VALENCE BOND CALCULATIONS FOR QUANTUM SPIN CHAINS

VALENCE BOND CALCULATIONS FOR QUANTUM SPIN CHAINS: FROM IMPURITY ENTANGLEMENT AND INCOMMENSURATE BEHAVIOUR TO QUANTUM MONTE CARLO

By

ANDREAS DESCHNER, Dipl.-Phys.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY, Physics and Astronomy, McMaster University (2014), Hamilton, Ontario

TITLE: Valence Bond Calculations for Quantum Spin Chains: From Impurity Entanglement and Incommensurate Behaviour to Quantum Monte Carlo

AUTHOR: Andreas Deschner, Dipl.-Phys. (Heidelberg University)

SUPERVISOR: Dr. Erik S. Sørensen

NUMBER OF PAGES: xi, 120

Abstract

In this thesis I present three publications about the use of valence bonds to gain information about quantum spin systems. Valence bonds are an essential ingredient of low energy states present in many compounds.

The first part of this thesis is dedicated to two studies of the antiferromagnetic J_1-J_2 chain with $S = 1/2$. We show how automated variational calculations based on valence bond states can be performed close to the Majumdar-Ghosh point (MG-point). At this point, the groundstate is a product state of dimers (valence bonds between nearest neighbours). In the dimerized region surrounding the MG-point, we find such variational computations to be reliable.

The first publication is about the entanglement properties of an impurity attached to the chain. We show how to use the variational method to calculate the negativity, an entanglement measure between the impurity and a distant part of the chain. We find that increasing the impurity coupling and a minute explicit dimerization, suppress the long-ranged entanglement present in the system for small impurity coupling at the MG-point.

The second publication is about a transition from commensurate to incommensurate behaviour and how its characteristics depend on the parity of the length of the chain. The variational technique is used in a parameter regime inaccessible to DMRG. We find that in odd chains, unlike in even chains, a very intricate and interesting pattern of level crossings can be observed.

The publication of the second part is about novel worm algorithms for a popular quantum Monte Carlo method called valence bond quantum Monte Carlo (VBQMC). The algorithms are based on the notion of a worm moving through a decision tree. VBQMC is entirely formulated in terms of valence bonds. In this thesis, I explain how the approach of VBQMC can be translated to the S_z-basis. The algorithms explained in the publication can be applied to this S_z-method.

Acknowledgements

It is done. My Ph.D. thesis is written. What is to do now, is to acknowledge those who have helped me during my years in Hamilton. Those who made it possible that I have grounds to claim that during my stay in Canada I have learned much and grown substantially, as a physicist and as a person.

The first person I want to thank is Erik S. Sørensen. He gave me the opportunity to come here, he encouraged me to trust my judgement and put it to use. He guided me throughout these years. He also payed me. Thank you Erik.

I also want to thank Catherine Kallin and Sung-Sik Lee, the other members of my committee. They have always been a pleasure to meet with to discuss my work. Sung-Sik Lee has often, and certainly not only in committee meetings, impressed me with his clarity of thought and his willingness to share his knowledge.

I am indebted to many people. More than I could be able to express my gratitude to. Also more than would be appropriate to express my gratitude to in this thesis. I nevertheless want to use this opportunity to acknowledge as well as I can.

Anyone who goes to a foreign country can only wish to be welcomed by a Joshua D. McGraw; by somebody as open, friendly and genuine. He has seen me at my strongest and my weakest. I could always rely on him.

This is also true for Jessica Thornton, who for many years was a source of warmth, trust, support and access to cats. Without her, this undertaking would have been much more difficult than it was already.

Having people in my life that I like and trust gave me strength. I have always had great friends around me. Some faces disappeared quickly, some I have had the opportunity to see age a tiny bit, some I have only known for a very short while. Thank you (in approximately alphabetical order) Allan, Anna, Annie, Alex, Andrew, Andrew, Brendan, Brit, Clare, Claire, Jane, Jackie, Jesse, Faiyaz, Laura, Martin, Mark, Mike, Paul, Phil, Prasanna, Patrick, Rafael, Ray, Rory, Rob, Sedigh, Spencer, Wen and all those who I have forgotten to mention here. You all have taught me something profound. You all have made me feel at home in Hamilton.

A home that I will now leave to go back to where I came from. To go where Diana Fichtner is, who believed in me when I needed somebody to believe in me while I was writing this thesis and who is giving me something to look forward to after my departure. To go where Joachim Hahn is, my dearest friend who has been there for me during the last 21 years. To go where my family is, who has always (and to my amazement) put up with my shortcomings. Every time I am among them, I know who is most important to me, regardless of how often I see them or speak with them.

For my parents this is not merely true. It is due to them and due to my mother, who I know will always do and has always done everything she could to support her children.

Thank you both, for giving me the opportunity to live and feel loved, to rejoice and be amazed.

List of Figures

List of Tables

Preface

In this sandwich thesis, I present the major results of my Ph.D.-research. The first chapter is an introduction that shall serve the purpose of setting the stage for the later chapters which contain the publications themselves as well as their own short introductions.

- **Publication I:**

Andreas Deschner and Erik S. Sørensen:
Impurity Entanglement in the J-J_2-δ Quantum Spin Chain;
J. Stat. Mech. (2011) P10023; doi:10.1088/1742-5468/2011/10/P10023;

Calculations: I performed all variational calculations. DMRG calculations were performed by Erik S. Sørensen.

Manuscript: I wrote the bulk of the manuscript and made all figures. The introduction was written in equal parts by Erik S. Sørensen and me. Furthermore, Erik S. Sørensen provided (partly substantial) edits, comments and supervision.

- **Publication II:**

Andreas Deschner and Erik S. Sørensen:
Incommensurability effects in odd length J_1-J_2 quantum spin chains: On-site magnetization and entanglement;
Physical Review B **8**7, 094415 (2013); doi: 10.1103/PhysRevB.87.094415

Calculations: I performed all variational calculations. DMRG calculations were performed by Erik S. Sørensen.

Manuscript: I wrote the bulk of the manuscript and made all figures.
Erik S. Sørensen wrote section IV. He also added several paragraphs throughout the manuscript. Furthermore, Erik S. Sørensen provided (partly substantial) edits, comments and supervision.

- **Publication III:**

Andreas Deschner and Erik S. Sørensen:
Valence Bond Quantum Monte Carlo Algorithms defined on Trees;
submitted to Physical Review E; pre-print: arXiv:1401.7592

Calculations: I performed all calculations.

Manuscript: I wrote the bulk of the manuscript and made all figures.
Erik S. Sørensen provided (partly substantial) edits, comments and supervision.

Contents

1. Introduction

Things in nature behave according to many complicated rules. This can make researching their behaviour very difficult indeed. What makes the work of a physicist possible is that the physical world has the wonderful property that most of the many degrees of freedom are not important at all. If you were to try to kick a ball, you would mostly be concerned with its center of mass, its weight and its radius. The specifics of the positions of the atoms and their constituents are of no importance at all. For the purpose of kicking a ball, they are only distractions. There are, of course, many problems that make it necessary to try to know more about the ball than its macroscopic properties. One could, for example, try to find out what the ball is made out of, how it burns if lit or where in the ball there are electrons. It depends on the question at hand, on the problem one wants to solve. If it is possible to focus on a limited set of variables, one should always do it; one should always use a model. Fewer variables usually mean less hassle. A good model captures the relevant physics while making the problem much simpler to solve. One should, however, not overdo it. Modelling a soccer ball as a massless hard sphere when trying to estimate its trajectory after kicking it will lead to incorrect results.

In condensed matter physics we are interested in the study of a huge number of atoms. With knowledge of the constituent atoms, it is not difficult to write down a model that contains all physics that is relevant at energy scales we usually are interested in. It is a Hamiltonian with the kinetic energies of the electrons and nuclei together with Coulomb interactions between all particles that make up the compound.

Macroscopic systems contain on the order of 10^{24} particles. Full analytic calculations without any approximations are therefore impossible. Even just trying to keep track of the wavefunction of all the particles is hopeless. With so many particles, it is also futile to attempt to tackle such a problem numerically. If we had a very big and very good module of random-access memory that we wanted to fill with one number per particle and that has a transfer rate of 1 TB/s (a rate that is outside of current capabilities), we would have to wait for roughly the age of the universe for it to be filled. Something we cannot do.

We thus have to remove everything that is not essential to the problem. Take a crystal for example. At low temperatures, the nuclei can be thought of as harmonic oscillators, each oscillating about a fixed position in a lattice. When one is only interested in the degrees of freedom of the electrons, it is often found to be worthwhile to ignore the motion of the nuclei altogether and take them into account with a background potential. One must, however, not get carried away and think that the motion of the nuclei does not matter for the electrons. It is the interaction between the motion of the nuclei and the electrons that is responsible for superconductivity in conventional superconductors.

1.1. Magnetic models

If one is considering a crystal in which the electrons are well localized at magnetic ions, it is often a good idea to only consider the degrees of freedom of the spins in order to capture the low-energy physics. Then, suitable models can be devised. Rich magnetic structures can be studied in this way.

How the electrons interact through their spins is highly interesting. The classical dipolar interaction, which one always has between magnetic dipoles, is not the main cause for the interactions. The Coulomb interaction is mostly responsible for the strength of the spin interaction in magnetic compounds, even though it is blind to the spin of particles. Only through the Pauli principle and the tendency of fermions to avoid each other does the spin become important for the Coulomb interaction. The strong interactions between spins that give rise to magnetic phenomena are thusly generated. They are called exchange interactions.

Exchange interactions come in two variants: Coulomb exchange and kinetic exchange. Coulomb exchange usually favours alignment of spins. It is caused by minimizing the potential energy of electrons on different sites. An intuitive picture is easily given and understood. Because of the Pauli principle, electrons have a tendency to stay far apart from each other if their spins are aligned. The potential energy in this case is thus lower than if their spins were not aligned.

The alignment of the spins also plays an important role for the kinetic energy. If the spins of two electrons are not aligned, it is possible for them to be in the same orbital. Thus, the electrons can roam around the crystal less restrictedly. This leads to a lower kinetic energy and the energy contribution stemming from this mechanism is fittingly called “kinetic exchange energy”. It usually favours misalignment of spins. A more rigorous and quantitative exposition of the these concepts can be found in Auerbach [1994] and many other textbooks.

Because of the inherently quantum mechanical origin of the interactions that lead to macroscopic magnetism, it is fair to say that the existence of magnetic materials is a distinctly quantum mechanical phenomenon.

The main model for magnetic insulators is the Heisenberg model. It contains only the interactions caused by spin degrees of freedom. Spin Hamiltonians are defined on lattices. At each point in the lattice there is a spin. The most basic Hamiltonian with spin rotational symmetry for such a model is given by

$$H = \sum_{ij} J_{ij}\, \mathbf{S}_i \cdot \mathbf{S}_j \ , \tag{1.1}$$

where the J_{ij} stand for the interaction strengths between pairs of spins. The Heisenberg model is studied in classical and quantum mechanical varieties. In classical models, the $\mathbf{S}_i$ are the components of the spins at lattice sites. In quantum mechanical systems, they are the spin operators.

Heisenberg models are used to describe phenomena in ferromagnets, in which the spins have the tendency to align, and in antiferromagnets, where they tend to antialign. It depends

on the sign of the interaction strength J_{ij} which case is modelled. If the Js are negative, one is looking at a ferromagnetic interaction. For a Hamiltonian with only such interactions, the system has a ferromagnetic groundstate with all the spins aligned in one direction. Because this state is an eigenstate of the Hamiltonian, it is a groundstate for the classical as well as the quantum mechanical problem.

If the Js are positive, one speaks of an antiferromagnetic Hamiltonian. Compared to the ferromagnetic Hamiltonian, the spectrum is inverted and the behaviour is more difficult to understand. If it is possible to stagger the spins in the system and thereby satisfy all bonds, i.e. if there is no frustration, the groundstate of the classical system shows a staggered magnetic structure. This state, the Néel state, is *not* an eigenstate of the Hamiltonian. Finding the groundstate of an antiferromagnet even in the simplest case is thus not easy.

With frustration, classical systems may have many degenerate groundstates. For quantum systems this means that there are many different possible groundstates, some of which are very quantum mechanical. The most prominent such groundstate is the spin liquid state, in which no symmetries of the Hamiltonian are broken. I will say more about such states in the next section. Frustrated models are much harder to understand. After decades of research they are still the topic of a very active area. To get a flavour of the breadth of current research, I recommend having a look at Lacroix *et al.* [2011].

For real systems, many different terms can appear in a spin Hamiltonian. It might, for example, happen that not all Js are positive or negative and that different interactions compete because of the sign. This can also introduce frustration. The interactions between the spins might also not be isotropic or only involve two spins. Interactions with magnetic fields may also play a role.

The antiferromagnetic Heisenberg model can be derived as a limiting case of the Hubbard model. The Hubbard model describes electrons that are in localized orbitals close to lattice sites and can hop from one site to the other (see Hubbard [1963]). The Hamiltonian is given by

$$H = -t \sum_{(i,j),\sigma} \left[c^{\dagger}_{i,\sigma} c_{j,\sigma} + h.c. \right] + U \sum n_{i,\uparrow} n_{i,\downarrow} \,, \tag{1.2}$$

where the c_i are annihilation operators at site i and the first sum is over sites in the lattice as well as the spin directions (as indicated by the σ). The Hamiltonian has two competing terms. The hopping term causes the electrons to delocalize. Its strength is given by t. The on-site repulsion term causes the electrons to stay far away from each other and thus limits how the electrons can delocalize. Its strength is given by U.

The ratio t/U determines the characteristics of the model. The smaller t/U is, the stronger is the tendency of any two electrons to be at different sites. At half-filling, where there are as many electrons as sites, one can freeze the positions of the electrons by making t/U smaller and smaller. This increases the penalty for an uneven distribution of electrons. The hopping term, however, is still important. By means of perturbation theory, one can see that it lowers the energy of some states on the basis of their spin arrangement. The lowest order effective

Hamiltonian that is obtained with perturbation theory is the antiferromagnetic Heisenberg Hamiltonian as introduced above. Terms of higher order include interactions between more than two spins (see e.g. Takahashi [1977]).

1.2. Different orders

As long as the dimension is big enough, classical systems order at a finite temperature. For one- or two-dimensional systems, this may not be true according to the theorem due to Mermin and Wagner [1966], which states that continuous symmetries cannot spontaneously be broken in systems with dimension smaller than three. Fluctuations play a much more important role in low dimensions and can mix states that are related by the symmetry that would otherwise be broken.

Groundstates of classical systems, however, always show some order. Orders of classical spin systems are usually very intuitive to understand. In the easiest cases, there may be ferromagnetic order, where all spins align, or antiferromagnetic order, where all spins anti-align. An illustration of antiferromagnetic order can be found in figure 1.1 (a). With competing interactions or interesting lattices, there may also be incommensurate spiral order, where neighbouring spins are neither parallel nor point into opposite directions, but are rotated with respect to each other by some fixed pitch angle.

Quantum mechanical systems can have (and often do have) groundstates that are fundamentally different. Especially in low dimensions and for spins with small spin quantum number S can such groundstates be observed. This is so for the same reason the theorem of Mermin and Wagner is true: fluctuations about an ordered classical groundstate are enhanced under these conditions. In classical systems, fluctuations matter only at finite temperature. In quantum systems they are also important at zero temperature. In many regards, quantum systems of dimension D can be thought of as classical systems of dimension $D+1$ at finite temperature. The role of thermal fluctuations is then played by quantum fluctuations. The correspondence between classical and quantum spin systems is very useful since it allows our classical intuition to be used on quantum mechanical systems. Care has to be taken to not rely on this too much, as there is much more to most quantum mechanical systems.

For the Heisenberg chain this correspondence can for example be exploited in the case of integer spin (see Haldane [1983]). For Heisenberg chains with half-integer spins this is not true. The spin spin correlation function of the $S=1$-Heisenberg chain (also called the Haldane chain) decays exponentially, as one would expect because of the correspondence and the theorem of Mermin and Wagner. In contrast, the spin spin correlation function of the $S=1/2$ Heisenberg chain decays algebraically.

The classical Néel-ordered state with opposing spins on neighbouring sites is *not* an eigenstate of the antiferromagnetic Heisenberg Hamiltonian. Antiferromagnetic quantum systems with dimensions bigger than one nevertheless often show Néel-order similar to classical antiferromagnets such that $\langle \mathbf{S}_i \rangle \neq 0$ and the rotational symmetry of the Hamiltonian is spontaneously broken.

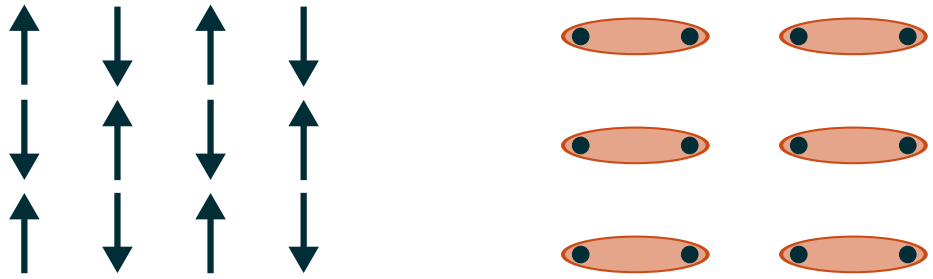
Many systems, however, show order which is very different to that of classical systems. Quantum mechanical systems have a means to minimize their energy that is not available to classical systems. In the presence of antiferromagnetic interactions, classical systems can only point their spins in as different directions as possible. The energy expectation value of such a state would be (with $J = \hbar = 1$)

$$\langle \uparrow_i \downarrow_j | \mathbf{S}_i \cdot \mathbf{S}_j | \uparrow_i \downarrow_j \rangle = -\frac{1}{4} \,. \tag{1.3}$$

Quantum spins can form a valence bond or dimer, given by the state

$$\frac{1}{\sqrt{2}} \left[| \uparrow_i \downarrow_j \rangle - | \downarrow_i \uparrow_j \rangle \right] \,. \tag{1.4}$$

This configuration is the groundstate for two spins with an antiferromagnetic interaction. The corresponding energy is $-3/4$. This is much lower than the energy for the antiparallel arrangement. There is a trade off to be taken into account: When two spins are in different dimers, the expectation value for the interaction term between them is zero. Because they dimerize, the two spins miss out on the opportunity to lower the energy through the interactions with other spins. The more neighbours a spins has, the less inclined it, thus, will be to form a dimer. Explicit dimerization (some of the bonds being stronger than others) aside, dimers or valence bonds therefore mostly play an important role in systems with one and two dimensions. An interesting feature of dimers is that they are spin singlets, i.e. that they are symmetric under spin-rotations. The spin rotational symmetry of the Hamiltonian is thus not broken in a valence bond state.



(a) Néel-order (b) Columnar dimer-order

Figure 1.1.: Illustrations of Néel-order (a) and dimer-order (b) on a square lattice of 12 sites. Which sites are connected by a dimer is indicated by an ellipse.

Dimerization can come in different shapes and forms. There is for example the valence bond crystal. The valence bonds are, in such a state, arranged in a repeating pattern. In this way, the state breaks symmetries of the underlying lattice. An example for such a state is illustrated in figure 1.1 (b).

With valence bonds as building blocks, one can also form states that do not break *any*

symmetry of the Hamiltonian. To get such a state, one has to add states with very many different arrangements of valence bonds (enough to have all states connected by symmetry operations covered). States with this property have been given the name resonating valence bond states (RVB). They are (and have been) of incredible interest. They were studied by Anderson [1973] and the interest in them rekindled when they were proposed to be important for the understanding of high temperature superconductivity (see Anderson [1987]). Because they, like a liquid, do not break any symmetries, such states are often called spin liquids.

Two reasonably realistic two-dimensional Heisenberg Hamiltonians have to date convincingly been shown to allow for such a state; for both this occurred recently. Firstly, the antiferromagnetic J_1-J_2 model on the square lattice, where nearest and next-nearest neighbours interact antiferromagnetically (see Jiang *et al.* [2012a]). The other system is the Kagome lattice with antiferromagnetic nearest neighbour interactions (see Yan *et al.* [2011]). This model is of relevance to the physics of a number of compounds (see Balents [2010]). With herbertsmithite there now also seems to be a material that is very well described by the Kagome Hamiltonian and, thus, considered to have a spin liquid groundstate (see Han *et al.* [2012]).

Some spin liquid groundstates are not entirely unorganized, even though no symmetries of the Hamiltonian are broken. These groundstates show topological order, as described by Wen [1989, 2013]. When such spin liquids are put on topologically non-trivial manifolds, there is a degeneracy. How severe this degeneracy is, depends on the genus of the manifold. One very unique quality is that this degeneracy cannot be eliminated by a local perturbation. A prominent example of such a model is the toric code introduced in Kitaev [2003]. The spin liquid state on the Kagome lattice is also believed to have topological order, as shown in Depenbrock *et al.* [2012]. Such states have long-ranged entanglement as was shown in Kitaev and Preskill [2006] and Levin and Wen [2006]. I will return to this issue in section 2.2.

1.3. The J_1-J_2 chain

In one dimension, one does not have to search for models that show inherently quantum mechanical order. The opposite is true. Even the simplest antiferromagnetic system with only nearest neighbour interaction has a spin liquid groundstate and adding next-nearest neighbour interactions induces dimerization. Two of the publications in this thesis are focused on exactly this spin model. In this section, I describe its phase diagram in some detail.

This spin chain is usually referred to as the J_1-J_2 chain. It is often also called the zigzag chain. Sadly, the also descriptive name “railroad trestle lattice” given to it in Anderson [1973] has somewhat gone out of style. The model was first considered by Majumdar and

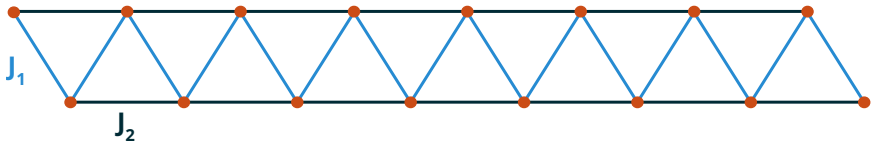


Figure 1.2.: A graphical representation of the J_1-J_2 chain.

Ghosh [1969a,b]. The Hamiltonian is

$$H = J_1 \sum_{(i,j)} \mathbf{S}_i \cdot \mathbf{S}_j + J_2 \sum_{((i,j))} \mathbf{S}_i \cdot \mathbf{S}_j \ , \tag{1.5}$$

where the subscripts (i, j) and $((i, j))$ denote pairs of nearest and next-nearest neighbours, respectively. A graphical representation is given in figure 1.2. There are many compounds that can be modelled with the J_1-J_2 chain. A (far from complete list) of such compounds can be found in Hase *et al.* [2004].

Depending on the interactions, the system may be frustrated. This makes for an interesting phase diagram, which is illustrated in figure 1.3. In the following an even number of sites is assumed (just as shown in figure 1.2).

I will go through the phase diagram in clockwise direction. I start at the point where both interactions are positive and $J_2/J_1 = J_c = 0.241167(5)$. For $J_2/J_1 < J_c$, the chain is in an antiferromagnetic phase (see White and Affleck [1996]; Eggert [1996]). It is labelled SAFM in figure 1.3. The systems is, in this phase, either slightly frustrated or not frustrated at all. In this phase, in which lies the Heisenberg chain with only nearest-neighbour interactions, one finds a gapless groundstate with short-ranged antiferromagnetic correlations. The magnetization at single sites is zero and there is no broken symmetry. This phase is, thus, an example of a spin liquid phase. The spin spin correlation function shows algebraic decay. In the lower right quadrant, the system is not frustrated. The phase just described is expected to be found in the whole lower right quadrant.

In the lower left quadrant, both interactions are negative. The interactions do not compete; J_2 in fact stabilizes the ferromagnetic polarization observed in the ferromagnetic Heisenberg chain without next-nearest-neighbour interaction. The ferromagnetic region extends into the upper left quadrant of slight frustration, where $J_2 > 0$. It is labelled with FM in figure 1.2.

At $|J_2/J_1| = J_{\text{UDRVB}} = 1/4$ there is a phase transition. The ferromagnetic groundstates become degenerate with an interesting, very unique state that is analytically known. It is a uniformly distributed resonating valence bond state (UDRVBS); a state given by the superposition of all product states with valence bonds between all pairs of sites (see Hamada *et al.* [1988]). For $|J_2/J_1| > 1/4$ the system has an incommensurate singlet groundstate which smoothly transforms into the UDRVBS as $|J_2/J_1| \searrow 1/4$ (see Tonegawa and Harada [1989]).

The maximum of the static magnetic structure factor starts to shift from $q = 0$ at J_{UDRVB}. It thus is a Lifshitz point. More information about such points can be found in the sections 3.1 and 3.2 of this thesis. The groundstate in this phase is believed to be dimerized.

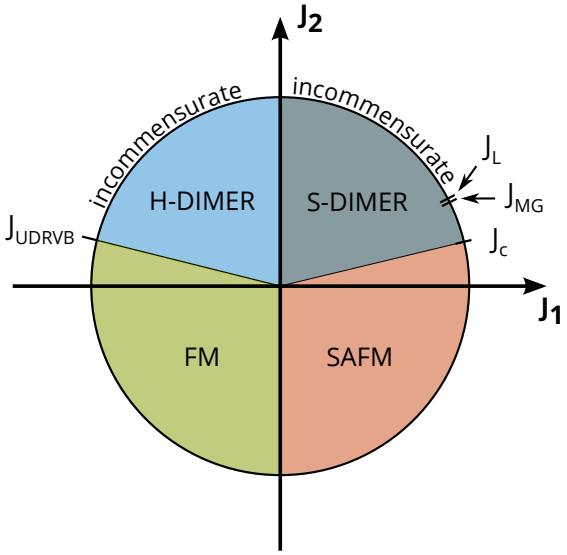


Figure 1.3.: The phase diagram of the J_1-J_2 chain as explained in the main text.

Appreciable values for the dimer-order parameter $\langle \mathbf{S}_{i-1} \cdot \mathbf{S}_i \rangle - \langle \mathbf{S}_i \cdot \mathbf{S}_{i+1} \rangle$ could, however, only be found for $|J_2/J_1| > 0.5$ (see Furukawa *et al.* [2012]). There likely is a very small gap in this region (see Itoi and Qin [2001]; Furukawa *et al.* [2012]). As is explained in Furukawa *et al.* [2012], one can likely think of the groundstate in this phase along the following lines. The ferromagnetic nearest neighbour interaction leads to effective degrees of freedom with S=1 that are located on each second bond. The constituents of the emergent degrees of freedom then form valence bonds. This is reminiscent of the situation in the Haldane chain (the S=1-Heisenberg chain). The phase is therefore called the Haldane dimer phase and the region is labelled with H-DIMER in figure 1.2. This phase has recently been of interest because of its relevance to ferroelectric compounds with chains of CuO_2 (see Furukawa *et al.* [2010]). Examples are $PbCuSO_4(OH)_2$ (see Willenberg *et al.* [2012]; Yasui *et al.* [2011]), $LiCu_2O_2$ (see Park *et al.* [2007]) and $LiCuVO_4$ (see Naito *et al.* [2007]).

When $J_1 = 0$ and $J_2 > 0$, the system is composed of two uncoupled antiferromagnetic Heisenberg chains. Thus, the model is gapless and there is antiferromagnetic short-ranged order with algebraic decay of the spin spin correlation function.

The J_1-J_2 chain is the subject of two papers presented in this thesis. Both studies concern a parameter range that lies in the upper right quadrant of figure 1.3. The phase is labelled with S-DIMER. For $J_2 > J_c = 0.241167(5)\, J_1$ the system is dimerized (see White and Affleck [1996]; Eggert [1996]). The spin spin correlations on alternating bonds have alternating sign in contrast to the Haldane dimer phase and the dimer-order parameter is about 40 times bigger than the dimer-order parameter in the Haldane dimer phase (see Furukawa *et al.* [2012]). There is a gap to the first exited state. The most well known compound whose low-energy properties can be modelled with the J_1-J_2 chain in this phase is probably

$CuGeO_3$ (see Castilla *et al.* [1995]).

The point at which $J_2 = J_{MG} = 0.5J_1$ is very special. It is called the Majumdar-Ghosh point (MG-point). Here, the groundstate is explicitly known. It is entirely composed of dimers, valence bonds between nearest neighbours. Since the arrangement of dimers is fixed and shows translational symmetry, this state is also called a dimer crystal. Under periodic boundary conditions this state is doubly degenerate. With open boundary conditions, there is only one such state. It is not too difficult to see that these states have to be groundstates.

Already in the first publications on the J_1-J_2 chain (Majumdar and Ghosh [1969b]), it was realized that the dimerized states at J_{MG} are exacts groundstates for a chain with 10 sites. I will, here, show that they are groundstates for any length of the chain. This will be done following the method used in van den Broek [1980]. The strategy of the proof is to first show that a lower limit for the groundstate energy exists and to then observe that the dimerized states saturate this limit.

We start by writing the Hamiltonian as a sum of terms that contain only interactions between 3 sites. Since $J_2 = \frac{1}{2} J_1$,

$$H = J_1 \sum H_i = \frac{1}{2} J_1 \sum_{i=1}^{N} \left(\mathbf{S}_i \cdot \mathbf{S}_{i+1} + \mathbf{S}_{i+1} \cdot \mathbf{S}_{i+2} + \mathbf{S}_i \cdot \mathbf{S}_{i+2} \right) , \tag{1.6}$$

where N is the number of sites in the chain. As is shown in the appendix A.1, the spectrum of single $H(i)$ is given by 3/4 and $-3/4$. The lowest eigenvalue of a sum of operators cannot be lower that the sum of the lowest eigenvalues of the single operators. Thus,

$$E_0 \geq J_1 \frac{N}{2} \left(-\frac{3}{4} \right) . \tag{1.7}$$

With a little algebra it is easy to see that a state composed of dimers has an energy of

$$E_{MG} = J_1 \frac{N}{2} \left(-\frac{3}{4} \right) \tag{1.8}$$

because every dimer (of which there are $N/2$) contributes an energy of $-J_1 3/4$. The dimerized states saturate the lower bound; they have to be groundstates.

The lowest lying excitations at the MG-point can be generated by breaking a dimer. The two unpaired spins can propagate freely (see Shastry and Sutherland [1981]; Sørensen *et al.* [1998]) and separate regions that have the dimerization pattern of the groundstate, which is why they are often called solitons. The gap and the simple structure of the lowest excited states makes variational calculations close to the MG-point very worthwhile.

At the MG-point, spins are only correlated with their next neighbour. The correlation length assumes its minimum at this point. It is thus also called a disorder point (see section 3.3 for more information on such points). Another special property of this point is that the spin spin correlation function starts to show incommensurate behaviour. This behaviour is present for all $J_2 > J_{MG}$. Only starting at a different point, the Lifshitz point where $J_2 =$

$J_L = 0.52063(6)\, J_1$ (see Bursill *et al.* [1995]), does the incommensurate behaviour become apparent in the static structure factor, the Fourier transform of the spin spin correlation function. More information about the Lifshitz point can be found in section 3.2.

1.4. Numerical approaches

It is very difficult to gain insight into the properties of spin systems analytically. Only for very few models is the groundstate known analytically. The J_1-J_2 chain at the Majumdar-Ghosh point (see last section) is an example. Hamiltonians analogous to the Majumdar-Ghosh chain can readily be constructed (see Klein [1982]). The idea behind this method is to, on each bond, project onto the state with a high total spin. Then, the low spin state minimizes the energy. The Hamiltonian of the $S = 1$ AKLT-chain, named after (Affleck, Kennedy, Lieb and Tasaki), falls into this class of Hamiltonians and thus has a known groundstate (see Affleck *et al.* [1987]).

One of the few systems that can be solved analytically is the nearest-neighbour Heisenberg chain. This can be done with the Bethe ansatz, a technique that was developed by Hans Bethe to solve exactly this problem (see Bethe [1931]). It has since been extended to other one-dimensional systems but all attempts to extend it to higher dimensions have failed.

Analytical approaches to spin systems almost always rely heavily on approximations and simplifications that are often hard to justify. It is, thus, not surprising that numerical methods play a central role in this branch of physics.

The main numerical techniques can be organised in three groups. Exact diagonalization, less exact diagonalization and quantum Monte Carlo.

For small systems it is possible to calculate whatever one desires just by diagonalizing the Hamiltonian in some basis. This is only possible for a small number of sites. Even if one removes redundant parts of the Hilbert space which are connected by operations under which the Hamiltonian is symmetric, this method can even in the best cases hardly be used to investigate groundstates for systems with more than about 45 sites (see Sandvik [2010] for a pedagogical introduction).

A popular way to alleviate this problem is to further restrict the Hilbert space in which the diagonalization is performed. Calculations that are performed in such a restricted Hilbert space are called variational. The state that has the smallest energy expectation value in this subspace is then taken to be the best approximation of the groundstate. How well this works (and if it works at all) depends heavily on the subspace that is used. Sometimes it is possible to just guess a good subspace. The papers in part I of this thesis rely on such an educated guess.

Often this is, however, not possible. Then, one needs to use a more inspired approach. The density matrix renormalization group (DMRG) relies on such an approach. Since its introduction in White [1992] this method has enjoyed widespread adoption especially for systems in one dimension. For a modern introduction to DMRG see Schollwöck [2011].

Systems with two dimensions are treated as one dimensional systems with irregular

interactions by consecutively labeling the sites in the lattice. That DMRG can be used for two-dimensional systems is attested to by its impressive use in studies that found spin liquid states in simple Heisenberg Hamiltonians and were mentioned in section 1.2 (Yan *et al.* [2011]; Jiang *et al.* [2012a]). The Hilbert space that the DMRG diagonalizes in can be written in terms of matrix product states (see again Schollwöck [2011]). It is now known that these states have different scaling of the entanglement entropy than the groundstates of many systems of interest. I will come back to this point in section 2.2. DMRG thus struggles to find such groundstates. The subspace it chooses is not ideal. Tensor network states allow for the implementation of variational calculations in terms of states with the correct scaling of the entanglement entropy (see Cirac and Verstraete [2009]).

Whenever possible, Monte Carlo methods are applied in spin lattice systems. Monte Carlo methods scale well with the sizes of the systems in question and allow for a robust estimate of uncertainties. They were first used during the Manhattan project and have developed into very popular tools to calculate statistical expectation values and many other things. One aspect all Monte Carlo methods have in common is that they make use of random numbers to approximately solve a problem.

The basic idea employed in calculations that are typical for condensed matter problems is easily explained in simple terms. If one wanted to calculate a statistical average

$$Q = \sum p_i Q_i \,, \tag{1.9}$$

but there are too many terms in the sum to actually perform the sum, there is still a way to approximately calculate it. One can generate an ensemble of the Q_i such that they are distributed according to the probabilities p_i and then just average over those Q_i. The central limit theorem tells us that the statistical uncertainty of this result will behave according to

$$\sigma \propto \frac{1}{\sqrt{N}} \,, \tag{1.10}$$

where N is the number of elements one uses to calculate the average. Many problems in statistical mechanics and quantum mechanics are exactly of the nature of equation 1.9.

Quantum Monte Carlo (QMC) is Monte Carlo of quantum systems. There are many different ways to employ Monte Carlo in the context of quantum mechanics. Many methods and algorithms have been proposed. I will restrict myself to the ones I consider the most important. In condensed matter systems, these mostly come in two variants.

The first variant is based on the partition function at finite T. Properties of the groundstates can be extracted by extrapolating to $T = 0$. The most well known and conceptually most intuitive method is called worldline Monte Carlo. The partition function is here written as a path integral via

$$Z = \mathrm{Tr} e^{-\beta H} = \mathrm{Tr}\left[e^{-\delta t\, H}\right]^L \,, \tag{1.11}$$

where $L\,\delta t = \beta$. The error one makes with this approximation can be controlled with δt. The lower the temperature is, the bigger L has to be chosen and the more expensive is the computation. Upon introducing many identities, this turns into a path integral with one compact dimension, the temperature- (or imaginary time-) dimension. Single contributions to the partition function that are sampled are the worldlines of the system. The temperature dimension plays the role of time. These worldlines form loops because in the temperature (time) direction periodic boundary conditions are employed. The worldlines here are analogous to the Q_i of equation 1.9. For a review of this method and different algorithms see Kawashima and Harada [2004].

Another very popular method for spin systems that is based on an expansion of the partition function is the stochastic series expansion (SSE) (see Sandvik and Kurkijärvi [1991] and for an introduction see Sandvik [2010]). Here, one Taylor expands the Boltzmann factor and stochastically evaluates a truncated version of the resulting sum. How early one can truncate depends crucially on the temperature because the expansion is basically a high-T expansion. The lower the temperature, the more terms in the series have to be kept and the more computationally expensive do calculations become.

The second variant of Quantum Monte Carlo is focused on $T=0$. One projects out the groundstate from a trial state with non-zero overlap with the groundstate by repeatedly acting on it with an appropriate operator. Popular choices for such a projection operator are the inverted and shifted Hamiltonian, $-H+C$, or $\exp(-H\tau)$. The second part of this thesis is about a projection method using the former operator. In projection QMC, one directly calculates groundstate properties and no extrapolation to $T=0$ is necessary. The quality of the results depends on the quality of the projection.

One class of such methods, Greens function Monte Carlo (see Ceperley and Kalos [1986]; Trivedi and Ceperley [1990]), performs the projection by evolving from one state to the next. One takes a trial state and acts on it with the projection operator. Of the many terms that are generated one only keeps a finite, manageable number. This is done in a stochastic way. Terms with high weight are more likely to be kept. By repeating this process one can then sample the groundstate. A pedagogical introduction can be found in Sorella *et al.* [2013].

It is also possible to do projective Monte Carlo in a very different way. If one expands operators such as $(-H+C)^N$, one is left with a sum of many different products of operators. These strings of operators can be the subject of sampling. This is usually done in a basis given by states in which all spins are in valence bonds. The method is therefore called "valence bond quantum Monte Carlo"(VBQMC). VBQMC was introduced in Liang [1990a], laid dormant for fifteen years, and was then further developed by Anders Sandvik (see Sandvik [2005]; Sandvik and Evertz [2010]). It is the subject of part II of this thesis and will there be explained in some detail. In section 4.1.2, I will explain that and how such projection QMC can also be done in different bases.

The generation of a suitable ensemble to average over is usually the hard part of Monte Carlo calculations. In very lucky cases it might be possible to avoid this issue and one might be able to just generate the ensemble from uniformly distributed random numbers via a direct mapping. An example of a method with this property and the associated algorithm

can be found in Gies *et al.* [2003]. If this is not possible (almost always), the generation of the ensembles is done using a Markov chain. One creates the next element of the ensemble from the current element such that it occurs with the probability

$$P = \min\left[1, \frac{p_{\text{next}}}{p_{\text{current}}}\right] , \tag{1.12}$$

where p_{next} and p_{current} are the probabilities with which one wishes to generate the next and the current state, respectively. If the next state is chosen according to equation 1.12, the algorithm is said to satisfy detailed balance and it produces a correct ensemble. This can most easily, most robustly, but sadly in many cases quite inefficiently be achieved by essentially changing the current element in a random way and then accepting the resulting element with the probability given in equation 1.12. If the random changes are too drastic, the acceptance rate is very low. Thus only minor changes can be performed in one step. This method is called "Metropolis-Hastings method" (see Metropolis *et al.* [1953]; Hastings [1970]).

The Metropolis-Hastings method works well in most cases. It is robust and applicable to many systems and problems. Yet, there are important cases where it is terribly inefficient. This is easy to see for the ferromagnetic 2D-Ising model close to its second order phase transition. Typical states of the system, here, contain large ordered domains. In a simulation one would now take such a state and flip a number of randomly chosen spins. The acceptance rate depends on the energy difference between the states before and after flipping the spins. The energy penalty for flipping spins would likely be big because one is likely to try to flip spins which lie in ordered domains. The probability of accepting moves is really small. Therefore, such domains are very difficult to revert and break up. In order to generate independent elements, one has to update very many times. Sampling the configuration space well takes a very long time if it is done in this way. This phenomenon is called critical slowing down. The degree of freedom that is changed (the direction of spin at randomly chosen sites) is not chosen well.

This problem was solved by the introduction of algorithms in which, instead of single spins, whole clusters of spins are flipped (see Swendsen and Wang [1987]; Wolff [1989]). By updating big parts of the system at once, updates are much more effectively changing the state.

In Evertz *et al.* [1993] an algorithm was developed for quantum mechanical systems that used the same strategy to alleviate some of the major problems of local updates, such as critical slowing down and non-ergodicity stemming from the conservation of winding numbers in the temperature as well as spatial directions. Their algorithm is applicable to worldline Monte Carlo. Here, just as in the algorithms for classical systems, big parts of the configurations that are sampled (the worldlines of the system) are chosen and then updated at once. Such algorithms are also known for other quantum Monte Carlo methods. Loop algorithms are, for example, known for VBQMC (see Sandvik and Evertz [2010]) and SSE.

There is another class of algorithm that do not suffer from the draw-backs of traditional

algorithms and can be used for worldline Monte Carlo. They are called worm algorithms and were first published in Prokof'ev and Svistunov [2001]. The idea here is to effect a large change in the configuration by doing many local updates of a special kind. One breaks up the configuration (we may think of closed worldlines) to create an open line instead of a loop. Then, the ends are moved around according to well chosen probabilities. The ends move around in configuration space like one might imagine little worms would. Thus the name worm algorithm. The configuration is only updated where the worm goes. Once the two ends meet again an update is complete. The local sampling is done in an extended space that includes configurations that are not present in the partition function. In this way it is possible to create algorithms that are competitive with non-local cluster updates. In state of the art quantum Monte Carlo, one usually uses either a cluster or a worm algorithm.

All QMC methods and algorithms have one issue in common. They cannot efficiently be used for frustrated systems. This is because, for frustrated system, all of them suffer from the sign problem – a truly hard problem (see Troyer and Wiese [2005]). It is caused by negative weights in the sum of equation 1.9. A probabilistic interpretation of the equation, then, becomes problematic and Monte Carlo approaches struggle.

Systems with more than 60,000 spins have been studied with quantum Monte Carlo methods (see Sandvik and Evertz [2010]). Further development is, however, still needed, as it may happen that data for very big system sizes is needed to obtain reliable results (as was e.g. a problem in Meng *et al.* [2010] and rectified in Sorella *et al.* [2012]). Often one might be interested in averages over a large number of simulations. Calculations in disordered systems, for example, make this necessary. Then, even minute improvements of algorithms can compound. Researchers will always push their tools as far as they can.

1.5. Outline

The remainder of this thesis is separated into two parts. In part I, two publications about variational calculations regarding the J_1-J_2 chain are presented. Both papers are preceded by an introduction of the main questions addressed in the paper.

I present a recent publication about novel algorithms for valence bond Monte Carlo in part II. It also contains an introduction in which the method itself and extensions of it are explained.

The thesis ends with a short conclusions in which the main findings of the three publications are summarized and remaining issues which could be the object of future research are highlighted.

Variational Calculations in the dimerized J_1-J_2 chain

In this part, two studies of the antiferromagnetic J_1-J_2 chain with $S = 1/2$ are presented. The two studies are linked by studying the same system with the same technique.

For both projects we heavily relied on an automated variational approach that we developed for these studies. In the dimerized phase of the J_1-J_2 chain such calculations have been shown to be very precise. In the past, these calculations could only be done at the MG-point. We can perform these calculations for any J_2/J_1. This made the whole dimerized phase accessible to the variational approach. How the calculations were done is explained in some detail in the publications themselves.

Otherwise, the two publications are quite distinct from each other because they are about very different properties of the J_1-J_2 chain. The first publication is about the entanglement properties of an impurity attached to the chain. The second publication is about a transition from commensurate to incommensurate behaviour that occurs in the J_1-J_2 chain and how the characteristics of this transition depend on the parity of the length of the chain.

2. Impurity entanglement in a dimerized quantum spin chain

Andreas Deschner and Erik S. Sørensen:
Impurity Entanglement in the J-J_2-δ Quantum Spin Chain;
J. Stat. Mech. (2011) P10023; doi:10.1088/1742-5468/2011/10/P10023;

Calculations: I performed all variational calculations. DMRG calculations were performed by Erik S. Sørensen.

Manuscript: I wrote the bulk of the manuscript and made all figures. The introduction was written in equal parts by Erik S. Sørensen and me. Furthermore, Erik S. Sørensen provided (partly substantial) edits, comments and supervision.

The first paper in this thesis is about entanglement. More specifically, it is about how an impurity site in a dimerized spin chain is entangled with the rest of the chain, how this should be studied and how a popular measure of entanglement can be calculated within a variational framework.

