by another universal constant $U_B(\tau = 0) = 1 - \frac{C_0}{3} \frac{f_4(\infty)}{(f_2(\infty))^2}$. This explains the clear crossing at T_c , shown in Figure 4.9. We have extracted the crossing between curves of size L and $2L$ and used it in Figure 4.7 as a more precise reference value for T_c . However, the highest accuracy can be gained from a data collapse of the Binder cumulant.

Remarkably, the scaling of the prefactors cancels for the Binder cumulant: we can rewrite the result (4.24) in analogy to eq. (4.16) as

$$U_B = \Phi_{U_B}\left(L\tau^{\frac{1}{\nu}}\right). \quad (4.25)$$

This scaling form allows a data collapse as discussed above with no more than two free parameters: T_c and ν ; one needs to rescale the x -axis only. In addition, the influence of ν and T_c on the curves of finite L is contrary in the following sense: ν rotates the curves such that they all have the same slope at $\tau = 0$, whereas an imprecise value of T_c leads to a vertical displacement near $\tau = 0$. Thus, a data collapse of the Binder cumulant is suited very well to obtain both ν and T_c to high precision. The result of such a fit is given in Figure 4.10.

4.4 1D-2D-Crossover

In the following we will use the tools described above to study the crossover between one and two dimensions. As discussed in section 2.5 and depicted in Figure 2.2(a), we will therefore vary the coupling strength J' in y -direction from $J' = 0$ to $J' = 1$ on a two dimensional square lattice. The coupling J within the chains is kept fix and serves as unit of energies.

While the isotropic Heisenberg model shows long-range order at least at $T = 0$ in two dimensions, it does not in one dimension [9]. As we have seen above, the anisotropic coupling $\Delta = 4$ of the XXZ -model (towards the Ising model) enhances the ordering to a finite temperature transition in two dimensions, which belongs to the universality class of the 2D Ising model. However, it is well known since its introduction in 1925 [3] that the one dimensional Ising model does not show long-range order at finite temperatures. Thus, we expect that the critical temperature T_c , at which the ordering sets in, vanishes for $J' \rightarrow 0$.

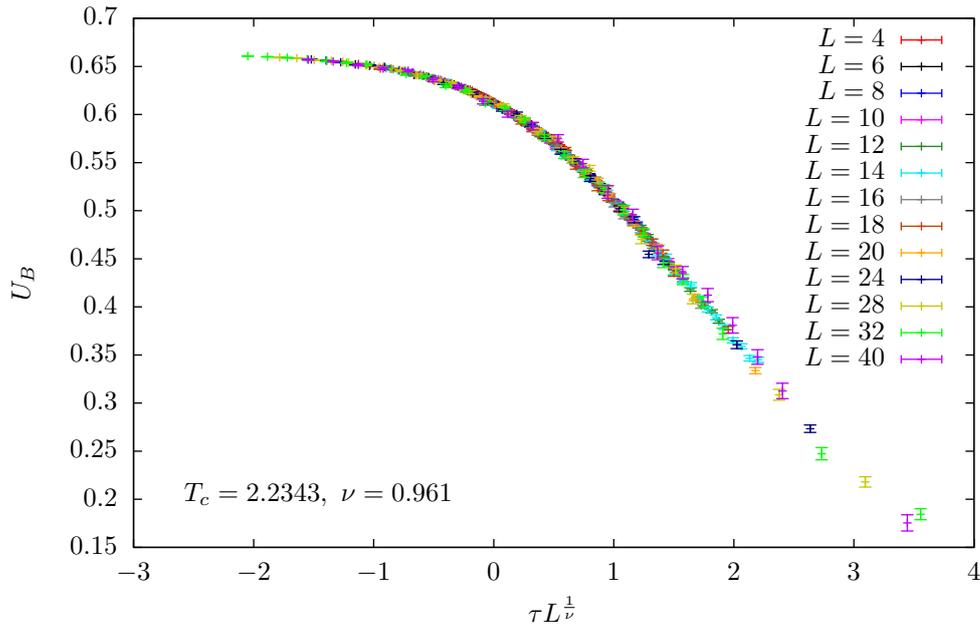


Figure 4.10: Data collapse of the Binder cumulant with best values for T_c and ν ; 2D square lattice, $\Delta = 4$. Note that only the x -axis is rescaled according to eq. (4.25). The obtained values are $T_c = 2.2343(2)$ and $\nu = 0.961(5)$.

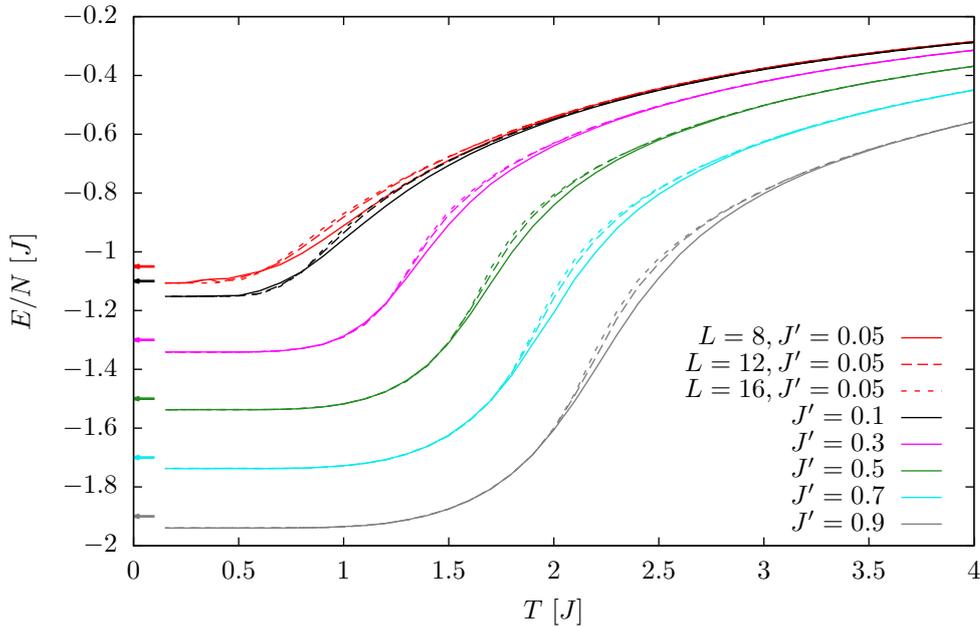


Figure 4.11: Energy per site of coupled chains on a 2D square lattice ($L \times L$), $\Delta = 4$. The color indicates the value of the inter-chain coupling J' , from top to bottom $J' \in \{0.05, 0.1, 0.3, 0.5, 0.7, 0.9\}$. For each J' , different line types indicate the length $L \in \{8, 12, 16\}$, revealing a stronger finite size dependence near the critical temperature. The arrows at the y -axis indicate the classical ground state energies for the different J' (with the same color key as the lines).

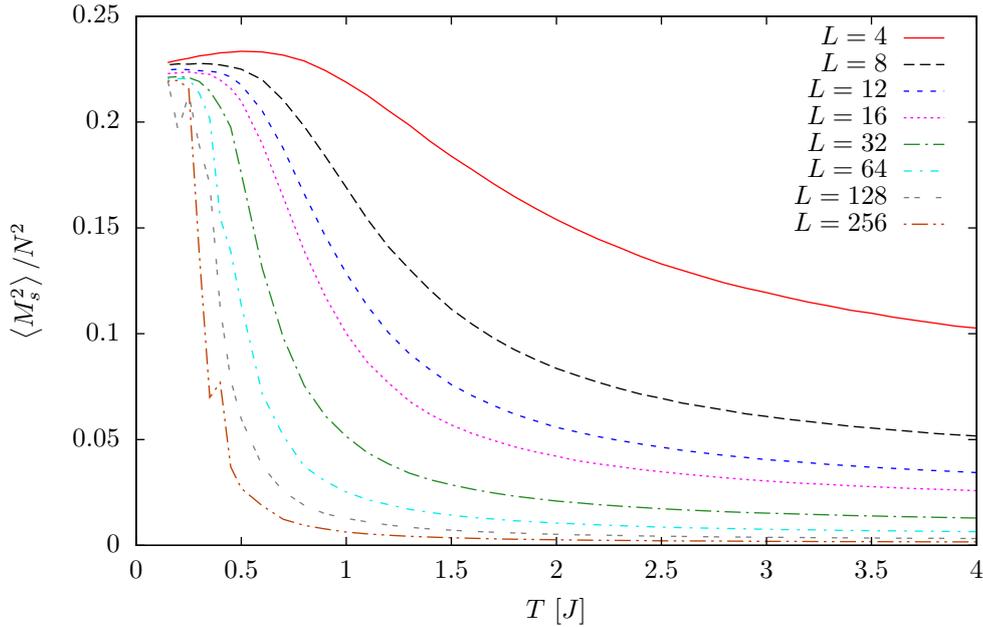


Figure 4.12: Squared staggered magnetization $\langle m_s^2 \rangle$ for a one dimensional chain with PBC, $\Delta = 4$. In the thermodynamic limit $L \rightarrow \infty$ there is no long-range order at any $T > 0$ [52]. The ‘kinks’ at low temperatures are unphysical and arise from statistical errors and a rough grid $\Delta T = 0.1J$ used to create the curves.

It turns out that we can understand the dimensional crossover at $\Delta = 4$ to a large extent from purely classical considerations, i.e., when neglecting the off-diagonal part in the Hamiltonian which leads to quantum fluctuations. The energy per site is depicted in Figure 4.11 for different couplings J' . With the same argument as given above for the isotropic 2D square lattice, the energy approaches 0 at high temperatures for any J' . The absolute value of the ground state energy per site increases with J' ; this is clear from the classical picture, where the energy per site is simply $E_{\text{gs.}}/N = -\frac{\Delta}{4}(J + J')$. Similarly, we understand the increase of the gap with J' , in the classical picture simply given by $E_{\text{gap}} = \Delta(J + J')$.

Note that the gap still persists in the one dimensional chain, i.e., for $J' = 0$. Clearly, at least the classical ground state shows long-range order. One might thus wonder how the order can be destroyed by thermal excitations if there is a gap. Although the probability for a single excitation is almost negligible for small T due to the gap, it is the huge number of excitations that destroys the long-range order at finite temperatures. Flipping a group of m neighboring spins has the same excitation energy as flipping a single spin. Thus, there are $L(L - 1)$ states with an excitation energy of $J\Delta$ in the 1D Ising chain of length L . While excitations with just a few flipped spins still contribute to long-range order, the states where m is a significant fraction of the total length do not. Although we have correlations on short length scales, the long-range order is thus destroyed in the thermodynamic limit at any $T > 0$. This is illustrated in Figure 4.12 with chains of length up to $L = 256$: for small length we see the tendency to order, but the temperature at which the ordering sets in decreases towards 0 with increasing L .

Each finite – but arbitrarily small – value $J' > 0$ leads to a fundamental change in that

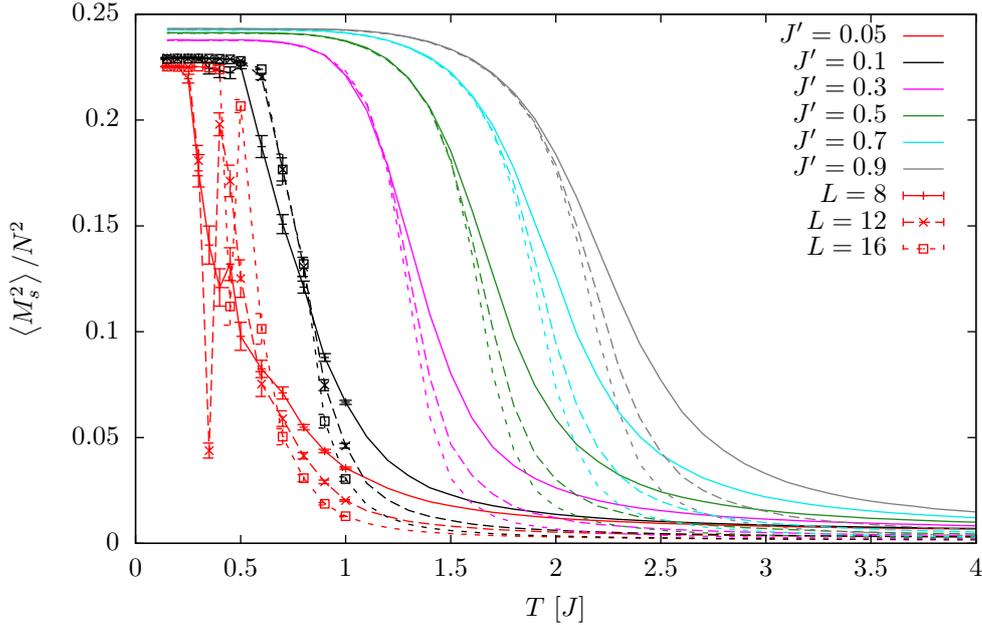


Figure 4.13: Squared staggered magnetization of coupled chains on a 2D square lattice, $\Delta = 4$. For small J' the data seems unreliable near the phase transition; an explanation is given in the main text. The error bars of single simulations with 10^6 sweeps are shown for $J' \in \{0.05, 0.1\}$ at low temperatures, but they heavily underestimate the real error.

picture: flipping a block of m neighboring spins in a chain now costs an energy of $m\Delta J'$ due to the inter-chain couplings in addition to the intra-chain contribution ΔJ . Thus, excitations where a significant number of spins is flipped are suppressed stronger. This consideration suggests that the finite temperature transition, which we observed in 2D, persists to any arbitrarily small $J' > 0$.

The performed SSE simulations seem to confirm this result. The staggered magnetization $\langle m_s^2 \rangle$ is shown in Figure 4.13 for different values of J' . Indeed, we see that the ordering sets in at temperatures $T_c(J')$ where the energy shows strong finite size dependence (see Figure 4.11). At low J' our algorithm obviously has problems to produce correct results: the fluctuations of the data points – which are supposed to be smooth in T – become very large near the phase transition, obviously even larger than the error bars: the auto-correlation time of $\langle m_s^2 \rangle$ diverges and thus the error bars are not estimated correctly. In that sense the problem has a similar origin like the one discussed in the beginning of this chapter that we do not obtain $\langle M_s \rangle = 0$. The coupling J in x -direction is stronger than in y -direction (J'). We can read off of Figure 4.12 that chains of the length we consider for the dimensional crossover (which requires the simulation of a complete square lattice with $L \times L$ spins irrespectively of J') tend to order at $\tilde{T} \approx 0.7 \dots 1.5$. Since $T_c \rightarrow 0$ with $J' \rightarrow 0$, there is an intermediate temperature regime for finite systems where the spins are ordered within the chains, i.e., in x -direction, but they are not (anti-)aligned in the y -direction with weak couplings. During the phase transition the spins have to align in y -direction as well. However, this requires either to flip all the L spins in a chain at once or to break the alignment within the strong couplings. Due to the large value of Δ the

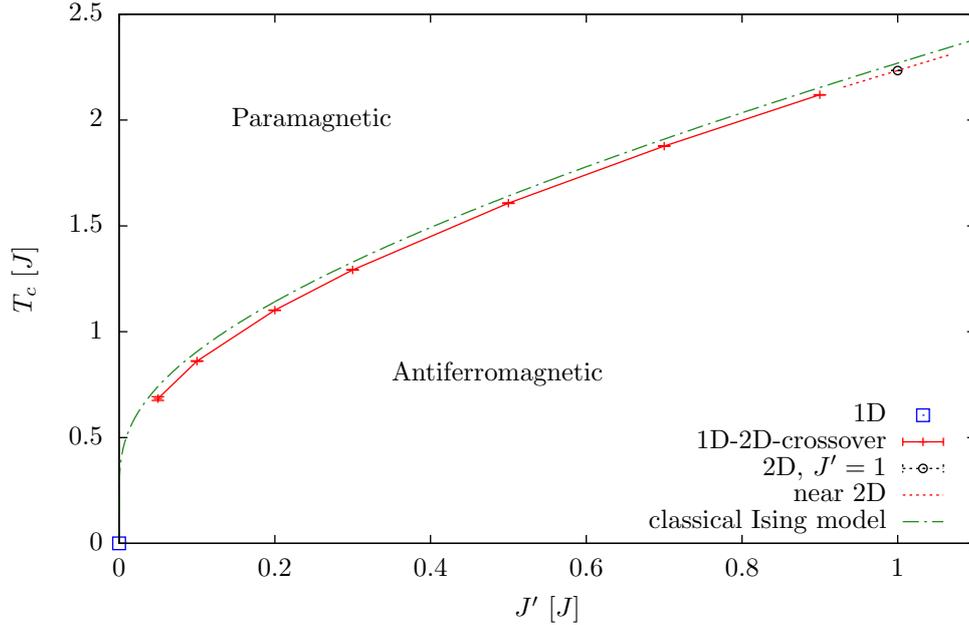


Figure 4.14: Phase diagram for the 1D-2D-crossover at $\Delta = 4$: critical temperature as a function of the coupling J' between the chains. The critical temperature has been determined from a data collapse of the Binder cumulant. Error bars are included but much smaller than the symbol size. The solid line is a guide to the eye connecting the data points. The dotted line near $J' = 1$ shows the result of the symmetry argument, eq. (4.29), using the value $T_c(J' = 1)$ in the isotropic 2D square lattice (black circle). The green dotted-dashed line shows the exact solution to the classical 2D Ising model by Onsager [4].

loops of the SSE algorithm are, however, quite small and local, thus these updates require many Monte Carlo sweeps and increase the autocorrelation time.

We extracted the critical temperature for different values of J' from a data collapse of the Binder cumulant and collected the results in Figure 4.14. Irrespectively of the algorithmic problems described in the previous paragraph, the values for $J' \geq 0.2$ should be reliable. The values for $J' \in \{0.05, 0.1\}$ should be considered with care. The statistics have been enhanced by simulations on a finer grid of temperature values and more (10^7) sweeps per data point. Although the Binder cumulant still shows a crossing for the small J' , it is by far not as clear as for the larger J' . However, these two values clearly show the correct tendency. In Figure 4.14, the values $T_c(J')$ are compared to the exact solution of the classical 2D Ising model, which was given by Onsager in 1944 [4]. Onsager derived a relation for the critical temperature in the thermodynamic limit which reads for our XXZ -model – neglecting the “quantum” off-diagonal part –

$$\sinh\left(\frac{\Delta J}{2T_c}\right) \sinh\left(\frac{\Delta J'}{2T_c}\right) = 1. \quad (\text{classical Ising model}) \quad (4.26)$$

As a further check, we can use a symmetry argument to get the critical temperature $T_c(J' = J(1 \pm \epsilon))$ near $J' = J$ for small $\epsilon \ll 1$. Indeed, we can split the Hamiltonian $\mathcal{H}(J, J') = J\mathcal{H}_x + J'\mathcal{H}_y$ in two parts in x - and y -direction and separate the coupling

strength J and J' , respectively. We have:

$$\begin{aligned}\mathcal{H}(1, J(1 + \epsilon)) &= \frac{1 + \epsilon}{1 + \epsilon} J\mathcal{H}_x + (1 + \epsilon)J\mathcal{H}_y \\ &= (1 + \epsilon)\mathcal{H}\left(J\frac{1}{1 + \epsilon}, 1\right).\end{aligned}\quad (4.27)$$

However, it does not matter which direction we call x or y ; $\mathcal{H}(J, J') \equiv \mathcal{H}(J', J)$. Thus, we have $\mathcal{H}(1, J' = 1 + \epsilon) = (1 + \epsilon)\mathcal{H}(1, J' = \frac{1}{1 + \epsilon})$ if we measure all energies in units of $J \equiv 1$. Accordingly, we must have $T_c(J' = 1 + \epsilon) = (1 + \epsilon)T_c(J' = \frac{1}{1 + \epsilon})$ for the critical temperature. With a Taylor expansion $T_c(J' = 1 + \epsilon) = T_c(1) + \epsilon\partial_\epsilon T_c + \mathcal{O}(\epsilon^2)$, we obtain to first order in ϵ :

$$T_c(1) + \epsilon\partial_\epsilon T_c = (1 + \epsilon)(T_c(1) - \epsilon\partial_\epsilon T_c) + \mathcal{O}(\epsilon^2) \quad \Rightarrow 2\partial_\epsilon T_c = T_c(1), \quad (4.28)$$

$$T_c(J' = 1 + \epsilon) = T_c(J' = 1) \left(1 + \frac{\epsilon}{2} + \mathcal{O}(\epsilon^2)\right). \quad (4.29)$$

This expansion is in good agreement with the Monte Carlo simulations.

All in all, the crossover between one and two dimensions of the XXZ -model at $\Delta = 4$ shows no fundamentally different features than the classical 2D Ising model; the phase diagram Figure 4.14 is almost identical. Quantitatively, the quantum fluctuations are larger in one dimension (and close to it), but they do not change the picture of the crossover qualitatively.

4.5 Dimers: Quantum Phase Transition

We now turn to the geometry of coupled dimers. When we talk of *dimers* we think of two neighboring spins coupled by a bond of unit strength $J \equiv 1$. As depicted in Figure 2.2(b), the dimers are regularly arranged on the 2D square lattice, parallel to the x -axis, and in an alternating pattern in y -direction. The remaining bonds have a strength $J' \leq 1$. While such a pattern is realized in the experiment with ultra-cold quantum gases by Greif *et al.* [2] in three dimensions, here we restrict ourselves to two dimensions where the numerical calculation is more easily feasible.

In contrast to the previous section, the off-diagonal quantum part in the Hamiltonian is much more relevant here: if one would neglect it, the ground state would be the Néel state with alternating spins on neighboring sites at any $J' > 0$. But we have already mentioned in section 2.5 that the ground state at $J' = 0$ is a product state of singlets on the dimers, which has no long-range order. The ground state remains a valence bond crystal up to some critical coupling J_c above which the ground state is Néel ordered and there is a finite temperature transition to the ordered phase. Thus, the off-diagonal terms in the Hamiltonian change the picture qualitatively and give rise to a quantum phase transition of the ground state, i.e., at zero temperature, which is driven by the coupling parameter J' in the Hamiltonian instead of the temperature. Of course, there are still finite-temperature transitions for $J' > J_c$ when the ground state is ordered.

To enhance our understanding of the quantum phase transition we use the following

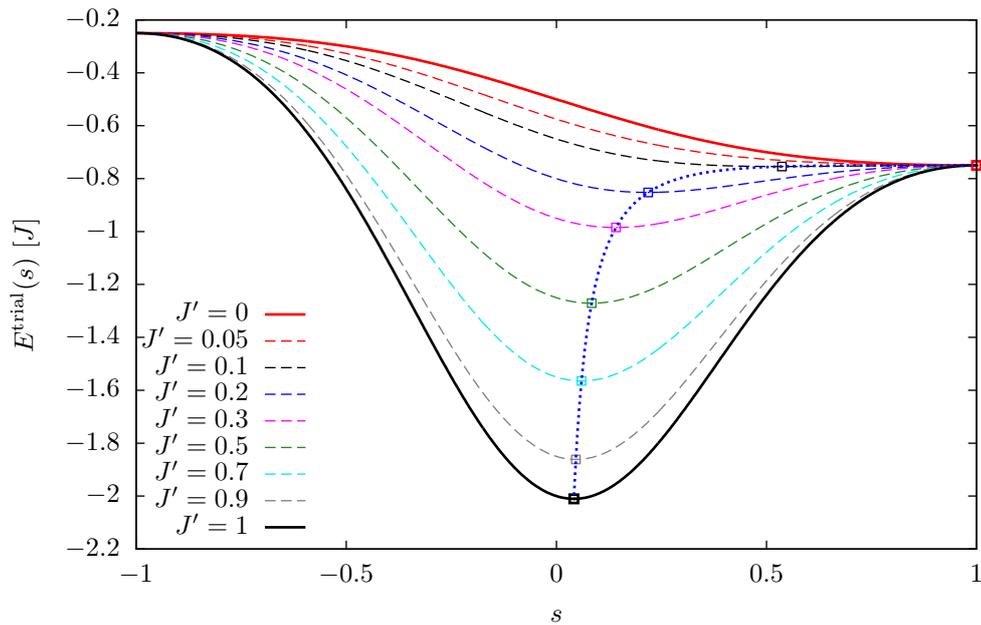


Figure 4.15: Energy per site $E^{\text{trial}}(s) = \langle \Psi^{\text{trial}} | \mathcal{H} | \Psi^{\text{trial}} \rangle / N$ of the trial state given in eq. (4.30) for $\Delta = 4$. The thick lines show the results for the extremal cases $J' = 0$ (decoupled dimers) and $J' = 1$ (spatially isotropic). The thin dashed lines correspond to values of J' shown in other figures in this section (with the same colors). The boxes show the locations of the minima; the blue dotted line indicates the movement of these with J' .

trial state as an approximation to the exact ground state [9]:

$$|\Psi^{\text{trial}}\rangle = \prod_{\text{dimers}\langle i,j\rangle} \frac{1}{\sqrt{1+|s|^2}} (|\uparrow_i\downarrow_j\rangle - s|\downarrow_i\uparrow_j\rangle). \quad (4.30)$$

Here the product is over all dimers with a strong bond $J \equiv 1$ and s is a free parameter which tunes between the types of ground states. With the choice $s = 0$ the trial state $|\Psi^{\text{trial}}\rangle$ is the classical Néel state with alternating spins which is close to the exact ground state in the isotropic lattice with $J' = 1$. On the other hand, the trial state is the exact ground state of the decoupled dimers for $J' = 0$ if $s = 1$ is chosen. The best approximation of the ground state is then obtained by minimization of $E^{\text{trial}}(s) := \langle \Psi^{\text{trial}} | \mathcal{H} | \Psi^{\text{trial}} \rangle / N$ by variation of s . Clearly, a trial state with $|s| > 1$ is equivalent to one with $1/s$ (up to a sign flip on all sites and a possible global phase); the limit $|s| \rightarrow \infty$ is well defined and the same as $s \rightarrow 0$. Thus, it suffices to consider $|s| < 1$. In general s might be chosen complex; but one can check that the extrema of $E^{\text{trial}}(s)$ lie on the real axis. For simplicity, we only consider real s in the following.

The calculation of $E^{\text{trial}}(s)$ is given in appendix C and the result is shown in Figure 4.15 for different J' . Further, the position of the minimum is calculated as a function of J' . As one can see in Figure 4.15, the trial energy $E^{\text{trial}}(s)$ is minimal at $s = 1$ if the inter-dimer coupling is weaker than some critical value

$$J_c^{\text{trial}} = \frac{J}{3\Delta} \stackrel{\Delta=4}{=} 0.08\bar{3}J. \quad (4.31)$$

This is the expected result: for weaker couplings $J' < J_c^{\text{trial}}$ the singlet formation on the dimers energetically dominates over the alignment of the spins between the dimers. Of course, the trial state is not the real ground state for $J' > 0$, but it is quite close: it neglects only the off-diagonal terms on the weak bonds of order J' which are by a factor Δ weaker than the next terms (namely the diagonal contributions on the weak bonds). At stronger couplings $J' > J_c^{\text{trial}}$ between the dimers the minimum shifts towards $s = 0$, i.e., towards the Néel state with long-range order. Although the minimum moves continuously, it shifts quite much over a very short range of J' : For $\Delta = 4$ at $J' = 0.1$, which is only slightly above J_c^{trial} , the minimum is at $s_{\text{min}} \approx 0.54$.

In the following we discuss the results of SSE simulations. The energy per site is shown in Figure 4.16. At large J' , similar as the coupled chains, the system is well described by classical considerations. The classical ground state energy is in this pattern $E_{\text{gs.}}/N = \frac{\Delta}{8}(J + 3J')$ and the classical spin gap is $E_{\text{gap}} = \frac{\Delta}{2}(J + 3J')$. As expected, the large energy gap decreases when J' is tuned from the isotropic couplings $J' = 1$ to smaller J' , and the strong finite size dependence – a sign of the divergence of the critical heat in the thermodynamic limit – indicates the drop of the critical temperature. As we have seen from our calculations with the trial state, the quantum nature – at least on the strong bonds – becomes important at low J' . In contrast to the chains (cf. Figure 4.11), the curves of $J' = 0.05$ and $J' = 0.1$ differ only by a few percent and show no detectable sign of a phase transition in the energy – at least not down to temperatures of $T = 0.15J$ as shown in the figures.

The squared staggered magnetization is depicted in Figure 4.17 and basically confirms the above picture. At large J' it qualitatively behaves like the isotropic 2D square lattice;

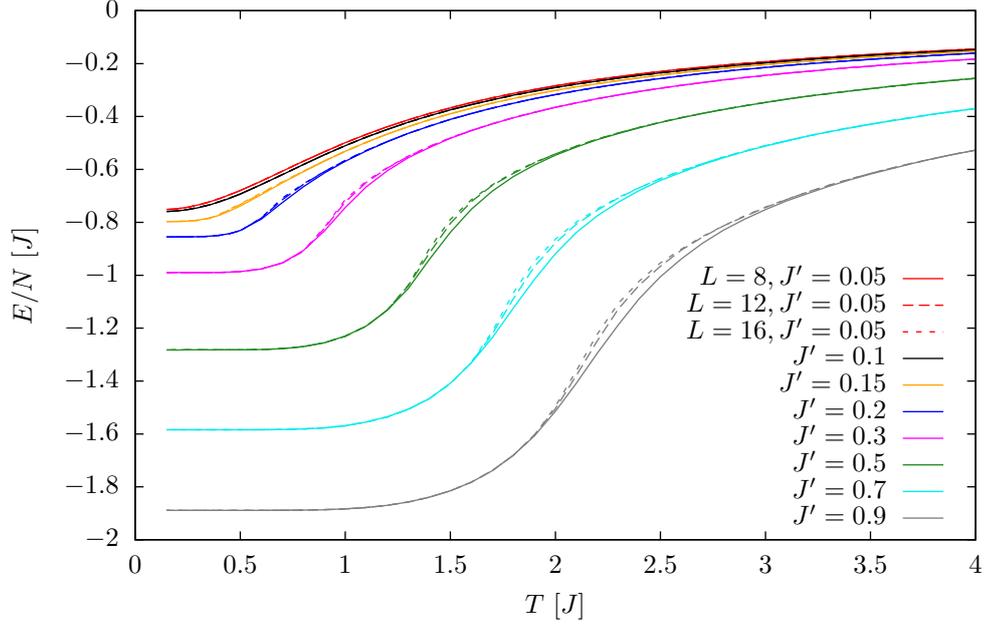


Figure 4.16: Energy per site of coupled dimers on a 2D square lattice, $\Delta = 4$. The color indicates the value of the inter-chain coupling constant J' , from top to bottom $J' \in \{0.05, 0.1, 0.2, 0.3, 0.5, 0.7, 0.9\}$. For each J' , different line types indicate the length $L \in \{8, 12, 16\}$ (different lengths are also plotted for $J' \leq 0.1$, but the curves lie on top of each other and can not be distinguished).

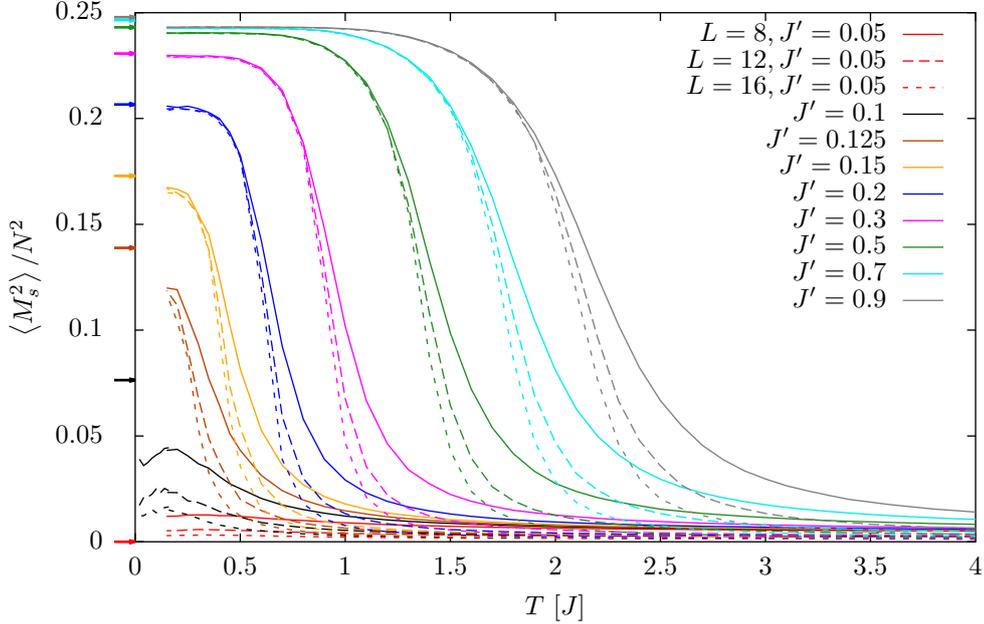


Figure 4.17: Squared staggered magnetization $\langle m_s^2 \rangle \equiv \langle M_s^2 \rangle / N^2$ of coupled dimers on a 2D square lattice, $\Delta = 4$. The arrows at the y -axis indicate the staggered magnetization of the trial state $\langle \Psi^{\text{trial}} | m_s^2 | \Psi^{\text{trial}} \rangle$ (see appendix, eq. (C.19)) in the thermodynamic limit $N \rightarrow \infty$ where the parameter s was chosen to minimize the energy for a given J' (with the same color key as the lines).

the critical temperature at which the order sets in behaves as expected. A weaker coupling between the dimers induces stronger quantum fluctuations on the strong bonds which manifest themselves in a decreasing maximum value of M_s in the ordered phase. Indeed, the arrows at the y -axis in Figure 4.17 indicate the squared staggered magnetization in the optimal trial state $\langle \Psi^{\text{trial}} | m_s^2 | \Psi^{\text{trial}} \rangle$ where s is determined by the minimum in Figure 4.15 and thus depends on J' . The calculation is given in appendix C. The trial state underestimates the quantum fluctuations – especially for large J' – since it takes into account the off-diagonal term of the Hamiltonian on the strong bonds only. Hence, the value m_s^2 estimated from the trial state is higher than in the exact ground state. However, it clearly gives the right qualitative tendencies. In turn, the good agreement justifies the choice of the trial state a posteriori. As expected from the considerations with the trial state, the quantum fluctuations destroy the long-range order completely at finite $J' > 0$. At $J' = 0.05$ we see only finite size fluctuations of M_s which vanish in the thermodynamic limit.

