Semi-Vortex Cluster Algorithm for the two dimensional XY Model

Master Thesis
Faculty of Science, University of Bern

handed in by
Marius Descombes

2021

Supervisor
Prof. Dr. Uwe-Jens Wiese

Abstract

In this thesis, the Wolff multi-cluster algorithm was used to simulate the two dimensional XY model. First improved estimators for the magnetic susceptibility and the spin-spin correlation function have been derived. Then, the concept of semi-vortices has been introduced in order to obtain an improved estimator for the vortex-vortex correlation function. This has then been used to observe the behavior of vortices and anti-vortices around the Berezinskii-Kosterlitz-Thouless phase transition.

Contents

1 Introduction

It was long believed that there is no order in two dimensional matter, even at zero temperature, because thermal fluctuations would destroy it. Without an order, there cannot be a transition to a phase where there is no order. David Thouless and Michael Kosterlitz took on this problem which lead to a complete new understanding of phase transitions in two dimensions. At the heart of their description are vortices which form tight pairs at low temperatures. When the temperature rises, the vortices suddenly unbind from each other and the Berezinskii-Kosterlitz-Thouless phase transition takes place. For their work on this transition, Kosterlitz and Thouless were awarded the 2016 Nobel prize in physics.

In this thesis we investigate the two dimensional XY model on a square lattice. In the second chapter we introduce the XY model, in particular the cluster representation of this model which gives rise to a highly efficient Monte Carlo algorithm, the Wolff cluster algorithm. Vortices and, most importantly, the concept of semi-vortices is also described in this chapter. In the third chapter, we familiarize ourselves with Monte Carlo simulations, in particular with the aforementioned algorithm, and the methods required to analyze data obtained from such a simulation. In the remainder of the thesis we derive several improved estimators, most important the vortex-vortex correlation function and investigate the behavior of vortices at the Berezinskii-Kosterlitz-Thouless transition.

2 The XY model

The classical XY model is a lattice model of statistical mechanics. It is a special case of the d-dimensional O(n) model with $n = 2$. The other special cases are the Ising model with $n = 1$ and the Heisenberg model with $n = 3$. As lattice one may choose any lattice type, for example a square or a honeycomb lattice, in any number of dimensions d, for example a spatial lattice ($d = 3$). In this thesis, however, a two dimensional square lattice is used.

2.1 The XY model on a square lattice

The XY model may be defined on a two dimensional square lattice Λ with unit lattice spacing. With each lattice site $x \in \Lambda$, a classical spin variable $\mathbf{s}_x \in S^1 = \{\mathbf{s} \in \mathbb{R}^2 | \, |\mathbf{s}| = 1\}$ is associated. The lattice and the spins are shown in Figure 2.1. Since a simulation of the XY model requires

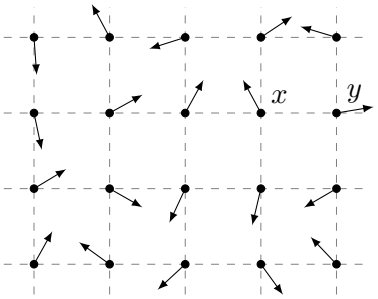


Figure 2.1: A two dimensional square lattice. The lattice sites are represented by dots and the associated spins by the arrows. The dashed lines indicate the lattice. The four lattice sites connected by the lattice to the site x are the nearest neighbors of x.

the lattice to be of finite extent, the question arises how to treat the boundaries of the lattice. In order to maintain translation invariance and to reduce finite-size effects, we introduce periodic boundary conditions. One may visualize these boundary conditions as wrapping the two dimensional lattice on a torus embedded in three dimensions. In this way, the first and the last lattice site on the same row become nearest neighbors of one another. The same holds for the top and bottom lattice sites of each column. A configuration of the system is completely defined by assigning a value $\mathbf{s}_x \in S^1$ to each of the spins on the lattice. We denote these spin configurations by $[\mathbf{s}]$. On a lattice of $|\Lambda|$ sites, the configuration space, i.e. the space of all configurations of the system is $(S^1)^{|\Lambda|}$.

We will consider the XY model with only nearest neighbor interactions, that is, governed by the Hamilton function

$$\mathcal{H}[\mathbf{s}] = -J \sum_{\langle xy \rangle} \mathbf{s}_x \cdot \mathbf{s}_y \tag{2.1}$$

with the coupling constant $J > 0$ that favors similarly aligned spins and where $\langle xy \rangle$ denotes that the sum runs over all nearest neighbors. This Hamilton function has a global symmetry, it is invariant under $O(2)$ transformations

$$\mathbf{s}'_x = O(\theta)\, \mathbf{s}_x, \qquad O(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in O(2) \tag{2.2}$$

where $\theta \in (-\pi, \, \pi]$. In the XY model, the system is in thermal equilibrium with its environment, that is, it is described by a canonical ensemble. The partition function, which depends

on the inverse temperature $\beta = 1/k_B T \geq 0$, is

$$Z(\beta) = \int \mathcal{D}\mathbf{s}\; e^{-\beta \mathcal{H}[\mathbf{s}]}, \tag{2.3}$$

where k_B is the Boltzmann constant, T the temperature, and we introduce the abbreviation

$$\int \mathcal{D}\mathbf{s} \;:= \int_{S^1} d\mathbf{s}_1 \cdots \int_{S^1} d\mathbf{s}_N = \prod_{x \in \Lambda} \int_{S^1} d\mathbf{s}_x. \tag{2.4}$$

The probability density to find the configuration $[\mathbf{s}]$ is given by the Boltzmann distribution

$$\rho[\mathbf{s}] = \frac{1}{Z(\beta)} e^{-\beta \mathcal{H}[\mathbf{s}]}. \tag{2.5}$$

For a classical observable, the expectation value is defined as

$$\langle O \rangle := \int \mathcal{D}\mathbf{s}\; \rho[\mathbf{s}]\; O[\mathbf{s}]. \tag{2.6}$$

For example, the spin-spin correlation function is defined as

$$\langle \mathbf{s}_x \cdot \mathbf{s}_y \rangle := \frac{1}{Z(\beta)} \int \mathcal{D}\mathbf{s}\; e^{-\beta \mathcal{H}[\mathbf{s}]}\; \mathbf{s}_x \cdot \mathbf{s}_y. \tag{2.7}$$

Because of translation invariance, the spin-spin correlation function does not depend on the coordinates $\mathbf{x}$, $\mathbf{y}$ of the lattice sites x and y, but only on the distance vector $\mathbf{y} - \mathbf{x}$ between them. For distances much larger than the lattice spacing, i.e. for $|\mathbf{y} - \mathbf{x}| \gg 1$, the spin-spin correlation function is of the form

$$\langle \mathbf{s}_x \cdot \mathbf{s}_y \rangle \propto \exp\left(-|\mathbf{y} - \mathbf{x}|/\xi\right), \tag{2.8}$$

where ξ is the correlation length.

2.2 Plaquettes and vorticity

The elementary lattice squares of a square lattice are called plaquettes. If we identify each plaquette with the site at its center, we find the so-called dual lattice which is also a square lattice. The lattice, its dual lattice, and the plaquettes are depicted in Figure 2.2.

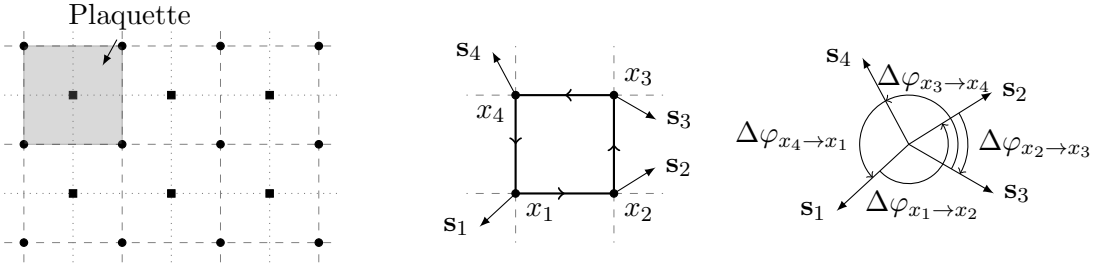


Figure 2.2: (left) A square lattice and its dual lattice. The lattice is indicated by dashed lines, the lattice sites by circles, the dual lattice by dotted lines, and the dual lattice sites by squares. (right) Graphical method to determine the vorticity of a plaquette.

To each plaquette a vorticity $\nu_\square \in \{-1, 0, 1\}$ can be assigned, defined as

$$\nu_\square = \frac{1}{2\pi} \left(\sum_{i=1}^{3} \Delta\varphi_{x_i \to x_{i+1}} + \Delta\varphi_{x_4 \to x_1} \right), \tag{2.9}$$

where $\Delta\varphi_{x\to y}$ is the oriented angle from spin $\mathbf{s}_x$ to $\mathbf{s}_y$, such that $\Delta\varphi_{x\to y} \in (-\pi, \pi]$ (see Figure 2.2). One may also calculate the vorticity of a plaquette by following the path around the edges of the plaquette in counterclockwise direction and summing up the oriented angles from one spin to the next one in the order they appear on the path. This path is also shown in Figure (2.2). Since the four spins at the corners of the plaquette account for the vorticity of the plaquette, we will sometimes refer to them and the corresponding lattice sites as the spins and lattice sites of the plaquette. Plaquettes with vorticity $+1$ are called vortices and those with vorticity -1 anti-vortices. Note that there is one configuration of the four spins of a plaquette which yields the vorticity $\nu_\square = 2$. Since the spin configurations which contain plaquettes with vorticity $\nu_\square = 2$ form a null set in configuration space, we can usually ignore these configurations. However, in a simulation of the XY model we generate spin configurations numerically and, therefore, such configurations may occur.

The notion of vorticity may be extended to arbitrary paths on the lattice from one lattice site to another by summing up the oriented angles from one spin to the next one as they appear on the path. For such an arbitrary path, the vorticity takes values in $\mathbb{R}$. From the definition, it follows that if we traverse a path $\gamma_{x\to y}$ from lattice site x to y in opposite direction, we find for the vorticity $\nu(\gamma_{x\to y}) = -\nu(\gamma_{y\to x})$. This simple observation allows for

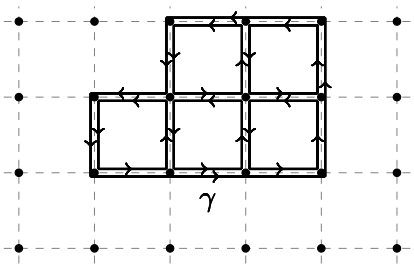


Figure 2.3: A closed path on the lattice and the paths around the plaquettes it encloses.

the calculation of the vorticity along a contractible closed path γ on the lattice. For the path depicted in Figure 2.3, the vorticity along γ is the same as the sum of the vorticities of the paths around each plaquette enclosed by the path γ. This is because the contributions of the two paths between each pair of neighboring plaquettes cancel each other. More generally we find

$$\nu(\gamma) = \sum_{\square} n(\gamma, \square)\, \nu_\square, \tag{2.10}$$

where we sum over all plaquettes $\square$ enclosed by the path γ and $n(\gamma, \square)$ is the winding number of γ around $\square$. As an application of this equation we may consider the path around the edges of the lattice traversed in counterclockwise direction. Because of the periodic boundary conditions, the vorticity along this path vanishes and we find $\sum_\square \nu_\square = 0$, that is, each configuration has the same number of vortices and anti-vortices.