2.1. Entanglement

Entanglement is one of the aspects of quantum mechanics that cannot be grasped with classical intuition. There just is nothing in the classical world that is quite the same. The canonical example for entangled states are the maximally entangled states of two-level systems, also called qubits. In a system of two spin ½, these states are given by

$$|\mathrm{ME}_1\rangle = \frac{1}{\sqrt{2}}\Big[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\Big] \qquad |\mathrm{ME}_2\rangle = \frac{1}{\sqrt{2}}\Big[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\Big]$$
$$|\mathrm{ME}_3\rangle = \frac{1}{\sqrt{2}}\Big[|\downarrow\downarrow\rangle + |\uparrow\uparrow\rangle\Big] \qquad |\mathrm{ME}_4\rangle = \frac{1}{\sqrt{2}}\Big[|\downarrow\downarrow\rangle - |\uparrow\uparrow\rangle\Big]\,. \tag{2.1}$$

If one only had access to one of the spins in such a state, a measurement of the z-component would yield up or down with a probability of one half. After this measurement, the spin would be in the respective state. The other spin would then also be in a state with fixed z-component. For this to be the case, they could be arbitrarily far apart; no shielding would do anything, as long as they are in one of the maximally entangled states in the beginning. A measurement on one of the spins always measures both. This is why they are called maximally entangled.

At second glance this might not be too unusual. After all, if somebody had just prepared two classical spins such that one is pointing up and the other is pointing down and chosen the first spin to point up with a probability of one half, the situation would indeed look quite similar to us. The two situations are however fundamentally different. Whereas the outcome of the measurement of both spins is determined before the measurement in the classical scenario, this is not the case for the quantum spins. Here, the measurement on the first spin changes the state of the second spin. A signature of entanglement and something the classical world has no analogue for.

Two systems, A and B, are entangled if it is not possible to write their state as a single product state of local basis states, $|j\rangle_\mathsf{A}$ and $|i\rangle_\mathsf{B}$, i.e. if their state only be written as

$$\begin{aligned} |\mathsf{AB}\rangle &= \sum_{ij} c_{ij} |i\rangle_\mathsf{A} |j\rangle_\mathsf{B} \\ &= \sum_{n} \tilde{c}_n |n\rangle_\mathsf{A} |n\rangle_\mathsf{B} \,, \end{aligned} \tag{2.2}$$

as a sum of product states. Note that in the last line I have chosen orthonormal bases for A and B such that the sum is over one index. This way of expressing a state in the product space of two Hilbert spaces is known as the Schmidt decomposition. Using the Schmidt decomposition makes for a simple looking density matrix:

$$\rho = \sum_{n,m} \tilde{c}_n \tilde{c}_m \; |n\rangle_\mathsf{A} \langle m|_\mathsf{A} \otimes |n\rangle_\mathsf{B} \langle m|_\mathsf{B} \,. \tag{2.3}$$

There is a convenient way to quantify the entanglement between two systems that are in a pure state. Imagine that some quantity has been measured in system A but we do not know the outcome of the measurement. System B then is in several states with some probability, i.e. system B is in a mixed state. The initial state of A and B was more entangled if, after the measurement of system A, we know less about which of the possible states system B is in, because then measuring on system A changed the state of system B more severely.

How little we know about which state system B is in is encoded in its von Neumann entropy, which was first used to calculate the entropy for quantum mechanical systems by von Neumann [1927]. This entropy is a good measure of the entanglement between the two

systems. It is given by

$$\begin{aligned} S_{\mathsf{B}} &= -\langle \ln \rho_{\mathsf{B}} \rangle \\ &= -\operatorname{tr}\left[\rho_{\mathsf{B}} \ln \rho_{\mathsf{B}}\right] \ . \end{aligned} \tag{2.4}$$

The reduced density matrix for system B is given by

$$\rho_{\mathsf{B}} = \operatorname{tr}_{\mathsf{A}} \rho = \sum_n |\tilde{c}_n|^2 \ |n\rangle_{\mathsf{B}} \langle n|_{\mathsf{B}} \ . \tag{2.5}$$

This measure of entanglement is called the (von Neumann)entanglement entropy. We know that $\operatorname{tr}\rho = \sum_n |\tilde{c}_n|^2 = 1$. Thus, $\operatorname{tr}\rho_{\mathsf{B}}^2 = \sum_n |\tilde{c}_n|^4 < 1$ if there are more than one terms in the sum in equation 2.2, i.e. if B is in a mixed state.

The entanglement entropy is given by $S_{\mathsf{B}} = -\sum_n |\tilde{c}_n|^2 \ln|\tilde{c}_n|^2$. If one calculates this entropy for system A, one gets the same result.

It may be instructive to look at a simple example that will be important in the publication that is the subject of this chapter. For the maximally entangled singlet state of two spins (as defined in equation 2.1), we find, in the usual product basis, that

$$\rho = |\mathrm{ME}_1\rangle\langle \mathrm{ME}_1| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/2 & -1/2 & 0 \\ 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \rho_{\mathsf{B}} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \tag{2.6}$$

and thus $S_{\mathsf{B}} = 2 \times 1/2 \times \ln 2 = \ln 2$.

There is a whole family of entropies that are also used to quantify bipartite entanglement of pure states: the Rényi entropies defined by

$$S_{\mathsf{A}}^r(\rho) = \frac{1}{r-1} \ln\left[\operatorname{tr}\left(\rho_{\mathsf{A}}^r\right)\right] \ . \tag{2.7}$$

For $r \to 1$ this is just the von Neumann entropy.

2.2. Entanglement is useful

Aside from being a highly interesting non-classical phenomenon, entanglement is also highly useful.

Entanglement plays a central role, for example, in the fields of quantum computation and quantum information. Entanglement's role is in fact so central that in a textbook on the subject the authors claim that it is "iron to the classical world's bronze age" of computation and information (Nielsen and Chuang [2000]). Quantum computation and quantum information algorithms and protocols rely on entanglement.

Fairly recently condensed matter physicists have also started to study entanglement in their

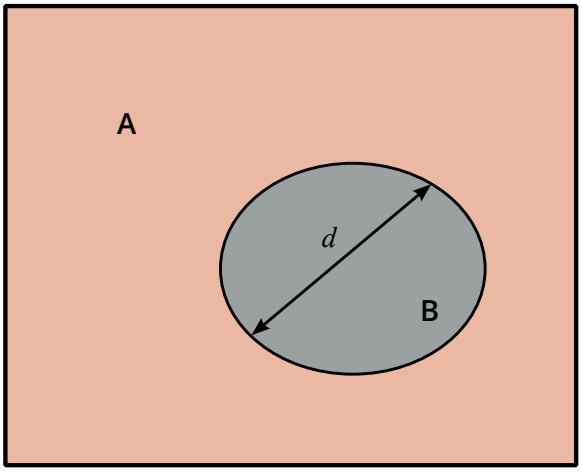


Figure 2.1.: A quantum mechanical system separated into two subsystems.

quest for the understanding of different phases in all kinds of quantum matter (for a review see Amico *et al.* [2008]).

One of the most important results of this research is the area law. It states that for groundstates of generic local Hamiltonians one can expect the entanglement entropy of a part of the system to depend on its size in a very specific way.

From thermodynamics we know the entropy to be extensive. It is proportional to the volume of a system. If one considers a subsystem of a bigger system in its groundstate, one, however, finds that the entanglement entropy of the subsystem does not scale with the volume as one may have guessed. For gapped systems, it scales with the length of the boundary surrounding the subsystem. For gapless systems, logarithmic corrections, which are sometimes universal to a whole class of systems, have to be applied.

It is not difficult to get an intuition for why the boundary might play an important role. The entanglement entropy is a measure of correlations between the subsystems A and B. In gapped systems, the correlations decay exponentially (Hastings [2004]). Thus, only close to the boundary can the two systems be entangled and the entanglement should be proportional to the size of the volume close to the boundary and thus be proportional to the area of the boundary.

For gapped 1D-systems, the area law has been proven to be valid (Hastings [2007]), i.e. it was shown that the entanglement entropy saturates to a finite value as the size of the subsystem is increased. For gapped systems with local Hamiltonian in higher dimensions it is generally expected to hold as well. A very interesting case are two-dimensional gapped systems with topological order. The area law is here given by $S = \alpha d - \gamma + \ldots$ (Kitaev and Preskill [2006]; Levin and Wen [2006]), where d is the linear dimension of the subsystem (as shown in figure 2.1) and γ, the *topological entanglement entropy*, is positive and depends solely on the type of topological order present. The topological entropy is the same for all Rényi entropies (see Flammia *et al.* [2009]). It is not easy to detect topological order since it

is an inherently non-local feature of the whole wavefunction. The topological entanglement entropy is therefore a very important tool which can and has been used to detect topological order (Isakov *et al.* [2011]; Jiang *et al.* [2012b]; Depenbrock *et al.* [2012]).

For generic gapless systems, the entanglement entropy is expected to scale like the length of the boundary plus logarithmic correction. One dimensional systems, here, also play a special role and can largely be considered understood.

By the use of conformal field theory, which is believed to describe gapless spin chains without internal length scale well, it was for example shown that (for periodic boundary conditions and an infinite system) the entanglement entropy scales like

$$S \sim c \ln(l/a) \,, \tag{2.8}$$

where a is a UV-cutoff length scale (for lattice models the lattice spacing) and l is the length of the subregion (Holzhey *et al.* [1994]; Calabrese and Cardy [2004]).

In higher dimensions the situation is more complex. Logarithmic corrections here also are necessary for many gapless systems (for a review see Eisert *et al.* [2010]).

The knowledge of the expected scaling of entanglement in groundstates is very useful if one should want to calculate a groundstate numerically. It is a hint that tells one where to search for the groundstate. This is especially important for variational techniques since they are in essence nothing but exact diagonalizations in a subspace of the whole Hilbert space. Their quality depends on how close the actual groundstate lies to this subspace. A method formulated in terms of states with the wrong scaling of the entanglement entropy will have grave difficulties to give the correct result. It is now understood that this is why DMRG (density matrix renormalization group) is difficult to apply to critical one-dimensional and to higher dimensional systems in general; it is a variational method in terms of matrix product states whose entanglement scaling is appropriate for gapped one-dimensional systems (see Schollwöck [2011] for a review). Now, tensor network states are used, because they have the desired scaling of the entanglement entropy. Efficient numerical methods based on them have been developed and implemented (see e.g. Cirac and Verstraete [2009]).

2.3. Measures of entanglement

To measure the entanglement between two systems which *together* are in a pure state is easy, as is described in section 2.1. One just has to calculate the entanglement entropy. Sadly, but obviously, this is not the only case in which one might want to be able to quantify entanglement. Just think of simple questions like, “How entangled are the spins on the sites 1 and 15 in a spin lattice?” or, “How entangled are the first 5 to the last 5 spins in a spin chain?”. A natural way to answer these questions is to first trace out the parts of the system one is not interested in and then calculate the entanglement between the remaining parts. After taking the trace, the remaining parts are in a *mixed* state. Thus, if we knew how to calculate entanglements for systems in mixed states, we could calculate the entanglement

between distant parts in a system. Sadly, no unique perfect measure of entanglement for systems in a mixed state is known. For different purposes there are different measures. In this section I will only give a very brief description of a few of them. For more information I refer the reader to Plenio and Virmani [2006].

Let $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$ be local bases for two systems A and B. A mixed state of A and B is *not* entangled (also called separable) if its density matrix can only be written as

$$\rho = \sum p_i |a_i b_i\rangle\langle a_i b_i| \,, \tag{2.9}$$

where $p_i > 0$. An example would be the state with the density matrix

$$\rho_{\text{sep}} = \frac{1}{2} |{\uparrow\downarrow}\rangle\langle{\uparrow\downarrow}| + \frac{1}{2} |{\downarrow\uparrow}\rangle\langle{\downarrow\uparrow}| \,, \tag{2.10}$$

of a system that is in the states $|{\uparrow\downarrow}\rangle$ and $|{\downarrow\uparrow}\rangle$ with equal probability.

One might be tempted to average the entanglements in the two contributing states according to the weights p_i to quantify the entanglement in the mixed state. This would not work. The reason is that density matrices can be separated in many different ways. Take for example the density matrix generated by mixing two maximally entangled states (as defined in equation 2.1)

$$\rho_{\text{m}} = \frac{1}{2} |\text{ME}_1\rangle\langle\text{ME}_1| + \frac{1}{2} |\text{ME}_2\rangle\langle\text{ME}_2| \,. \tag{2.11}$$

As calculated above in section 2.1, both parts by themselves have the entanglement entropy $S = \ln 2$. So one might also expect the state of equation 2.11 to have this entropy and be entangled. Writing ρ_{m} in terms of the product basis, we see that this result would be very wrong;

$$\begin{aligned}
\rho_{\text{m}} &= \frac{1}{2}\left[\frac{1}{2}\Big(|{\uparrow\downarrow}\rangle - |{\downarrow\uparrow}\rangle\Big)\Big(\langle{\uparrow\downarrow}| - \langle{\uparrow\downarrow}|\Big)\right] + \frac{1}{2}\left[\frac{1}{2}\Big(|{\uparrow\downarrow}\rangle + |{\downarrow\uparrow}\rangle\Big)\Big(\langle{\uparrow\downarrow}| + \langle{\uparrow\downarrow}|\Big)\right] \\
&= \quad \frac{1}{4}\Big[|{\uparrow\downarrow}\rangle\langle{\uparrow\downarrow}| - |{\uparrow\downarrow}\rangle\langle{\downarrow\uparrow}| - |{\downarrow\uparrow}\rangle\langle{\uparrow\downarrow}| + |{\downarrow\uparrow}\rangle\langle{\downarrow\uparrow}|\Big] \\
&\quad + \frac{1}{4}\Big[|{\uparrow\downarrow}\rangle\langle{\uparrow\downarrow}| + |{\uparrow\downarrow}\rangle\langle{\downarrow\uparrow}| + |{\downarrow\uparrow}\rangle\langle{\uparrow\downarrow}| + |{\downarrow\uparrow}\rangle\langle{\downarrow\uparrow}|\Big] \\
&= \quad \frac{1}{2}\Big[|{\uparrow\downarrow}\rangle\langle{\uparrow\downarrow}| + |{\downarrow\uparrow}\rangle\langle{\downarrow\uparrow}|\Big] \,,
\end{aligned} \tag{2.12}$$

which is equal to the density matrix of the separable state of equation 2.10. Mixing the two maximally entangled states resulted in state with no entanglement! Just looking at one possible way of writing the density matrix thus does not work.

There is, however, an entanglement measure in use that is based on the strategy I just dismissed: the entanglement of formation E_{f} (introduced in Bennett *et al.* [1996]). It is defined as the minimum value of entanglement among all decompositions of the density

matrix,

$$E_\mathrm{f} = \min\left[\sum p_i \tilde{S}_\mathsf{A}(\rho_i)\right] . \tag{2.13}$$

The entanglement entropy commonly used for this definition uses $\log_2$ instead of ln (which is usually true in publications in the quantum information community). The minimization is performed among all possible ways to decompose the density matrix of the system. In most realistic scenarios this makes E_f not practical. If the two systems under scrutiny are two qubits (like spins), however, there is a way to calculate E_f (see Wootters [1998]). It turns out that it can be written in terms of a quantity called "concurrence", denoted by C, via

$$E_\mathrm{f} = h\left(\frac{1+\sqrt{1-C^2}}{2}\right) , \quad \text{where} \quad h(x) = x\log_2(x) - (1-x)\log_2(1-x) . \tag{2.14}$$

Similarly to E_f, the concurrence for a density matrix $\rho = \sum p_i|\varphi_i\rangle\langle\varphi_i|$ of two qubits is given by a minimization over all possible decompositions of the density matrix (see e.g. Wootters [2001]),

$$C(\rho) = \min\left[\sum p_i|\langle\varphi|\sigma_y\otimes\sigma_y|\varphi^*\rangle|\right] . \tag{2.15}$$

This time the minimization is not nearly as detrimental as in the case of E_f. It can be shown that knowledge of the density matrix of the two qubits allows one to calculate the concurrence. It is given by

$$C(\rho) = \max[0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4] , \tag{2.16}$$

where the λ_i are the eigenvalues of $\rho\ (\sigma_y\otimes\sigma_y)\rho^*(\sigma_y\otimes\sigma_y)$ in descending order. Conveniently, the concurrence itself can serve as an entanglement measure. In this way pairwise entanglement of distant spins in a spin-½ system can be calculated.

A very easily computable entanglement measure for mixed states is the negativity introduced in Vidal and Werner [2002]. All eigenvalues of the partial transpose with respect to a subsystem of a density matrix of a separable state are positive (see Peres [1996]). In other words: If the partial transpose has negative eigenvalues, the state has to be entangled. The more negative its eigenvalues are, the more entanglement is present. This leads us straight to the definition of the entanglement measure negativity:

$$\mathcal{N}(\rho) = \frac{\|\rho^{\mathsf{T}_\mathsf{A}}\|_1 - 1}{2} , \tag{2.17}$$

where $\rho^{\mathsf{T}_\mathsf{A}}$ is the partial transpose of ρ with respect to subsystem A and $\|\cdot\|_1$ is the trace norm defined by $\|\mathsf{M}\|_1 = \mathrm{tr}\sqrt{\mathsf{M}^\dagger\mathsf{M}}$, or, with M's eigenvalues λ_i, $\|\mathsf{M}\|_1 = \sum|\lambda_i|$.

To see how a partial transposition acts on a density matrix, we look at a simple example of two spins in the usual product space with the basis ordered as $|{\uparrow_A}{\uparrow_B}\rangle$, $|{\uparrow_A}{\downarrow_B}\rangle$, $|{\downarrow_A}{\uparrow_B}\rangle$, $|{\downarrow_A}{\downarrow_B}\rangle$. The 4×4 Hermitian density matrix can be separated into four submatrices belonging to the four different combinations of the basis vectors in subsystem A,

$$\rho = \left(\begin{array}{c|c} \mathsf{C} & \mathsf{D}^\dagger \\ \hline \mathsf{D} & \mathsf{F} \end{array}\right) , \tag{2.18}$$

where C and F are Hermitian. The partial transpose with regards to system A is given by

$$\rho^{\mathsf{T}_\mathsf{A}} = \left(\begin{array}{c|c} \mathsf{C} & \mathsf{D} \\ \hline \mathsf{D}^\dagger & \mathsf{F} \end{array}\right) . \tag{2.19}$$

It is Hermitian and has the same trace as ρ. The trace of the partial transpose is, thus, equal to one. If some eigenvalues are negative, the sum of the absolute value of the eigenvalues has to be bigger than one. The negativity directly measures to what extend the partial transpose of a density matrix fails to be positive. Partial transposition with regards to subsystem B yields

$$\rho^{\mathsf{T}_\mathsf{B}} = \left(\begin{array}{c|c} \mathsf{C}^\mathsf{T} & \mathsf{D}^* \\ \hline \mathsf{D}^\mathsf{T} & \mathsf{F}^\mathsf{T} \end{array}\right) = (\rho^{\mathsf{T}_\mathsf{A}})^\mathsf{T} . \tag{2.20}$$

Thus, the negativity is the same regardless of with respect to which system the partial transpose is performed. These results are also true for systems that contain more than two spins.

It should be noted that a zero negativity is *not* a sufficient criterion for entanglement. There are entangled states with no negativity. The great strength of the negativity is the possibility of computing it.

2.4. The paper

In the publication presented in this chapter, we variationally calculate the negativity in the J_1-J_2 spin chain with explicit dimerization and an impurity attached to one end. This work is the natural extension of prior research. Erik S. Sørensen and collaborators published two papers about the entanglement of an impurity attached to the J_1-J_2 chain (see Sørensen *et al.* [2007a,b]).

For $J_2/J_1 < J_c \approx 0.241$, where the chain is gapless, this model shows the same physics as the low-energy regime of the Kondo model, where a spin interacts with itinerant electrons via a Kondo coupling and is screened by them (see e.g. Laflorencie *et al.* [2008]). Therefore, their work was mostly focused on this parameter regime.

They found that the entanglement entropy between the part of the chain that includes the impurity and the rest of the chain scales in a very particular way. This is so because the size

of the screening cloud surrounding the impurity provides a second length scale in addition to the length of the chain. To see this they defined the *impurity entanglement entropy* S_{imp}. It is the difference between the entanglement entropy with and without the impurity present in the chain. One of their main results was that $S_{\text{imp}} = S_{\text{imp}}(r/\xi_K, r/N)$, where r is the distance of the bipartition to the impurity, ξ_K is the size of the screening cloud derived from the decay of S_{imp} with r and N is the length of the chain.

They also studied the impurity entanglement entropy in the gapped dimerized regime with $J_2/J_1 > J_c$, for which they developed a very simple heuristic picture around the notions of "single particle entanglement" (SPE) and the "impurity valence bond" (IVB). This picture is based on the same ideas as the variational calculations with which many of the results in this thesis have been achieved.

The SPE is particularly important in the dimerized chain. With an odd number of spins, one spin cannot be in a dimer. It stays unpaired and acts as a domain wall between dimerized regions. It is thus usually called the soliton. The groundstate then roughly consist of a superposition of states that have this unpaired spin on different sites. The entanglement this induces is called "single particle entanglement".

The IVB can also easily be understood. If the impurity is coupled to the members of the chain just like all the other spins in the system (or equivalently if we have a chain without impurity), all spins will be in a dimer with their neighbour (aside from a soliton if applicable). Lowering the coupling on the impurity sites makes the valence bond from the impurity to its neighbour less favourable. It will thus terminate at a site farther away. This bond is called the impurity valence bond. The IVB contributes to the entanglement entropy, whenever it crosses the line dividing the two sectors between which the entropy is calculated.

It was also shown that the dimerized chain is, at the Majumdar-Ghosh point, very well treatable with a simple variational ansatz and that the entanglement entropy can be calculated in this way.

To get a better grasp of the entanglement of the impurity, it would of course be nice to be able to directly calculate how entangled the impurity is with a distant block in the chain. The negativity is a quantity that can exactly fulfill this need, as was outlined in section 2.3.

Bayat *et al*. were able to calculate the negativity for the J_1-J_2 spin chain with impurity using DMRG (see Bayat *et al.* [2010]). They showed that the negativity is indeed a very useful tool to understand how far into the chain the impurity is entangled. The gapless Kondo regime was of most interest to them. Some results were also presented for the dimerized phase. Based on DMRG data of the negativity, the same authors published a study in which they proposed the chain in the Kondo regime to be a good candidate to create long-ranged entanglement between the ends of the chain by quenching the interactions on one end (Sodano *et al.* [2010]).

We set out to answer the following questions:

- How do the impurity entanglement entropy and the negativity compare?
- How can variational calculations be used for a wider range of parameters?

- How can the negativity be calculated within the variational framework?
- What happens when one introduces an explicit dimerization?

The answers to these questions can be found in the first publication of this thesis.

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Impurity entanglement in the J–J_2–δ quantum spin chain

Andreas Deschner and Erik S Sørensen

Department of Physics and Astronomy, McMaster University,
1280 Main Street West, Hamilton, ON L8S 4M1, Canada
E-mail: deschna@mcmaster.ca and sorensen@mcmaster.ca

Received 15 July 2011
Accepted 23 September 2011
Published 24 October 2011

Online at stacks.iop.org/JSTAT/2011/P10023
doi:10.1088/1742-5468/2011/10/P10023

Abstract. The contribution to the entanglement of an impurity attached to one end of a J–J_2–δ quantum spin chain ($S = 1/2$) is studied. Two different measures of the impurity contribution to the entanglement have been proposed: the impurity entanglement entropy $S_{\rm imp}$ and the negativity $\mathcal{N}$. The first, $S_{\rm imp}$, is based on a subtractive procedure where the entanglement entropy in the absence of the impurity is subtracted from results with the impurity present. The other, $\mathcal{N}$, is the negativity of a part of the system separated from the impurity and the impurity itself. In this paper we compare the two measures and discuss their similarities and the differences between them. In the J–J_2–δ model it is possible to perform very precise variational calculations close to the Majumdar–Ghosh point ($J_2 = J/2$ and $\delta = 0$) where the system is gapped with a dimerized ground state. We describe in detail how such calculations are done and how they can be used to calculate $\mathcal{N}$ as well as $S_{\rm imp}$ for any impurity coupling $J_{\rm K}$. We then study the complete crossover in the impurity entanglement as $J_{\rm K}$ is varied between 0 and 1 close to the Majumdar–Ghosh point. In particular, we study the impurity entanglement when a staggered nearest neighbour interaction proportional to δ is introduced. In this case we observe a very rapid reduction in the impurity entanglement as δ is increased.

Keywords: dimers (theory), spin chains, ladders and planes (theory), entanglement in extended quantum systems (theory)

ArXiv ePrint: 1107.2907

Contents

1. Introduction

Entanglement in quantum spin chains has received considerable attention. For gapless quantum spin chains, detailed predictions [1] of many aspects of entanglement have been obtained from conformal field theory (CFT). For these (1 + 1)-dimensional models a precise understanding of the entanglement has proven invaluable for the understanding of numerical techniques such as the density matrix renormalization group [2] (DMRG) approach as well as for the development of new techniques. The contribution to the entanglement arising from impurities has also attracted considerable interest [3]–[13]. Here the term 'impurities' is used in the general sense and includes the effects of boundary magnetic fields [3], boundaries [4], qubits interacting with a decohering environment [5]–[8] as well as Kondo-like impurities [9]–[13]. The presence of the impurity can lead to different conformally invariant boundary conditions which can dramatically alter the entanglement. For a review see [14]. Quantum spin models have been used for studies of qubit teleportation and quantum state transfer [15]–[23], where the change in entanglement arising from the impurities plays a crucial role.

Usually the entanglement is defined in terms of the von Neumann entanglement entropy of a subsystem A of size l and reduced density matrix ρ_A defined by [24, 25]

$$S(l, L) \equiv -\operatorname{Tr}[\rho_A \ln \rho_A], \tag{1}$$

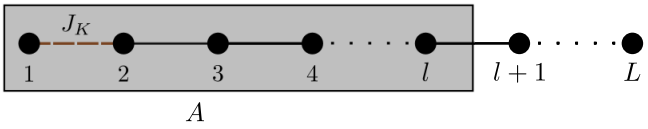


Figure 1. An example of a spin chain with only nearest neighbour interactions and an impurity on the first site. The interaction of the impurity site is multiplied by J_K. The chain is separated into part A (the grey box), which includes the impurity and has the length l, and the rest of the chain.

where L stands for the total system size. It is known [4] that the entanglement entropies of Heisenberg spin chains with open boundary conditions have a uniform as well as an alternating part:

$$S(l,L) = S_\mathrm{u}(l,L) + (-1)^l S_\mathrm{a}(l,L), \tag{2}$$

where the subscripts u and a stand for uniform and alternating respectively.

On the basis of this observation, the impurity entanglement entropy $S_\mathrm{imp}(J_\mathrm{K},l,L)$ was introduced to quantify the contribution that an impurity makes to the entanglement of a spin chain [9, 10]. To model an impurity, the interactions on one site are scaled by a factor of J_K. The impurity entanglement entropy is given by the difference of the uniform parts of the entanglement entropies of the chain with the impurity present and the chain without the impurity. The subsystem A in the definition of the entanglement entropy in equation (1) is chosen to contain the first l sites of the chain and thus contains the impurity site, too (see figure 1). Upon removing the impurity site, A and the whole chain are both one element shorter. The uniform entanglement entropy in a system where the impurity is absent is thus given by $S_\mathrm{u}(1,l-1,L-1)$. This leads to the following definition of S_imp:

$$\begin{aligned} S_\mathrm{imp}(J_\mathrm{K},l,L) &\equiv S_\mathrm{u}(\text{with impurity}) - S_\mathrm{u}(\text{no impurity}) \\ &\equiv\ S_\mathrm{u}(J_\mathrm{K},l,L) - S_\mathrm{u}(1,l-1,L-1), \qquad l>1. \end{aligned} \tag{3}$$

This definition is similar in spirit to experimental procedures for extracting impurity contributions to, for instance, susceptibilities. It is also possible to consider the alternating part of the impurity entanglement, but we shall not do that here.

However, it was later pointed out [11, 12] that a more consistent definition of the impurity contribution to the entanglement can be obtained from the negativity [26]. The negativity between a subsystem A and the rest of the system is defined by

$$\mathcal{N} = \frac{\sum|\lambda_i| - 1}{2}, \tag{4}$$

where the λ_i are the eigenvalues of the partial transpose of the density matrix with respect to either A or the rest of the system. In contrast to the entanglement entropy it can also be used to quantify the entanglement if the system is not in a pure state. This makes it possible to use the negativity to directly quantify the entanglement of the impurity spin and another part of the system.

It is an important question to what extent these two quantities yield the same information about the impurity entanglement. In this paper we compare these two

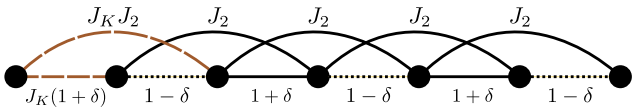


Figure 2. The J–J_2–δ chain with an impurity coupled with strength J_K.

measures and show that they agree for many universal features but are fundamentally different, with $\mathcal{N}$ being the more general measure of the impurity entanglement.

Entanglement and in particular impurity entanglement in gapped quantum spin chains has received comparatively little attention. In part this is due to the fact that for non-critical spin chains with a finite correlation length ξ, one expects [1] for a partition into two semi-infinite chains

$$S \sim \log \xi. \tag{5}$$

For instance, for the $S = 1$ Affleck–Kennedy–Lieb–Tasaki (AKLT) chain [27, 28] where the entanglement can be calculated analytically [29]–[35] it is known that $S(l, L)$ approaches a constant for large l, L in an exponential manner [34, 35] on a length scale equal to the bulk correlation length $\xi = 1/\ln(3)$. At the Majumdar–Ghosh (MG) point [36] of the J–J_2 spin chain, where $J_2 = J/2$, one finds that $S(l, L)$ is either 0 or $\ln(2)$ for any l, L. However, the impurity contribution to the entanglement can still be long range in gapped spin chains. For the J–J_2 spin chain this was shown to be the case at the MG point [9, 10]. This is due to the twofold degeneracy of the singlet ground state, corresponding to the two different dimerization patterns, and the associated presence of solitons separating regions with different dimerizations. The model that we study is the slightly more general J–J_2–δ quantum spin chain defined as

$$\begin{aligned} H = J_\mathrm{K}[(1+\delta)\vec{S}_0 \cdot \vec{S}_1 + J_2\, \vec{S}_0 \cdot \vec{S}_2] + \sum_{i=1}^{L-3}[(1+(-\delta)^i)\vec{S}_i \cdot \vec{S}_{i+1} + J_2 \vec{S}_i \cdot \vec{S}_{i+2}] \\ + (1+(-\delta)^{L-2})\vec{S}_{L-2} \cdot \vec{S}_{L-1}, \end{aligned} \tag{6}$$

where we implicitly have set the nearest neighbour coupling $J \equiv 1$. The coefficient J_K describes the coupling of the impurity spin and δ is a staggering of the nearest neighbour coupling inducing dimerization. Figure 2 may serve to clarify the definition of the Hamiltonian. We focus exclusively on chains with open boundary conditions and the impurity at one end of the chain. With $\delta = 0$ this model undergoes a transition from the gapless Heisenberg phase to a dimerized state at $J_2^\mathrm{c} = 0.241\,167$ [37, 38]. In the Heisenberg phase a correspondence to the low energy physics of a Kondo impurity can be established [39, 40]. This correspondence becomes exact at J_2^c. The mapping to the Kondo problem requires one to use open boundary conditions and we therefore do not consider periodic boundary conditions here. However, entanglement in models similar to the one considered here but with periodic boundary conditions has also been studied; see section 4 of [14] for a review.

With periodic boundary conditions and an even number of sites, the system at the MG point has an exactly known twofold-degenerate dimerized ground state and a gapped spectrum [36, 41]. (With open boundary conditions and L even the ground state is unique with the same energy.) The lowest lying excitations of the MG chain are pairs of unbound

solitons and upon introducing a finite dimerization the solitons become confined [42]. For a chain with an odd number of sites, solitons separate regions with the ground state dimerization pattern of the chain with an even number of sites [41, 42]. For the gapped phase, in the proximity of the MG point, it is possible to perform very precise variational calculations [41, 43, 44] for the J–J_2 model (corresponding to $\delta = 0$ in this study) since appropriate states spanning the variational subspace can be identified with relative ease, as the ground state at the MG point is known. Within this variational framework it is possible to evaluate the entanglement [10]. Here, we explain in detail how similar calculations can be performed for the negativity and discuss the appropriate subspaces used in these computations. Along the so called disorder line, where $J_2 = (1-\delta)/2$, one of the two degenerate ground states of the Majumdar–Ghosh chain remains the exact ground state. Close to the disorder line, one therefore expects the same variational methods to yield very precise results.

In the past, these variational calculations have been performed more or less by hand [41, 43, 44]. If the variational problem is, however, formulated as a generalized eigenvalue problem, it is not difficult to automate the calculation, as for any two states, φ_i and φ_j, with all but no, one or two sites being in dimers, $\langle\varphi_i|\vec{S}_i\vec{S}_j|\varphi_j\rangle$ as well as the overlaps between different φ can be computed in an automated fashion. One advantage of the automated computation of these matrix elements is that one avoids lengthy and error-prone calculations. Secondly, solving a generalized eigenvalue problem can be done using standard libraries. This approach is not limited to the specific Hamiltonian studied in this paper. It can in fact be used to perform variational calculations using dimerized states forming the subspace in which the diagonalization is carried out for any Heisenberg spin Hamiltonian.

In addition to providing a more direct analytical insight into the physics, the main advantages that the variational approach has over other methods such as the DMRG technique are computational cost and simplicity. These advantages do not, however, come for free. To identify a good subspace to use for variational calculations, one always has to rely on knowledge gained obtained by other means. This role is often played by the DMRG technique.

The paper is organized as follows. In sections 2–5 the variational approach is discussed along with the different subspaces used in the calculations. Section 6 describes the calculation of the negativity once the variational ground state is known and sample results for the negativity are compared to DMRG data. In section 7 the negativity and the impurity entanglement entropy are compared. Section 8 presents our results for the negativity and S_{imp} for general J_{K} at the MG point. In section 9 we describe how the impurity entanglement is affected by the presence of a non-zero explicit dimerization, δ.

2. The variational calculation as a generalized eigenvalue problem

For completeness, we review here how the variational problem can be formulated as a generalized eigenvalue problem. Consider a set of states $\{\varphi_i\}$ to be used for the variational calculation. We then wish to minimize

$$\langle H\rangle = \frac{(\varphi|H\varphi)}{(\varphi|\varphi)}, \tag{7}$$

where the trial state φ is given by

$$\varphi = \sum c_j \varphi_j. \tag{8}$$

The minimization is to be performed with respect to the c_j. In variational calculations we make use of the fact that

$$\forall \varphi : \; E_0 \leq \frac{(\varphi|H\varphi)}{(\varphi|\varphi)}, \tag{9}$$

where E_0 is the ground state energy. One strives to find a well chosen set of φ_is as well as the state $\varphi_{\min} \in \mathrm{span}\{\varphi_i\}$ that satisfies

$$\frac{(\varphi|H\varphi)}{(\varphi|\varphi)}\Big|_{\varphi_{\min}} = \min_{\mathrm{span}\{\varphi_i\}} \frac{(\varphi|H\varphi)}{(\varphi|\varphi)}. \tag{10}$$

This state is the variational estimate for the ground state. Let us denote the energy eigenbasis as $\{\mathfrak{e}_i\}$. The variational state φ can evidently be written as $\varphi = \sum a_i \mathfrak{e}_i$. Thus

$$\frac{(\varphi|H\varphi)}{(\varphi|\varphi)} = \sum_{i,j} \bar{a}_i a_j \frac{(\mathfrak{e}_i|H\mathfrak{e}_j)}{\sum_i |a_i|^2} = \frac{\sum_i |a_i|^2 E_i}{\sum_i |a_i|^2} \geq E_v, \tag{11}$$

where E_v is the energy of the eigenstate $\mathfrak{e}_v$ of non-zero projection onto $\mathrm{span}\{\varphi_i\}$ that has the lowest energy. With any orthonormal basis $\{\mathfrak{b}_i\}$ of $\mathrm{span}\{\varphi_i\}$, the variational guess for the ground state $\varphi_{\min}$ can thus be found by diagonalizing H:

- Minimization problem $\equiv$ Finding the lowest eigenvalue of $(\mathfrak{b}_i|H\mathfrak{b}_j)$.

If we do not have an orthonormal basis at our disposal, we have to do a little more work. Let us suppose that the φ_i are linearly independent and the $\mathfrak{e}_i$ are the orthonormal basis of energy eigenstates. For the variational ground state given by $\varphi_{\min} = \sum_i a_i \varphi_i$ and the variational estimate of the ground state energy E_v we know from the equation (11) that

$$H \sum_j a_j \varphi_j = E_v \sum_j a_j \varphi_j. \tag{12}$$

We will now show that a vector $\mathfrak{a} = (a_1, \ldots, a_n)$ is a solution to the variational problem if and only if

$$\forall i : \; \sum_j a_j (\varphi_i|H\varphi_j) = E_v \sum_j a_j (\varphi_i|\varphi_j). \tag{13}$$

With the definitions

$$\mathfrak{H}_{ij} = (\varphi_i|H\varphi_j) \qquad \text{and} \qquad \mathfrak{B}_{ij} = (\varphi_i|\varphi_j), \tag{14}$$

one can write equation (13) as

$$\mathfrak{H}\mathfrak{a} = E_v \mathfrak{B}\mathfrak{a}, \tag{15}$$

which defines a generalized eigenvalue problem. That equation (12) implies equation (13) can be seen by taking the scalar product of both sides in equation (12) with φ_i.

That equation (13) implies (12) will now be proved. We do this by proving that for any basis (linearly independent set of states that span the space) $\{\varphi_i\}$,

$$\forall i: \qquad (\varphi_i|\mathfrak{q}) = (\varphi_i|\mathfrak{r}) \Rightarrow \mathfrak{q} = \mathfrak{r}, \tag{16}$$

for any vectors $\mathfrak{q}$, $\mathfrak{r}$. With $\mathfrak{q} = H\sum_j a_j\varphi_j$ and $\mathfrak{r} = E_v\sum_j a_j\varphi_j$ the relation (16) becomes the relation equation (13) $\Rightarrow$ (12).

Proof:

$$\forall i: \qquad (\varphi_i|\mathfrak{q}) = (\varphi_i|\mathfrak{r}) \tag{17}$$

$$\Rightarrow$$

$$\forall i: \qquad (\varphi_i|\mathfrak{q} - \mathfrak{r}) = 0 \tag{18}$$

$$\Rightarrow$$

$$\forall i: \qquad \varphi_i \perp (\mathfrak{q} - \mathfrak{r}) \tag{19}$$

$$\Rightarrow$$

$$\mathfrak{q} - \mathfrak{r} = 0. \qquad \square \tag{20}$$

We have thus proven that the minimization problem of equation (10) can be replaced by the generalized eigenvalue problem of equation (15) if the states used to span the subspace in which the minimization takes place are linearly independent.

3. The Hilbert spaces used in the calculation

The implementation of the variational technique as a generalized eigenvalue problem allows us to just solve a linear algebra problem instead of solving a set of non-linear equations which is necessary if one approaches the minimization naively. The one requirement for this implementation to work is for the states that span the Hilbert space in which the minimization is done to be linearly independent. This flexibility makes it possible to use different bases according to the specific needs of the problem. Here we are concerned with the two cases of even and odd length chains.

(i) A chain with an odd number of sites.
The simplest ensemble of variational states for a chain with an odd number of sites are the so called single-soliton states [41, 43, 44, 10]. These states can be generated by taking $L-1$ of the sites of the chain to be maximally dimerized while leaving the site left over (the soliton) to be in the $S_z = 1/2$ state. Here and in the remainder of the paper we use L to denote the length of the chain. The number of dimers between the undimerized site and the left end of the chain can be used to label the states. In figure 3 we show a pictorial representation of the first three states. When drawing a dimer we use an arrow to denote the order in $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$. For an odd length chain with L spins there exist $N = (L+1)/2$ such states. Note that here we do not include states with the soliton in the $S_z = -1/2$ state.

(ii) A chain with an even number of sites.
In studies of the entanglement of the impurity site with the rest of a chain with an even number of sites, a suitable set of states can be defined by leaving the bulk

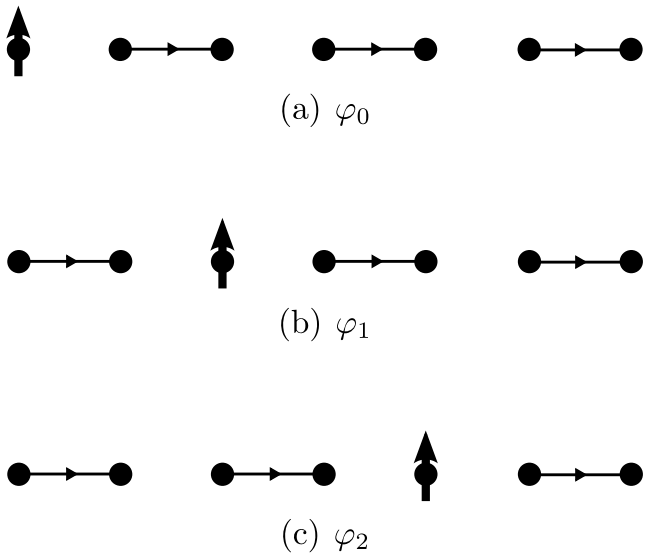


Figure 3. The first three variational states used in the calculations with the single-soliton space for a chain with an odd number of sites. Note the arrows on the dimers that indicate their direction.

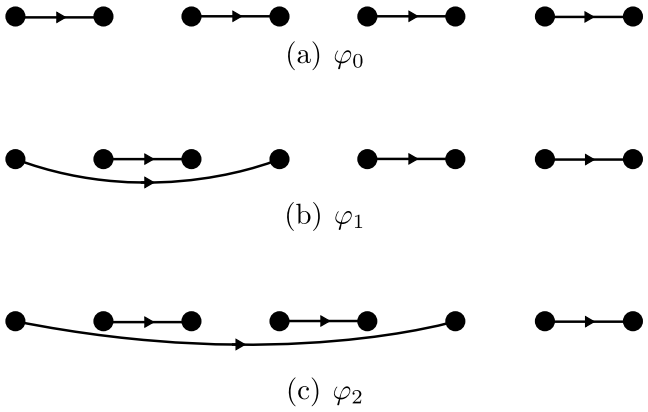


Figure 4. The first three variational states used in the calculation of the negativity for a chain with an even number of sites.

$(L-2)$ sites of the chain dimerized while one of the spins is in a valence bond with the impurity. Example states are given in figure 4. These states are chosen in order to reflect the presence of an impurity valence bond (IVB) [10], connecting the impurity site to the bulk of the chain.

4. The matrix elements of the Hamiltonian $\mathfrak{H}$

To set up the generalized eigenvalue problem it is necessary to calculate the matrices $\mathfrak{H}$ and $\mathfrak{B}$ which were defined in equation (14). In this section we will explain how to compute

$\mathfrak{H}$. In order to allow for the most general Hamiltonian possible, we write it as follows:

$$H = \sum_{\{m,l\}} J_{ml} \vec{S}_m \vec{S}_l. \tag{21}$$

Here, the summation $\{m, l\}$ is over 'bonds' of the Hamiltonian and the coupling constants J_{ml} can take any value. It is convenient to write the matrix elements $\mathfrak{H}_{ij}$ in terms of the operators $h_{ml} = \vec{S}_m \vec{S}_l - 1/4$, because they have a simple action on the φs. This gives

$$(\varphi_i|H\varphi_j) = \sum_{\{m,l\}} J_{ml}(\varphi_i|\vec{S}_m\vec{S}_l\varphi_j) = \sum_{\{m,l\}} J_{ml}(\varphi_i|[h_{ml} + 1/4]\varphi_j). \tag{22}$$

The following description of this action is an extension of results presented earlier [45] in the context of quantum Monte Carlo simulations in the valence bond basis where it is necessary to take the h_{ml} to have the sign opposite to what we use here. The action of h_{ml} on states relevant to us is given by

$$h_{ml}[m;l][k;n] = -[m;l][k;n] \tag{23}$$

$$h_{mk}[m;l][k;n] = \tfrac{1}{2}[m;k][n;l] \tag{24}$$

$$h_{ml} \uparrow_m [l;k] = \tfrac{1}{2}[m;l] \uparrow_k \tag{25}$$

$$h_{mk}[m;l] \uparrow_k = -\tfrac{1}{2}[\uparrow_m [l;k] + [m;l] \uparrow_k], \tag{26}$$

where

$$[m;k] := \frac{1}{\sqrt{2}}(\uparrow_m\downarrow_k - \uparrow_k\downarrow_m)$$

are the spins at m and k in a singlet state.