While the trial state calculation suggested that $J' = 0.1$ is above the critical coupling and should have a finite temperature transition, we were not able to determine a critical temperature for this coupling strength. For this J' the staggered magnetization increases at low temperatures, but has still a very strong finite size dependence at low temperatures – in contrast to larger J' where we observe a plateau with weak finite size dependence in the ordered phase. The value of m_s^2 suggested by the trial state seems to be too high in any case, which holds true for $J' \in \{0.125, 0.15\}$ as well. However, we observed a phase transition to an ordered phase at $T_c = 0.270(1)J$ for $J' = 0.125$, which is thus an upper limit to J_c .

Once more, we used a data collapse of the Binder cumulant to determine the critical temperature for the different values of J' . The resulting phase diagram is shown in Figure 4.18. The slope of the critical temperature $T_c(J')$ as a function of the coupling J' at $J' \in \{0.125, 0.15, 0.175\}$ also suggests that the critical coupling J_c of the quantum phase transition is then the value of the trial state $J_c^{\text{trial}} = 0.08\bar{3}$. For an exact determination of J_c one would need to carry out calculations for more different values of J' . Strictly speaking, J_c is determined by the ground state properties, and another sampling scheme – or even a completely different numerical method – might be better suited to determine J_c to high precision. We will here only give the conservative estimate $J_c = 0.10(2)$ based on the following three facts: at $J' = 0.125$ we clearly have a finite T_c below which the system orders, the trial state gives a lower bound at $J' = 0.8\bar{3}$, and finally we can not detect a finite temperature transition at $J' = 0.1$ but see large finite size dependence with increasing ordering at low temperatures.

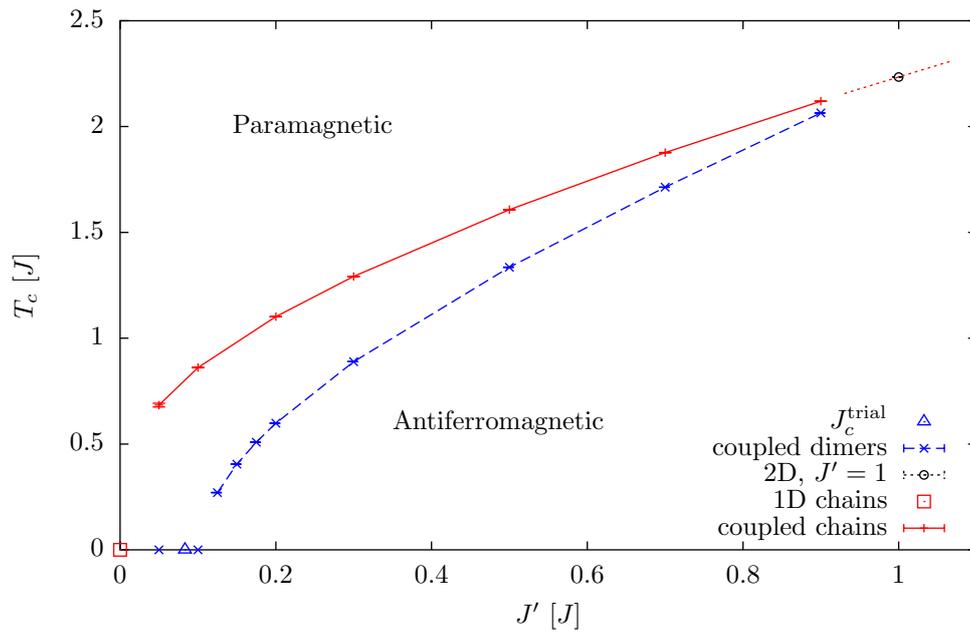


Figure 4.18: Phase diagram for the coupled dimers (blue) at $\Delta = 4$: critical temperature as a function of the coupling J' between the dimers. The critical temperature was determined by a data collapse of the Binder cumulant. Error bars are included but much smaller than the symbol size. For reference, the 1D-2D-crossover of Figure 4.14 has also been included (red). The solid line is a guide to the eye connecting the data points. The blue triangle shows the critical coupling J_c^{trial} obtained from the trial state.

5 Mutual Information

In this chapter we will change our point of view and turn to the fundamental level of information theory. Therefore, we will first review some basic notations, in particular we discuss different entropies, which measure something as fundamental as the amount of information in a random variable. While these measures are comparatively easily accessible in ED and DMRG, it requires a trick to obtain them in Monte Carlo simulations. We will thus first discuss the necessary modifications for the so-called ‘replica trick’ [1] in section 5.2, before we give results from SSE simulations. Remarkably, we will see that the mutual information can be used to detect phase transitions, as well. Finally, we will discuss a modification of the replica trick in a similar spirit, namely the so-called ‘ratio trick’ [80].

5.1 Some Basics of Information Theory

The mathematical foundations of information theory have been laid by Shannon in 1948 [26]. Consider a statistical experiment with discrete possible results $\nu \in \mathcal{C}$, where \mathcal{C} is a finite set, and denote the probabilities for the outcome ν with w_ν , i.e., $0 \leq w_\nu \leq 1$ and $\sum_\nu w_\nu = 1$. Shannon quantified the amount of information in the result of the experiment by the entropy

$$S(\{w_\nu\}) = - \sum_{\nu \in \mathcal{C}} w_\nu \log(w_\nu). \quad (5.1)$$

The entropy measures the amount of information in the result of the experiment. Note that we need to know the probabilities of the different results. For example, we know that a fair coin toss has two possible results, namely head and tail, which occur with the equal probability of 50%. The entropy measures the information we gain from a single coin toss; in this case it is called one bit.

Shannon motivated this quantity as a measure of information with three postulates which determine the form of eq. (5.1) up to a prefactor. Fixing the prefactor can be seen as the choice of units for the entropy. Note that we can absorb the prefactor into the base of the logarithm. While in information theory usually the binary logarithm to the base of 2 is used – and the unit of information is a bit –, we use the natural logarithm in this work. Through this choice and the natural units $k_B \equiv 1$, eq. (5.1) coincides with the thermodynamic entropy in classical statistical mechanics established by Boltzmann and Gibbs in the 19th century if we identify the probabilities w_ν with the Boltzmann weights, e.g., as for the Ising model in eq. (3.3).

If we know the macro state of a thermodynamic system, in particular the temperature, we know – at least theoretically – the probabilities for the single micro states and can give predictions for the micro state. The entropy quantifies the additional information in

the micro state, i.e., the amount of information one gets by determining the actual micro state. Equivalently, the entropy is a measure for the uncertainty of not knowing the micro state. Statements defining the entropy as the ‘measure of disorder’ have to be understood in this context and should not be confused with the order parameter quantifying a phase transition.

Let us recall some basic properties of the entropy [26, 28]:

- The entropy is non-negative, $S \geq 0$. It vanishes if and only if only one of the w_ν is nonzero (and thus equals 1). Indeed, it is clear that we do not get information from an experiment which has only one possible result and is thus completely predictable.
- The entropy is maximal if all results are equally likely, i.e., $w_\nu = 1/|\mathcal{C}|$. In that case it takes the value $S = S_\alpha = \log(|\mathcal{C}|)$. Clearly, this is strictly monotonically increasing with the number of possible states $|\mathcal{C}|$. This is actually one of the postulates used by Shannon [26]. From a physical point of view, this case corresponds to infinite temperature. In this case, the entropy is extensive: the number of states – or the dimension of the Hilbert space in quantum mechanics – grows exponentially with the volume, i.e., the number of lattice sites N . For example, the Ising model has $|\mathcal{C}| = 2^N$ possible states and thus $S(T \rightarrow \infty) = N \log(2)$.
- If we have two independent probability distributions $\{w_\nu\}$ and $\{\tilde{w}_\mu\}$ (where ν and μ do not have to be in the same set \mathcal{C}), we can form the product $\{w_\nu \tilde{w}_\mu\}$. Then we have

$$S(\{w_\nu \tilde{w}_\mu\}) = S(\{w_\nu\}) + S(\{\tilde{w}_\mu\}). \quad (5.2)$$

This is known as additivity of entropies.

In quantum mechanics, we have to replace eq. (5.1) with the von Neumann entropy [27]

$$S(\hat{\rho}) = -\text{Tr} \{ \hat{\rho} \log(\hat{\rho}) \}, \quad (5.3)$$

where $\hat{\rho}$ is the density matrix describing the system. For a pure state $\hat{\rho} = |\Psi\rangle \langle\Psi|$ the entropy clearly vanishes. At finite temperatures we have to use the mixed density matrix

$$\hat{\rho} = \sum_{\nu} w_{\nu} |\nu\rangle \langle\nu| \quad \text{where } \mathcal{H} |\nu\rangle = E_{\nu} |\nu\rangle \quad \text{and } w_{\nu} = \frac{e^{-\beta E_{\nu}}}{\mathcal{Z}}. \quad (5.4)$$

We can evaluate the trace of eq. (5.3) in the energy eigenbasis and recover the Shannon entropy. If we have a unique ground state $|\Psi_{\text{gs.}}\rangle$ without degeneracy, the density matrix $\hat{\rho} = |\Psi_{\text{gs.}}\rangle \langle\Psi_{\text{gs.}}|$ at $T = 0$ is pure, i.e., $\hat{\rho}^2 = \hat{\rho}$. We see immediately that $S(|\Psi_{\text{gs.}}\rangle \langle\Psi_{\text{gs.}}|) = 0$. However, we will see shortly that the entropy in a subregion does not vanish in quantum mechanical systems – in contrast to classical systems.

Suppose we divide our system into two regions. We denote them with A and B and assume that B is the complement of A , i.e., $A \cap B = \emptyset$ and $A \cup B$ is the complete system. Thanks to the lattice structure this is easily possible. In this work, we consider two different shapes for region A , namely the strips and squares depicted in Figure 5.1. If



Figure 5.1: Examples for different regions of the 2D square lattice with dimensions $L \times L$, here $L = 6$. The colors indicate different $h = 1, 2, 3$ (for red, blue, and green, respectively). For given h we say all sites within the corresponding dashed line belong to region A , while the complement – i.e., all sites outside of the dashed line – is denoted with B . The boundary of a strip in (a) cuts $2L$ bonds (independent of h : it does not cut the bond of the periodic boundary in y -direction); the boundary of the squares (b) cuts $4h$ bonds.

are only interested in region A , we integrate out region B to obtain the reduced density matrix $\hat{\rho}_A$. This is accomplished by a partial trace over region B :

$$\hat{\rho}_A = \text{Tr}_B \{ \hat{\rho} \} \equiv \sum_{\beta} \langle \beta | \hat{\rho} | \beta \rangle \equiv \sum_{\alpha, \alpha', \beta} \rho_{\alpha, \beta; \alpha', \beta} |\alpha\rangle \langle \alpha'|, \quad \text{where } \rho_{\alpha, \beta; \alpha', \beta} := \langle \alpha | \langle \beta | \hat{\rho} | \alpha' \rangle | \beta' \rangle. \quad (5.5)$$

Here, the sums run over orthonormal basis sets $\{|\alpha\rangle\}$ and $\{|\beta\rangle\}$ in region A and B , accordingly. Since the total Hilbert space is a tensor product over all sites, it is natural to use the product basis of the local basis states, in our case

$$\{|\alpha\rangle\} = \left\{ \prod_{i \in A} |\sigma_i\rangle, \text{ where } \sigma_i \in \{\uparrow, \downarrow\} \right\}. \quad (5.6)$$

For a classical system as the Ising model, the energy eigenstates $|\nu\rangle$ are just a simple products, $|\nu\rangle = |\alpha\rangle |\beta\rangle$. The reduced density matrix then takes the simple, diagonal form

$$\hat{\rho}_A^{\text{clas.}} = \sum_{\alpha} \left\{ \underbrace{\left(\sum_{\beta} w_{\alpha, \beta} \right)}_{=: w_{\alpha}} |\alpha\rangle \langle \alpha| \right\}. \quad (5.7)$$

The von Neumann entropy (5.3) then clearly reduces to the Shannon entropy for the reduced weights w_{α} . In particular, the entropy $S(\hat{\rho}_A^{\text{clas.}})$ of the reduced system vanishes for $T \rightarrow 0$ if the ground state is not degenerate: in this case $w_{\alpha, \beta}$ is nonzero only for one combination of α and β . In the 2D Ising model the ground state is two fold degenerate. Thus, at zero temperature $w_{\alpha, \beta}$ is $\frac{1}{2}$ only if either all spins in α and β point up or all spins point down; otherwise $w_{\alpha, \beta}$ is zero. Note that in this case the knowledge of α implies the knowledge of β since the two regions are not independent of each other¹. Accordingly,

¹ If we had no bonds between the two regions A and B , they would be independent and the entropy would just sum up, $S(\hat{\rho}) = S(\hat{\rho}_A) + S(\hat{\rho}_B)$, according to eq. (5.2).

the entropy of the Ising model at zero temperature is $S(\hat{\rho}^{\text{clas.}}) = S(\hat{\rho}_A^{\text{clas.}}) = \log(2)$ for any (non-empty) region A .

As mentioned above, quantum systems behave differently: even with a unique ground state $|\Psi_{\text{gs.}}\rangle$ at zero temperature, such that we have a pure density matrix $\hat{\rho} = |\Psi_{\text{gs.}}\rangle\langle\Psi_{\text{gs.}}|$, the reduced density matrix will be mixed. As a simple example consider the singlet $|S\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ – i.e., the ground state of a two site Heisenberg dimer discussed in section 2.4 – and one of the two sites as region A . The corresponding density matrices are

$$\hat{\rho} = \frac{1}{2}(|\uparrow\downarrow\rangle\langle\uparrow\downarrow| - |\uparrow\downarrow\rangle\langle\downarrow\uparrow| - |\downarrow\uparrow\rangle\langle\uparrow\downarrow| + |\downarrow\uparrow\rangle\langle\downarrow\uparrow|), \quad \hat{\rho}_A = \frac{1}{2}(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|). \quad (5.8)$$

Although $\hat{\rho}$ itself is pure by construction, the reduced density matrix $\hat{\rho}_A$ clearly is not. Thus, the quantum entanglement leads to a nonzero entropy of the reduced system, although the system as a whole is in a well defined, unique ground state.

We have noted above that the entropy is extensive at high temperatures. Thus, one might also expect that the entropy of a reduced system in the region A scales with the number of sites $|A|$ within the considered region. Clearly, this is true at high temperatures. However, this is usually not the case at low temperatures. Instead, the entanglement entropy of many physical ground states seems to follow an area law [29]

$$S(\hat{\rho}_A) \propto |\partial A|, \quad (5.9)$$

where $|\partial A|$ is the surface area of the boundary ∂A between the two regions A and B . Intuitively, we can argue the entanglement entropy emerges from the quantum correlations along the boundary of the region, which should be short-ranged as they emerge from short-range interactions. Nonetheless, there are situations where the area law does not hold: at quantum critical points in 1D the entropy scales with the logarithm of the region length [29, 81], For a review of rigorous results and the area laws in general see Ref. [29]. Beside its physical relevance the entanglement entropy is also of importance for numerical simulations of correlated quantum many-body systems: one can easily convince oneself that the entanglement of matrix product states (MPS) – which are used in the density matrix renormalization group (DMRG) method – and their higher dimensional analogs, the projected entangled pair states (PEPS), obey an area law [29]. The great success of DMRG can thus be related to the fact that the entanglement entropy of many physical ground states in one dimension does not increase with the size of the region², and thus MPS can be used to approximate the ground state very well.

5.1.1 Mutual information

Above, we discussed the entropy $S(A) := S(\rho_A)$ of the region A where the state of the complement is neglected. One may also ask how much information is contained in the region A if the state of region B is known. We discuss this first for classical systems with weights $w_{\alpha,\beta}$. As mentioned above, $S(A)$ is just the Shannon entropy of the reduced weights:

$$S(A) = - \sum_{\alpha} w_{\alpha} \log(w_{\alpha}), \quad \text{where } w_{\alpha} = \sum_{\beta} w_{\alpha,\beta}. \quad (5.10)$$

² Clearly, the surface area of a connected one dimensional region is a constant.

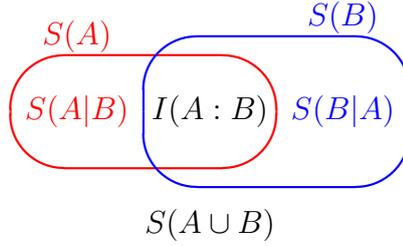


Figure 5.2: Pictorial representation of relations between entropies in regions A and B , namely between the joint entropy $S(A \cup B)$, the individual entropies $S(A)$ and $S(B)$ in each region, the conditional entropies $S(A|B)$ and $S(B|A)$ and the mutual information $I(A : B)$. All these quantities are positive. Reproduced from Ref. [82].

Now suppose that we fix the state of the system in region B to some β instead of taking the partial trace. The remaining distribution in A is then given by the conditional probabilities $w_{\alpha,\beta}/(\sum_{\alpha} w_{\alpha,\beta}) = w_{\alpha,\beta}/w_{\beta}$. Accordingly, the entropy is

$$S(A|\beta) = - \sum_{\alpha} \frac{w_{\alpha,\beta}}{w_{\beta}} \log \left(\frac{w_{\alpha,\beta}}{w_{\beta}} \right). \quad (5.11)$$

With an average over β , we obtain the so-called conditional entropy [26]:

$$\begin{aligned} S(A|B) &:= \sum_{\beta} w_{\beta} S(A|\beta) \equiv - \sum_{\alpha,\beta} w_{\alpha,\beta} \log \left(\frac{w_{\alpha,\beta}}{w_{\beta}} \right) \\ &= - \sum_{\alpha,\beta} w_{\alpha,\beta} \log(w_{\alpha,\beta}) + \sum_{\beta} \sum_{\alpha} w_{\alpha,\beta} \log(w_{\beta}) = S(A \cup B) - S(B). \end{aligned} \quad (5.12)$$

Here, we labeled the entropy of the complete system (sometimes called joined entropy) with $S(A \cup B) \equiv S(\rho_{A \cup B})$ to distinguish it from the entropy of the single regions. The conditional entropy measures the remaining uncertainty in region A if the state on region B is known. Clearly, this quantity is not symmetric with respect to the regions, $S(A|B) \neq S(B|A)$. Instead, we see from eq. (5.12) that $S(A \cup B) = S(A|B) + S(B) = S(B|A) + S(A)$. These relations are pictorially represented in Figure 5.2. The figure makes it easy to identify a further measure of information regarding the two regions: the mutual information.

As the name suggests, the mutual information measures the amount of information that the knowledge of the state in region A gives us about the region B , or vice versa. If we have no information about the system in B , our uncertainty about A is given by $S(A)$; the knowledge of B reduces this uncertainty to the conditional entropy $S(A|B)$. It is intuitively clear that $S(A|B) \leq S(A)$, which can in fact be proven [82]. The mutual information is then the (positive) difference between these quantities:

$$I(A : B) := S(A) - S(A|B) \equiv S(A) + S(B) - S(A \cup B). \quad (5.13)$$

In terms of weights for classical distributions, we can write the mutual information as:

$$I(A : B) = \sum_{\alpha,\beta} w_{\alpha,\beta} \log \left(\frac{w_{\alpha} w_{\beta}}{w_{\alpha,\beta}} \right). \quad (5.14)$$

However, the definition eq. (5.13) is more advantageous here since it can directly be used in quantum mechanics: we can just substitute the classical Shannon entropies with the quantum mechanical von Neumann entropies of the (reduced) density matrices.

Area Law for the Mutual Information

The mutual information measures only the relative information between two regions, i.e., it measures the *correlations* between A and B . If we have a finite correlation length (much smaller than the size of the regions), the sites which are separated by a distance larger than the correlation length will not be influenced by the state in the other region. This heuristic argument suggests an area law for the mutual information, which can in fact be proven for both quantum and classical systems [83]. In the following, we give the proof along the lines of Ref. [83] for quantum systems in the thermodynamic Gibbs ensemble defined in eq. (5.4).

We can split the Hamiltonian into three parts $\mathcal{H} = \mathcal{H}_A + \mathcal{H}_\partial + \mathcal{H}_B$ where \mathcal{H}_A contains all terms which depend only on the state of region A , \mathcal{H}_B those of region B , respectively, and \mathcal{H}_∂ contains the terms depending on both regions, i.e., bonds with one site in each of A and B . It is a well known fact of statistical physics that the Gibbs ensemble minimizes the free energy $F(\hat{\rho})$. In particular, we can compare the Gibbs ensemble of the complete system $\hat{\rho}_{AUB}$ with the tensor product of the reduced density matrices:

$$F(\hat{\rho}_{AUB}) \leq F(\hat{\rho}_A \otimes \hat{\rho}_B). \quad (5.15)$$

Now, we evaluate both sides further with $F(\rho) = \text{Tr} \{ \hat{\rho} \mathcal{H} \} - \frac{1}{\beta} S(\hat{\rho})$. By definition $\mathcal{H}_A = \mathcal{H}_A \otimes \mathbb{1}$ is independent of the state in B . Thus, we obtain on the left hand side:

$$\begin{aligned} F(\hat{\rho}_{AUB}) &= \text{Tr} \{ \hat{\rho}_{AUB} (\mathcal{H}_A + \mathcal{H}_B + \mathcal{H}_\partial) \} - \frac{1}{\beta} S(\hat{\rho}_{AUB}) \\ &= \text{Tr}_A \{ \hat{\rho}_A \mathcal{H}_A \} + \text{Tr}_B \{ \hat{\rho}_B \mathcal{H}_B \} + \text{Tr} \{ \hat{\rho}_{AUB} \mathcal{H}_\partial \} - \frac{1}{\beta} S(\hat{\rho}_{AUB}). \end{aligned} \quad (5.16)$$

By definition of the tensor product $\hat{\rho}_A \otimes \hat{\rho}_B$ the distributions in the two regions are independent, and thus we can use the additivity of the entropy eq. (5.2) to evaluate the right hand side of eq. (5.15):

$$\begin{aligned} F(\hat{\rho}_A \otimes \hat{\rho}_B) &= \text{Tr} \{ \hat{\rho}_A \otimes \hat{\rho}_B (\mathcal{H}_A + \mathcal{H}_B + \mathcal{H}_\partial) \} - \frac{1}{\beta} S(\hat{\rho}_A \otimes \hat{\rho}_B) \\ &= \text{Tr}_A \{ \hat{\rho}_A \mathcal{H}_A \} + \text{Tr}_B \{ \hat{\rho}_B \mathcal{H}_B \} + \text{Tr} \{ \hat{\rho}_A \otimes \hat{\rho}_B \mathcal{H}_\partial \} - \frac{1}{\beta} (S(\hat{\rho}_A) + S(\hat{\rho}_B)). \end{aligned} \quad (5.17)$$

In the last line, we have used the normalization of the reduced density matrices. Comparison with eq. (5.16) yields from eq. (5.15):

$$\begin{aligned} I(A : B) &\equiv S(\hat{\rho}_A) + S(\hat{\rho}_B) - S(\hat{\rho}_{AUB}) \leq \beta \text{Tr} \{ (\hat{\rho}_A \otimes \hat{\rho}_B - \rho_{AUB}) \mathcal{H}_\partial \} \\ &\leq \beta |\partial A| |\Delta E_{\max}|. \end{aligned} \quad (5.18)$$

The last inequality assumes that \mathcal{H} contains only nearest neighbor interactions – such that the number of bonds crossing the boundary equals the ‘area’ of the boundary $|\partial A|$ between

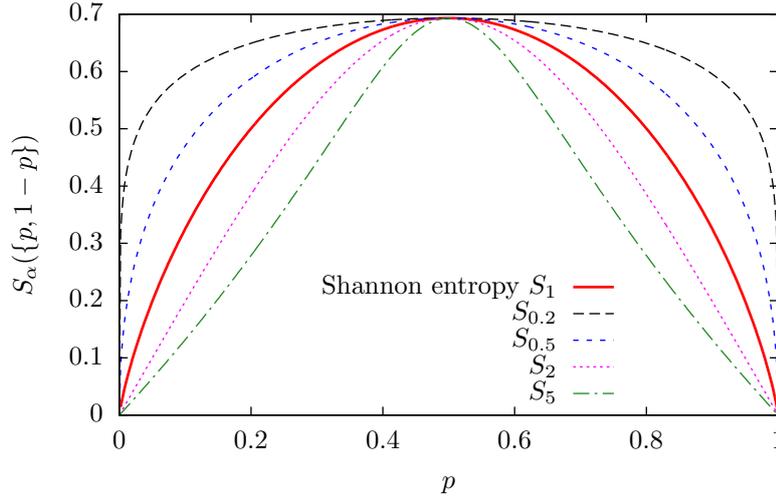


Figure 5.3: Rényi entropies of different order m for a classical distribution $\{p, 1 - p\}$.

the regions A and B – and that the energy spectrum of a single bond is finite, the range given by $|\Delta E_{\max}|$, the difference between the maximum and the minimum eigenenergy of a single bond.

This shows that the mutual information obeys an area law at finite temperatures. This law should not be confused with the area law of the quantum entanglement entropy, eq. (5.9). In fact, all the individual, joint, and conditional entropies are extensive at high temperatures. Note also the prefactor of the inverse temperature β in eq. (5.18). This prefactor makes the estimate useless for the ground state at zero temperature, $T \rightarrow 0$. The joint entropy $S(A \cup B) = 0$ vanishes in this case, such that the mutual information is basically just the entanglement entropy $S(A) + S(B)$. As mentioned in the discussion of eq. (5.9), the area law of these quantities is violated at quantum critical points.

5.1.2 Rényi Entropies

Unfortunately, there is – at least to the knowledge of the author – no way to calculate the Shannon (or von Neumann) entropy in (quantum) Monte Carlo methods directly. However, there are generalizations – the so-called Rényi entropies, introduced by Rényi in 1961 [28] – which can be calculated with the help of the replica trick discussed in the next section.

The Rényi entropies are defined for different orders³ $m > 0, m \neq 1$, which is indicated by an index to S . For discrete classical distributions, they are defined as [28]

$$S_m(\{w_\nu\}) = \frac{1}{1 - m} \log \left(\sum_{\nu \in \mathcal{C}} w_\nu^m \right). \quad (5.19)$$

The Rényi entropies of different order are compared in Figure 5.3 for the simplest (non-trivial) distribution: a decision between just two possibilities, one of them chosen with

³ In literature (e.g., by Rényi himself), the order is often labeled with α . However, in this work we use m to distinguish it from the states $|\alpha\rangle$ in region A .

probability p and the other with $1 - p$. It can easily be seen⁴, that the limit $m \rightarrow 1$ recovers the Shannon entropy from the Rényi entropies. Therefore, we may write S_1 for the Shannon (or von Neumann) entropy to distinguish it from Rényi entropies of different order. The Rényi entropies are mainly based on the same postulates as the Shannon entropy⁵, thus they share many of their properties, and one can also use the Rényi entropies as a measure for information content instead of the Shannon entropy. Further, we note that the Rényi entropies are monotonically decreasing in the order m , i.e., $S_m \leq S_{\tilde{m}}$ for $m > \tilde{m}$. Thus, the Rényi entropy of order 2 – which we can explicitly calculate with the replica trick – is a lower bound to the Shannon entropy $S \equiv S_1$.

The definition of the classical Rényi entropies eq. (5.19) can easily be generalized to the quantum mechanical case of density matrices:

$$S_m(\hat{\rho}) := \frac{1}{1-m} \log(\text{Tr}\{\hat{\rho}^m\}). \quad (5.22)$$

Obviously, this definition can be used both for the joint entropy with $\hat{\rho}_{A \cup B}$ and for the individual entropies in the regions A and B with the reduced density matrices $\hat{\rho}_A$ and $\hat{\rho}_B$, respectively. From these, the conditional entropy and mutual entropy can be calculated with the relations depicted in Figure 5.2, e.g., the Rényi mutual information of order m is – in analogy with eq. (5.13) – given by

$$I_m(A : B) := S_m(\hat{\rho}_A) + S_m(\hat{\rho}_B) - S_m(\hat{\rho}_{A \cup B}). \quad (5.23)$$

5.2 Replica Trick

The replica trick was used by Calabrese and Cardy in 2004 [81] in the context of conformal field theories. The basic idea is to use multiple copies of the system, and ‘sew’ (or ‘glue’) them together in the region A , which allows to obtain the trace of the reduced density matrix $\hat{\rho}_A$ to some power $m \in \{2, 3, \dots\}$, as we will see shortly. Calabrese and Cardy used this trick to calculate the von Neumann entropy from $\partial_m \text{Tr}\{\hat{\rho}_A^m\}|_{m \rightarrow 1}$.

In the context of (quantum) Monte Carlo methods, replicas were used for the first time by Hastings *et al.* [84] to calculate the Rényi entropy of order two (or more general, $m =$

⁴ Indeed, we can use l’Hôpital’s rule and rewrite $w_\nu^m = \exp(m \log w_\nu)$ to obtain:

$$\lim_{m \rightarrow 1} S_m = \frac{1}{-1} \lim_{m \rightarrow 1} \frac{\sum_\nu [\exp(m \log(w_\nu)) \log(w_\nu)]}{\sum_\nu w_\nu^m} = - \sum_\nu w_\nu \log(w_\nu) \quad (5.20)$$

⁵ Rényi replaced the following postulate of the Shannon entropy with the weaker additivity condition (5.2) for independent distributions [28]:

$$S(\{tw_1, (1-t)w_1, w_2, w_3 \dots\}) = S(\{w_1, w_2, \dots\}) + w_1 S(\{t, (1-t)\}). \quad (5.21)$$

This postulate can be interpreted as follows [26]: The complete process to determine a random variable consists of two subprocesses: First a variable is chosen according to the probabilities $\{w_1, w_2, \dots\}$. Then, if we got the result corresponding to w_1 , we make a second, independent decision between two alternatives with probabilities t and $(1-t)$. The left hand side of eq. (5.21) measures the information gained by the whole process. The right hand side distinguishes the two subprocesses and measures the gained information separately; the factor w_1 in the second term appears since the second decision is only necessary with probability w_1 .

$\{2, 3, \dots\}$) with a so-called ‘swap operator’ in the valence bond basis at zero temperature. Half a year later, Melko, Kallin, and Hastings [1] outlined the calculation of the mutual information $I_m(A : B)$ at finite temperatures in the formulation of SSE. In the following, we discuss this method in detail and apply it to the spatially anisotropic models.