2.3 Berezinskii-Kosterlitz-Thouless phase transition

The XY model does not have any long-range order since the ground configuration is unstable against low-energy spin-wave excitations. As there is no ordered phase, there cannot be a phase transition of the usual type. Nonetheless, the XY model does have a phase transition, the Berezinskii-Kosterlitz-Thouless phase transition (BKT-transition), from a quasi-ordered or spin-wave phase below some critical temperature T_c to a disordered phase above the critical temperature. Kosterlitz and Thouless argued in [5, 4] that this phase transition is due to the

unbinding of bound vortex-anti-vortex pairs below the critical temperature to free vortices and anti-vortices above the critical temperature. This unbinding can be seen from a fairly simple argument. The energy of a single vortex is

$$E \approx \pi J \log\left(\frac{A}{A_0}\right), \tag{2.11}$$

where A is the area of the system and $A_0 = a^2$ with a, the lattice spacing. Since there are A/A_0 possible positions for the vortex, its entropy is

$$S \approx k_B \log\left(\frac{A}{A_0}\right) + O(1). \tag{2.12}$$

The free energy then is

$$F \approx (\pi J - k_B T) \log\left(\frac{A}{A_0}\right). \tag{2.13}$$

At low temperatures, the free energy is dominated by the contribution of the energy. Since the free energy is minimized at equilibrium, the occurrence of free vortices is unfavorable. With increasing temperature, the entropy contribution to the free energy takes over and free vortices occur. This simple derivation completely neglects interactions between vortices. It, however, suggests that a suitable description of the system is to consider spin-wave excitations and vortices separately. In order to proceed, it is assumed that the system is large compared to the lattice spacing and that one can restrict oneself to configurations in which the angle between adjacent spins are small. After all, these are the configurations which contribute most to the partition function. For temperatures $T < T_c$, the configurations are determined by spin-wave excitations. Vortices are closely bound in vortex-anti-vortex pairs and hence their contribution to the configurations is negligible. For temperatures $T > T_c$, the configurations are determined by both, spin-wave excitations and vortices.

Let us consider the effects of the BKT-Transition onto the spin-spin correlation function (2.7). In the vicinity of the critical temperature, for $T < T_c$, spin-wave excitations contribute predominantly to the spin-spin correlation function. One finds for $|\mathbf{x} - \mathbf{y}| \gg a$ that

$$\langle \mathbf{s}_x \cdot \mathbf{s}_y \rangle \propto \left|\frac{\mathbf{x} - \mathbf{y}}{a}\right|^{-\eta(T)}, \qquad \text{where } \eta(T) = \frac{k_B T}{4\pi J}. \tag{2.14}$$

We see that although the spin-wave excitations are responsible for the destruction of any long-range order in the system, they are not responsible for the phase transition. At temperatures $T > T_c$ the spin-spin correlation function is found to be

$$\langle \mathbf{s}_x \cdot \mathbf{s}_y \rangle \propto \exp\left(-\frac{|\mathbf{x} - \mathbf{y}|}{\xi(T)}\right), \qquad \text{where } \xi(T) \propto \exp\left(b\sqrt{\frac{T_c}{T - T_c}}\right) \tag{2.15}$$

and b is some constant. Hence, the correlation length, defined through (2.8), is given by

$$\begin{aligned} \xi(T) &= \infty, & T < T_c \\ \xi(T) &\propto \exp\left(b\sqrt{\tfrac{T_c}{T - T_c}}\right), & T > T_c. \end{aligned} \tag{2.16}$$

Not only T_c but every temperature below T_c corresponds to a critical point. According to the simple argumentation above, the critical temperature is reached when the entropy term takes over, i.e. when $F = 0$ in (2.13). Hence, $T_c \approx \pi J/k_B$. If we insert this into (2.14), we find the critical exponent $\eta(T_c) = 1/4$. If we take into account that at the critical temperature, interactions between vortices cannot be neglected altogether, the critical temperature is lowered and so is $\eta(T_c)$. But now, the vortices also contribute to the critical exponent which brings it back to $\eta(T_c) = 1/4$.

2.4 Cluster representation of the XY model

A highly efficient Monte Carlo method for the simulation of the XY model is the Wolff cluster algorithm. This algorithm is based on the cluster representation of the XY model. Before proceeding to this representation of the XY model, it is instructive to first study the cluster representation of the Ising model. A derivation can be found in [7]. The cluster representation of the XY model can, to some extent, be reduced to that of the Ising model by an approach known as the Wolff embedding trick. The idea is that for every choice of direction, the projection of the XY model spins onto that direction form an Ising model with inhomogeneous coupling.

Let $\mathbf{n}_W \in S^1$ be some arbitrary vector. This vector defines the direction mentioned above. The plane, or rather the line, orthogonal to $\mathbf{n}_W$ is called the Wolff plane. The reflection $\mathbf{s}'$ of a spin $\mathbf{s} \in S^1$ at the Wolff plane is given by

$$\mathbf{s}' = \mathbf{s} - 2(\mathbf{s} \cdot \mathbf{n}_W)\mathbf{n}_W, \tag{2.17}$$

and we will refer to the transformation $\mathbf{s} \to \mathbf{s}'$ as flipping the spin. For the dot product of two spins, where one of both is flipped, we get

$$\mathbf{s}_x \cdot \mathbf{s}'_y = \mathbf{s}_x \cdot \mathbf{s}_y - 2(\mathbf{s}_x \cdot \mathbf{n}_W)(\mathbf{s}_y \cdot \mathbf{n}_W) = \mathbf{s}'_x \cdot \mathbf{s}_y. \tag{2.18}$$

From this, it follows that $\mathbf{s}'_x \cdot \mathbf{s}'_y = \mathbf{s}_x \cdot \mathbf{s}_y$.

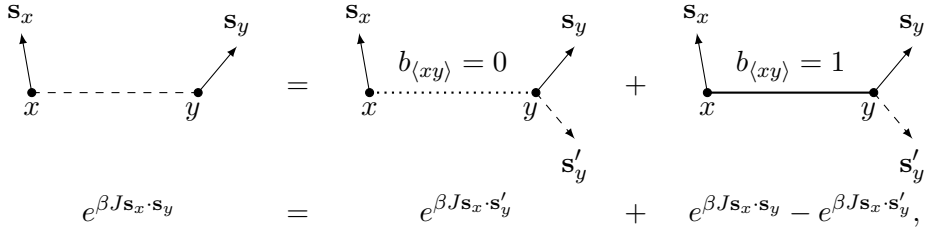
We begin to rewrite the XY model in terms of spin and bond variables by introducing the latter. Bond variables $b_{\langle xy \rangle}$ are variables that are defined on the connection between neighboring lattice sites. They can either be activated ($b_{\langle xy \rangle} = 1$) or deactivated ($b_{\langle xy \rangle} = 0$). Since the Hamilton function (2.1) is a sum of contributions from neighboring spins $\mathbf{s}_x, \mathbf{s}_y$, the Boltzmann factor $\exp(-\beta\mathcal{H}[\mathbf{s}])$ in the partition function may be written as

$$\exp(-\beta\mathcal{H}[\mathbf{s}]) = \exp(\beta J \sum_{\langle xy \rangle} \mathbf{s}_x \cdot \mathbf{s}_y) = \prod_{\langle xy \rangle} \exp(\beta J \mathbf{s}_x \cdot \mathbf{s}_y) = \prod_{\langle xy \rangle} \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y)), \tag{2.19}$$

where $h(\mathbf{s}_x, \mathbf{s}_y) = -J\mathbf{s}_x \cdot \mathbf{s}_y$. Thus, each pair of neighboring spins contributes a factor of $\exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y))$ to the partition function (2.3). These factors are now split into a contribution of the activated and a contribution of the deactivated bond connecting the nearest neighbors $\mathbf{s}_x$ and $\mathbf{s}_y$, i.e.

$$\exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y)) = \sum_{b_{\langle xy \rangle}=0,1} \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y, b_{\langle xy \rangle})). \tag{2.20}$$

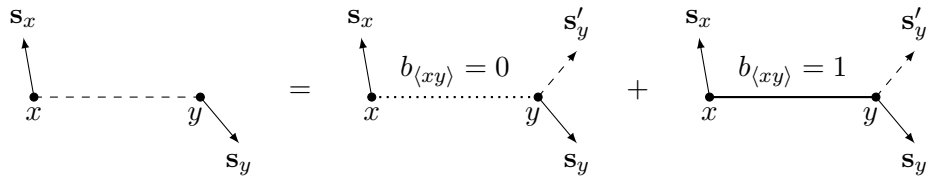
If the two spins are on the same side of the Wolff plane, that is, if $(\mathbf{s}_x \cdot \mathbf{n}_W)(\mathbf{s}_y \cdot \mathbf{n}_W) > 0$, then we split the Boltzmann factor $\exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y)) = \exp(\beta J \mathbf{s}_x \cdot \mathbf{s}_y)$ according to



(2.21)

where $\mathbf{s}'_y$ is the reflection of $\mathbf{s}_y$ at the Wolff plane. If, on the other hand, the two spins are on different sides of the Wolff plane, that is, if $(\mathbf{s}_x \cdot \mathbf{n}_W)(\mathbf{s}_y \cdot \mathbf{n}_W) \leq 0$, then we split the

Boltzmann factor as



(2.22)

The function $h(\mathbf{s}_x, \mathbf{s}_y, b_{\langle xy\rangle})$ is thus defined by

$$\exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y, b_{\langle xy\rangle})) = \begin{cases} \exp(\beta J \mathbf{s}_x \cdot \mathbf{s}'_y) & \text{if } (\mathbf{s}_x \cdot \mathbf{n}_W)(\mathbf{s}_y \cdot \mathbf{n}_W) > 0, b_{\langle xy\rangle} = 0, \\ \exp(\beta J \mathbf{s}_x \cdot \mathbf{s}_y) - \exp(\beta J \mathbf{s}_x \cdot \mathbf{s}'_y) & \text{if } (\mathbf{s}_x \cdot \mathbf{n}_W)(\mathbf{s}_y \cdot \mathbf{n}_W) > 0, b_{\langle xy\rangle} = 1, \\ \exp(\beta J \mathbf{s}_x \cdot \mathbf{s}_y) & \text{if } (\mathbf{s}_x \cdot \mathbf{n}_W)(\mathbf{s}_y \cdot \mathbf{n}_W) \leq 0, b_{\langle xy\rangle} = 0, \\ 0 & \text{if } (\mathbf{s}_x \cdot \mathbf{n}_W)(\mathbf{s}_y \cdot \mathbf{n}_W) \leq 0, b_{\langle xy\rangle} = 1. \end{cases} \tag{2.23}$$

In terms of spins and bonds the partition function becomes

$$\begin{aligned} Z(\beta) &= \int \mathcal{D}\mathbf{s}\ \exp(-\beta \mathcal{H}[\mathbf{s}]) = \int \mathcal{D}\mathbf{s} \prod_{\langle xy\rangle} \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y)) \\ &= \int \mathcal{D}\mathbf{s} \prod_{\langle xy\rangle} \sum_{b_{\langle xy\rangle}=0,1} \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y, b_{\langle xy\rangle})) \\ &= \int \mathcal{D}\mathbf{s} \sum_{[b]} \prod_{\langle xy\rangle} \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y, b_{\langle xy\rangle})) = \int \mathcal{D}\mathbf{s} \sum_{[b]} \exp(-\beta \mathcal{H}[\mathbf{s}, b]) \end{aligned} \tag{2.24}$$

where

$$\mathcal{H}[\mathbf{s}, b] = \sum_{\langle xy\rangle} h(\mathbf{s}_x, \mathbf{s}_y, b_{\langle xy\rangle}). \tag{2.25}$$