In figure 5 a pictorial representation of the equations (23)–(26) is shown. To fix the phase of the states that we include in the calculation, we again represent $[m;k]$ by an arrow pointing from m to k. For the calculations that we present here, these are the only rules needed. However, when considering the action of the Hamiltonian it is sometimes convenient to apply the simple rules

$$[m;k] = -[k;m] \tag{27}$$

$$[k;l] \uparrow_m = [k;m] \uparrow_l + \uparrow_k [m;l], \tag{28}$$

in order to reduce the action of the h_{ml} to one of the above.

We want to stress that even though the single-soliton states can be orthogonalized with relative ease [46], this is not generally the case and even after orthogonalization, the resulting eigenvalue problem is still non-trivial for a general set of J_{lm}. Whether or not it is possible to proceed with analytical calculations often depends on specific choices for the coupling constants in the Hamiltonian. Here, we can solve the variational problem for any values of the couplings.

As can be seen in equation (24) and figure 5(b), the application of h_{ml} to a given state φ_i may result in a state that is not an element of span$\{\varphi_i\}$. Therefore, in order to evaluate

$$(\varphi_i|H\varphi_j) = \sum_{\{m,l\}} J_{ml}(\varphi_i|h_{ml}\varphi_j)$$

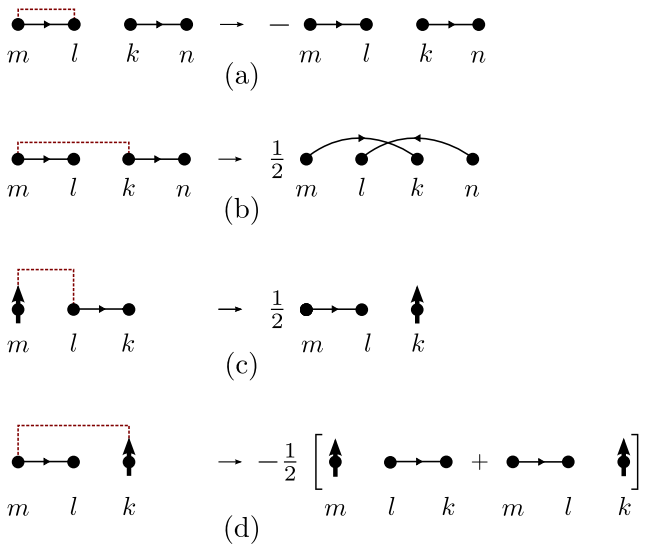


Figure 5. The action of h_{ml} on the variational states. The dotted line signifies the sites acted upon by h_{ml}. Arrows fix the phase of singlets between sites of the lattice. The convention is explained in the main text.

we need to compute the overlap of $h_{ml}\varphi_j$ with all the variational states. This slows down our calculations since it is not sufficient to build a table of all the overlaps of the φ_i. We now turn to a discussion of how the overlaps are calculated.

5. The matrix elements of the overlap matrix $\mathfrak{B}$

As is explained in [47, 45] the magnitude of the overlaps of valence bond states depends solely on the length of loops created by overlaying the two states. If one chooses l to be the length of such a loop then the overlap is of magnitude $2^{-l/2+1}$. To get the sign of the overlap we, in addition to overlaying them, reverse all arrows on one of the states. Then we follow the loop and count how often one has to go against the arrow. If the resulting number is odd, the sign of the overlap is negative. If the resulting number is even, the sign of the overlap is positive. Because the total number of bonds in the loop is even, this leads to a well defined sign of the overlap. To get the overlaps of two states one only has to follow this prescription for every loop and multiply contributions from different loops. Figure 6(c) shows the loop structure for the states shown in figures 6(a) and (b). The left loop is six bonds long and going around it one has to go against the direction of the arrow three times. The right loop is two bonds long and one has to go against the direction of the arrow once. The overlap is therefore given by

$$(-2^{-2})(-2^{0}) = \tfrac{1}{4}.$$

In the case of unpaired spins being part of the states, the rules have to be modified slightly. If there is one unpaired spin in both states, upon overlaying the two states, there will in addition to loops be a string that goes from one unpaired spin to the other. As is the case for the loops, for the string the magnitude of the contribution to the overlap depends only on the length. One finds the contribution to be $2^{-l/2}$. The contribution to the sign of the overlap can be determined in exactly the same way as in the case of complete loops.

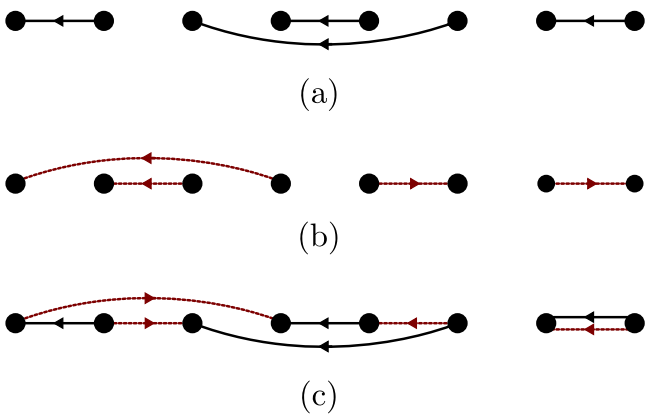


Figure 6. The two states (a) and (b) together with the corresponding loop structure created by reversing the arrows of one state and overlaying the two states (c). Both loops contribute a minus sign. The overlap is given by 1/4.

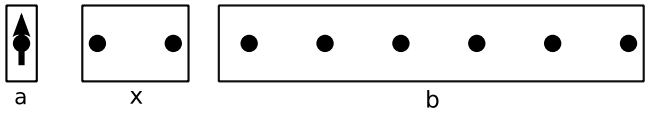


Figure 7. The chain separated into three parts: the impurity spin (a), the part whose entanglement with the impurity spin is to be quantified (b) and the part that separates the two (x).

6. The calculation of the reduced density matrix and the negativity

For the calculation of the negativity describing the entanglement between the impurity spin and a distant part of the chain, it is necessary to separate the chain into three parts [26, 11]: the impurity spin, the part whose entanglement with the impurity spin is to be quantified and the part that separates the two (the parts will in the remainder of the paper also be referred to as regions a, b and x respectively. The negativity is invariant under interchange of b and x. For clarification, see figure 7). To calculate the negativity between the impurity spin and a part of the chain that is separated from the impurity spin by a different part, we have to compute the reduced density matrix that results from tracing out the separating part of the chain. We will now explain how this is done within our variational framework.

Let the result of the generalized eigenvalue problem be the variational ground state φ_g given in the variational basis: $\varphi_g = \sum a_i \varphi_i$. To take trace of the density matrix over the region x in figure 7, we express the density matrix in a product basis formed by states of region x and states of the rest of the chain. This requires us to represent the variational ground state φ_g in this product basis. In doing so we follow the approach of [10], where the entanglement entropy was calculated from a variational ground state. For a chain with an odd number of spins these bases will, in the following, be introduced. We start with the case where the region x contains an even number of sites followed by the case where the region x contains an odd number of sites. For a chain with an even total number of sites an analogous calculation was performed.

6.1. An even number of sites to be traced out

(i) For the impurity spin we choose the S_z basis and denote the two states by $f_1 = \uparrow$ and $f_2 = \downarrow$.

(ii) For the region x we denote the basis by g_i. We take the region to contain at least two sites. The states are chosen to be

$$g_1 = \sum_{n=1}^{R} a_n \underbrace{\uparrow \overbrace{- \cdots -}^{n-1 \text{ dimers}} \uparrow - \cdots -}_{2R \text{ sites}}$$

$$g_2 = \sum_{n=1}^{R} a_n \underbrace{\uparrow \overbrace{- \cdots -}^{n-1 \text{ dimers}} \downarrow - \cdots -}_{2R \text{ sites}},$$

$$g_3 = \underbrace{\uparrow - \cdots - \uparrow}_{2R \text{ sites}}, \qquad g_4 = \underbrace{\downarrow - \cdots - \downarrow}_{2R \text{ sites}}, \qquad g_5 = \underbrace{\downarrow - \cdots - \uparrow}_{2R \text{ sites}},$$

$$g_6 = \underbrace{\uparrow - \cdots - \downarrow}_{2R \text{ sites}}, \qquad g_7 = \underbrace{- \cdots -}_{2R \text{ sites}},$$

where the symbol '$-$' is used for valence bonds between neighbouring sites and R is the highest possible number of dimers that can could be formed in region x.

(iii) For region b we define the states in the following way:

$$\tilde{h}_1 = \underbrace{- \cdots -}_{L-2R-1 \text{ sites}} \qquad \tilde{h}_2 = \sum_{n=R+1}^{N-1} a_n \underbrace{\uparrow \overbrace{- \cdots -}^{n-R-1 \text{ dimers}} \uparrow - \cdots -}_{L-2R-1 \text{ sites}}$$

$$\tilde{h}_3 = \sum_{n=R+1}^{N-1} a_n \underbrace{\downarrow \overbrace{- \cdots -}^{n-R-1 \text{ dimers}} \uparrow - \cdots -}_{L-2R-1 \text{ sites}}.$$

In order to express the density matrix of the impurity spin and the region b in an orthonormal basis it is necessary to orthonormalize the basis $\{\tilde{h}_i\}$ as it is not orthogonal in the way in which it is defined above:

$$(\tilde{h}_1|\tilde{h}_3) = \sum_{n=R+1}^{N-1} a_n \, (-2)^{-n+R+1} \left(\frac{-1}{\sqrt{2}}\right). \tag{29}$$

We do this following the usual Gram–Schmidt procedure. Our orthonormalized version is given by

$$h_1 = \frac{\tilde{h}_1}{\sqrt{(\tilde{h}_1|\tilde{h}_1)}} \qquad h_2 = \frac{\tilde{h}_2}{\sqrt{(\tilde{h}_2|\tilde{h}_2)}} \qquad h_3 = \frac{\tilde{h}_3 - (\tilde{h}_3|h_1)h_1}{\sqrt{(\tilde{h}_3|\tilde{h}_3) - |(\tilde{h}_3|h_1)|^2}}. \tag{30}$$

For the product space of the impurity spin and region b we use the basis states:

$$k_1 = f_1 h_1, \qquad k_2 = f_1 h_2, \qquad k_3 = f_1 h_3,$$
$$k_4 = f_2 h_1, \qquad k_5 = f_2 h_2, \qquad k_6 = f_2 h_3.$$

With these definitions, the variational ground state can be written as

$$\varphi_{\mathrm{g}} = \sum a_n \varphi_n = \sum_{i,j} C_{ij} k_i g_j, \tag{31}$$

where the matrix C is given by

$$C = \begin{pmatrix} 0 & \frac{\|\tilde{h}_1\|}{\sqrt{2}} & 0 & 0 & 0 & \frac{(\tilde{h}_3|h_1)}{2} & a_0\|\tilde{h}_1\| \\ 0 & 0 & 0 & -\frac{\|\tilde{h}_2\|}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\kappa}{2} & 0 & 0 \\ -\frac{\|\tilde{h}_1\|}{\sqrt{2}} & 0 & -\frac{(\tilde{h}_3|h_1)}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{\|\tilde{h}_2\|}{2} & 0 \\ 0 & 0 & -\frac{\kappa}{2} & 0 & 0 & 0 & 0 \end{pmatrix},$$

where $\kappa = \sqrt{\|\tilde{h}_3\|^2 - |(\tilde{h}_3|h_1)|^2}$ and $\|h\| = \sqrt{(h|h)}$.

With the matrix $G_{ij} := (g_i|g_j)$, whose matrix elements can be computed using the prescription given in section 5, one can write the reduced density matrix of the impurity spin and region b as

$$(\rho_{\mathrm{ab}})_{ij} = (\mathrm{tr}_{\mathrm{x}}\ \rho_{\mathrm{axb}})_{ij} = \sum_{n,m} C_{in} C_{jm} (g_n|g_m), \tag{32}$$

or, in matrix form,

$$\rho_{\mathrm{ab}} = CGC^{\mathrm{T}}. \tag{33}$$

6.2. An odd number of sites to be traced out

(i) For the impurity spin we choose the S_z basis and denote the two states by $f_1 = \uparrow$ and $f_2 = \downarrow$.

(ii) For the region to be traced out we denote the basis by g_i. The states are chosen to be

$$g_1 = \underbrace{- \ \cdots \ -\downarrow}_{2R+1 \text{ sites}}, \qquad g_2 = \underbrace{- \ \cdots \ -\downarrow}_{2R+1 \text{ sites}},$$

$$g_3 = \sum_{n=1}^{R} a_n \underbrace{\uparrow \overbrace{- \ \cdots \ -}^{n-1 \text{ dimers}} \uparrow - \ \cdots \ -\uparrow}_{2R+1 \text{ sites}},$$

$$g_4 = \sum_{n=1}^{R} a_n \underbrace{\downarrow \overbrace{- \ \cdots \ -}^{n-1 \text{ dimers}} \uparrow - \ \cdots \ -\uparrow}_{2R+1 \text{ sites}},$$

$$g_5 = \sum_{n=1}^{R} a_n \underbrace{\uparrow \overbrace{- \ \cdots \ -}^{n-1 \text{ dimers}} \uparrow - \ \cdots \ -\downarrow}_{2R+1 \text{ sites}},$$

$$g_6 = \sum_{n=1}^{R} a_n \underbrace{\downarrow \overbrace{- \ \cdots \ -}^{n-1 \text{ dimers}} \uparrow - \ \cdots \ -\downarrow}_{2R+1 \text{ sites}},$$

$$g_7 = \underbrace{\uparrow - \ \cdots \ -}_{2R+1 \text{ sites}}, \qquad g_8 = \underbrace{\downarrow - \ \cdots \ -}_{2R+1 \text{ sites}},$$

where the symbol '−' is used for valence bonds between neighbouring sites and R is the highest possible number of dimers that can be formed in region x.

(iii) For region b we define the states in the following way:

$$\tilde{h}_1 = \underbrace{\uparrow - \ \cdots \ -}_{L-2R-2 \text{ sites}}, \qquad \tilde{h}_2 = \underbrace{\downarrow - \ \cdots \ -}_{L-2R-2 \text{ sites}}, \qquad \tilde{h}_3 = \sum_{n=R+1}^{N-1} a_n \overbrace{\underbrace{- \ \cdots \ -}_{n-R-1 \text{ dimers}} \uparrow - \ \cdots \ -}^{L-2R-2 \text{ sites}}.$$

We orthonormalize and build a product basis for the impurity spin and region b just as is done in the case of an even number of sites that have to be traced out.

With $G_{ij} := (g_i|g_j)$ and

$$C = \begin{pmatrix} -\frac{a_0}{\sqrt{2}} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & \frac{(\tilde{h}_1|\tilde{h}_3)}{\sqrt{2}} \\ 0 & \frac{a_0}{\sqrt{2}} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\kappa}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & -\frac{(\tilde{h}_1|\tilde{h}_3)}{\sqrt{2}} & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{\kappa}{\sqrt{2}} & 0 \end{pmatrix},$$

where $\kappa = \sqrt{(\tilde{h}_3|\tilde{h}_3) - |(\tilde{h}_3|h_1)|^2}$, we can again write the reduced density matrix as

$$\rho_{\text{ab}} = CGC^{\text{T}}. \tag{34}$$

For a chain with an even total number of sites we followed an analogous procedure. In order to save space, we have not included it here.

6.3. The negativity

With the expression for the reduced density matrix of the impurity spin and region b, ρ_{ab}, at hand, it is now straightforward to calculate the negativity according to its definition [26]

$$\mathcal{N} = \frac{\sum |\lambda_i| - 1}{2}, \tag{35}$$

where the λ_i are the eigenvalues of the partial transpose of ρ_{ab}.

6.4. The precision of the variational approach

The variational states that form our starting point should yield almost exact results close to the MG point ($J_2 = J/2$) where a chain with an even number of spins is fully dimerized. However, the states that we include in our variational calculation (see section 3) form only a subset of the states in which the bulk of the chain is dimerized. At the MG point the contribution arising from these neglected states is negligible and the variational approach is very good. However, as J_2 is decreased towards the critical point J_2^{c} we expect the contribution from the neglected states to grow in importance and at J_2^{c}, where the ground state is no longer dimerized, we expect our variational approach to fail. Nevertheless, for intermediate J_2, $J_2^{\text{c}} < J_2 \leq J/2$, we still expect the variational approach to yield qualitatively good results.

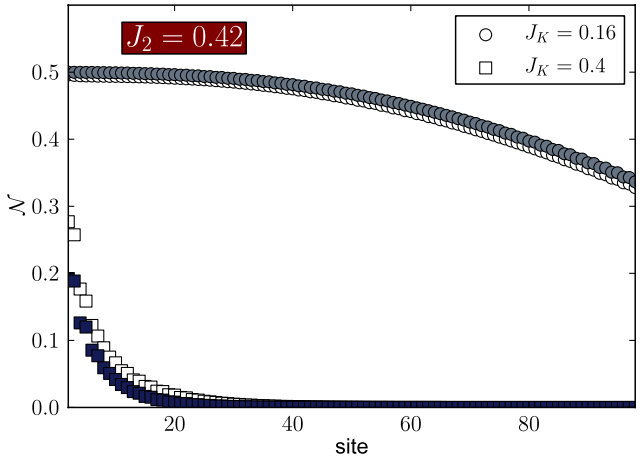


Figure 8. Variational results for the negativity (filled symbols) together with DMRG data from [11] (empty symbols) for two values of the impurity coupling J_K. The calculations were done for a chain of length $L = 250$, with frustration $J_2 = 0.42$ and dimerization $\delta = 0$.

It is possible to calculate the negativity, $\mathcal{N}$, using DMRG techniques, as was shown by Bayat *et al* [11]. In figure 8 we compare our variational results for $\mathcal{N}$ to DMRG results from [11] at $J_2 = 0.42J$ for two different values of J_K: $J_\mathrm{K} = 0.16$ and 0.4 (here $\delta = 0$). The definition of $\mathcal{N}$ used in [11] differs from the one used by us by a factor of 2. To account for this difference we multiplied their data by 2. For both values of J_K we observe a very good agreement between the DMRG and variational results.

7. $\mathcal{N}$ compared to S_imp at the MG point

Having established the validity of the variational approach we now turn to a comparison of the two measures for the impurity entanglement, $\mathcal{N}$ [11] and S_imp [9]. In figure 9 we show results for both $\mathcal{N}$ and S_imp at the MG point for a system with $L = 200$ sites and an impurity coupling $J_\mathrm{K} = 0$. Shown also is the impurity entanglement arising from a single impurity valence bond (IVB) [9] between the impurity spin and the bulk of the chain. For $J_\mathrm{K} = 0$ and L even, the idea of an impurity valence bond is easy to understand. The uncoupled impurity must form a singlet with the unpaired $S = 1/2$ in the bulk of the chain. If the entire system (impurity and bulk) is divided into two parts A and B, this impurity valence bond will contribute either $\ln(2)$ or 0 depending on whether the unpaired spin is in part B or A. With the wavefunction for the unpaired spin $\varphi_\mathrm{sol}(x)$ and $p = \int_A |\varphi_\mathrm{sol}(x)|^2 \,\mathrm{d}x$, the impurity entanglement is then simply [9, 10]

$$S_\mathrm{imp}^\mathrm{IVB} = (1 - p)\ln(2). \tag{36}$$

To a first approximation $\varphi_\mathrm{sol}(x)$ can be taken to be the wavefunction of a free particle in a box [10]:

$$\varphi_\mathrm{sol}(i) = \sqrt{\frac{2}{L}}\sin\left(\frac{\pi i}{L}\right). \tag{37}$$

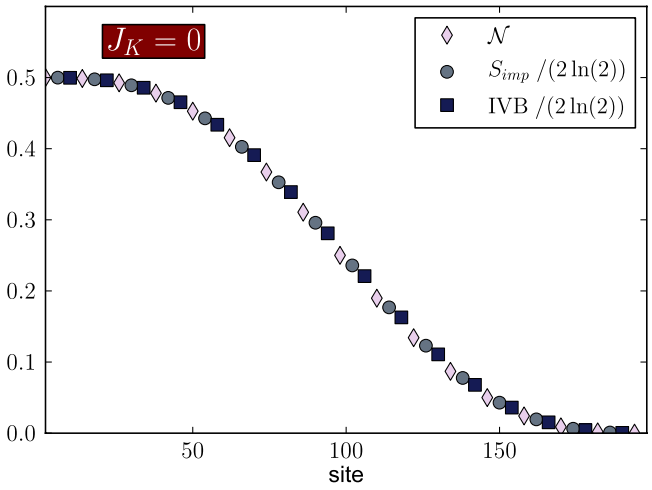


Figure 9. Variational results for the negativity $\mathcal{N}$ and the impurity entanglement S_{imp} scaled by $1/(2\ln(2))$ at impurity coupling $J_{\mathrm{K}} = 0$ and $\delta = 0$. Shown also is the impurity entanglement arising from an impurity valence bond (IVB) scaled by $1/(2\ln(2))$. The calculations were done for a chain of length $L = 200$ at the MG point. For clarity, only a fraction of all data points are shown.

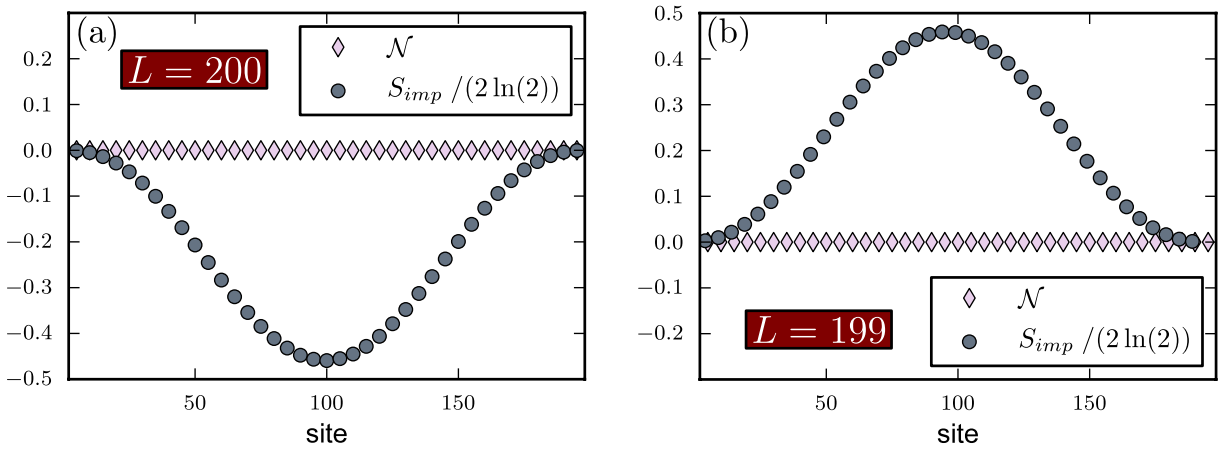


Figure 10. Variational results for the negativity $\mathcal{N}$ and the impurity entanglement S_{imp} scaled by $1/(2\ln(2))$ at impurity coupling $J_{\mathrm{K}} = 1$. The calculations were done at the MG point for a chain of length $L = 200$ (a) as well as $L = 199$ (b). For clarity only a fraction of all data points are shown.

It is also possible to obtain more precise estimates of φ_{sol} [10]. Apart from an overall scaling factor of $2\ln(2)$, the two measures are in the case shown in figure 9 essentially identical.

We now compare the two measures of the impurity entanglement in the limit $J_{\mathrm{K}} = 1$. In figure 10(a) we show variational results for both quantities at $J_{\mathrm{K}} = 1$. Since the impurity here is coupled with the same strength as the remaining sites and since we consider a chain with an even number of sites, the chain is in its dimerized ground state. The dimerization of the state implies that every spin is maximally entangled with one of

its neighbours. It follows that the impurity is only entangled with its neighbour and not with any other spin. While this can be inferred from the negativity shown in figure 10(a) which is exactly 0 for $l > 0$, this is not the case for $S_{\rm imp}$ which in fact becomes negative (!). The difference of the two measures arises from the subtractive procedure used to define $S_{\rm imp}$ where the term $S_U(1, l-1, L-1)$, in addition to terms cancelling out $S_U(1, l, L)$, also contains a contribution arising from the single unpaired spin which is present because $L-1$ is odd. This contribution from $S_U(1, l-1, L-1)$ can be identified as the single-particle entanglement [10] (SPE). Hence, $S_{\rm imp}$ fails in this case to be a good measure of the impurity entanglement since $S_U(1, l-1, L-1)$ is not a good reference state and cannot be identified with S (no impurity). It would be very interesting to define a measure of the impurity entanglement based on a relative entropy, but again one would need a well defined reference state with the impurity absent.

For a system with an odd number of sites the SPE is the only contribution to $S_{\rm imp}$ at $J_{\rm K} = 1$. Here the impurity is viewed in the general sense as arising from the boundary conditions yielding the non-zero $S_{\rm imp}$. However, the negativity, since it is really a tripartite measure, is not sensitive to this and is negligible for $l > 1$, as shown in figure 10(b), at the MG point.

It may be conjectured that $\mathcal{N}$ and $S_{\rm imp}$ agree well as long as the IVB contribution dominates $S_{\rm imp}$ as is the case in figure 9. We also note that the SPE vanishes at $J_2^{\rm c}$ [10] and we expect $\mathcal{N}$ and $S_{\rm imp}$ to exhibit the same scaling behaviour at $J_2^{\rm c}$, as has indeed been demonstrated [9, 11]. For $J_2 > J_2^{\rm c}$, $\mathcal{N}$ is the more faithful measure of the impurity entanglement and in the following we mainly use this measure.

8. The negativity and the SPE for general $J_{\rm K}$ at the MG point

As we saw in section 7, the impurity entanglement for an even length system with $J_{\rm K} = 0$ can be seen as arising from an impurity valence bond (IVB). At the MG point for $J_{\rm K} = 0$ this was calculated within the variational approach in [10]. Here we focus on the negativity $\mathcal{N}$ and the complete crossover as $J_{\rm K}$ is varied between 0 and 1. Variational results for $\mathcal{N}$ at different $J_{\rm K}$ for a system of length $L = 200$ (including the impurity site) are shown in figure 11(a). At $J_{\rm K} = 0$ the impurity entanglement is long range, extending throughout the chain, but as $J_{\rm K}$ is turned on, it quickly becomes suppressed and for $J_{\rm K}$ larger than $\sim$0.3, it has all but disappeared, reflecting the fact that the impurity valence bond now preferentially terminates close to the impurity site. For these non-zero values of $J_{\rm K}$ the appropriate $\varphi_{\rm sol}$ to use would then no longer be the wavefunction of a free particle, as in equation (37), but instead a wavefunction describing a localized state bound to the impurity site. The presence of such a localized state should be reflected in an exponentially decaying impurity entanglement away from the impurity site. For $J_{\rm K} > 0.15$ we have verified that the negativity $\mathcal{N}$ shown in figure 11(a) does indeed have a tail decaying exponentially with the site index. The length scale, $\xi_{\rm loc}$, associated with this exponential decay is therefore clearly distinct from the bulk spin–spin correlation length in the system which here is effectively zero. The sharpness of the crossover is illustrated in figure 11(b) where for a chain of 200 spins the negativity with the region x containing 40 spins is plotted versus $J_{\rm K}$. As can be seen, $\mathcal{N}$ transitions from $\sim$0.5 to 0 very close to $J_{\rm K} = 0.15$. For $J_{\rm K} \leq 0.15$ we have not been able to identify any exponentially decaying part in $\mathcal{N}$ as obtained within the variational approach even for systems substantially longer than

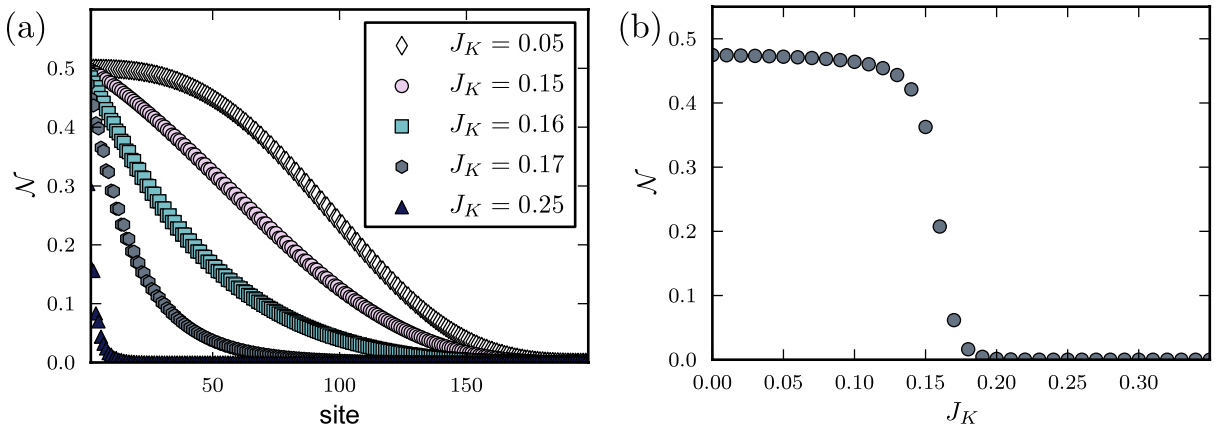


Figure 11. (a) Variational results for the negativity for different J_K with $L = 200$. (b) Variational results for the negativity for fixed size, $l = 40$, of the subsystem x as a function of J_K with $L = 200$. All results are obtained at the MG point with $\delta = 0$.

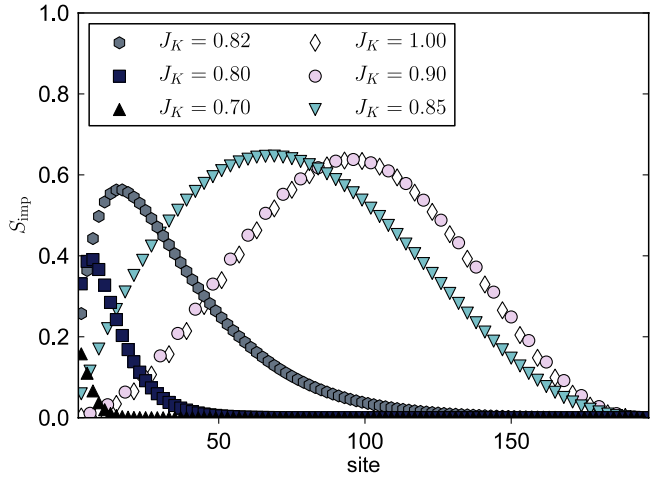


Figure 12. Variational results for the S_imp for different J_K with $L = 199$. All results are obtained at the MG point with $\delta = 0$. For clarity only a fraction of all data points are shown.

200 and it appears that a finite $J_\mathrm{K}^\mathrm{c} \sim 0.15$ is needed to induce an exponential decay in $\mathcal{N}$. This would imply that the presence of a non-zero J_K cannot be modelled as a simple one-dimensional potential well, since this would always have a bound state independent of the depth of the potential well. Probably, an enlarged variational space for the calculation will change the value for J_K^c and could possibly drive it all the way to zero. However, Bayat *et al* [11], using the DMRG, do not find an exponential decay for $J_2 = 0.46$ and $J_\mathrm{K} = 0.16$ consistent with a non-zero J_K^c. Due to the complexity of the calculation of $\mathcal{N}$ in larger subspaces we have, however, been unable to complete such calculations.

In a similar manner, we can study the single-particle entanglement (SPE) for general J_K within the variational approach. Our results are shown in figure 12. The single-particle

entanglement arises from the presence of a single unpaired spin for odd L which is always present in the ground state for odd L. The unpaired spin moves freely throughout the system when $J_K = 1$, but as J_K is decreased away from 1 towards zero it rapidly becomes localized on the impurity site, thereby quenching the impurity entanglement. As can be seen in figure 12, the impurity entanglement rapidly transitions around $J_K \sim 0.8$–0.85. While the variational approach agrees closely with DMRG data away from this transition region, some quantitative differences appear in this region due to the limitation of the Hilbert space used in the variational approach. The variational and DMRG results are however qualitatively the same and we have for clarity not included DMRG results in figure 12. We have verified that for $J_K \leq 0.85$ the data for S_{imp} shown in figure 12 develop an exponentially decaying part with an associated length scale, ξ_{loc}, different from the bulk spin–spin correlation length. Also here, it appears that ξ_{loc} diverges before J_K reaches 1 but this effect could possibly be due to finite-size effects in the DMRG calculations or limitations in the subspace used in the variational calculations.

9. Impurity entanglement with $\delta \neq 0$

The variational approach that we employ here is straightforward to use for any J_2, δ and J_K and we now discuss our results for the impurity entanglement for non-zero explicit dimerization $\delta \neq 0$. As we stressed in section 1, the impurity entanglement can be long range in these systems due to the near degeneracy of the two singlet ground states of the chain with periodic boundary conditions. With a non-zero δ this degeneracy is lifted and we expect the impurity entanglement to decrease dramatically. Essentially, this is because when δ is increased it is more costly for the impurity to be entangled with far away parts of the systems: if the impurity is in a singlet with a spin far away, there are more dimers on weak bonds than if the impurity bonds with a site that is close to it. If we consider L even, then, in order to have sizable entanglement, the bulk of the chain must have a site in a valence bond with the impurity spin and this spin is then bound to one end of the chain [42].

For a chain with an even number of spins we expect to be able to describe $\mathcal{N}$ in terms of an IVB picture for which an estimate of φ_{sol} is needed for non-zero δ. Such an estimate has been derived by Uhrig *et al* [46]. They showed that to good approximation $\varphi_{sol}(i) \propto \mathrm{Ai}(i/\xi + z)$, where $\mathrm{Ai}(x)$ is the Airy function and $z_1 = -2.3381$ is its biggest root. Here, the size of the area that the impurity spin binds to is roughly given by $\xi = (3m\delta/2)^{-1/3}$, where $m \approx 1/(1 + 7/\sqrt{65})$. Our variational approach directly yields the lowest state vector which is simply φ_{sol}. In figure 13(a) we compare this directly to $\mathrm{Ai}(i/\xi + z)$ where it can be seen that the agreement is good. It is noteworthy that the maximum in φ_{sol} is quite distant from the end of the chain, in agreement with previous results.

We can use this result to estimate $\mathcal{N}$ using equation (36). In figure 13(b) we see typical graphs of the behaviour of the negativity as δ is increased with $J_K = 0$ and $L = 500$. The range over which the impurity site is entangled is decreased upon increasing the dimerization. Figure 13(b) also includes the impurity valence bond entropy (IVB) obtained from equation (36) using the Airy functions for φ_{sol}. For easier comparison with the negativity we scaled by a factor of $1/2\ln 2$. For $\delta = 0$ we used equation (37) for φ_{sol}. A reasonable agreement between $\mathcal{N}$ and the IVB contribution (solid lines) is

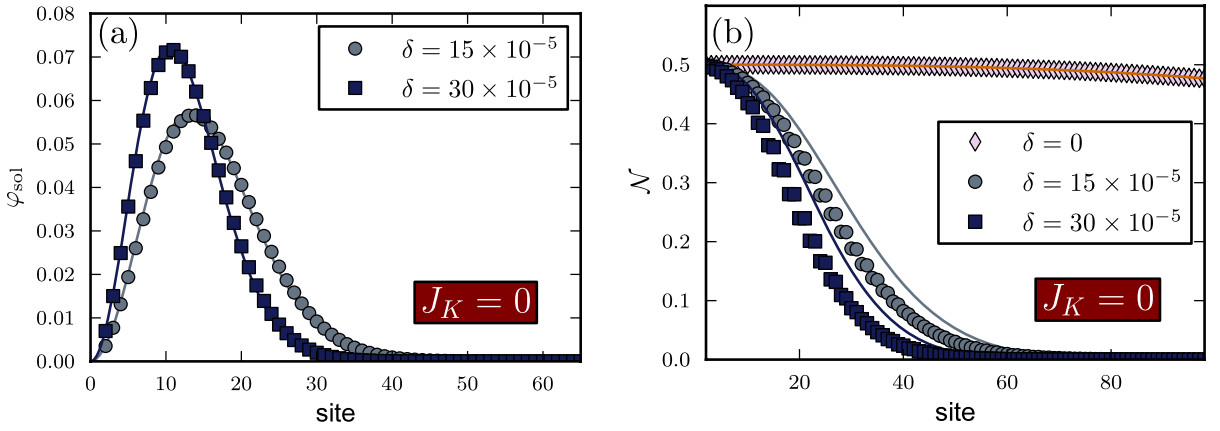


Figure 13. (a) Variational results for $\varphi_{\rm sol}$ (markers) compared to $\mathrm{Ai}(i/\xi + z)$ (lines); see the main text. (b) Variational results for the negativity at impurity coupling $J_{\rm K} = 0$ for three values of the dimerization δ plotted along with the scaled IVB entanglement (lines). The calculations were done for a chain of length $L = 500$, with $J_2 = J/2$.

observed. Overall, as one would expect, the impurity entanglement is now dramatically suppressed. Even for $\delta = 15 \times 10^{-5}$ the impurity entanglement is effectively zero beyond 60 sites. Following the above discussion we expect the extent over which the impurity entanglement is non-zero to diverge as $\delta^{-1/3}$ as dictated by ξ.

10. Conclusion

We have presented detailed discussion of how variational calculations can be used to calculate the negativity. This approach is quite generally applicable and should be very precise close to the MG point. A complete characterization of the impurity entanglement for any $J_{\rm K}$ can then be obtained. For $J_{\rm K} = 0$ (L even) the impurity entanglement is long range and extends throughout the system, while for larger $J_{\rm K} > 0.15$ it decays exponentially with a length scale clearly different from the correlation length in the system. For L odd the SPE present at $J_{\rm K} = 1$ is rapidly destroyed once $J_{\rm K}$ is decreased below $J_{\rm K} = 0.8$–0.85. In both cases it would be valuable to perform higher precision calculations in order to determine whether or not a critical $J_{\rm K}$ is needed to destroy the impurity entanglement. When a non-zero dimerization is introduced and the ground state degeneracy is lifted, the impurity entanglement rapidly disappears. It would be interesting to investigate impurity entanglement in other gapped systems with degenerate singlet ground states.

Acknowledgments

We are grateful to I Affleck and N Laflorencie for interesting discussions and prior collaborations that initiated this work. We also acknowledge fruitful discussions with A Bayat, S Bose and P Sodano and thank them for making their DMRG data for the negativity available to us. This research was supported by NSERC and CFI and was

made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET:www.sharcnet.ca).

References

[1] Calabrese P and Cardy J, *Entanglement entropy and quantum field theory*, 2004 *J. Stat. Mech.* P06002
[2] White S R, *Density matrix formulation for quantum renormalization groups*, 1992 *Phys. Rev. Lett.* **69** 2863
[3] Zhou H-Q, Barthel T, Fjaerestad J and Schollwöck U, *Entanglement and boundary critical phenomena*, 2006 *Phys. Rev.* A **74** 050305(R)
[4] Laflorencie N, Sørensen E S, Chang M-S and Affleck I, *Boundary effects in the critical scaling of entanglement entropy in 1D systems*, 2006 *Phys. Rev. Lett.* **96** 100603
[5] Cho S Y and McKenzie R H, *Quantum entanglement in the two-impurity Kondo model*, 2006 *Phys. Rev.* A **73** 012109
[6] Kopp A and Le Hur K, *Universal and measurable entanglement entropy in the spin-boson model*, 2007 *Phys. Rev. Lett.* **98** 220401
[7] Le Hur K, Doucet-Beaupre P and Hofstetter W, *Entanglement and criticality in quantum impurity systems*, 2007 *Phys. Rev. Lett.* **99** 126801
[8] Le Hur K, *Entanglement entropy, decoherence, and quantum phase transitions of a dissipative two-level system*, 2007 *Ann. Phys., NY* **323** 2208
[9] Sørensen E S, Chang M-S, Laflorencie N and Affleck I, *Impurity entanglement entropy and the Kondo screening cloud*, 2007 *J. Stat. Mech.* L01001
[10] Sørensen E S, Chang M-S, Laflorencie N and Affleck I, *Quantum impurity entanglement*, 2007 *J. Stat. Mech.* P08003
[11] Bayat A, Sodano P and Bose S, *Negativity as the entanglement measure to probe the Kondo regime in the spin-chain Kondo model*, 2010 *Phys. Rev.* B **81** 064429
[12] Sodano P, Bayat A and Bose S, *Kondo cloud mediated long-range entanglement after local quench in a spin chain*, 2010 *Phys. Rev.* B **81** 100412
[13] Eriksson E and Johannesson H, *Impurity entanglement entropy in Kondo systems from conformal field theory*, 2011 arXiv:1102.2492
[14] Affleck I, Laflorencie N and Sørensen E S, *Entanglement entropy in quantum impurity systems and systems with boundaries*, 2009 *J. Phys. A: Math. Theor.* **42** 504009
[15] Bose S, *Quantum communication through an unmodulated spin chain*, 2003 *Phys. Rev. Lett.* **91** 207901
[16] Christandl M, Datta N, Ekert A and Landahl A J, *Perfect state transfer in quantum spin networks*, 2004 *Phys. Rev. Lett.* **92** 187902
[17] Christandl M, Datta N, Dorlas T C, Ekert A, Kay A and Landahl A J, *Perfect transfer of arbitrary states in quantum spin networks*, 2005 *Phys. Rev.* A **71** 032312
[18] Burgarth D and Bose S, *Perfect quantum state transfer with randomly coupled quantum chains*, 2005 *New J. Phys.* **7** 135
[19] Burgarth D, Giovannetti V and Bose S, *Efficient and perfect state transfer in quantum chains*, 2005 *J. Phys. A: Math. Gen.* **38** 6793
[20] Wójcik A, Luczak T, Kurzyński P, Grudka A, Gdala T and Bednarska M, *Unmodulated spin chains as universal quantum wires*, 2005 *Phys. Rev.* A **72** 034303
[21] Plenio M B and Semião F L, *High efficiency transfer of quantum information and multiparticle entanglement generation in translation-invariant quantum chains*, 2005 *New J. Phys.* **7** 73
[22] Campos Venuti L, Degli Esposti Boschi C and Roncaglia M, *Long-distance entanglement in spin systems*, 2006 *Phys. Rev. Lett.* **96** 247206
[23] Campos Venuti L, Degli Esposti Boschi C and Roncaglia M, *Qubit teleportation and transfer across antiferromagnetic spin chains*, 2007 *Phys. Rev. Lett.* **99** 060401
[24] von Neumann J, *Thermodynamik quantenmechanischer Gesamtheiten*, 1927 *Nachr. Ges. Wiss Göttingen* 273
[25] Wehrl A, *General properties of entropy*, 1978 *Rev. Mod. Phys.* **50** 221
[26] Vidal G and Werner R F, *Computable measure of entanglement*, 2002 *Phys. Rev.* A **65** 032314
[27] Affleck I, Kennedy T, Lieb E H and Tasaki H, *Rigorous results on valence-bond ground states in antiferromagnets*, 1987 *Phys. Rev. Lett.* **59** 799
[28] Affleck I, Kennedy T, Lieb E H and Tasaki H, *Valence bond ground states in isotropic quantum antiferromagnets*, 1988 *Commun. Math. Phys.* **115** 477
[29] Tribedi A and Bose I, *Quantum critical point and entanglement in a matrix-product ground state*, 2007 *Phys. Rev.* A **75** 042304

[30] Tribedi A and Bose I, *Entanglement and fidelity signatures of quantum phase transitions in spin liquid models*, 2008 *Phys. Rev.* A **77** 032307

[31] Hirano T and Hatsugai Y, *Entanglement entropy of one-dimensional gapped spin chains*, 2007 *J. Phys. Soc. Japan* **76** 074603

[32] Fáth G, Legeva Ö, Lajkó P and Iglói F, *Logarithmic delocalization of end spins in the $s = \frac{3}{2}$ antiferromagnetic Heisenberg chain*, 2006 *Phys. Rev.* B **73** 214447

[33] Fan H, Korepin V and Roychowdhury V, *Entanglement in a valence-bond solid state*, 2004 *Phys. Rev. Lett.* **93** 227203

[34] Alipour S, Karimipour V and Mermarzadeh L, *Entanglement and quantum phase transitions in matrix-product spin-1 chains*, 2007 *Phys. Rev.* A **75** 052322

[35] Geraedts S D and Sørensen E S, *Exact results for the bipartite entanglement entropy of the AKLT spin-1 chain*, 2010 *J. Phys. A: Math. Theor.* **43** 185304

[36] Majumdar C K, *Antiferromagnetic model with known ground state*, 1970 *J. Phys. C: Solid State Phys.* **3** 911

[37] Haldane F D M, *Spontaneous dimerization in the $S = 1/2$ Heisenberg antiferromagnetic chain with competing interactions*, 1982 *Phys. Rev.* B **25** 4925

[38] Eggert S, *Numerical evidence for multiplicative logarithmic corrections from marginal operators*, 1996 *Phys. Rev.* B **54** 9612

[39] Eggert S and Affleck I, *Magnetic impurities in half-integer-spin Heisenberg antiferromagnetic chains*, 1992 *Phys. Rev.* B **46** 10866

[40] Laflorencie N, Sørensen E S and Affleck I, *The Kondo effect in spin chains*, 2008 *J. Stat. Mech.* P02007

[41] Shastry B S and Sutherland B, *Excitation spectrum of a dimerized next-neighbor antiferromagnetic chain*, 1981 *Phys. Rev. Lett.* **47** 964

[42] Sørensen E S, Affleck I, Augier D and Poilblanc D, *Soliton approach to spin–Peierls antiferromagnets: large-scale numerical results*, 1998 *Phys. Rev.* B **58** R14701

[43] Caspers W J and Magnus W, *Some exact excited states in a linear antiferromagnetic spin system*, 1982 *Phys. Lett.* A **88** 103

[44] Caspers W J, Emmett K M and Magnus W, *The Majumdar–Ghosh chain. Twofold ground state and elementary excitations*, 1984 *J. Phys. A: Math. Gen.* **17** 2687

[45] Beach K S D and Sandvik A W, *Some formal results for the valence bond basis*, 2006 *Nucl. Phys.* B **750** 142

[46] Uhrig G S, Schönfeld F, Laukamp M and Dagotto E, *Unified quantum mechanical picture for confined spinons in dimerized and frustrated spin chains*, 1999 *Eur. Phys. J.* B **7** 67

[47] Sutherland B, *Systems with resonating-valence-bond ground states: correlations and excitations*, 1988 *Phys. Rev.* B **37** 3786

3. Incommensurability effects in an odd length quantum spin chain

Andreas Deschner and Erik S. Sørensen:
Incommensurability effects in odd length J_1-J_2 quantum spin chains: On-site magnetization and entanglement;
Physical Review B **8**7, 094415 (2013); doi: 10.1103/PhysRevB.87.094415

Calculations: I performed all variational calculations. DMRG calculations were performed by Erik S. Sørensen.