The key achievement of the replica trick is the evaluation of $\text{Tr} \{\hat{\rho}_A^m\}$ for $m \in \{2, 3, \dots\}$. With the orthonormal bases $\{|\alpha\rangle\}$ and $\{|\beta\rangle\}$ (see eq. (5.6)) and the definition of the reduced density matrix $\hat{\rho}_A$ in eq. (5.5), we obtain:

$$\begin{aligned} \text{Tr} \{\hat{\rho}_A^m\} &= \sum_{\alpha_1} \langle \alpha_1 | \left(\sum_{\beta_1} \langle \beta_1 | \hat{\rho} | \beta_1 \rangle \cdots \sum_{\beta_m} \langle \beta_m | \hat{\rho} | \beta_m \rangle \right) | \alpha_1 \rangle \\ &\equiv \sum_{\alpha_1, \beta_1, \dots, \alpha_m, \beta_m} \rho_{\alpha_1, \beta_1; \alpha_2, \beta_1} \cdot \rho_{\alpha_2, \beta_2; \alpha_3, \beta_2} \cdots \rho_{\alpha_m, \beta_m; \alpha_1, \beta_m}. \end{aligned} \quad (5.24)$$

In the second line we have used the completeness of $\{|\alpha\rangle\}$ in the Hilbert space of region A where ρ_A is defined, see eq. (5.5). We can interpret this as m copies of the system – i.e., m density matrices $\hat{\rho}$ – with special boundary conditions: In region B the k -th density matrix is sandwiched between states β_k , i.e., we have a periodicity (with respect to the imaginary time direction) within each copy. Clearly, this periodicity arises from the partial trace in the definition of ρ_A . In contrast, the copies are glued together in region A : the final state of the first copy is the initial state of the second copy and so on. Thus, the copies are not independent of each other, but they effectively interact. The outer trace is over the Hilbert space corresponding to the region A (since $\hat{\rho}_A$ acts on this region only). Thus, we have a periodicity in region A as well, but only over *all* copies of the system: the final state of the last copy has to coincide (in region A) with the initial state of the first copy. This boundary structure is depicted in Figure 5.4 for $m = 2$.

To proceed, we separate the normalization, i.e., the partition function $\mathcal{Z}[T] = \text{Tr} \{e^{-\beta\mathcal{H}}\}$ of the ‘normal’ system, from the density matrix $\hat{\rho} = e^{-\beta\mathcal{H}}/\mathcal{Z}[T]$. We define a partition function $\mathcal{Z}[A, m, T] := \text{Tr}_A \{(\text{Tr}_B \{e^{-\beta\mathcal{H}}\})^m\}$ for the modified system of replicas as well such that we have:

$$\text{Tr} \{\hat{\rho}_A^m\} = \frac{\mathcal{Z}[A, m, T]}{(\mathcal{Z}[T])^m}, \quad (5.25)$$

$$S_m(A) = \frac{1}{1-m} \log(\text{Tr} \{\hat{\rho}_A^m\}) = \frac{1}{1-m} (\log(\mathcal{Z}[A, m, T]) - m \log(\mathcal{Z}[T])). \quad (5.26)$$

Although the partition function is not directly available in Monte Carlo simulations (a key point of the Metropolis algorithm is to use only the ratio of the weights such that the normalization does not matter), there are ways to obtain it indirectly. We will discuss a trick to directly evaluate the ratio of the partition functions in eq. (5.25) in the next section; but first we focus on the integration method outlined in Ref. [1].

This method simply uses the fact that the thermodynamic average of the energy is related to the partition function by a derivative:

$$E = \langle \mathcal{H} \rangle = -\frac{d}{d\beta} \log(\mathcal{Z}), \quad (5.27)$$

$$\Rightarrow \log(\mathcal{Z}[T]) = \log(\mathcal{Z}[T = \infty]) - \int_0^\beta d\tilde{\beta} E(\tilde{\beta}). \quad (5.28)$$

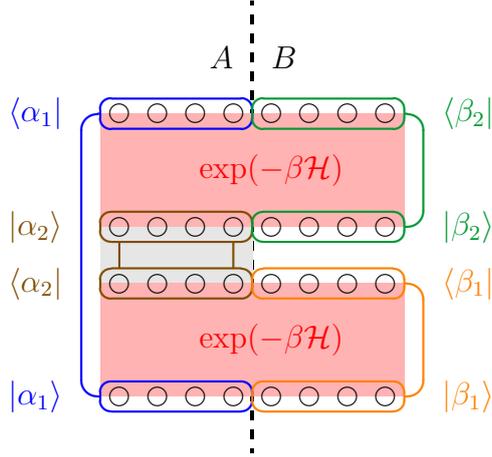


Figure 5.4: Geometry and boundary conditions used in the replica trick – i.e., the evaluation of $\mathcal{Z}[A, m, T] = \text{Tr}_A \{ \text{Tr}_B \{ e^{-\beta \mathcal{H}} \} \}$, see eq. (5.24) – for $m = 2$. The red shaded regions correspond to $e^{-\beta \mathcal{H}}$ and are expanded to operator strings as in Figure 3.2; each of them is periodic in region B (the right four sites), but not in region A . Instead, the operator strings are connected continuously in region A as indicated by the gray shaded area. This leads to an operator string m times as long as usual, with period $m\beta$ in region A . Reproduced from Ref. [80].

With this equation, we can evaluate the logarithms of the partition functions in eq. (5.26) with two separate Monte Carlo simulations: we run one simulation (with a fine grid of temperature ranges) with the usual geometry for the evaluation of $\log(\mathcal{Z}[T])$, and a second simulation with the modified geometry and boundary conditions depicted in Figure 5.4 for the evaluation of $\log(\mathcal{Z}[A, m, T])$.

We start the integration at $\beta = 0$ since this is the only point where we know the value of the partition function a priori. In general, one could also start the integration from other points where the partition function (or Rényi entropy) is known, e.g., from a separate calculation with the ratio trick, or in the ground state of a classical system (for which the replica trick can also be used [85]). At infinite temperature all states are weighted equally, $e^{-\beta \mathcal{H}} \rightarrow \mathbb{1}$. Thus, evaluating the partition functions involves simply counting the number of states; in our case we have

$$\log(\mathcal{Z}[T = \infty]) = \log \left(\sum_{\alpha, \beta} 1 \right) = (|A| + |B|) \log(2), \quad (5.29)$$

$$\log(\mathcal{Z}[A, m, T = \infty]) = \log \left(\sum_{\alpha} \left(\sum_{\beta} 1 \right)^m \right) = (|A| + m|B|) \log(2). \quad (5.30)$$

Here, $|A|$ and $|B|$ denote the number of sites in region A and its complement B , respectively; the sum $|A| + |B| = N$ is the number of sites.

We now discuss the necessary modifications in the SSE algorithm to implement the modified geometry with the replicas. For each replica we have to evaluate $e^{-\beta \mathcal{H}}$, but in this case the initial and final state in region A do not have to coincide. The mapping to the classical configuration space can be done in complete analogy to section 3.2. Thus, we have m operator strings – one for each replica. As before, we fix the length of each

operator string to some cutoff L . Recall that we fixed L with a constant factor times the maximal appearing number of (non-identity) operators n during the thermalization. While the number of (non-identity) operators n can fluctuate and needs not to be the same in different replicas, it is not necessary to use different cut-offs L : due to the overall periodicity in imaginary time, the replicas are equivalent and thus the distribution of n is identical on the different replicas. Actually, we have just used one large operator string of length $m \cdot L$ in the program code and identify the operator string of the replica $\tilde{m} \in \{1, 2, \dots, m\}$ with the part of the L positions $p \in \{(\tilde{m} - 1)L, \dots, \tilde{m}L - 1\}$. This storage scheme reflects directly the fact that the replicas are connected in region A and the ‘world lines’ are continuous in this region, see Figure 5.4.

Although it would basically suffice to store the initial state of region A only for the first replica due to this continuity, we store the complete initial states for both regions for each replica, i.e., all $|\alpha_{\tilde{m}}\rangle$ and $|\beta_{\tilde{m}}\rangle$ for all $\tilde{m} \in \{1, \dots, m\}$. The scaling of the computation time is dominated by the length of the operator string and thus not significantly changed, the additional memory requirements are nowadays not a problem – the overall SSE simulation uses 30 to 200 megabytes — and the ratio trick discussed in the next section needs less adjustments with this scheme.

The sampling of the diagonal update needs only small adjustments for the replica trick. First, recall that we generate the propagated states on the fly during the iteration over the propagation index. Of course, this procedure has to reflect the modified boundary conditions of Figure 5.4: whenever the increase of the propagation index changes between two replicas, we have to take care of the discontinuity of the states in region B . Further, one has to take care with the acceptance probabilities in eq. (3.37): one may not use the total number of non-identity operators n , but only the number within the corresponding replica $n_{\tilde{m}}$, and also the cut-off L is not the overall length of the total operator string, but of one replica only. This is necessary to reflect the fact that we expand $e^{-\beta\mathcal{H}}$ for each replica separately.

Recall further that we connect the vertices in the operators string with links (depicted as vertical lines in Figure 3.6(b)) before the loop updates for the off-diagonal sampling. For the modified geometry of the replicas, it suffices to adjust the construction of these links. Remarkably, the actual loop update can then be done as before, irrespective of whether the deterministic loop update or directed loops are used. The necessary modifications are straightforwardly taken from the modified topology of Figure 5.4; one simply has to reflect the two different periodicities in the two regions: In region A the states (or ‘world lines’) are continuous and the vertices are linked over the complete operator string, i.e., all replicas, exactly as before. In contrast, the periodicity within each replica in the complement B is directly transferred to the links.

Note that also the flips of free spins, i.e., sites on which no operator acts, have to be adjusted. This update is necessary for the sampling over the initial states. It only concerns the first propagated state $|\alpha(p=0)\rangle$ in region A , but all the initial states of the replicas $|\beta(p=(\tilde{m}-1)L)\rangle$, $\tilde{m} \in \{1, \dots, m\}$ in region B . It requires some additional book keeping for the replicas. For example it might happen that an operator acts on some site in region B in the first replica of the system, but not in the second replica. In that case, one should flip the spin in the initial state of the second replica (with probability $\frac{1}{2}$ since the weight does not change), but not in the initial state of the first replica.

The only observable we calculate in the replica system is the energy. To be more precise,

we want to calculate $-\partial_\beta \log(\mathcal{Z}[A, m, T])$ in order to use eq. (5.28). Note that this does not coincide with the ‘energy’ in the sense of $\langle \mathcal{H} \rangle_{A,m}$, where the average $\langle \cdot \rangle_{A,m}$ is taken with the sampling of the replica system as described above, i.e.,

$$\langle \hat{A} \rangle_{A,m} = \frac{\text{Tr}_A \left\{ \text{Tr}_B \left\{ e^{-\beta \mathcal{H}} \right\}^{m-1} \text{Tr}_B \left\{ e^{-\beta \mathcal{H}} \hat{A} \right\} \right\}}{\mathcal{Z}[A, m, T]}. \quad (5.31)$$

Instead, a straightforward calculation shows $-\partial_\beta \log(\mathcal{Z}[A, m, T]) = m \langle \mathcal{H} \rangle_{A,M}$. In complete analogy to eq. (3.31) we simply have to count the numbers of (non-identity) operators:

$$-\frac{d}{d\beta} \log(\mathcal{Z}[A, m, T]) = - \left\langle \frac{n_1 + \dots + n_m}{\beta} \right\rangle_{A,m} = - \left\langle \frac{n}{\beta} \right\rangle_{A,m}. \quad (5.32)$$

This quantity has to be used as the ‘energy’ for the integration in eq. (5.28), i.e., the only observable used in the SSE simulation with the modified topology is the number of (non-identity) operators in the complete operator string (over all replicas).

We finally give a short remark on the integration in eq. (5.28). Clearly, we have to run SSE simulations on a fine grid of temperatures. We chose β as integration variable since we start the integration at infinite temperature and make the grid uniform with respect to β . It lies in the nature of SSE that the resulting data points have statistical errors. Due to these errors, Newton-Cotes formulas of high order do not perform better than the ones of low order. In addition to the statistical errors, we have a systematic error which stems from the finite grid in β used for the integration. Clearly, a finer grid reduces the systematic error of the integration, but on the other hand this requires longer CPU times for the simulations, and it also increases the statistical errors: The integration is basically a sum; assuming statistically independent errors, the error of the sum is basically a random walk in 1D. Thus, the statistical error of the integration grows with the square root of the number of points used for the integrations⁶. Assuming fixed computer resources, i.e., a fixed CPU time, one has to balance between a fine grid, which reduces the systematic error, and a rougher grid with more CPU time per data point to reduce the statistical errors. To give a rough estimate, we used $\approx 10^6$ sweeps for each data point and a grid $\beta J \in \{0.001, 0.002, \dots\}$ to reach temperatures down to the order of $1J$, and a rougher grid and more sweeps to reach lower temperatures. Further, we reduced the systematic error a little bit⁷ with the following method: for each interval between two (neighboring) data points, we use the six nearest data points for a least square fit to a parabola; then we integrate the parabola on the corresponding interval. Note that the error bars given in this work only reflect the statistical errors, not the systematic error from the integration. They were obtained from a propagation of the bin averages with the jackknife method discussed in the end of section 3.1.

Now we have discussed all necessary ingredients for the replica trick with the integration method. A useful check of the implementation can be made by considering the two

⁶ However, this does not imply that the statistical error of the integration increases if a finer grid (with more data points) is used (while the statistical error of a single point is kept fix): the integration is normalized by the distance between two points in the grid. This follows the intuition that increasing the total number of Monte Carlo sweeps should decrease the statistical error.

⁷ We checked this by comparison with ED data. While it definitively helps, the effect is by far not as dramatic as the use of a finer grid in β .

extremal cases for the definition of the region A . First, we consider an empty region $A = \emptyset$. In that case, the partial trace over the complement B is already the trace over the complete Hilbert space. Consequently, the replicas are in fact just copies which are completely independent of each other and we have $\rho_A = \text{Tr}_B \{\hat{\rho}\} = 1$ and $\mathcal{Z}[A = \emptyset, m, T] = (\mathcal{Z}[T])^m$ (which implies $S_m(A = \emptyset) = 0$). Thus, we can relate the observable, eq. (5.32), to the average energy in the usual geometry, i.e., without replicas:

$$-\frac{d}{d\beta} \log(\mathcal{Z}[A, m, T]) = \left\langle \frac{n_1 + \dots + n_m}{\beta} \right\rangle_{A=\emptyset, m} [T] = m \langle \mathcal{H} \rangle [T]. \quad (5.33)$$

Here, we emphasized with the square brackets that the average is taken at the same temperature.

This is different in the second extremal choice – namely that the region A includes all sites and the complement $B = \emptyset$ is empty. Pictorially, the geometry of Figure 5.4 reduces to a single operator string with the normal boundary conditions, i.e., periodicity over the complete string, but m times longer than in the normal string at the same temperature. More formally, the partial trace $\text{Tr}_B \{\cdot\}$ has no effect and we have $\mathcal{Z}[B = \emptyset, m, T] = \text{Tr}_A \{(\text{Tr}_\emptyset \{e^{-\beta \mathcal{H}}\})^m\} = \text{Tr} \{e^{-m\beta \mathcal{H}}\} = \mathcal{Z}[\frac{T}{m}]$. In other words, we can absorb the larger length of the operator string in a rescaling of the temperature. Consequently, the analogous equation to eq. (5.33) reads:

$$-\frac{d}{d\beta} \log(\mathcal{Z}[A, m, T]) = -\frac{d}{d\beta} \log \left(\mathcal{Z} \left[\frac{T}{m} \right] \right) = m \langle \mathcal{H} \rangle \left[\frac{T}{m} \right]. \quad (5.34)$$

The Rényi entropy S_2 is shown in Figure 5.5 for different regions A . The curves approach constant values at high temperatures. Of course, this is no surprise since we constructed them in such a way. The limits are readily evaluated with equations (5.26) - (5.30),

$$S_m(T \rightarrow \infty) = \frac{\log(2)}{1-m} ((|A| + m|B|) - m(|A| + |B|)) = |A| \log(2). \quad (5.35)$$

When the system is cooled down, the entropy decreases monotonically to a minimum value of approximately one bit, i.e., $\log(2) \approx 0.693$. This value should be interpreted as the logarithm $\log(\Omega_0)$ of the degeneracy of the ground state: the ground states is two fold degenerate, the states are the two Néel states plus a few quantum fluctuations. In the previous chapter, we have seen that the quantum fluctuations are very weak for $\Delta = 4$, i.e., the ground states have huge overlap with the corresponding classical Néel states, but almost none with other states. In the limit $T \rightarrow 0$, the operator $e^{-\beta(\mathcal{H}-E_{\text{gs.}})}$ is just a projection operator onto the ground state manifold. Consequently, the partition functions for both the usual geometry $\mathcal{Z}[T]$ and for the replicas $\mathcal{Z}[A, m, T]$ basically⁸ take the value Ω_0 , i.e., the degeneracy in the ground state. Thus, by eq. (5.26) all Rényi entropies are given by $\log(\Omega_0)$. Deviations from this value are caused by the quantum entanglement. However, we will not discuss this here since the errors are significant at low temperatures: As discussed above, the statistical errors increase with the interval $[0, \beta]$ of the integration; additionally, both the partition functions diverge with β (compare the last footnote) such that a cancellation error appears. The ratio trick discussed in the next section is better suited to study the entanglement entropy in the limit of zero temperature.

⁸ This is correct up to the prefactor: $\mathcal{Z}[T] \equiv \text{Tr} \{e^{-\beta \mathcal{H}}\} = e^{-\beta E_{\text{gs.}}} \text{Tr} \{e^{-\beta(\mathcal{H}-E_{\text{gs.}})}\}$, and similarly $\mathcal{Z}[A, m, T] = e^{-m\beta E_{\text{gs.}}} \text{Tr}_A \{(\text{Tr}_B \{e^{-\beta(\mathcal{H}-E_{\text{gs.}})}\})^m\}$. The prefactors cancel in the calculation of the Rényi entropies, see eq. (5.26).

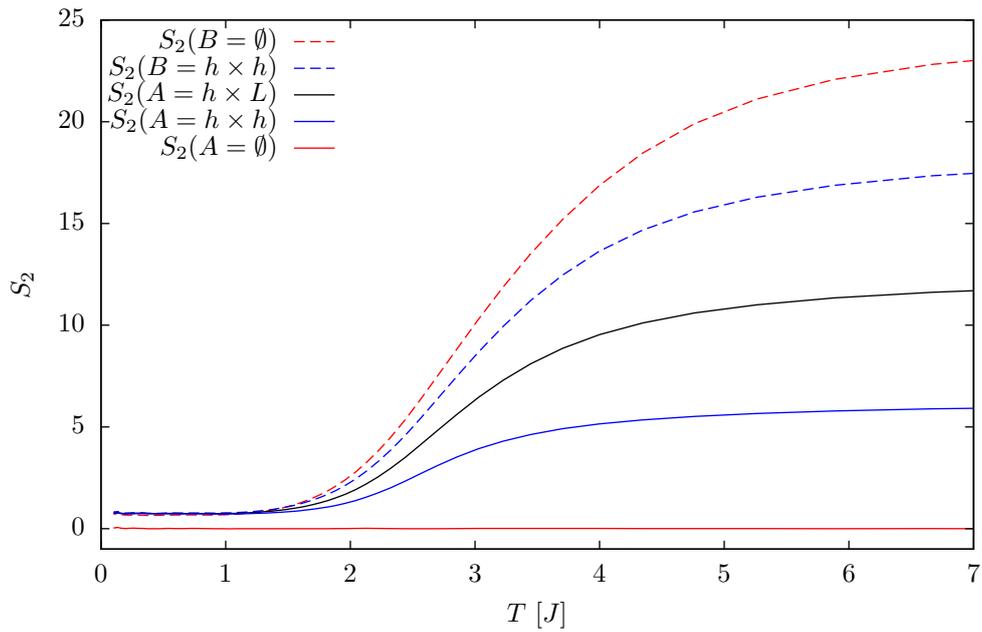
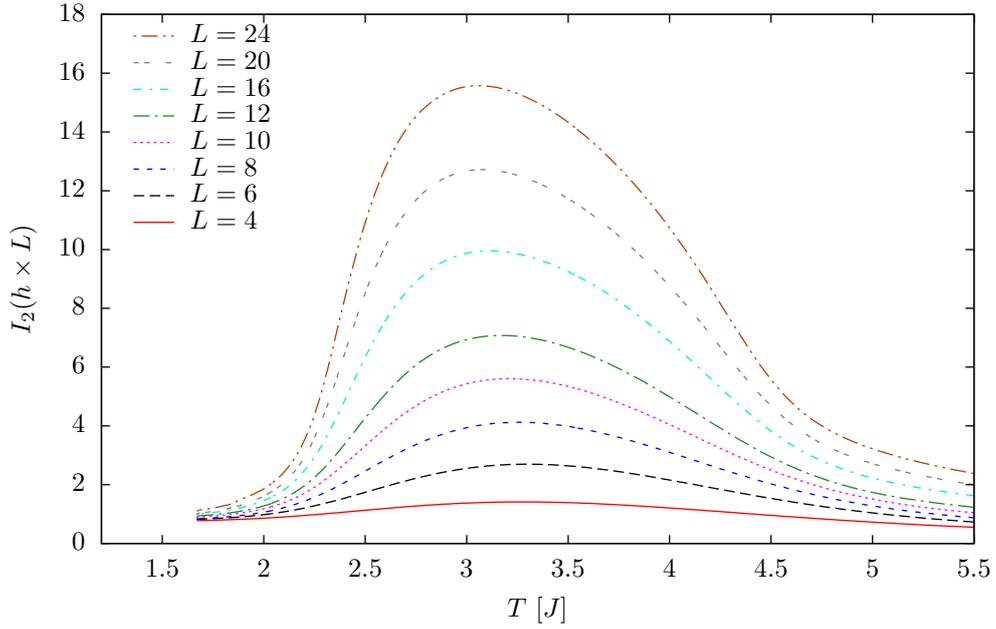


Figure 5.5: Rényi entropy of second order for different shapes of region A on a 2D square lattice of fixed length $L = 6$, $\Delta = 4$. The different shapes are labeled as in Figure 5.1 with $h = \frac{L}{2} = 3$. Note that $S_2(B = \emptyset) \equiv S_2(A \cup B)$ is the joint entropy of the complete system. As a consistency check, $S_2(A = \emptyset)$ has been calculated as well which should vanish at any temperature (see discussion above eq. (5.33)).



(a) Rényi mutual information

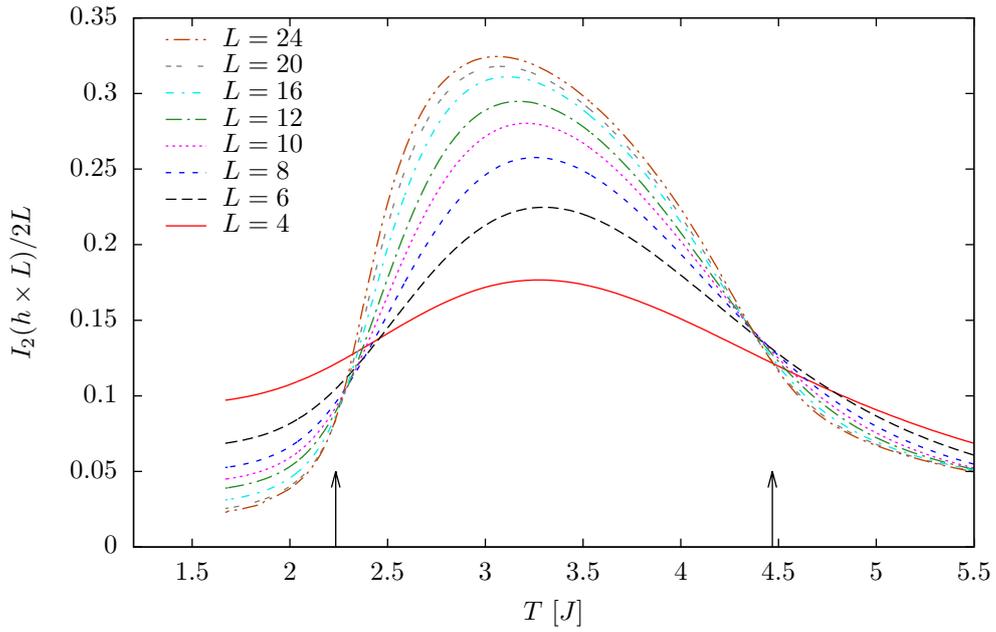
(b) Rényi mutual information, divided by the boundary length $|\partial A| = 2L$

Figure 5.6: Rényi mutual information of order 2 where the region A is one half ($h = \frac{L}{2}$ in Figure 5.1(a)) of the 2D square lattice with different length L ; $\Delta = 4$. The lower panel shows the same data as the upper one, but divided by the boundary length $|\partial A| = 2L$. The arrows in the lower panel indicate the critical temperature T_c and $2T_c$.

The Rényi mutual information compares the Rényi entropies of the region A and its complement B with the joint entropy, see eq. (5.23). In Figure 5.5, we thus show the entropy of the complement as well with a dashed line for each shape of A . Note that the complement of a region with the shape $h \times L$ depicted in Figure 5.1(a) has the same form with $h' = L - h$, i.e., it is identical for $h = \frac{L}{2}$. Once the Rényi entropies have been calculated, it is straightforward to take the difference to obtain the Rényi mutual information $I_2(A : B) = S_2(A) + S_2(B) - S(A \cup B)$. The result is depicted in Figure 5.6(a) for different system sizes.

At large temperatures, the mutual information decays as $\frac{1}{T^2}$: The infinite temperature limits of the entropies cancel as one can check straightforwardly. Further, recall that the energy decays as β for large temperatures; and the integration $\int_0^\beta d\tilde{\beta}$ according to eq. (5.26) leads to the behavior with $\beta^2 \equiv \frac{1}{T^2}$. When the system is cooled down, the mutual information increases up to a maximum value at $T \approx 3J$ and decreases at lower temperatures. Below that maximum, the mutual information rapidly drops to smaller values. If we ignore the quantum entanglement at zero temperature for a moment, we see that the classical, thermodynamic contribution of the entropies $S_m(A) = \log(\Omega_0)$ from the degeneracy of the ground state (valid in any (non-trivial) region A) leads to the same value for the mutual information $I_m(A) = \log(\Omega_0)$. However, the entanglement entropy of the ground state follows the area law eq. (5.9); thus it will dominate over this constant contribution for large systems.

More generally, the mutual information increases monotonically with the system size (i.e., L) at any fixed temperature. We have constructed the mutual information in such a way that the bulk contribution of the individual entropies $S(A)$ and $S(B)$ cancels with the one of the joint entropy $S(A \cup B)$ since we are interested in the correlations between the two regions only. Consequently, we were able to derive the area law eq. (5.18), i.e., the leading scaling with the system size is given by the surface area $|\partial A|$ of the boundary between the regions – in 2D clearly proportional to L . The region A may also have corners. The number of corners, which we denote with $n_c(A)$, depends only on the shape, not on the size of the region; thus we expect a constant part in the mutual information, which is independent of L and proportional to $n_c(A)$. For example, the squares in Figure 5.1(b) have $n_c(h \times h) = 4$ corners. The strips in Figure 5.1(a) have no corners, $n_c(h \times L) = 0$, i.e., this shape avoids these sub-leading contributions. However, we have already seen above that the symmetry breaking, i.e., the degeneracy of the ground state, gives rise to a constant $\log(\Omega_0)$, independent of L . On these general grounds, we expect the following finite size scaling of the mutual information in 2D [86]:

$$I_m(A : B) = |\partial A| \cdot a_m + n_c(A) \cdot b_m + c_m. \quad (5.36)$$

The coefficients a_m, b_m and c_m for the different geometrical contributions are unknown and in general depend on the temperature and the order m of the mutual information. In Figure 5.6(b), the rescaled mutual information $I_m(A : B) / |\partial A|$ is shown such that the curves approach the coefficient of the area law $a_m[T]$ in the thermodynamic limit $L \rightarrow \infty$.

Remarkably, the finite size behavior of these curves changes direction twice, such that two crossings occur between each pair of two curves; one of them near T_c and another one near $2T_c$. More generally, the crossing at higher temperatures occurs near mT_c , as we will discuss in the following along the lines of Ref. [85, 86]. Since the region A considered in Figure 5.6(b) has no corners, we can trace these crossings back to two sign changes in c_m ,

which occur at T_c and mT_c . As we have discussed above, in the completely ordered phases, $T < T_c$, the symmetry breaking leads to $c_m(T \lesssim T_c) = \log(\Omega_0) > 0$. As a consequence, we see that the finite- L curves in Figure 5.6(b) approach the thermodynamic limit from above in this temperature regime.

In the regime $T_c < T < mT_c$ we observe the contrary behavior, so we expect that c_m is negative in this regime. To understand this, we first express the mutual information with eq. (5.26) in terms of the partition functions:

$$I_m(A : B) = \frac{1}{1-m} \left(\log(\mathcal{Z}[A, m, T]) + \log(\mathcal{Z}[B, m, T]) - m \log(\mathcal{Z}[T]) - \log(\mathcal{Z}[A \cup B, m, T]) \right). \quad (5.37)$$

Recall that the operator strings in the replica geometry of $\mathcal{Z}[A, m, T]$ are connected in region A , see Figure 5.4. This leads to an effective temperature $\frac{T}{m}$ in region A – for the derivation of eq. (5.34) we have seen the extremal case $\mathcal{Z}[A \cup B, m, T] \equiv \mathcal{Z}[B = \emptyset, m, T] = \mathcal{Z}[\frac{T}{m}]$, while the unconnected copies in region B are at the temperature T . From an algorithmic point of view, the temperature only enters in the acceptance probabilities of (diagonal) operators, such that it determines the number of operators in the operator string. The system orders when the average number of operators per site reaches a certain level, since the operators lead to correlations of the neighboring sites during the loop updates. Thus, the spins in region A are effectively at the temperature $\frac{T}{m}$ and order at higher temperatures due to the larger period in imaginary time direction.

To estimate the values of the partition functions in eq. (5.37), we ignore⁹ their ‘unphysical’, divergent part such that they just count the number of thermally accessible states. Then we have $\mathcal{Z}[A \cup B, m, T] = \mathcal{Z}[\frac{T}{m}] = \Omega_0$ in the ordered phase for $T < mT_c$. In the unordered phase, each of the $N = |A| + |B|$ spins has $\sigma = 2$ possible values, but a finite correlation length ξ effectively reduces the number of degrees of freedom to N/ξ^D in D dimensions, such that the partition function is $\mathcal{Z}[T] = \log(\sigma^{N/\xi^D})$. In the replica geometry for the simulation of $\mathcal{Z}[A, m, T]$, the spins (anti)align in region A to one of the Ω_0 states. The sites of region B near the boundary in a shell of thickness $\approx \xi$ will feel this ordering as well, which reduces the number of free degrees in region B to $(|B| - |\partial A| \xi)/\xi^D$ in each of the m copies, such that $\mathcal{Z}[A, m, T] = \Omega_0 \sigma^{m(|B| - |\partial A| \xi)/\xi^D}$. As expected, the bulk contributions cancel in the mutual information and we obtain [85]:

$$\begin{aligned} I_m(A : B) &= \frac{1}{1-m} \left(\log \left(\Omega_0 \sigma^{m(|B| - |\partial A| \xi)/\xi^D} \right) + \log \left(\Omega_0 \sigma^{m(|A| - |\partial A| \xi)/\xi^D} \right) \right. \\ &\quad \left. - m \log \left(\sigma^{N/\xi^D} \right) - \log(\Omega_0) \right) \\ &= \frac{1}{1-m} \left(\log(\Omega_0) - \frac{2m |\partial A|}{\xi^{D-1}} \log(\sigma) \right). \end{aligned} \quad (5.38)$$

The term independent of L has a negative sign since $m > 1$, which explains the finite size behavior of Figure 5.6(b) in this temperature regime.

⁹ Formally, we achieve this by measuring all energies relative to the ground state energy: $\mathcal{H} \rightarrow \mathcal{H} - E_{\text{gs}}$. It is well known that such a constant shift does not influence the physical behavior of a system, as we have seen above, the prefactors from this shift cancel.

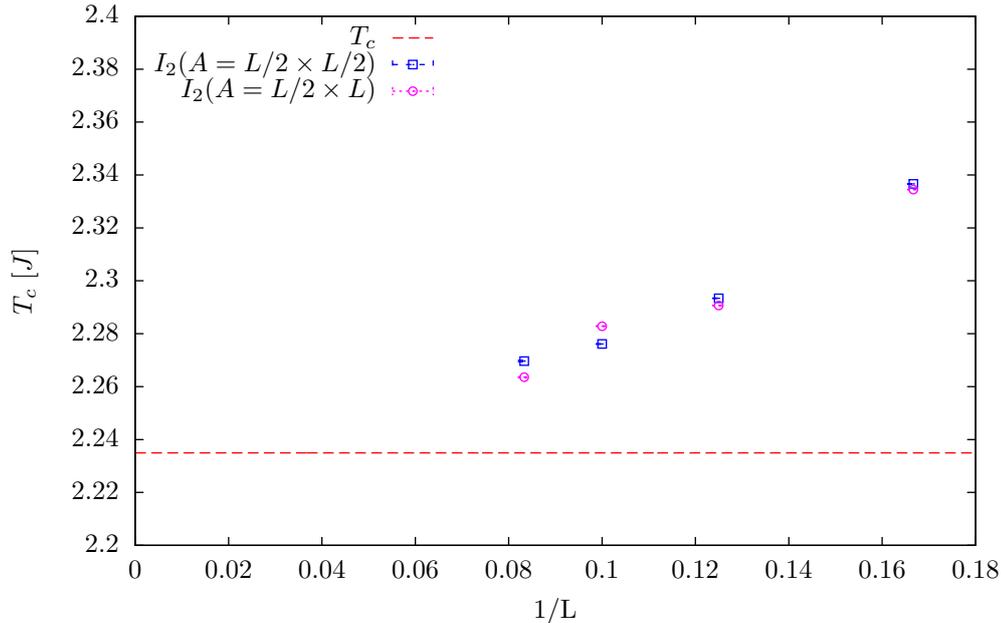


Figure 5.7: Shift of the crossing (near T_c) between curves of length L and $2L$ in Figure 5.6(b) as a function of $1/L$. The red dashed line shows the critical temperature obtained from the crossings of the Binder cumulant.