As we have only split the contribution of each pair of neighboring spins to the Boltzmann factor $\exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y))$ into a contribution of the activated and deactivated bond, the partition function remains unchanged. This is important to note since the partition function plays an essential role in the description of the system. In particular, it determines the probability density to find the spin and bond configuration $[\mathbf{s}, b]$:

$$\rho[\mathbf{s}, b] = \frac{1}{Z(\beta)} \exp(-\beta \mathcal{H}[\mathbf{s}, b]) \tag{2.26}$$

Since the introduction of bond variables does not alter the partition function, one expects to find the spin configuration $[\mathbf{s}]$ with probability density (2.5). As the following calculation shows, this is indeed the case

$$\begin{aligned} \sum_{[b]} \rho[\mathbf{s}, b] &= \frac{1}{Z(\beta)} \sum_{[b]} \prod_{\langle xy\rangle} \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y, b_{\langle xy\rangle})) \\ &= \frac{1}{Z(\beta)} \prod_{\langle xy\rangle} \sum_{b_{\langle xy\rangle}=0,1} \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y, b_{\langle xy\rangle})) \\ &= \frac{1}{Z(\beta)} \prod_{\langle xy\rangle} \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y)) = \frac{1}{Z(\beta)} \exp(-\beta \mathcal{H}[\mathbf{s}])) = \rho[\mathbf{s}]. \end{aligned} \tag{2.27}$$

The space of all spin and bond configurations is given by $(S^1)^{|\Lambda|} \times \{0,1\}^{2|\Lambda|}$. However, in contrast to the XY model described by spin configurations only, not all of these spin and bond configurations have a non-vanishing probability density. As follows from the definition of the function $h(\mathbf{s}_x, \mathbf{s}_y, b_{\langle xy \rangle})$ in (2.23), configurations where at least one activated bond connects spins on different sides of the Wolff plane, have probability density $\rho[\mathbf{s}, b] = 0$. The converse is also true. Furthermore, up to a null set of configurations in which at least one spin lies in the Wolff plane, each configuration, where at least one activated bond connects spins on different sides of the Wolff plane, has a neighborhood in which the probability density vanishes. Thus, we may conclude, that configurations with vanishing probability density have no influence on any observable of the system. Either, because they are part of a null set or because they have an environment in which the probability density vanishes. The configuration space is therefore the subset of $(S^1)^{|\Lambda|} \times \{0,1\}^{2|\Lambda|}$ of configurations with non-zero probability density. In all of these configurations, activated bonds only connect spins on the same side of the Wolff plane while deactivated bonds may connect both, spins on the same and on different sides of the Wolff plane.

Spins connected by activated bonds form regions of spins which are all on the same side of the Wolff-plane. These regions are called clusters. More precisely, a cluster can be defined as a subset of the lattice sites such that:

1. Each pair of lattice sites in the cluster is connected by a path on the lattice along activated bonds.
2. No lattice site outside the cluster is connected to one inside the cluster by an activated bond.

A lattice site that is not connected to any other lattice site by an activated bond forms a cluster by itself. In this way, each lattice site belongs to exactly one cluster. Let us denote the set of all clusters by $\mathcal{C}$. An example of a bond configuration and the resulting clusters is shown in Figure 2.4.

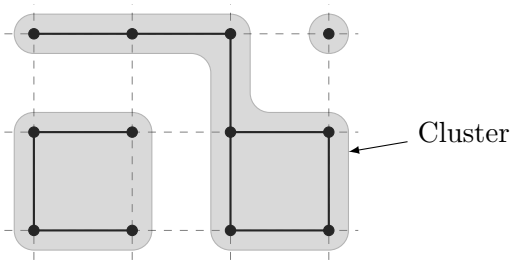


Figure 2.4: Clusters on the lattice. The solid lines represent activated bonds and the gray areas indicate the resulting clusters.

By definition, lattice sites of different clusters can only be connected by deactivated bonds. As follows from (2.23) by using (2.18), the Boltzmann factor of such a connection is invariant under spin flips of the two involved spins

$$\begin{aligned} \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y, 0)) &= \exp(-\beta h(\mathbf{s}_x, \mathbf{s}'_y, 0)) = \exp(-\beta h(\mathbf{s}'_x, \mathbf{s}_y, 0)) \\ &= \exp(-\beta h(\mathbf{s}'_x, \mathbf{s}'_y, 0)) = \exp(\beta J \mathbf{s}_x \cdot \mathbf{s}'_y) = \exp(\beta J \mathbf{s}'_x \cdot \mathbf{s}_y). \end{aligned} \tag{2.28}$$

Thus, if we flip an entire cluster $c \in \mathcal{C}$, that is, if we apply the transformation $\mathbf{s}_x \to \mathbf{s}'_x \; \forall x \in c$, the probability density of the configuration remains the same. We can do this with every

cluster $c \in \mathcal{C}$ which results in $2^{|\mathcal{C}|}$ configurations with the same probability density. Among these $2^{|\mathcal{C}|}$ configurations, there is one configuration in which $\mathbf{s}_x \cdot \mathbf{n}_W > 0 \ \forall x \in \Lambda$. This configuration is known as a reference configuration. Note that we are only allowed to flip entire clusters. If we were to flip any other subset of the lattice spins we would end up with a configuration in which at least one activated bond connects spins on different sides of the Wolff plane. Such a configuration is not part of the configuration space.

We may visualize the lattice spins as being composed of clusters of alike spins, where alike means that the spins are on the same side of the Wolff plane.

2.5 Semi-vortices

Let us consider some configuration $[\mathbf{s}, b]$ and the set of clusters $\mathcal{C}$, that results from the bond configuration. Within a cluster, all spins are on the same side of the Wolff plane. Hence, a vortex or anti-vortex cannot be contained in a single cluster. Instead, plaquettes with non-zero vorticity only occur between clusters, that is, the lattice sites of such a plaquette do not all belong to the same cluster.

Let us now consider the $2^{|\mathcal{C}|}$ configurations, obtained from $[\mathbf{s}, b]$ by independently flipping entire clusters, and a plaquette. The plaquette is chosen such that the lattice sites of the plaquette do not all belong to the same cluster. We call the clusters, the lattice sites of the plaquette belong to, the clusters of the plaquette. There is not necessarily a configuration among the $2^{|\mathcal{C}|}$ configurations in which this plaquette is a vortex or anti-vortex. However, if there is such a configuration in which the plaquette has vorticity $\nu_\square$, then we find, by flipping all clusters of the plaquette, a configuration in which the vorticity of the plaquette is $-\nu_\square$. Also, if we flip all clusters of the plaquette such that the spins of the clusters are on the same side of the Wolff plane, then $\nu_\square = 0$.

In general, a plaquette can have up to four clusters, which results in up to $2^4 = 16$ different configurations of the spins of the plaquette. Let us focus on those configurations where only one of the clusters is flipped out of the reference configuration of the plaquette, i.e. the configuration in which for all spins of the plaquette $\mathbf{s} \cdot \mathbf{n}_W > 0$. If in one of these configurations the vorticity of the plaquette is $+1$, then we call the spins of the plaquette that belong to the flipped cluster a positive semi-vortex of the plaquette and if the vorticity is -1, then we call them a negative semi-vortex. It turns out that, provided the plaquette is in one of the configurations of its spins, a vortex, there are always exactly two semi-vortices, one positive and one negative. The vorticity of the plaquette then only depends on those two clusters to which the positive and negative semi-vortex of the plaquette belong. The possible configuration of these two clusters and the resulting vorticity of the plaquette is given in Table 2.1. Flipping the remaining clusters, if there are any, has no influence on the vorticity

c_+	c_-	$\nu_\square$
$\uparrow$	$\uparrow$	0
$\uparrow$	$\downarrow$	-1
$\downarrow$	$\uparrow$	1
$\downarrow$	$\downarrow$	0

Table 2.1: Vorticity of a plaquette in dependence of the two clusters to which the positive (c_+) and negative (c_-) semi-vortex belong. $\uparrow$ means that the cluster c_+ or c_- are flipped such that for the spins of the cluster $\mathbf{s} \cdot \mathbf{n}_W > 0$ holds and $\downarrow$ such that $\mathbf{s} \cdot \mathbf{n}_W < 0$ holds.

of the plaquette.

3 Monte Carlo simulation of the XY model

The partition function (2.3) cannot be calculated numerically by means of brute force. An alternative is a so-called Monte Carlo method. This is an algorithm that numerically generates spin configurations from which one can estimate expectation values of observables.

3.1 The Monte Carlo method

In principle, we could generate spin configurations completely randomly and weight their contribution to an estimate of the expectation value of some observable by the Boltzmann factor $\exp(-\beta\mathcal{H}[\mathbf{s}])$. Such an approach, however, suffers from the drawback that we would mostly generate configurations with low weights. This problem can be addressed by importance sampling. The idea is to predominantly generate those configurations with the largest contributions to the partition function.

3.1.1 Markov chain

In a Monte Carlo simulation of the XY model, a sequence of spin configurations is generated by the application of some algorithm

$$[\mathbf{s}^{(1)}] \to [\mathbf{s}^{(2)}] \to \ldots \to [\mathbf{s}^{(i)}] \to [\mathbf{s}^{(i+1)}] \to \cdots \to [\mathbf{s}^{(N+M)}]. \tag{3.1}$$

If the configuration $[\mathbf{s}^{(i+1)}]$ is generated solely from its immediate predecessor $[\mathbf{s}^{(i)}]$, the sequence (3.1) is called a Markov chain and the algorithm by which this chain is generated can be described by a probability kernel $\omega([\mathbf{s}] \to [\tilde{\mathbf{s}}])$. Since the algorithm generates a new configuration $[\tilde{\mathbf{s}}]$ out of every configuration $[\mathbf{s}]$, the probability kernel is normalized as

$$\int \mathcal{D}\tilde{\mathbf{s}}\, \omega([\mathbf{s}] \to [\tilde{\mathbf{s}}]) = 1. \tag{3.2}$$

The probability kernel completely characterizes the algorithm used to turn the configuration $[\mathbf{s}^{(i)}]$ into $[\mathbf{s}^{(i+1)}]$.

The initial configuration $[\mathbf{s}^{(1)}]$ is either picked at random or chosen by some other means. The probability density to find the configuration $[\mathbf{s}]$ at this stage is $\rho_1[\mathbf{s}] = \delta([\mathbf{s}] - [\mathbf{s}^{(1)}])$. After one Monte Carlo iteration, that is, one application of the algorithm, the probability density is $\rho_2[\mathbf{s}] = \omega([\mathbf{s}^{(1)}] \to [\mathbf{s}])$. In general, after i Monte Carlo iterations, the probability density to find the configuration $[\mathbf{s}]$ is given by

$$\begin{aligned} \rho_{i+1}[\mathbf{s}_{i+1}] &= \int \mathcal{D}\mathbf{s}_i\; \rho_i([\mathbf{s}_i])\, \omega([\mathbf{s}_i] \to [\mathbf{s}_{i+1}]) \\ &= \left[\prod_{j=2}^{i} \int \mathcal{D}\mathbf{s}_j\right] \omega([\mathbf{s}^{(1)}] \to [\mathbf{s}_2]) \prod_{j=2}^{i} \omega([\mathbf{s}_j] \to [\mathbf{s}_{j+1}]). \end{aligned} \tag{3.3}$$

The questions which now arise are: does the sequence of probability densities ρ_i converge towards some (stationary) equilibrium distribution ρ, do they converge towards the same equilibrium distribution for any initial configuration and if so, is this distribution the Boltzmann distribution (2.5). As discussed in Section 3.1.2, this is indeed the case if the algorithm is ergodic and satisfies detailed balance with respect to the Boltzmann distribution. The process of converging towards the equilibrium distribution is depicted in Figure 3.1.