Manuscript: I wrote the bulk of the manuscript and made all figures.
Erik S. Sørensen wrote section IV. He also added several paragraphs throughout the manuscript. Furthermore, Erik S. Sørensen provided (partly substantial) edits, comments and supervision.

The publication that is the subject of this chapter contains variational results concerning the onset of incommensurate behaviour in a dimerized spin chain of odd length, which turned out to be much more fascinating than expected (at least by me).

3.1. Incommensurate order

When different interactions compete, it is natural for a spin system to show incommensurate order. This means that the system shows order that is not compatible with the underlying lattice. This is a generic feature of many systems. A very simple and for this thesis most fitting example is the classical antiferromagnetic J_1-J_2 chain at $T=0$. At $T>0$ this system is not ordered because of the theorem of Mermin and Wagner [1966].
The energy is given by

$$\begin{aligned} E &= \sum \left[J_1 \ \mathbf{S}_i \cdot \mathbf{S}_{i+1} + J_2 \ \mathbf{S}_i \cdot \mathbf{S}_{i+2} \right] \\ &\propto \sum \left[J_1 \ \cos\left(\Theta_{i,i+1}\right) + J_2 \ \cos\left(\Theta_{i,i+2}\right) \right] , \end{aligned} \tag{3.1}$$

where $\Theta_{i,j}$ is the relative angle between spins on the sites i and j. For $J_2/J_1 \to \infty$, the

(a) The spin chain for $J_2/J_1 < 0.25$.

(b) The spin chain for $J_2/J_1 \to \infty$.

Figure 3.1.: The classical J_1-J_2 chain in the two limiting cases

groundstate is given by two antiferromagnetically ordered chains of next-nearest neighbours, i.e. $\Theta_{i,i+1} = \pi/2$ (or $-\pi/2$) and $\Theta_{i,i+2} = \pi$ (or $-\pi$). If $J_2/J_1 = 0$, the groundstate is antiferromagnetically ordered, i.e. $\Theta_{i,i+1} = \pi$ (or $-\pi$) and $\Theta_{i,i+2} = 2\pi$. The two extreme cases are shown in figure 3.1. Upon increasing J_2/J_1 from zero, the second nearest neighbours will have a tendency to not be parallel anymore. They will lean. This of course comes with an energy cost associated with nearest neighbours. If this cost is too high, nothing happens. If a high enough J_2/J_1 makes it worthwhile, something does happen. Because it is energetically favourable for any spin to be misaligned with both of their nearest neighbours by the same amount, all spins will lean by the same amount. The groundstate will therefore show constant shifts $\Theta_{i,i+1} = \Theta$ between neighbouring spins. In other words, there will be a wave vector of magnitude Θ associated with the order.

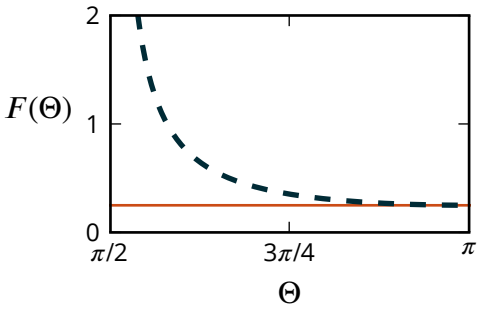


Figure 3.2.: The function $F(\Theta)$ (broken line) and a solid line which indicates the minimum value of J_2/J_1 for the energy to be stationary.

Θ can easily be derived. For the energy of a system with constant Θ we find

$$E \propto J_1 \cos(\Theta) + J_2 \cos(2\Theta) \tag{3.2}$$

Upon differentiation, we find the energy to be stationary when

$$F(\Theta) := -\frac{\sin(\Theta)}{2\sin(2\Theta)} = -\frac{1}{4\cos(\Theta)} = \frac{J_2}{J_1} . \tag{3.3}$$

As figure 3.2 tells us, there are no solutions when $J_2/J_1 \leq 0.25$. When $J_2/J_1 > 0.25$, Θ is not equal to π and may take any value between $\pi/2$ and π, i.e. the system is incommensurately ordered. The wave vector of the incommensurate order is given by

$$\Theta = \cos^{-1}\left(\frac{4J_2}{J_1}\right) . \tag{3.4}$$

The point $J_2/J_1 = 1/4$, thus, is a point at which three phases meet. A disordered phase at finite temperature and two ordered phases at $T = 0$, one of which shows incommensurate order. Such points are called Lifshitz points. Starting with Hornreich *et al.* [1975], Lifshitz

points have been studied since the 1970s. This research mostly focused on classical Ising systems such as the axial next-nearest neighbour Ising (ANNNI) model (see Selke [1988]). Lifshitz points have since been proposed for many different kinds of systems ranging from ferroelectrics (see Slivka *et al.* [1990]) to diblock copolymers (see Bates *et al.* [1995]). Chapter 6 in the book of Henkel and Pleimling [2010] might serve as a good source of more information.

In the publication of this chapter, we will be concerned with a quantum chain that shows behaviour similar to that of a classical system with a Lifshitz point.

3.2. Quantum effects

The classical chain has long-range order. Its spin spin correlation function oscillates with a certain wavelength; the wavenumber is π in the commensurate regime and between $\pi/2$ and π in the incommensurate regime. The static structure factor (the Fourier transform of the spin spin correlation function) is a delta function because there is only one ordering wave vector and no decay of the correlations. Immediately where incommensurate order starts, the peak in the static structure factor starts to shift. There is no doubt that the incommensurate order begins at the Lifshitz point.

The quantum J_1-J_2 chain's behaviour is different. Firstly, there is only short-ranged order as the chain is gapped in the region close to the Lifshitz point. Starting at $J_2 = 0.5\, J_1$, the Majumdar-Ghosh point, incommensurate order develops, but it is only apparent in real space (see Bursill *et al.* [1995]). This point is called the disorder point (about which I will say more in the next section). The static structure factor's maximum stays at π until another point, the Lifshitz point, is reached at $J_2 \approx 0.52063(6)\, J_1$ as was also found in Bursill *et al.* [1995].

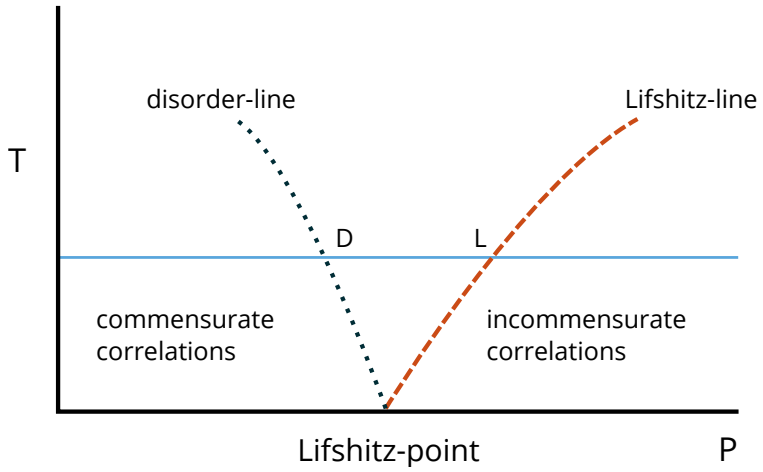


Figure 3.3.: Schematic generic phase diagram at a Lifshitz point (here at $T = 0$). The parameter P stands for any quantity that drives the transition.

The role of the AKLT point in the $S = 1$ chain with biquadratic term is very similar to the role of the Majumdar-Ghosh point in the J_1-J_2 chain. In Schollwöck *et al.* [1996] a study was undertaken of a transition from commensurate to incommensurate behaviour in the $S = 1$ chain with the Hamiltonian

$$H = \sum \left[\cos(\theta)\, \mathbf{S}_i \cdot \mathbf{S}_{i+1} + \sin(\theta)\, \left(\mathbf{S}_i \cdot \mathbf{S}_{j+1} \right)^2 \right] . \tag{3.5}$$

which shares many properties with the transition in the J_1-J_2 chain. There is a disorder point (the AKLT point at $\tan(\theta) = 1/3$) and also a Lifshitz point (at $\tan(\theta) \approx 0.44$). In their publication they provided an heuristic intuitive explanation for this observation. They draw heavily on the correspondence between classical systems and quantum systems whose dimension is smaller by one. The following exposition of the mechanism follows their original presentation given in Schollwöck *et al.* [1996].

That the transition happens at different points in real and reciprocal space, can be considered to be caused by quantum fluctuations effectively introducing a finite temperature. A schematic phase diagram at a Lifshitz point is shown in figure 3.3. The figure is drawn for systems that are ordered only at zero temperature. For other systems the mechanism presented here also works as long as one stays above but close to the critical temperature. Close to the ordered states at $\mathsf{T} = 0$, one expects the short-ranged correlation to still have their characteristics. To the left of the Lifshitz point, the correlations should be commensurate and to the right incommensurate. Close to the Lifshitz point, the two orderings will be competing. When one increases P at some finite temperature (moving along the blue line) there is a point, in the figure called D, at which incommensurate correlations appear in real space. Due to the competing orders, at this point, one naturally expects the correlation length to be minimal; thus, it is a disorder point. For a bigger P, at L, the incommensurate correlations will also be evident in the static structure factor.

Quantum fluctuations effectively make it as if there was a finite temperature. When changing P, one thus, even at $\mathsf{T} = 0$, moves along the blue line for the quantum chain.

That higher spin versions of the J_1-J_2 chain also exhibit a disorder point and a Lifshitz point (as seen in Roth and Schollwöck [1998]) was taken to be further evidence that this simple picture captures the main physics. Another strong point of this picture is that the AKLT point and the MG-point share many characteristics with disorder points in classical systems, as one would expect. I will dwell on these characteristics a little more in section 3.3.

It may be surprising that the structure factor shows no trace of incommensurate behaviour while it is, in real space quantities, plainly present between the disorder point and the Lifshitz point. Yet, it is not very difficult to understand how this can be the case. The following explanation follows Schollwöck *et al.* [1996].

We start with a very simple model static structure factor given by

$$\tilde{S}_{\text{sim}}(q) = \frac{1}{a(q-\pi)^2 + (q-\pi)^4 + \tilde{\xi}^{-2}} , \tag{3.6}$$

where $\tilde{\xi}^{-2}$ controls the width of the maximum. If the quartic term were absent and $a = 1$, then $\tilde{\xi}$ would be the correlation length of the system. The maximum of $\tilde{S}_{\text{gen}}$ is at $q_{\text{max}} = \pi$, if $a > 0$. It is more convenient to have the maximum at $q_{\text{max}} = 0$. We thus shift all momenta and work with

$$S_{\text{sim}}(q) = \frac{1}{aq^2 + q^4 + \tilde{\xi}^2} \,. \tag{3.7}$$

If $a > 0$, the static structure has a maximum at $q = 0$; if $a < 0$, the static structure has a local minimum at $q = 0$ and a maximum at a different q (see figure 3.4).

To find out about the corresponding spin spin correlation functions,

$$\langle \mathbf{S}(y) \cdot \mathbf{S}(y+x) \rangle = \int \mathrm{d}q \; S_{\text{sim}}(q) \;\; e^{-iqx} \,, \tag{3.8}$$

we have to have a look at the pole structure of the static structure factor, as this is what determines the long distance behaviour of the correlation functions. The integral can be calculated using Cauchy's integral formula by closing the arch in the lower half of the complex plane. Thus, the two poles with negative imaginary part are important. The poles are at

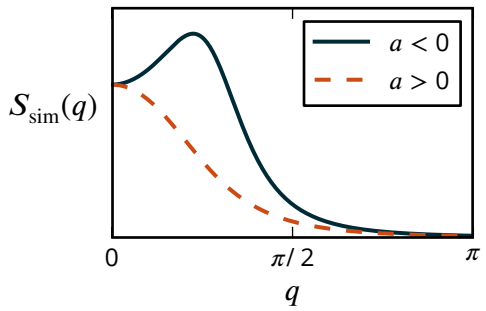


Figure 3.4.: The maximum of the generic structure factor of equation 3.7 is at zero for $a > 0$ and shifted from zero for $a < 0$.

$$u_{1,2,3,4} = \sqrt{g_{\pm}(a)} = \sqrt{-\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - \tilde{\xi}^2}} \,. \tag{3.9}$$

The structure of the poles for different a can be seen in figure 3.5. If $a > 0$ and $\frac{a^2}{4} > \tilde{\xi}^2$, $g_{\pm}(a)$ is negative and real. The poles of S_{sim} thus are purely imaginary. The two poles, say u_1 and u_2, with negative imaginary parts both contribute one term to the correlation function. These parts are proportional to $\exp(-iu_i x)$. Since the u_i are negative and purely imaginary this implies exponential decay. The correlation length is then given by $\xi^{-1} = \min\left[|u_i|\right]$.

If $a > 0$ and $\frac{a^2}{4} < \tilde{\xi}^2$, then the $g_{\pm}(a) = \frac{a}{2} \pm ik$, where k is a real positive constant. The two poles of S_{sim} with negative imaginary part thus have the same imaginary part but opposite real parts. The imaginary parts lead to exponential decay. The non-zero real part leads to oscillations in the correlation function. The maximum of S_{sim} is, for this value of a, still at the origin (see figure 3.4), yet incommensurate behaviour is apparent in the real space correlation function.

For not too negative $a < 0$ (where $\frac{a}{4} < \tilde{\xi}^2$), the poles have the same structure and incommensurate behaviour is prevalent in real space and in the static structure factor.

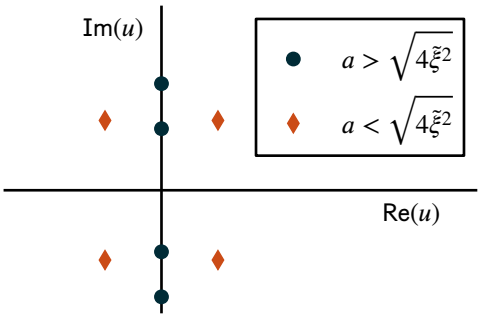


Figure 3.5.: The positions of the poles in the complex plane for two different settings of a and $\tilde{\xi}^2$.

For large and negative $a < 0$, this simple example is not applicable and S_{sim} develops poles on the real axis.

The simple example tells us the following. If the static structure factor has finite width, it can very well be that incommensurate correlations appear in real space while the structure factor shows no sign of this occurring. This is not possible in a classical system with its sharp maximum of the structure factor.

3.3. Disorder points

A disorder point is a point at which the character of short-ranged correlations changes. It was first defined in Stephenson [1970a]. The correlation function decays monotonically on one side and shows oscillations on the other.

The disorder point in the $S = 1$ chain of equation 3.5 (the AKLT point) very much resembles disorder points of the first kind found in classical systems (see Schollwöck *et al.* [1996]). Disorder points of the first kind are disorder points where the wave vector of the oscillation changes with the temperature. At disorder points of the second kind, the wave number of the oscillation does not depend on the temperature.

For such a point in classical J_1-J_2 Ising systems with ferromagnetic J_1 and antiferromagnetic J_2, it was, for example, found (see Stephenson [1970b]; Garel and Maillard [1986]) that everywhere, aside from the disorder point, the correlation function for big distances followed the Ornstein-Zernike form one might expect for a classical spin system and which is given by

$$\langle S_0 S_x \rangle \propto \left(\frac{1}{x}\right)^{(d-1)/2} e^{-x/\xi} , \tag{3.10}$$

where d is the dimension of the system. At the disorder point, the correlation function is of the same form but with the dimension reduced by one. Furthermore it was found that the wave number of the incommensurate oscillations grows like

$$q(T) \propto (T - T_D)^{1/2} , \tag{3.11}$$

where T_D is the temperature at the disorder point.

The behaviour of the correlation length at the disorder point is also remarkable. Its derivative diverges at the disorder point if one approaches the disorder point from the commensurate side. It is finite on the incommensurate side. This creates a cusp at the disorder point as shown in figure 3.6. All these results are also true for the disorder point in the $S = 1$ chain as long as one uses $d = 1 + 1$ for the dimension and J_2/J_1 as the

temperature (see Schollwöck *et al.* [1996]; Fáth and Sütő [2000]).

For the J_1-J_2 chain with $S = ½$ the same behaviour is expected. This is mostly because it also has a disorder point and a Lifshitz point.

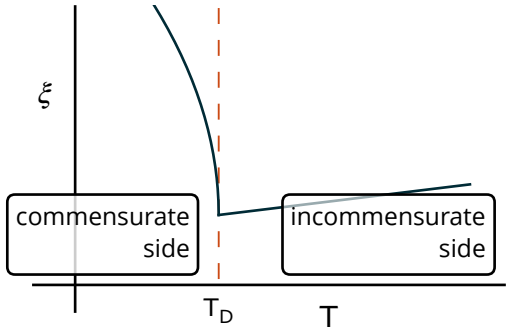


Figure 3.6.: Expected behaviour of the correlation length at a disorder point of the first kind. At the disorder point T_D (indicated by a dotted line) there is a cusp.

3.4. The paper

The J_1-J_2 chain has been studied extensively. The transitions through the disorder point and the Lifshitz point are no exceptions. Yet, all studies (most notably Bursill *et al.* [1995]; White and Affleck [1996]) of the transition from commensurate to incommensurate behaviour suffered from the same omission. They entirely focused on chains with an even number of sites. In dimerized systems it can, however, be very important whether or not the chain has an even or an odd number of spins. An odd system is for example not gapped and even at the MG-point its groundstate is not known. The paper in the last chapter is a nice illustration of how the parity of the length of the chain can be very important.

One could say that these differences become less and less important as the thermodynamic limit is approached. This consideration, while true in many regards, does not at all imply that odd-length systems are not important. The thermodynamic limit is not reached in real compounds, typical chains are only a few hundred to a few thousand spins long. For such chains it matters if they are of odd or even length. If one wants to have a full understanding of the J_1-J_2 chain as it might appear in a real system, an understanding of the odd-length chain is mandatory.

When working on the paper of chapter 2, I saw intriguing oscillations in the on-site magnetization in the odd-length J_1-J_2 chain develop at the Lifshitz point (see figure 10 in the paper).

Investigating these results with DMRG surprisingly proved to be very challenging. It was clear that we did not fully understand the physics at the transition from commensurate to incommensurate behaviour in the odd-length chain. We thus set out to study the transition in the odd-length J_1-J_2 chain with DMRG and variational calculations.

While working on the project, we noticed that there seemed to have been something else that never had been done. To our knowledge, it had never been checked whether the correlation length, at the MG-point, shows the behaviour expected at a classical disorder point (see last section).

The main questions we wanted to address for the odd-length chain were

- Why can DMRG not be used close to the Lifshitz point?

- How different (and in which way different) is the Lifshitz point for an odd-length chain?
- Does the correlation length show the expected behaviour at the MG-point (the disorder point)?
- Does the behaviour at the disorder point depend on the parity of the chain?

The answers to these questions can be found in the second publication in this thesis.

Incommensurability effects in odd length J_1-J_2 quantum spin chains: On-site magnetization and entanglement

Andreas Deschner* and Erik S. Sørensen
Department of Physics and Astronomy, McMaster University 1280 Main Street West, Hamilton, Ontario, Canada, L8S 4M1
(Received 18 December 2012; revised manuscript received 12 February 2013; published 14 March 2013)

For the antiferromagnetic J_1-J_2 quantum spin chain with an *even* number of sites, the point $J_2^d = J_1/2$ is a disorder point. It marks the onset of incommensurate real space correlations for $J_2 > J_2^d$. At a distinct larger value of $J_2^L = 0.520\,36(6)J_1$, the Lifshitz point, the peak in the static structure factor begins to move away from $k = \pi$. Here, we focus on chains with an *odd* number of sites. In this case, the disorder point is also at $J_2^d = J_1/2$ but the behavior close to the Lifshitz point, $J_2^L \simeq 0.538J_1$, is quite different: starting at J_2^L, the ground state goes through a sequence of level crossings as its momentum changes away from $k = \pi/2$. An even length chain, on the other hand, is gapped for any $J_2 > 0.24J_1$ and has the ground-state momentum $k = 0$. This gradual change in the ground-state wave function for chains with an odd number of sites is reflected in a dramatic manner directly in the ground-state on-site magnetization as well as in the bipartite von Neumann entanglement entropy. Our results are based on DMRG calculations and variational calculations performed in a restricted Hilbert space defined in the valence bond picture. In the vicinity of the point $J_2 = J_1/2$, we expect the variational results to be very precise.

DOI: 10.1103/PhysRevB.87.094415 PACS number(s): 75.10.Jm, 75.10.Pq, 75.50.Ee

I. INTRODUCTION

Disorder points were first discussed by Stephenson in models described by classical statistical mechanics.[1–4] On one site of a disorder point, the correlation function shows monotonic decay, on the other oscillatory decay. Depending on how the wavelength of the oscillation depends on the temperature, one distinguishes between two kinds of disorder points. If the wavelength of the oscillation depends on the temperature, one speaks of a disorder point of the first kind, if it does not, one speaks of a disorder point of the second kind.[2] In the first studies, disorder points were found where the paramagnetic phase of frustrated two-dimensional Ising models starts to show incommensurate instead of commensurate behavior. In models with competing commensurate and incommensurate order, one might expect such a point to occur where the short-range correlations with the largest correlation length change from being commensurate to being incommensurate. Such a point should then be associated with a cusp in the correlation length, a fact that was quickly established.[5] Schollwöck, Jolicoeur and Garel first investigated disorder points in a quantum spin chain for the bilinear-biquadratic $S = 1$ quantum spin chain,[6] which has $H = \cos\theta \sum \mathbf{S_i} \cdot \mathbf{S_{i+1}} + \sin\theta \sum \mathbf{S_i} \cdot \mathbf{S_{i+1}}$. They pointed out that the disorder point in this gapped quantum model coincides with the AKLT point where $\tan\theta_{\mathrm{VBS}} = 1/3$ and the known ground state is a valence bond solid (VBS) with correlation length $\xi = 1/\ln(3)$. They also identified another distinct point, the Lifshitz point, at $\tan\theta_L = 1/2$, where the peak in the structure factor is displaced away from π due to incommensurability effects. A third distinct point in this $S = 1$ model, $\tan\theta_{\mathrm{disp}} \simeq 0.4$, has also been located[7] where the minimum in the magnon dispersion shifts away from π and the curvature (velocity) vanishes. These three points can be distinct since no phase transition occurs and the correlation length remains finite. Subsequently, it was confirmed[8] that for the $S = 1/2$ J_1-J_2 spin chain with Hamiltonian

$$H = J_1 \sum_i \mathbf{S}_i \cdot \mathbf{S}_{i+1} + J_2 \sum_i \mathbf{S}_i \cdot \mathbf{S}_{i+2}, \quad J_1, J_2 > 0\,, \qquad (1)$$

the situation is similar. For the calculations presented in this paper, we set $J_1 \equiv 1$ and vary the remaining parameter J_2. The disorder point, with minimal correlation length ($\xi \simeq 0$), occurs at the Majumdar-Ghosh[9] (MG) point, $J_2^d = J_1/2$, and the Lifshitz point at $J_2^L = 0.520\,36(6)J_1$.[8]

At the disorder points, the ground states of these two quantum spin models share important features: the system is gapped and an exact wave function is known. For both, the momentum of the lowest excitations changes at a distinct point. Yet, there are also important differences between the two systems. While the $S = 1$ VBS state remains an exact state for a chain with an *odd* number of spins, this is *not* the case for the $S = 1/2$ J_1-J_2 chain at the MG point where no analytical expression for the odd length ground-state wave function is known. Moreover, the odd length J_1-J_2 chain is gapless in the thermodynamic limit within the $S_{\mathrm{Tot}} = 1/2$ subspace, and a large spin gap exists. The onset of incommensurability effects for odd length chains must then be quite distinct from the onset in even length chains. Here, we show that this is indeed the case. While the disorder point remains unchanged, the nature of the Lifshitz point, $J_2^L \simeq 0.538$, is rather different. At J_2^L, a sequence of level crossings starts, changing the ground-state momentum away from $k = \pi/2$. Although the correlation length remains small, the change in ground-state momentum induces pronounced oscillations directly in the on-site magnetization as well as the entanglement entropy. The modulations in the on-site magnetization are potentially observable in experiments. This scenario is reminiscent of a real Lifshitz transition[10–13] in which the ground state becomes modulated. The scaling of the entanglement entropy at Lifshitz transitions recently has been the subject of interest.[14,15]

The $S = 1/2$ J_1-J_2 antiferromagnetic (AF) spin chain is one of the simplest frustrated Heisenberg spin models, but it has a rich phase diagram. The system undergoes a transition[16] from a gapless Luttinger liquid to a dimerized phase at a critical value of $J_2^c = 0.241167J_1$.[17–19] For *even* length chains, the ground-state wave function at the MG point $J_2 = J_1/2$ is known to be formed by nearest-neighbor dimers.[9,20,21]

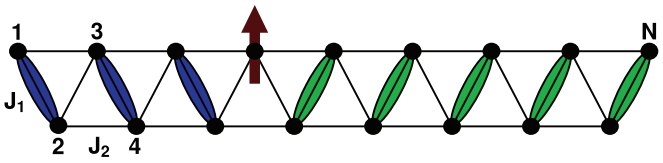


FIG. 1. (Color online) The odd length J_1-J_2 chain. Two different dimerization patterns are separated by a soliton.

It is twofold degenerate, corresponding to the two possible nearest-neighbor dimerization patterns, indicated in Fig. 1. As is evident from Fig. 1, an unpaired spin, a soliton[22,23] can act as a "domain wall" and separate regions of different dimerization patterns. In the Luttinger liquid phase, unpaired spins are more commonly called spinons since they do not act as domain walls. Spin excitations in the even length chain correspond to introducing two solitons, and it is known[24] that in the vicinity of the MG point, the solitons do *not* bind and a large spin gap of $\Delta = 2\Delta_{\rm sol}$ (at the MG point $\Delta_{\rm sol}/J_1 = 0.1170(2)$)[24] exists. The spin gap for even length chains is known to remain sizable[25,26] beyond J_2^L. The presence of a large soliton mass $\Delta_{\rm sol}$ renders variational calculations based on a reduced Hilbert space consisting of soliton states very precise;[22,23] a fact that we shall exploit here.

In contrast, for *odd* length chains it is not possible for the chain to be fully dimerized and the ground-state wave function is not known for any value of J_2. An $S = 1/2$ soliton that effectively behaves as a free particle[27] is always present in the ground state and gives rise to gapless excitations. Depending on the quantity in question, odd and even length chains can show very different behavior. Under open boundary conditions (OBC), this has, for example, already been seen in the on-site magnetization,[24] the entanglement entropy,[28,29] and the negativity.[30] As mentioned above, here we focus on odd length chains.

While it is possible to perform highly precise density matrix renormalization group (DMRG) calculations well beyond the onset of incommensurability for even chains,[25,26] the sequence of level crossings that we encounter for odd length chains for $J_2 > J_2^L$ significantly restrains the usefulness of the DMRG technique in a large region of parameter space for $J_2 > J_2^L$. Fortunately, using the picture of Shastry and Sutherland,[22] it is possible to quite efficiently perform very precise variational calculations for both open and periodic boundary conditions (PBC). Here, we mainly present results of such variational calculations and supplement them with DMRG results.

A number of spin-Peierls compounds, which to some extent realize the J_1-J_2 spin chain, have been identified. One of the most well known is $CuGeO_3$.[31] In these materials, impurities often cut the chains at random points. Therefore both odd and even length chains are present. A particular point of focus has been the study of $S = 1/2$ solitons[32–35] in these systems. Thus our results might be directly verifiable if materials with sufficiently large $J_2 > J_2^d$ can be found.

The outline of the paper is as follows. In Sec. II, the variational approach is described. Section III begins with a presentation of our DMRG results for the correlation functions, correlation lengths, and the structure factor. In Sec. III A, we discuss our variational results for the J_1-J_2 with periodic boundary conditions and show the change in ground-state momentum developing at the Lifshitz point. Section III B contains variational and DMRG results for the on-site magnetization and level crossings occurring with open boundary conditions. Variational and DMRG results for the entanglement entropy for a range of J_2 for odd length chains (OBC) are presented in Sec. IV and contrasted with results for *even* length chains (OBC). Finally, estimates for the location of the Lifshitz point are presented in Sec. V.

In the following, we shall take $J_1 \equiv 1$. This leaves us with only one parameter J_2 that governs the properties of the system.

II. THE VARIATIONAL METHOD

Most of the results presented in this paper were generated using variational calculations,[22,23,28,30,36] i.e., the results were obtained by minimizing the expectation value of the Hamiltonian within a reduced Hilbert space:

$$\langle H\rangle = \frac{(\varphi|H\varphi)}{(\varphi|\varphi)}\,, \tag{2}$$

where

$$\varphi = \sum c_j \varphi_j \tag{3}$$

and the minimization is done with respect to the c_j. To get a good estimate of the true ground state of the system, it is necessary that the ground state has a sizable projection onto the subspace one diagonalizes in. The quality of the result of a variational calculation thus depends very strongly on the choice of subspace. Often one has to rely on physical insight and intuition to choose well. For the J_1-J_2 model, which we consider, the selection of an appropriate subspace is straightforward as long as one stays in the dimerized phase. In contrast, in the Luttinger liquid phase, selecting an appropriate subspace seems intractable.

The first variational calculations on the J_1-J_2 model were done in a space that we in the following shall call R_0.[22,23] It is spanned by the states in which there are domains that have one of the two ground-state configurations of the MG chain and which are separated by one soliton. Examples can be seen in Figs. 1 and 2. The arrows in Fig. 2 serve to fix the phase of the dimers that make up the ground state. Our convention is such that if the arrow goes from site i to site j, the spins are in the state: $\frac{1}{\sqrt{2}}(|\uparrow\rangle_i|\downarrow\rangle_j - |\downarrow\rangle_i|\uparrow\rangle_j)$. For a chain with an odd number of sites, a set of single soliton states can be generated by leaving the chain maximally dimerized and taking the remaining site to be in the S_z=1/2 state. For a chain with open boundary conditions, the soliton can only reside on every second site. The dimension of this variational

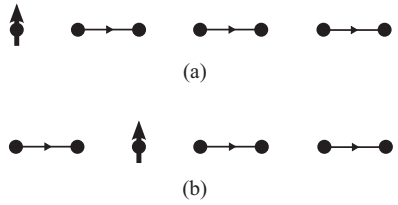


FIG. 2. Two variational states used in the calculations within R_0 for a chain with an odd number of sites. The arrows between sites are used to fix the phase of the dimers (see main text).

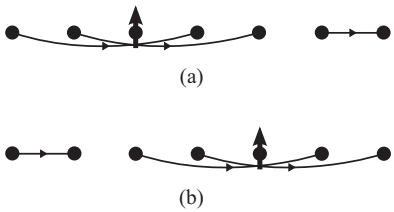


FIG. 3. Two variational states used in the calculations within R_1 for a chain with an odd number of sites. The arrows between sites are used to fix the phase of the dimers (see main text).

subspace is then $D = (N + 1)/2$. We use N to denote the length of the chain. For calculations on odd length chains with periodic boundary conditions, it is necessary to allow a nearest-neighbor dimer across the boundary and to let the soliton cross the boundary by going from site N to site 2. In this case, R_0 has dimension N and incorporates states with the soliton at every site with the remaining spins paired in nearest neighbor dimers. (For odd N and PBC, it becomes difficult to distinguish the two dimerization patterns since they twist into each other at the boundary. Still, the soliton clearly denotes a domain wall between the two patterns.)

To improve upon R_0, it is natural to act with the Hamiltonian onto the space as doing this repeatedly generates a space that contains the ground state if the starting space had any overlap with and all symmetries of the ground state. It was shown that acting onto R_0 with the Hamiltonian only once is at the MG point already enough to make the calculation almost exact.[23] For the J_1-J_2 model with $J_2 \neq 0.5J_1$, the linearly independent states generated by acting with the Hamiltonian onto R_0 fall into three classes, each of which corresponds to a variational subspace. (1) The variational space R_0 itself. (2) The variational space R_1 which is spanned by states in which sites to the left and right of the soliton are connected by a valence bond. Pictorial representations of example states are shown in Fig. 3. (3) The variational space R_2 which is spanned by states in which two neighboring sites are in a valence bond with their next-nearest neighbor. These states are generated by the action of the nearest-neighbor-terms and the next-nearest neighbor terms in the Hamiltonian on adjacent dimers in the states in R_0. Pictorial representations of example-states are shown in Fig. 4. In the case of the MG chain, J_1 and J_2 are

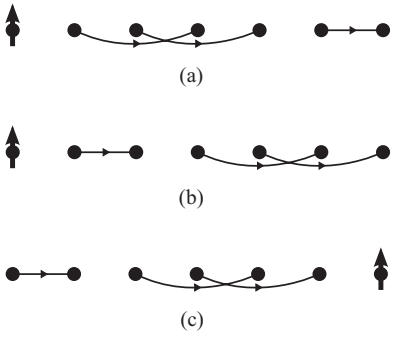


FIG. 4. Three states that are generated by acting with the Hamiltonian onto R_0. The arrows between sites are used to fix the phase of the dimers (see main text).

balanced in such a way that these states are not generated because they occur with a weight of $2J_2 - J_1$.

The number of states in R_0 and R_1 scales linearly with the size of the chain, whereas the number of states in R_2 scales quadratically. Due to computational cost, we have thus not found it practical to use the union of the three as the variational subspace for chains longer than 101 sites. We performed calculations using the union of R_0 and R_1 (in the following called Z_s) for chains up to 1001 and the union of all three (in the following called Z_b) for a chain of 101 sites. In this way, we could go to long chains and also check the validity through the comparison at $N = 101$. We found that while there were small quantitative differences between calculations done in Z_s and Z_b, the overall qualitative features of the results where the same. Therefore we chose to use Z_s, the union of R_0 and R_1, or just R_0 for the variational calculations shown in this paper.

All the states in these spaces have $S_z^{\mathrm{Tot}} := \sum_i S_z^i = 1/2$. We could equally well have worked in the $S_z^{\mathrm{Tot}} = -1/2$ space. States of higher total spin are of little importance to the low-energy physics since they contain more solitons and are thus gapped by at least twice the soliton mass Δ_{sol}. Since Δ_{sol} is sizable[24] in the regime of our study, such states can be disregarded for both odd and even N.

A variational description of a chain with an even number of sites can be done along the same lines. Again, the states are chosen in order to leave all but two spins in the favored dimerized state. In this way, one can gain insight into the low-energy singlet as well as triplet excitations by choosing the two spins to be in singlet or the triplet states, respectively.

If one considers a subspace with an orthogonal basis, one can just diagonalize the Hamiltonian. While an easy way to orthogonalize R_0 is known,[27] this is generally not true for other subspaces. Importantly, for R_1, no such method is known. We thus have to solve the generalized eigenvalue problem given by

$$\mathfrak{H}\,\vec{\varphi} = \lambda\,\mathfrak{B}\vec{\varphi}\,, \tag{4}$$

where $\mathfrak{H}_{ij} := (\varphi_i|H\varphi_j)$ and $\mathfrak{B}_{ij} := (\varphi_i|\varphi_j)$. Such generalized-eigenvalue problems can be solved numerically by standard routines. We calculate $\mathfrak{H}$ and $\mathfrak{B}$ by evaluating their defining expressions. This is possible because for valence-bond states the action of H on them as well as the overlap between them can straightforwardly be calculated in an automated manner. How to do all other calculations necessary to get the results presented in this paper has already been described in an earlier publication.[30] We took the coefficients c_j in Eq. (3) to be real. Also for PBC, the resulting wave function is not an eigenstate of the translational operator, which would have required the use of complex c_j's. Effectively, we obtain states that are linear combinations of translationally invariant states with k and $-k$, degenerate in energy. While this has no bearing on the obtained energies, it does affect real-space quantities like the on-site magnetization and entanglement, which cannot be translational invariant.

III. THE INCOMMENSURATE BEHAVIOR

Previous numerical studies of disorder points in $S = 1/2$[8,17,25,26,37,38] and $S = 1$[6,7,39–41] quantum spin chains have

concentrated on the behavior of even length chains. For the $S = 1/2$ J_1-J_2 chain, it has been shown that the disorder point of this one-dimensional quantum system can be understood as a $1 + 1$ dimensional classical disorder point. In particular, it was shown[42–44] that in the "commensurate" region of the phase diagram the correlation function behaves asymptotically, with $r \gg \xi$, as

$$\langle S_i S_{i+r} \rangle \sim (-1)^r \frac{e^{-r/\xi}}{\sqrt{r}}, \tag{5}$$

and in the "incommensurate" region of the phase diagram as

$$\langle S_i S_{i+r} \rangle \sim (-1)^r \frac{e^{-r/\xi}}{\sqrt{r}} \cos\left[(q - \pi)r + \phi\right]. \tag{6}$$

Here, q is the wave vector of the incommensurate correlations and ϕ a phase shift. However, right at the disorder point separating commensurate and incommensurate correlations the correlation function is asymptotically purely exponential:

$$\langle S_i S_{i+r} \rangle \sim (-1)^r e^{-r/\xi}. \tag{7}$$

For these quantum spin models it appears that this purely exponential behavior is in part connected to the fact that the ground state is an exact nearest-neighbor dimer state. Interestingly, as we shall see, the correlation functions at the MG point for odd length chains display the same behavior in the absence of a unique nearest-neighbor dimer ground state. Furthermore, it is known[42,43] that as the disorder point is approached from the commensurate side, the derivative of the correlation length with respect to the driving coupling becomes infinite, while it is finite on the incommensurate side. It is also known that the disorder point has special degeneracies that are exact for any system size N. For instance, for the J_1-J_2 chain with periodic boundary conditions and an even length, the two dimerization patterns are degenerate at the disorder point while their symmetric and antisymmetric combinations are split with an exponentially small gap away from this point.

We now present our results for the incommensurate effects in odd length $S = 1/2$ J_1-J_2 chains. Our first point of focus is the location of the disorder point. As stressed above, when N is odd, the nearest-neighbor dimer wave function is *not* an exact solution,[22] and there are also no special degeneracies. There is therefore no reason to expect that the behavior of the correlation length at the MG point is in any way unique. However, as we shall see, it is unique indeed. DMRG results for $C(r) = \langle S_i S_{i+r} \rangle$ for an open chain with $N = 201$ are shown in Fig. 5(a) for $J_2 = 0.5$ and 0.51. The correlation function follows a *purely* exponential decay at the MG point, $J_2 = 0.5$, with a finite correlation length:

$$\xi_{\mathrm{MG}} \sim 2.8 \ \ (\text{odd } N). \tag{8}$$

Distant spins in odd chains are correlated even at the MG point, because the soliton is present in the chain. The correlations can be thought of as correlations in the soliton wave function. Secondly, as can be seen in Fig. 5(a), incommensurate correlations are clearly present for $J_2 = 0.51$. They were present in every calculation we performed with $J_2 > 0.5$. We conclude that the disorder point remains at $J_2 = J_1/2$, albeit with a finite correlation length compared to the case of even N where the correlation length is nominally zero.

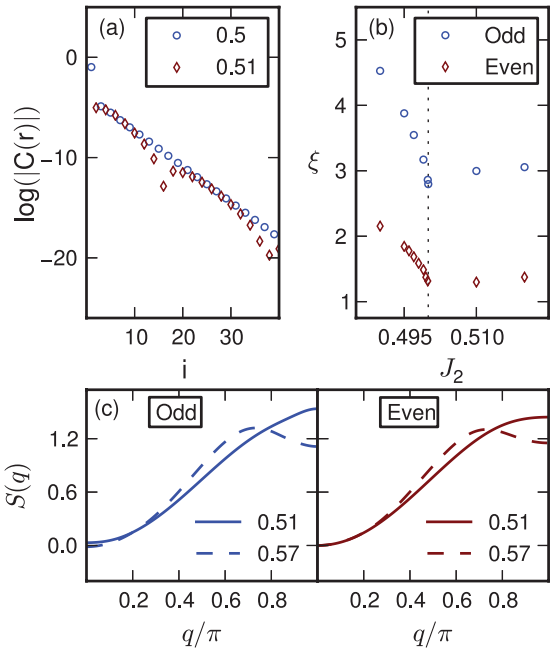


FIG. 5. (Color online) (a) The spin-spin correlation function $C(r) = \langle S_i S_{i+r} \rangle$ for a chain with 201 sites and $J_2 = 0.5$,0.51. The correlation length (b) and the static structure factor (c) for chains with an odd (201) or an even (200) number of sites. The data were obtained with DMRG. Open boundary conditions were employed and $m = 256$ states kept. For both, odd and even length chains, the correlation length displays a minimum at the MG point. The static structure factors of odd and even are shown for $J_2 = 0.51$ and $J_2 = 0.57$. The maximum remains at $q = \pi$ until the Lifshitz point is crossed (not shown in figure).

The precise behavior of the correlation length around the disorder point $J_2 = 1/2$ appears to have been studied neither for even length nor for odd length chains. Results for larger $J_2 > 0.6$ are available for even N.[26] By fitting DMRG results for chains of 200 and 201 sites to the forms of Eqs. (5) and (6) we have determined ξ as a function of J_2 for both even and odd N [see Fig. 5(b)]. The results for the even and the odd length chain are remarkably similar. At the disorder point, there is a discontinuity in the slope of ξ and on the commensurate side the slope of ξ approaches $-\infty$. We found that close to the disorder point in the commensurate region combined forms like $|C(r)| \sim C \exp(-r/\xi_C)/\sqrt{r} + D \exp(-r/\xi_D)$ with $\xi_C > \xi_D$ fit the data better than the single forms of Eqs. (5) and (6) because the dominant short-ranged correlations change at the disorder point. The results presented in Fig. 5 do not use such combined forms. We also note that for odd N and for a range of $J_2 > 0.538$ it becomes very difficult to obtain reliable DMRG results due to the appearance of many almost degenerate states.

The structure factor for even chains has been studied in some detail previously,[8,17] and the Lifshitz point has been located, $J_2^L = 0.520\,36(6) J_1$.[8] Our DMRG results are shown in Fig. 5(c). In agreement with previous studies for even N, we observe that the maximum in the structure factor remains at $q = \pi$ for $J_2 = 0.51$ but has clearly moved away from π at $J_2 = 0.57$. This is clearly also the case for odd N. Due to the above mentioned difficulties in obtaining reliable

DMRG results for odd N and $J_2 > 0.538$, we have not been able to determine the precise point where the peak in the structure factor is displaced from $q = \pi$. Using the variational techniques outlined above, it is possible to understand in detail what happens close to $J_2 \simeq 0.538$.