For temperatures above mT_c , the spins in region A are disordered as well, but with a different correlation $\zeta := \xi(T/m) \gg \xi(T)$ larger than in region B . Note that for $\Delta = 4$ the off-diagonal operators are strongly suppressed, and additionally not many operators appear in the operator string, such that the spin state on a single site changes quite rarely in the imaginary time direction. Thus, the above argument for the scaling with the boundary is still valid. In analogy to above we obtain:

$$\begin{aligned}
 I_m(A : B) &= \frac{1}{1-m} \left(\log \left(\sigma^{|A|/\zeta^D + m(|B| - |\partial A| \xi)/\xi^D} \right) + \log \left(\sigma^{|B|/\zeta^D + m(|A| - |\partial A| \xi)/\xi^D} \right) \right. \\
 &\quad \left. - m \log \left(\sigma^{N/\xi^D} \right) - \log \left(\sigma^{N/\zeta^D} \right) \right) \\
 &= \frac{1}{1-m} \left(-\frac{2m |\partial A| \xi}{\xi^D} \log(\sigma) \right). \tag{5.39}
 \end{aligned}$$

For correlations ξ smaller than the lattice constant $a \equiv 1$ we should remove the ξ^D in the denominator of the exponents (since the correlations do not reduce the degrees of freedom further), but not in the numerator. This leads to a decay in the same fashion as ξ at large temperatures, as one can expect from the definitions of the mutual information.

Clearly, these heuristic arguments are not valid close to a phase transition when $\xi \gg L$. Thus, there are significant corrections to the above scaling form near T_c and mT_c [86]. These corrections lead to a shift such that the crossings occur not exactly at T_c as it is the case for the Binder cumulant (cf. section 4.3.1). The crossings between curves corresponding to L and $2L$ are shown in Figure 5.7. Unfortunately, we were not able to obtain the crossings with the necessary accuracy and for large enough systems to make a reliable extrapolation to $L \rightarrow \infty$ possible. Comparison to literature [87] reveals a

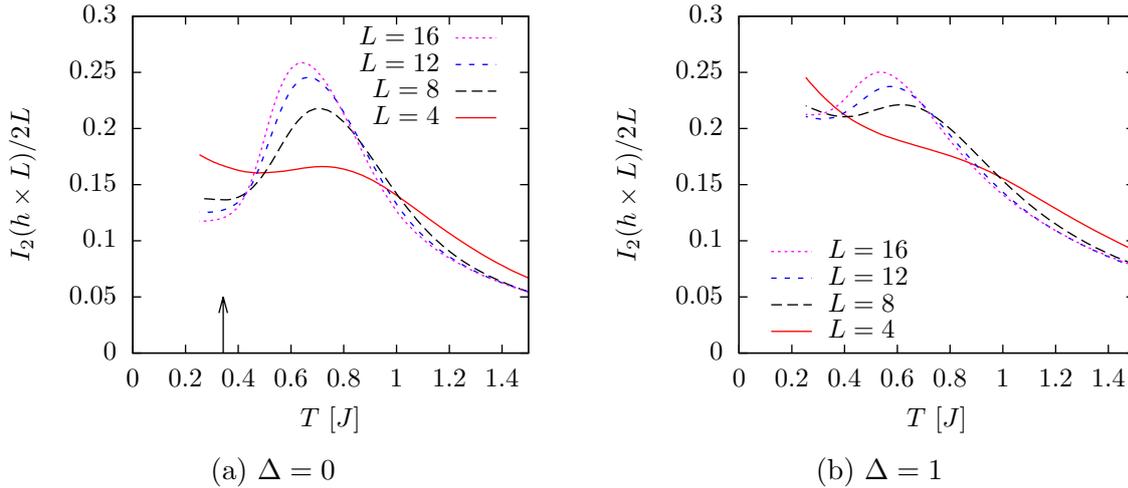


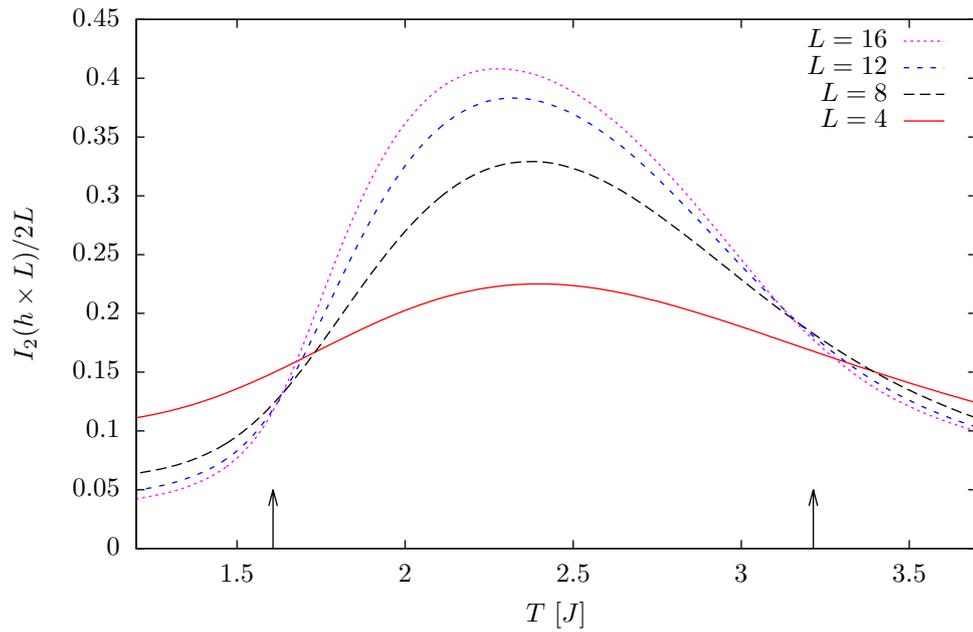
Figure 5.8: Rényi mutual information of order 2, divided by the boundary length $|\partial A| = 2L$, where the region A is one half ($h = \frac{L}{2}$ in Figure 5.1(a)) of the 2D square lattice with different length L . Panel (a) shows the quantum XY model for $\Delta = 0$, which obeys a Kosterlitz-Thouless transition at the temperature $T_{KT} = 0.3803(8)$ (indicated by the arrow) [88]. Panel (b) shows the $SU(2)$ symmetric Heisenberg model, $\Delta = 1$.

systematic error – probably from the integration – which is small yet too large enough that the extraction of the crossings is unreliable and prevents us from a detailed scaling analysis at T_c .

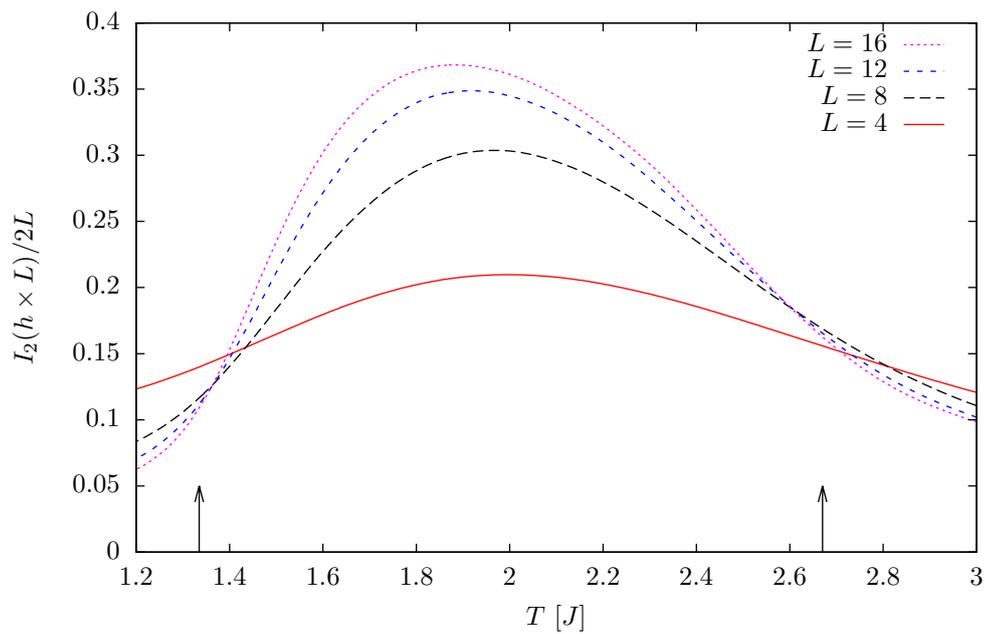
Remarkably, one can observe similar crossings in the XY model at $\Delta = 0$ [1], which is depicted in Figure 5.8(a). In other words, the mutual information can detect the Kosterlitz-Thouless transition, too – a transition to a quasi-long-range ordered phase with a power law decay of correlations. While one can use the spin stiffness to determine the transition temperature [88], it has no local order parameter. It is a big advantage of the mutual information that it does not rely on the definition of an order parameter, but can be used to detect nontrivial phase transitions, based on the very fundamental grounds of information theory. However, the crossings seem to settle at some temperature above the actual Kosterlitz-Thouless temperature, and the crossings at higher temperature are weakly distinct and can only hardly be localized. Note that we observe crossings for $SU(2)$ symmetric couplings (i.e., $\Delta = 1$) as well, see Figure 5.8(b). But in that case the (local) maximum in $I_2(A : B)/|\partial A|$ strongly shifts – as well as the crossings themselves – to lower temperatures with increasing L . Any way, the finite size dependence is clearly too strong to allow a reliable extrapolation to the thermodynamic limit here.

The use of the replica trick in Monte Carlo simulations is not restricted to quantum models, but has also been successfully applied to the Ising model and the classical XY model by Iaconis *et al.* [85]. They observed crossings in these models as well, and – since much larger system sizes can be reached – even a reliable extrapolation $L \rightarrow \infty$ of the crossings is possible.

We have also calculated the mutual information on the spatially anisotropic lattices, i.e., weakly coupled chains and dimers; results for $J' = 0.5$ are shown in Figure 5.9. All in all, we observe a very similar behavior as in the spatially isotropic lattice, and quite



(a) Coupled chains, see Figure 2.2(a)



(b) Coupled dimers, see Figure 2.2(b).

Figure 5.9: Rényi mutual information of order 2, divided by the boundary length $|\partial A| = 2L$, where the region A is one half ($h = \frac{L}{2}$ in Figure 5.1(a)) of the 2D square lattice with different length L ; for spatially anisotropic couplings, exemplary for $J' = 0.5$. The arrows indicate the T_c and $2T_c$, obtained from the Binder cumulant.

clear crossings slightly above the corresponding T_c and below mT_c . This comes as no great surprise since the anisotropic lattices have the same well-behaved order-disorder transition; the arguments given above for the scaling in the case of a symmetry broken, degenerate ground state still go through.

5.3 Ratio Trick

The ratio trick is based on the ideas of the replica trick discussed in the previous section; it uses the same geometry with the replicas depicted in Figure 5.4 in order to evaluate $\text{Tr} \{\hat{\rho}_A^2\} = \mathcal{Z}[A, m, T] / (\mathcal{Z}[T])^m = \mathcal{Z}[A, m, T] / \mathcal{Z}[\emptyset, m, T]$. However, the ratio of the partition functions is evaluated directly such that no integration is necessary. In their pioneering paper, Hastings *et al.* [84] used the SWAP operators, defined in the valence bond basis at zero temperature, to evaluate this ratio. Humeniuk *et al.* [80] explained how to calculate the ratio of the partition functions with general (quantum) Monte Carlo simulations at finite temperatures. In any case, a naive evaluation of the ratio is not successful: the ratio gets very small with increasing size of region A , which leads to problems in the ergodicity and efficiency of the simulation; we will discuss the reasons for this in detail later. These problems are cured by a simple trick [84]: one increases the size of the region A in steps, only a few sites at once. Denoting increasing regions with $A_i, i \in \{0, 1, \dots, I\}$ such that $A_0 = \emptyset$ and $A_I = A$, we can summarize this trick in the following, trivial equation:

$$\frac{\mathcal{Z}[A, m, T]}{\mathcal{Z}[\emptyset, m, T]} \equiv \frac{\mathcal{Z}[A_I, m, T]}{\mathcal{Z}[A_0, m, T]} = \frac{\mathcal{Z}[A_1, m, T]}{\mathcal{Z}[A_0, m, T]} \cdot \frac{\mathcal{Z}[A_2, m, T]}{\mathcal{Z}[A_1, m, T]} \dots \frac{\mathcal{Z}[A_I, m, T]}{\mathcal{Z}[A_{I-1}, m, T]}. \quad (5.40)$$

Before we go into details why this trick is necessary and how it works, we discuss the evaluation of the ratios $R_m(A_i, A_{i-1}) := \mathcal{Z}[A_i, m, T] / \mathcal{Z}[A_{i-1}, m, T]$ in SSE, following the discussion of Ref. [80].

Recall that we discussed the mapping of the SSE to the classical configuration space \mathcal{C} which consists of the (initial) basis states $|\alpha\rangle$ and the operator strings. A weight w_ν is associated with each configuration ν . For convenience, we change the notation slightly: we use w_ν for the unnormalized weights and explicitly write the normalization with the sum of the weights, i.e., the partition function $\mathcal{Z} = \sum_{\nu \in \mathcal{C}} w_\nu$. Further, we also consider configurations in which the operator string does not fulfill the boundary conditions and simply define that they have weight 0. Similarly, the configuration space for the replica geometry, which we denote with $\mathcal{C}_m(A)$, consists of the basis states $\{\alpha_0, \beta_0, \dots, \alpha_m, \beta_m\}$ and m operator strings with the corresponding boundary conditions as depicted in Figure 5.4. We can write

$$\mathcal{Z}[A, m, T] = \sum_{\nu \in \mathcal{C}_m(A)} w_\nu^A. \quad (5.41)$$

The key idea of Humeniuk and Roscilde [80] was to use the extended ensemble space $\mathcal{C}_m(A_i) \cup \mathcal{C}_m(A_{i-1})$ for the evaluation of the ratio $R_m(A_i, A_{i-1})$. Note that we have an one to one correspondence of the configurations ν between the two configurations spaces $\mathcal{C}_m(A_i)$ and $\mathcal{C}_m(A_{i-1})$ – at least when including the configurations with vanishing weight due to wrong boundary conditions: we simply have to identify the initial basis states

$|\alpha_{\bar{m}}\rangle |\beta_{\bar{m}}\rangle$ for the different regions with each other, which is straightforward since we use the product basis over the single sites, see eq. (5.6):

$$|\alpha_{\bar{m}}\rangle |\beta_{\bar{m}}\rangle = \prod_{i \in A} |\sigma_i\rangle \prod_{j \in B} |\sigma_j\rangle = \prod_{i \in A \cup B} |\sigma_i\rangle. \quad (5.42)$$

This suggests to sample between the two configuration spaces with an update, which changes the region from A_{i-1} to A_i via the identification of configurations. Such an update can be performed during each Monte Carlo sweep in addition to the usual updates. However, one should respect the detailed balance condition (3.10), which is simplest done with the Metropolis scheme with acceptance probability $P^{\text{acc.}}(A_{i-1} \rightarrow A_i) = \min(1, w_\nu^{A_i}/w_\nu^{A_{i-1}})$, see eq. (3.20).

In our case of SSE, the ratio of these weights takes a particularly simple form: it is either¹⁰ 1 or 0. The reason is simple: the update does not change the configuration, but only the boundary condition, i.e., if a configuration ν of initial states and operator strings is compatible with the boundary conditions for both A_i and A_{i-1} , the weight does not change during the update (otherwise, the configuration for the other size of A has weight 0). This means we can always perform the update, if the configuration is compatible with the boundary conditions. To be specific, consider an update $A_{i-1} \rightarrow A_i$, where the region A is enlarged by the sites $i \in A_i \cap B_{i-1}$. A configuration is compatible with the boundary conditions in both cases if and only if the initial states on these sites $i \in A_i \cap B_{i-1}$ – initially in the complement B_i – coincide on all m replicas: only then the operator string is continuous for the larger A_i , see Figure 5.4. The same condition is necessary for the opposite direction $A_i \rightarrow A_{i-1}$; in this case to ensure the periodicity within the replicas in region B .

The average time a (Markov chain) Monte Carlo algorithm spends in a configuration ν is proportional to the weight¹¹. Thus, the ratio of the partition functions is simply the ratio of the times \mathcal{T}_A (i.e., number of sweeps) the algorithm spends in each of the configuration spaces $\mathcal{C}_m(A)$ [80]:

$$R_m(A_i, A_{i-1}) = \left\langle \frac{\mathcal{T}_{A_i}}{\mathcal{T}_{A_{i-1}}} \right\rangle_{\mathcal{C}_m(A_i) \cup \mathcal{C}_m(A_{i-1})}. \quad (5.43)$$

In fact, we do not even need to change between the two configurations spaces $\mathcal{C}_m(A_{i-1})$ and $\mathcal{C}_m(A_i)$, but it suffices to measure how often this would be possible in two separate simulations, one in $\mathcal{C}_m(A_{i-1})$ and one in $\mathcal{C}_m(A_i)$. Therefore, we define the ‘observable’ $X_\nu(A)$ as 1 if the configuration ν is compatible with the boundary conditions corresponding to region A , and 0 otherwise. By definition, the weighted average over $\mathcal{C}_m(A)$ includes only configurations ν with $X_\nu(A) = 1$ since otherwise the weight vanishes. Moreover, only configurations compatible with (the boundary conditions of) both regions A_{i-1} and

¹⁰ Strictly speaking, the weight in the denominator can also be 0, but since this weight corresponds to the configuration *before* the update, the algorithm will never be in such a configuration and we do not have to take care of that case.

¹¹ More precisely, the number of sweeps \mathcal{T}_ν which the algorithm spends in configuration ν divided by the total number of sweeps \mathcal{T} is the normalized weight.

A_i contribute to the average of $X_\nu(A_i)$ over the configuration space $\mathcal{C}_m(A_{i-1})$:

$$\begin{aligned} \langle X_\nu(A_i) \rangle_{\mathcal{C}_m(A_{i-1})} &= \frac{1}{\mathcal{Z}[A_{i-1}, m, T]} \sum_{\nu \in \mathcal{C}_m(A_{i-1})} w_\nu^{A_{i-1}} X_\nu(A_i) \\ &= \frac{1}{\mathcal{Z}[A_{i-1}, m, T]} \sum_{\nu \in \mathcal{C}_m(A_{i-1})} w_\nu^{A_{i-1}} X_\nu(A_{i-1}) X_\nu(A_i) =: \frac{X_\nu^\Sigma(A_{i-1}; A_i)}{\mathcal{Z}[A_{i-1}, m, T]}. \end{aligned} \quad (5.44)$$

Due to the one to one correspondence of the configuration spaces $\mathcal{C}_m(A_{i-1})$ and $\mathcal{C}_m(A_i)$, and since the corresponding weights are equal if compatible with the boundary conditions, $w_\nu^{A_{i-1}} X_\nu(A_{i-1}) X_\nu(A_i) = w_\nu^{A_i} X_\nu(A_{i-1}) X_\nu(A_i)$, we obtain the same sum $X_\nu^\Sigma(A_i; A_{i-1}) = X_\nu^\Sigma(A_{i-1}; A_i)$ also in the converse direction:

$$\langle X_\nu(A_{i-1}) \rangle_{\mathcal{C}_m(A_i)} = \frac{1}{\mathcal{Z}[A_i, m, T]} \sum_{\nu \in \mathcal{C}_m(A_i)} w_\nu^{A_{i-1}} X_\nu(A_i) = \frac{X_\nu^\Sigma(A_i; A_{i-1})}{\mathcal{Z}[A_i, m, T]}. \quad (5.45)$$

Thus, the sums cancel in the ratio of the averages and only the normalizations, i.e., the partition functions, remain:

$$R_m(A_i, A_{i-1}) \equiv \frac{\mathcal{Z}[A_i, m, T]}{\mathcal{Z}[A_{i-1}, m, T]} = \frac{\langle X_\nu(A_i) \rangle_{\mathcal{C}_m(A_{i-1})}}{\langle X_\nu(A_{i-1}) \rangle_{\mathcal{C}_m(A_i)}}. \quad (5.46)$$

With the above framework at hand we can understand the problem with the naive ansatz to calculate the ratio of the partition functions between regions which differ a lot, i.e., directly between $A_0 = \emptyset$ and another large region $A_I = A$. Such a simulation would immediately give the Rényi entropy $S_m(A) = \frac{1}{1-m} \log(R_m(A, \emptyset))$. But we can also invert this equation to obtain the scaling of the ratios: $R_m(A, \emptyset) = \exp((1-m)S_m(A))$. Assuming an area law in $2D$, $S_m(A) \propto L$, we see that the ratio is exponentially small in L ; the scaling of $S_m(A)$ with the bulk at finite temperatures makes it even worse. This means that, if we allow to switch between the two configurations, the algorithm will spend only very few Monte Carlo sweeps in the configuration space $\mathcal{C}_m(A)$, but will almost always be in the space $\mathcal{C}_m(A = \emptyset)$. This leads to very long autocorrelations for the measurement of the ratio, which becomes even more clear in the formulation with the X_m : Obviously, since X_ν takes only the values 0 and 1, the average is in the interval $[0, 1]$. Thus, as one can see from the scaling argument, at least the average in the nominator of eq. (5.46) must be small. In other words, there are only very few states compatible with the boundary conditions for different regions – this is also intuitively clear from the condition that the initial states of the replicas have to coincide on the sites by which the region A changes, i.e., for $A_0 = \emptyset$ on the complete region A . A small average of X means that during the simulation X has to be 0 for a long time before the next 1 may appear, which obviously leads large a autocorrelation and thus large statistical errors¹². Note that this problem may even appear if the ratio $R_m(A, \emptyset)$ is not large, e.g., for non-critical 1D systems in the ground state where the entropy is independent of L : the canceling sum X_ν^Σ may still be much smaller than the partition functions, such that both expectation values of X_ν are small.

¹² This transfers directly to the algorithm in the extended phase space $\mathcal{C}_m(\emptyset, m, T) \cup \mathcal{C}_m(A, m, T)$, since X measures the probability for the update changing the region A in one Monte Carlo sweep.

4	8	12	16
3	7	11	15
2	6	10	14
1	5	9	13

(a) $h \times L$

13	14	15	16
7	8	9	12
3	4	6	11
1	2	5	10

(b) $h \times h$

Figure 5.10: Scheme for increment of region A on a 4×4 lattice: For an increment of just one site in each step, A_i contains the sites marked with $1, \dots, i$.

It is also clear from these considerations that the intermediate steps of the increment trick cure the problem. We use different schemes to increase the region A , depending on whether we are interested in shapes $h \times L$ or $h \times h$; they are depicted in Figure 5.10. However, this makes $\mathcal{O}(|A|) = \mathcal{O}(N)$ separate simulations necessary for each temperature, i.e., the computational cost increases dramatically. On the other hand, the ratio trick does not require the fine grid in the temperature which is necessary for the integration discussed in the previous section. Especially to reach to reach low temperatures, in particular the ground state entropies, the ratio trick may thus perform better than the integration method discussed above. As mentioned above, one may also combine the ratio trick with the integration method: one can use the results of the ratio trick as starting point for the integration. However, we have not done it for this thesis due to time restrictions.

Error due to non-ergodicity

Unfortunately, we are faced with another problem at low temperatures in the ordered phase: our algorithm is not ergodic for large Δ . In the correct thermodynamic average, each of the two degenerate ground states $|\text{gs.}; \uparrow\rangle$ and $|\text{gs.}; \downarrow\rangle$ is equally probable for $T \rightarrow 0$. But our algorithm spontaneously chooses one of them, exactly as a single physical system would do. We argued in the beginning of chapter 4 that the missing ergodicity does not influence most variables due to the symmetry between the ground states. But as we have seen, the symmetry breaking leads to additive constants in the entropies, i.e., we obtain wrong results for the entropies in the ordered phase where the algorithm gets non-ergodic.

Suppose we simulate two independent replicas of the system for $A = \emptyset$. In that case, both replicas are actually independent of each other. Since we do not apply a magnetic field, both of them will have the first spin point up or down with 50%. Analytically, we obtain immediately the correct thermodynamic average $\langle X_\nu(A=1) \rangle_{\mathcal{C}_m(A=\emptyset)} = 0.5$ since the spins on site 1 are parallel in two out of four equally probable cases. But in a single SSE simulation the sign of the staggered magnetization M_s in each of the replicas basically is fixed due to the missing ergodicity. Quantum fluctuations may flip the spin on site 1 during the simulation with some probability p which, however, is quite small for $\Delta = 4$, $p \approx 1.5\%$. We obtain exactly this value from a single SSE simulation for $X_\nu(A = \{1\})$ if the sign of M_s differs in the replicas; otherwise – if the signs are equal – we obtain $1 - p$. In the average of many simulations we would obtain the correct value 0.5. This single simulation of $X_\nu(A = \{1\})$ for $A = \emptyset$ leads to the largest error due to the non-ergodicity.

If A is non-empty, the two replicas are not independent any more and the exact value can

not be obtained from such simple arguments. Note that configurations with different signs of M_s on the two replicas have much smaller weight: they require either parallel spins along the boundary of A , or quantum fluctuations (i.e., off-diagonal operators) within the complete region A . These configurations correspond to cross terms in the square of $\hat{\rho}_A = \frac{1}{2} |\text{gs}; \uparrow\rangle \langle \text{gs}; \uparrow| + \frac{1}{2} |\text{gs}; \downarrow\rangle \langle \text{gs}; \downarrow|$.

Although it basically might be possible to calculate the relative factor of weights between configuration with different or equal sign of M_s in the replicas and manually add an update for ergodicity, we will here choose a different way: We calculate the entanglement entropy of a single state $|\text{gs}; \uparrow\rangle$ only. Therefore, we ensure that the sign of M_s is the same on all replicas, i.e., we bias our simulation by setting the initial state to the same classical Néel state in all replicas and start the simulation at low temperatures in the ordered phase. The results of such a biased calculation are shown in Figure 5.11(b). It is clear, that we can not capture the additive constant $\log(\Omega_0) \equiv \log(2)$ due to the degeneracy of the ground state. However, we basically capture the scaling of the entropy with different sizes and shapes of the region A – up to contributions from cross terms between the different states. The upper figure Figure 5.11(a) shows the entropy for the intermediate regions also. The step structure for the red curve and the plateau for black curve arise from the shapes of the A_i : the jumps in the red curve occur when we increase the boundary lengths by adding a single site to a rectangular shape. Similarly, the outliers of the plateau of the black curve correspond to the rectangular shapes $h \times L$ with boundary lengths $2L$, while the regions of the plateau have boundary lengths $2L + 1$. When A fills the entire lattice, there is no boundary any more; the entanglement entropy of a single state is then 0, as one can see immediately from the definition. Indeed, we obtain that result within the error-bars which also confirms that the temperature was chosen low enough – thanks to the large energy gap, $T = 0.5J$ was sufficiently low. Note also that the errors increase as expected with the region size. We have also performed a perturbation theory for large Δ ; the result eq. (D.9) is an area law, which compares very well to our simulations, as it is shown in Figure 5.11(b).

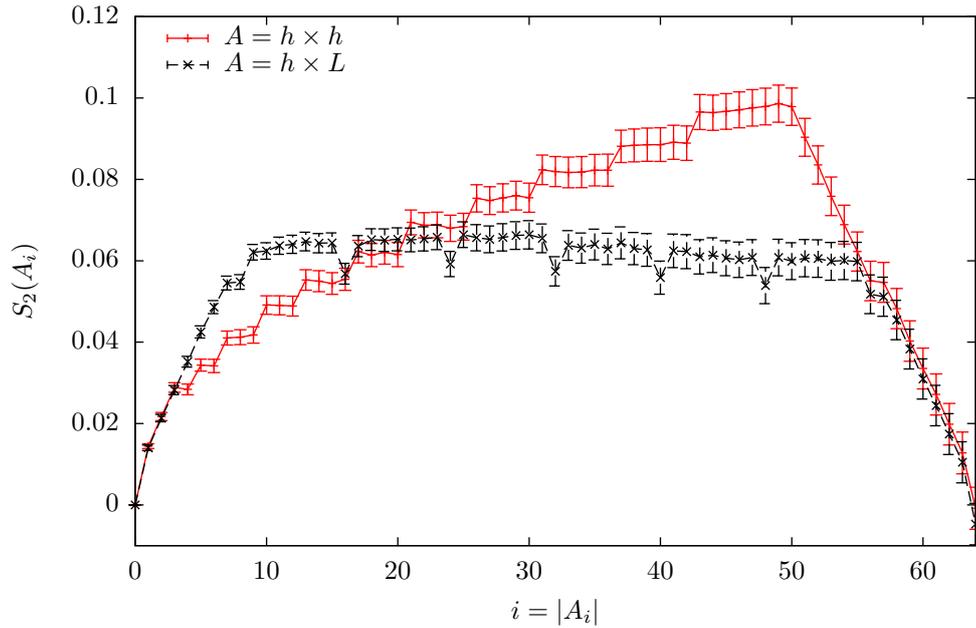
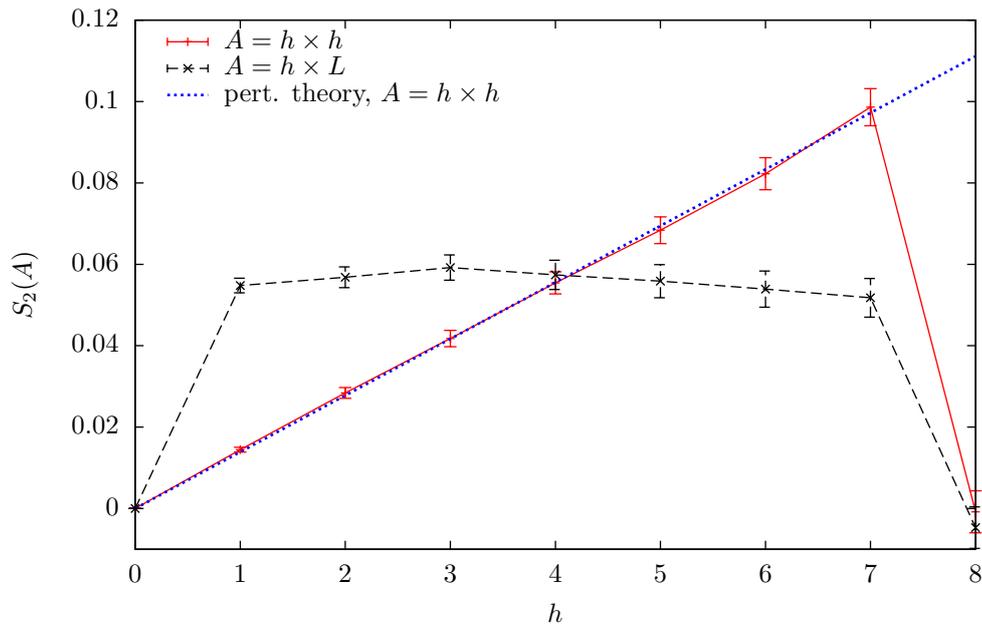
(a) For all intermediate regions A_i (b) For regular shapes of Figure 5.1 vs h

Figure 5.11: Scaling of the Rényi entropy at $T = 0.5$ (i.e., due to the gap mostly in the ground state), $\Delta = 4$ on a 2D square lattice of 8×8 sites, obtained with the ratio trick and with a bias fixing the sign of M_s , see main text. We increased the A_i by just one site in each step, following the scheme indicated in Figure 5.10. In the lower panel, only the points corresponding to rectangular shapes $h \times h$ and $h \times L$ for A are shown versus h . The dotted blue line shows the result of the perturbation theory calculation in appendix D for the boundary lengths of the $h \times h$ regions, but without the additive constant $\log(2)$ from the degeneracy of the ground state.