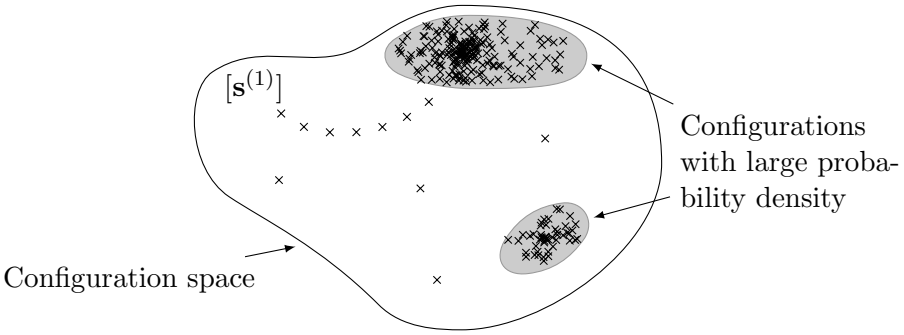


Figure 3.1: A Markov chain in configuration space. The crosses indicate the generated configurations. Those from $[\mathbf{s}^{(1)}]$ to one of the regions with large probability density, illustrate the equilibration process. Note, however, that just because the algorithm generates a configuration with large probability density, does not necessarily mean that equilibrium has been reached.

After a possibly large amount N of Monte Carlo iterations, an equilibrium distribution of the configurations is reached, or rather, the probability density ρ_N is sufficiently close to the equilibrium distribution ρ for us to start measuring observables. At this stage the Monte Carlo simulation generates mainly those configurations with the largest contributions to the partition function. The expectation value $\langle O \rangle$ of some observable O can then be estimated from the average over the configurations generated after equilibrium is reached

$$\bar{O} = \frac{1}{M} \sum_{i=N+1}^{N+M} O[\mathbf{s}^{(i)}], \tag{3.4}$$

where M is the number of measurements and $N + M$ is the total number of Monte Carlo iterations executed during the simulation. In the limit $M \to \infty$, this approximation becomes exact, i.e. $\langle O \rangle = \lim_{M\to\infty} \bar{O}$. Note that in (3.4), the contributions of the individual configurations in the Markov Chain are not weighted with the Boltzmann factor $\exp(-\beta\mathcal{H}[\mathbf{s}])$. This is not necessary since the configurations are generated in accordance with the Boltzmann distribution.

3.1.2 Ergodicity and Detailed Balance

A Monte Carlo method is only of use to us, if it reaches at some point an equilibrium distribution, if this distribution is the Boltzmann distribution (2.5) and if it is reached from any initial configuration. We can show that this is indeed the case if the algorithm is ergodic and obeys detailed balance with respect to the Boltzmann distribution. Note that the proof presented here does not rely on the fact that the equilibrium distribution is the Boltzmann distribution. We focus here on this distribution because we intend to simulate the XY model, but Monte Carlo methods can be applied to a variety of different problems.

Ergodicity means that, starting from an arbitrary configuration $[\mathbf{s}^{(1)}]$, any configuration $[\mathbf{s}]$ can be reached with non-vanishing probability density in a finite number of iterations. This implies that, although the set of all generated configurations is a null set in configuration space, every subset of configuration space will sooner or later be reached by the algorithm. This is apparently an important feature since the average of some observable over the generated configurations can only approach its expectation value, if configurations from all over configuration space are taken into account.

A Monte Carlo method obeys detailed balance with respect to the probability density ρ,

if it satisfies

$$\rho[\mathbf{s}]\,\omega([\mathbf{s}] \to [\tilde{\mathbf{s}}]) = \rho[\tilde{\mathbf{s}}]\,\omega([\tilde{\mathbf{s}}] \to [\mathbf{s}]), \tag{3.5}$$

for all configurations $[\mathbf{s}], [\tilde{\mathbf{s}}]$ in configuration space. The condition of detailed balance states that the probability kernel for a transition from a configuration with low probability density to one with high probability density must be larger than for the reverse transition.

Let us now show that a Monte Carlo method has an equilibrium distribution, if it satisfies detailed balance. Note that detailed balance is sufficient but not necessary for the Monte Carlo method to have an equilibrium distribution. Consider the map

$$\rho_{i+1}[\tilde{\mathbf{s}}] = \int \mathcal{D}\mathbf{s}\; \rho_i[\mathbf{s}]\,\omega([\mathbf{s}] \to [\tilde{\mathbf{s}}]). \tag{3.6}$$

If the sequence of probability densities ρ_i generated by this map approaches an equilibrium distribution ρ^*, then this equilibrium distribution must be a fixed-point of this map, that is,

$$\rho^*[\tilde{\mathbf{s}}] = \int \mathcal{D}\mathbf{s}\; \rho^*[\mathbf{s}]\,\omega([\mathbf{s}] \to [\tilde{\mathbf{s}}]). \tag{3.7}$$

If we assume that the algorithm satisfies detailed balance with respect to some probability density ρ, then we find, using the normalization of the probability kernel (3.2), that

$$\begin{aligned} \int \mathcal{D}\mathbf{s}\; \rho[\mathbf{s}]\,\omega([\mathbf{s}] \to [\tilde{\mathbf{s}}]) &= \int \mathcal{D}\mathbf{s}\; \rho[\tilde{\mathbf{s}}]\,\omega([\tilde{\mathbf{s}}] \to [\mathbf{s}]) \\ &= \rho[\tilde{\mathbf{s}}] \int \mathcal{D}\mathbf{s}\; \omega([\tilde{\mathbf{s}}] \to [\mathbf{s}]) = \rho[\tilde{\mathbf{s}}]. \end{aligned} \tag{3.8}$$

Thus, ρ is stationary, that is, it is an equilibrium distribution. One can then show, using ergodicity, that ρ is unique and that the sequence of probability densities ρ_i converges to the equilibrium distribution ρ for any initial configuration $[\mathbf{s}^{(1)}]$.

3.2 The Wolff cluster algorithm

The Wolff cluster algorithm is a Monte Carlo method based on the cluster representation of the XY model (see Section 2.4). It generalizes the Swendsen-Wang algorithm for the Ising model to the XY model. In this Thesis, we use the multi-cluster variant of the algorithm. Given some spin configuration $[\mathbf{s}^{(i)}]$, the subsequent configuration $[\mathbf{s}^{(i+1)}]$ in the Markov chain is generated as follows.

1. First a vector $\mathbf{n}_W \in S^1$ is chosen randomly. This vector is the normal vector to the Wolff plane and spins will be flipped with respect to this plane. Thus, only components of the spins parallel to $\mathbf{n}_W$ will be changed in this iteration.

2. Each bond $b_{\langle xy\rangle}$, connecting the spins $\mathbf{s}_x^{(i)}$ and $\mathbf{s}_y^{(i)}$ is activated or deactivated based on the orientation of the two spins in relation to the Wolff plane. If the two spins are on different sides of the Wolff plane, that is, if $(\mathbf{s}_x^{(i)} \cdot \mathbf{n}_W)(\mathbf{s}_y^{(i)} \cdot \mathbf{n}_W) \leq 0$, then the bond is deactivated. If, on the other hand, the two spins are on the same side of the Wolff plane, that is, if $(\mathbf{s}_x^{(i)} \cdot \mathbf{n}_W)(\mathbf{s}_y^{(i)} \cdot \mathbf{n}_W) > 0$, then the bond is activated with probability

$$\begin{aligned} p_{\langle xy\rangle} &= \frac{\exp(-\beta h(\mathbf{s}_x^{(i)}, \mathbf{s}_y^{(i)}, 1))}{\exp(-\beta h(\mathbf{s}_x^{(i)}, \mathbf{s}_y^{(i)}, 0)) + \exp(-\beta h(\mathbf{s}_x^{(i)}, \mathbf{s}_y^{(i)}, 1))} \\ &= 1 - \exp(-2\beta J(\mathbf{s}_x^{(i)} \cdot \mathbf{n}_W)(\mathbf{s}_y^{(i)} \cdot \mathbf{n}_W)) \end{aligned} \tag{3.9}$$

   and deactivated with probability $1 - p_{\langle xy\rangle}$.

3. Once all bonds are either activated or deactivated, the clusters can be identified. Each cluster is then flipped with probability 1/2. This completes one sweep of the algorithm.

The Wolff cluster algorithm generates a Markov chain of spin configurations with equilibrium distribution (2.5). However, at each sweep of the algorithm a bond configuration, which is consistent with the spin configuration after flipping the clusters, i.e. no activated bond connects spins on different sides of the Wolff plane, is generated. So the algorithm generates a Markov chain of spin and bond rather than just spin configurations. The information contained in the bond configurations can be used to evaluate improved estimators (see Section 3.3) but in order to do so, we must make sure that the equilibrium distribution of the spin and bond configurations is (2.26). Let us consider a spin configuration $[\mathbf{s}^{(i)}]$ generated by the algorithm at equilibrium, i.e. with probability density

$$\rho[\mathbf{s}^{(i)}] = \frac{1}{Z(\beta)} \prod_{\substack{\langle xy\rangle \mid \\ (\mathbf{s}_x^{(i)}\cdot\mathbf{n}_W)(\mathbf{s}_y^{(i)}\cdot\mathbf{n}_W)\leq 0}} \exp(\beta J \mathbf{s}_x^{(i)}\cdot\mathbf{s}_y^{(i)}) \prod_{\substack{\langle xy\rangle \mid b_{\langle xy\rangle}=0, \\ (\mathbf{s}_x^{(i)}\cdot\mathbf{n}_W)(\mathbf{s}_y^{(i)}\cdot\mathbf{n}_W)>0}} \exp(\beta J \mathbf{s}_x^{(i)}\cdot\mathbf{s}_y^{(i)}) \prod_{\substack{\langle xy\rangle \mid b_{\langle xy\rangle}=1, \\ (\mathbf{s}_x^{(i)}\cdot\mathbf{n}_W)(\mathbf{s}_y^{(i)}\cdot\mathbf{n}_W)>0}} \exp(\beta J \mathbf{s}_x^{(i)}\cdot\mathbf{s}_y^{(i)}) \tag{3.10}$$

The probability to generate the bond configuration $[b]$ given the spin configuration $[\mathbf{s}^{(i)}]$ is

$$p([b]|[\mathbf{s}^{(i)}]) = \prod_{\substack{\langle xy\rangle \mid b_{\langle xy\rangle}=0, \\ (\mathbf{s}_x^{(i)}\cdot\mathbf{n}_W)(\mathbf{s}_y^{(i)}\cdot\mathbf{n}_W)>0}} (1-p_{\langle xy\rangle}) \prod_{\substack{\langle xy\rangle \mid b_{\langle xy\rangle}=1, \\ (\mathbf{s}_x^{(i)}\cdot\mathbf{n}_W)(\mathbf{s}_y^{(i)}\cdot\mathbf{n}_W)>0}} p_{\langle xy\rangle}. \tag{3.11}$$