A. Variational results in periodic boundary conditions

We now turn to a discussion of our variational results obtained using the method outlined in Sec. II. We begin by focusing on the case of odd length chains and periodic boundary conditions. The case of open boundary conditions will be the subject of the next section. The results shown in this subsection were obtained using the space R_0 (see Sec. II), consisting of all single soliton states with $S_z^{\mathrm{Tot}} = 1/2$.

At the MG point, the spectrum of the J_1-J_2 model has been studied extensively. The feature that is most important to us is the low-lying dispersive line that is well separated from the continuum[24] and roughly follows a cosine as found in previous variational studies:[22,45]

$$E(k) = \tfrac{1}{8}[5 + 4\cos(2k) - 3N]. \tag{9}$$

Our variational method reproduces this estimate and agrees well with the low-energy data of an exact diagonalization of a chain of 23 sites (see Fig. 6). It may be surprising that the minimum of the dispersion relation is not at $k = \pi$ but at $k = \pi/2$. This is a natural consequence of the effective doubling of the unit cell that occurs because the action of the Hamiltonian displaces the soliton by *two* sites.

One of the strengths of the variational method is that within the limits of the approximation it is possible to easily access not only the ground state but also the entire energy spectrum within the subspace of $S_z^{\mathrm{Tot}} = 1/2$ states. Computing the spectrum through the transition region reveals very surprising behavior (see Fig. 7). All the states are twofold degenerate corresponding to the energetically degenerate k and $-k$. As one approaches the transition, the excited states linearly move closer and closer to the ground state. At the Lifshitz point $J_2^L \approx 0.53$, the energy of the first excited state crosses the ground-state energy. This level crossing marks the first shift in the ground-state momentum and is followed by a series of other level crossings at larger J_2 that further shift the ground-state momentum. Clearly, the presence of the many adjacent level crossings hinders the effectiveness of DMRG calculations.

This is in stark contrast to the spectrum of even length chains—the ground state of even length chains is exactly twofold degenerate at the MG point for any N, whereas for larger J_2, the symmetric and antisymmetric combinations are split with an exponentially small gap in N. The excited states are separated from these two states by a large gap of approximately $2\Delta_{\mathrm{sol}}$. This gap persists throughout the transition region and no level crossings are observed.[17,26]

It is very instructive to look at how the dispersion relation in Fig. 6 evolves with J_2. As can be seen in Fig. 8, the dispersion relation changes its shape when J_2 is increased. The minimum at $k = \pi/2$ first becomes flat very close to $J_2 = 0.53$ and then becomes a local maximum. In the process, two minima are

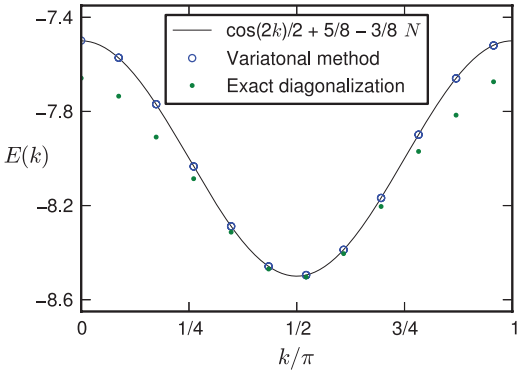


FIG. 6. (Color online) Comparison of the estimate by Shastry and Sutherland,[22] the variational method and exact diagonalization (ED) data.[24] Data were taken for a chain with 23 sites and $J_2 = 0.5$. Only the lowest energy of a spin 1/2 excitation for every momentum is shown. The deviations to ED occur at higher energies where the dispersive mode enters the continuum and the variational calculation is only of limited value.

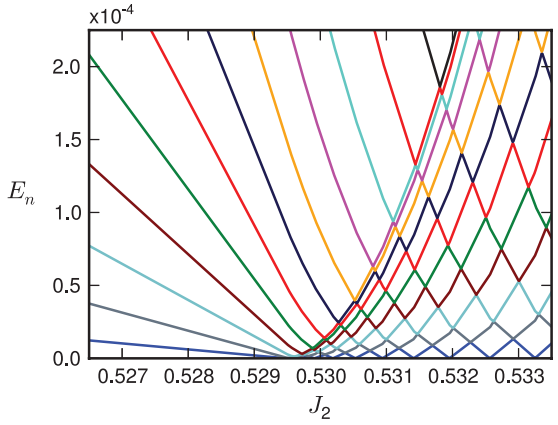


FIG. 7. (Color online) The excitation spectrum with periodic boundary conditions in the variational subspace with $S_z^{\mathrm{Tot}} = 1/2$. The energies of the first few excited states are shown. The ground-state energy was set to zero. The energies of the excited states approaches the ground-state energy until, at $J_2 = J_2^L$, level crossings start to occur. All energy levels are doubly degenerate. The data were taken for a chain with 401 sites.

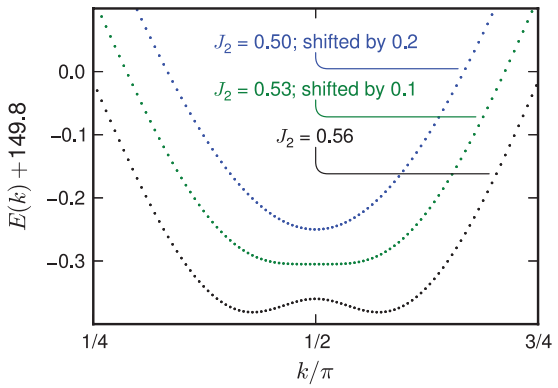


FIG. 8. (Color online) Dispersion relation for varying J_2. The data were taken for a chain with 401 sites. To avoid cluttering, the dispersion relations for the smaller two J_2 were shifted. As J_2 is increased, the minimum first flattens and then turns into a local maximum.

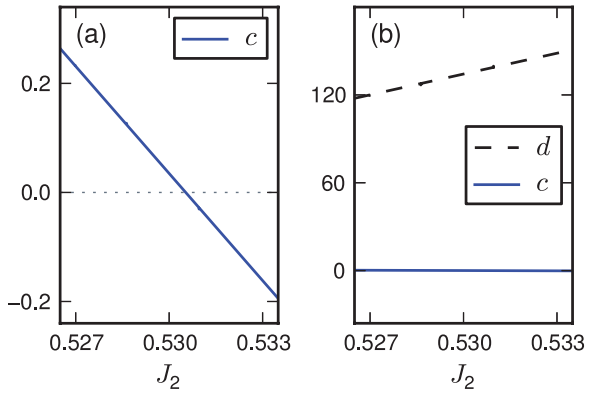


FIG. 9. (Color online) The second-order coefficient of the dispersion relation c changes sign [(a) and (b)], while the forth-order coefficient d stays positive (b). The data are results off a fit to data of the kind shown in Fig. 8. The dotted horizontal line in (a) separates positive from negative values.

created, which move away from $k = \pi/2$ with increasing J_2. The ground-state momentum is then clearly changing away from $k = \pi/2$ beyond $J_2 = 0.53$ and we may identify the point where this happens with a real Lifshitz transition[11,46–50] as opposed to the corresponding point in the $S = 1$ bilinear biquadratic chain where the ground-state momentum remains unchanged and the shift is in the excited magnon dispersion. Due to the shift in the ground-state momentum, we conclude that the maximum of the structure factor will shift away from $k = \pi$ at the same point. This is consistent with the data in Fig. 5. We therefore in the following refer to this point as the Lifshitz point J_2^L. The precise behavior of the dispersion relation close to J_2^L is analyzed in Fig. 9. For a range of J_2, we fitted the dispersion relation to the form[7]

$$E(k) = E(k_0) + \frac{c}{2}(k - k_0)^2 + \frac{d}{24}(k - k_0)^4 \qquad (10)$$

and confirmed that the second-order coefficient c changes its sign at a J_2 close to 0.53, while the fourth-order coefficient d stays positive (see Fig. 9). This behavior is typical of a Lifshitz transition and if the coefficient $c = v^2/\Delta_{\rm sol}$ is associated with a velocity v, the Lifshitz point signals the vanishing of this velocity.[7]

The variational calculations with periodic boundary conditions presented in this section were limited to the subspace R_0 described in Sec. II. This basis only includes nearest-neighbor valence bonds and it is quite noteworthy that the physics of the Lifshitz point along with the associated level crossings are captured within this simple basis set. However, as we discuss in Sec. V, we do not expect the precise location of the Lifshitz point to be accurately determined within R_0.

B. Open boundary conditions

In materials that realize the J_1-J_2 spin chain, impurities are always present. They often act as nonmagnetic impurities effectively breaking the linear chains into finite segments. The use of open boundary conditions is therefore closer to the experimental situation than the use of periodic boundary conditions. Furthermore, it is natural to expect half of the

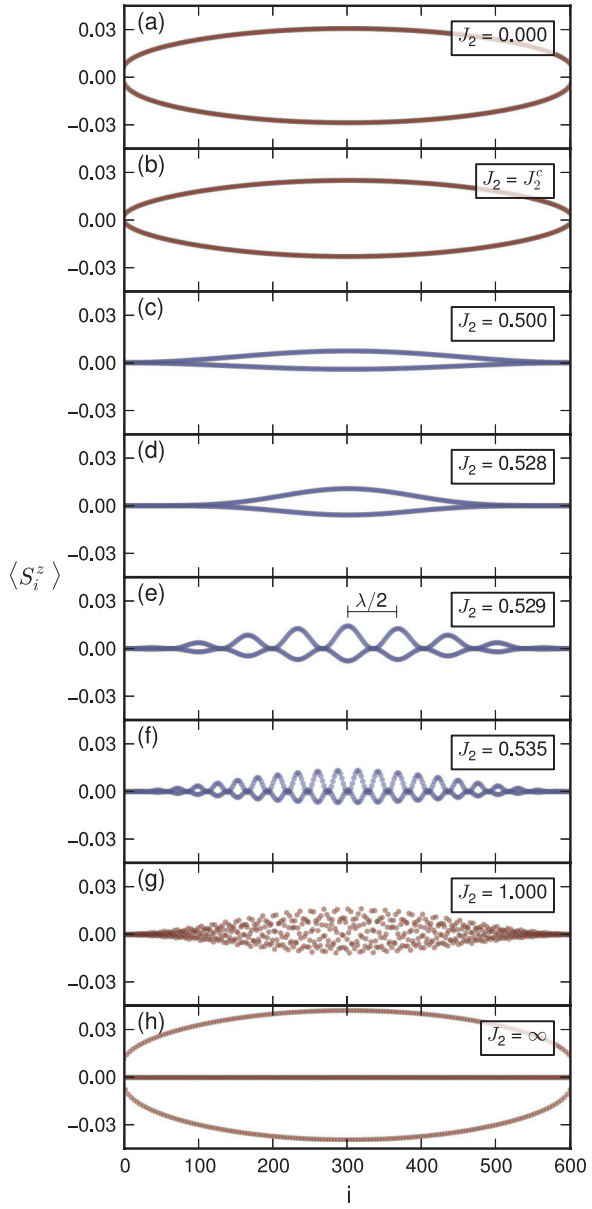


FIG. 10. (Color online) The on-site magnetization at a range of values of J_2 for $N = 601$ sites. Additional structure appears beyond J_2^L. For $J_2 = 0$, J_2^c, 1, and ∞ DMRG data (obtained with $m = 256$ states kept) are shown (red). For the remaining J_2 variational calculations are shown (blue).

chain segments to have an odd number of sites. In this section, we therefore focus on odd length chains with open boundary conditions. In particular, we describe the change that switching from periodic to open boundary conditions causes. The variational results shown in this section were obtained using the space Z_s (see Sec. II).

One quantity that very directly shows the qualitative difference between PBC and OBC is the ground-state on-site or local magnetization $\langle S_i^z \rangle$, which is of importance to, for instance, NMR measurements.[32,33] Figure 10 shows the on-site magnetization at eight different values of the frustrating interaction J_2 between $J_2 = 0$ and $J_2 = \infty$ in a chain of 601 sites. Figures 10(c)–10(f) show variational calculations through the Lifshitz point J_2^L where DMRG calculations are

less effective, the remaining results, Figs. 10(a), 10(b), 10(g), and 10(h), are obtained with DMRG.

a. The Luttinger liquid phase ($J_2 \leqslant J_2^c$). The transition to the dimerized phase occurs at J_2^c, see Fig. 10(b). At this point, as well as throughout the Luttinger liquid phase ($J_2 < J_2^c$), the on-site magnetization agrees very well with the prediction for the on-site magnetization in the ground state with $S_z^{\mathrm{Tot}} = 1/2$ from conformal field theory:[51]

$$\langle S_i^z \rangle = C(-1)^i \sqrt{\frac{\pi}{2N} \sin\left(\frac{\pi i}{N}\right)} + \frac{1}{2N}, \qquad (11)$$

where C is a constant. In this phase, $\langle S_i^z \rangle$ increases with the characteristic behavior $\langle S_i^z \rangle \sim \sqrt{i}$ for small i close to the boundary.

b. Dimerized phase with $J_2 < J_2^L$. Once the dimerized phase is entered, $\langle S_i^z \rangle$ is drastically altered. The on-site magnetization roughly follows the behavior of a massive particle in a box[28] with $\langle S_i^z \rangle \sim i^2$ close to the boundary. This behavior is visible at the MG point [see Fig. 10(c)]. As J_2 is increased beyond the MG point towards the Lifshitz point, J_2^L, the central peak sharpens [see Fig. 10(d)].

c. "Incommensurate" phase $J_2 > J_2^L$. At the Lifshitz point, there is another dramatic change in $\langle S_i^z \rangle$: additional maxima develop and the magnetization is modulated by an oscillating function [see Fig. 10(e)]. Upon increasing J_2 further, more such maxima form and the wavelength of the modulation decreases [see Figs. 10(f) and 10(g)]. If J_2 is fine-tuned for a given N, it is possible to find a point where two maxima occur in $\langle S_i^z \rangle$, then three maxima, and so forth.

It is natural to expect this behavior based on the results for PBC presented in Sec. III A. The local magnetization is effectively modulated with the momentum of the ground state. The running wave found under periodic boundary conditions is converted to a standing wave under open boundary conditions. Then, as the momentum of the ground states changes with growing J_2, the wavelength of the modulation shrinks. Finally, in Fig. 10(h), we show results for $J_2 \to \infty$. In this limit, the odd length chain with N sites is split into two chains with $(N-1)/2$ and $(N+1)/2$ sites, one of which will have an even number of sites and hence $\langle S_i^z \rangle \equiv 0$. The on-site magnetization of the other chain can be found by calculating $\langle S_i^z \rangle$ for a chain with $J_2 = 0$ of the same length. The results shown in Fig. 10(h) were obtained in this way, i.e., from data for a chain with $N = 301$ and $J_2 = 0$ that was then interspersed with zeros from the half of the chain that had an even number of sites.

To estimate the wavelength of the incommensurate modulation, we make use of the fact that, if our system had translational invariance, the distance between maxima in the on-site magnetization would be equal to half of the wavelength, as indicated in Fig. 10(e). Thus, by calculating the mean distance of the central maxima, we are able to determine an estimate for the wavelength of the incommensurate modulation. The inverse of this quantity can then be used to calculate the wave number, $q_{\mathrm{est}} = 2\pi/\lambda_{\mathrm{est}}$. In Fig. 11(a), we show how q_{est} varies with J_2 for four chains whose length ranges from 301 to 1001 sites.

Since the incommensurate behavior can only be seen if the wavelength is shorter than the system, it starts later in smaller chains. Aside from small deviations, which can be attributed

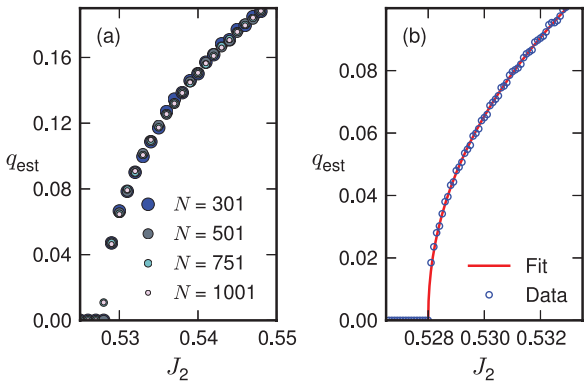


FIG. 11. (Color online) The wave number q_{est} against J_2. Data are shown for four chains of different length (a) and with the result of a fit $q_{\mathrm{est}} = 1.2062\ \Theta(J_2 - 0.528)(J_2 - 0.528)^{0.4806}$ for a chain of 601 sites (b).

to finite size effects, the wavelength does only depend on J_2 and not the length of the chain [see Fig. 11(a)]. In the limit of infinite J_2, the next-neighbor interaction J_1 can be neglected and the chain be partitioned into two subchains that do not interact. As mentioned, the J_1-J_2 model in this limit approaches two uncoupled chains with intrachain coupling J_2. The wavelength of the incommensurate behavior in this limit reaches its minimum with $\lambda = 4$ lattice spacings.

For $J_2 > J_2^L$, one expects the wave number q to behave as $q \propto (J_2 - J_2^L)^\alpha$, where $0 < \alpha < 1$.[6] In a study of correlations functions around the disorder point in the $S = 1/2$ J_1-J_2 model with an even number of sites and modified interactions on the edge of the chain, the exponent was reported to have been calculated to be $\alpha = 1/2$.[52] This is consistent with calculations on classical Lifshitz points,[11,46] which at the mean-field level find $\alpha = 1/2$. In the present case, where the ground-state momentum is changing, one might also expect corrections to the mean-field value of $\alpha = 1/2$ as described in Refs. 11 and 46.

Our calculations indeed confirm that $q_{\mathrm{est}}(J_2)$ follows a power law with exponent smaller than 1 [see Fig. 11(b)]. The line in Fig. 11(b) is a fit of the three-parameter function $f(x) = c_1\Theta(x - c_2)|x - c_2|^\alpha$, where $\Theta(x)$ is the Heaviside step function, to the blue data points also shown in the plot. Using this form, we find a value for the exponent $\alpha = 0.4806$. The data in Fig. 11 show steplike features. The cause of the steps is the introduction of new maxima: every time a new maximum appears, q_{est} jumps abruptly in order to accommodate the new maximum and there is a step. Between the appearance of new maxima, the maxima that are present move closer together and q_{est} increases smoothly. As one increases the system size, this effect affects the mean distance between maxima less and thus leads to less pronounced steps. Due to the different range of J_2 values, the steplike features explained above are more pronounced in Fig. 11(b) than in Fig. 11(a). Because of the inaccuracies, the steplike features introduce to the fitting procedure, we cannot comment on whether or not the corrections mentioned above are necessary.

While the on-site magnetization could relatively easily be understood from the results obtained with PBC, this is not the case for the energy spectrum. To the left of the transition

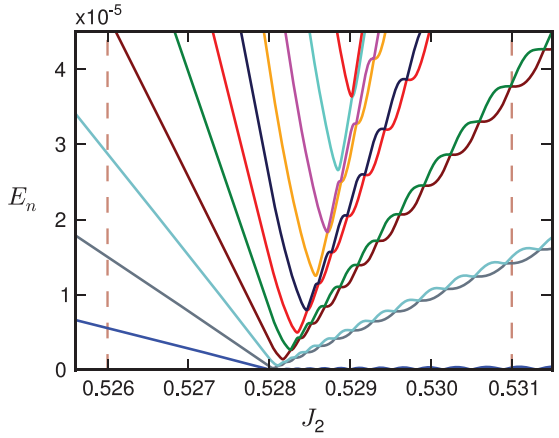


FIG. 12. (Color online) The excitation spectrum with open boundary conditions in the variational subspace with $S_z^{\mathrm{Tot}} = 1/2$. The energies of the first few excited states are shown. The ground-state energy was set to zero. The values at which the scaling with N is studied in Fig. 14 are indicated by dashed vertical lines. The data were taken for a chain with 601 sites.

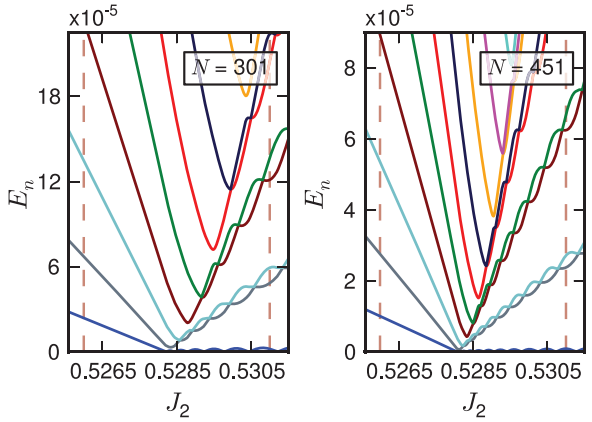


FIG. 13. (Color online) Spectrum with open boundary conditions for chains of 301 and 451 sites. The values at which the scaling with N is studied in Fig. 14 are indicated by dashed vertical lines.

($J_2 < J_2^L$), the spectrum for OBC (see Fig. 12) looks exactly like the spectrum for PBC (see Fig. 7)—yet there is an important difference: the spectrum for OBC is not degenerate. Introducing the boundary splits the degenerate states. On the other side of the transition ($J_2 > J_2^L$), the behavior of the energy of the first excited state also looks familiar: it hits the ground-state energy, grows, approaches it again, and another level crossing occurs. Repeated level crossings of just the two states follow. Higher-excitation levels, however, do not cross many other levels as they do for PBC. They approach the ground state, then turn around and form a pair with the state they would have been degenerate with under PBC. The two states exhibit a repeated pattern of intertwining level crossings, while their mean energy difference to the ground state grows. We do not know of an intuitive way of understanding the spectrum for OBC from the spectrum with PBC. Modifying the couplings at the boundary of the chain by a multiplicative factor of λ and varying λ between 0 and 1, we have studied the crossover from PBC to OBC. A low-energy spectrum similar to the one for OBC is observed until $\lambda \approx 0.9$.

It is reasonable to ask if the Lifshitz point is a well defined point in the spectrum. In order to answer this question, we show results in Fig. 13 for chains of length 301 and 451 sites for a range of J_2 close to J_2^L. As can be clearly seen, the turnarounds of the higher energy levels occur much closer to the first level crossing of the ground state for $N = 451$ than for $N = 301$. In the thermodynamic limit we expect the turnarounds for all higher lying levels to occur at J_2^L.

We next focus on the scaling of the energy levels with N. In Fig. 14(a), we show data for $N^2 E_n$ taken at $J_2 = 0.526$, to the left of the transition as indicated in Figs. 12 and 13. As can be seen, it converges to a constant value indicating that for this value of J_2 $E_n \propto N^{-2}$. For the first excited state, this behavior is apparent for quite short chains already and it seems plausible that for higher-excited states longer chains would lead to the same decay proportional to N^{-2}. This scaling is not surprising since the soliton behaves like a massive particle in a box. We therefore expect the low-energy spectrum to be approximated by $\hbar^2 k^2/(2\Delta_{\mathrm{sol}})$ with $k = \pi n/2N$, $n = 1,2,\ldots$, yielding the expected scaling of the energies as N^{-2}.

In Fig. 14(b), we show data taken on the other side of J_2^L at $J_2 = 0.531$ (again indicated in Fig. 12 and 13). For the smallest N shown, the higher excited states still show signs of the transition at this value of J_2. For short chains the second, third, forth, and fifth excited states thus have minimum in Fig. 14(b). For chains with more than roughly 160 sites, we see of intertwining pairs of states familiar from Figs. 12 and 13. The average energy of the pair at big N also scales proportionally to N^{-2}. We therefore conclude that sufficiently far away from the transition point, for the first few energy states, $\Delta E_n \propto J_2/N^2$.

In order to study the scaling of the spectrum at the Lifshitz point J_2^L, we focus on the minimum in the second excited state. Although this minimum occurs at slightly different J_2, as N is varied, it serves as the best possible definition of an excited energy scale at the Lifshitz point. Specifically, we define the minimal energy difference of the ground and the second excited states as $E_2^{\mathrm{min}} = \min_{J_2}(E_2 - E_0)$. Our results for E_2^{min} are shown in Fig. 15. As can be clearly seen in this

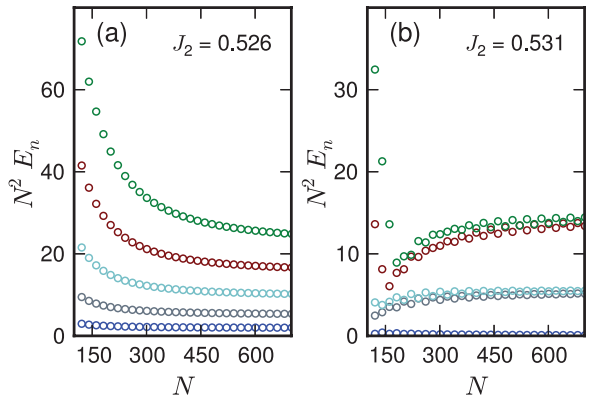


FIG. 14. (Color online) The energy scales proportionally to N^{-2}. Plot of $N^2 \Delta E_n$ to the left of the transition (a) and to the right of the transition (b) for the first five excited states.

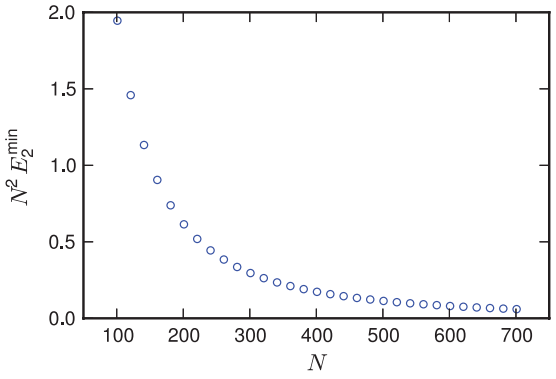


FIG. 15. (Color online) The energy at which the second excited state turns around goes to zero faster than N^{-2}. The value of J_2 of $E_2^{\min}$ was found up to $\Delta J_2 = 10^{-5}$. The resulting uncertainty of the value of the minimum energy is smaller than the size of the symbols in the plot.

figure, $E_2^{\min}$ goes to zero faster than N^{-2} violating the simple scaling found elsewhere.

We now turn to an estimate of the location of J_2^L within the variational approach. The level crossing of the first excited and the ground state allows for an easy way to define the value of J_2^L for a given N. As one could already see in Figs. 12 and 13, J_2^L varies slightly with the length of the chain. Our results are shown in Fig. 16 for chains out to $N = 701$. The main panel in Fig. 16 shows that J_2^L converges to approximately $J_2^L = 0.528$ as one increases the length of the chain.

The value of J_2 at which the second excited state has its minimum also approaches J_2^L. To show this, we use the value of J_2 at which the nth state reaches its first minimum for a given N. We call this quantity C_n. The inset in Fig. 16 shows $\Delta C_{21} = C_2 - J_2^L$. As can be seen, this quantity approaches 0 and the minimum for big N thus lies at the Lifshitz point.

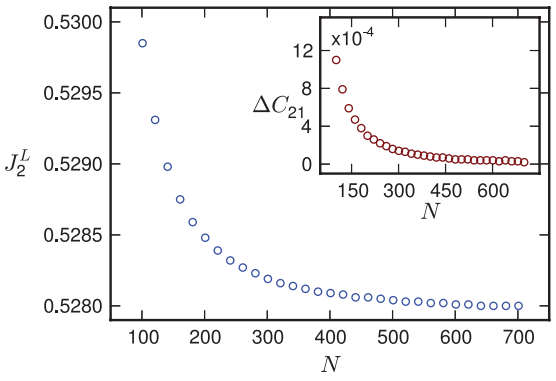


FIG. 16. (Color online) J_2^L as a function of the length of the chain. (Inset) The difference between J_2^L and the C_2 at which the minimum of the second excited state occurs. All values were determined up to $\Delta J_2 = 10^{-5}$. The resulting uncertainty of the value of the minimum energy is smaller than the size of the symbols in the plot and causes deviations from very smooth behavior.

IV. INCOMMENSURATE BEHAVIOR IN THE ENTANGLEMENT ENTROPY

The scaling of the entanglement entropy at a (quantum) Lifshitz transition has recently been the subject of interest.[14,15] In free fermion models, analogous to the spin chain model discussed here, the Lifshitz transition is associated with a change in the topology of the Fermi surface. In one dimension, new Fermi points appear at the Lifshitz transition and, analogously, new patches appear in higher dimensional models. If one associates a chiral conformal field theory with each patch, it can be argued[53] that, when the number of points (patches) increases by a factor K, the entanglement should be multiplied with the same factor K. For a free fermion model with next-nearest-neighbor hopping, t_2, one expects the number of Fermi points to double at the Lifshitz transition $t_2^L = 1/2$ at half-filling with a corresponding doubling in the entanglement entropy. This behavior is well confirmed in numerical calculations.[15]

In this section, we discuss our results for the entanglement entropy across the Lifshitz point in the odd length J_1-J_2 quantum spin chain which is the quantum spin analog of the model considered in Ref. 15. We study the entanglement in terms of the von Neumann entanglement entropy of a subsystem A of size l and reduced density-matrix ρ_A defined by[54,55]

$$S(l,N) \equiv -\mathrm{Tr}(\rho_A \ln \rho_A)\,, \tag{12}$$

where N again stands for the total system size. We consider exclusively open boundary conditions.

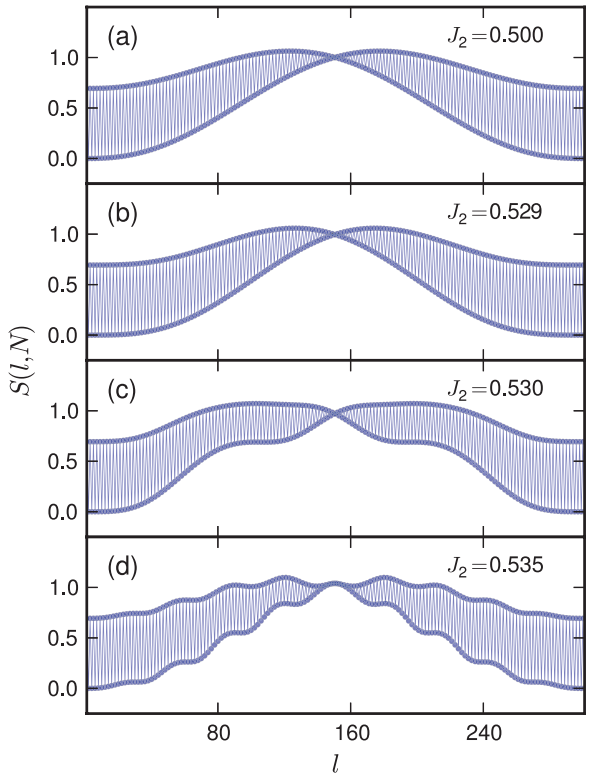
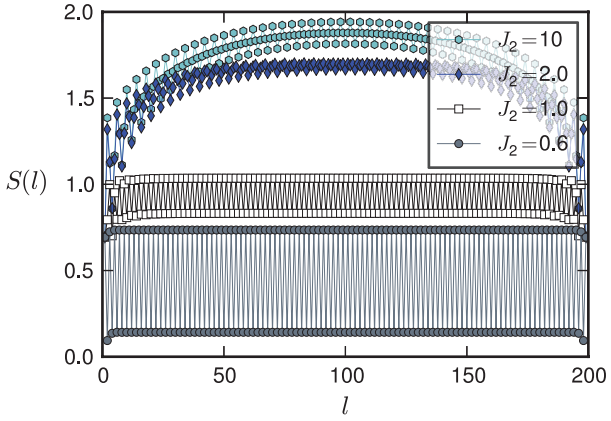


FIG. 17. (Color online) The entanglement entropy of a bipartition of an odd chain. The data were taken for chain of 301 sites and the variational subspace R_0 (see Sec. II) was used.

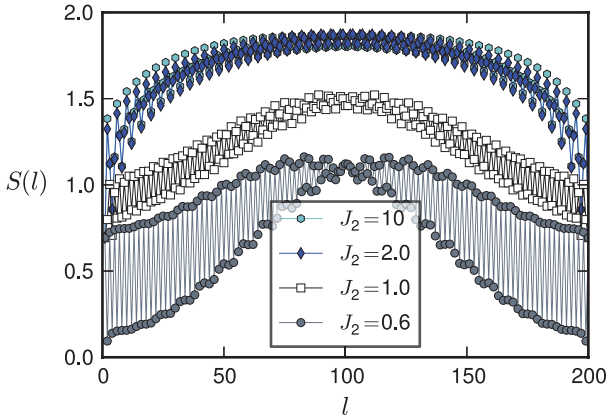
If one uses the restricted space R_0, which was introduced in Sec. II, as the variational subspace, one can also calculate the entanglement entropy using the method employed in this paper.[28] Away from MG and Lifshitz points, where the variational method is not reliable, we complement the variational results with DMRG calculations.

We first discuss the variational results for odd length chains close to the Lifshitz point shown in Fig. 17 for $N = 301$. The entanglement entropy at the MG point for the odd length chain, shown in Fig. 17(a), has previously been discussed in detail.[28] Since the entanglement entropy is very directly connected to the wave function of the state, drastic changes of the wave function should also be present in the entanglement entropy when the Lifshitz point is reached. This is clearly the case as can be seen in Fig. 17. As the Lifshitz point, J_2^L, is reached, the entanglement entropy develops *plateaus* [see Fig. 17(c)]. As J_2 is increased more plateaus appear [see Fig. 17(d)]. For the free fermion model studied in Ref. 15, analogous oscillations in the entanglement entropy are observed beyond $t_2 > 1/2$. Because a different subspace was used in the previous parts of this paper, the transition begins at $J_2 \sim 0.529$, which is slightly higher than $J_2 \sim 0.528$, which could be inferred from Fig. 13.

For an even length system, *no* such plateaus are visible [see Fig. 18(a)]. As $J_2 \to \infty$ the entanglement *increases* towards that of *two* independent gapless Heisenberg chains as it must. A similar increase is seen for an odd number of sites but with pronounced signatures of the incommensurability [see Fig. 18(b)].



(a) Even number of sites



(b) Odd number of sites

FIG. 18. (Color online) The entanglement entropy of a bipartition of an even chain (200 sites) (a) and an odd chain (201 sites) (b). The data were obtained using DMRG with $m = 256$ states kept.

V. THE TRANSITION POINT

The numerical value of J_2^L for the Lifshitz point depends not only on the length of the chain but also on the basis set that one uses in the variational calculation. While this is a small concern when one looks at qualitative features, it is of course detrimental if one is interested in a precise estimate of the Lifshitz point. Just using the different basis sets introduced in Sec. II, this is evident. Using the smallest basis, R_0, for a chain with $N = 301$ sites, we obtained $J_2^L \approx 0.5295$ (see Fig. 17) for the onset of oscillations in the entanglement. This is a slightly bigger value than what was found in Fig. 13, $J_2^L \approx 0.528$, based on the calculations with the larger basis Z_s.

DMRG can give us a more reliable estimate for J_2^L at least for small chains. For a chain with 201 sites, we found the first indications of incommensurate behavior in the local magnetization at

$$J_2^L \approx 0.538(1). \tag{13}$$

We expect this estimate to depend on N in roughly the same way as the variational estimate does in Fig. 16. If this is the case, an eventual extrapolation to the $N \to \infty$ limit might change this estimate by 0.0005, which is smaller than the uncertainty to which we have determined the point.

VI. CONCLUSION

We have studied incommensurability effects as they occur in the odd length antiferromagnetic J_1-J_2 chain. Even though no exact ground-state wave function is known at the MG point, $J_2 = J_1/2$, this point is the disorder point with minimal correlation length. The Lifshitz point $J_2^L = 0.538 J_1$ marks the onset of significant modulations directly in the ground state $\langle S_i^z \rangle$ as well as a shift in the ground-state momentum. A series of intertwining level crossings causing the shift in the ground-state momentum starts at the Lifshitz point. The shift in the ground-state momentum and the associated modulations directly affect the entanglement entropy, which shows distinct plateaus developing for $J_2 > J_2^L$.

In realistic compounds with chain breaking impurities, one would expect half the chain segments to be of odd length. The experimentally well studied compound $CuGeO_3$ has a $J_2 \sim 0.36 J_1 < J_1/2$.[56] If compounds with a J_2 in excess of $J_1/2$ can be identified, it would be very interesting to experimentally look for the odd length effects that we have detailed here. In particular, the effects on the on-site magnetization shown in Fig. 10 might be observable using NMR techniques or other local probes.

ACKNOWLEDGMENTS

We acknowledge many helpful discussions with Sung-Sik Lee as well as with H. Francis Song, Marlon Rodney, and Karyn Le Hur. This work is supported by NSERC.

*deschna@physics.mcmaster.ca

[1] J. Stephenson, Can. J. Phys. **47**, 2621 (1969).
[2] J. Stephenson, Phys. Rev. B **1**, 4405 (1970).
[3] J. Stephenson and D. Betts, Phys. Rev. B **2**, 2702 (1970).
[4] J. Stephenson, Can. J. Phys. **48**, 1724 (1970).
[5] J. Stephenson, Can. J. Phys. **48**, 2118 (1970).
[6] U. Schollwöck, T. Jolicoeur, and T. Garel, Phys. Rev. B **53**, 3304 (1996).
[7] O. Golinelli, T. Jolicoeur, and E. S. Sørensen, Eur. Phys. J. B **11**, 199 (1999).
[8] R. Bursill, G. A. Gehring, D. J. J. Farnell, J. B. Parkinson, T. Xiang, and C. Zeng, J. Phys.: Condens. Matter **7**, 8605 (1995).
[9] C. K. Majumdar, J. Phys. C **3**, 911 (1970).
[10] I. Lifshitz, Sov. Phys. JETP **11**, 1130 (1960).
[11] R. M. Hornreich, M. Luban, and S. Shtrikman, Phys. Rev. Lett. **35**, 1678 (1975).
[12] Y. Blanter, M. Kaganov, A. Pantsulaya, and A. Varlamova, Phys. Rep. **245**, 159 (1994).
[13] Y. Yamaji, T. Misawa, and M. Imada, J. Phys. Soc. Jpn. **75**, 094719 (2006).
[14] E. Fradkin, J. Phys. A **42**, 504011 (2009).
[15] M. Rodney, H. F. Song, S.-S. Lee, K. Le Hur, and E. S. Sørensen, arXiv:1210.8403.
[16] F. D. M. Haldane, Phys. Rev. B **25**, 4925 (1982).
[17] T. Tonegawa and I. Harada, J. Phys. Soc. Jpn. **56**, 2153 (1987).
[18] K. Okamoto and K. Nomura, Phys. Lett. A **169**, 433 (1992).
[19] S. Eggert, Phys. Rev. B **54**, R9612 (1996).
[20] C. K. Majumdar and D. K. Ghosh, J. Math. Phys. **10**, 1388 (1969).
[21] P. M. v. d. Broek, Phys. Lett. A **77**, 261 (1980).
[22] B. S. Shastry and B. Sutherland, Phys. Rev. Lett. **47**, 964 (1981).
[23] W. J. Caspers, K. M. Emmett, and W. Magnus, J. Phys. A **17**, 2687 (1984).
[24] E. S. Sørensen, I. Affleck, D. Augier, and D. Poilblanc, Phys. Rev. B **58**, R14701 (1998).
[25] R. Chitra, S. Pati, H. R. Krishnamurthy, D. Sen, and S. Ramasesha, Phys. Rev. B **52**, 6581 (1995).
[26] S. R. White and I. Affleck, Phys. Rev. B **54**, 9862 (1996).
[27] G. Uhrig, F. Schönfeld, M. Laukamp, and E. Dagotto, Eur. Phys. J. B **7**, 67 (1999).
[28] E. S. Sørensen, M. Chang, N. Laflorencie, and I. Affleck, J. Stat. Mech. (2007) P08003.
[29] I. Affleck, N. Laflorencie, and E. S. Sørensen, J. Phys. A **42**, 504009 (2009).
[30] A. Deschner and E. S. Sørensen, J. Stat. Mech. (2011) P10023.
[31] M. Hase, I. Terasaki, and K. Uchinokura, Phys. Rev. Lett. **70**, 3651 (1993).
[32] Y. Fagot-Revurat, M. Horvatić, C. Berthier, P. Ségransan, G. Dhalenne, and A. Revcolevschi, Phys. Rev. Lett. **77**, 1861 (1996).
[33] Y. Fagot-Revurat, M. Horvatić, C. Berthier, J.-P. Boucher, P. Ségransan, G. Dhalenne, and A. Revcolevschi, Phys. Rev. B **55**, 2964 (1997).
[34] M. Horvatić, Y. Fagot-Revurat, C. Berthier, G. Dhalenne, and A. Revcolevschi, Phys. Rev. Lett. **83**, 420 (1999).
[35] G. S. Uhrig, F. Schönfeld, J.-P. Boucher, and M. Horvatić, Phys. Rev. B **60**, 9468 (1999).
[36] C. Zeng and J. B. Parkinson, Phys. Rev. B **51**, 11609 (1995).
[37] K. Nomura and K. Okamoto, J. Phys. Soc. Jpn. **62**, 1123 (1993).
[38] K. Nomura and K. Okamoto, J. Phys. A **27**, 5773 (1994).
[39] A. Kolezhuk, R. Roth, and U. Schollwöck, Phys. Rev. Lett. **77**, 5142 (1996).
[40] R. Roth and U. Schollwöck, Phys. Rev. B **58**, 9264 (1998).
[41] E. Polizzi, F. Mila, and E. S. Sørensen, Phys. Rev. B **58**, 2407 (1998).
[42] T. Garel and J. M. Maillard, J. Phys. C **19**, L505 (1986).
[43] G. Fáth and A. Sütö, Phys. Rev. B **62**, 3778 (2000).
[44] K. Nomura, J. Phys. Soc. Jpn. **72**, 476 (2003).
[45] D. P. Arovas and S. M. Girvin, *Recent Progress in Many Body Theories III* (Plenum, New York, 1992), pp. 315–345.
[46] R. Hornreich and A. Bruce, J. Phys. A **11**, 595 (1978).
[47] R. Hornreich, J. Magn. Magn. Mater. **15**, 387 (1980).
[48] A. Michelson, Phys. Rev. B **16**, 577 (1977).
[49] A. Michelson, Phys. Rev. B **16**, 585 (1977).
[50] A. Michelson, Phys. Rev. B **16**, 5121 (1977).
[51] S. Eggert, I. Affleck, and M. D. P. Horton, Phys. Rev. Lett. **89**, 047202 (2002).
[52] K. Nomura and T. Murashima, J. Phys. Soc. Jpn. **74**, 42 (2005).
[53] B. Swingle, Phys. Rev. Lett. **105**, 050502 (2010).
[54] J. von Neumann, Nachr. Ges. Wiss. Göttingen, 273 (1927).
[55] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).
[56] J. Riera and A. Dobry, Phys. Rev. B **51**, 16098 (1995).

Concerning Valence Bond Quantum Monte Carlo

This part of the thesis is about work that was done in order to develop novel worm algorithms for a quantum Monte Carlo method called valence bond quantum Monte Carlo (VBQMC). In addition to a publication about the algorithms, this part contains an explanation of how VBQMC works and how its general scheme can be implemented in the S_z-basis.

4. Worm algorithms for valence bond quantum Monte Carlo

Andreas Deschner and Erik S. Sørensen:
Valence-Bond Quantum Monte Carlo Algorithms defined on Trees;
submitted to Physical Review E; pre-print: arXiv:1401.7592

Calculations: I performed all calculations.

Manuscript: I wrote the bulk of the manuscript and made all figures.
Erik S. Sørensen provided (partly substantial) edits, comments and supervision.

In the publication that is the subject of this part, we show that, using the notion of a worm travelling through a decision tree, intriguing algorithms can be developed for valence bond quantum Monte Carlo.

4.1. Valence bond quantum Monte Carlo

In this section, I introduce the method that is the subject of the publication that is the topic of this part. It is a quantum Monte Carlo method formulated solely in terms of valence bonds. Thus, it is called valence bond quantum Monte Carlo (VBQMC).

The first publications using VBQMC were published by Shoudan Liang in the very early nineties (see Liang [1990b,a]). These studies followed and made use of variational calculations performed in the valence bond basis published in Liang *et al.* [1988]. The method then laid dormant for many years until it was popularised by Anders Sandvik and then systematically further developed – mostly also by Sandvik (see e.g. Sandvik [2005], Sandvik and Beach [2007] and Sandvik and Evertz [2010]).

The unique features of VBQMC have been used in a number of very interesting studies. One great example are studies of deconfined criticality (see Sandvik [2007]; Pujari *et al.* [2013]). For those studies, a Hamiltonian was constructed that is perfectly suited for the use of VBQMC. The Hamiltonian is made up of two parts: the usual 2D-Heisenberg Hamiltonian on the square (Sandvik [2007]) or the hexagonal lattice (Pujari *et al.* [2013]), which favours

Néel-order, and a term which favours a valence bond crystal. This term lends itself naturally to be implemented with VBQMC. As the relative strength of the two terms is varied, the system undergoes a quantum phase transition.