6 Singlet and Triplet Projectors

In their experiments with ultra-cold quantum gases, Greif *et al.* [2] project on the singlet (S) and triplet (T_0) states on particular bonds with a clever detection scheme in order to measure the spin correlations. In other words, they measure the expectation values of projection operators \hat{P}_b^S and $\hat{P}_b^{T_0}$. In a general setup, we define the singlet and triplet states for any two neighboring sites i and j , i.e., for any bond $b = \langle i, j \rangle$:

$$|S; i, j\rangle \equiv \frac{|\uparrow_i \downarrow_j\rangle - |\downarrow_i \uparrow_j\rangle}{\sqrt{2}}, \quad |T_0; i, j\rangle \equiv \frac{|\uparrow_i \downarrow_j\rangle + |\downarrow_i \uparrow_j\rangle}{\sqrt{2}}. \quad (6.1)$$

Note that a local projection operator – e.g., onto a singlet $\hat{P}_b^S := |S; i, j\rangle \langle S; i, j|$, or analogous onto triplets T_0 and T_{\pm} , respectively – is the identity operator on all other sites but i and j ; it projects only in the local Hilbert space spanned by the two sites. Clearly, the expectation value $\langle \hat{P}_b^S \rangle$, i.e., the fraction of singlets on bond $\langle i, j \rangle$, may depend on the bond in a spatially anisotropic lattice. For example for the geometry of coupled dimers depicted in Figure 2.2(b), Greif *et al.* [2] measured the projections on the strong bonds only, i.e., on the dimers. In a separate measurement of the projections along weak bonds parallel to the dimers, the observed fraction of singlets was as high as the fraction of triplets: along the weak bonds the spins were not correlated.

With these measurements in mind, we wondered whether the singlet and triplet projectors on neighboring sites can be used to detect a phase transition between the ordered and disordered phase. Do they show critical behavior at a phase transition? Do singularities occur at T_c ? We address these questions in this chapter. However, we emphasize that the results of our model calculations can not directly be transferred to the experiment of Greif *et al.* [2]. First of all, the experiment is described by the Hubbard model such that there is a competition between spin and charge excitations, while we restrict ourselves to the Heisenberg model. Secondly, the experiment was done in three dimensions, while we use a 2D lattice. Last but not least, the spin interaction in the experiment is $SU(2)$ symmetric, while we use an anisotropic coupling $\Delta = 4$ as in the previous chapters in order to have the finite temperature transitions. Numerical calculations with direct comparison to the experiment of Greif *et al.* are available in the literature: Sciolla *et al.* [40] compared to DMRG results in 1D, and Imriška *et al.* [41] performed a dynamical cluster approximation in 3D. However, they do not discuss the criticality of the projectors.

6.1 SSE Estimators for Singlet and Triplet Projectors

The projection operators are off-diagonal since the states on which they project – the singlet S and triplet T_0 – are not in the product basis of the local spins. Explicitly, we

can simply multiply their definitions from eq. (6.1) out:

$$\hat{P}_b^S = |S; i, j\rangle \langle S; i, j| = \frac{1}{2} (|\uparrow\downarrow\rangle \langle\uparrow\downarrow| - |\uparrow\downarrow\rangle \langle\downarrow\uparrow| - |\downarrow\uparrow\rangle \langle\uparrow\downarrow| + |\downarrow\uparrow\rangle \langle\downarrow\uparrow|), \quad (6.2)$$

$$\hat{P}_b^{T_0} = |T_0; i, j\rangle \langle T_0; i, j| = \frac{1}{2} (|\uparrow\downarrow\rangle \langle\uparrow\downarrow| + |\uparrow\downarrow\rangle \langle\downarrow\uparrow| + |\downarrow\uparrow\rangle \langle\uparrow\downarrow| + |\downarrow\uparrow\rangle \langle\downarrow\uparrow|). \quad (6.3)$$

Thus, one can not simply apply eq. (3.34) to calculate their expectation values. Fortunately, the projectors are closely related to the bond operators $\mathcal{H}_b = \mathcal{H}_b^{\text{diag}} - \mathcal{H}_b^{\text{offd}}$ in terms of which we formulated the SSE. Taking simple linear combinations of eq. (6.2) and (6.3), we obtain:

$$\hat{P}_b^S + \hat{P}_b^{T_0} = |\uparrow\downarrow\rangle \langle\uparrow\downarrow| + |\downarrow\uparrow\rangle \langle\downarrow\uparrow|, \quad (6.4)$$

$$\hat{P}_b^{T_0} - \hat{P}_b^S = |\uparrow\downarrow\rangle \langle\downarrow\uparrow| + |\downarrow\uparrow\rangle \langle\uparrow\downarrow| = \frac{2}{J_b} \mathcal{H}_b^{\text{offd}}. \quad (6.5)$$

The latter combination is the off-diagonal operator $\mathcal{H}_b^{\text{offd}} = \frac{J_b}{2} (S_i^+ S_j^- + S_i^- S_j^+)$ on the bond $b = \langle i, j \rangle$, while the former combination is a part of $\mathcal{H}^{\text{diag}}$, namely the part acting on anti-parallel spins. Recall that we separately sample the off-diagonal and diagonal actions of \mathcal{H}_b on different states with different vertices, see Table 3.1. The generalization of eq. (3.31) given in appendix A shows that counting the number of vertices in the operator string during an SSE simulation gives – up to a factor β – the corresponding expectation values, e.g.,

$$\langle \mathcal{H}_b^{\text{diag}} |\uparrow\downarrow\rangle \langle\uparrow\downarrow| \rangle = \underbrace{\langle \uparrow\downarrow | \mathcal{H}_b^{\text{diag}} | \uparrow\downarrow \rangle}_{\in \mathbb{R}} \cdot \langle |\uparrow\downarrow\rangle \langle\uparrow\downarrow| \rangle = \frac{1}{\beta} \langle n_b^{\text{diag}}(i = \uparrow, j = \downarrow) \rangle_{SSE}, \quad (6.6)$$

where $n_b^{\text{diag}}(i = \uparrow, j = \downarrow)$ is the number of vertices on bond b in the operator string which correspond to the matrix elements $\langle \uparrow\downarrow | \mathcal{H}_b^{\text{diag}} | \uparrow\downarrow \rangle$. This leads to the following estimator for $\hat{P}_b^S + \hat{P}_b^{T_0}$:

$$\langle \hat{P}_b^S + \hat{P}_b^{T_0} \rangle = \frac{\langle n_b^{\text{diag}}(i = \uparrow; j = \downarrow) + n_b^{\text{diag}}(i = \downarrow; j = \uparrow) \rangle_{SSE}}{\beta \left(\langle \uparrow\downarrow | \mathcal{H}_b^{\text{diag}} | \uparrow\downarrow \rangle + \langle \downarrow\uparrow | \mathcal{H}_b^{\text{diag}} | \downarrow\uparrow \rangle \right)}. \quad (6.7)$$

This estimator involves counting all diagonal operators acting on anti-parallel spins of bond b . The normalization with β and the weights of these vertices can be evaluated post-simulation.

Care has to be taken with the sign of the off-diagonal operators: since $\mathcal{H}_b = \mathcal{H}_b^{\text{diag}} - \mathcal{H}_b^{\text{offd}}$, we have

$$\langle \hat{P}_b^S - \hat{P}_b^{T_0} \rangle = \left\langle -\frac{2\mathcal{H}_b^{\text{offd}}}{J_b} \right\rangle = +\frac{2 \langle n_b^{\text{offd}} \rangle_{SSE}}{J_b \beta}. \quad (6.8)$$

We immediately see that there are more singlets than triplets since $n_b^{\text{offd}} \geq 0$, in agreement with our physical intuition as the energy of singlets is lower than the one of triplets. We evaluate the two estimators eq. (6.7) and (6.8) during SSE simulations and calculate the expectation values of the singlet and triplet projectors post-simulation.

Since we will often use the two linear combinations of the singlet and triplet projectors, we denote them by (the notation will be clear in a moment):

$$\hat{P}_b^{\text{ap}} := \hat{P}_b^S + \hat{P}_b^{T_0}, \quad \hat{Q}_b := \hat{P}_b^S - \hat{P}_b^{T_0}. \quad (6.9)$$

Moreover, we identify them as proper physical observables: Since the singlet and triplet states are orthogonal, the sum \hat{P}_b^{ap} is itself a projector of rank two – as one can see from eq. (6.4), it projects onto the space where the spins are anti-parallel in the S^z -basis; hence we chose the notation with ‘ap’ for ‘anti-parallel’. Thus, it is closely related to the nearest neighbor correlators $S_i^z S_j^z$:

$$\hat{P}_b^{\text{ap}} = 2 \left(\frac{1}{4} - S_i^z S_j^z \right). \quad (6.10)$$

On the other hand, the difference of the singlet and triplet projectors (which itself is not a projector) measures the quantum fluctuations (\Rightarrow notation Q), i.e., the spin correlation in the transverse x - and y -directions:

$$\hat{Q}_b = -\frac{2\mathcal{H}_b^{\text{offd}}}{J_b} = -2 (S_i^+ S_j^- + S_i^- S_j^+) = -4 (S_i^x S_j^x + S_i^y S_j^y). \quad (6.11)$$

For $SU(2)$ symmetric couplings, $\Delta = 1$, the expectation values $\langle S_i^x S_j^x \rangle = \langle S_i^y S_j^y \rangle = \langle S_i^z S_j^z \rangle$ clearly coincide and thus both \hat{P}_b^{ap} and \hat{Q}_b (and consequently also \hat{P}_b^S and $\hat{P}_b^{T_0}$) are completely determined by a single correlator $\langle \vec{S}_i \cdot \vec{S}_j \rangle = \frac{1}{3} \langle S_i^z S_j^z \rangle$ [40, 41]. However, for the symmetry broken couplings, $\Delta \neq 1$, they are clearly independent of each other.

Since \hat{P}_b^{ap} is diagonal in our basis, we could basically also use the simple estimator given in eq. (3.34). However, the estimator (6.7) has the advantage that it averages over the propagated states such that the statistical noise is reduced, and the necessary counting is cheap and easy to implement. An additional estimator for the temperature derivative of such estimators is given in eq. (A.7) in appendix A. Nevertheless, a ‘usual’ estimator of the nearest neighbor spin correlations can be used as a cross-check.

Up to now we have written down the projectors for a single bond. We can reduce the statistical noise further with a spatial average over all bonds for which we expect the same results due to spatial symmetries¹. For example, for the weakly coupled chains depicted in Figure 2.2(a), we distinguish between the strong bonds with coupling strength J and the weak bonds with coupling strength J' connecting the chains.

The results for the singlet and triplet projectors \hat{P}_b^S and $\hat{P}_b^{T_0}$ on an isotropic square lattice are depicted in Figure 6.1(a) and compared to the expectation values for a single dimer of two sites, which we discussed in section 2.4. It is eye-catching that even the finite size curves \hat{P}_b^S and $\hat{P}_b^{T_0}$ basically keep themselves in a constant distance; in other words, \hat{Q}_b is almost constant over the shown temperature range, as one can see in the bottom panel (b). Note, however, that at infinite temperatures all states are equally likely and thus both \hat{P}_b^S and $\hat{P}_b^{T_0}$ approach $\frac{1}{4}$, i.e., $\hat{P}_b^{\text{ap}} \rightarrow \frac{1}{2}$ and $\hat{Q}_b \rightarrow 0$. But the energy scale

¹ In experiments, the measurement changes the states, such that one can not simultaneously measure the projectors on two bonds sharing a site. In our SSE simulations, this is not a problem. However, one should be aware that one does not measure the simultaneous action of the projectors $\langle \hat{P}_1 \hat{P}_2 \dots \rangle$ – which is ill-defined in the sense that the order matters – but the sum $\sum_b \langle \hat{P}_b \rangle$.

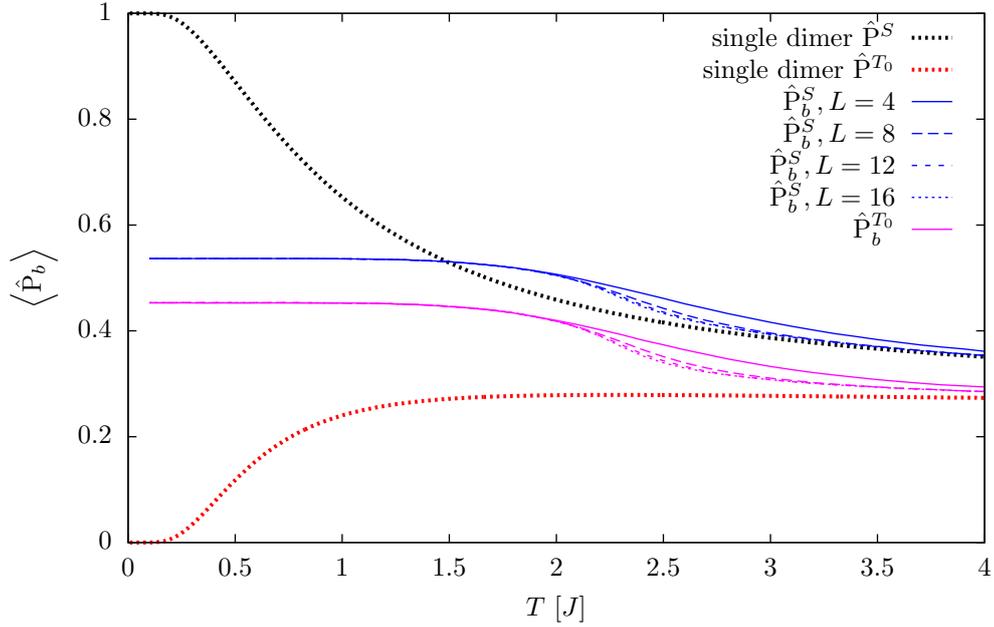
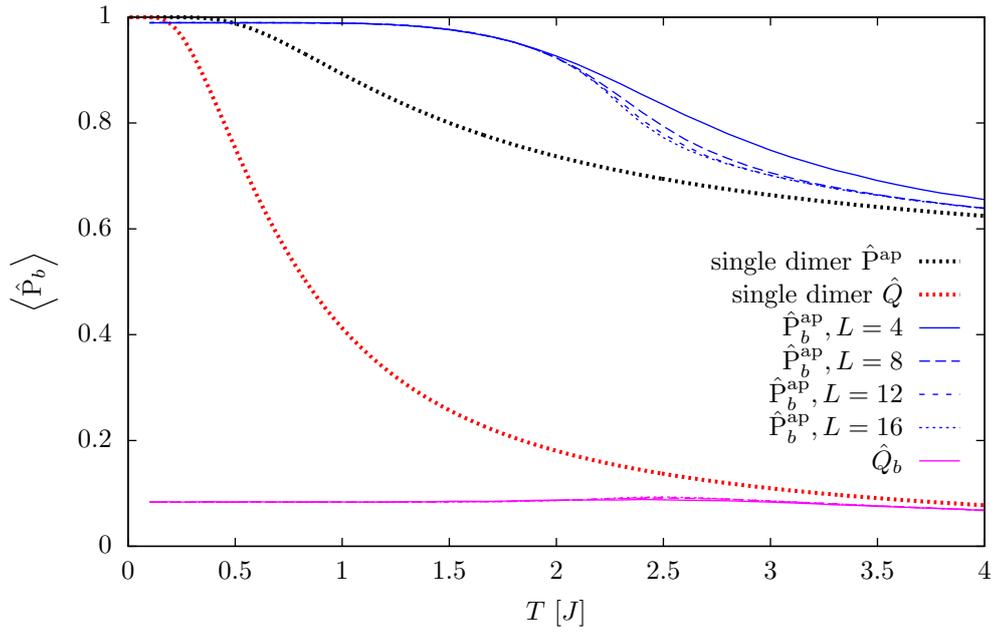
(a) \hat{P}_b^S and $\hat{P}_b^{T_0}$ (b) The linear combinations $\hat{P}_b^{\text{ap}} = \hat{P}_b^S + \hat{P}_b^{T_0}$ and $\hat{Q}_b = \hat{P}_b^S - \hat{P}_b^{T_0}$

Figure 6.1: Singlet and triplet projectors on an isotropic 2D square lattice, $\Delta = 4$. The thick, dotted curves show the expectation values of the projectors on a single dimer, see section 2.4. For each observable, different line styles indicate different lengths $L \in \{4, 8, 12, 16\}$, revealing the finite size dependence.

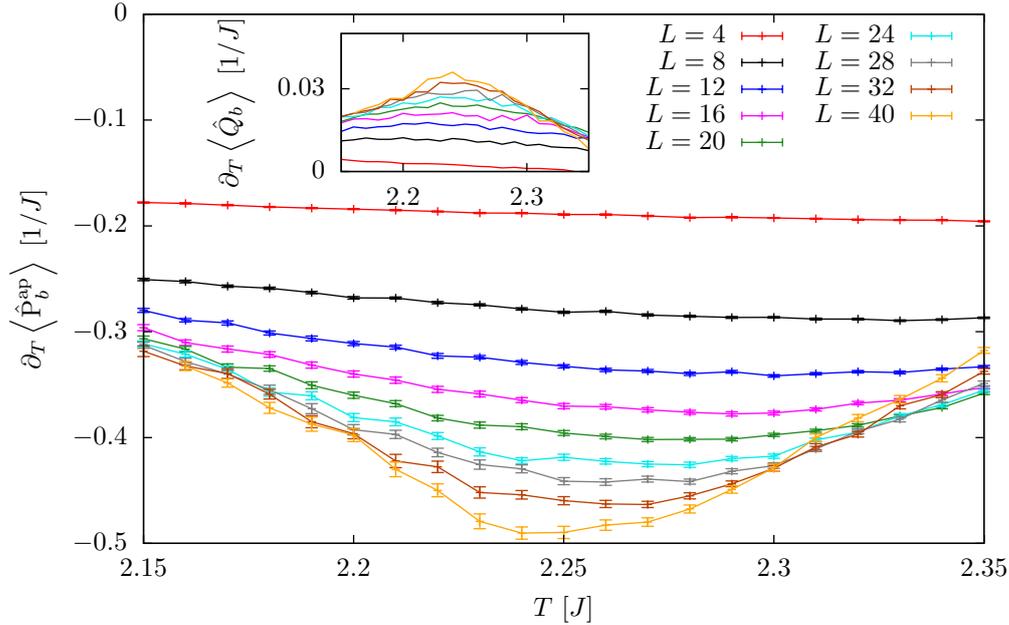


Figure 6.2: Derivative of $\langle \hat{P}_b^{\text{ap}} \rangle$ with respect to the temperature T , on a 2D square lattice, $\Delta = 4$, similar to Figure 6.1(b). The inset shows the temperature derivative of $\langle \hat{Q}_b \rangle$ in the same temperature regime. The (relative) statistical errors in the inset are much larger such that the curves do not look smooth.

on which these limits are approached is large: $E_{T_{\pm}} - E_S = \frac{\Delta+1}{2}J = 2.5J$. It is not a big surprise that the values of \hat{P}_b^S and $\hat{P}_b^{T_0}$ deviate strongly from the calculation with a single dimer in the ordered phase: we expect the spins on two neighboring sites to be in the states $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$, but not in the superposition given by the singlet and triplet states. Indeed, we find that the fraction of anti-aligned spins $\langle \hat{P}_b^{\text{ap}} \rangle$ is almost 1 at low temperatures, only 1.04(2)% of neighboring spins have the same sign of S^z – caused by quantum fluctuations. As we have seen above, the quantum fluctuations – i.e., spin flips in the S^z basis in the imaginary time direction – are measured by \hat{Q}_b ; it takes the value $\langle \hat{Q}_b \rangle = 8.33(1)\%$ in the ground state, which is quite low due to the large $\Delta = 4$.

6.2 Critical Behavior of Projectors at Phase Transitions

Remarkably, we indeed observe some kind of ‘critical behavior’ in the projectors, most strikingly a strong finite size dependence in the temperature regime around T_c . The majority of the observed finite size dependence arises from \hat{P}_b^{ap} . Clearly, the projection operators are bounded, so the ‘most singular’ behavior we can expect is a divergent derivative. The temperature derivative of $\langle \hat{P}_b^{\text{ap}} \rangle$ is shown in Figure 6.2. Indeed, the curves do not converge with increasing L , but seem to diverge. However, note that the divergence is very weak compared to the power law observed e.g., for the susceptibility, cf. Figure 4.2(b). We argue in the following that we basically observe the same divergence in Figure 6.2 as we do for the specific heat in Figure 4.4, which happens to be logarithmically divergent – i.e., the power law exponent $\alpha = 0$ vanishes – in the universality class of the 2D Ising model. Moreover, the projectors are directly related to the energy.

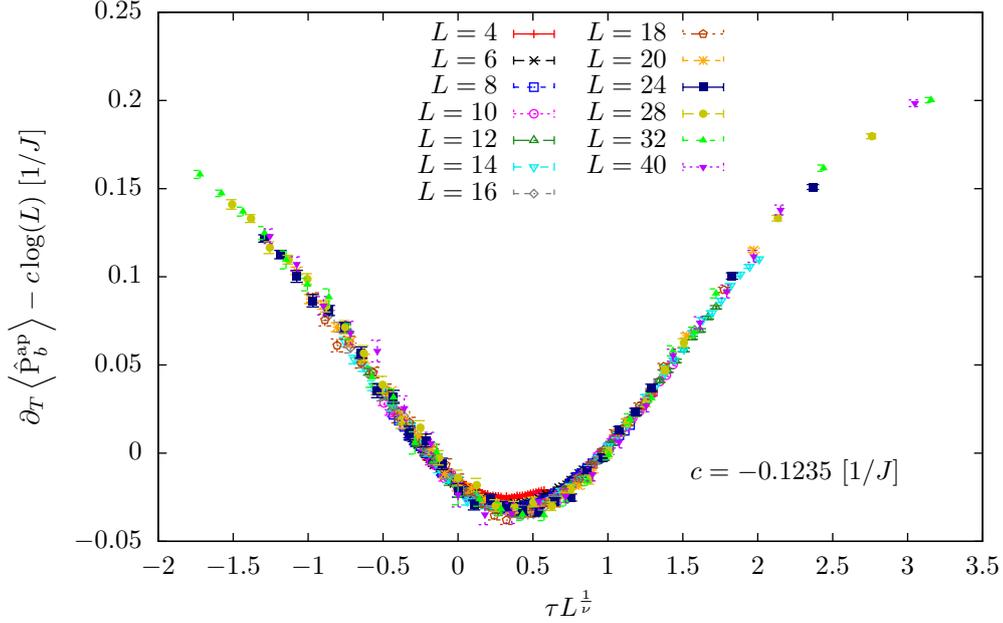


Figure 6.3: Collapse of the data $\partial_T \langle \hat{P}_b^{\text{ap}} \rangle$ shown in Figure 6.2. The parameters $T_c = 2.234J$ and $\nu = 1$ have been fixed, as well as the scaling with the logarithm; the only used fit-parameter is $c = -0.1235(2)$. Rescaling the y -axis with a power law instead of the logarithm does not lead to a collapse.

For the SSE mapping in section 3.2 we have extensively used the fact that our Hamiltonian $\mathcal{H} = \sum_b \mathcal{H}_b$ splits into terms \mathcal{H}_b describing the interaction of neighboring sites. By definition, the singlet S and the three triplets T_0 , T_+ and T_- are the eigenstates of the bond operator \mathcal{H}_b . In particular, these states form an orthonormal eigenbasis on the local Hilbert space of the two neighboring sites of b , which is (by construction) common for \mathcal{H}_b and all four projectors \hat{P}_b^S , $\hat{P}_b^{T_0}$, $\hat{P}_b^{T_+}$, and $\hat{P}_b^{T_-}$ onto these eigenstates – and arbitrary linear combinations of them, e.g., \hat{P}_b^{ap} and \hat{Q}_b . Even more, we can write \mathcal{H}_b as a linear combination of the projectors:

$$\mathcal{H}_b = E_S \hat{P}_b^S + E_{T_0} \hat{P}_b^{T_0} + E_{T_+} \hat{P}_b^{T_+} + E_{T_-} \hat{P}_b^{T_-}. \quad (6.12)$$

$$\Rightarrow \frac{E}{N_b} = \frac{1}{N_b} \sum_b \langle \mathcal{H}_b \rangle = E_S \langle \hat{P}_b^S \rangle + E_{T_0} \langle \hat{P}_b^{T_0} \rangle + E_{T_+} \langle \hat{P}_b^{T_+} \rangle + E_{T_-} \langle \hat{P}_b^{T_-} \rangle. \quad (6.13)$$

This shows rigorously that any singularity in the energy appears in at least one of the projections. In other words, at least one of the projections must behave singularly at the critical temperature.

Note that the sum of the four projectors is the identity, $\mathbb{1} = \hat{P}_b^S + \hat{P}_b^{T_0} + \hat{P}_b^{T_+} + \hat{P}_b^{T_-} \Rightarrow \hat{P}_b^{T_+} + \hat{P}_b^{T_-} = \mathbb{1} - \hat{P}_b^{\text{ap}}$. Since the two triplets T_{\pm} are degenerate, we can go even further and nail the singularity of the energy down to the combinations \hat{P}_b^{ap} and \hat{Q}_b :

$$\begin{aligned} \frac{E}{N_b} &= \frac{E_S + E_{T_0}}{2} \langle \hat{P}_b^{\text{ap}} \rangle + \frac{E_S - E_{T_0}}{2} \langle \hat{Q}_b \rangle + E_{T_{\pm}} \langle \mathbb{1} - \hat{P}_b^{\text{ap}} \rangle \\ &= \frac{J\Delta}{4} - \frac{J\Delta}{2} \langle \hat{P}_b^{\text{ap}} \rangle - \frac{J}{2} \langle \hat{Q}_b \rangle. \end{aligned} \quad (6.14)$$

In the last step, we plugged in the eigenenergies of the bond operator \mathcal{H}_b (given in Table 2.1) – the previous arguments were quite generic and apply to other models as well.

We conclude that the temperature derivative of $\langle \hat{P}_b^{\text{ap}} \rangle$ depicted in Figure 6.2 diverges in the same manner as the derivative of the energy, i.e., the specific heat. In particular, we can apply the machinery of finite size scaling near the critical temperature given in section 4.3: For example, we can collapse the finite length data, which is depicted in Figure 6.3. As mentioned above, the specific heat of the 2D Ising model does not diverge with a power law, but logarithmically [89]:

$$\frac{C_V}{N} = A_0 \log(L) + \tilde{\Phi}(\tau, \xi) = A_0 \log(L) + \Phi(\tau L^{\frac{1}{\nu}}). \quad (6.15)$$

Here, $A_0 = \frac{2}{\pi} (\log(1 + \sqrt{2}))^2 \approx 0.4945$ is a constant, Φ is a scaling function (which can explicitly be given for the 2D Ising model [89]), and we ignore sub-leading contributions. We found the collapse of $\partial_T \langle \hat{P}_b^{\text{ap}} \rangle$ assuming the same scaling form, but a different prefactor c instead of A_0 and another scaling function Φ . The best collapse was obtained for $c = -0.1235(2)/J$. According to eq. (6.14), the singular behavior of C_V (i.e., the prefactor of the divergence) has to split between $\partial_T \langle \hat{P}_b^{\text{ap}} \rangle$ and $\partial_T \langle \hat{Q}_b \rangle$; by comparison we obtain (with $N_b/N = 2$) $A_0 = -J\Delta c - Jc'$, where c' is the corresponding prefactor in the divergence of $\partial_T \langle \hat{Q}_b \rangle$. Within error bars, this relation holds for $c' = 0$, which suggests that $\partial_T \langle \hat{Q}_b \rangle$ does not diverge at all. However, the inset of Figure 6.2 gives numerical evidence that c' is a small *positive* number. Nevertheless, it is an order of magnitude smaller than c such that the statistical errors are too large to extract c' precisely.

Further, the minimum of $\partial_T \langle \hat{P}_b^{\text{ap}} \rangle$ moves towards T_c with increasing L in exactly the same fashion as the maximum of the specific heat does. This is depicted in Figure 6.4. Note that the minimum of the derivative $\partial_T \langle \hat{P}_b^{\text{ap}} \rangle$ corresponds to an inflection point of the projections themselves. However, the extraction of these inflection points gives worse results than the minima of the derivative, see Figure 6.4. This may be related to the fact that we fitted the results to a third degree polynomial in both cases.

Susceptibilities

Above, we have seen the close relation between the projectors and the energy and concluded that they share the same singular behavior at the critical temperature. In other words, we have rigorously answered the question whether the projectors have singular behavior at T_c with *yes*. Nevertheless, we go on and study in addition spatial correlations of them between different bonds. However, we restrict ourselves to the linear combinations \hat{P}_b^{ap} and \hat{Q}_b .

Although the projectors on a bond b share the eigenbasis with \mathcal{H}_b , they do not commute with the total Hamiltonian. Therefore, we measure the correlations with the generalized susceptibilities (see eq. (3.58)):

$$\frac{\chi_{\hat{A}\hat{B}}}{\beta} = \frac{1}{\beta} \left. \frac{\partial}{\partial b} \right|_{b=0} \langle \hat{A} \rangle_{\mathcal{H}+b\hat{B}} = \frac{1}{\beta} \int_0^\beta \langle \hat{A}(\tau)\hat{B} \rangle d\tau - \langle \hat{A} \rangle \langle \hat{B} \rangle. \quad (6.16)$$

Since the projectors are basically parts of the Hamiltonian, the susceptibility boils down to a connected correlator for numbers of vertices in the operator string, see eq. (B.13) in appendix B.

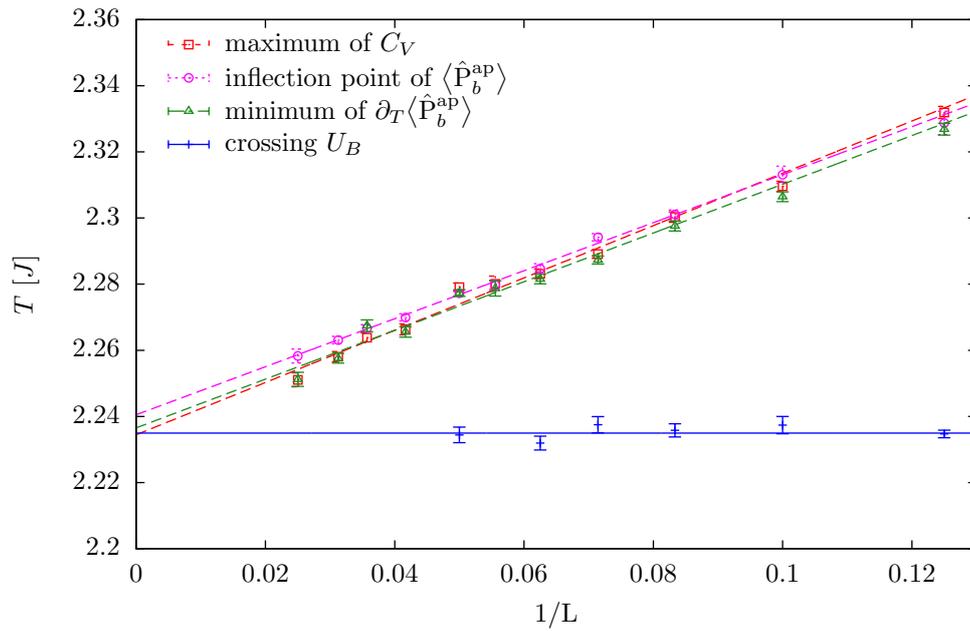


Figure 6.4: Finite size scaling of the minima of $\partial_T \langle \hat{P}_b^{\text{ap}} \rangle$ and the inflection points of $\langle \hat{P}_b^{\text{ap}} \rangle$, each extracted by a fit to a polynomial of third degree. Due to time restrictions we did not fully propagate the statistical errors (but used `gnuplot`); the error bars seem to be too small: clearly, the inflection points of $\partial_T \langle \hat{P}_b^{\text{ap}} \rangle$ and the minima of $\langle \hat{Q}_b \rangle$ should coincide. For comparison, the maxima of the specific heat and the crossing of the Binder cumulant are shown as well, see Figure 4.7.

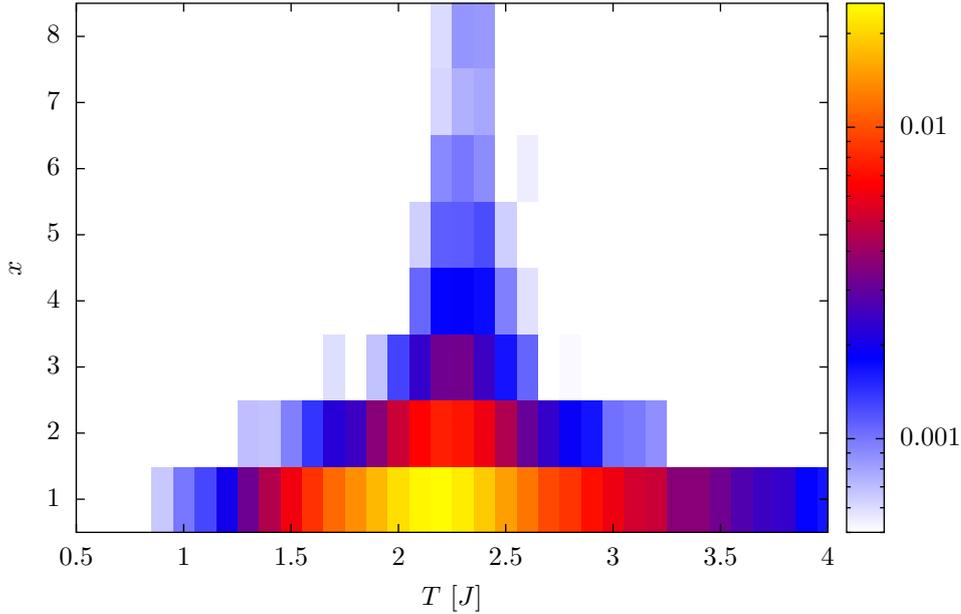


Figure 6.5: Spatial correlation of projectors \hat{P}_b^{ap} , i.e., $\chi_{\hat{P}_b^{\text{ap}}, \hat{P}_{b'}^{\text{ap}}}/\beta$ on a 16×16 2D square lattice, $\Delta = 4$, where the bonds $b = \langle 1, 2 \rangle$ and $b' = \langle 1+x, 2+x \rangle$ are both parallel to the x -axis and b' is shifted by x sites (in x -direction) relative to b . Note the logarithmic scale for the colors; statistical errors are less than 5×10^{-4} .