For the product of the corresponding factors in $\rho[\mathbf{s}^{(i)}]$ and $p([b]|[\mathbf{s}^{(i)}])$ with $b_{\langle xy\rangle} = 0$ and $(\mathbf{s}_x^{(i)}\cdot\mathbf{n}_W)(\mathbf{s}_y^{(i)}\cdot\mathbf{n}_W) > 0$ we find

$$\begin{aligned} \exp(\beta J \mathbf{s}_x^{(i)}\cdot\mathbf{s}_y^{(i)})(1-p_{\langle xy\rangle}) &= \exp(\beta J(\mathbf{s}_x^{(i)}\cdot\mathbf{s}_y^{(i)} - 2(\mathbf{s}_x^{(i)}\cdot\mathbf{n}_W)(\mathbf{s}_y^{(i)}\cdot\mathbf{n}_W))) \\ &= \exp(\beta J \mathbf{s}_x^{(i)}\cdot\mathbf{s}_y^{(i)\prime}) = \exp(-\beta h(\mathbf{s}_x^{(i)}, \mathbf{s}_y^{(i)}, 0)), \end{aligned} \tag{3.12}$$

and for those with $b_{\langle xy\rangle} = 1$ and $(\mathbf{s}_x^{(i)}\cdot\mathbf{n}_W)(\mathbf{s}_y^{(i)}\cdot\mathbf{n}_W) > 0$ we find

$$\begin{aligned} \exp(\beta J \mathbf{s}_x^{(i)}\cdot\mathbf{s}_y^{(i)})\, p_{\langle xy\rangle} &= \exp(\beta J \mathbf{s}_x^{(i)}\cdot\mathbf{s}_y^{(i)}) - \exp(\beta J \mathbf{s}_x^{(i)}\cdot\mathbf{s}_y^{(i)\prime}) \\ &= \exp(-\beta h(\mathbf{s}_x^{(i)}, \mathbf{s}_y^{(i)}, 1)). \end{aligned} \tag{3.13}$$

For the remaining factors in $\rho[\mathbf{s}^{(i)}]$ it clearly holds that

$$\exp(\beta J \mathbf{s}_x^{(i)}\cdot\mathbf{s}_y^{(i)}) = \exp(-\beta h(\mathbf{s}_x^{(i)}, \mathbf{s}_y^{(i)}, 0)). \tag{3.14}$$

So in total we find that

$$\rho[\mathbf{s}^{(i)}]\, p([b]|[\mathbf{s}^{(i)}]) = \frac{1}{Z(\beta)} \prod_{\langle xy\rangle} \exp(-\beta h(\mathbf{s}_x^{(i)}, \mathbf{s}_y^{(i)}, b_{\langle xy\rangle})) = \rho[\mathbf{s}^{(i)}, b] = \rho[\mathbf{s}^{(i+1)}, b], \tag{3.15}$$

where in the last step we used that the spin configurations $[\mathbf{s}^{(i)}]$ and $[\mathbf{s}^{(i+1)}]$ only differ by flipping entire clusters, and we have executed this step because we will take the configuration $[\mathbf{s}^{(i+1)}, b]$ as an element of the Markov chain. Hence, we see that at equilibrium, the Wolff cluster algorithm generates the spin and bond configurations with probability density (2.26) as desired.

Let us now show that the Wolff cluster algorithm is ergodic and obeys detailed balance with respect to the probability density (2.5). Ergodicity is guaranteed by the fact that there

is always a non-vanishing probability that each lattice site forms a cluster by itself and that only one of them is flipped. Since for any two spins, there is always a Wolff plane such that one spin can be flipped into the other, we can reach, at least in principle, any spin configuration in at most $|\Lambda|$ sweeps.

To show that the algorithm obeys detailed balance with respect to the probability density (2.5), we must show that it satisfies the condition (3.5) for any two spin configurations $[\mathbf{s}]$ and $[\tilde{\mathbf{s}}]$ in configuration space. First of all, if $[\mathbf{s}] = [\tilde{\mathbf{s}}]$, then detailed balance is clearly satisfied. Also, if there are two different lattice sites such that flipping $\mathbf{s}_x$ into $\tilde{\mathbf{s}}_x$ requires a different Wolff plane than flipping $\mathbf{s}_y$ into $\tilde{\mathbf{s}}_y$, then $\omega([\mathbf{s}] \to [\tilde{\mathbf{s}}]) = \omega([\tilde{\mathbf{s}}] \to [\mathbf{s}]) = 0$ and detailed balance is satisfied. Thus, we can assume that there is one Wolff plane such that we can flip $[\mathbf{s}]$ into $[\tilde{\mathbf{s}}]$ and vice versa. It is sufficient to only consider one pair of neighboring spins. We distinguish between the following cases based on the orientation of the two spins in relation to the Wolff plane.

1. The two spins are on different sides of the Wolff plane in $[\mathbf{s}]$. They necessarily belong to different clusters and the bond between them is deactivated with probability 1.

   (a) The two spins are on different sides of the Wolff plane in $[\tilde{\mathbf{s}}]$. They also belong to different clusters in $[\tilde{\mathbf{s}}]$ and the bond between them is deactivated with probability 1. From $[\mathbf{s}]$ we reach $[\tilde{\mathbf{s}}]$ with probability $1/2$ and with the same probability we reach $[\mathbf{s}]$ from $[\tilde{\mathbf{s}}]$, thus we find

   $$\begin{aligned} [\mathbf{s}] \to [\tilde{\mathbf{s}}] : &\qquad \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y))\frac{1}{2} = \exp(\beta J \mathbf{s}_x \cdot \mathbf{s}_y)\frac{1}{2} \\ [\tilde{\mathbf{s}}] \to [\mathbf{s}] : &\qquad \exp(-\beta h(\tilde{\mathbf{s}}_x, \tilde{\mathbf{s}}_y))\frac{1}{2} = \exp(\beta J \tilde{\mathbf{s}}_x \cdot \tilde{\mathbf{s}}_y)\frac{1}{2}. \end{aligned} \tag{3.16}$$

   Since either $\tilde{\mathbf{s}}_x = \mathbf{s}_x$, $\tilde{\mathbf{s}}_y = \mathbf{s}_y$ or $\tilde{\mathbf{s}}_x = \mathbf{s}'_x$, $\tilde{\mathbf{s}}_y = \mathbf{s}'_y$, both factors are, according to (2.18), the same and detailed balance is satisfied.

   (b) The two spins are on the same side of the Wolff plane in $[\tilde{\mathbf{s}}]$. We reach $[\tilde{\mathbf{s}}]$ from $[\mathbf{s}]$ with probability $1/2$. From $[\tilde{\mathbf{s}}]$ we can only reach $[\mathbf{s}]$, if the two spins are in different clusters. For this to happen, the bond between the two spins must be deactivated, which happens with probability $1 - p_{\langle xy \rangle}$. We then return to $[\mathbf{s}]$ with probability $1/2$. Thus, we find

   $$\begin{aligned} [\mathbf{s}] \to [\tilde{\mathbf{s}}] : &\qquad \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y))\frac{1}{2} = \exp(\beta J \mathbf{s}_x \cdot \mathbf{s}_y)\frac{1}{2} \\ [\tilde{\mathbf{s}}] \to [\mathbf{s}] : &\qquad \exp(-\beta h(\tilde{\mathbf{s}}_x, \tilde{\mathbf{s}}_y))(1 - p_{\langle xy \rangle})\frac{1}{2} = \exp(\beta J \tilde{\mathbf{s}}_x \cdot \tilde{\mathbf{s}}'_y)\frac{1}{2}. \end{aligned} \tag{3.17}$$

   Since either $\tilde{\mathbf{s}}_x = \mathbf{s}'_x$, $\tilde{\mathbf{s}}_y = \mathbf{s}_y$ or $\tilde{\mathbf{s}}_x = \mathbf{s}_x$, $\tilde{\mathbf{s}}_y = \mathbf{s}'_y$, both factors are the same and detailed balance is satisfied.

2. The two spins are on the same sides of the Wolff plane in $[\mathbf{s}]$.

   (a) The two spins are on different sides of the Wolff plane in $[\tilde{\mathbf{s}}]$. This case has been discussed above but with $\mathbf{s}_x$, $\mathbf{s}_y$ replaced by $\tilde{\mathbf{s}}_x$, $\tilde{\mathbf{s}}_y$ and vice versa. Hence, detailed balance is satisfied in this case.

   (b) The two spins are on the same sides of the Wolff plane in $[\tilde{\mathbf{s}}]$. We reach $[\tilde{\mathbf{s}}]$ from $[\mathbf{s}]$ in one of three ways. If the two spins are connected through other activated bonds, they belong to the same cluster. In this case it doesn't matter if the bond between them is activated or not, and we reach $[\tilde{\mathbf{s}}]$ with probability 1. For the

other two cases, we assume that the two spins are not connected through other activated bonds. If the bond between the two spins is activated, they belong to the same cluster, and we reach $[\tilde{\mathbf{s}}]$ with probability $p_{\langle xy\rangle}$. If, on the other hand, the bond between the two spins is deactivated, they belong to different clusters. We then reach $[\tilde{\mathbf{s}}]$ with probability $(1-p_{\langle xy\rangle})\frac{1}{2}$. The situation is the same for the reverse transition, hence we find

$$\begin{aligned} [\mathbf{s}] \to [\tilde{\mathbf{s}}]: &\qquad \exp(-\beta h(\mathbf{s}_x, \mathbf{s}_y)) \left(1 + p_{\langle xy\rangle} + (1-p_{\langle xy\rangle})\frac{1}{2}\right) \\ [\tilde{\mathbf{s}}] \to [\mathbf{s}]: &\qquad \exp(-\beta h(\tilde{\mathbf{s}}_x, \tilde{\mathbf{s}}_y)) \left(1 + p_{\langle xy\rangle} + (1-p_{\langle xy\rangle})\frac{1}{2}\right) \end{aligned} \tag{3.18}$$

.

Since $\mathbf{s}'_x \cdot \mathbf{s}'_y = \mathbf{s}_x \cdot \mathbf{s}_y$ and either $\tilde{\mathbf{s}}_x = \mathbf{s}_x$, $\tilde{\mathbf{s}}_y = \mathbf{s}_y$ or $\tilde{\mathbf{s}}_x = \mathbf{s}'_x$, $\tilde{\mathbf{s}}_y = \mathbf{s}'_y$, both factors are the same and detailed balance is satisfied.

In all cases, detailed balance is satisfied and thus, the algorithm obeys detailed balance with respect to the Boltzmann distribution.

3.3 Improved estimators

The Markov chain generated by the Wolff cluster algorithm is a chain of spin and bond configurations. While we could just ignore the bond configurations and estimate expectation values of observables from the generated spin configurations only, using both allows us to estimate them much more efficiently.

Consider the expression for the expectation value of an observable

$$\langle O\rangle := \int \mathcal{D}\mathbf{s}\ \rho[\mathbf{s}]\ O[\mathbf{s}]. \tag{3.19}$$

The expectation value is estimated from the average over the measured values $O[\mathbf{s}^{(i)}]$, as given by equation (3.4). If we can evaluate (3.19) in parts exactly, then chances are that we find a quantity I, called an improved estimator for the observable O, which has the same expectation value as the observable but with a smaller standard deviation in the estimate of the expectation value from its average

$$\bar{I} = \frac{1}{M}\sum_{i=N+1}^{N+M} I[\mathbf{s}^{(i)}, b^{(i)}]. \tag{3.20}$$

As in (3.4), M is the number of measurements and N denotes the Monte Carlo iteration at which we assume the probability density ρ_N to be sufficiently close to the Boltzmann distribution.