VBQMC was also used in studies of entanglement. It was used to calculate the valence bond entanglement entropy. The valence bond entanglement entropy of two regions is proportional to the expectation value of the number of valence bonds that is cross the border of the two, a quantity that is computable with VBQMC. In Alet *et al.* [2007], it was proposed as a good computable measure of entanglement. Its scaling with the size of the smaller part of a bipartition was found to be different from the scaling of the von Neumann entropy (see section 2.1) in Kallin *et al.* [2009], yet it still shows different behaviour for gapped and gapless phases, making it potentially useful to detect phase transitions. The valence bond entanglement entropy has been reformulated in terms of the S_z-basis. Thus, methods without the sign problem can be used to calculate it for many frustrated systems (see Alet *et al.* [2010]).

Furthermore, in Hastings *et al.* [2010], it was shown that VBQMC can be used to calculate the Rényi entropy S^2 (see section 2.1).

As was shown in Beach *et al.* [2009], a generalized version of VBQMC can be used for SU(N) symmetric Heisenberg Hamiltonians. For the Heisenberg Hamiltonian on the square lattice, a phase transition at $N = 4.57(5)$ between a Néel-ordered and a bond ordered state was found.

4.1.1. How it works

VBQMC is a Monte Carlo method that can be used to calculate groundstate properties of Heisenberg models that are free of the sign problem (i.e. not frustrated) and have a groundstate that is a total spin singlet. It works by projecting the groundstate out of a trial state. Therefore, it is not a finite temperature but a $T = 0$ method.

In this section, I introduce VBQMC in a way that is directly applicable to systems with any spin. This approach is novel and has yet not been published. VBQMC can and has also been used on systems with $S \neq 1/2$ by mapping the original system onto a bigger system with $S = 1/2$. Then, the usual spin ½ algorithm can be used. An $S = 1$ chain of length L is then for example mapped into a $S = 1/2$ system of size $2 \times L$ (see Liang [1990a]). The method presented here circumvents a mapping of any kind. It rather naturally extends all expressions used for VBQMC directly to higher spin. Almost everything contained in this section that has nothing to do with whether or not we have $S = 1/2$, is a summary of the standard technique as it is, for example, published in Sandvik and Beach [2007].

To see how VBQMC is done and why it works, we start with the antiferromagnetic Heisenberg model on a bi-partite lattice, i.e. on a lattice where all interaction are between members of different sublattices:

$$H = \sum \mathbf{S}_i \cdot \mathbf{S}_j \ . \tag{4.1}$$

We will use Schwinger bosons to represent the spins on the lattice. This means that we introduce two kinds of operators which correspond to two kinds of bosons, a and b. They have the usual commutation relations for bosonic operators; $[a, a^\dagger] = [b, b^\dagger] = 1$ and $[a, b] = 0$. The states that can be generated by applying $b^\dagger$s and $a^\dagger$s to the empty state $|0\rangle$ and that fulfill the constraint

$$a^\dagger a + b^\dagger b = \hat{n}_a + \hat{n}_b = 2S \tag{4.2}$$

form a representation of SU(2) of total spin S (or equivalently dimension $2S$+1). For $S = 1/2$, the two species of bosons correspond to spin up and spin down, i.e.

$$a^\dagger|0\rangle \leftrightarrow |\uparrow\rangle \quad \text{and} \quad b^\dagger|0\rangle \leftrightarrow |\downarrow\rangle \; . \tag{4.3}$$

The spin operators are given by

$$\begin{aligned} S^x &= \frac{1}{2} \left(a^\dagger b + b^\dagger a\right) \; , \\ S^y &= \frac{1}{2i} \left(a^\dagger b - b^\dagger a\right) \; , \\ S^z &= \frac{1}{2} \left(a^\dagger a - b^\dagger b\right) \; . \end{aligned} \tag{4.4}$$

It is convenient to perform a rotation about the y axis on one of the sublattices, such that $S^x \to -S^x$, $S^y \to S^y$ and $S^z \to -S^z$. In terms of the operators a and b, this means that $a^\dagger \to -\tilde{b}^\dagger$ and $b^\dagger \to \tilde{a}$. Let us take the sublattice that corresponds to the index j in equation 4.1 to be the one that is rotated.

After the rotation the original Hamiltonian can then be written as

$$H = -\frac{1}{2} \sum_{\{i,j\}} \left[A_{ij}^\dagger A_{ij} - 2S^2\right] , \tag{4.5}$$

where

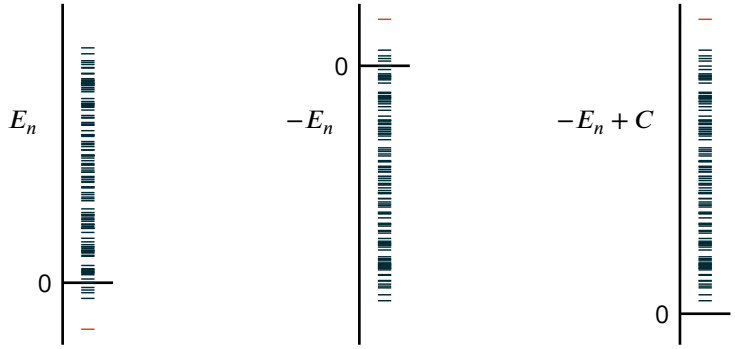
$$A_{ij}^\dagger = a_i^\dagger \tilde{a}_j^\dagger + b_i^\dagger \tilde{b}_j^\dagger \; . \tag{4.6}$$

The operators $A_{ij}^\dagger$ generate (unnormalized) singlet states (valence bonds) as

$$\begin{aligned} A_{ij}^\dagger|0\rangle &= (a_i^\dagger \tilde{a}_j^\dagger + b_i^\dagger \tilde{b}_j^\dagger)|0\rangle \\ &= (a_i^\dagger b_j^\dagger - b_i^\dagger a_j^\dagger)|0\rangle \\ &= |\uparrow_i\downarrow_j\rangle - |\downarrow_i\uparrow_j\rangle \; . \end{aligned} \tag{4.7}$$

In the second line I took back the rotation.

How the actual projection works is easy to understand. Take a Hamiltonian H with



(a) Spectrum of H. (b) Spectrum of $-H$. (c) Spectrum of $-H+C$.

Figure 4.1.: The projection operator is constructed by first flipping the spectrum (b) and then shifting it (c). The eigenvalue of the groundstate is indicated by the lone orange line.

bounded spectrum. From this operator, we construct an operator whose eigenstate with the eigenvalue of biggest magnitude is the groundstate of H.

First we multiply H with -1 to move the groundstate energy from the bottom to the top of the spectrum. Then we add an appropriate constant C to shift the spectrum up so that all eigenvalues are positive. Figure 4.1 illustrates this process.
Now we take a state $|\varphi\rangle$ (which has to have finite overlap with the groundstate). Repeatedly acting with $-H+C$ results in a projection onto the groundstate. This can easily be seen by expanding the state in energy eigenstates.

Then, we get

$$\begin{aligned}(-H+C)^n|\varphi\rangle =&(-H+C)^n \sum_{i=0}^{N} c_i \, |E_i\rangle \\ =&(-E_0+C)^n c_0 \left(|E_0\rangle + \sum_{i=1}^{N} \left[\frac{-E_i+C}{-E_0+C}\right]^n \frac{c_i}{c_0} \, |E_i\rangle + \dots \right) , \end{aligned} \tag{4.8}$$

where N is the dimension of the Hilbert space. The projective power n determines the quality of the projection.
Because $-E_i + C < -E_0 + C$ (by construction, see figure 4.1), the contributions of all excited states are suppressed relative to the groundstate.

In many cases the most sensible (and thus customary) choice for C is $C = \mathcal{N} S^2$, where $\mathcal{N}$ is the number of summands in the Hamiltonian. This makes the projection operator especially simple and we get

$$-H + C = \frac{1}{2} \sum A_{ij}^\dagger A_{ij} \, . \tag{4.9}$$

All different valence bond coverings of the lattice form a basis of the singlet subspace (see

e.g. Beach and Sandvik [2006]). If one knows that the groundstate is a singlet state, it thus suffices to do this projection in the subspace spanned by all valence bond coverings.

These states can be generated by acting with $A_{ij}^\dagger$s onto the empty state in all possible ways. The only condition is that every index appears with the correct multiplicity such that the sites are populated with the correct number of bosons and the constraint from equation 4.2 is not violated. For $S=1/2$, this means that each index should be acted on once; for $S=1$ this means that each index should be acted on twice.

The general idea is now to take a valence bond state α and to act on it with the projective operator as often as necessary.

With bond operators defined by

$$h_{ij} := \frac{1}{2} A_{ij}^\dagger A_{ij} \ , \tag{4.10}$$

this takes the form

$$\begin{aligned}(-H+C)^n|\alpha_{\text{trial}}\rangle &= \Big(\sum h_{ij}\Big)^n|\alpha\rangle \\ &= \sum_k \prod_{u=1}^{n} h_{i_{uk}j_{uk}}|\alpha\rangle \\ &= \sum_k s_k|\alpha\rangle \\ &= \sum_k \sum_r^{m(k)} w_{kr}|\alpha_{kr}\rangle \ ,\end{aligned} \tag{4.11}$$

where the operator string $s_k := \prod_{u=1}^{n} h_{i_{uk}j_{uk}}$ was introduced and $|\alpha_{kr}\rangle$ are valence bond states. The sum over r in the last line accounts for the possibility of a string generating more than one state.

It should be noted that $k \neq k'$ and $r \neq r'$ does not imply $|\alpha_{kr}\rangle \neq |\alpha_{k'r'}\rangle$. There are many operator strings that generate the same state.

The action of the projective operator is the same as the action of a sum of strings of n bond operators. The action of each of the summands results in states $|\alpha_{kr}\rangle$ and weights w_{kr}. It turns out that these constants are positive. Equation 4.11 is a good example for an expression that can be evaluated with Monte Carlo. One just has to be able to generate the $|\alpha_{kr}\rangle$ in relative proportions of w_{kr}.

In the $S=1/2$ case, acting with a bond operator onto any valence bond state returns exactly one valence bond state. Therefore acting onto any valence bond state with any operator string only returns one state times a weight w_k. The sum over r in equation 4.11 thus contains one summand and equation 4.11 becomes

$$(-H+C)^n|\alpha\rangle = \sum_k s_k|\alpha\rangle = \sum_k w_k|\alpha_k\rangle \ . \tag{4.12}$$

The action of the operator strings can easily be calculated if one knows the action of the bond operators. The action of the bond operators in turn follows simple rules that can be calculated. The following commutation rules are a good starting point.

$$\begin{aligned}\left[A_{ij}, A_{ij}^\dagger\right] &= 2 + \hat{n}_i + \hat{n}_j \ , \\ \left[A_{kj}, A_{ij}^\dagger\right] &= a_i^\dagger a_k + b_i^\dagger b_k := \chi_{ik} \ , \\ \left[\chi_{ik}, A_{km}^\dagger\right] &= A_{im}^\dagger \ . \end{aligned} \tag{4.13}$$

(a) Acting on a bond returns the state.

(b) Acting on sites without bond rearranges bonds and multiplies the state with ½.

Figure 4.2.: The action of a bond operator on valence bond states is given by two simple rules for systems with $S = 1/2$. Lines between sites stand for valence bonds connecting the two sites. The dotted ellipse marks the two sites that are acted on by the bond operator.

If $S = 1/2$, one has to deal with two different scenarios when acting with h_{ij} on a valence bond state: Either the sites i and j are in a valence bond, which we indicate by $|[i,j]\rangle$, or not. If they are in a valence bond, we find

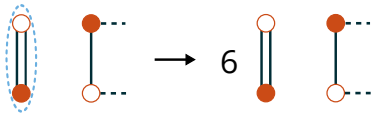
$$\begin{aligned} h_{ij}|[i,j]\rangle =& \frac{1}{2} A_{ij}^\dagger A_{ij} A_{ij}^\dagger |0\rangle \\ =& \frac{1}{2} A_{ij}^\dagger (2 + \hat{n}_i + \hat{n}_j)|0\rangle \\ =& A_{ij}^\dagger |0\rangle \\ =& |[i,j]\rangle \ . \end{aligned} \tag{4.14}$$

In this case, the state is just returned.

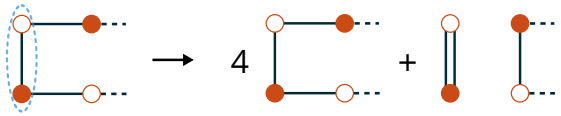
If they are not in a valence bond, the situation is slightly more difficult:

$$\begin{aligned} h_{kj}|[i,j][k,m]\rangle =& \frac{1}{2} A^\dagger_{kj} A_{kj} A^\dagger_{ij} A^\dagger_{km} |0\rangle \\ =& \frac{1}{2} A^\dagger_{kj} \left[A^\dagger_{ij} A_{kj} + \chi_{ik} \right] A^\dagger_{km} |0\rangle \\ =& \frac{1}{2} A^\dagger_{kj} \left[A^\dagger_{ij} A^\dagger_{km} A_{kj} + \chi_{mj} \right] |0\rangle \\ &+ \frac{1}{2} A^\dagger_{kj} \left[A^\dagger_{km} \chi_{ik} + A^\dagger_{im} \right] |0\rangle \\ =& \frac{1}{2} A^\dagger_{kj} A^\dagger_{im} |0\rangle \\ =& \frac{1}{2} |[k,j][i,m]\rangle \,, \end{aligned} \tag{4.15}$$

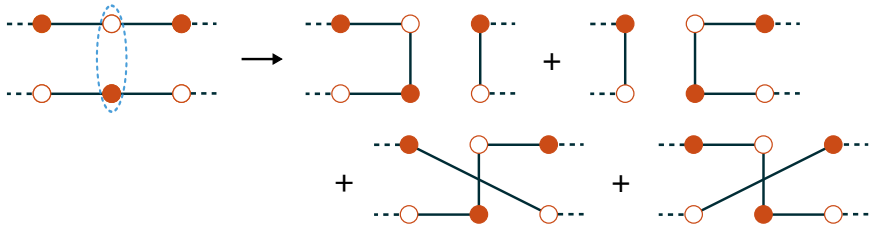
where χ_{ik} was defined in equation 4.13. In this case, the bonds are rearranged. After the action, the sites that were acted on are in a bond with each other. The sites they originally were in a bond with are then also in a bond with each other. The whole state is multiplied with 1/2 (see figure 4.2 for clarification).



(a) The bond operator acts on two sites that are connected with two bonds.



(b) The bond operator acts on two sites that are connected with one bond.



(c) The bond operator acts on two sites that are not connected.

Figure 4.3.: The action of a bond operator on a covering state of systems with $S=1$. The dotted ellipse marks the sites the bond operator acts on. Because we are considering spins with $S=1$, each site has two bonds connected to it.

These rules are more complicated for higher spin. The rules one obtains for $S=1$ are shown in figure 4.3. The most important new feature is that the action of a bond operator can return more than one state.

The most straightforward implementation of VBQMC works as follows. Firstly, one chooses a trial state $|\alpha_{\text{trial}}\rangle$. To make for a very efficient calculation, this state should have

very big overlap with the groundstate. Then one creates an operator string. By acting with the string on the state according to rules like the ones from figures 4.2 and 4.3, one then generates the weight and the state $|\alpha_{kr}\rangle$ from equation 4.11.

Since the action of each bond operator in the string only returns one state, this is easy to do for $S = 1/2$. One just collects the factors of 1/2 that are the result of hitting two sites that are not connected by a bond, while acting on the state with the operators in the string. This weight can then directly be used in a Metropolis rejection step. To get the next operator string (and thus state), one can then attempt to change a few of the bond operators and repeat the process.

If $S \neq 1/2$, this is not a trivial task because one operator string can return a large number of states. This problem can be overcome by introducing weighted branching. If the action of one bond operator returns more than one state with weights k_i, we choose one of the states with probability $k_i / \sum k_i$. The weight to be used in the Metropolis step is then given by $\prod_1^n \left(\sum k_i\right)$, where the product is taken over all bond operators in the string. How and why this sampling works is described in appendix A.2. Branching would also occur for $S = 1/2$ if one would not choose the offset C in the prescribed way. The suppression of excited states is controlled by the ratio of the shifted groundstate energy and the shifted excited energy. Choosing C as small as possible thus increases the suppression of the low-lying excited states. Choosing C very small will potentially lead to contamination by higher excited states that might no longer be suppressed sufficiently. Using an optimized trial state that does not contain such states might alleviate this problem. There have not yet been any attempts at increasing the projection quality using this approach.

There is no hard rule for how many operators one should change when going from one operator string to the next. Changing very few operators will change the string only very little but lead to high acceptance rates. Changing very many (or even all) will make it very unlikely that the new string is accepted. Thus, many update attempts are necessary in order to appreciably change the operator string. Maximising the average number of changed operators per update attempt will typically lead to acceptance rates around 0.2–0.4 and on the order of 4–10 changed operators in a Monte Carlo step. To make sure that consecutive measurements of the quantity one wants to calculate are sufficiently independent to allow for a reliable error estimate, one defines the Monte Carlo sweep as the smallest practical unit of calculation length. One sweep contains as many steps as are necessary to on average have attempted to change as many operators as there are in the string. The number of steps in a sweep thus scales linearly with the length of the operator string. How long a step takes is also proportional to the length of the operator string. Thus the computational time is of order $O(n^2)$. The acceptance rate is not very sensitive to both the length of the string and the size of the system under consideration (Sandvik and Beach [2007]).

There is no way to determine the necessary projective power for a given accuracy theoretically. Simulating and looking for how the result converges with a growing projective power is the way to go. There is however a quite useful rule of thumb for how the result should

converge. The second excited state is suppressed by a factor of

$$\left(\frac{-E_1+C}{-E_0+C}\right)^n = \left(1-\frac{\Delta E}{-E_0+C}\right)^n \approx \exp\left[-n\frac{\Delta E}{-E_0+C}\right] , \tag{4.16}$$

where we have introduced ΔE, which is the gap in the singlet sector. The singlet sector is all that matters because the trial state has no overlap with any states with $S \neq 0$. The gap can be expected to be proportional to its value in the thermodynamic limit plus some power of the size of the system, or in fewer words $\Delta E \propto \delta + a_1 N^{-\gamma} + a_2 N^{-\mu} + \ldots$.

The groundstate energy and the constant are both roughly proportional to the size of the system: $-E_0 + C \approx kN$. For equation 4.16 this means that

$$\left(\frac{-E_1+C}{-E_0+C}\right)^n \propto \exp\left[-\delta\frac{n}{N}\right]\exp\left[-k\frac{n}{N^{\gamma+1}}\right] , \tag{4.17}$$

where k is a constant. The required projection power to get some fixed accuracy is thus proportional to the size of the system in the case of a system whose singlet sector has a gap in the thermodynamic limit and proportional to $N^{\gamma+1}$ for gapless systems. For a system with a gap in the singlet sector, one finds (for big enough N) that

$$\langle E_0\rangle \approx E_0 + k\exp\left[-\delta\frac{n}{-E_0+C}\right]E_1 , \tag{4.18}$$

where k is again a constant.

The easiest quantity to calculate with VBQMC is the groundstate energy. There is a trick involved that works with operators to whom the groundstate is an eigenvalue. It is based on the Néel state having equal overlap with all valence bond states. In our rotated-boson basis, the Néel state, $|\text{Néel}\rangle$, is created by acting with $a^\dagger$ or $\tilde{a}^\dagger$ on all sites. If one expands the products of $A_{ij}^\dagger = a_i^\dagger \tilde{a}^\dagger + b_i^\dagger \tilde{b}^\dagger$ that generate valence bond states, there is always exactly one term that only contains $a^\dagger$s and $\tilde{a}^\dagger$s, regardless of which valence bond state one is looking at. This term is responsible for the overlap with the Néel state.

The groundstate energy can be calculated according to

$$\begin{aligned} E_0 &= \frac{\langle\text{Néel}|H|E_0\rangle}{\langle\text{Néel}|E_0\rangle} \\ &= \frac{\langle\text{Néel}|H\lim_{n\to\infty}(-H+C)^n|\alpha_{\text{trial}}\rangle}{\langle\text{Néel}|\lim_{n\to\infty}(-H+C)^n|\alpha_{\text{trial}}\rangle} , \end{aligned} \tag{4.19}$$

where $|E_0\rangle$ stands for the groundstate. With equation 4.11 we see that, if we are able to generate a set of $|\alpha_j\rangle$ with the correct probabilities, we can approximate the groundstate

energy according to

$$\begin{aligned} E_0 &\approx \frac{\langle \text{Néel}|H|\sum_k^n \sum_r^{m(k)} w_{kr}\alpha_{kr}\rangle}{\langle \text{Néel}|\sum_k^n \sum_r^{m(k)} w_{kr}\alpha_{kr}\rangle} \\ &\approx \frac{\langle \text{Néel}|H|\sum_j \alpha_j\rangle}{\langle \text{Néel}|\sum_j \alpha_j\rangle} \\ &= \frac{\sum_j \langle \text{Néel}|H|\alpha_j\rangle}{\sum_j \langle \text{Néel}|\alpha_j\rangle} \\ &= \frac{\sum_j \left[\sum_l \varepsilon_{jl}\langle \text{Néel}|\alpha\rangle\right]}{M\langle \text{Néel}|\alpha\rangle} \\ &= \frac{\sum_j \left[\sum_l \varepsilon_{jl}\right]}{M}, \end{aligned} \tag{4.20}$$

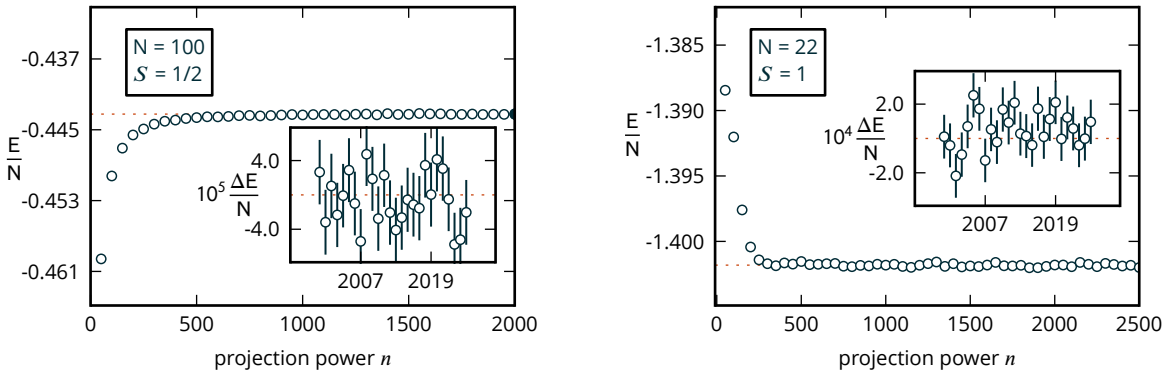
where the ε_{jl} are defined by $H|\alpha_j\rangle = \sum_l \varepsilon_{jl}|\alpha_{jl}\rangle$ and M is the number of terms in the sum on the second line. On the first line the "≈" is used instead of a "=" because n is not infinite. The sum in the second line is the Monte Carlo estimate of the sum in the first line; hence the "≈" in the second line.

Figure 4.4 contains the results for the groundstate energy of the Heisenberg model with $S = 1/2$ together with the correct result calculated with the Bethe ansatz by Erik S. Sørensen [panel (a)] and with $S = 1$ together with the result of an exact diagonalization (published in Golinelli *et al.* [1994]) [panel (b)]. The Monte Carlo data is presented as a function of the projection power n. Thus, to the left the values disagree with the correct value as the projection is bound to be insufficient. To the right however the Monte Carlo data agree with the known results nicely. It is noteworthy that there is no particular importance to the direction the groundstate energy is approached from.

To calculate operators' groundstate expectation value to whom the groundstate is not an eigenvalue is slightly more complicated. One has to perform two projections at the same time and approximate an expectation value of an operator O as

$$\begin{aligned} \langle O\rangle =& \frac{\langle E_0|O|E_0\rangle}{\langle E_0|E_0\rangle} \\ =& \frac{\langle \beta_{\text{trial}}|(\lim_{n\to\infty}(-H+C)^n)^\dagger \;\; O \;\; \lim_{n\to\infty}(-H+C)^n|\alpha_{\text{trial}}\rangle}{\langle \beta_{\text{trial}}|(\lim_{n\to\infty}(-H+C)^n)^\dagger \lim_{n\to\infty}(-H+C)^n|\alpha_{\text{trial}}\rangle}, \end{aligned} \tag{4.21}$$

where $|E_0\rangle$ again stands for the groundstate and the trial states β_{trial} and α_{trial} would usually be chosen to be the same. This is called a double projection. In a real calculation, n can of course only be finite. With $(-H+C)^n|\alpha_{\text{trial}}\rangle = \sum_k w_k|\alpha_k\rangle$ (and the equivalent for $|\beta_{\text{trial}}\rangle$



(a) $S = 1/2$, $N = 100$ (b) $S = 1$, $N = 22$

Figure 4.4.: With a big enough projective power, n, the Monte Carlo results agree well with the correct value for $S = 1/2$ (a) and $S = 1$ (b). The correct values were calculated with the Bethe ansatz (by Erik S. Sørensen) (a) and exact diagonalization (taken from Golinelli *et al.* [1994]) (b) and are indicated by a dotted line. The insets contain a magnification of a parameter range in which the calculation is almost fully converged and the data is almost statistically scattered around the groundstate energy per site.

just with the weights v_k), we get

$$\begin{aligned}\langle O\rangle &\approx \frac{\sum_j \sum_k v_j w_k \langle\beta_j|O|\alpha_k\rangle}{\sum_j \sum_k v_j w_k \langle\beta_j|\alpha_k\rangle} \\ &= \sum_j \sum_k \frac{v_j w_k \langle\beta_j|\alpha_k\rangle}{\sum_j \sum_k v_j w_k \langle\beta_j|\alpha_k\rangle} \frac{\langle\beta_j|O|\alpha_k\rangle}{\langle\beta_j|\alpha_k\rangle} \\ &= \sum_j \sum_k a_{jk} \frac{\langle\beta_j|O|\alpha_k\rangle}{\langle\beta_j|\alpha_k\rangle}\,, \end{aligned} \tag{4.22}$$

where in the last line the weights a_{ij} defined by

$$a_{ij} = \frac{v_j w_k \langle\beta_j|\alpha_k\rangle}{\sum_j \sum_k v_j w_k \langle\beta_j|\alpha_k\rangle} \tag{4.23}$$

were introduced. They are positive and normalised. One can thus use $v_j w_k \langle\beta_j|\alpha_k\rangle$ as weights in the Monte Carlo calculation.

The quality of the calculation very much depends on the choice of the trial state. One would like to maximise the overlap with the groundstate. A really simple way of doing this is to start with a randomly chosen state, run the simulation for a while and then take the result of the application of an operator string as the actual trial state. This does boost the performance. It can, however, be very much worthwhile to try a little harder to optimize the

groundstate.

The idea is to use a more complicated trial state and optimize it. Since this method was not used in the publication, which is the main focus of the chapter, the treatment here is rather brief. For more information I recommend the reader to have a look at Sandvik and Beach [2007].

The trial state is a superposition of valence bond states:

$$|\alpha_{\text{opt}}\rangle = \sum c_k |\alpha^k\rangle \ . \tag{4.24}$$

With such a state one can just as well perform the Monte Carlo projection; one just has more sums to deal with:

$$\langle O\rangle \approx \frac{\sum_i \sum_l \sum_j \sum_k c_i c_l v_j w_k \langle \beta_j^i | O | \alpha_k^l\rangle}{\sum_i \sum_l \sum_j \sum_k c_i c_l v_j w_k \langle \beta_j^i | \alpha_k^l\rangle} \ , \tag{4.25}$$

where $|\alpha_k^l\rangle$ is defined by $(-H+C)^n|\alpha^l\rangle = \sum_k w_k |\alpha_k^l\rangle$. The weights to be used in the calculation are $c_i c_l v_j w_k \langle \beta_j^i | \alpha_k^l\rangle$ and the estimator (the quantity that is evaluated and averaged) is

$$\frac{\langle \beta_j^i | O | \alpha_k^l\rangle}{\langle \beta_j^i | \alpha_k^l\rangle} \ . \tag{4.26}$$

Now one also has to update the state.

The c_k are taken to be functions of the length of the bonds in the state. They are defined as $c_k = \prod d(x_k^i, y_k^i)$, where the product is over all bonds in the state and $d(x, y)$ is some function of the distance of the two endpoints of a bond. One approach would be to now variationally choose the function $d(x, y)$ that minimizes the energy expectation value and use this state in the projection Monte Carlo.

A more efficient way is to stochastically improve the state and thus the $d(x, y)$ using the projection. The probability distribution of bond lengths in the state $|\alpha_{\text{opt}}\rangle$ is proportional to $d(x, y)$. After projection onto the groundstate we expect the distribution of bond lengths to be close to the optimal distribution. If $d(x, y)$ was originally chosen very well, the distribution of bond lengths should be the same before and after a projection.

By evaluating the probability distributions before and after the projection and adjusting $d(x, y)$ to be higher or smaller depending on whether or not there were too many or too few bonds with a specific length, it is possible to achieve a $d(x, y)$ such that the distributions do not change anymore. The state with this $d(x, y)$ is then used in the actual projection.

For most of the important studies an improved algorithm was used in conjunction with double projection. This loop algorithm was published in Sandvik and Evertz [2010] but by then had already been used by various groups.

Since it was not used in my work, I will not explain how it works but refer the interested reader to the original publication.

To use the algorithm described in the last section, one has to do $O(n)$ updates to change the operator string significantly. The time necessary for one update also scales with n, the length of the operator string, because for each update one has to act on the state with the whole string. Overall we thus are confronted with a computational time of $O(n^2)$.

The time one loop update takes scales linearly with the length of the operator string, just as one update step in the local implementation. A single loop update, however, already changes the string tremendously. The duration of one full loop update, thus, also scales linearly with the length of the string. Compared to the local update, much bigger systems can be accessed if one is using it.

The implementation of the loop update turns out to be very similar to that of loop updates for SSE, the very popular finite-T Monte Carlo method based on an expansion of the partition function (see Sandvik [1999]; Sandvik and Evertz [2010] and section 1.4).

4.1.2. Calculations in the S_z-basis

The projective technique described in subsection 4.1.1 is not only useful in the valence bond basis. It can, in fact, also be used in the product basis of eigenstates to the S_z-operator. In this subsection, I present previously unpublished results of a work in progress.

The main advantage of using the S_z-basis is that all basis states are orthonormal. Overlaps between basis states do not have to be calculated. Another difference is that one is not restricted to the spin singlet sector. The main implementational difference to standard VBQMC is that branching (see subsec 4.1.1 and appendix A.2) occurs very frequently.

The algorithm is very simple to implement and understand. We call this method S_z-projector Monte Carlo (SPMC).

Where VBQMC is applicable, it is faster than the SPMC. This advantage comes with a built-in restriction: everything takes place in the subspace with $S_{\text{tot}} = 0$.

We explain the algorithm using the example of the XXZ-model with

$$H = \sum_{(i,j)} \left[S_i^x S_j^x + S_i^y S_j^y + \Delta S_i^z S_j^z \right] =: \sum_{(i,j)} \tilde{B}_{ij} \,, \tag{4.27}$$

where the sum is over all pairs of nearest neighbours. After a sublattice rotation about the z-axis on the sites on one sublattice, each contribution is, in the standard S_z-basis, given by

$$\tilde{B}_{ij} = \frac{1}{4} \begin{pmatrix} \Delta & 0 & 0 & 0 \\ 0 & -\Delta & -2 & 0 \\ 0 & -2 & -\Delta & 0 \\ 0 & 0 & 0 & \Delta \end{pmatrix} , \tag{4.28}$$

which has the eigenvalues $\Delta/4$, $(2-\Delta)/4$ and $-(2+\Delta)/4$. The constant to add in order to get the projection operator (see equation 4.9) can thus be chosen to be $\Delta/4$. In fact, there is

more freedom. Bond operators to be used for the projection can be of the form

$$\begin{aligned} \boldsymbol{B}_{ij}(r) =& \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2\Delta & 2 & 0 \\ 0 & 2 & 2\Delta & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + r\, \mathbf{1}_4 \\ =& \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2\Delta + r & 0 & 0 \\ 0 & 0 & 2\Delta + r & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ &+ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r \end{pmatrix} \\ =& \boldsymbol{B}_{ij}^{\text{diag}} + \boldsymbol{B}_{ij}^{\text{flip}} + \boldsymbol{B}_{ij}^{\text{par}} \,, \end{aligned} \tag{4.29}$$

where $\mathbf{1}_4$ stands for the 4×4 identity matrix and I have also introduced the operators I call the diagonal, the flip and the parallel bond operators.

Acting on sites with parallel spins with a diagonal or a flip operator, destroys the state, while the parallel operator will return the state multiplied by r. A state with anti-parallel spins on the two sites is destroyed by the parallel operator. The diagonal operator returns states with anti-parallel spins multiplied with $2\Delta + r$; the flip operator flips the spins and multiplies the state by 2.

The parameter r shifts all eigenvalues of $\boldsymbol{B}_{ij}$. Since the quality of the projection is controlled by the fraction of the eigenvalues of the excited states and the groundstate, one wants to choose it as small as possible.

Just like in VBQMC, I make use of the fact that there exists a state that has equal overlap with all basis states in order to calculate the groundstate energy. I exploit this fact in the same way. In each measurement, I then calculate the numerical factors coming from the action of the Hamiltonian on the resulting state and add them. The average of these measurements is the estimate of the groundstate energy. If a term in the Hamiltonian hits sites with parallel spins, a value of Δ is added. If a term hits sites with anti-parallel spins, I add $-(2 + \Delta)$.

The Implementation

Depending on the value one chooses for r, there are two qualitatively different possibilities: $r = 0$ and $r \neq 0$.

Setting $r = 0$ When $r = 0$, acting with an operator on two sites with parallel spins destroys the state (leads to an operator string of zero weight). Such an operator string would certainly be rejected. This, of course, needs to be avoided. A naive implementation where one just

changes a number of operators in the string, propagates through it and then decides whether or not to accept, is not practical because already for short operator lengths the probability for this to occur is big.

It is possible to implement a branching algorithm (see subsection 4.1.1 and appendix A.2). While propagating through the operator string, at each node, we have to keep the number of operators that do *not* destroy the state and multiply the weights stemming from the bond operators themselves: 2Δ and 2 for the diagonal and the flip operators, respectively. The parallel operators do not have to be taken into account because they destroy all states if $r = 0$. If $\Delta = 1$ all weights are the same and only the branching weights need to be kept track of. Then, a Metropolis decision can be made. This approach is however not very efficient, because the weights can be very big and fluctuate substantially. The algorithm can be implemented very similarly to the driven worm algorithm, which is subject of the publication presented in this part of the thesis. Implementing a rejection free worm algorithm along the lines of the bouncing worm algorithm, which is also subject of the publication, should also be possible.

Setting $r \neq 0$ If one chooses $r \neq 0$, it is easy to set up a branching algorithm that lets one reach reasonable acceptance rates while updating the string at random positions.

We now first choose a bond and then, depending on whether or not we act on parallel spins, consider two different cases. If we act on parallel spins, the state is not changed and we collect the weight r. If we act on anti-parallel spins, I randomly choose between the diagonal or the flip operator. Which of the operators is chosen is, in the case of acting on anti-parallel spins, stored as part of the operator string and updated just like the bond the operator-string acts on. If we act with the diagonal operator, we collect a weight of $2 \times (2\Delta + r)$. If we act with the flip-operator, we collect a weight of 2×2. The factor of 2 in front of the weights is necessary because of branching.

I found it to be useful to update both the bond-position and the decision between diagonal and flip operator at the node that is updated. Per step I do this at two nodes. Then, the acceptance-rate is approximately 0.1. As the number of steps per sweep, I chose the fraction of the length of the operator-string and the number of operators I try to update. Between measurements I do as many sweeps as are necessary to on average have roughly as many operators successfully updated as there are operators in the string. For the calculations here this means that we do one measurement every four or five sweeps.

For the XXZ-model we find agreement with the result from the Bethe ansatz calculated by Erik S. Sørensen (see table 4.1).

Δ	Length	SPMC	Bethe ansatz
0.5	20	-7.536(5)	-7.534
0.5	100	-37.50(1)	-37.51
0.25	20	-6.935(4)	-6.934
0.25	100	-34.519(9)	-34.524

Table 4.1.: The estimates for the groundstate energy from SPMC agree with results from the Bethe ansatz.

Branching with a magnetic field

Other interactions can also be included quite easily. We show this with the example of the inclusion of a magnetic field acting in the z-direction:

$$H = H_{\mathrm{xxz}} + h \sum_i S_i^z \,. \tag{4.30}$$

The total S^z_{tot}-sector of which one finds the lowest energy is chosen with the trial state.

For the magnetic field term, it is surely impossible to increase the energy by more than $h/4$ times the length of the string. If we increase C in the definition of the projection operator of equation 4.9 by (at least) this number, we can be sure to create a valid projection operator. The projection operator then contains new operators D_i which act on each site (I call them site operators) and which for the purpose of the projection are treated just like the usual bond operators. Again we have the freedom to shift up with a unit matrix, which yields the family of site operators

$$D_i = \begin{pmatrix} q & 0 \\ 0 & 2h+q \end{pmatrix} \,. \tag{4.31}$$

Depending on whether or not the operator hits an up or a down spin the weight in the Monte Carlo will be q or $2h+q$, respectively. If $q=0$, it is, just as with $r=0$ above, very likely that operator strings of weight zero are generated. This leads to an inefficient projection and needs to be avoided.

Alternatively to introducing a new set of operators, one can also include the weights stemming from the magnetic field in the weights of the diagonal and parallel operators. Then q can be set to zero without causing any ill effects. With $q=0$, the modified operators are

$$B_{ij}^{\mathrm{diag}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2\Delta + r + 2h & 0 & 0 \\ 0 & 0 & 2\Delta + r & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.32}$$

and

$$B_{ij}^{\text{par}} = \begin{pmatrix} r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r+2h \end{pmatrix}. \tag{4.33}$$

SPMC, in its current form, cannot, for practical purposes, compete with mature state of the art methods. Should one succeed to find a working loop algorithm for it, this may not be the case anymore. Such an algorithm also seems to be necessary in order to effectively use double projection. The overlaps between different states appearing in equation 4.22 make an efficient implementation only possible if states without overlap are avoided. Valence bond states always have non zero overlap. Thus, they are not affected by this issue. In SPMC we deal with an orthonormal basis and this is an important problem.

Sadly, a loop algorithm that solves this problem is not known.

4.2. The paper

The work on the publication which is the topic of this chapter was started in the hopes of combining VBQMC and the notion of a worm (see section 1.4) travelling through the operator string. It is the natural extension of earlier work done by Erik S. Sørensen. In Alet and Sørensen [2003a] and Alet and Sørensen [2003b], Fabien Alet and Erik S. Sørensen published a successful attempt at creating a worm algorithm for the quantum rotor model. The publication of this part is about our work of trying to do the same for VBQMC.

The main questions we wanted to address were

- Is it possible to use a worm algorithm with VBQMC?
- How efficient can such an algorithm be?

The answers to these questions can be found in the third publication in this thesis.

Valence-Bond Quantum Monte Carlo Algorithms Defined on Trees

Andreas Deschner[1,*] and Erik S. Sørensen[1,†]

[1]*Department of Physics and Astronomy, McMaster University, Hamilton, Canada L8S 4M1*

We present a new class of algorithms for performing valence-bond quantum Monte Carlo of quantum spin models. Valence-bond quantum Monte Carlo is a projective T=0 Monte Carlo method based on sampling of a set of operator-strings that can be viewed as forming a tree-like structure. The algorithms presented here utilize the notion of a worm that moves up and down this tree and changes the associated operator-string. In quite general terms we derive a set of equations whose solutions correspond to a new class of algorithms. As specific examples of this class of algorithms we focus on two cases. The *bouncing worm* algorithm, for which updates are *always accepted* by allowing the worm to bounce up and down the tree and the *driven worm* algorithm, where a single parameter controls how far up the tree the worm reaches before turning around. The latter algorithm involves only a *single bounce* where the worm turns from going up the tree to going down. The presence of the control parameter necessitates the introduction of an acceptance probability for the update.

I. INTRODUCTION

Projective techniques are often used for determining the ground-state properties of strongly correlated models defined on a lattice. They were initially developed for non-lattice models [1] and then used for the study of fermionic lattice models [2]. They were subsequently applied to quantum spin models [3–8] as well as other models. The underlying idea is easy to describe. For a lattice Hamiltonian H, it is possible to choose a constant c such that the dominant eigenvalue E of $c\mathbb{1} - H$ corresponds to the ground-state wavefunction of H, $|\Psi_0\rangle$. We can then use $P = c\mathbb{1} - H$ as a projective operator in the sense that the repeated application of P to a trial wave function, $P^n|\Psi_T\rangle$, will approach $E^n|\Psi_0\rangle$ for large n. Hence, if n can be taken large enough, $|\Psi_0\rangle$ can be projected out in this manner provided that $\langle\Psi_0|\Psi_T\rangle \neq 0$. Some variants of this approach are often referred to as Green's functions Monte Carlo (GFMC) [2, 5–8]. Other projective operators such as $\exp(-\tau H)$ can be used depending on the model and its spectrum. For a review see Ref. 9–11. The convergence of such projective techniques may be non-trivial as can be shown by analyzing simple models [12]. If $P|\Psi_T\rangle$ can be evaluated exactly, this projective scheme is equivalent to the power method as used in exact diagonalization studies. As the number of sites in the lattice model is increased, exact evaluation quickly becomes impossible and Monte Carlo methods (projector Monte Carlo) have to be used.

The efficiency of the Monte Carlo sampling is crucial for the performance of implementations of the projective method and detailed knowledge of such Monte Carlo methods is of considerable importance. Here, we have investigated a new class of Monte Carlo algorithms for projective methods for lattice models. We discuss these algorithms within the context of quantum Monte Carlo where the projection is performed in the valence bond basis [3, 4, 13, 14], so called valence bond quantum Monte Carlo (VBQMC). The algorithms are, however, applicable to projective techniques in any basis.

VBQMC was first developed by Liang [3, 4] and then, starting fifteen years later, significantly further developed by Sandvik and collaborators [13, 14] and it is now widely used. Since its inception, VBQMC has been improved and generalized in several ways: it can be used on systems with spins with $S \neq 1/2$ [4] and states with total $S_z = 1/2$ [15]. An efficient sampling algorithm with loop updates is known for systems with $S = 1/2$ [14].

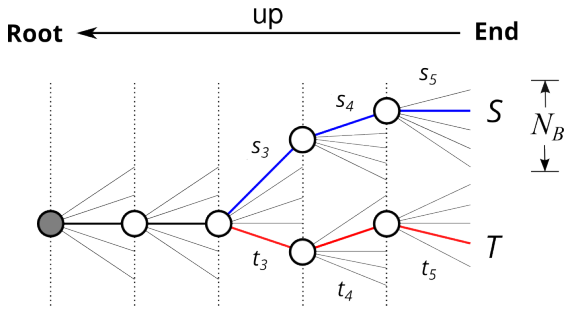


FIG. 1. The branching tree of length 5 for the selection of an operator string in a system with a Hamiltonian of N_B (here 5) terms. The operator that acts on the state first, is chosen at the first node on the left. This node is called the root and the direction towards the root we define to be up. The operator that acts on the state last is at the end of the string. The two colored paths differ in the choice of the last three operators. The last three branches, thus, contribute different operators and weights (s_i, t_i). The resulting strings S and T are different.

As outlined above, VBQMC works by projecting onto the ground-state by repeatedly acting on a trial-state $|\Psi_T\rangle$ with $P = c\mathbb{1} - H$, where the constant c is chosen such that the ground-state has the biggest eigenvalue. For Hamiltonians with bounded spectrum such a c can always be found. For a simple quantum spin model defined on a lattice we have

$$H = J \sum_{<i,j>} \mathbf{S}_i \cdot \mathbf{S}_j = \sum_{<i,j>} h_{ij} \qquad (1)$$

and we can write $P = c\mathbb{1} - H = \sum O_{ij}$ as a sum over N_B bond-operators O_{ij}. Taking P to the nth power then results in a sum over products of these bond-operators O_{ij}:

$$P^n = \sum_a \underbrace{O_{i(a,1)j(a,1)} \dots O_{i(a,n)j(a,n)}}_{n\text{-operators}} := \sum_a S_a \,. \qquad (2)$$

* deschna@mcmaster.ca
† sorensen@mcmaster.ca

Each instance of this product then forms a string S_a of bond-operators of length n. When selecting such a string of length n, one has to make a choice between the N_B bond-operators at each position in the string. It is possible to view the construction of such a string as a specific path in a decision-tree (see Fig. 1).