The correlation of projectors \hat{P}_b^{ap} on different bonds is shown in Figure 6.5. Due to the periodic boundary conditions – which allows an average over the positions of b reducing the statistical errors – the correlations are symmetric around $x = \frac{L}{2} = 8$; the other half is not shown. We clearly see long-range correlations around the critical temperature, strongest slightly above T_c due to the finite system size. The decay with the distance right at the critical temperature is shown in Figure 6.6 in detail. We observe a power law, yet with a much faster decay than the longitudinal two-point spin correlations. Note that each of \hat{P}_b^{ap} involves the correlations on two neighboring sites (see eq. (6.10)) such that the susceptibility is a four-site correlator (the constant shift cancels):

$$\frac{\chi_{\langle i,j \rangle, \langle k,l \rangle}^{\text{ap}}}{\beta} = 4 \left[\langle S_i^z(\tau) S_j^z(\tau) S_k^z S_l^z \rangle - \langle S_i^z S_j^z \rangle \langle S_k^z S_l^z \rangle \right]. \quad (6.17)$$

Thus, it is not a big surprise that the susceptibility of \hat{P}_b^{ap} decays much faster (as a function of distance) than the two-point correlations. In the 2D Ising model, this correlation function decays as r^{-2} with the distance $r = |k - i|$ between the bonds, which compares well with our results.

We have not shown the data for $x = 0$ in Figure 6.5, but they are shown separately in Figure 6.7. In the high temperature limit, we have $e^{-\beta\mathcal{H}} \rightarrow \mathbb{1}$ such that $\chi_{\hat{P}_b^{\text{ap}}, \hat{P}_b^{\text{ap}}}/\beta \rightarrow \langle (\hat{P}_b^{\text{ap}})^2 \rangle - \langle \hat{P}_b^{\text{ap}} \rangle^2$ reduces to the connected correlator over all states. Since \hat{P}_b^{ap} is a projector, $(\hat{P}_b^{\text{ap}})^2 = \hat{P}_b^{\text{ap}}$, the connected correlator does not vanish at infinite temperature, but approaches a constant value $\frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$. On the other hand, \hat{Q}_b is not a projector,

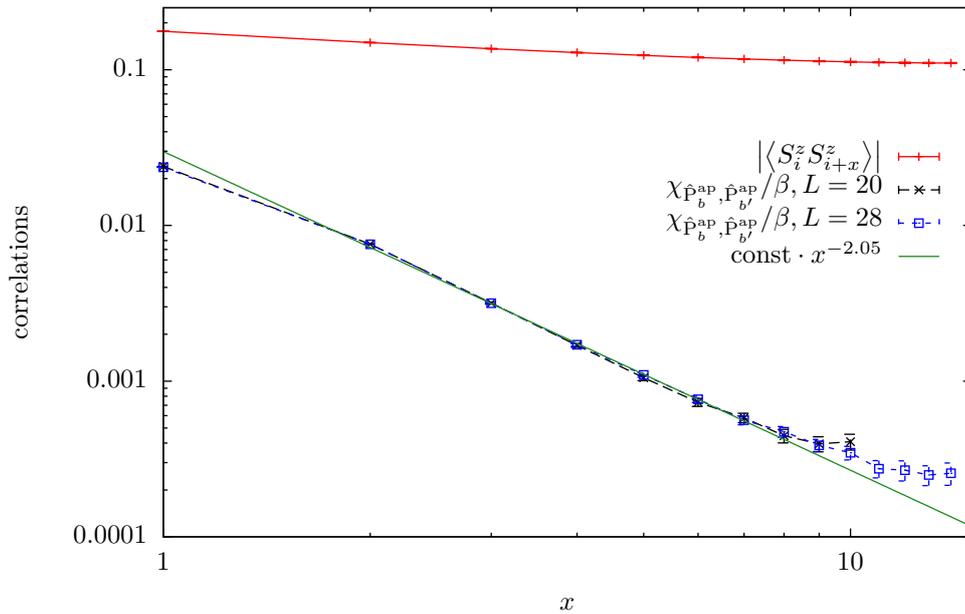


Figure 6.6: Decay of susceptibilities $\chi_{\hat{p},\hat{p}}/\beta$ with increasing distance between the bonds, on an $L \times L$ 2D square lattice, $\Delta = 4$, $T = 2.24J \approx T_c$. Note the logarithmic axes, such that power laws appear as straight lines. For comparison the longitudinal two-point spin correlations are shown as well, see Figure 4.5. The green line shows a fit to a power law for intermediate distances $3 \leq x \leq 8$ of the $L = 28$ data, yielding the exponent 2.05(3) which compares well to the expected value 2.

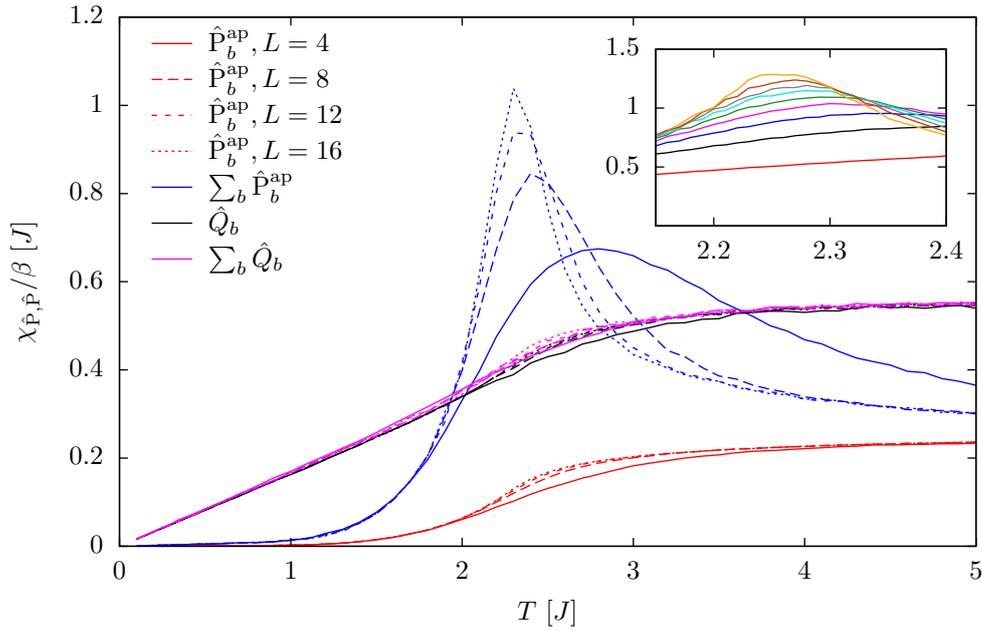


Figure 6.7: Susceptibilities $\chi_{\hat{P}, \hat{P}}/\beta$ on an $L \times L$ 2D square lattice, $\Delta = 4$: locally on a single bond, $\hat{P} = \hat{P}_b^{\text{ap}}$ (red) and $\hat{P} = \hat{Q}_b$ (black), or for the global sums $\hat{P} = \sum_b \hat{P}_b^{\text{ap}}$ (blue) and $\hat{P} = \frac{1}{N_b} \sum_b \hat{Q}_b$ (purple), see eq. (6.18). For each of them, different line styles indicate different lengths $L \in \{4, 8, 12, 16\}$, revealing the finite size dependence. The inset shows the global susceptibility $\chi_{\hat{P}^{\text{ap}}, \hat{P}^{\text{ap}}}/\beta$ near the critical temperature for $L = 4, 8, \dots, 28, 32, 40$ from bottom to top.

but we have $(\hat{Q}_b)^2 = (\hat{P}_b^S)^2 + (\hat{P}_b^{T_0})^2 = \hat{P}_b^{\text{ap}}$, such that at high temperatures $\chi_{\hat{Q}_b, \hat{Q}_b} / \beta \rightarrow \frac{1}{2} - 0^2 = \frac{1}{2}$. However, this clearly implies that the susceptibilities themselves vanish proportionally to β at high temperatures.

Due to the energy gap, the system is effectively only in the ground state over a wide temperature range in the ordered phase. While $\chi_{\hat{P}_b^{\text{ap}}, \hat{P}_b^{\text{ap}}}$ vanishes exponentially fast as expected in a gapped system, we find that $\chi_{\hat{Q}_b, \hat{Q}_b} / \beta$ is linear with T , i.e., $\chi_{\hat{Q}_b, \hat{Q}_b}$ is a constant in the ground state. This can be understood from the fact that applying \hat{Q} twice in a row does not destroy the Néel state, while a single application of \hat{Q} on the Néel state flips the spins on both sites such that the spins of the adjacent bonds have the same signs. In the picture of the SSE operator string, we mostly find the vertices of \hat{Q}_b – i.e., $\mathcal{H}_b^{\text{offd}}$, see eq. (6.8) – in pairs following each other. Single vertices are much less probable.

It is remarkable that the phase transition – a phenomenon requiring the thermodynamic limit and characterized by divergent correlations – manifests itself with a divergent derivative in the local projectors on a single bond. To capture the correlations of the projectors on different bonds in a single quantity, we define ‘global’ susceptibilities

$$\chi_{\hat{P}, \hat{P}} := \frac{1}{N_b} \chi_{(\sum_b \hat{P}_b), (\sum_{b'} \hat{P}_{b'})} = \frac{1}{N_b} \sum_{b, b'} \chi_{\hat{P}_b, \hat{P}_{b'}}. \quad (6.18)$$

We normalize with the number of bonds $N_b = \sum_b 1$ such that we obtain the local susceptibility $\chi_{\hat{P}_b, \hat{P}_b}$ if there are no spatial correlations, i.e., $\chi_{\hat{P}_b, \hat{P}_{b'}} \propto \delta_{b, b'}$. For spatially anisotropic systems we take the sum over equivalent bonds only (stating explicitly which ones) and normalize with their number, respectively.

The global susceptibilities are shown in Figure 6.7 in addition to their local counterparts. The curves for the local and global susceptibility of \hat{Q}_b lie almost on top of each other. In other words, there are – at least for the considered system sizes – almost no spatial correlations between different bonds for \hat{Q}_b . The largest nevertheless small difference between the curves is found near the critical temperature. In contrast, the local and global susceptibility of \hat{P}^{ap} differ dramatically at intermediate temperatures: as we have seen in Figure 6.5, the projectors develop relevant spatial correlations near T_c . Once more, we find the same logarithmic divergence as for the specific heat $C_V = \beta \chi_{\mathcal{H}, \mathcal{H}}$: it contains all the terms present in the $\chi_{\hat{P}_b^{\text{ap}}, \hat{P}_b^{\text{ap}}}$ – yet with a different (global) prefactor, and beside further terms from the other projectors $\hat{P}_b^{T_+} + \hat{P}_b^{T_-}$.

We can also relate this divergence to the decay of the local susceptibilities $\chi_{\hat{P}_b^{\text{ap}}, \hat{P}_{b'}^{\text{ap}}}$ shown in Figure 6.6: due to translation symmetry only the distance of the bonds matters and one of the sums in eq. (6.18) cancels against $\frac{1}{N_b}$. In the thermodynamic limit, we can further replace the remaining sum with an integral over the distance r between the bonds (using the lattice constant $a \equiv 1$ as an ultraviolet cutoff):

$$\chi_{\hat{P}, \hat{P}} = \sum_b \chi_{\hat{P}_0 \hat{P}_b} \rightarrow \int_a^L dr 2\pi r \underbrace{\chi_{\hat{P}_0 \hat{P}_r}}_{\propto r^{-2}} \propto \log(L) \quad (6.19)$$

Here, the factor $2\pi r$ stems from the volume element in two dimensions. Further, we ignored details in the borders of the integral (we integrated over a disc instead of a square), as they do not matter for the scaling in the thermodynamic limit.

6.3 1D-2D-Crossover

In the previous section we have seen a close connection between the energy and the projections on a spatially isotropic 2D square lattice. The argument in the beginning of section 6.2 shows that at least some of the projectors have the same critical behavior at T_c as the energy; yet it does not specify which one – we found only very weak divergencies for \hat{Q} . In the presence of spatial anisotropies we expect that the projections on different bonds are not the same, which raises the question how the singularity splits over the different types of bonds.

In this section, we focus on the projections in the 1D-2D-crossover of the coupled chains depicted in Figure 2.2(a), which we discussed in section 4.4. Clearly, this geometry consists of only two groups of equivalent bonds: the ones within the chains with the strong coupling strength $J \equiv 1$, and the weak bonds of strength $J' < J$ coupling the chains.

The expectation values $\langle \hat{P}_b^{\text{ap}} \rangle$ and $\langle \hat{Q}_b \rangle$ for bonds of each of these two groups are shown in Figure 6.8 for different coupling strengths J' between the chains. As expected, the projections on the different bonds are very similar for $J' = 0.9$ and close to the isotropic 2D square lattice, see Figure 6.1(b). With decreasing J' towards the 1D limit, the projections on the two groups of bonds behave differently.

The quantum fluctuations measured by $\hat{P}_b^{\text{ap}} = \frac{2}{J_b} \mathcal{H}_b^{\text{offd}}$ even show a different monotonic behavior: they get stronger within the chains towards the 1D limit, while they decrease on the bonds between the chains at the same time. To understand the monotonic behavior, we consider quantum fluctuations as virtual processes on the Néel state where two neighboring spins of a certain bond are flipped for a short (imaginary) time – this is close to what we find in our SSE simulations. By definition, the kinetic energy $\mathcal{H}_b^{\text{offd}}$ of a weak bond scales with J' , but the energy of the intermediate virtual state has a term proportional to J due to the intra-chain coupling. Thus, virtual processes along weak bonds get suppressed for small J' , i.e., there are less quantum fluctuations along the weak bonds for small J' . On the other hand, we keep the kinetic energy gained for flips within the chains constant, while the energy of the virtual state decreases with smaller J' ; thus, such virtual processes are more probable, i.e., there are more quantum fluctuations within the chains. Further, we find a local maximum in $\langle \hat{Q}_b \rangle$ around the critical temperature, which is present but very weakly pronounced for any J' and on all bonds. It can be seen clearest for small J' on the strong bonds.

The curves of $\langle \hat{P}_b^{\text{ap}} \rangle$ have the same monotonic behavior with J' on the different bonds; yet they differ for small J' as well. At intermediate to large temperatures, we find much stronger nearest-neighbor correlations along the chains than in-between them; the inter-chain coupling becomes relevant at much lower temperatures, $\langle \hat{P}_b^{\text{ap}} \rangle$ at these bonds is close to the high-temperature value $\frac{1}{2}$ down to temperatures $T \approx 1.5J$ for $J' = 0.1$. We conclude that for small J' the phase transition to long-range order is mainly driven by the onset of correlations between the chains, as we have already argued in section 4.4.

The global susceptibilities of the two groups of bonds are depicted in Figure 6.9. First, note a tremble of the curves – i.e., large statistical errors – at high temperatures, especially for $J' = 0.1$. These large errors appear since the number of vertices in the operator string, especially on the weak bonds, is quite small such that the discrete nature of the numbers gets relevant; it takes many Monte Carlo sweeps to approximate the exact continuous distribution. The high and low temperature limits of all curves coincide and are the same

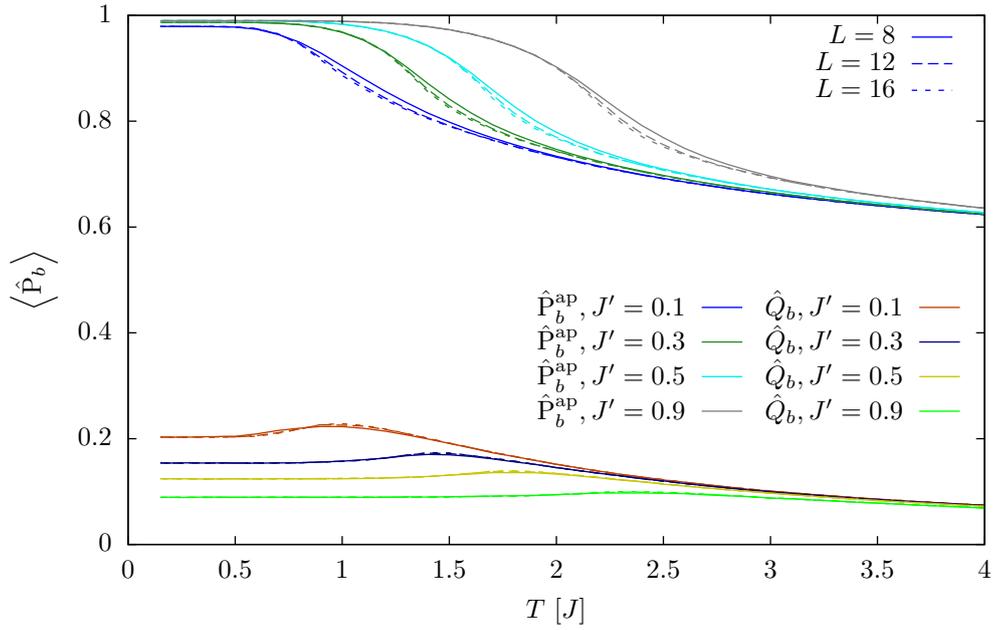
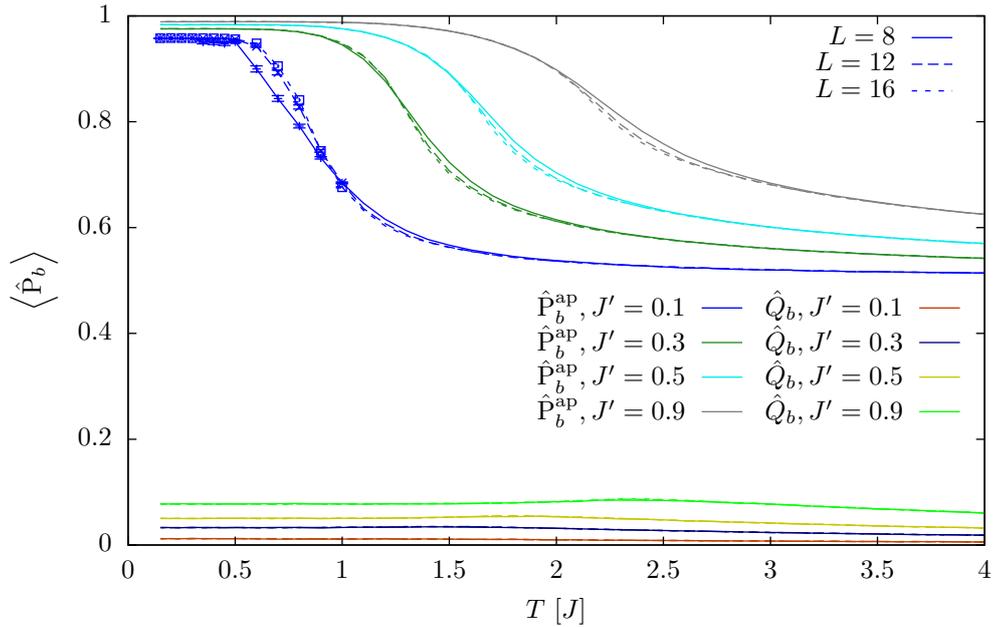
(a) on a strong bond of the intra-chain coupling $J \equiv 1$ (b) on a weak bond of the inter-chain coupling $J' < J$

Figure 6.8: Local projections \hat{P}_b^{ap} and \hat{Q}_b for coupled chains on an $L \times L$ square lattice (see Figure 2.2(a)), $\Delta = 4$. Different colors indicate different $J' \in \{0.1, 0.3, 0.5, 0.9\}$. For each of them, different line styles indicate different lengths $L \in \{8, 12, 16\}$, revealing the finite size dependence. We marked the ordered phase of $J' = 0.1$ for \hat{Q}_b on a weak bond with error bars to emphasize problems caused by missing ergodicity, see Figure 4.13. Elsewhere the errors are much smaller.

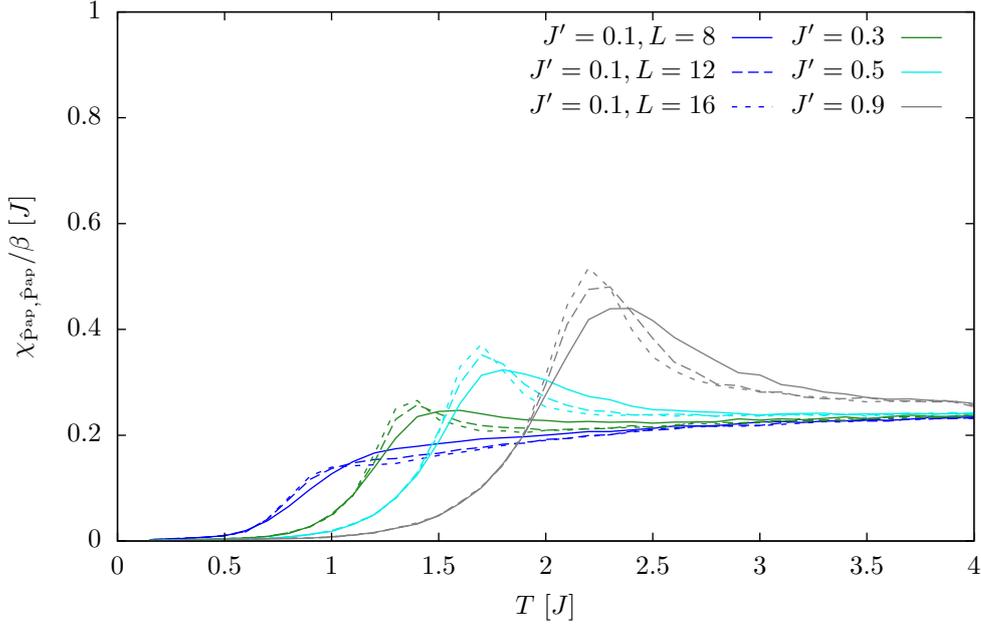
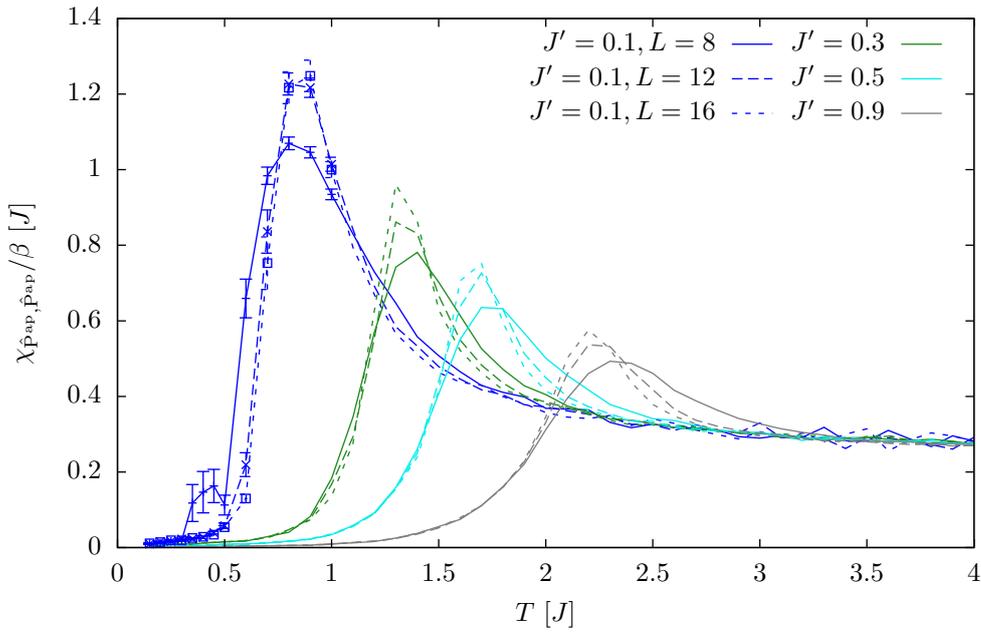
(a) for the strong bonds of the intra-chain coupling $J \equiv 1$ (b) for the weak bonds of the inter-chain coupling $J' < J$

Figure 6.9: Global susceptibility of \hat{P}^{ap} for coupled chains on an $L \times L$ square lattice, $\Delta = 4$. Here, the sums in eq. (6.18) run either over all strong bonds (top), or over all weak bonds (bottom). Different colors indicate different $J' \in \{0.1, 0.3, 0.5, 0.9\}$. For each of them, different line styles indicate different lengths $L \in \{8, 12, 16\}$, revealing the finite size dependence. We marked the ordered phase of $J' = 0.1$ with error bars to emphasize problems caused by missing ergodicity, see the discussion of Figure 4.13. Further statistical errors are visible as a tremble of the curves at high temperatures, especially for small J' in the bottom panel, see main text.

as in the isotropic 2D lattice, see Figure 6.7. However, compared to Figure 6.7, the peak around T_c is much less pronounced in the limit $J' \rightarrow 1$ (which is the same in the top and bottom panel), since we only consider half of the bonds and thus do not capture the correlation between bonds in x - and y -direction here.

In Figure 6.9(a), i.e., for the strong bonds within the chains, the peaks around T_c get much smaller with decreasing J' and vanish almost completely for $J' = 0.1J$. In that case, the spins order to a large amount within the chains even above the critical temperature (see the discussion in section 4.4). But this order is only driven by local alignment, which is subtracted in the connected correlation functions and thus not captured by the susceptibility. In contrast, the peaks in Figure 6.9(b) grow larger towards the 1D limit, which confirms above picture that the phase transition to the ordered phase is driven by the correlations in y -direction.

6.4 Coupled Dimers

Finally, we consider the coupled dimers depicted in Figure 2.2(b) as a second example for spatially anisotropic lattices. As we have discussed and seen in section 4.5, the critical temperature T_c vanishes for finite $J' \approx 0.1J$; there is a quantum phase transition in addition to the finite temperature transitions for $J' > J_c$.

This geometry has three groups of equivalent bonds: (i) the strong bonds of coupling strength $J \equiv 1$ which form the dimers, (ii) the weak bonds of coupling strength $J' < J$ parallel to the dimers, and (iii) the weak bonds of coupling strength $J' < J$ perpendicular to the dimers. However, we expect very similar results for the groups (ii) and (iii), thus we consider only the groups (i) and (ii) in this work.

In Figure 6.11, the projections $\langle \hat{P}_b^{\text{ap}} \rangle$ and $\langle \hat{Q}_b \rangle$ are shown, for bonds of the first group in panel (a) and for bonds of the second group in (b). Of course, the limit $J' \rightarrow 1$ corresponds to the 2D square lattice as before. Further, the value of $\langle \hat{Q}_b \rangle$ on the weak bonds decreases for smaller J' as well, which can be understood with the same argument of virtual processes given above, yet with a slight modification: in this case we do not start from a Néel state. However, two spins belonging to a dimer are always (up to less than 2%) anti-aligned at low temperatures, as one can clearly see in the top panel (a). A weak bond connects between two dimers. If the spins of a weak bond are parallel, a virtual process flipping those two spins is not possible. If they are anti-parallel, the spin flip is possible; but – as before – the intermediate state has a high energy on the order of J since it involves parallel spins on the two dimers connected by the weak bond. Thus, in both cases the quantum fluctuations are suppressed on the weak bonds.

On the strong bonds – i.e., the dimers – the quantum fluctuations $\langle \hat{P}_b^{\text{ap}} \rangle$ get stronger with decreasing J' : the ground state for $J' = 0$ is a product state of singlets on the dimers; thus both $\hat{P}_b^{\text{ap}} = \hat{P}_b^S + \hat{P}_b^{T_0}$ and $\hat{Q}_b = \hat{P}_b^S - \hat{P}_b^{T_0}$ are (almost) 1 in the ground state for small J' . For $J' = 0.1 \approx J_c$, we find that $\langle \hat{Q}_b \rangle \approx 0.9$ on the dimers. It drops rapidly if J' is increased further, and we observe local maxima around the critical temperature as for the coupled chains in the previous section.

It is straightforward to calculate the projections in the trial state $|\Psi^{\text{trial}}\rangle$ defined in eq. (4.30). We obtain:

$$\langle \Psi^{\text{trial}} | \hat{P}_b^{\text{ap}} | \Psi^{\text{trial}} \rangle = \begin{cases} 1 & \text{on strong bonds, } J_b = J, \\ \frac{1+s^4}{(1+s^2)^2} & \text{on weak bonds, } J_b = J', \end{cases} \quad (6.20)$$

$$\langle \Psi^{\text{trial}} | \hat{Q}_b | \Psi^{\text{trial}} \rangle = \begin{cases} \frac{2s}{1+s^2} & \text{on strong bonds, } J_b = J, \\ 0 & \text{on weak bonds, } J_b = J'. \end{cases} \quad (6.21)$$

The arrows in Figure 6.10 indicate these results for the s minimizing the trial state energy for a given coupling J' , i.e., for the best candidate for the ground state, see the discussion of Figure 4.15. The trial state fails for \hat{Q}_b on the weak bonds – it ignores this part of the Hamiltonian. However, it describes the drop of both \hat{Q} on the strong bonds and \hat{P}^{ap} on the weak bonds surprisingly well.

We find that $\langle \hat{P}_b^{\text{ap}} \rangle$ on the dimers is nearly on top of the curve for a single dimer at all temperatures even for inter-dimer couplings as large as $J' = 0.2 > J_c$. In other words, the projections \hat{P}_b^{ap} on the strong bonds are in no way singular at T_c for small J' . Instead, we find the strong finite size dependence indicating a divergent derivative in the curves $\langle \hat{P}_b^{\text{ap}} \rangle$

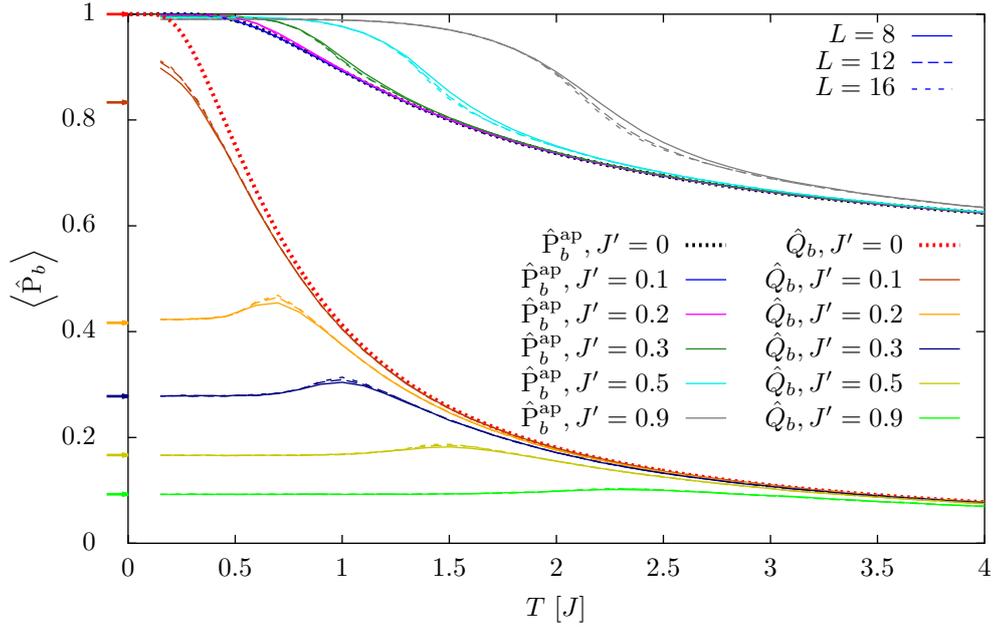
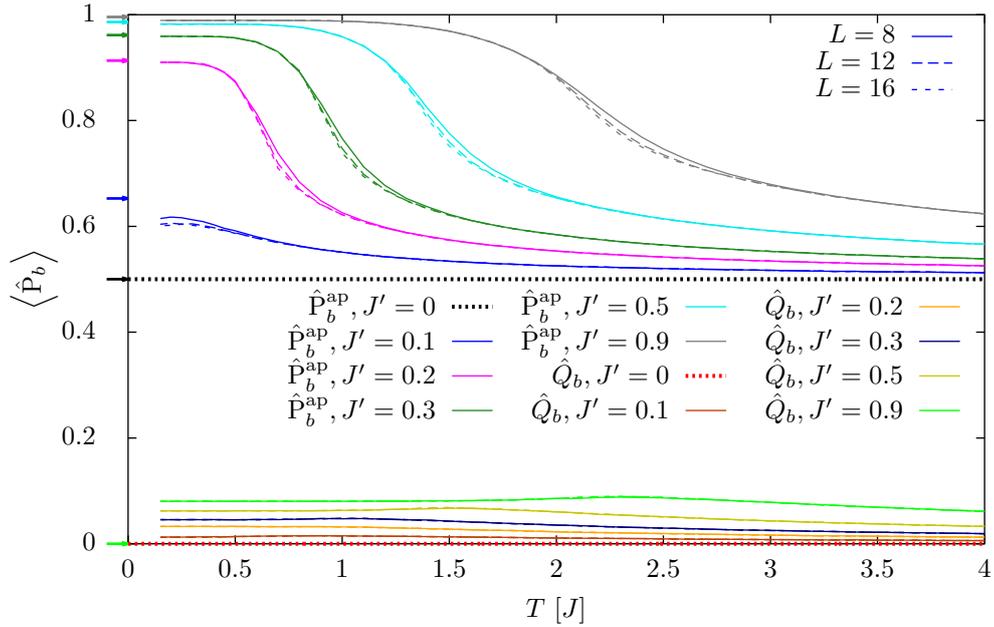
(a) on a strong bond $J \equiv 1$, i.e., a dimer(b) on a weak bond $J' < J$ parallel to the dimers

Figure 6.10: Local projections \hat{P}_b^{ap} and \hat{Q}_b for coupled dimers on an $L \times L$ square lattice (see Figure 2.2(b)), $\Delta = 4$. Different colors indicate different $J' \in \{0.1, 0.2, 0.3, 0.5, 0.9\}$. For each of them, different line styles indicate different lengths $L \in \{8, 12, 16\}$, revealing the finite size dependence. The thick, dotted curves show the limit $J' \rightarrow 0$, i.e., completely decoupled dimers. The arrows at the y -axis indicate the results for the trial state $|\Psi^{\text{trial}}\rangle$ (see eq. (6.20) and (6.21)) where the parameter s was chosen to minimize the energy for given J' (with the same color key as the lines).

on the weak bonds – similar as for the 1D-2D-crossover of the previous section. These curves show a strong increase at the critical temperatures.