To be a little more concrete about how an improved estimator could be found, consider the $2^{|\mathcal{C}|}$ configurations obtained from a configuration $[\mathbf{s}, b]$ with $|\mathcal{C}|$ clusters by independently flipping entire clusters. Since all of these configurations have the same probability density, an observable which has canceling contributions among these configurations, can be improved by defining a quantity that only measures the non-canceling contributions. If we simulate long enough, all of these $2^{|\mathcal{C}|}$ configurations will be generated, hence, the improved estimator and the observable have the same expectation value.

The concept of improved estimators for observables will be much more clear once we derive some of them in the next chapter.

3.4 Error analysis

The measurement of an observable, or a corresponding improved estimator, O during a Monte Carlo simulation at equilibrium generates a sample of data $x_i = O[\mathbf{s}^{(N+i)}, b^{(N+i)}]$, $i = 1, \ldots, M$. From this data sample we would like to estimate the expectation value of the observable. One problem we are facing is, that in a Monte Carlo simulation, the configurations are generated from their immediate predecessors. The data in the sample will, therefore, not be statistically independent of each other. This problem can be addressed by binning the data. Furthermore, we might not just be interested in some directly observable quantity but in a function of expectation values of observables. A systematic method to estimate the expectation value of such a function is known as jackknife.

3.4.1 Basic analysis

Suppose we have a sample x_i, $i = 1, \ldots, M$ of M data points which we assume to be statistically independent of each other. The data points x_i are not all equal but rather drawn according to some normalized distribution $P(x)$, which is characterized by the mean μ and the variance σ^2

$$\mu = \langle x \rangle, \quad \sigma^2 = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2, \quad \text{where } \langle x^n \rangle = \int_{-\infty}^{\infty} x^n P(x) dx. \tag{3.21}$$

The standard deviation σ is defined as the square root of the variance. The data sample can be characterized by the sample mean $\bar{x}$ and sample standard deviation s, defined by

$$\bar{x} = \frac{1}{M} \sum_{i=1}^{M} x_i, \qquad s^2 = \frac{1}{M} \sum_{i=1}^{M} (x_i - \bar{x})^2. \tag{3.22}$$

Note that we denote an average over an exact distribution by angular brackets and an average over a sample of data by an over-bar.

Suppose we could repeat the measurement of the data sample many times. From the number of times we would find a sample mean $\bar{x}$ in $d\bar{x}$, we could construct a distribution $\tilde{P}(\bar{x})$. This distribution is characterized by the (exact) average of the sample mean $\langle \bar{x} \rangle$ and the standard deviation in the sample mean $\sigma_{\bar{x}}$. It can be shown that these are related to the sample mean and the sample variance by

$$\mu = \langle x \rangle = \langle \bar{x} \rangle, \qquad \sigma_{\bar{x}}^2 = \frac{1}{M-1} \langle s^2 \rangle. \tag{3.23}$$

Thus, the exact average of the sample mean is the exact average of x. Hence, the mean value $\bar{x}$, obtained from one sample of data, typically deviates from its exact average $\mu = \langle x \rangle$ by an amount of order $\sigma_{\bar{x}}$. Equation (3.23) shows that the best estimate of the exact average μ is $\bar{x}$ and the best estimate for the standard deviation $\sigma_{\bar{x}}$ is

$$\frac{s}{\sqrt{M-1}} = \left(\frac{1}{M(M-1)} \sum_{i=1}^{M} (x_i - \bar{x})^2 \right)^{1/2}. \tag{3.24}$$

For this estimate of the standard deviation to be useful, we need to know the probability for μ to lie within $\bar{x} \pm \frac{s}{\sqrt{M-1}}$. If for the distribution of x, μ and σ are finite, then the central limit theorem tells us that for large M, the distribution of $\bar{x}$ is Gaussian.

3.4.2 Jackknife

The method discussed in the previous section also applies to linear combinations of different average values $\mu_x, \mu_y, \ldots$. For a non-linear function f of averages, however, this method does not produce reliable results because of potential cross-correlations in the random variables $x, y, z, \ldots$. For simplicity, we assume that f is a non-linear function of two average values μ_x and μ_y, but the generalization to more than two is immediate. From the two data samples x_i and y_i, $i = 1, \ldots M$, we define the i-th jackknife estimate x_i^J, $i = 1, \ldots, M$, to be

$$x_i^J := \frac{1}{M-1} \sum_{j=1, j\neq i}^{M} x_j = \bar{x} + \frac{1}{M-1}(\bar{x} - x_i) \tag{3.25}$$

and analogous for y_i. The jackknife estimates are the sums over all but one data point in the sample. We also define the corresponding jackknife estimate of the function f, f_i^J, and the overall jackknife estimate $\overline{f^J}$ of $f(\mu_x, \mu_y)$ which is the average over the M jackknife estimates f_i^J:

$$f_i^J := f(x_i^J, y_i^J), \qquad \overline{f^J} := \frac{1}{M} \sum_{i=1}^{M} f_i^J. \tag{3.26}$$

To find the standard deviation in f, we define the variance in the jackknife average values by

$$s_{f^J}^2 := \overline{(f^J)^2} - \left(\overline{f^J}\right)^2, \qquad \text{where } \overline{(f^J)^2} = \frac{1}{M} \sum_{i=1}^{M} (f_i^J)^2. \tag{3.27}$$

It can be shown that $f(\mu_x, \mu_y)$ and σ_f are related to $f(\bar{x}, \bar{y})$, $\overline{f^J}$ and s_{f^J} by

$$f(\mu_x, \mu_y) = M\langle f(\bar{x}, \bar{y})\rangle - (M-1)\langle \overline{f^J}\rangle + O\left(\frac{1}{M^2}\right), \qquad \sigma_f^2 = (M-1)s_{f^J}^2. \tag{3.28}$$

This shows that one should estimate $f(\mu_x, \mu_y)$ from $Mf(\bar{x}, \bar{y}) - (M-1)\overline{f^J}$ and σ_f from $\sqrt{M-1}s_{f^J}$. If, however, M is sufficiently large, one can also estimate $f(\mu_x, \mu_y)$ from either $f(\bar{x}, \bar{y})$ or $\overline{f^J}$.

3.4.3 Binning

Subsequent configurations, generated by a Monte Carlo method, are not statistically independent of each other, and neither are the measured values of observables. However, measurements, that differ by a sufficiently large number of Monte Carlo sweeps, are independent. This is exploited when binning the measurements. Rather than considering each individual measurement, we consider average values of subsequent measurements, called bins. If we choose the bin size large enough, the average values of measurements in subsequent bins are practically independent. Although there are methods to estimate the bin size required to ensure statistically independent bin averages, we can circumvent this problem by only binning pairs of subsequent measurements

$$\underbrace{x_1, x_2}_{B_1}, \underbrace{x_3, x_4}_{B_2}, \ldots, \underbrace{x_{2i-1}, x_{2i}}_{B_i}, \underbrace{x_{2(i+1)-1}, x_{2(i+1)}}_{B_{i+1}}, \ldots, \underbrace{x_{M-3}, x_{M-2}}_{B_{M/2-1}}, \underbrace{x_{M-1}, x_M}_{B_{M/2}}, \tag{3.29}$$

where B_i denotes the average over the two measurements x_{2i-1} and x_{2i}. By repeating this procedure, but now applied on the bins B_i, we double the bin size. If we restrict the number of

measurements to $M = c\,2^m$, with c and m some positive integers, we can execute a total of m binning steps. Since subsequent measurements are not independent, the standard deviation in the mean of the measurements underestimates the true statistical error. By binning the data, the average values over the measurements of subsequent bins become more and more independent and the standard deviation in the mean of the bin averages increases, as shown in Figure 3.2. If we proceed, the average values over the measurements of subsequent bins

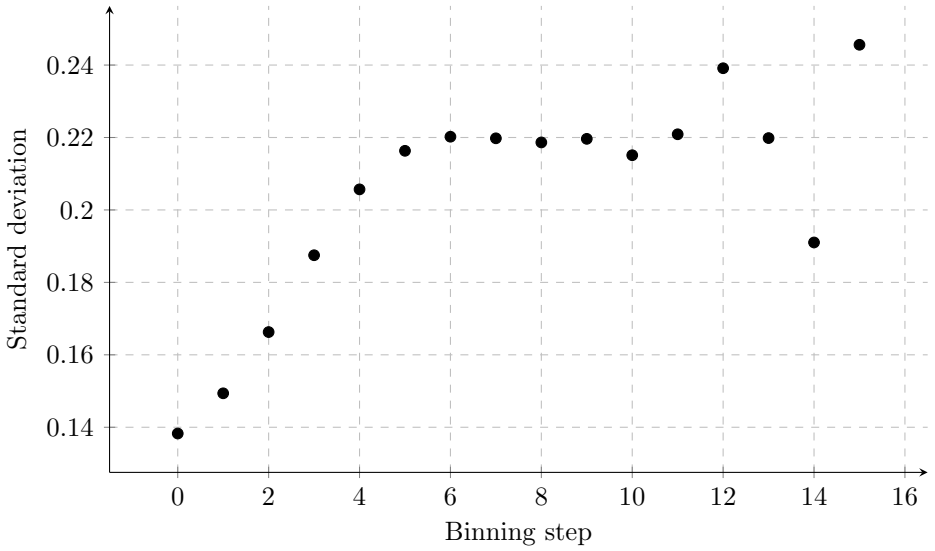


Figure 3.2: Standard deviation in the means of the bins as a function of the binning step. After six binning steps a plateau is reached. The value of the standard deviation at this plateau then is a good estimate of the true statistical error.

become, sooner or later, independent and the standard deviation reaches a plateau. The value of the standard deviation at this plateau then is a good estimate for the true statistical error. If we continue with binning, the number of bins becomes too small for a reliable calculation of the standard deviation.

4 Simulation results

For the simulation we use a lattice of $|\Lambda| = L_1 L_2$ sites, where $L_i,\ i = 1, 2$ are the number of lattice sites in the two directions. It is convenient for the discussion in this section to introduce coordinates on the lattice such that the coordinate axes are aligned with the lattice and the lattice sites x take coordinates in $\mathbf{x} \in \{1, \ldots, L_1\} \times \{1, \ldots, L_2\}$. Since we use a square lattice, there are just as many plaquettes as there are lattice sites. Each plaquette $\square_u$ is identified with the dual lattice site u at its center which takes values in $\mathbf{u} \in \{\frac{3}{2}, \ldots, L_1 + \frac{1}{2}\} \times \{\frac{3}{2}, \ldots, L_2 + \frac{1}{2}\}$.

The inverse temperature β for which $J\beta = 1.12$ is expected to be close to the critical inverse temperature β_c [9]. For the simulation results presented in this chapter we have used $J = 1$ as coupling constant, hence, $\beta_c \approx 1.12$.

4.1 Cluster sizes

The Wolff cluster algorithm builds clusters in each spin configuration by activating bonds with probability (3.9) between neighboring spins. The probability for activating a bond depends on the projection of the two spins onto the Wolff plane normal. Consider a region where the spins are nearly orthogonal to the Wolff plane normal. The probability to activate a bond is low and we typically find many small clusters in such a region. On the other hand, if the spins are close to being parallel or anti-parallel to the Wolff plane normal, then the probability to activate a bond is large. The clusters then typically grow until they reach a surface along which the projection of the spins onto the Wolff plane normal is small. The probability to activate a bond also depends on the temperature. At low temperatures, the probability is larger than at high temperatures.