Although the algorithms we present can be extended to higher spin models, we shall restrict the discussion to quantum spin models with $S = 1/2$ where one usually takes $c = JN_B/4$. The action of the bond-operators then takes an attractively simple form.

In a valence bond basis state spins are paired into singlets. A specific pairing of all spins is usually referred to as a covering. All such coverings form an over complete basis for the singlet sub-space of the model. We shall only be concerned with models defined on a bi-partite lattice in which case a given valence bond covering, C, for a lattice with N spins can be denoted by listing all $N/2$ pairs of $[i, j]$ with i on sub-lattice A and j on sub-lattice B. Here, $[i, j] = (|\uparrow_{i_A}\downarrow_{j_B}\rangle - |\downarrow_{i_A}\uparrow_{j_B}\rangle)/\sqrt{2}$. We label the initial covering (trial-state) as C_0. The action of an operator O_{ij} can take two forms [3, 4, 13]:

- The sites i and j are in a singlet before the action of the operator. Then, the action of the operator does not change the state and we can associate a weight of $w = 1$:

$$O_{ij}[i, j] = 1[i, j]\,. \tag{3}$$

- The sites i and j are not in a singlet before the action of the operator. Then, after the action of the operator, the sites i and j form a singlet. The sites they were originally connected to are also returned in a singlet-state. Furthermore, the state is multiplied by a weight equal to $w = \frac{1}{2}$:

$$O_{ij}[i, k][l, j] = \frac{1}{2}[i, j][l, k]\,. \tag{4}$$

A particularly nice feature is that the application of any of the O_{ij} to any given covering yields a unique other covering and *not* a linear combination of coverings. Although convenient, this feature of projections in the valence bond basis is not strictly necessary for the algorithms we discuss here as they can be adapted to the case where a linear combination of states are generated [16]. For a given operator string $S_a = \prod_k O_k$, we can associate a weight given by $W_a = \prod w_k$. The state $S_a C_0$ will contribute to the final projected estimate of the ground-state with this weight. One can then sample the ground-state by performing a random walk in the space of all possible strings S_a [3, 4, 13] according to the weight W_a. This way of sampling is quite different from GFMC even though VBQMC and GFMC are closely related projective techniques. GFMC, as it is used in for instance Ref. 6, is usually performed in the S^z basis but can also be done in terms of the valence bond basis [17]. In GFMC the projection is done by stochastically evaluating the action of the whole projection-operator on a trial-state. This is done by introducing probabilistic "walkers". In contrast, as mentioned, in VBQMC a single state results and the strings S_a are sampled according to their weight. Clearly, the efficient sampling of states resulting from the stochastic projection of the trial-state is a difficult problem. Here, we propose to use worm (cluster) algorithms for this purpose.

In Monte Carlo calculations one averages over many configurations of the system which are generated with appropriate probabilities. Usually, this is done in a Markov-chain, where one configuration is chosen as a variation of the last. One important feature of an efficient algorithm is that these consecutive configurations are as uncorrelated as possible. This led to the introduction of algorithms where whole clusters and not just single elements are changed going from one configuration to the next [18, 19] or where all elements in the path of a *worm* are changed [20].

Here, we show how it is possible to adapt such worm algorithms for projections in the context of VBQMC. The algorithms we have studied are based on the notion of a worm moving around in the decision tree described above. As in earlier worm algorithms, the change of many elements is achieved by moving the worm based on local conditions [20–24] and one might refer to the algorithms as tree-worm algorithms. In general, the algorithms can be viewed as *directed* [23] algorithms.

When we update the string, we start with a worm at the end of the tree and move it up the tree. See Fig. 1. The worm then moves around in the tree and where it goes the operator-string is changed. When the worm finds its way back to the bottom of the tree the update is complete. We derive a set of simple equations governing the movement of the worm. The solution of these equations lead to parameters defining a new class of algorithms. Quite generally, many solutions are possible leaving significant room for choosing parameters that will lead to the most optimal algorithm.

We focus on two specific choices of parameters corresponding to two different algorithms. The *bouncing worm* algorithm, for which every update is accepted and the *driven worm* algorithm, for which the update is accepted with some probability. With the driven worm algorithm one can choose at will how much of the operator-string is on average changed in a successful update.

In order to test the algorithms, we calculate the ground-state energy of the isotropic Heisenberg-chain. This quantity is easy to calculate with VBQMC and can be exactly computed using the Bethe-ansatz. It is thus a very convenient quantity to test the algorithms with. The algorithms presented in this paper can, however, be used for the same calculations as other VBQMC implementations (see e.g. [13]).

In section II we derive the general equations governing the movement of the worm. Section III contains a description of the specific implementation corresponding to the two choices of parameter solutions we have studied. The *bouncing worm* is detailed in section III A while the *driven worm* algorithm is described in section III B. The algorithms are then compared in section IV. We present our conclusions in section V.

II. TREE ALGORITHMS

We now turn to a discussion of the general framework for the algorithms we have investigated. We begin by deriving the equations governing their behavior in a general way.

Let us take the Hamiltonian to have N_B terms. We now imagine a tree where each node indicates the decision to chose one of the N_B bond operators composing the string (see Fig. 1). Each branch of the tree corresponds to one of the N_B bond operators. A given operator string then corresponds to selecting a path in the tree. Consider 2 such paths S and T that are identical for the part of the operator string first applied to the trial-state. The last 3 operators, however, differ. This leads to different weights, which we denote with s_i and t_i.

As it is done in most Monte Carlo methods, we set out to construct a Markov-chain. Here it is a chain of different strings. If the probabilities to go from one string to the next have detailed balance, the Markov-chain contains the strings with the desired probability. For detailed balance, the probabilities for starting from operator string S and going to operator string T and reverse have to satisfy

$$\frac{P(S \to T)}{P(T \to S)} = \frac{t_3 t_4 t_5}{s_3 s_4 s_5} \,. \tag{5}$$

We can achieve this ratio of probabilities by imagining a worm (tree-worm) working its way up the tree to the point p where it turns around and then working its way down again.

Let us call the valence-bond covering of the trial-state C_0. Up to numerical factors, the application of an operator string S of length n will yield a new valence-bond covering $SC_0 \propto C_n$. The worm is started by removing the last applied bond operator and considering the resulting covering C_{n-1}. A decision now has to be made if the worm is to continue "up" the tree by removing more bond operators from the string or if it should instead go "down" the tree by adding a new bond operator to the string. At each node in the tree the decision to continue up or turn around is made according to a set of *conditional probabilities* $P(\mathrm{up}|s)$ and $P(t|s)$. Here, $P(\mathrm{up}|s)$ denotes the probability for going up after coming from a bond operator that carried weight s and $P(t|s)$ is the probability for turning around by applying a bond operator of weight t coming from an operator with weight s. Likewise, $P(s|\mathrm{up})$ denotes the probability of choosing an operator with weight s given that the worm is coming from further up the tree. With these conditional probabilities the left-hand side of Eq. (5) can be written as

$$\frac{P(S \to T)}{P(T \to S)} = \frac{P(t_5|\mathrm{up})P(t_4|\mathrm{up})P(t_3|s_3)P(\mathrm{up}|s_4)P(\mathrm{up}|s_5)}{P(\mathrm{up}|t_5)P(\mathrm{up}|t_4)P(s_3|t_3)P(s_4|\mathrm{up})P(s_5|\mathrm{up})} \,. \tag{6}$$

Clearly, Eq. (6) is satisfied if we choose

$$\frac{P(\mathrm{up}|s)}{P(s|\mathrm{up})} = \frac{c}{s} \qquad \text{and} \qquad \frac{P(t|s)}{P(s|t)} = \frac{t}{s} \,, \tag{7}$$

where c is an additional free parameter included for later optimization of the probabilities. If we can choose conditional probabilities with these properties, we can go between different operator strings always accepting the new string. The rejection probability is then zero. This is a very desirable property of any Monte Carlo Algorithm since it indicates that the algorithm is sampling. We mostly focus on so called *zero bounce* algorithms for which if the worm turns around the probability for replacing a bond operator with the same operator is zero. Then the two operator strings S and T are always different. This means that

$$P(s|s) = 0 \,. \tag{8}$$

Quite generally, it is easy to find *many* solutions to the equations (7) leading to many Monte Carlo algorithms which can be tuned for efficiency.

We now focus on S=1/2-Heisenberg models defined on bipartite lattices. As has been described above, for these models only 2 weights can occur: $1, 1/2$. The two weights correspond to the two different actions the bond-operators can have on the state. It is 1 if the operator acts on two sites that are in a valence-bond. The state is not altered under the action of such an operator. We call such operators diagonal. The weight is $1/2$ if the operator acts on two sites that are *not* in a valence-bond. After the action of the operator the two sites are connected by a bond as well as the sites they were connected to. We call such operators non-diagonal.

If a decision has to be made at the node at position m, the conditional probabilities depend on how many of the N_B bond-operators will yield a weight of 1 (are diagonal) or $\frac{1}{2}$ (are non-diagonal) when applied to the present covering C_{m-1}. We shall denote these numbers by N_1 and $N_{\frac{1}{2}}$ respectively. When the worm is started N_1 and $N_{\frac{1}{2}}$ therefore have to be calculated for C_{n-1}, if they are not already known from an earlier update. It is thus sensible to store N_1 or $N_{\frac{1}{2}}$ at all nodes. N_1 can only be zero at the node furthest up the tree (the root) and only if the trial-state is chosen to not contain any diagonal bonds. In Fig. 1 it is the gray node on the very left. $N_{\frac{1}{2}}$ cannot be smaller than $N_B/2$.

We can now write down an $(N_B+1)\times(N_B+1)$ matrix M of conditional probabilities for each node of the tree. The j'th column of the matrix describes the probability for going in any of the N_B+1 directions when coming from the direction j. For clarity we order the rows and columns such that the first $N_{\frac{1}{2}}$ correspond to the non-diagonal operators and the next N_1 to the diagonal operators. The last column contains the probabilities for going down the tree when coming from above; the last row the probabilities for going up the tree when coming from below. The remaining part of the matrix describe the probabilities for

replacing one operator with another when the worm turns from going up to going down. The matrix M has the form

$$M = \left(\begin{array}{ccccc|ccccc} \overbrace{} & & & & \overbrace{} & \overbrace{} & & & \overbrace{} & \overbrace{} \\ P(\frac{1}{2}|\frac{1}{2}) & P(\frac{1}{2}'|\frac{1}{2}) & P(\frac{1}{2}'|\frac{1}{2}) & \cdots & P(\frac{1}{2}'|\frac{1}{2}) & P(\frac{1}{2}|1) & P(\frac{1}{2}|1) & \cdots & P(\frac{1}{2}|1) & P(\frac{1}{2}|\mathrm{up}) \\ P(\frac{1}{2}'|\frac{1}{2}) & P(\frac{1}{2}|\frac{1}{2}) & P(\frac{1}{2}'|\frac{1}{2}) & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ P(\frac{1}{2}'|\frac{1}{2}) & P(\frac{1}{2}'|\frac{1}{2}) & P(\frac{1}{2}'|\frac{1}{2}) & \cdots & P(\frac{1}{2}|\frac{1}{2}) & P(\frac{1}{2}|1) & P(\frac{1}{2}|1) & \cdots & P(\frac{1}{2}|1) & P(\frac{1}{2}|\mathrm{up}) \\ \hline P(1|\frac{1}{2}) & P(1|\frac{1}{2}) & P(1|\frac{1}{2}) & \cdots & P(1|\frac{1}{2}) & P(1|1) & P(1'|1) & \cdots & P(1'|1) & P(1|\mathrm{up}) \\ \cdot & \cdot & \cdot & & P(1|\frac{1}{2}) & P(1'|1) & P(1|1) & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ P(1|\frac{1}{2}) & P(1|\frac{1}{2}) & P(1|\frac{1}{2}) & \cdots & P(1|\frac{1}{2}) & P(1'|1) & P(1'|1) & \cdots & P(1|1) & P(1|\mathrm{up}) \\ \hline P(\mathrm{up}|\frac{1}{2}) & P(\mathrm{up}|\frac{1}{2}) & P(\mathrm{up}|\frac{1}{2}) & \cdots & P(\mathrm{up}|\frac{1}{2}) & P(\mathrm{up}|1) & P(\mathrm{up}|1) & \cdots & P(\mathrm{up}|1) & P(\mathrm{up}|\mathrm{up}) \end{array}\right) , \quad \overbrace{}^{N_{\frac{1}{2}}} \quad \overbrace{}^{N_1} \quad \overbrace{}^{\mathrm{up}}$$

where $P(s'|s)$ refers to the conditional probability of coming from an operator with weight s and going to a *different* operator with the same weight. As mentioned above, $P(\mathrm{up}|s)$ denotes the probability for going up coming from an operator with weight s and $P(t|s)$ is the probability for turning around by choosing a bond operator of weight t coming from an operator with weight s. Likewise, $P(s|\mathrm{up})$ denotes the probability of choosing an operator with weight s coming from further up the tree.

To shorten the notation we introduce the short-hand

$$x = P(1/2'|1/2), \;\; y = P(1/2|1), \;\; z = P(1'|1) \; . \tag{9}$$

Furthermore we define the 'bounce' probabilities

$$b_{\frac{1}{2}} = P(1/2|1/2), \;\; b_1 = P(1|1), \;\; b_u = P(\mathrm{up}|\mathrm{up}) \; . \tag{10}$$

Here it is implied that the probabilities are for going from one operator to the *same* operator. Finally we also need to define the branching probabilities

$$u = P(1/2|\mathrm{up}), \;\; w = P(1|\mathrm{up}) \; , \tag{11}$$

from which it follows (using Eq. (7)) that:

$$2cu = P(\mathrm{up}|1/2), \;\; cw = P(\mathrm{up}|1) \; . \tag{12}$$

The matrix M is then given by

$$= \begin{pmatrix} b_{\frac{1}{2}} & x & x & \cdots & x & y & y & \cdots & y & u \\ x & b_{\frac{1}{2}} & x & \cdots & x & y & y & \cdots & y & u \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ x & \cdot & \cdot & \cdots & b_{\frac{1}{2}} & y & y & \cdots & y & u \\ 2y & 2y & 2y & \cdots & 2y & b_1 & z & \cdots & z & w \\ \cdot & \cdot & \cdot & & 2y & z & b_1 & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ 2y & 2y & 2y & \cdots & 2y & z & z & \cdots & b_1 & w \\ 2cu & 2cu & 2cu & \cdots & 2cu & cw & cw & \cdots & cw & b_u \end{pmatrix} \; . \tag{13}$$

The requirement that this matrix be stochastic (i.e. some branch is chosen with probability one) means that the entries in each column have to sum to 1. This leads to the set of equations

$$\begin{aligned} 1 &= N_{\frac{1}{2}} u + N_1 w + b_u \\ 1 &= N_{\frac{1}{2}} y + (N_1 - 1) z + cw + b_1 \\ 1 &= (N_{\frac{1}{2}} - 1) x + N_1 2y + 2cu + b_{\frac{1}{2}} \; . \end{aligned} \tag{14}$$

These simple equations are the central equations governing the behavior of the algorithms. To find an algorithm, we need to solve these 3 equations with the constraints that $0 \leq x, y, z, b_{\frac{1}{2}}, b_1, b_u, u, w \leq 1$; a straight forward problem.

At the root, the equations are modified slightly: since it is not possible to go further up the tree, $2cu$, cw, b_u are not meaningful and can be set to zero. For convenience we set $b_1 = x$ and $b_2 = z$ at the root. This allows one to just choose diagonal operators twice as often as non-diagonal operators. Since the number of diagonal operators does not change at the root, a table generated at the beginning of the calculation suffices to perform this task.

It can be very useful to choose different c's at different nodes. Then, calculating the probabilities to choose operators according to the rules introduced in this section will not lead to an algorithm with detailed balance, because c_i from different strings will not cancel in Eq. (6). It is necessary to work with an acceptance probability. We find

$$\frac{P(\mathcal{S} \to \mathcal{T})}{P(\mathcal{T} \to \mathcal{S})} = \frac{t_3 t_4 t_5}{s_3 s_4 s_5} \frac{c_4^{\mathcal{S}} c_5^{\mathcal{S}}}{c_4^{\mathcal{T}} c_5^{\mathcal{T}}} \frac{P_{\mathrm{acc}}(\mathcal{S} \to \mathcal{T})}{P_{\mathrm{acc}}(\mathcal{T} \to \mathcal{S})} \; . \tag{15}$$

Here $c_i^{\mathcal{S}}$ and $c_i^{\mathcal{T}}$ denote c at the different nodes in the strings $\mathcal{S}$ and $\mathcal{T}$, respectively. To validate the algorithm, we must therefore introduce an acceptance probability that must cancel the factor $(c_4^{\mathcal{S}} c_5^{\mathcal{S}})/(c_4^{\mathcal{T}} c_5^{\mathcal{T}})$. This can be achieved by choosing

$$P_{acc}(\mathcal{S} \to \mathcal{T}) = \min\left(1, \frac{c_4^{\mathcal{T}} c_5^{\mathcal{T}}}{c_4^{\mathcal{S}} c_5^{\mathcal{S}}}\right) , \tag{16}$$

meaning that when a new string is generated through a worm move it is accepted with this probability.

Since we always start from the bottom of the tree (the last operator applied), the worm algorithms presented in this paper always change a block of consecutive branches at the end of the string. This is favorable to changes across the whole string because changes far up the string might be undone by changes closer to the end of the string [13]. In this way the most important part of the string is updated most substantially.

It is also important to note that the algorithm will conserve certain topological numbers. For instance, for a two-dimensional system $S = 1/2$ Heisenberg model the number of valence bond crossing a cut in the x- or y-direction is either odd or even. Hence, the initial covering, C_0 is characterized by these 2 parities. It is easy to see that the application of P to any covering can not change these parities and they are therefore preserved under the projection.

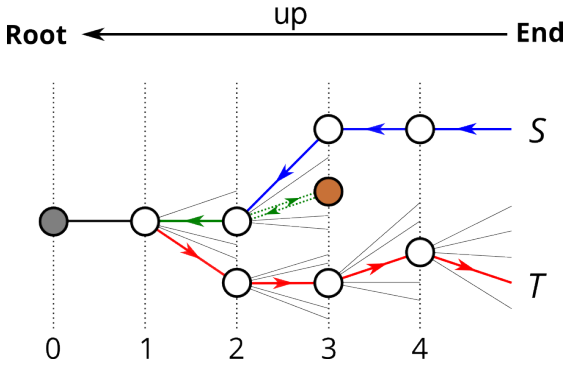


FIG. 2. A possible path that contains one bounce and connects the string S and the string T. The worm first goes up to the node 2 where it turns to go down to node 3. The worm bounces back and goes all the way to node 1. Then the worm turns around and does not bounce again.

III. IMPLEMENTATIONS OF TREE WORM ALGORITHMS

As is explained in the last section, many different algorithms can be found because many different solutions to the equations (14) exist.

In this section we present two different algorithms. One pure worm algorithm where every update is accepted (the *bouncing worm* algorithm) and an algorithm that allows for control over how far in the tree updates are attempted (the *driven worm* algorithm). To test and compare the different algorithms, we calculate the ground-state energy of the antiferromagnetic Heisenberg chain.

The Néel-state $|\text{Néel}\rangle$ has equal overlap with all valence-bond states. This can be used to very directly estimate the ground-state energy, E_0 [13]:

$$\begin{aligned} E_0 &= \frac{\langle \text{Néel}|H|\Psi_0\rangle}{\langle \text{Néel}|\Psi_0\rangle} \\ &= \lim_{n\to\infty} \frac{\langle \text{Néel}|HP^n|C_0\rangle}{\langle \text{Néel}|P^n|C_0\rangle} \\ &= \lim_{n\to\infty} \sum_{a=1}^{N_B^n} \frac{\langle \text{Néel}|HS_a|C_0\rangle}{\sum_{a=1}^{N_B^n}\langle \text{Néel}|S_a|C_0\rangle} \\ &= \lim_{n\to\infty} \sum_{a=1}^{N_B^n} \frac{W_a}{\sum_{a=1}^{N_B^n} W_a} \frac{\langle \text{Néel}|H|C_a\rangle}{\langle \text{Néel}|C_a\rangle}. \end{aligned} \tag{17}$$

If we take $E_a C_b = HC_a$ and assume that the Monte Carlo sampling will visit strings according to their weight W_a, then for a Monte Carlo sequence of length N of independent strings we find:

$$E_0 = \frac{1}{N}\sum_{a=1}^{N} E_a, \tag{18}$$

where again we have used the fact that $\langle \text{Néel}|C\rangle$ is independent of the covering C.

To analyze the correlation-properties of the worm algorithms we use the energy-autocorrelation-time, which we take to be the number of updates it takes the energy-autocorrelation-function

$$A_E(t) = \frac{\langle E_i E_{i+t}\rangle - \langle E\rangle^2}{\langle E^2\rangle - \langle E\rangle^2} \tag{19}$$

to decay to 0.1. The results of all update-attempts enter the calculation of the expectation-values. The shorter the autocorrelation-time is, the fewer steps have to be done between consecutive measurements.

If not stated otherwise an operator-string of 20,000 operators was used for calculations with worm algorithms.

A. The bouncing worm algorithm

The first algorithm we discuss is the bouncing worm algorithm. Only a few of the variables that appear in the equations (14) are chosen to be non-zero. We choose to set:

$$x = y = 0, \quad b_1 = b_{1/2} = 0, \tag{20}$$

while $z = P(1'|1) \neq 0$ as is u, w. We leave b_u as a parameter that can be zero or non-zero allowing for tuning of the algorithm. With this choice, when the worm is moving up the tree the only possibility for it to turn around is by opting to replace one diagonal operator with another diagonal operator. The c_i are chosen to be the same at all nodes: $c_i = c$.

The equations for the non-zero parameters are then

$$\begin{aligned} u &= \frac{1}{2c} \\ w &= \frac{1}{N_1}\left(1 - \frac{N_{\frac{1}{2}}}{2c} - b_u\right) \\ z &= \frac{1 - cw}{N_1 - 1}\,. \end{aligned} \tag{21}$$

The requirements that $z, w > 0$ imply that

$$\frac{N_{\frac{1}{2}}}{2(1-b_u)} \leq c \leq \frac{N_B + N_1}{2(1-b_u)}\,. \tag{22}$$

To satisfy Eq. (22) with node-independent c, we set

$$c = \frac{N_B}{2(1-b_u)} \,. \tag{23}$$

With this choice of parameters, we find the probability to go up the tree if the worm is at a node with a non-diagonal operator to be

$$P(\text{up}|1/2) = 2cu = 1 \,, \tag{24}$$

for *any* b_u. Likewise, if the worm is at a node with a diagonal operator the probability to go up is given by

$$P(\text{up}|1) = cw = 1/2 \,, \tag{25}$$

independent of b_u. The probability for going up the tree is therefore independent of b_u.

We define the penetration depth (p.-depth), which we denote by r, as the maximal height that the worm reaches. The actual length of the worm is denoted by l and with $b_u = 0$ we find $l = 2r$. The penetration depth r will determine how much the operator-string is changed. Obviously, it is desirable to have the worm reach as far up the tree as possible. It is possible to force the worm farther up the tree by having it bounce back to going up after it has turned to go down (see Fig. 2). In that case, the actual *length* of the worm, l, will then be substantially different from twice the penetration depth since the worm can turn many times, a point we shall return to later. Such bounces occurs with a likelihood of b_u which was left as a free parameter and can now be used as a tuning parameter.

The algorithm is straight forward to implement and the acceptance probability for a worm update is 1. The move is *always* accepted. Specific details of an implementation of the bouncing worm update can be found in appendix A 1.

We begin by discussing the case of $b_u = 0$. In this case the worm first moves up the operator string, turns around once and then proceeds down to the bottom of the tree. It does not go back up the operator string since $b_u = 0$. In order to measure the performance of the algorithm we did calculations on an antiferromagnetic Heisenberg chain with 50 sites using an operator string of length $100,000$. As can be seen in table I, this leads to a rather small mean penetration-depth (p.-depth) of about 5. The maximal penetration-depth of 50 is substantially larger. Both these numbers are, however, substantially smaller than the length of the operator string $(100,000)$ and it appears that the algorithm with b_u is not very effective.

We now turn to the case $b_u \neq 0$. In this case the worm can now switch directions many times during construction (see Fig. 2). The results for the mean and maximal penetration-depth are also listed in table I. As b_u is increased from zero, the maximal penetration-depth first increases very slowly until about $b_u = 0.25$. It then grows dramatically and, not surprisingly, reaches the length of the operator string. This occurs at $b_u \approx 0.2790$. At the same time the mean penetration-depth only increases by a factor of roughly 4, from 5 to about 20. For bounce-probabilities bigger than $b_u \approx 1/4$ the program is slowed down significantly compared to the algorithm with $b_u = 0$ as indicated in the last column in table I. Thus, even though a large penetration-depth is desirable the computational cost can become so big that increasing b_u might not be worthwhile.

b_u	mean p.-depth	max p.-depth	slowdown
0.0000	4.561(4)	50	1
0.2500	7.38(1)	305	1.7
0.2750	10.44(3)	1,465	9.7
0.2789	15.64(9)	41,010	316.7
0.2790	19.7(5)	>100,000	<5,535.7

TABLE I. Data for several runs at different b_u. At $b_u \approx 0.25$ increasing the bounce-probability starts to significantly slow down the algorithm. The last column contains the run-times divided by the runtime for $b_u = 0$. The data were generated with an operator-string of 100,000 operators. The maximal penetration and the expected slowdown could thus not be resolved for $b_u = 0.2790$. We used 10^6 measurements and a chain with 50 sites.

In contrast to the maximal penetration-depth, the mean penetration-depth grows very slowly for the values of b_u we have been able to study. For computations of reasonable computational cost it never reaches the size of the system and thus also not the length of the operator-string which has to be chosen to be several times the size of the system. That only a small part of the string is updated regularly is directly reflected in the energy-autocorrelation-time (see Fig. 3). The number of bonds that can change in one update of the worm-calculations is twice the penetration-depth. Typical updates never reach far into the operator-string. Thus, the bigger the system is, the less it is perturbed by the update and the more correlated are the energies measured after consecutive updates. As shown in Fig. 3,

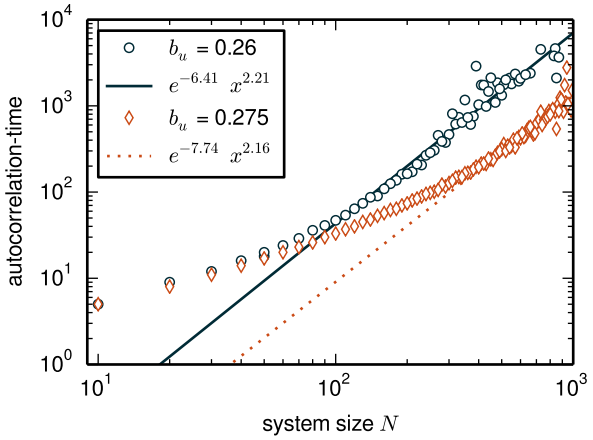


FIG. 3. The autocorrelation-time of the energy as a function of system size N. The lines indicate tentative power-law fits to the data at large N with a power of 2.21 for $b_u = 0.26$ and 2.16 for $b_u = 0.275$. Operator-strings of 20,000 operators were used.

increasing b_u decreases the autocorrelation-time. However, for large system sizes the overall scaling of the autocorrelation-time with the system size appears independent of b_u. At $b_u = 0.26$ we find a power-law with an exponent of 2.21 while a slightly larger $b_u = 0.275$ yields a power of 2.16.

Even though the mean penetration depth remains small, one

can still obtain high quality results. In particular, it is *not* necessary for the mean penetration depth to reach a value close to the length of the operator-string (the projection power) in order to get reliable results. Since the maximal penetration-depth is substantially larger than the mean, the operator-string is often updated deeper than the mean penetration-depth. Hence, the mean penetration-depth can be much smaller than the length of the operator-string has to be for otherwise equivalent calculations with conventional VBQMC. We discuss this effect in more detail in section IV.

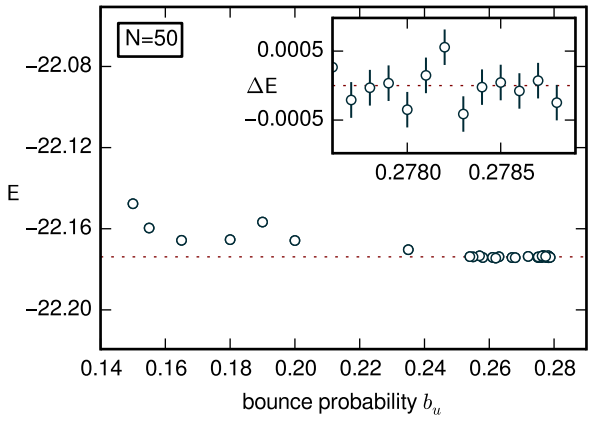


FIG. 4. With a bigger probability to bounce b_u, the bouncing worm algorithm yields a better approximation of the ground-state energy. The ground-state energy was calculated using the Bethe-ansatz and is indicated by a dotted line. Since some operators in the string are never updated, calculations with different b_u were effectively done with different trial-states. The calculation was done for a chain with 50 sites. Operator-strings of 20,000 operators were used.

As one increases the probability to bounce up the tree a longer part of the string actually partakes in the projection. Thus, also the quality of the projection is better (see Fig. 4). If operators are never updated, they do not contribute to the projection; they do however modify the trial-state used in the projection. This leads to the irregularly scattered pattern the data shows for small b_u. In this sense one can think of rare updates that go high up the tree as effectively changing the trial-state and the whole calculation as an averaging over these trial-states.

As is evident from Fig. 4, the bouncing worm algorithm yields good results for $b_u > 0.25$. It is an attractively simple algorithm with zero bounce probability and a probability of 1 for accepting a new string. The autocorrelation-time can be reduced by increasing b_u whereas the overall power of the growth at large system sizes appears independent of b_u. However, the increased computational cost associated with increasing b_u is considerable and we have therefore investigated another parameter choice leading to a different algorithm, the driven worm algorithm. We now turn to a discussion of this algorithm.

B. The driven worm algorithm

Clearly, it is desirable to have all updates result in a substantial change of the operator-string. Then, fewer updates have to be performed. For the problem at hand, this means that we need the worm to go far up the tree as often as possible without increasing the computational cost too much. Direct control over the associated probability would be very convenient. We achieve this by setting the probabilities to go up the tree to be

$$2cu = cw = \alpha\,. \tag{26}$$

The value of α is the probability to, at each node, decide to go up the tree. Since u and w depend on N_1 and N_2, this is only possible by allowing c to vary with the node. As explained at the end of section II, the acceptance step of Eq. (16) thus has to be introduced. Updated strings may be rejected.

We set all bounce-probabilities to be zero, $b_1 = b_{1/2} = b_u = 0$. Hence, the worm will move up the tree and then turn around once. To get a working algorithm, we have to find solutions to the equations (14) which will determine the transition-probabilities (see Eq. (13)). If $N_{\frac{1}{2}}, N_1 > 1$ the solutions to equations (14) are given by:

$$\begin{aligned} x =& \frac{1}{N_{\frac{1}{2}}(N_{\frac{1}{2}}-1)}\Big[2N_1(N_1-1)\,z \\ &+ (1-\alpha)(N_{\frac{1}{2}}-2N_1)\Big]\,, \\ y =& \frac{1}{N_{\frac{1}{2}}}\Big[(1-N_1)z+1-\alpha\Big]\,, \\ u =& 1/(N_{\frac{1}{2}}+2N_1)\,, \\ w =& 2u\,, \end{aligned} \tag{27}$$

where

$$\frac{1-\alpha}{N_1-1}\Big[1-\frac{N_{\frac{1}{2}}}{2N_1}\Big] \le z \le \frac{1-\alpha}{N_1-1}\,. \tag{28}$$

If $N_1 \neq 1$, we set

$$z = \frac{1-\alpha}{N_1-1}\Big[1-\frac{1}{2}\frac{N_{\frac{1}{2}}}{2N_1}\Big]\,, \tag{29}$$

if it results in $z > 0$ or

$$z = \frac{1}{2}\frac{1-\alpha}{N_1-1} \tag{30}$$

otherwise. In this way Eq. (28) is always satisfied and $z \geq 0$. If $N_1 = 1$ we set $z = 0$. Finally, we note that the worm update in this case has to be accepted/rejected according to the probability Eq. (16). Specific details of an implementation of this driven worm algorithm can be found in appendix A 2.

How far up the tree updates are attempted can in this case easily be calculated. The probability for the worm to have length l and turn around after going up $r = l/2$ nodes is given by $P(r) = \alpha^r(1-\alpha)$. The expectation-value of r is given by

$$\langle r\rangle = \frac{1}{1-\alpha}\,. \tag{31}$$

The probability distribution for the worm to penetrate the tree r nodes deep during a computation of m updates, is given by

$$P_{m,\alpha}(r_{\max}) = \underbrace{\left(1-(1-\alpha)\sum_{q=r}^{\infty}\alpha^q\right)^m}_{\substack{\text{probability that in } m \text{ attempts } \textit{no} \text{ worm} \\ \text{turns at a node with } r > r_{\max}}} - \underbrace{\left((1-\alpha)\sum_{q=1}^{r-1}\alpha^q\right)^m}_{\substack{\text{probability that in } m \text{ attempts } \textit{all} \text{ worms} \\ \text{turn at a node with } r < r_{\max}}}$$

$$=(1-\alpha^r)^m-(1-\alpha^{r-1})^m\ . \qquad (32)$$

How far up the tree is updated, is *not* given by how far the worm goes up the tree since the update might be rejected. The mean penetration-depth is therefore not equal to $\langle r\rangle$. In Fig. 5 we show results for the mean and maximal penetration-depth for two different system sizes, $N = 50, 1000$ as a function of $1/(1-\alpha)$. As expected, both the mean and maximal penetration-depth increase monotonically with $1/(1-\alpha)$.

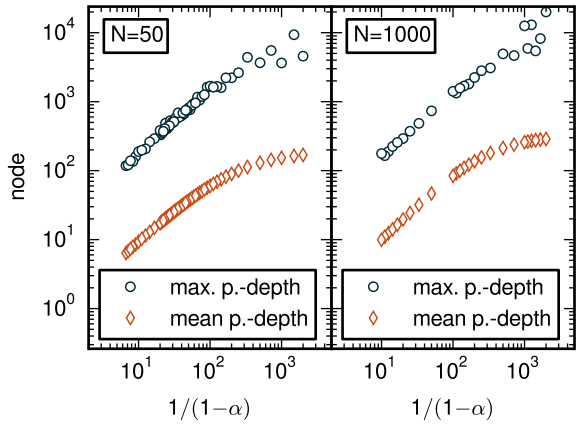


FIG. 5. The mean and the maximal penetration-depths grow as α approaches one. The mean penetration-depth is always smaller than $\langle r\rangle = 1/(1-\alpha)$. This behavior is independent of the system size. For the chain with 50 sites 3×10^8 and for the chain with 1000 sites 10^7 updates were performed. Operator-strings of 20,000 operators were used.

Another measure of the performance of the algorithm can be established by simply looking at the calculated ground-state energy and its error. This is done in Fig. 6 where the ground-state energy is shown as a function of $1/(1-\alpha)$. Operators that are never updated, only change the effective trial-state the ground-state is projected out of. By forcing the worm further up the tree, one can have a bigger part of the operator-string partake in the projection (see Fig. 5). This leads to a better approximation of the ground-state energy as can be seen in Fig. 6.

For the driven worm algorithm we have also studied the behavior of the autocorrelation-time of the energy. Our results

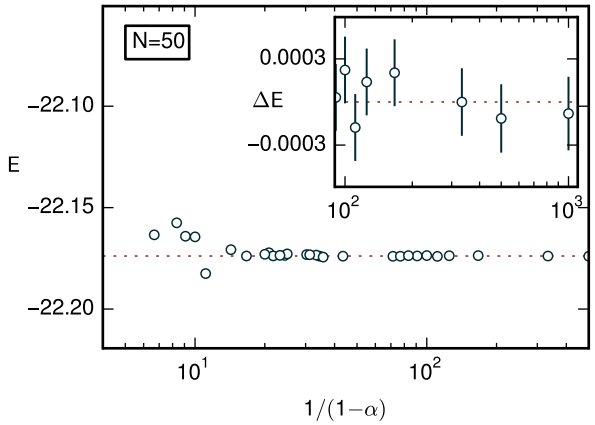


FIG. 6. The bigger the penetration probability α, the better the approximation of the ground-state energy given by the driven worm algorithm. The ground-state energy was calculated using the Bethe-ansatz. It is indicated by a dotted line. Since some operators in the string are never updated, calculations with different α were effectively done with different trial-states. The calculation was done for a chain with 50 sites. Operator-strings of 20,000 operators were used.

are shown in Fig. 7 as a function of $1/(1-\alpha)$. The behavior is in this case not monotonic. At first the autocorrelation time decreases, but then it starts to *grow* at larger $1/(1-\alpha)$.

This can be understood in the following way: As long as $\langle r\rangle$ is much smaller than the size of the system, the autocorrelation-time *decreases* with increasing α. This follows naturally from the fact that increasing $1/(1-\alpha)$ will increase $\langle r\rangle$ and therefore lead to larger and more effective updates. This decreases the correlations between operator-strings. The farther the worm travels up the string, the smaller is the probability that an update is accepted (see table II). For bigger α, and thus also $\langle r\rangle$, this effect dominates and the autocorrelation-time grows. A characteristic minimum in the autocorrelation-time as $1/(1-\alpha)$ is increased can then be identified as is clearly evident in Fig. 7.

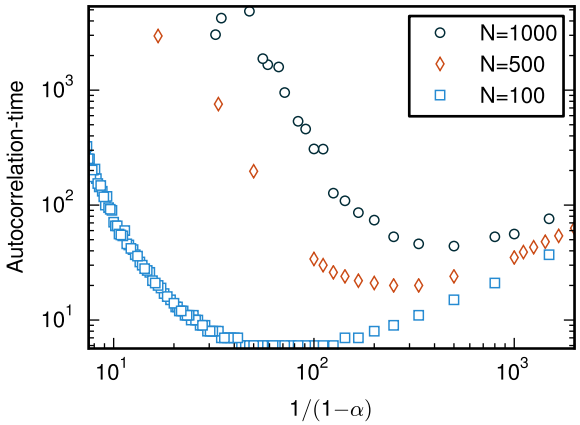


FIG. 7. The autocorrelation-time decreases with $\langle r\rangle$, if typical updates are smaller than the system size. Since bigger $\langle r\rangle$ means better projection, this implies that the autocorrelation-time *decreases* as the quality of the projection is improved. Operator-strings of 20,000 operators were used.

$1/(1-\alpha)$	acceptance-rate
200	0.41
1000	0.13
5000	0.03

TABLE II. The acceptance rate drops when the string is updated more substantially. The calculations were done for a chain with 50 sites.

IV. COMPARISON OF ALGORITHMS

In the following we compare worm-updates to simple conventional VBQMC-updates as described for example in reference [13]. This means that for VBQMC we attempt to change 4 randomly selected operators during one update. We do not compare to loop-updates as introduced in reference [14], since we anticipate the worm algorithms to be of most with algorithms for which loop-updates are not known although our current implementations of them are similar to conventional VBQMC.

We first consider the convergence of the energy with the projection power (the length of the operator strings). Our results for a 50 site Heisenberg chain are shown in Fig. 8. It turns out that if the worm algorithms penetrate the tree sufficiently deeply, the results do not depend on the type of algorithm in use. In particular, the dependence of the results on the length of the string is the same for *all three* algorithms (see Fig. 8), just as one might have expected since the power method underlies all three algorithms.

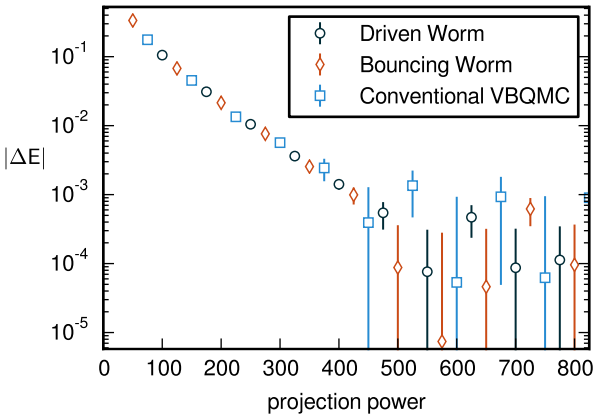


FIG. 8. Upon increasing the quality of the projection by using longer operator strings, the estimate of the ground-state energy converges to the correct value in the same way for worm- and conventional VBQMC-algorithm as long as the string is penetrated sufficiently deeply. The driven worm algorithm was run with the probability to go up the tree $\alpha = 0.995$, which corresponds to a mean penetration-depth of about 90 and full penetration of the string. The bouncing worm algorithm was run with a bounce probability $b_u = 0.2675$, which corresponds to a mean penetration-depth of roughly 8 and full penetration of the string. The calculation was done for a chain with 50 sites.

When using the worm algorithms, the operator string is usually chosen so long that the worm never or very rarely reaches the root of the tree. This means that there are almost always nodes close to the root with operators that are never updated and thus act on the trial-state after every update. In this way, we are effectively using an optimized trial-state. The effect is similar to generating the trial-state by performing several updates on a randomly chosen trial-state and taking the resulting state for the actual calculation. We used such a trial-state for the conventional VBQMC-calculations shown in this section.

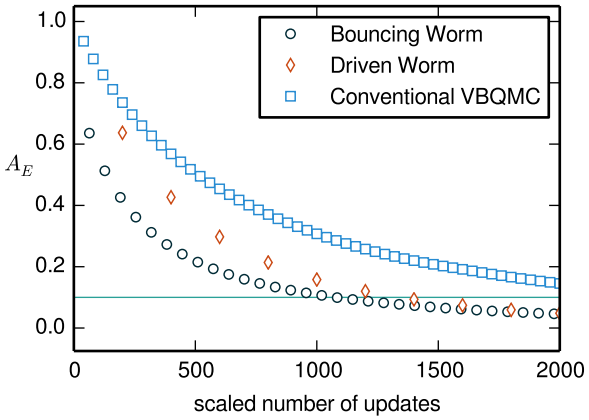


FIG. 9. The autocorrelation-function versus scaled number of updates for the two worm algorithms and conventional VBQMC. The number of updates was scaled by the number of operators attempted to be changed in an update, $\langle l/2 \rangle$. Hence, data for conventional VBQMC updates, the driven worm algorithm and the bouncing worm algorithm were multiplied by 4, 200 and 63.758, respectively. For the worm algorithms the same parameters as in Fig. 8 were used. This means that $\alpha = 0.995$ and $b_u = 0.2675$. An operator-string of length 1000 was used for all three algorithms. The horizontal line at 0.1 was added to allow for easy visual estimation of the scaled autocorrelation-time.

A useful measure of the effectiveness of an algorithm can be obtained from the autocorrelation function. If simply measured as a function of the number of updates it decreases dramatically faster for the worm algorithms when compared to conventional VBQMC. However, just using one update as the temporal unit puts conventional VBQMC at an unfair disadvantage. The reason is, that in calculations with conventional VBQMC one attempts to change 4 operators per update while for the worm algorithms it could be many more. The number of updated operators in a given worm update varies greatly with the length of the worm, $\langle l \rangle$, which can easily be hundreds of operators long. Since a single worm update is, therefore, computationally more expensive to perform than a single 4 operator update with conventional VBQMC, it seems fairer to compare autocorrelation functions with this difference taken into account. That is, a fair comparison would ask which algorithm has the smallest correlations when on average the same number of changes has been attempted. We can take this into account by simply scaling the temporal axis with the average size of the attempted update.

In Fig. 9 we therefore show results for the energy autocorrelation function for the two worm algorithms as well as for conventional VBQMC with the temporal axis rescaled by the number of operators one attempts to change in a single update. During one update with worm-algorithms one tries to update $l/2$ operators. The scaled number of updates is simply $\#\text{updates} \times \langle l/2 \rangle$ with $\langle l/2 \rangle = 4$ for conventional VBQMC

and $\langle l/2\rangle = \langle r\rangle = 1/(1-\alpha)$ for the driven worm algorithm. For the bouncing worm algorithm, $\langle l\rangle$ has to be measured during the simulation, since the bouncing worm can go up and down the tree many times. Thus, $\langle l/2\rangle$ can be orders of magnitudes bigger than the mean penetration-depth. For instance, for the calculations shown in Fig. 9 the mean penetration-depth was approximately 7.8 whereas $\langle l/2\rangle = 63.758$. Even including such a rescaling of the temporal axis, it is clear that the autocorrelation-times are much shorter for the worm algorithms, as shown in Fig. 9.