As in the previous section, the global susceptibilities $\chi_{\hat{p}_{\text{ap}},\hat{p}_{\text{ap}}}$ depicted in Figure 6.11 confirm that the phase transition is driven by the correlations of the weak bonds. Even for $J' = 0.9$, close to the isotropic 2D square lattice, the peak of the susceptibility on the weak bonds is much smaller than the peak on the strong bonds². On the dimers, the ‘peak’ is hardly visible for $J' = 0.5$ and vanishes completely on the other J' . We argue in the same manner as before: The phase transition is driven by the weak bonds, the spins on the dimers are (pairwise) anti-aligned even at temperatures above T_c . Slightly above T_c we have thus many fluctuations on the dimers, but few fluctuations between the dimers. At the phase transition, the correlations diverge and the fluctuations within the dimers get suppressed due to the weak bonds connecting the dimers, such that long-range order develops. Indeed, we find that the peaks of the susceptibilities for the weak bonds in Figure 6.11(b) grow for decreasing J' – until we find no finite temperature transition any more at $J' = 0.1 \approx J_c$.

² The two groups of bonds (i) and (ii) are equivalent for $J' = J$ – in contrast to the group (iii) of bonds in the y -direction. Thus, the curves in the top and bottom panel of Figure 6.11 indeed coincide for $J' = J \equiv 1$.

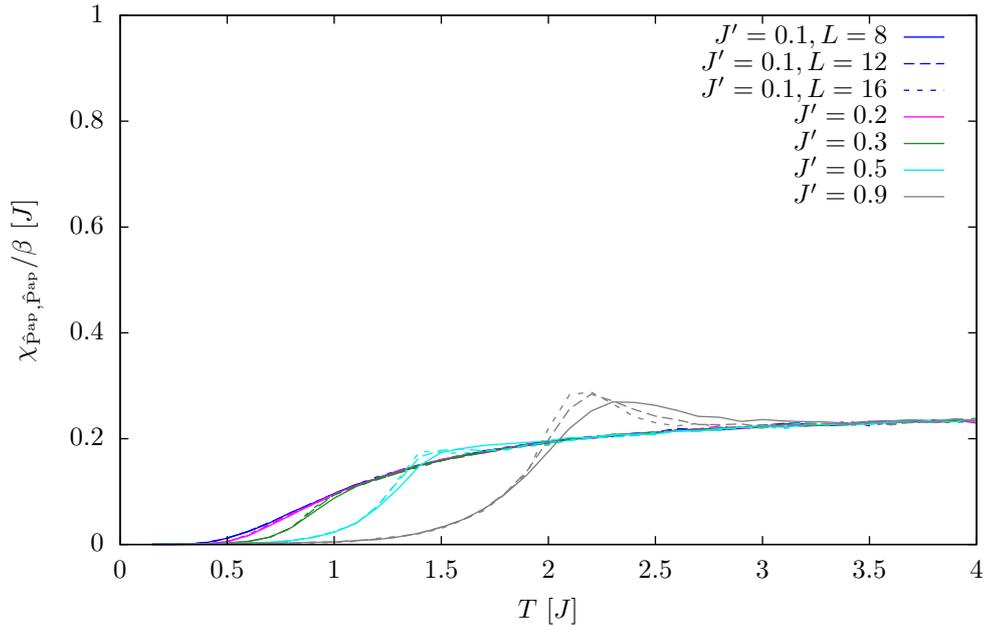
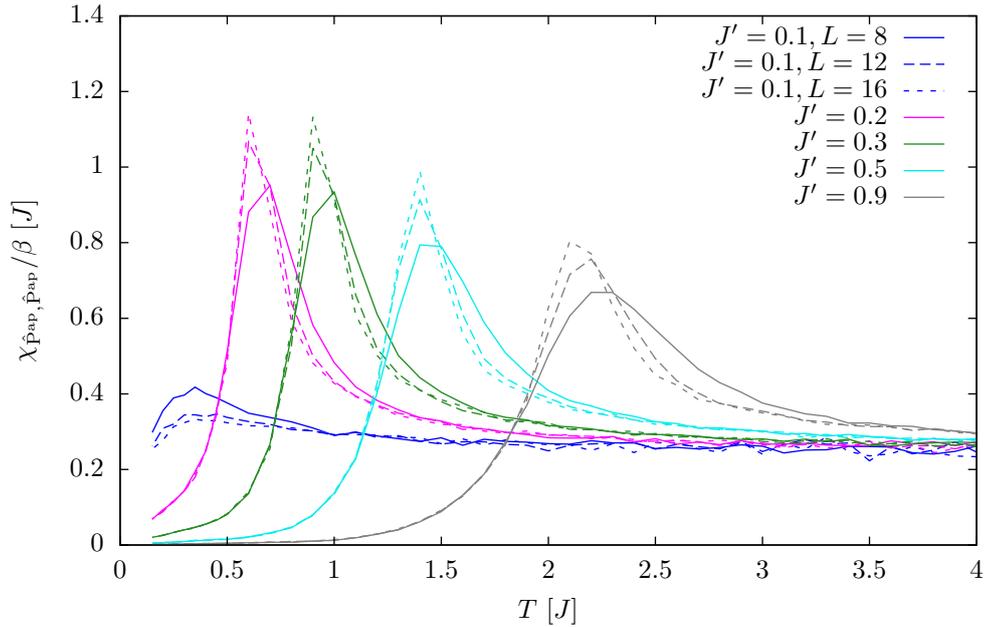
(a) for the strong bonds $J \equiv 1$, i.e., all dimers(b) for all weak bonds $J' < J$ parallel to the dimers

Figure 6.11: Global susceptibility of \hat{P}^{ap} for coupled dimers on an $L \times L$ square lattice, $\Delta = 4$. The sums in eq. (6.18) run either over all strong bonds (top), or over all weak bonds parallel to the dimers (bottom). Different colors indicate different $J' \in \{0.1, 0.2, 0.3, 0.5, 0.9\}$. For each of them, different line styles indicate different lengths $L \in \{8, 12, 16\}$, revealing the finite size dependence. Like in Figure 6.9, statistical errors cause a tremble of the curves at high temperature, especially on the weak bonds for small J' .

7 Summary

In the introduction, we gave an overview on the realization of magnetism in experiments with ultracold atoms. We motivated the isotropic Heisenberg model as an effective Hamiltonian for the low-energy physics of the Hubbard model in the Mott insulating phase. Since the isotropic Heisenberg model has no finite temperature phase transition in two dimension according to the Mermin-Wagner theorem, we studied the XXZ -model for $\Delta = 4$.

We have successfully implemented the stochastic series expansion from scratch, with both deterministic loop updates for $SU(2)$ symmetric couplings and the more general directed loop updates for the XXZ -model. In chapter 3, we reviewed the theoretical aspects how and why this quantum Monte Carlo method works and commented on some practical issues of the implementation.

In chapter 4, we discussed the finite temperature phase transition to the antiferromagnetic ordered phase in the two dimensional XXZ -model for $\Delta = 4$. We found that our algorithm is not ergodic in the ordered phase, but we argued that we nevertheless obtain correct results for (most of) the considered observables. We discussed the divergence of the correlation lengths (and other quantities) as the critical temperature is approached and used finite size scaling to extract T_c with various methods. The best results are obtained from a data collapse of the Binder cumulant, which yields $T_c = 2.2343(2)J$. As expected, we found the critical exponents of the universality class of the 2D Ising model. We examined the phase diagram of the 1D-2D-crossover and found good qualitative agreement with the Ising model. For weakly coupled dimers, we found a quantum phase transition between a valence bond crystal with singlets on the dimers and antiferromagnetic Néel order, when the weak bonds have a coupling strength $J_c \approx 0.10(2)J$. We discussed the quantum phase transition with a trial state approximating the true ground state and compared it to our numerical results from SSE simulations.

In chapter 5, we discussed the concept of entropy and mutual information. We implemented the replica trick to obtain the Rényi entropies, reproduced crossings close to T_c and $2T_c$ in the Rényi mutual information of second order from Ref. [1] and found them also in for the spatially anisotropic geometries of coupled chains or dimers. As an alternative to the thermodynamic integration, we discussed the ratio trick and calculated the entanglement entropy in the ground state. Since our numerical simulations are not ergodic, we did not obtain the constant $\log(2)$ stemming from spontaneous symmetry breaking. Nevertheless, the numerical results agree with the area law scaling which we obtained by a separate perturbation theory calculation starting from the Ising model.

In chapter 6, we examined projectors onto singlet and triplet states on certain bonds, as well as the sum and difference \hat{P}_b^{ap} and \hat{Q}_b , which we identified as the projector on anti-parallel spin states and the off-diagonal spin-flip term of the Hamiltonian which causes the quantum fluctuations. We illustrated a close relation of the projections and the energy, which proves rigorously that at least one of the projectors has the same singular behavior

at T_c as the energy. Indeed, we find numerical evidence that the derivatives of both \hat{P}_b^{ap} and \hat{Q}_b diverge logarithmically with the system size, i.e., in the same way as the specific heat. However, the prefactor of this divergence is an order of magnitude smaller for \hat{Q}_b . We considered also spatial correlations measured by generalized susceptibilities. We find strong (connected) correlations of \hat{P}_b^{ap} near the critical temperature, where they decay with a power law in the distance. The global susceptibility for \hat{P}^{ap} diverges logarithmically as the specific heat, too. In the spatially anisotropic geometries, we found that the phase transition is driven by the correlations of the weaker bonds; we found the cleanest signs of the phase transition in \hat{P}_b^{ap} on weak bonds. The peaks at T_c observed in the global susceptibility of \hat{P}_b^{ap} grow on the weak bonds when their strength J' is decreased. In contrast, the peaks vanish almost completely on the strong bonds.

A SSE Estimators for parts of the Hamiltonian

We assume a general Hamiltonian of the form $\mathcal{H} = \text{const} - \sum_k \mathcal{H}_k$, where we sample each of the parts \mathcal{H}_k as own vertices in the operator string, i.e., each $\mathcal{H}_k \propto |\alpha_k\rangle \langle \alpha_k|$ has just a single non-vanishing matrix element. To get estimators for $\langle \mathcal{H}_k \rangle$, we use a generating functional $\mathcal{Z} [\{\beta_k\}]$, where $\mathcal{Z} [\{\beta_k = \beta\}]$ coincides with the usual partition function:

$$\langle \mathcal{H}_k \rangle = \frac{\partial}{\partial \beta_k} \log (\mathcal{Z} [\{\beta_k\}]) \Big|_{\beta_k = \beta} \quad \text{for} \quad \mathcal{Z} [\{\beta_k\}] := \text{Tr} \left\{ \exp \left(- \sum_k \beta_k \mathcal{H}_k \right) \right\}. \quad (\text{A.1})$$

A straightforward generalization of section 3.2 leads to (cf. eq. (3.29)):

$$\mathcal{Z} [\{\beta_k\}] = \sum_{n=0}^{\infty} \sum_{|\alpha\rangle} \sum_{S_n} \frac{1}{n!} \langle \alpha | \prod_{p=0}^{n-1} \beta_{k(p)} \mathcal{H}_{k(p)} | \alpha \rangle \quad (\text{A.2})$$

$$= \sum_{n=0}^{\infty} \sum_{|\alpha\rangle} \sum_{S_n} \frac{1}{n!} \left(\prod_k (\beta_k)^{n_k} \right) \langle \alpha | \prod_p \mathcal{H}_{k(p)} | \alpha \rangle. \quad (\text{A.3})$$

Here, $n_{k'} = \sum_p \delta_{k(p), k'}$ is the number of vertices in the operator string corresponding to \mathcal{H}_k such that $\sum_k n_k = n$. Taking the derivative is straightforward. We obtain:

$$\langle \mathcal{H}_k \rangle = \frac{\langle n_k \rangle}{\beta}. \quad \Rightarrow \quad \langle \mathcal{H} \rangle = \text{const} - \frac{\langle n \rangle}{\beta}. \quad \checkmark \quad (\text{A.4})$$

Moreover, we can calculate another estimator for the temperature derivative:

$$\frac{\partial}{\partial \beta} \langle \mathcal{H}_k \rangle = \frac{\partial}{\partial \beta} \left[\frac{1}{\mathcal{Z} [\beta]} \sum_{n=0}^{\infty} \sum_{|\alpha\rangle} \sum_{S_n} \frac{\beta^{n-1}}{n!} n_k \langle \alpha | \prod_p \mathcal{H}_{k(p)} | \alpha \rangle \right] \quad (\text{A.5})$$

$$= \left\langle \frac{n_k(n-1)}{\beta^2} \right\rangle - \left\langle \frac{n_k}{\beta} \right\rangle \left\langle \frac{n}{\beta} \right\rangle. \quad (\text{A.6})$$

$$\frac{\partial}{\partial T} \langle \mathcal{H}_k \rangle = -(\langle n_k n \rangle - \langle n_k \rangle (1 + \langle n \rangle)). \quad (\text{A.7})$$

Using a separate estimator for the temperature derivative is basically possible for any observable and usually leads to less statistical noise than estimating it from separate simulations at (slightly) different temperatures.

B Derivation of Formulas for Generalized Susceptibilities

In the following we derive an estimator in the SSE framework for the generalized susceptibility

$$\chi_{\hat{A}\hat{B}} = \left. \frac{\partial}{\partial b} \right|_{b=0} \langle \hat{A} \rangle_{\mathcal{H}+b\hat{B}}. \quad (\text{B.1})$$

Using the definition of the thermodynamic average $\langle \cdot \rangle$ we have

$$\chi_{\hat{A}\hat{B}} = \left. \frac{\partial}{\partial b} \right|_{b=0} \frac{\text{Tr} \left\{ e^{-\beta\mathcal{H}-\beta b\hat{B}} \hat{A} \right\}}{\mathcal{Z}(b)} \quad \text{with } \mathcal{Z}(b) = \text{Tr} \left\{ e^{-\beta\mathcal{H}-\beta b\hat{B}} \right\} \quad (\text{B.2})$$

$$= \frac{1}{\mathcal{Z}(b)} \left. \frac{\partial}{\partial b} \right|_{b=0} \text{Tr} \left\{ e^{-\beta\mathcal{H}-\beta b\hat{B}} \hat{A} \right\} - \frac{1}{(\mathcal{Z}(b))^2} \text{Tr} \left\{ e^{-\beta\mathcal{H}} \beta \hat{B} \right\} \text{Tr} \left\{ e^{-\beta\mathcal{H}} \hat{A} \right\} \quad (\text{B.3})$$

$$= \int_0^\beta \langle \hat{A}(\tau) \hat{B} \rangle d\tau - \beta \langle \hat{A} \rangle \langle \hat{B} \rangle \quad \text{with } \hat{A}(\tau) = e^{\tau\mathcal{H}} \hat{A} e^{-\tau\mathcal{H}}. \quad (\text{B.4})$$

The last equation is the Kubo formula. We will first focus on \hat{A} and \hat{B} diagonal in the chosen S^z basis. Now we basically need to repeat the steps of section 3.2 and Taylor-expand both appearing exponentials:

$$\langle \hat{A}(\tau) \hat{B}(0) \rangle = \sum_{|\alpha\rangle} \langle \alpha | e^{-(\beta-\tau)\mathcal{H}} \hat{A} e^{\tau\mathcal{H}} \hat{B} | \alpha \rangle \quad (\text{B.5})$$

$$= \sum_{|\alpha\rangle} \sum_{\tilde{n}=0}^{\infty} \sum_{\tilde{m}=0}^{\infty} \frac{(\beta-\tau)^{\tilde{n}} \tau^{\tilde{m}}}{\tilde{n}! \tilde{m}!} \langle \alpha | \mathcal{H}^{\tilde{n}} \hat{A} \mathcal{H}^{\tilde{m}} \hat{B} | \alpha \rangle \quad (\text{B.6})$$

We can reorder the sums over \tilde{n} and \tilde{m} by relabeling $(n, m) := (\tilde{m} + \tilde{n}, \tilde{m})$, where we restrict the sum over m to numbers $0, \dots, n$. Then we can identify n as the total operator string length. With the shorthand for the matrix elements $A_p := \langle \alpha(p) | \hat{A} | \alpha(p) \rangle$ we have

$$\langle \hat{A}(\tau) \hat{B}(0) \rangle = \sum_{|\alpha\rangle} \sum_{n=0}^{\infty} \sum_{m=0}^n \frac{(\beta-\tau)^{n-m} \tau^m}{(n-m)! m!} \langle \alpha | \mathcal{H}^{n-m} \hat{A} \mathcal{H}^m \hat{B} | \alpha \rangle \quad (\text{B.7})$$

$$= \sum_{|\alpha\rangle} \sum_{n=0}^{\infty} \sum_{S_n} \sum_{m=0}^n \frac{(\beta-\tau)^{n-m} \tau^m}{(n-m)! m!} \langle \alpha | \prod_{p=m}^n \mathcal{H}_{b(p)}^{a(p)} \underbrace{\hat{A}}_{A_m} \prod_{p=0}^{m-1} \mathcal{H}_{b(p)}^{a(p)} \underbrace{\hat{B}}_{B_0} | \alpha \rangle \quad (\text{B.8})$$

$$= \left\langle \sum_{m=0}^n \frac{(\beta-\tau)^{n-m} \tau^m}{(n-m)! m!} \frac{n!}{\beta^n} A_m B_0 \right\rangle. \quad (\text{B.9})$$

In the last line we introduced the weighted average of the SSE by comparison to eq. (3.30). The integral over τ can be performed with a repeated integration by parts and yields

$$\int_0^\beta (\beta - \tau)^{n-m} \tau^m d\tau = \frac{m!(n-m)!}{n!} \frac{\beta^{n+1}}{n+1}. \quad (\text{B.10})$$

Thus, the combinatorial prefactors cancel to $\frac{\beta}{n+1}$. We can further average over the propagated states as in eq. (3.54) and split the sum $\sum_{m=0}^n \rightarrow \sum_{m=0}^{n-1} + \delta_{m,n}$, such that we arrive at

$$\int_0^\beta \langle \hat{A}(\tau) \hat{B}(0) \rangle d\tau = \frac{\beta}{n(n+1)} \left\langle \left(\sum_{p=0}^{n-1} A_p \right) \left(\sum_{p=0}^{n-1} B_p \right) + \sum_{p=0}^{n-1} A_p B_p \right\rangle. \quad (\text{B.11})$$

A similar calculation can be carried out for $\hat{A} = \mathcal{H}_b^a$ and $\hat{B} = \mathcal{H}_{\tilde{b}}$. In this case care has to be taken during the introduction of the operator string: we use $n-2$ operators for the weight and 2 as observables themselves. Since m labels the power of \mathcal{H} between A and B , it runs only from 0 to $n-2$ and the analog to eq. (B.9) is

$$\langle \hat{A}(\tau) \hat{B}(0) \rangle = \left\langle \sum_{m=0}^{n-2} \frac{(\beta - \tau)^{n-2-m} \tau^m}{(n-2-m)! m!} \frac{n!}{\beta^n} \delta_{a(m+1),a} \delta_{b(m+1),b} \delta_{a(0),\tilde{a}} \delta_{b(0),\tilde{b}} \right\rangle \quad (\text{B.12})$$

The Kronecker deltas make sure that the considered operators \hat{A} and \hat{B} are at the places $p=0$ and $p=m+1$. As before we can average over p . Then (B.12) involves counting for any position p of \hat{B} the number of operators \hat{A} at any *other* position $p' \neq p$. In analogy to the diagonal operators we derive [23, 52]

$$\chi_{\mathcal{H}_b^a, \mathcal{H}_{\tilde{b}}} = \frac{1}{\beta} \langle (n_b^a - \delta_{a,\tilde{a}} \delta_{b,\tilde{b}}) n_{\tilde{b}}^{\tilde{a}} \rangle - \frac{1}{\beta} \langle n_b^a \rangle \langle n_{\tilde{b}}^{\tilde{a}} \rangle. \quad (\text{B.13})$$

As a check we can relate the specific heat to the susceptibility with $\hat{A} = \hat{B} = \mathcal{H}$:

$$C_V = \frac{\partial \langle \mathcal{H} \rangle}{\partial T} = \beta \chi_{\mathcal{H}, \mathcal{H}} = \beta \sum_{a,b} \sum_{\tilde{a}, \tilde{b}} \chi_{\mathcal{H}_b^a, \mathcal{H}_{\tilde{b}}} \quad (\text{B.14})$$

$$= \left\langle \left(\sum_{a,b} n_b^a - 1 \right) \sum_{\tilde{a}, \tilde{b}} n_{\tilde{b}}^{\tilde{a}} \right\rangle - \left\langle \sum_{a,b} n_b^a \right\rangle \left\langle \sum_{\tilde{a}, \tilde{b}} n_{\tilde{b}}^{\tilde{a}} \right\rangle. \quad \checkmark \quad (\text{B.15})$$

C Energy of the Trial State for Coupled Dimers

In this appendix we calculate the energy $E^{\text{trial}}(s) := \langle \Psi^{\text{trial}} | \mathcal{H} | \Psi^{\text{trial}} \rangle / N$ for the trial state given in eq. (4.30) as:

$$|\Psi^{\text{trial}}\rangle = \prod_{\text{dimers } \langle i,j \rangle} \frac{1}{\sqrt{1+s^2}} (|\uparrow_i \downarrow_j\rangle - s |\downarrow_i \uparrow_j\rangle) \quad (\text{C.1})$$

$$\equiv \prod_{\text{dimers } \langle i,j \rangle} |(i,j); s\rangle \quad \text{with } |(i,j); s\rangle := |\uparrow_i \downarrow_j\rangle - s |\downarrow_i \uparrow_j\rangle. \quad (\text{C.2})$$

Here, we write the Hamiltonian in the following form:

$$\mathcal{H} = \sum_{\text{dimers } \langle i,j \rangle} J \mathcal{H}_{i,j} + \sum_{\text{non-dimers } \langle i,j \rangle} J' \mathcal{H}_{i,j} \quad \text{with} \quad (\text{C.3})$$

$$\mathcal{H}_{i,j} := \frac{1}{2} (S_i^- S_j^+ + S_i^+ S_j^-) + \Delta S_i^z S_j^z. \quad (\text{C.4})$$

We consider first the part of the Hamiltonian $\mathcal{H}_{i,j}$ where i and j belong to the same dimer. We obtain:

$$\langle \Psi^{\text{trial}} | \mathcal{H}_{i,j} | \Psi^{\text{trial}} \rangle = \frac{1}{1+s^2} \langle (i,j); s | \mathcal{H}_{i,j} | (i,j); s \rangle \quad (\text{C.5})$$

$$= \frac{1}{1+s^2} \left[\frac{1}{2} (-s \langle \downarrow_i \uparrow_j | S_i^- S_j^+ | \uparrow_i \downarrow_j \rangle - s \langle \uparrow_i \downarrow_j | S_i^+ S_j^- | \downarrow_i \uparrow_j \rangle) \right. \\ \left. + \Delta (\langle \uparrow_i \downarrow_j | S_i^z S_j^z | \uparrow_i \downarrow_j \rangle + s^2 \langle \downarrow_i \uparrow_j | S_i^z S_j^z | \downarrow_i \uparrow_j \rangle) \right] \quad (\text{C.6})$$

$$= \frac{-s}{1+s^2} - \frac{\Delta}{4}. \quad (\text{C.7})$$

This part has two extrema at $s = \pm 1$. It favors the singlet state at $s = +1$ over the triplet ($s = -1$).

Next, we consider a weak bond $\langle j, k \rangle$ between two dimers $\langle i, j \rangle$ and $\langle k, l \rangle$. Note that the weak bonds in the considered geometry connect always a left spin in one dimer with a right spin in the other dimer (where "left" and "right" refer to the position within the dimer), see Figure 2.2(b). The off-diagonal part of the bond vanishes: the action of single S_j^+ or S_j^- leads to parallel spins on the dimer $\langle i, j \rangle$ (or vanishes immediately), which has no overlap with the original trial state itself. We have:

$$\langle \Psi^{\text{trial}} | \mathcal{H}_{j,k} | \Psi^{\text{trial}} \rangle = \frac{\Delta}{(1+s^2)^2} \langle (i,j)(k,l); s | S_j^z S_k^z | (i,j)(k,l); s \rangle, \quad (\text{C.8})$$

$$|(i,j)(k,l); s\rangle = |\uparrow_i \downarrow_j \uparrow_k \downarrow_l\rangle + s^2 |\downarrow_i \uparrow_j \downarrow_k \uparrow_l\rangle - s |\uparrow_i \downarrow_j \downarrow_k \uparrow_l\rangle - s |\downarrow_i \uparrow_j \uparrow_k \downarrow_l\rangle, \quad (\text{C.9})$$

$$\langle \Psi^{\text{trial}} | \mathcal{H}_{j,k} | \Psi^{\text{trial}} \rangle = \frac{\Delta}{(1+s^2)^2} \left(-\frac{1}{4} - \frac{s^4}{4} + \frac{s^2}{4} + \frac{s^2}{4} \right) = -\frac{\Delta}{4} \left(\frac{1-s^2}{1+s^2} \right)^2. \quad (\text{C.10})$$

This part has two maxima at $s = \pm 1$ and one minimum at $s = 0$.

The unit cell of the lattice consists of one dimer (i.e., two sites and a strong bond) and three weak bonds. The total energy per site is thus

$$E^{\text{trial}}(s) \equiv \frac{\langle \Psi^{\text{trial}} | \mathcal{H} | \Psi^{\text{trial}} \rangle}{N} = -\frac{J}{2} \left(\frac{s}{1+s^2} + \frac{\Delta}{4} \right) - \frac{3J'\Delta}{2 \cdot 4} \left(\frac{1-s^2}{1+s^2} \right)^2. \quad (\text{C.11})$$

This result is shown for different J' (and $\Delta = 4$) in Figure 4.15. A straightforward calculation gives the derivative:

$$\frac{\partial E(s)}{\partial s} = -\frac{1-s^2}{(1+s^2)^3} \left(\frac{J}{2}(1+s^2) - 6J'\Delta s \right). \quad (\text{C.12})$$

Beside the two zeros at $s = \pm 1$, there may exist two further zeros at

$$s_{\pm} = \frac{3J'\Delta}{J} \pm \sqrt{\left(\frac{3J'\Delta}{J} \right)^2 - 1} \quad \text{for } J' > \frac{J}{3\Delta} =: J_c^{\text{trial}}. \quad (\text{C.13})$$

The energy $E^{\text{trial}}(s)$ is always maximal at $s = -1$. For $J < J_c^{\text{trial}}$, only $s = 1$ is the single minimum. Above J_c^{trial} , two equivalent (see above) minima exists – one for the negative sign in eq. (C.13) ($s_- < 1$) and one for the positive sign ($s_+ = \frac{1}{s_-} > 1$). The minima move continuously away from $s = 1$, but over a very short range of J' they move quite far.

Staggered Magnetization For completeness, we also calculate the staggered magnetization of the trial state. We find:

$$\phi_i S_i^z |(i, j)(s)\rangle = \phi_j S_j^z |(i, j)(s)\rangle = \frac{1}{2} |\uparrow_i \downarrow_j\rangle - s \left(-\frac{1}{2} \right) |\downarrow_i \uparrow_j\rangle, \quad (\text{C.14})$$

$$\langle \Psi^{\text{trial}} | m_s | \Psi^{\text{trial}} \rangle = \frac{1}{N} \sum_i \frac{1}{1+s^2} \langle (i, j)(s) | \phi_i S_i^z |(i, j)(s)\rangle = \frac{1}{2} \frac{1-s^2}{1+s^2}. \quad (\text{C.15})$$

For $M_s^2 = \sum_i \sum_j \phi_i S_i^z \phi_j S_j^z$ we need to distinguish whether i and j belong to the same dimer. On a single dimer, the spins are anti-aligned for any s :

$$\langle (i, j)(s) | \phi_i S_i^z \phi_j S_j^z |(i, j)(s)\rangle = \langle (i, j)(s) | (\phi_i S_i^z)^2 |(i, j)(s)\rangle = \frac{1}{4} (1+s^2). \quad (\text{C.16})$$

For i and j on different dimers, we find:

$$\langle (i, k)(j, l)(s) | \phi_i S_i^z \phi_j S_j^z |(i, k)(j, l)(s)\rangle = \frac{1}{4} (1+s^4 - 2s^2) = \frac{1}{4} (1-s^2)^2. \quad (\text{C.17})$$

For each i , there are two spins j on the same dimer as i and $N-2$ spins on another dimer.