At low temperatures the system is calm, that is, angles between neighboring spins are small. Typically, one then find one large along with several small clusters. With increasing temperature, the system becomes more turbulent. A typical cluster configuration then consists of many small clusters. This can be seen from the expectation value of the mean cluster size

$$\left\langle \overline{|c|} \right\rangle = \left\langle \frac{1}{|\mathcal{C}|} \sum_{c \in \mathcal{C}} |c| \right\rangle = \left\langle \frac{|\Lambda|}{|\mathcal{C}|} \right\rangle, \tag{4.1}$$

however, a more complete picture is provided by the relative frequency of cluster sizes. For two different temperatures, the relative frequency of cluster sizes is shown in Figure 4.1.

4.2 Magnetic susceptibility

The magnetic susceptibility is defined by

$$\chi := \frac{1}{L_1 L_2} \left(\langle |\boldsymbol{\mathcal{M}}|^2 \rangle - \langle \boldsymbol{\mathcal{M}} \rangle^2 \right), \tag{4.2}$$

where $\boldsymbol{\mathcal{M}}[\mathbf{s}] := \sum_{x \in \Lambda} \mathbf{s}_x$ is the magnetization. Since the XY model is invariant under $O(2)$ transformations, the spin configurations $[\mathbf{s}]$ and $[\tilde{\mathbf{s}}]$, where $[\tilde{\mathbf{s}}]$ is obtained from $[\mathbf{s}]$ by rotating every spin by an angle of π, have the same probability density but opposite magnetization, i.e. $\boldsymbol{\mathcal{M}}[\tilde{\mathbf{s}}] = -\boldsymbol{\mathcal{M}}[\mathbf{s}]$. As a consequence, $\langle \boldsymbol{\mathcal{M}} \rangle = 0$ and the magnetic susceptibility becomes

$$\chi = \frac{1}{L_1 L_2} \langle |\boldsymbol{\mathcal{M}}|^2 \rangle. \tag{4.3}$$

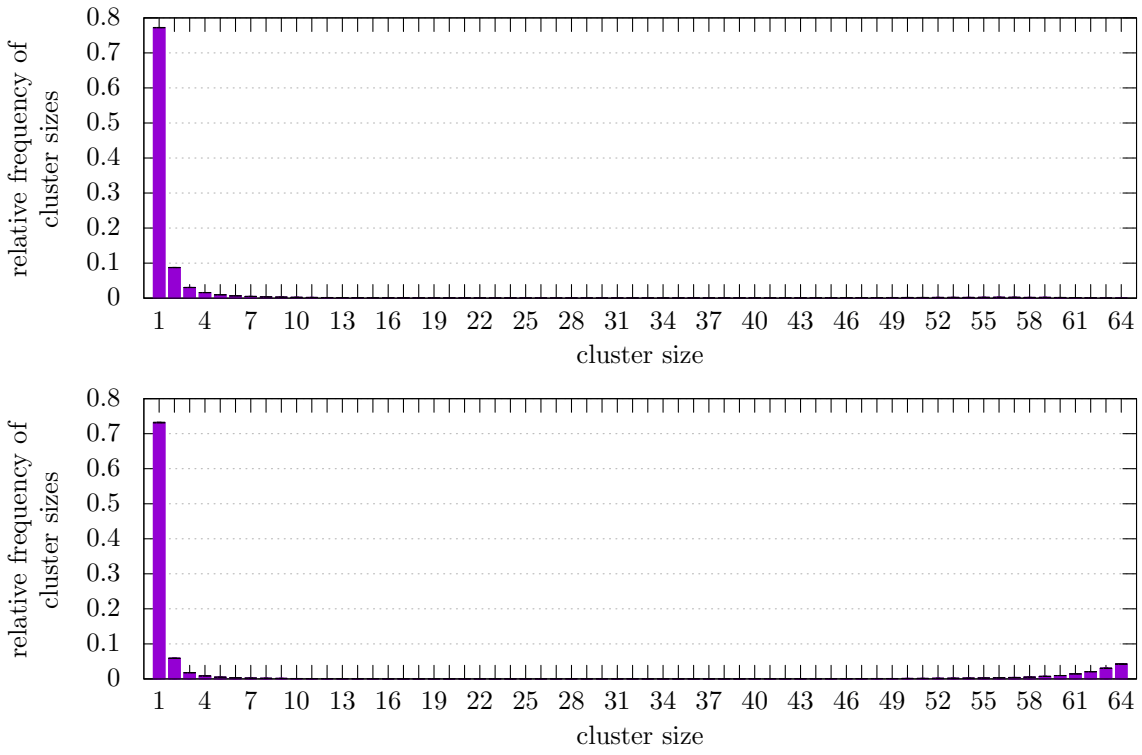


Figure 4.1: Relative frequency of cluster sizes for a system of $|\Lambda| = 8^2$ lattice sites at $\beta = 0.9$ (top) and $\beta = 1.3$ (bottom). 10′000 sweeps were executed to equilibrate the system. After that 65′536 measurements were taken.

The decomposition of the lattice spins into clusters of alike spins allows us to derive an improved estimator for the magnetic susceptibility. For this purpose note that the magnetization can be written as a sum of cluster magnetizations

$$\boldsymbol{\mathcal{M}}[\mathbf{s}] := \sum_{x\in\Lambda} \mathbf{s}_x = \sum_{c\in\mathcal{C}} \boldsymbol{\mathcal{M}}_c[\mathbf{s}], \qquad \text{where } \boldsymbol{\mathcal{M}}_c[\mathbf{s}] := \sum_{x\in c} \mathbf{s}_x. \tag{4.4}$$

The magnetization can be decomposed into a component orthogonal and one parallel to the Wolff plane normal

$$\boldsymbol{\mathcal{M}} = \boldsymbol{\mathcal{M}}_\parallel + \boldsymbol{\mathcal{M}}_\perp = \mathcal{M}_\parallel \mathbf{n}_W + \mathcal{M}_\perp \mathbf{t}_W, \quad \text{where } \mathcal{M}_\parallel = \boldsymbol{\mathcal{M}} \cdot \mathbf{n}_W,\ \mathcal{M}_\perp = \boldsymbol{\mathcal{M}} \cdot \mathbf{t}_W \tag{4.5}$$

and $\mathbf{t}_W$ is a unit vector orthogonal to the Wolff plane normal $\mathbf{n}_W$. Since it doesn't matter which one we choose, let $\mathbf{t}_W$ be the vector obtained by rotating $\mathbf{n}_W$ in counterclockwise direction by an angle of $\pi/2$. An analogous decomposition holds for the cluster magnetization with $\mathcal{M}_\parallel = \sum_{c\in\mathcal{C}} \mathcal{M}_{\parallel c}$ and $\mathcal{M}_\perp = \sum_{c\in\mathcal{C}} \mathcal{M}_{\perp c}$. For the magnetization squared one then finds

$$\langle |\boldsymbol{\mathcal{M}}|^2 \rangle = \langle \mathcal{M}_\parallel^2 + 2\boldsymbol{\mathcal{M}}_\parallel \cdot \boldsymbol{\mathcal{M}}_\perp + \mathcal{M}_\perp^2 \rangle = \langle \mathcal{M}_\parallel^2 \rangle + \langle \mathcal{M}_\perp^2 \rangle = 2\langle \mathcal{M}_\parallel^2 \rangle. \tag{4.6}$$

In the last step we used that because of the $O(2)$ invariance of the system and since $\mathcal{M}_\parallel$ and $\mathcal{M}_\perp$ are two components of the vector $\boldsymbol{\mathcal{M}}$, $\langle \mathcal{M}_\parallel^2 \rangle$ must be equal to $\langle \mathcal{M}_\perp^2 \rangle$. For the component

in direction of the Wolff plane normal we find

$$\begin{aligned}\langle \mathcal{M}_{\parallel}^2 \rangle &= \int \mathcal{D}\mathbf{s}\ \rho[\mathbf{s}]\ \mathcal{M}_{\parallel}^2[\mathbf{s}] = \int \mathcal{D}\mathbf{s} \sum_{[b]} \sum_{c_1,c_2\in\mathcal{C}} \rho[\mathbf{s},b]\ \mathcal{M}_{\parallel c_1}[\mathbf{s}]\ \mathcal{M}_{\parallel c_2}[\mathbf{s}] \\ &= \left[\prod_{x\in\Lambda}\int_{S^1} d\mathbf{s}_x\right] \sum_{[b]} \sum_{c_1,c_2\in\mathcal{C}} \rho[\mathbf{s},b]\ \mathcal{M}_{\parallel c_1}[\mathbf{s}]\ \mathcal{M}_{\parallel c_2}[\mathbf{s}] \\ &= \left[\prod_{x\in\Lambda}\int_{S^1|\mathbf{s}_x\cdot\mathbf{n}_W>0} d\mathbf{s}_x\right] \left[\prod_{x\in\Lambda}\sum_{\tilde{\mathbf{s}}_x=\mathbf{s}_x,\mathbf{s}'_x}\right] \sum_{[b]} \sum_{c_1,c_2\in\mathcal{C}} \rho[\tilde{\mathbf{s}},b]\ \mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}]\ \mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}].\end{aligned} \tag{4.7}$$

The sum over all bonds introduced in the first line above is actually only a sum over those bond configurations for which $[\tilde{\mathbf{s}},b]$ is in configuration space. We can interchange the sum over the spin flips, i.e. over $\tilde{\mathbf{s}}_x = \mathbf{s}_x, \mathbf{s}'_x\ \forall x\in\Lambda$, with the sum over the bonds by replacing the restriction on the bonds by one on the allowed spin flips, i.e. we only flip entire clusters. Mathematically we can understand that this interchange is allowed by noting that for configurations which are not in configuration space, $\rho[\tilde{\mathbf{s}},b]=0$. Hence, we can remove the restriction on the bonds, interchange the summations and remove the configurations with $\rho[\tilde{\mathbf{s}},b]=0$ by imposing a restriction on the allowed spin flips. Hence, we find

$$\begin{aligned}&\left[\prod_{x\in\Lambda}\sum_{\tilde{\mathbf{s}}_x=\mathbf{s}_x,\mathbf{s}'_x}\right] \sum_{[b]} \sum_{c_1,c_2\in\mathcal{C}} \rho[\tilde{\mathbf{s}},b]\ \mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}]\ \mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}] \\ &\qquad = \sum_{[b]} \sum_{c_1,c_2\in\mathcal{C}} \left[\prod_{x\in\Lambda}\sum_{\tilde{\mathbf{s}}_x=\mathbf{s}_x,\mathbf{s}'_x}\right] \rho[\tilde{\mathbf{s}},b]\ \mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}]\ \mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}].\end{aligned} \tag{4.8}$$

For some given bond configuration, let us consider the $2^{|\mathcal{C}|}$ spin configurations obtained by independently flipping entire clusters. All of these configurations have the same probability density $\rho[\tilde{\mathbf{s}},b]$. Since under a flip of the cluster c, $\mathcal{M}_{\parallel c}[\tilde{\mathbf{s}}]$ changes sign, there are, for $c_1\neq c_2$, $2^{|\mathcal{C}|-2}$ among the $2^{|\mathcal{C}|}$ spin configurations with

$$\begin{aligned}&\mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}] = \left|\mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}]\right| \text{ and } \mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}] = \left|\mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}]\right|, \\ &\mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}] = -\left|\mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}]\right| \text{ and } \mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}] = \left|\mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}]\right|, \\ &\mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}] = \left|\mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}]\right| \text{ and } \mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}] = -\left|\mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}]\right|, \\ &\mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}] = -\left|\mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}]\right| \text{ and } \mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}] = -\left|\mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}]\right|.\end{aligned} \tag{4.9}$$