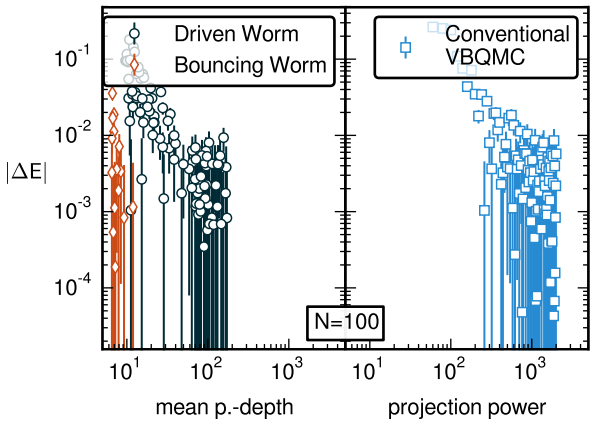


FIG. 10. The deviation, $|\Delta E|$, from the exact Bethe-ansatz results for a chain with 100 sites. The results are shown for the driven and bouncing worm algorithms versus the mean penetration-depth and for conventional VBQMC updates versus the projective power (length of operator string). The worm algorithms reach the same small value of $|\Delta E|$ with a mean penetration-depth an order of magnitude smaller than the projective power used for the calculation with conventional VBQMC updates. The bars on the markers indicated the statistical uncertainty. The colored (dark) surfaces are due to overlapping error-bars. Operator-strings of 20,000 operators were used for the calculations with the worm algorithms.

The two worm algorithms change operators of the string starting from one end while the conventional VBQMC selects 4 operators at random to be changed. As mentioned in subsection III A, the mean and the maximum penetration-depth are usually much smaller than the length of the operator-string (the projection power). It is therefore natural to ask if one can reach a similar quality of results using worm algorithms and conventional VBQMC.

That this is so can be seen by plotting the absolute deviation from the ground-state energy, $|\Delta E|$, versus the mean penetration-depth. As shown in Fig. 10, the mean penetration-depth can, in fact, be much smaller than the projection power of a conventional VBQMC-calculation and still yield results of the same accuracy.

Finally, we look at how the scaled autocorrelation-time depends on the size of the system studied. For convenience, we define the scaled autocorrelation-time to be the point where the autocorrelation function has decreased to the value 0.1 (see Fig. 9). Since in realistic calculations one would use a fixed (large) length of operator string with the worm algorithms, while one would scale it with the size of the system in conventional VBQMC, we here only compare the two worm algorithms. Our results are shown in Fig. 11 for a fixed length operator string of $20,000$.

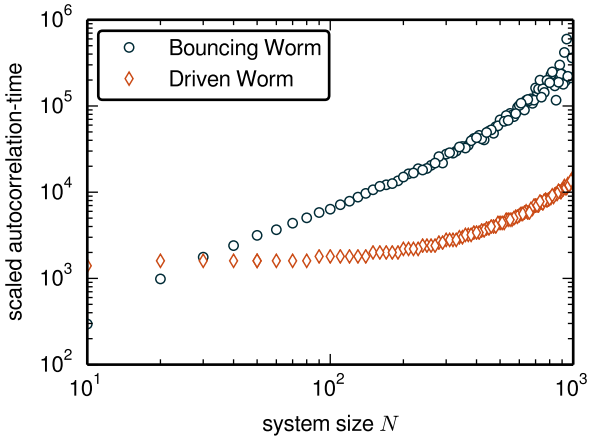


FIG. 11. The scaled autocorrelation-time for the driven and bouncing worm algorithms as a function of system size. A fixed operator string of length $20,000$ was used in the calculations. In the calculations shown, the bouncing worm algorithm was run with $b_u = 0.275$ and the driven worm algorithm was run with $\alpha = 0.995$. For the scaling we use $\langle l/2\rangle = 200$ for the driven worm algorithm and an $\langle l/2\rangle$ between 60 for $N = 10$ and 221 for $N = 1000$ for the bouncing worm algorithm.

For the simulations shown in Fig. 11 the mean penetration-depths for the driven worm algorithm were about 100. The autocorrelation-time for the driven worm algorithm starts to increase appreciably at this system size, while it is initially are almost flat. We conclude that a significant increase in the autocorrelation-time appears once the system size significantly exceeds the mean penetration-depth. A similar effect can be observed for the bouncing worm algorithm. The mean penetration-depths for the bouncing worm algorithm are, however, much smaller (around 9; see Fig. 10). Results for N smaller than the mean penetration-depth are therefore not shown in Fig. 11. The autocorrelation-times remain manageable for the system sizes studied, even though it is consistently increasing.

Compared to simple implementations of VBQMC, the worm algorithms have significant overhead. This is largely compensated by the large number of operators that can be changed in an update and resulting shorter autocorrelation-times, as we found in all computations. Given the somewhat different properties of the two worm algorithms, a realistic implementation could combine the two by performing updates with the driven worm algorithm mixed with updates using the bouncing worm algorithm (and perhaps conventional VBQMC updates).

V. CONCLUSION

We have shown that valence-bond quantum Monte Carlo can be implemented with an update build around the notion of a worm propagating through a tree. Many different such algorithms are possible. We studied the validity and efficiency of

two of them. One for which no update is rejected (the bouncing worm algorithm) and one for which big parts of the operator-string are updated (the driven worm algorithm). Both algorithms are attractively simple and straight forward to implement and produce high quality results.

While they may not be computationally competitive with state of the art loop update algorithms [14] for VBQMC, the algorithms presented here are intrinsically interesting since they represent a new class of algorithms that should be generally applicable to projective methods. These algorithms are not restricted to the valence bond basis and preliminary results show that they can be quite efficient in the S^z-basis [16] method and might spark further development of it. We also note that many other algorithms can easily be found with the results contained in this paper and that it is possible that the parameter space allows for much more efficient algorithms than the two we have studied here.

In terms of further optimizing the algorithms several directions may be interesting to pursue. Not updating some of the operators the worm visits, might boost the acceptance ratio of the driven worm algorithms and thereby reduce the autocorrelation-times. This could be combined with attempting to reduce the overhead of the driven worm calculations by always forcing the worm all the way down to the root. This would eliminate the need to keep track of the state at each node. With the current practice of updating all operators after turning around, going all the way to the node during every update leads to very small acceptance ratios.

We acknowledge computing time at the Shared Hierarchical Academic Research Computing Network (SHARCNET:www.sharcnet.ca) and research support from NSERC.

Appendix A: Pseudocode for implementations of the tree worm algorithm

This appendix contains pseudocode that shall serve to clarify the algorithms proposed in this paper. To simplify notation we refer to diagonal operator as DOP and non-diagonal operators as NDOP.

1. Bouncing worm algorithm

In this section we give detailed information on a straightforward (albeit not optimized) implementation of the bounce-algorithm (see Subsec. III A). Shown is an outline of the central part of the algorithm: the update of the operator-string and the state.

The algorithm works its way up the tree. It starts at the last branch which is assigned the nth position. At each position it is decided if the worm goes up the tree or down, in which case a new operator is chosen for the branch at this position. The necessary probabilities are calculated according to the expressions given in Eq. 21 and Eq. 23. If a new operator is chosen for the nth branch, the update is complete.

It is assumed that the tree is so high (the operator-string so long) that the root is never reached. If the root is reached, one has to choose an operator for the first branch according to the probabilities outlined in Sec. II after Eq. 14.

Schematically, using pseudocode, a bouncing worm update of a tree with n nodes can be outlined as follows:

```
pos = n                                          ▷ start at last branch
going_up = TRUE
while pos != n + 1 do
    ran = uniform(0, 1)
    if going_up then
        if operator at pos is NDOP then
            pos = pos − 1
        else if ran < cw then
            pos = pos − 1
        else
            going_up = FALSE
            choose new DOP at pos
            update state, w and N_1 at pos
            pos = pos + 1
        end if
    else
        if ran < b_u then
            going_up = TRUE
        else
            if ran − b_u < wN_1 then
                choose DOP at pos
            else
                choose NDOP at pos
            end if
            update state, w and N_1 at pos
            pos = pos + 1
        end if
    end if
end while
```

The weights w, u and b_u, N_1 as well as the state are stored at each node.

2. Driven worm algorithm

We now turn to a description of a (not optimized) implementation of the driven worm algorithm (see Subsec. III B). As above, we show an outline of the central part of the algorithm: the update of the operator-string and the state.

The worm works its way up the tree. It starts at the last branch which is assigned the position n. While going up the tree, the worm, at each node, goes further up the tree with probability α or turns around with probability $1-\alpha$. After turning around, the worm keeps going down until it reaches the end. At the nodes the worm visits new operators are chosen. When the worm reaches the end, it has to be decided whether or not the update should be accepted. The associated probabilities are calculated according to the expressions given in the main text (see Eq. 26, Eq. 27 and Eq. 16).

As above, we assume that the tree is so high (the operator-strings so long) that the root is never reached. If that the root is reached, one has to choose an operator for the first branch according to the probabilities outlined in Sec. II after Eq. 14.

Shown is the driven worm update of a tree with n nodes. Using pseudocode language, a driven worm update then takes the following form for a tree with n nodes:

```
pos = n                                              ▷ start at last branch
going_up = TRUE
while pos ! = n + 1 do
    ran = uniform(0, 1)
    while going_up do
        if ran < α then
            pos = pos − 1
        else
            going_up = FALSE
            if operator at pos is DOP then
                if ran − α < yN_{1/2} then
                    choose new NDOP at pos
                else
                    choose DOP at pos
                end if
            else
                if ran − α < zN_1 then
                    choose new DOP at pos
                else
                    choose NDOP at pos
                end if
                update state, weights, c, and N_1 at pos
                pos = pos + 1
            end if
        end if
    end while
    if ran < wN_1 then
        choose DOP at pos
    else
        choose NDOP at pos
    end if
    update state, weights, c, and N_1 at pos
    pos = pos + 1
end while
Accept or reject using old and new c's.
```

The weights, c, N_1 as well as the state are stored at each node. A new string is not always accepted.

[1] D. Ceperley and M. Kalos, in *Monte Carlo Methods in Statistical Physics*, Topics in Current Physics, Vol. 7, edited by K. Binder (Springer Berlin Heidelberg, 1986) pp. 145–194.

[2] R. Blankenbecler and R. L. Sugar, Physical Review D **27**, 1304 (1983).

[3] S. Liang, B. Doucot, and P. W. Anderson, Phys. Rev. Lett. **61**, 365 (1988); S. Liang, Phys. Rev. B **42**, 6555 (1990).

[4] S. Liang, Phys. Rev. Lett. **64**, 1597 (1990).

[5] N. Trivedi and D. M. Ceperley, Physical Review B **41**, 4552 (1990).

[6] K. J. Runge, Physical Review B **45**, 7229 (1992).

[7] M. Calandra, F. Becca, and S. Sorella, Physical Review Letters **81**, 5185 (1998).

[8] L. Capriotti, A. E. Trumper, and S. Sorella, Physical Review Letters **82**, 3899 (1999).

[9] W. von der Linden, Physics Reports **220**, 53 (1992).

[10] N. J. Cerf and O. C. Martin, International Journal of Modern Physics C **06**, 693 (1995).

[11] S. Sorella, G. E. Santoro, and F. Becca, "SISSA Lecture notes on Numerical methods for strongly correlated electrons," (2013).

[12] J. H. Hetherington, Physical Review A **30**, 2713 (1984).

[13] A. W. Sandvik, Phys. Rev. Lett. **95**, 207203 (2005); K. Beach and A. W. Sandvik, Nuclear Physics B **750**, 142 (2006); A. W. Sandvik and K. S. D. Beach, 0704.1469 (2007).

[14] A. W. Sandvik and H. G. Evertz, Phys. Rev. B **82**, 024407 (2010).

[15] A. Banerjee and K. Damle, Journal of Statistical Mechanics: Theory and Experiment **2010**, P08017 (2010).

[16] A. Deschner and E. S. Sørensen, unpublished (2014).

[17] G. Santoro, S. Sorella, L. Guidoni, A. Parola, and E. Tosatti, Physical Review Letters **83**, 3065 (1999).

[18] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).

[19] H. G. Evertz, G. Lana, and M. Marcu, Phys. Rev. Lett. **70**, 875 (1993).

[20] N. Prokof'ev and B. Svistunov, Phys. Rev. Lett. **87** (2001).

[21] N. Prokof'ev and B. Svistunov, *Worm Algorithm for Problems of Quantum and Classical Statistics*, arXiv e-print 0910.1393

(2009).
[22] F. Alet and E. S. Sørensen, Phys. Rev. E **67**, 015701 (2003).
[23] F. Alet and E. S. Sørensen, Phys. Rev. E **68**, 026702 (2003); P. Hitchcock, E. S. Sørensen, and F. Alet, Phys. Rev. E **70**, 016702 (2004).
[24] J.-S. Wang, Phys. Rev. E **72**, 036706 (2005).

5. Conclusions

In part I of this thesis, I presented two variational studies of the dimerized J_1-J_2 chain. The first publication was about the entanglement between the chain and an impurity attached to one of its ends. We explained how it is possible to calculate a genuine entanglement measure, the negativity, between a block of spins at one end of the J_1-J_2 chain and an impurity spin attached to the other end, within a conceptually very simple variational framework that was known to be applicable to the J_1-J_2 chain at the Majumdar-Ghosh point (MG-point). Via comparison with results of DMRG and other techniques, we were able to show that this variational method does not only work at the MG-point but also in the surrounding region.

The results for the negativity were used to compare it to a different quantity, S_{imp}, that had been introduced earlier to quantify the impurity entanglement in this model. It was found that the negativity is better suited to help understand the impurity entanglement. For small values of the coupling of the impurity to a chain with an odd number of sites, we, in accordance with prior studies, found the entanglement at the MG-point to be long-ranged. Increasing the impurity coupling was observed to quickly reduce the range of the entanglement. Another result contained in this publication is that a very small explicit dimerization can also suppress long-range entanglement very effectively.

The second publication was about a study of the J_1-J_2 chain with an odd number of sites. We found DMRG to not be applicable close to the Lifshitz point, because of a massive degeneracy developing at this point. Thus, the same variational method as in the first publication was used in the second publication. We showed that there is a sequence of level crossings starting at the Lifshitz point. Such level crossings are absent in the J_1-J_2 chain with an even number of sites, which is gapped throughout the region of interest. The correlation length was shown to have the same characteristic behaviour close to the disorder point for even and odd chains; at the disorder point it is, for example, minimal in both cases. For periodic boundary conditions, we were able to develop a robust understanding of the nature of the level crossing. For odd boundary conditions we were unable to do so. It would be truly great to understand why and how the intriguing pattern of the intertwining levels forms. The on-site magnetization and the entanglement entropy show dramatic changes while one is going through the level crossings.

The variational method used in the two publication makes no reference to the specifics of the system or the Hamiltonian. As long as the groundstate can reasonably well be approximated by a manageable number of dimerized states, it can be used and good results can be expected.

In the second part, I presented one publication about new algorithms for valence bond quantum Monte Carlo. In this publication, we showed that it is possible to implement worm

algorithms to generate substantial updates of the operator strings which are sampled in this method. Two algorithms with very different characteristics were introduced and studied. Several ways to potentially improve the worm algorithms were mentioned in the publication. I believe that it would be very interesting and possible to only update a fraction of all the nodes that a worm passes and thereby have updates that penetrate the string even more substantially.

In the sections accompanying the publication we explained how the general scheme of VBQMC can be translated to the S_z-basis. The algorithms of the publication are applicable to the S_z-method. Testing such an implementation would be very interesting, since efficient loop or cluster algorithms are not known for the VBQMC in the S_z-basis. It would also be very worthwhile to try to develop a loop algorithm. There is a lot of room for further research and it is far too early to make any decisive statement about the prospects of the worm algorithms and S_z-projector Monte Carlo.

A. Appendix

A.1. The spectrum of $H(i)$

In the following, the spectrum of the operator $H(1) = \mathbf{S}_1 \cdot \mathbf{S}_2 + \mathbf{S}_2 \cdot \mathbf{S}_3 + \mathbf{S}_1 \cdot \mathbf{S}_3$ acting on three spin ½ is calculated. I start by decomposing the 8 dimensional Hilbert space into three subspaces, which have definite total spin. Then, it is not difficult to calculate the eigenvalues of $H(1)$.

First we look at the subspace with $S_{\text{tot}} = 3/2$. The basis states are (in declining order of S_z-eigenvalue) given by:

$$\begin{aligned} e_8 =& |\uparrow\uparrow\uparrow\rangle\,, \\ e_7 =& \frac{1}{\sqrt{3}}\left[|\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle\right], \\ e_6 =& \frac{1}{\sqrt{3}}\left[|\uparrow\downarrow\downarrow\rangle + |\downarrow\downarrow\uparrow\rangle + |\downarrow\uparrow\downarrow\rangle\right], \\ e_5 =& |\downarrow\downarrow\downarrow\rangle\,. \end{aligned} \tag{A.1}$$

For the corresponding eigenvalues one finds that

$$H(1)\,e_5 = H(1)\,e_6 = H(1)\,e_7 = H(1)\,e_8 = \frac{3}{4}\,. \tag{A.2}$$

With the shorthand

$$|-\rangle = \frac{1}{\sqrt{2}}\left[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right],$$

eigenstates of the first subspace of $S_{\text{tot}} = 1/2$ can be written as

$$\begin{aligned} e_4 =& |\uparrow\rangle|-\rangle + |-\rangle|\uparrow\rangle\,, \\ e_3 =& |\downarrow\rangle|-\rangle + |-\rangle|\downarrow\rangle\,. \end{aligned} \tag{A.3}$$

Eigenstates of the second subspace of $S_{\text{tot}} = 1/2$ are given by

$$\begin{aligned} e_2 =& |\uparrow\rangle|-\rangle - |-\rangle|\uparrow\rangle\,, \\ e_1 =& |\downarrow\rangle|-\rangle - |-\rangle|\downarrow\rangle\,. \end{aligned} \tag{A.4}$$

For both subspaces it does not matter which spins are in the singlet state as long as they are the same ones for both states. Let it be the third spin that is in the state $|\uparrow\rangle$. Then,

$$\begin{aligned} H(1)\,|-\rangle|\uparrow\rangle &= \left[-3/4 + \mathbf{S}_1\cdot\mathbf{S}_3 + \mathbf{S}_2\cdot\mathbf{S}_3\right]\frac{1}{\sqrt{2}}\left[|\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle\right] \\ &= -3/4|-\rangle|\uparrow\rangle + \frac{1}{\sqrt{2}}\left[\frac{1}{4}\uparrow\downarrow\uparrow\rangle + \frac{1}{4}|\downarrow\uparrow\uparrow\rangle - \frac{1}{2}|\uparrow\uparrow\downarrow\rangle - \frac{1}{4}\uparrow\downarrow\uparrow\rangle - \frac{1}{4}|\downarrow\uparrow\uparrow\rangle + \frac{1}{2}|\uparrow\uparrow\downarrow\rangle\right] \\ &= -3/4|-\rangle|\uparrow\rangle\,. \end{aligned} \tag{A.5}$$

Thus,

$$H(1)\,e_1 = H(1)\,e_2 = H(1)\,e_3 = H(1)\,e_4 = -3/4\,. \tag{A.6}$$

We have thus shown that the spectrum of $H(1)$ (and thus all $H(i)$ from equation 1.6) is given by 3/4 and −3/4 and that both eigenvalues are fourfold degenerate.

A.2. Valence bond quantum Monte Carlo sampling for $S > 1/2$

Trying to simulate spins with spin quantum number bigger than $S = 1/2$ introduces difficulties that are absent when $S = 1/2$, as is explained in section 4.1 of chapter 4 in part II. One has to deal with the fact that a bond operator creates more than one states and that thus the sum over r in

$$\begin{aligned} (-H + C)^n|\alpha_{\text{trial}}\rangle &= \Big(\sum h_{ij}\Big)^n|\alpha\rangle \\ &= \sum_k \prod_{u=1}^{n} h_{i_{uk}j_{uk}}|\alpha\rangle \\ &= \sum_k s_k|\alpha\rangle \\ &= \sum_k \sum_r^{m(k)} w_{kr}|\alpha_{kr}\rangle \end{aligned} \tag{A.7}$$

has more than one term. We call this branching. Equation A.7 is the same as equation 4.11. It was copied for the readers convenience. It is the equation that we want to calculate (the projection onto the groundstate) and the quantity in the last line is the quantity we evaluate with Monte Carlo.

The goal is to generate the α_{kr} with a probability proportional to the w_{rk}. The weight of an operator string is calculated by acting on $|a_{\text{trial}}\rangle$ with one of the bond operators after the other.

Just choosing a state from the superposition that is created by an operator string and using the corresponding w_{rk} in the acceptance step does not work. The reason is that the states that are compared are not *chosen* with the same probability. Imagine a string s_1 that creates a superposition of 20 states. A single state is then chosen with probability 0.05. Now imagine the next string to generate a superposition of 100 states. A single state is then chosen with a probability of 0.01. Thus, the probability to consider a state from s_1 as a new state coming from s_2 is $5 = 100/20$ times higher than the probability to consider a state from s_2 coming from s_1. This imbalance has to be taken into account.

Firstly, I explain a very simple way of overcoming this problem and show that it is valid, i.e. that it has detailed balance. Then, I explain how the simple method can be improved significantly.

Here is the simple way of doing it. If acting on the state generates c branches, one of the branches is chosen randomly and the weight of the operator string given by the coefficient a state is generated with is multiplied with c.

It will now be shown that with this procedure detailed balance is achieved. For this purpose it is necessary to distinguish between the weight w_{ij} of a state in the sum in equation A.7 which is given by the numerical factors generated by the action of an operator string and the weight g_{ij} that is used during the acceptance rejection step during the generation of the ensemble of states that is used in the simulation.

Let us start with an arbitrary state $|\alpha_{1r}\rangle$ generated by the operator string s_1. Its weight in the sum in equation A.7 shall be w_{1r}. Its weight in the simulation is denoted by g_{1r}. Let c_1^1 be number of terms in the superposition created if we act with the first bond operator on $|\alpha_{\text{trial}}\rangle$. If we choose a state $|\alpha_1^1\rangle$ with coefficient k_1^1 randomly from the superposition, it is chosen with a probability of $1/c_1^1$. Acting on the chosen state $|\alpha_1^1\rangle$ with the next operator in s_1, we get a superposition of c_1^2 states back. If we randomly choose one state $|\alpha_1^2\rangle$ that contributes the coefficient k_1^2, it will be chosen with the probability $1/c_1^2$. After application of the first two operators in s_1, we therefore get $|\alpha_1^2\rangle$ with a probability of $1/(c_1^1 c_1^2)$. The weight from the coefficients is given by $k_1^1 k_1^2$. After the application of all of s_1's n bond operators, the state $|\alpha_{1r}\rangle = |\alpha_1^n\rangle$ will be the state we have to accept or reject with the probability of

$$P_{1r} = \frac{1}{\prod c_1^i} \,. \tag{A.8}$$

The weight of $|\alpha_{1r}\rangle$ in the sum in equation A.7 is

$$w_{1r} = \prod k_1^i \,. \tag{A.9}$$

The weight we use in subsequent acceptance rejection steps is

$$g_{1r} = \frac{w_{1r}}{P_{1r}} \,. \tag{A.10}$$

Now we calculate the transition amplitude to a second state $|\alpha_{2l}\rangle$ generated by an operator

string s_2 which itself is generated by randomly changing some of the bond operators making up s_1. The probability for s_2 to be chosen is denoted by p. It is the same as choosing s_1 coming from the string s_2. The necessary quantities w_{2l} and g_{2l} can be calculated following the recipe given for $|\alpha_{1r}\rangle$.

The acceptance probability $u_{1r\to 2l}$ of $|\alpha_{2l}\rangle$ is given by $u_{1r\to 2l} = \min\left[\frac{g_{2l}}{g_{1r}}, 1\right]$. Combining the probabilities for s_2 to be chosen, for $|\alpha_{2l}\rangle$ to be chosen from the superposition created by s_2 and the acceptance probability, we find that the transition probability to go from $|\alpha_{1r}\rangle$ to $|\alpha_{2l}\rangle$ is given by

$$t_{1r\to 2l} = p\, P_{2l}\, u_{1r\to 2l}\ . \tag{A.11}$$

Similarly we can get the transition probability to go from $|\alpha_{2l}\rangle$ to $|\alpha_{1r}\rangle$ that

$$t_{2l\to 1r} = p\, P_{1r}\, u_{2l\to 1r}\ . \tag{A.12}$$

It follows that

$$\begin{aligned}\frac{t_{2l\to 1r}}{t_{1r\to 2l}} &= \frac{p\, P_{1r}\, u_{2l\to 1r}}{p\, P_{2l}\, u_{1r\to 2l}} = \frac{P_{1r}\, g_{1r}}{P_{2l}\, g_{2l}} \\ &= \frac{w_{1r}}{w_{2l}}\ .\end{aligned} \tag{A.13}$$

Detailed balance is observed.

The better way to do it

It is desirable to keep fluctuation of the weight used in the acceptance rejection step as small as possible. This enhances the acceptance ratio and thus makes the algorithm much more efficient. It can be achieved by choosing states that have big weight more often. We should thus choose branches with a probability proportional to their coefficients.

Let us call the coefficients of states created by the jth operator in the first string k^j_{1i}. We choose the ith branch after acting with the jth operator with the probability $k^j_{1i}/\sum_i k^j_{1i}$. The probability for a branch to be chosen from the superposition generated by the first string is then

$$P_{1r} = \prod_j \left[\frac{k^j_{1i}}{\sum_i k^j_{1i}}\right]\ . \tag{A.14}$$

In this way, branches with big coefficients are chosen more often and fewer states have to be rejected. The introduction of this probability has to be balanced through a division of the

rejection weight by P_{1r}. We get

$$\begin{aligned} g_{1r} &= \frac{w_{1r}}{P_{1r}} \\ &= \prod_j \sum_i k_{1i}^j \ . \end{aligned} \tag{A.15}$$

Doing these steps for a second string and following the steps that led to equation A.13 directly leads to the conclusion that this method has detailed balance.

Bibliography

Assa Auerbach, *Interacting Electrons and Quantum Magnetism* (Springer, 1994).

C. Lacroix, P. Mendels, and F. Mila, eds., *Introduction to Frustrated Magnetism* (Springer, 2011).

J. Hubbard, "Electron Correlations in Narrow Energy Bands," Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **276**, 238–257 (1963).

M. Takahashi, "Half-filled Hubbard model at low temperature," Journal of Physics C: Solid State Physics **10**, 1289 (1977).

N. D. Mermin and H. Wagner, "Absence of Ferromagnetism or Antiferromagnetism in One- or Two-Dimensional Isotropic Heisenberg Models," Physical Review Letters **17**, 1133–1136 (1966).

F. D. M. Haldane, "Nonlinear Field Theory of Large-Spin Heisenberg Antiferromagnets: Semiclassically Quantized Solitons of the One-Dimensional Easy-Axis Néel State," Physical Review Letters **50**, 1153–1156 (1983).

P. W. Anderson, "Resonating valence bonds: A new kind of insulator?" Materials Research Bulletin **8**, 153–160 (1973).

P. W. Anderson, "The Resonating Valence Bond State in La_2CuO_4 and Superconductivity," Science **235**, 1196–1198 (1987).

H.-C. Jiang, H. Yao, and L. Balents, "Spin liquid ground state of the spin-½ square J_1-J_2 Heisenberg model," Physical Review B **86**, 024424 (2012a).

S. Yan, D. A. Huse, and S. R. White, "Spin-Liquid Ground State of the S = ½ Kagome Heisenberg Antiferromagnet," Science **332**, 1173 –1176 (2011).

L. Balents, "Spin liquids in frustrated magnets," Nature **464**, 199–208 (2010).

T.-H. Han, J. S. Helton, S. Chu, D. G. Nocera, J. A. Rodriguez-Rivera, C. Broholm, and Y. S. Lee, "Fractionalized excitations in the spin-liquid state of a kagome-lattice antiferromagnet," Nature **492**, 406–410 (2012).

X.-G. Wen, "Vacuum degeneracy of chiral spin states in compactified space," Physical Review B **40**, 7387–7390 (1989).

X.-G. Wen, "Topological Order: From Long-Range Entangled Quantum Matter to a Unified Origin of Light and Electrons," ISRN Condensed Matter Physics **2013**, 1–20 (2013).

A. Kitaev, "Fault-tolerant quantum computation by anyons," Annals of Physics **303**, 2–30 (2003).

S. Depenbrock, I. P. McCulloch, and U. Schollwöck, "Nature of the Spin-Liquid Ground State of the $S = 1/2$ Heisenberg Model on the Kagome Lattice," Physical Review Letters **109**, 067201 (2012).

A. Kitaev and J. Preskill, "Topological Entanglement Entropy," Physical Review Letters **96**, 110404 (2006).

M. Levin and X.-G. Wen, "Detecting Topological Order in a Ground State Wave Function," Physical Review Letters **96**, 110405 (2006).

C. K. Majumdar and D. K. Ghosh, "On Next-Nearest-Neighbor Interaction in Linear Chain. II," Journal of Mathematical Physics **10**, 1399–1402 (1969a).

C. K. Majumdar and D. K. Ghosh, "On Next-Nearest-Neighbor Interaction in Linear Chain. I," Journal of Mathematical Physics **10**, 1388–1398 (1969b).

M. Hase, H. Kuroe, K. Ozawa, O. Suzuki, H. Kitazawa, G. Kido, and T. Sekine, "Magnetic properties of $Rb_2Cu_2Mo_3O_{12}$ including a one-dimensional spin-½ Heisenberg system with ferromagnetic first-nearest-neighbor and antiferromagnetic second-nearest-neighbor exchange interactions," Physical Review B **70**, 104426 (2004).

S. R. White and I. Affleck, "Dimerization and incommensurate spiral spin correlations in the zigzag spin chain: Analogies to the Kondo lattice," Physical Review B **54**, 9862 (1996).

S. Eggert, "Numerical evidence for multiplicative logarithmic corrections from marginal operators," Physical Review B **54**, R9612–R9615 (1996).

T. Hamada, J-i. Kane, S-i. Nakagawa, and Y. Natsume, "Exact Solution of Ground State for Uniformly Distributed RVB in One-Dimensional Spin-½ Heisenberg Systems with Frustration," Journal of the Physical Society of Japan **57**, 1891–1894 (1988).

T. Tonegawa and I. Harada, "One-Dimensional Isotropic Spin-½ Heisenberg Magnet with Ferromagnetic Nearest-Neighbor and Antiferromagnetic Next-Nearest-Neighbor Interactions," Journal of the Physical Society of Japan **58**, 2902–2915 (1989).

S. Furukawa, M. Sato, S. Onoda, and A. Furusaki, "Ground-state phase diagram of a spin-½ frustrated ferromagnetic XXZ chain: Haldane dimer phase and gapped/gapless chiral phases," Physical Review B **86**, 094417 (2012).

C. Itoi and S. Qin, "Strongly reduced gap in the zigzag spin chain with a ferromagnetic interchain coupling," Physical Review B **63**, 224423 (2001).

S. Furukawa, M. Sato, and S. Onoda, "Chiral Order and Electromagnetic Dynamics in One-Dimensional Multiferroic Cuprates," Physical Review Letters **105**, 257205 (2010).

B. Willenberg, M. Schäpers, K. C. Rule, S. Süllow, M. Reehuis, H. Ryll, B. Klemke, K. Kiefer, W. Schottenhamel, B. Büchner, B. Ouladdiaf, M. Uhlarz, R. Beyer, J. Wosnitza, and A. U. B. Wolter, "Magnetic Frustration in a Quantum Spin Chain: The Case of Linarite $PbCuSO_4(OH)_2$," Physical Review Letters **108**, 117202 (2012).

Y. Yasui, Y. Yanagisawa, M. Sato, and I. Terasaki, "Relationship between ferroelectricity and magnetic structure of $PbCuSO_4(OH)_2$ with CuO_2 ribbon chains," Journal of Physics: Conference Series **320**, 012087 (2011).

S. Park, Y. J. Choi, C. L. Zhang, and S-W. Cheong, "Ferroelectricity in an S=½ Chain Cuprate," Physical Review Letters **98**, 057601 (2007).

Y. Naito, K. Sato, Y. Yasui, Y. Kobayashi, Y. Kobayashi, and M. Sato, "Ferroelectric Transition Induced by the Incommensurate Magnetic Ordering in $LiCuVO_4$," Journal of the Physical Society of Japan **76**, 023708 (2007).

G. Castilla, S. Chakravarty, and V. J. Emery, "Quantum Magnetism of $CuGeO_3$," Physical Review Letters **75**, 1823–1826 (1995).

P. M. van den Broek, "Exact value of the ground state energy of the linear antiferromagnetic Heisenberg chain with nearest and next-nearest neighbour interactions," Physics Letters A **77**, 261–262 (1980).

B. S. Shastry and B. Sutherland, "Excitation Spectrum of a Dimerized Next-Neighbor Antiferromagnetic Chain," Physical Review Letters **47**, 964 (1981).

E. Sørensen, I. Affleck, D. Augier, and D. Poilblanc, "Soliton approach to spin-Peierls antiferromagnets: Large-scale numerical results," Physical Review B **58**, R14701 (1998).

R. Bursill, G. A. Gehring, D. J. J. Farnell, J. B. Parkinson, T. Xiang, and C. Zeng, "Numerical and approximate analytical results for the frustrated spin-½ quantum spin chain," Journal of Physics: Condensed Matter **7**, 8605–8618 (1995).

D. J. Klein, "Exact ground states for a class of antiferromagnetic heisenberg models with short-range interactions," Journal of Physics A: Mathematical and General **15**, 661 (1982).

I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, "Rigorous results on valence-bond ground states in antiferromagnets," Physical Review Letters **59**, 799 (1987).

H. Bethe, "Zur Theorie der Metalle," Zeitschrift für Physik **71**, 205–226 (1931).

A. W. Sandvik, "Computational studies of quantum spin systems," AIP Conference Proceedings **1297**, 135–338 (2010), pre-print: arXiv:1101.3281.

Steven R. White, "Density matrix formulation for quantum renormalization groups," Physical Review Letters **69**, 2863–2866 (1992).

U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states," Annals of Physics **326**, 96–192 (2011).

J. I. Cirac and F. Verstraete, "Renormalization and tensor product states in spin chains and lattices," Journal of Physics A: Mathematical and Theoretical **42**, 504004 (2009).

N. Kawashima and K. Harada, "Recent Developments of World-Line Monte Carlo Methods," Journal of the Physical Society of Japan **73**, 1379–1414 (2004).

A. W. Sandvik and J. Kurkijärvi, "Quantum Monte Carlo simulation method for spin systems," Physical Review B **43**, 5950–5961 (1991).

D.M. Ceperley and M.H. Kalos, "Quantum many-body problems," in *Monte Carlo Methods in Statistical Physics*, Topics in Current Physics, Vol. 7, edited by Kurt Binder (Springer Berlin Heidelberg, 1986) pp. 145–194.

N. Trivedi and D. M. Ceperley, "Ground-state correlations of quantum antiferromagnets: A green-function monte carlo study," Physical Review B **41**, 4552–4569 (1990).

S. Sorella, G. E. Santoro, and F. Becca, "SISSA Lecture notes on Numerical methods for strongly correlated electrons," Trieste: SISSA (2013).

Shoudan Liang, "Monte Carlo calculations of the correlation functions for Heisenberg spin chains at $T=0$," Phys. Rev. Lett. **64**, 1597–1600 (1990a).

Anders W. Sandvik, "Ground state projection of quantum spin systems in the valence-bond basis," Phys. Rev. Lett. **95**, 207203 (2005).

A. W. Sandvik and H. G. Evertz, "Loop updates for variational and projector quantum Monte Carlo simulations in the Valence-Bond Basis," Physical Review B **82**, 024407 (2010).

H. Gies, K. Langfeld, and L. Moyaerts, "Casimir effect on the worldline," Journal of High Energy Physics **2003**, 018 (2003).

N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," The Journal of Chemical Physics **21**, 1087–1092 (1953).

W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," Biometrika **57**, 97–109 (1970).

R. H. Swendsen and J.-S. Wang, "Nonuniversal critical dynamics in Monte Carlo simulations," Physical Review Letters **58**, 86–88 (1987).

U. Wolff, "Collective Monte Carlo Updating for Spin Systems," Physical Review Letters **62**, 361–364 (1989).

H. G. Evertz, G. Lana, and M. Marcu, "Cluster algorithm for vertex models," Physical Review Letters **70**, 875–879 (1993).

N. Prokof'ev and B. Svistunov, "Worm Algorithms for Classical Statistical Models," Physical Review Letters **87**, 160601 (2001).

M. Troyer and U.-J. Wiese, "Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations," Physical Review Letters **94** (2005).

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

S. Sorella, Y. Otsuka, and S. Yunoki, "Absence of a Spin Liquid Phase in the Hubbard Model on the Honeycomb Lattice," Scientific Reports **2** (2012).

J. von Neumann, "Thermodynamik quantenmechanischer Gesamtheiten," Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen , 273 (1927).

M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

L. Amico, R. Fazio, A. Osterloh, and V. Vedral, "Entanglement in many-body systems," Reviews of Modern Physics **80**, 517–576 (2008).

M. B. Hastings, "Lieb-Schultz-Mattis in higher dimensions," Physical Review B **69**, 104431 (2004).

M. B. Hastings, "An area law for one-dimensional quantum systems," Journal of Statistical Mechanics: Theory and Experiment **2007**, P08024 (2007).

S. T. Flammia, A. Hamma, T. L. Hughes, and X.-G. Wen, "Topological Entanglement Rényi Entropy and Reduced Density Matrix Structure," Physical Review Letters **103**, 261601 (2009).

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

H.-C. Jiang, Z. Wang, and L. Balents, "Identifying topological order by entanglement entropy," Nature Physics **8**, 902–905 (2012b).

C. Holzhey, F. Larsen, and F. Wilczek, "Geometric and renormalized entropy in conformal field theory," Nuclear Physics B **424**, 443–467 (1994).

P. Calabrese and J. Cardy, "Entanglement entropy and quantum field theory," Journal of Statistical Mechanics: Theory and Experiment **2004**, P06002 (2004).

J. Eisert, M. Cramer, and M. B. Plenio, "*Colloquium*: Area laws for the entanglement entropy," Reviews of Modern Physics **82**, 277–306 (2010).

M. B. Plenio and S. Virmani, *An introduction to entanglement measures*, arXiv e-print (2006) published as Quant.Inf.Comput. **7**, 1 (2007).

C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed-state entanglement and quantum error correction," Physical Review A **54**, 3824 (1996).

W. K. Wootters, "Entanglement of Formation of an Arbitrary State of Two Qubits," Physical Review Letters **80**, 2245 (1998).

W. K. Wootters, "Entanglement of formation and concurrence," Quant.Inf.Comput. **1**, 27–44 (2001).

G. Vidal and R. F. Werner, "Computable measure of entanglement," Physical Review A **65**, 032314 (2002).

A. Peres, "Separability Criterion for Density Matrices," Physical Review Letters **77**, 1413 (1996).

E. S. Sørensen, M.-S. Chang, N. Laflorencie, and I. Affleck, "Impurity entanglement entropy and the Kondo screening cloud," Journal of Statistical Mechanics: Theory and Experiment **2007**, L01001 (2007a).

E. S. Sørensen, M.-S. Chang, N. Laflorencie, and I. Affleck, "Quantum impurity entanglement," Journal of Statistical Mechanics: Theory and Experiment **2007**, P08003 (2007b).

N. Laflorencie, E. S. Sørensen, and I. Affleck, "The Kondo effect in spin chains," Journal of Statistical Mechanics: Theory and Experiment **2008**, P02007 (2008).

A. Bayat, P. Sodano, and S. Bose, "Negativity as the entanglement measure to probe the Kondo regime in the spin-chain Kondo model," Physical Review B **81**, 064429 (2010).

P. Sodano, A. Bayat, and S. Bose, "Kondo cloud mediated long-range entanglement after local quench in a spin chain," Physical Review B **81**, 100412 (2010).

R. M. Hornreich, M. Luban, and S. Shtrikman, "Critical Behavior at the Onset of $\vec{k}$-Space Instability on the λ Line," Physical Review Letters **35**, 1678–1681 (1975).

W. Selke, "The ANNNI model - theoretical analysis and experimental application," Physics Reports **170**, 213–264 (1988).

A. G. Slivka, E. I. Gerzanich, P. P. Guranich, and V. S. Shusta, "Phase p, T, x-diagram and peculiarities of physical properties of $Sn_2P_2(Se_xS_{1-x})_6$ ferroelectric crystals near Lifshitz point," Ferroelectrics **103**, 71–82 (1990).

F. S. Bates, W. Maurer, T. P. Lodge, M. F. Schulz, M. W. Matsen, K. Almdal, and K. Mortensen, "Isotropic Lifshitz Behavior in Block Copolymer-Homopolymer Blends," Physical Review Letters **75**, 4429–4432 (1995).

M. Henkel and M. Pleimling, *Non-Equilibrium Phase Transitions Volume 2: Ageing and Dynamical Scaling Far from Equilibrium* (Springer, 2010).

U. Schollwöck, Th. Jolicoeur, and T. Garel, "Onset of incommensurability at the valence-bond-solid point in the $S = 1$ quantum spin chain," Physical Review B **53**, 3304–3311 (1996).

R. Roth and U. Schollwöck, "Frustrated antiferromagnetic quantum spin chains for spin length $S > 1$," Physical Review B **58**, 9264–9268 (1998).

J. Stephenson, "Ising Model with Antiferromagnetic Next-Nearest-Neighbor Coupling: Spin Correlations and Disorder Points," Physical Review B **1**, 4405–4409 (1970a).

J. Stephenson, "Two one-dimensional Ising models with disorder points," Canadian Journal of Physics **48**, 1724–1734 (1970b).

T. Garel and J. M. Maillard, "RPA approach to disorder points," Journal of Physics C: Solid State Physics **19**, L505–L511 (1986).

G. Fáth and A. Sütő, "Commensurate and incommensurate correlations in Haldane-gap antiferromagnets," Physical Review B **62**, 3778–3785 (2000).

S. Liang, "Existence of Néel order at T=0 in the spin-½ antiferromagnetic Heisenberg model on a square lattice," Physical Review B **42**, 6555–6560 (1990b).

S. Liang, B. Doucot, and P. W. Anderson, "Some New Variational Resonating-Valence-Bond-Type Wave Functions for the Spin-½ Antiferromagnetic Heisenberg Model on a Square Lattice," Physical Review Letters **61**, 365 (1988).

A. W. Sandvik and K. S. D. Beach, "Monte Carlo Simulations of Quantum Spin Systems in the Valence Bond Basis," Computer Simulation Studies in Condensed Matter Physics XX (2007).

A. W. Sandvik, "Evidence for Deconfined Quantum Criticality in a Two-Dimensional Heisenberg Model with Four-Spin Interactions," Physical Review Letters **98**, 227202 (2007).

S. Pujari, K. Damle, and F. Alet, "Néel-State to Valence-Bond-Solid Transition on the Honeycomb Lattice: Evidence for Deconfined Criticality," Physical Review Letters **111**, 087203 (2013).

F. Alet, S. Capponi, N. Laflorencie, and M. Mambrini, "Valence Bond Entanglement Entropy," Physical Review Letters **99**, 117204 (2007).

A. B. Kallin, I. González, M. B. Hastings, and R. G. Melko, "Valence Bond and von Neumann Entanglement Entropy in Heisenberg Ladders," Physical Review Letters **103**, 117203 (2009).

F. Alet, I. P. McCulloch, S. Capponi, and M. Mambrini, "Valence-bond entanglement entropy of frustrated spin chains," Physical Review B **82**, 094452 (2010).

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

K. S. D. Beach, F. Alet, M. Mambrini, and S. Capponi, "SU(N) Heisenberg model on the square lattice: A continuous-N quantum Monte Carlo study," Physical Review B **80**, 184401 (2009).

K. S. D. Beach and A. W. Sandvik, "Some formal results for the valence bond basis," Nuclear Physics B **750**, 142–178 (2006).

O. Golinelli, Th. Jolicoeur, and R. Lacaze, "Finite-lattice extrapolations for a Haldane-gap antiferromagnet," Physical Review B **50**, 3037 (1994).

A. W. Sandvik, "Stochastic series expansion method with operator-loop update," Physical Review B **59**, R14157–R14160 (1999).

F. Alet and E. S. Sørensen, "Cluster Monte Carlo algorithm for the quantum rotor model," Physical Review E **67**, 015701 (2003a).

F. Alet and E. S. Sørensen, "Directed geometrical worm algorithm applied to the quantum rotor model," Physical Review E **68**, 026702 (2003b).