All in all, we have for the square staggered magnetization the following expression:

$$\langle \Psi^{\text{trial}} | m_s^2 | \Psi^{\text{trial}} \rangle = \frac{1}{N} \sum_i \left(\frac{1}{N} \underbrace{\sum_{j \in \text{dimer } i}}_{=2} \frac{1}{4} \frac{1+s^2}{1+s^2} + \frac{1}{N} \underbrace{\sum_{j \notin \text{dimer } i}}_{=N-2} \frac{1}{4} \frac{(1-s^2)^2}{(1+s^2)^2} \right) \quad (\text{C.18})$$

$$= \left(\frac{1}{2} \frac{1-s^2}{1+s^2} \right)^2 \left(1 - \frac{2}{N} \left(\left(\frac{1+s^2}{1-s^2} \right)^2 - 1 \right) \right). \quad (\text{C.19})$$

D Perturbation Theory for large Δ

In this appendix, we apply standard perturbation theory to the ground state of the (spatially isotropic) Heisenberg Hamiltonian with anisotropic couplings for large Δ , aiming at the Rényi entanglement entropy at zero temperature. We expand around $\Delta = \infty$ corresponding to the classical Ising model such that $\epsilon := \frac{1}{\Delta} \ll 1$ is a small parameter. With the normalization used before, the energies diverge in the limit $\Delta \rightarrow \infty$, thus we measure all energies in terms of $J\Delta \equiv 1$ instead of J in this appendix. We split the Hamiltonian as (cf. eq. (2.19)):

$$\frac{\mathcal{H}}{J\Delta} = \frac{1}{\Delta} \frac{1}{2} \underbrace{\sum_{\langle i,j \rangle} (S_i^- S_j^+ + S_i^+ S_j^-)}_{\mathcal{H}_1} + \underbrace{\sum_{\langle i,j \rangle} S_i^z S_j^z}_{\mathcal{H}_0} = \mathcal{H}_0 + \epsilon \mathcal{H}_1 \quad (\text{D.1})$$

The ground states of \mathcal{H}_0 are the classical Néel states $|\text{gs.}^0; \uparrow\rangle = |\uparrow\downarrow\uparrow \dots\rangle$ and $|\text{gs.}^0; \downarrow\rangle = |\downarrow\uparrow\downarrow \dots\rangle$. The first excited states are obtained by flipping a single spin on site i relative to the Néel states, which we denote with $|i; \uparrow\rangle$ and $|i; \downarrow\rangle$, respectively. They have an excitation energy $\frac{z}{2}$, where z is the number of next neighbors. For simplicity, we restrict ourselves here to 2D square lattices of size $L \times L = N$ with periodic boundary conditions such that $z = 4$ (and $N^b = \frac{zN}{2} = 2N$). However, these excitations are in the $S^z = 1$ sector and \mathcal{H}_1 commutes with S^z ; thus, these excitations do not change the ground states to first order in perturbation theory. Consequently, we consider the next excitations where a pair of spins on neighboring sites is flipped; which we denote with $|i, j; \uparrow\rangle$ and $|i, j; \downarrow\rangle$. On the 2D square lattice, they have an excitation energy of $\frac{6}{2}J\Delta = 3$ above the ground state. Note that these excitations are the only states created by a single application of \mathcal{H}_1 on the classical Néel states, thus we have to first order in ϵ :

$$\begin{aligned} |\text{gs.}; \uparrow\rangle &= |\text{gs.}^0\rangle + \epsilon |\text{gs.}^1; \uparrow\rangle + \mathcal{O}(\epsilon^2) = |\text{gs.}^0; \uparrow\rangle + \epsilon \sum_{\langle i,j \rangle} |i, j; \uparrow\rangle \frac{\langle i, j; \uparrow | \mathcal{H}_1 | \text{gs.}, \uparrow \rangle}{E_{\text{gs.}}^0 - E_{|i,j;\uparrow\rangle}^0} + \mathcal{O}(\epsilon^2) \\ &= |\text{gs.}^0; \uparrow\rangle - \frac{\epsilon}{6} \sum_{\langle i,j \rangle} |i, j; \uparrow\rangle. \end{aligned} \quad (\text{D.2})$$

Here, we have assumed that the system has more than 4 sites such that $\langle i, j; \uparrow | \mathcal{H}_1 | \text{gs.}; \downarrow \rangle = 0$. An analogous equation holds for $|\text{gs.}; \downarrow\rangle$. Note that $|\text{gs.}; \uparrow\rangle$ is not normalized, instead we have

$$\mathcal{N}^2 := \langle \text{gs.}; \uparrow | \text{gs.}; \uparrow \rangle = 1 + N_{A \cup B}^b \frac{\epsilon^2}{36}. \quad (\text{D.3})$$

The density matrix at zero temperature is a mixture between the two degenerate ground states:

$$\hat{\rho} \mathcal{N}^2 = \frac{1}{2} |\text{gs.}; \uparrow\rangle \langle \text{gs.}; \uparrow| + \frac{1}{2} |\text{gs.}; \downarrow\rangle \langle \text{gs.}; \downarrow|. \quad (\text{D.4})$$

For the second Rényi entropy $S_2(A) = -\log(\hat{\rho}_A^2)$ of a region A , we need the reduced density matrix $\hat{\rho}_A = \text{Tr}_B \{\hat{\rho}\}$. First, we focus on $|\text{gs.}; \uparrow\rangle \langle \text{gs.}; \uparrow|$ and simplify notation by leaving away \uparrow , which indicates the degeneracy of the ground state. We denote the states in region A with a subindex A , e.g., $|\text{gs.}^0\rangle_A$ is the classical Néel state in region A such that $\text{Tr}_B \{|\text{gs.}^0\rangle \langle \text{gs.}^0|\} = |\text{gs.}^0\rangle_A \langle \text{gs.}^0|_A$. We obtain:

$$\begin{aligned} \text{Tr}_B \{|\text{gs.}\rangle \langle \text{gs.}|\} &= \text{Tr}_B \left\{ |\text{gs.}^0\rangle \langle \text{gs.}^0| - \frac{\epsilon}{6} \sum_{\langle i,j \rangle} (|\text{gs.}^0\rangle \langle i,j| + \text{h.c.}) + \frac{\epsilon^2}{36} \sum_{\langle i,j \rangle, \langle kl \rangle} |i,j\rangle \langle k,l| \right\} \\ &= |\text{gs.}^0\rangle_A \langle \text{gs.}^0|_A - \frac{\epsilon}{6} \sum_{\langle i,j \rangle_A} (|\text{gs.}^0\rangle \langle i,j| + \text{h.c.}) \\ &\quad + \frac{\epsilon^2}{36} \left(N_B^b |\text{gs.}^0\rangle \langle \text{gs.}^0| + \sum_{\langle i,j \rangle_A} |i,j\rangle_A \langle i,j|_A \right. \\ &\quad \left. + \sum_{\langle i,j \rangle_{\partial A}} |i\rangle_A \langle i|_A + \sum_{\langle i,j \rangle_{\partial A} \langle k,j \rangle_{\partial A}; i \neq j} |i\rangle_A \langle k|_A \right). \end{aligned} \quad (\text{D.5})$$

Here, $\langle i,j \rangle_A$ indicates any pair of two neighboring sites in A and $\langle i,j \rangle_{\partial A}$ means that i is in A and j in B , i.e., the last sum contains corner terms and the second last sum corresponds to boundary terms. $N_B^b = \sum_{\langle i,j \rangle_B} 1$ is the number of bonds within region B . Note that the sum proportional to ϵ in the first line is purely off-diagonal such that it produces no boundary term. To proceed, we have to calculate the square $\hat{\rho}_A^2$. The boundary terms in the last line will then be of order ϵ^4 – there are no cross terms with lower orders –, and we omit them:

$$\begin{aligned} (\text{Tr}_B \{|\text{gs.}\rangle \langle \text{gs.}|\})^2 &= \left(1 + N_B^b \frac{\epsilon^2}{36} \right)^2 |\text{gs.}^0\rangle_A \langle \text{gs.}^0|_A - \frac{\epsilon}{6} \sum_{\langle i,j \rangle_A} (|\text{gs.}^0\rangle_A \langle i,j|_A + \text{h.c.}) \\ &\quad + N_A^b \frac{\epsilon^2}{36} |\text{gs.}^0\rangle_A \langle \text{gs.}^0|_A + \frac{\epsilon^2}{36} \sum_{\langle i,j \rangle_A} |i,j\rangle_A \langle i,j|_A + \mathcal{O}(\epsilon^3). \end{aligned} \quad (\text{D.6})$$

Finally, we take the trace over the remaining states. The terms proportional to ϵ vanish since they are off-diagonal. One can easily see that the terms proportional to ϵ^3 – which we omitted – are off-diagonal as well, such that the result is actually correct up to ϵ^4 :

$$\begin{aligned} \text{Tr} \{(\text{Tr}_B \{|\text{gs.}, \uparrow\rangle \langle \text{gs.}, \uparrow|\})^2\} &= \left(1 + N_B^b \frac{\epsilon^2}{36} \right)^2 + 2N_A^b \frac{\epsilon^2}{36} + \mathcal{O}(\epsilon^4) \\ &= 1 + 2(N_B^b + N_A^b) \frac{\epsilon^2}{36} + \mathcal{O}(\epsilon^4). \end{aligned} \quad (\text{D.7})$$

Including the normalization eq. (D.3), we obtain the Rényi entropy corresponding to one of the two ground states:

$$\begin{aligned} S_2 \left(\frac{|\text{gs.}\rangle_A \langle \text{gs.}|_A}{\mathcal{N}^2} \right) &= \log \left(\frac{\mathcal{N}^4}{\text{Tr} \{(\text{Tr}_B \{|\text{gs.}, \uparrow\rangle \langle \text{gs.}, \uparrow|\})^2\}} \right) \\ &= 2 \frac{\epsilon^2}{36} \underbrace{(N_{A \cup B}^b - N_A^b - N_B^b)}_{=N_{\partial A}^b} + \mathcal{O}(\epsilon^4). \end{aligned} \quad (\text{D.8})$$

In the last step, we have expanded the logarithm around 1 for small ϵ . As expected, we obtain an area law; the entanglement entropy comes from the bonds between A and B . However, the prefactor is quite small; for $\Delta = 4$ it is just $\frac{1}{288}$.

If we assume that the region A is large enough such the states $|i, j, \uparrow\rangle_A$ and $|i, j, \downarrow\rangle_A$ have no overlap (i.e., the region A is larger than four sites), the degeneracy $\Omega_0 = 2$ of the ground state leads only to an additive constant: in that case there appear no cross terms in the square of the reduced density matrix, $\hat{\rho}_A^2 = (\text{Tr}_B \{|\text{gs.}; \uparrow\rangle \langle \text{gs.}; \uparrow|\})^2/4 + (\text{Tr}_B \{|\text{gs.}; \downarrow\rangle \langle \text{gs.}; \downarrow|\})^2/4$. Clearly, the traces are identical for \uparrow and \downarrow , such that we obtain:

$$S_2(\hat{\rho}_A^2) = \log(2) + \frac{\epsilon^2}{18} N_{\partial A}^b + \mathcal{O}(\epsilon^4). \quad (\text{D.9})$$

E List of Tables

2.1	Eigenstates and -energies of a single bond	18
3.1	Action and pictorial representation of operators in operatorstring	27
4.1	Moments of a normal distribution	57

F List of Figures

2.1	Possible hoppings of electrons	14
2.2	Spatially anisotropic couplings	19
3.1	Outer algorithmic structure for Monte Carlo simulations	25
3.2	Example of an operator string.	28
3.3	Distribution of operator string length n during a Monte Carlo simulation	30
3.4	Legs of a vertex	31
3.5	Example of a diagonal update	32
3.6	Example of deterministic loop updates	34
3.7	Example of the choice of an exit leg	36
3.8	Coupling between directed loop equations	37
3.9	Comparison of largest bounce probabilities	39
3.10	Energy on an 1D chain for different Δ , compared to ED	43
3.11	Difference between ED and SSE data	44
4.1	Staggered magnetization, 2D square lattice	45
4.2	Squared staggered magnetization and susceptibility, 2D square lattice	46
4.3	Energy, 2D square lattice	47
4.4	Specific heat, 2D square lattice	48
4.5	Decay of spin-spin correlations, 2D square lattice	49
4.6	Correlation length, 2D square lattice	50
4.7	Finite size scaling to obtain T_c , 2D square lattice	54
4.8	Data collapse of susceptibility through rescaling of axes, 2D square lattice	55
4.9	Binder cumulant, 2D square lattice	56
4.10	Rescaled Binder cumulant, 2D square lattice	59
4.11	Energy, coupled chains	59
4.12	Squared staggered magnetization, 1D chain, PBC	60
4.13	Squared staggered magnetization, coupled chains	61
4.14	Phase diagram for 1D-2D-crossover, coupled chains	62
4.15	Energy of trial state for coupled dimers	64
4.16	Energy, coupled dimers	66
4.17	Squared staggered magnetization, coupled dimers	66
4.18	Phase diagram for coupled dimers	68
5.1	Partion of the 2D square lattice into different regions	71
5.2	Addition of entropies on regions A and B	73
5.3	Rényi entropy for distribution $\{p, 1 - p\}$	75
5.4	Geometry of the replica trick	78
5.5	Rényi entropy for different subregions	82
5.6	Rényi mutual information, $\Delta = 4$, $A = h \times L$	83

5.7	Finite size dependence of the crossings of the mutual information	86
5.8	Mutual information divided by boundary length for $\Delta \in \{0, 1\}$	87
5.9	Mutual information divided by boundary length for coupled chains and dimers	88
5.10	Scheme for increment of subregions for ratio trick	92
5.11	Rényi entropy $T = 0.5$ for different sizes of A	94
6.1	Singlet and Triplet projectors, 2D square lattice	98
6.2	Derivative $\langle \hat{P}_b^{\text{ap}} \rangle$, 2D square lattice	99
6.3	Data collapse of derivative of projectors	100
6.4	Finite Size Scaling to extract T_c for projections, 2D square lattice	102
6.5	Correlation of projectors on different bonds, 2D square lattice	103
6.6	Spatial decay of projector-correlations at T_c , 2D square lattice	104
6.7	Local and global susceptibility of projectors, 2D square lattice	105
6.8	Local projections \hat{P}^{ap} and \hat{Q} , coupled chains	108
6.9	Global susceptibility of \hat{P}^{ap} , coupled chains	109
6.10	Local projections \hat{P}^{ap} and \hat{Q} , coupled dimers	112
6.11	Global susceptibility of \hat{P}^{ap} , coupled dimers	114

G Bibliography

- [1] R. G. Melko, A. B. Kallin, and M. 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. doi: 10.1103/PhysRevB.82.100409.
- [2] D. Greif, T. Uehlinger, G. Jotzu, L. Tarruell, and T. Esslinger. Short-Range Quantum Magnetism of Ultracold Fermions in an Optical Lattice. *Science*, 340(6138):1307, 2013. doi: 10.1126/science.1236362.
- [3] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31:253, 1925. doi: 10.1007/BF02980577.
- [4] L. Onsager. Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition. *Phys. Rev.*, 65:117, 1944. doi: 10.1103/PhysRev.65.117.
- [5] J. Keeling. Lecture Notes: Quantum Magnetism, 2008. URL <http://www.st-andrews.ac.uk/~jmjk/keeling/teaching/magnetism-notes.pdf>.
- [6] H. Bethe. Zur Theorie der Metalle. *Zeitschrift für Physik*, 71:205, 1931. doi: 10.1007/BF01341708.
- [7] N. D. Mermin and H. Wagner. Absence of Ferromagnetism or Antiferromagnetism in One- or Two-Dimensional Isotropic Heisenberg Models. *Phys. Rev. Lett.*, 17:1133, 1966. doi: 10.1103/PhysRevLett.17.1133.
- [8] E. Manousakis. The spin- $\frac{1}{2}$ Heisenberg antiferromagnet on a square lattice and its application to the cuprous oxides. *Rev. Mod. Phys.*, 63:1, 1991. doi: 10.1103/RevModPhys.63.1.
- [9] U. Schollwöck, J. Richter, D. J. J. Farnell, and R. F. Bishop. *Quantum Magnetism*. Lecture notes in physics. Springer, 2004. ISBN 9783540214229.
- [10] P. W. Anderson. Resonating valence bonds: A new kind of insulator? . *Materials Research Bulletin*, 8:153, 1973. doi: 10.1016/0025-5408(73)90167-0.
- [11] R. R. P. Singh, M. P. Gelfand, and D. A. Huse. Ground States of Low-Dimensional Quantum Antiferromagnets. *Phys. Rev. Lett.*, 61:2484, 1988. doi: 10.1103/PhysRevLett.61.2484.
- [12] S. Sachdev. Quantum magnetism and criticality. *Nat Phys*, 4:173, 2008. doi: 10.1038/nphys894.
- [13] J. G. Bednorz and K. A. Müller. Possible high T_c superconductivity in the Ba-La-Cu-O system. *Zeitschrift für Physik B Condensed Matter*, 64:189, 1986. doi: 10.1007/BF01303701.

- [14] P. W. Anderson. The Resonating Valence Bond State in La_2CuO_4 and Superconductivity. *Science*, 235:1196, 1987. doi: 10.1126/science.235.4793.1196.
- [15] Y. Kamihara, T. Watanabe, M. Hirano, and H. Hosono. Iron-Based Layered Superconductor $\text{La}[\text{O}_{1-x}\text{F}_x]\text{FeAs}$ ($x = 0.05\text{--}0.12$) with $T_c = 26$ K. *Journal of the American Chemical Society*, 130:3296, 2008. doi: 10.1021/ja800073m.
- [16] A. J. Leggett. What DO we know about high T_c ? *Nat Phys*, 2:134, 2006. doi: 10.1038/nphys254.
- [17] M. Reehuis, C. Ulrich, K. Prokeš, A. Gozar, G. Blumberg, S. Komiyama, Y. Ando, P. Pattison, and B. Keimer. Crystal structure and high-field magnetism of La_2CuO_4 . *Phys. Rev. B*, 73:144513, 2006. doi: 10.1103/PhysRevB.73.144513.
- [18] J. S. Miller and A. J. Epstein. Molecule-Based Magnets—An Overview. *MRS Bulletin*, 25:21, 2000. doi: 10.1557/mrs2000.221.
- [19] M. Troyer and U.-J. Wiese. Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations. *Phys. Rev. Lett.*, 94:170201, 2005. doi: 10.1103/PhysRevLett.94.170201.
- [20] U. Schollwöck. The density-matrix renormalization group. *Rev. Mod. Phys.*, 77:259, 2005. doi: 10.1103/RevModPhys.77.259.
- [21] D. C. Handscomb. The Monte Carlo method in quantum statistical mechanics. *Proc. Cambridge Phil. Soc.*, 58:594, 1962.
- [22] A. W. Sandvik and J. Kurkijärvi. Quantum Monte Carlo simulation method for spin systems. *Phys. Rev. B*, 43:5950, 1991. doi: 10.1103/PhysRevB.43.5950.
- [23] A. W. Sandvik. A generalization of Handscomb’s quantum Monte Carlo scheme—application to the 1D Hubbard model. *Journal of Physics A: Mathematical and General*, 25:3667, 1992.
- [24] V. L. Berezinskii. Destruction of Long-range Order in One-dimensional and Two-dimensional Systems having a Continuous Symmetry Group I. Classical Systems. *Soviet Journal of Experimental and Theoretical Physics*, 32:493, 1971.
- [25] J. M. Kosterlitz and D. J. Thouless. Ordering, metastability and phase transitions in two-dimensional systems. *Journal of Physics C: Solid State Physics*, 6:1181, 1973.
- [26] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379, 1948. doi: 10.1002/j.1538-7305.1948.tb01338.x.
- [27] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer, Berlin, second edition, 1996.
- [28] A. Rényi. *On Measures of Entropy and Information*, page 547. University of California Press, Berkeley, Calif., 1961.
- [29] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium. *Rev. Mod. Phys.*, 82:277, 2010. doi: 10.1103/RevModPhys.82.277.

- [30] W. Ketterle, D. S. Durfee, and D. M. Stamper-Kurn. Making, probing and understanding Bose-Einstein condensates, April 1999. URL <http://arxiv.org/abs/cond-mat/9904034v2>.
- [31] R. A. Hart, P. M. Duarte, T.-L. Yang, X. Liu, T. Paiva, E. Khatami, R. T. Scalettar, N. Trivedi, D. A. Huse, and R. G. Hulet. Observation of antiferromagnetic correlations in the Hubbard model with ultracold atoms, 2014. URL <http://arxiv.org/pdf/1407.5932v1>.
- [32] C. Chin, R. Grimm, P. Julienne, and E. Tiesinga. Feshbach resonances in ultracold gases. *Rev. Mod. Phys.*, 82:1225, 2010. doi: 10.1103/RevModPhys.82.1225.
- [33] W. Ketterle and M. W. Zwierlein. Making, probing and understanding ultracold Fermi gases. *La Rivista del Nuovo Cimento*, page 247, 2008. doi: 10.1393/ncr/i2008-10033-1.
- [34] T. Esslinger. Fermi-Hubbard Physics with Atoms in an Optical Lattice. *Ann. Rev. of Cond. Mat. Phys.*, 1:129, 2010. doi: 10.1146/annurev-conmatphys-070909-104059.
- [35] J. Hubbard. Electron Correlations in Narrow Energy Bands. *Proc. of Royal Soc. of London A*, 276:238, 1963. doi: 10.1098/rspa.1963.0204.
- [36] M. Greiner, O. Mandel, T. Esslinger, T. Hansch, and I. Bloch. Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms. *Nature*, 415(6867):39, 2002. doi: 10.1038/415039a.
- [37] R. Jordens, N. Strohmaier, K. Gunter, H. Moritz, and T. Esslinger. A Mott insulator of fermionic atoms in an optical lattice. *Nature*, 455(7210):204, 2008. doi: 10.1038/nature07244.
- [38] U. Schneider, L. Hackermüller, S. Will, T. Best, I. Bloch, T. A. Costi, R. W. Helmes, D. Rasch, and A. Rosch. Metallic and Insulating Phases of Repulsively Interacting Fermions in a 3D Optical Lattice. *Science*, 322(5907):1520, 2008. doi: 10.1126/science.1165449.
- [39] S. Trotzky, P. Cheinet, S. Fölling, M. Feld, U. Schnorrberger, A. M. Rey, A. Polkovnikov, E. A. Demler, M. D. Lukin, and I. Bloch. Time-Resolved Observation and Control of Superexchange Interactions with Ultracold Atoms in Optical Lattices. *Science*, 319(5861):295, 2008. doi: 10.1126/science.1150841.
- [40] B. Sciolla, A. Tokuno, S. Uchino, P. Barmettler, T. Giamarchi, and C. Kollath. Competition of spin and charge excitations in the one-dimensional Hubbard model. *Phys. Rev. A*, 88:063629, Dec 2013. doi: 10.1103/PhysRevA.88.063629. URL <http://link.aps.org/doi/10.1103/PhysRevA.88.063629>.
- [41] J. Imriška, M. Iazzi, L. Wang, E. Gull, D. Greif, T. Uehlinger, G. Jotzu, L. Tarruell, T. Esslinger, and M. Troyer. Thermodynamics and Magnetic Properties of the Anisotropic 3D Hubbard Model. *Phys. Rev. Lett.*, 112:115301, 2014. doi: 10.1103/PhysRevLett.112.115301.

- [42] C. J. M. Mathy, D. A. Huse, and R. G. Hulet. Enlarging and cooling the Néel state in an optical lattice. *Phys. Rev. A*, 86:023606, 2012. doi: 10.1103/PhysRevA.86.023606.
- [43] T. Fukuhara, A. Kantian, M. Endres, M. Cheneau, P. Schausz, S. Hild, D. Bellem, U. Schollwock, T. Giamarchi, C. Gross, I. Bloch, and S. Kuhr. Quantum dynamics of a mobile spin impurity. *Nat. Phys.*, 9:235, 2013. doi: 10.1038/nphys2561. 10.1038/nphys2561.
- [44] T. Fukuhara, P. Schausz, M. Endres, S. Hild, M. Cheneau, I. Bloch, and C. Gross. Microscopic observation of magnon bound states and their dynamics. *Nature*, 502(7469):76, 2013. doi: 10.1038/nature12541.
- [45] J. Simon, W. S. Bakr, R. Ma, M. Eric Tai, P. M. Preiss, and M. Greiner. Quantum simulation of antiferromagnetic spin chains in an optical lattice. *Nature*, 472:307, 2011. doi: 10.1038/nature09994.
- [46] J. Struck, C. Ölschläger, R. Le Targat, P. Soltan-Panahi, A. Eckardt, M. Lewenstein, P. Windpassinger, and K. Sengstock. Quantum Simulation of Frustrated Classical Magnetism in Triangular Optical Lattices. *Science*, 333:996, 2011. doi: 10.1126/science.1207239.
- [47] J. R. Schrieffer and P. A. Wolff. Relation between the Anderson and Kondo Hamiltonians. *Phys. Rev.*, 149:491, 1966. doi: 10.1103/PhysRev.149.491.
- [48] S. Bravyi, D. P. DiVincenzo, and D. Loss. Schrieffer–Wolff transformation for quantum many-body systems. *Annals of Physics*, 326:2793, 2011. doi: 10.1016/j.aop.2011.06.004.
- [49] N. N. Bogoliubov. *Physik. Abhandl. Sowjetunion*, 6(113):229, 1962.
- [50] P. C. Hohenberg. Existence of Long-Range Order in One and Two Dimensions. *Phys. Rev.*, 158:383, 1967. doi: 10.1103/PhysRev.158.383.
- [51] J. Goldstone, A. Salam, and S. Weinberg. Broken Symmetries. *Phys. Rev.*, 127:965, 1962. doi: 10.1103/PhysRev.127.965.
- [52] A. W. Sandvik. Computational Studies of Quantum Spin Systems. *AIP Conf. Proc.*, 1297:135, 2010.
- [53] M. L. Bellac, F. Mortessagne, and G. G. Batrouni. *Equilibrium and Non-Equilibrium Statistical Thermodynamics*. Cambridge University Press, 2004. ISBN 9780521821438.
- [54] K. S. D. Beach and A. W. Sandvik. Some formal results for the valence bond basis. *Nuclear Physics B*, 750:142, 2006. doi: 10.1016/j.nuclphysb.2006.05.032. URL <http://www.sciencedirect.com/science/article/pii/S0550321306004214>.
- [55] A. W. Sandvik. Ground State Projection of Quantum Spin Systems in the Valence-Bond Basis. *Phys. Rev. Lett.*, 95:207203, 2005. doi: 10.1103/PhysRevLett.95.207203.

- [56] S. Wenzel, L. Bogacz, and W. Janke. Evidence for an Unconventional Universality Class from a Two-Dimensional Dimerized Quantum Heisenberg Model. *Phys. Rev. Lett.*, 101:127202, 2008. doi: 10.1103/PhysRevLett.101.127202.
- [57] M. Campostrini, M. Hasenbusch, A. Pelissetto, P. Rossi, and E. Vicari. Critical exponents and equation of state of the three-dimensional Heisenberg universality class. *Phys. Rev. B*, 65:144520, 2002. doi: 10.1103/PhysRevB.65.144520.
- [58] L. Fritz, R. L. Doretto, S. Wessel, S. Wenzel, S. Burdin, and M. Vojta. Cubic interactions and quantum criticality in dimerized antiferromagnets. *Phys. Rev. B*, 83:174416, 2011. doi: 10.1103/PhysRevB.83.174416.
- [59] F.-J. Jiang. Monte Carlo simulations of an unconventional phase transition for a two-dimensional dimerized quantum Heisenberg model. *Phys. Rev. B*, 85:014414, 2012. doi: 10.1103/PhysRevB.85.014414.
- [60] O. F. Syljuåsen and A. W. Sandvik. Quantum Monte Carlo with directed loops. *Phys. Rev. E*, 66:046701, 2002. doi: 10.1103/PhysRevE.66.046701.
- [61] O. F. Syljuåsen. Directed loop updates for quantum lattice models. *Phys. Rev. E*, 67:046701, 2003. doi: 10.1103/PhysRevE.67.046701.
- [62] 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, 1953. doi: 10.1063/1.1699114.
- [63] H. G. Evertz. The loop algorithm. *Advances in Physics*, 52:1, 2003. doi: 10.1080/0001873021000049195.
- [64] U. Wolff. Monte Carlo errors with less errors. *Computer Physics Communications*, 156:143, 2004. doi: 10.1016/S0010-4655(03)00467-3.
- [65] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97, 1970. doi: 10.1093/biomet/57.1.97.
- [66] U. Wolff. Collective Monte Carlo Updating for Spin Systems. *Phys. Rev. Lett.*, 62:361, 1989. doi: 10.1103/PhysRevLett.62.361.
- [67] A. W. Sandvik. Stochastic series expansion method with operator-loop update. *Phys. Rev. B*, 59:R14157, 1999. doi: 10.1103/PhysRevB.59.R14157.
- [68] J. Hauschild. Exakte Diagonalisierung für das Drude-Gewicht der Spin-1/2 Heisenbergkette. Bachelor thesis, LMU Munich, 2012.
- [69] R. H. Swendsen and J.-S. Wang. Replica Monte Carlo Simulation of Spin-Glasses. *Phys. Rev. Lett.*, 57:2607, 1986. doi: 10.1103/PhysRevLett.57.2607.
- [70] A. Pelissetto and V. Ettore. Critical phenomena and renormalization-group theory. *Physics Reports*, 368:549, 2002. doi: 10.1016/S0370-1573(02)00219-3.

- [71] K. G. Wilson. Renormalization Group and Critical Phenomena. I. Renormalization Group and the Kadanoff Scaling Picture. *Phys. Rev. B*, 4:3174, 1971. doi: 10.1103/PhysRevB.4.3174.
- [72] K. G. Wilson. Renormalization Group and Critical Phenomena. II. Phase-Space Cell Analysis of Critical Behavior. *Phys. Rev. B*, 4:3184, 1971. doi: 10.1103/PhysRevB.4.3184.
- [73] L. P. Kadanoff. Scaling laws for Ising models near $T(c)$. *Physics*, 2:263, 1966.
- [74] B. Widom. Surface Tension and Molecular Correlations near the Critical Point. *The Journal of Chemical Physics*, 43:3892, 1965.
- [75] Michael E. Fisher and Michael N. Barber. Scaling Theory for Finite-Size Effects in the Critical Region. *Phys. Rev. Lett.*, 28:1516, 1972. doi: 10.1103/PhysRevLett.28.1516.
- [76] Somendra M Bhattacharjee and Flavio Seno. A measure of data collapse for scaling. *Journal of Physics A: Mathematical and General*, 34:6375, 2001.
- [77] S. G. Johnson. The NLOpt nonlinear-optimization package. <http://ab-initio.mit.edu/nlopt>. URL <http://ab-initio.mit.edu/nlopt>.
- [78] K. S. D. Beach, L. Wang, and A. W. Sandvik. Data collapse in the critical region using finite-size scaling with subleading corrections, 2005. URL <http://arxiv.org/pdf/cond-mat/0505194v1>.
- [79] K. Binder. Finite size scaling analysis of ising model block distribution functions. *Zeitschrift für Physik B Condensed Matter*, 43:119, 1981. doi: 10.1007/BF01293604.
- [80] S. Humeniuk and T. Roscilde. Quantum Monte Carlo calculation of entanglement Rényi entropies for generic quantum systems. *Phys. Rev. B*, 86:235116, 2012. doi: 10.1103/PhysRevB.86.235116.
- [81] P. Calabrese and J. Cardy. Entanglement entropy and quantum field theory. *J. Stat. Mech.: Theory and Experiment*, 2004:P06002, 2004. doi: 10.1088/1742-5468/2004/06/P06002.
- [82] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Information*. Cambridge Univ. Press, Cambridge, 10th anniversary edition, 2010.
- [83] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac. Area Laws in Quantum Systems: Mutual Information and Correlations. *Phys. Rev. Lett.*, 100:070502, 2008. doi: 10.1103/PhysRevLett.100.070502.
- [84] M. B. Hastings, I. González, A. B. Kallin, and R. G. Melko. Measuring Renyi Entanglement Entropy in Quantum Monte Carlo Simulations. *Phys. Rev. Lett.*, 104:157201, 2010. doi: 10.1103/PhysRevLett.104.157201.
- [85] J. Iaconis, S. Inglis, A. B. Kallin, and R. G. Melko. Detecting classical phase transitions with Renyi mutual information. *Phys. Rev. B*, 87:195134, 2013. doi: 10.1103/PhysRevB.87.195134.

-
- [86] R. R. P. Singh, M. B. Hastings, A. B. Kallin, and R. G. Melko. Finite-Temperature Critical Behavior of Mutual Information. *Phys. Rev. Lett.*, 106:135701, 2011. doi: 10.1103/PhysRevLett.106.135701.
- [87] S. Inglis and R. G. Melko. Wang-Landau method for calculating Rényi entropies in finite-temperature quantum Monte Carlo simulations. *Phys. Rev. E*, 87:013306, 2013. doi: 10.1103/PhysRevE.87.013306.
- [88] 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*, 69:014509, 2004. doi: 10.1103/PhysRevB.69.014509.
- [89] A. E. Ferdinand and M. E. Fisher. Bounded and Inhomogeneous Ising Models. I. Specific-Heat Anomaly of a Finite Lattice. *Phys. Rev.*, 185:832, 1969. doi: 10.1103/PhysRev.185.832.

Acknowledgements

First of all, I gratefully thank Priv. Doz. Dr. Fabian Heidrich-Meisner for great mentoring and assistance; You kept a good balance between giving the direction of work and nevertheless leaving me the freedom to choose the topics of my interest. I thank you for investing much of your time in many and regular discussions with me.

Beside Fabian I thank Marius Fischer and Lukas Weidinger for a careful proofreading of my thesis. The constructive criticism of all you three helped me to improve both the English and the work itself.

Further, I thank Dr. Stephen Inglis for assistance with the implementation of directed loop updates and the replica trick in SSE, Dr. Vincenzo Alba for fruitful discussions on the projection operators, as well as Prof. Dr. Lode Pollet for general discussions and supervision.

I acknowledge the chair of Prof. Dr. Ulrich Schollwoeck for providing an office and access to computational clusters and for the possibility to visit a fall school on computational numerical methods in Würzburg.

Beside the academic support, I thank all my friends (in- and outside the university) and my family: on the one hand for covering my back when I had work to be done, and on the other hand for keeping me from work from time to time if I neglected you and my life too much. I hope you know how grateful I am to knowing you, even if I do not list all your names.

Statement of Authorship

Herewith, I certify that this thesis has been composed by myself and describes my own work unless otherwise acknowledged in the text. I created all figures on my own unless explicitly stated otherwise. Parts that are direct quotes or paraphrases are identified as such.

Munich, October 31, 2014

(Johannes Hauschild)