Therefore, if $c_1\neq c_2$, we find

$$\left[\prod_{x\in\Lambda}\sum_{\tilde{\mathbf{s}}_x=\mathbf{s}_x,\mathbf{s}'_x}\right] \rho[\tilde{\mathbf{s}},b]\ \mathcal{M}_{\parallel c_1}[\tilde{\mathbf{s}}]\ \mathcal{M}_{\parallel c_2}[\tilde{\mathbf{s}}] = 0. \tag{4.10}$$

Thus, only those terms in the sum (4.8) survive, for which $c_1=c_2$, and (4.7) evaluates to

$$\begin{aligned}\langle \mathcal{M}_{\parallel}^2 \rangle &= \left[\prod_{x\in\Lambda}\int_{S^1|\mathbf{s}_x\cdot\mathbf{n}_W>0} d\mathbf{s}_x\right] \left[\prod_{x\in\Lambda}\sum_{\tilde{\mathbf{s}}_x=\mathbf{s}_x,\mathbf{s}'_x}\right] \sum_{[b]}\sum_{c\in\mathcal{C}} \rho[\tilde{\mathbf{s}},b]\ \mathcal{M}_{\parallel c}^2[\tilde{\mathbf{s}}] \\ &= \int \mathcal{D}\mathbf{s} \sum_{[b]}\sum_{c\in\mathcal{C}} \rho[\mathbf{s},b]\ \mathcal{M}_{\parallel c}^2[\mathbf{s}] = \left\langle \sum_{c\in\mathcal{C}} \mathcal{M}_{\parallel c}^2 \right\rangle.\end{aligned} \tag{4.11}$$

Hence, the magnetic susceptibility is given by

$$\chi = \frac{2}{L_1 L_2} \langle \mathcal{M}_{\parallel}^2 \rangle = \frac{2}{L_1 L_2} \left\langle \sum_{c \in \mathcal{C}} \mathcal{M}_{\parallel c}^2 \right\rangle . \tag{4.12}$$

Rather than measuring the magnetization squared $\boldsymbol{\mathcal{M}}^2[\mathbf{s}]$ for each generated configuration, we measure the quantity $2 \sum_{c \in \mathcal{C}} \mathcal{M}_{\parallel c}^2[\mathbf{s}]$. Clearly, both quantities have the same expectation value, but since we have removed the canceling contributions (4.10), we expect an estimate of the magnetic susceptibility from the average over the latter quantity, to have a smaller standard deviation than an estimate from the average over $\boldsymbol{\mathcal{M}}^2[\mathbf{s}]$.

4.3 Implementation test

In order to test the implementation of the Wolff cluster algorithm we can make use of the fact that Ulli Wolff lists simulation results in Table I of [9]. The table contains the values of several observables at different values of the inverse temperature β and different lattice sizes of a square lattice. In particular, it lists the magnetic susceptibility and the cluster size.

The algorithm used in [9] is the single-cluster variant of the algorithm used here. In the single-cluster variant, instead of activating or deactivating every bond and then identifying the clusters, only one cluster is built. This is done by randomly selecting one lattice site and activating the bonds to nearest neighbors with probability (3.9). Lattice sites that connect to the randomly chosen site by an activated bond are added to the cluster. One then continues by probing all bonds connecting lattice sites of the cluster to sites outside the cluster. This is continued until there are no bonds left to probe.

The mean cluster size $\overline{|c|}_{sc}$ of clusters generated by the single-cluster variant of the Wolff cluster algorithm does not agree with the mean cluster size (4.1) but we can, nonetheless, measure it. The cluster sizes we would like to measure are obtained by randomly choosing a lattice site and measure the size of the cluster that contains that site. The cluster that we select in this way, is reached from any of its sites and consequently the probability of choosing a particular cluster c is given the fraction $|c|/|\Lambda|$. Hence, we find

$$\langle |c|_{sc} \rangle = \left\langle \sum_{i=1}^{|\mathcal{C}|} \frac{|c_i|}{|\Lambda|} |c_i| \right\rangle = \left\langle \frac{1}{|\Lambda|} \sum_{i=1}^{|\mathcal{C}|} |c_i|^2 \right\rangle . \tag{4.13}$$

Table 4.1 lists the results of simulations executed on lattices of the same size, with the same values of inverse temperature and with the number of measurements chosen such that the error bars in the magnetic susceptibility are about the same as in Table I of [9]. For all but two measurements of the magnetic susceptibility, the differences in the mean values from Table 4.1 and Table I of [9] are within the error bars. Also, only two of the differences in the mean values of the mean cluster sizes differ by more than the error bars. Thus, it appears that the algorithm has been implemented correctly.

4.4 Spin-spin correlation

Let x and y be two arbitrary lattice sites. The spin-spin correlation $\langle \mathbf{s}_x \cdot \mathbf{s}_y \rangle$ is defined as the expectation value of the dot product of the spins at lattice site x and y. An improved estimator for the spin-spin correlation can be derived similarly as the magnetic susceptibility.

$\lvert\Lambda\rvert$	β	M	$\left\langle \overline{\lvert c\rvert} \right\rangle$	$\left\langle \overline{\lvert c\rvert}_{sc} \right\rangle \lvert\Lambda\rvert^{-1}$	$\chi\lvert\Lambda\rvert^{-1}$
32^2	1.12	$5 \cdot 2^{19}$	4.425(2)	0.3584(2)	0.4414(2)
64^2	1.12	2^{20}	4.093(2)	0.3046(3)	0.3756(3)
128^2	1.12	$3 \cdot 2^{18}$	3.940(2)	0.2586(3)	0.3190(3)
32^2	1.07	2^{20}	4.077(2)	0.3231(3)	0.3982(4)
64^2	1.07	2^{19}	3.829(2)	0.2629(3)	0.3242(4)
128^2	1.07	$7 \cdot 2^{15}$	3.724(2)	0.2122(4)	0.2618(5)
128^2	1.04	2^{18}	3.592(1)	0.1646(4)	0.2030(4)

Table 4.1: Estimates of the magnetic susceptibility χ, the mean cluster size of the multi-cluster variant $\left\langle \overline{\lvert c\rvert} \right\rangle$ and the single-cluster variant $\left\langle \overline{\lvert c\rvert}_{sc} \right\rangle$ of the Wolff cluster algorithm. $\lvert\Lambda\rvert$ indicates the size of the lattice used, β the inverse temperature and M the number of measurements. The measurements were taken after an equilibration of $10'000$ sweeps.

We begin by decomposing the spins into a component along and one orthogonal to the Wolff plane normal

$$\mathbf{s} = \mathbf{s}_\parallel + \mathbf{s}_\perp = s_\parallel \mathbf{n}_W + s_\perp \mathbf{t}_W, \qquad \text{where } s_\parallel = \mathbf{s}\cdot\mathbf{n}_W,\ s_\perp = \mathbf{s}\cdot\mathbf{t}_W \tag{4.14}$$

and $\mathbf{t}_W$ is again a unit vector orthogonal to the Wolff plane normal $\mathbf{n}_W$. For the spin-spin correlation one then finds

$$\langle \mathbf{s}_x \cdot \mathbf{s}_y \rangle = \langle s_{\parallel x}\ s_{\parallel y} + \mathbf{s}_{\parallel x}\cdot\mathbf{s}_{\perp y} + \mathbf{s}_{\perp x}\cdot\mathbf{s}_{\parallel y} + s_{\perp x}\ s_{\perp y}\rangle = \langle s_{\parallel x}\ s_{\parallel y}\rangle + \langle s_{\perp x}\ s_{\perp y}\rangle = 2\langle s_{\parallel x}\ s_{\parallel y}\rangle \tag{4.15}$$

In the last step we used that, by $O(2)$ invariance of the system, we can interchange the values of $s_{\parallel x}\ s_{\parallel y}$ and $s_{\perp x}\ s_{\perp y}$, thus, $\langle s_{\perp x}\ s_{\perp y}\rangle = \langle s_{\parallel x}\ s_{\parallel y}\rangle$. For the product of the components in the direction of the Wolff plane normal we find

$$\begin{aligned} \langle s_{\parallel x}\ s_{\parallel y}\rangle &= \int \mathcal{D}\mathbf{s}\ \rho[\mathbf{s}]\ s_{\parallel x}\ s_{\parallel y} = \int \mathcal{D}\mathbf{s} \sum_{[b]} \rho[\mathbf{s}, b]\ s_{\parallel x}\ s_{\parallel y} \\ &= \left[\prod_{x\in\Lambda} \int_{S^1 \mid \mathbf{s}_x\cdot\mathbf{n}_W > 0} d\mathbf{s}_x\right] \left[\prod_{x\in\Lambda} \sum_{\tilde{\mathbf{s}}_x = \mathbf{s}_x, \mathbf{s}'_x}\right] \sum_{[b]} \rho[\tilde{\mathbf{s}}, b]\ s_{\parallel x}\ s_{\parallel y}. \end{aligned} \tag{4.16}$$

As in equation (4.8) we can exchange the sum over the bonds with the one over the spin flips by replacing the restriction on the bonds by the restriction that we only flip entire clusters, i.e.

$$\left[\prod_{x\in\Lambda} \sum_{\tilde{\mathbf{s}}_x = \mathbf{s}_x, \mathbf{s}'_x}\right] \sum_{[b]} \rho[\tilde{\mathbf{s}}, b]\ s_{\parallel x}\ s_{\parallel y} = \sum_{[b]} \left[\prod_{x\in\Lambda} \sum_{\tilde{\mathbf{s}}_x = \mathbf{s}_x, \mathbf{s}'_x}\right] \rho[\tilde{\mathbf{s}}, b]\ s_{\parallel x}\ s_{\parallel y}. \tag{4.17}$$

For a given bond configuration, there are $2^{|\mathcal{C}|}$ spin configurations obtained by independently flipping entire clusters. All of these configurations have the same probability density $\rho[\tilde{\mathbf{s}}, b]$. Since under a flip of the cluster to which the lattice site x (or y) belongs, $s_{\parallel x}$ (or $s_{\parallel y}$) changes sign, we find, if x and y belong to different clusters, that

$$\left[\prod_{x\in\Lambda} \sum_{\tilde{\mathbf{s}}_x = \mathbf{s}_x, \mathbf{s}'_x}\right] \rho[\tilde{\mathbf{s}}, b]\ s_{\parallel x}\ s_{\parallel y} = 0. \tag{4.18}$$

On the other hand, if the lattice sites x and y belong to the same cluster, then the product $s_{\parallel x}\ s_{\parallel y}$ remains the same among the $2^{|\mathcal{C}|}$ configurations obtained by independently flipping

clusters. Hence, (4.16) evaluates to

$$\begin{aligned}\langle s_{\parallel x}\ s_{\parallel y}\rangle &= \left[\prod_{x\in\Lambda}\int_{S^1|\mathbf{s}_x\cdot\mathbf{n}_W>0} d\mathbf{s}_x\right]\left[\prod_{x\in\Lambda}\sum_{\tilde{\mathbf{s}}_x=\mathbf{s}_x,\mathbf{s}'_x}\right]\sum_{[b]}\sum_{c\in\mathcal{C}}\rho[\tilde{\mathbf{s}},b]\ s_{\parallel x}\ s_{\parallel y}\ \Theta(c,x)\ \Theta(c,y)\\ &= \int\mathcal{D}\mathbf{s}\sum_{[b]}\sum_{c\in\mathcal{C}}\rho[\mathbf{s},b]\ s_{\parallel x}\ s_{\parallel y}\ \Theta(c,x)\ \Theta(c,y) \qquad (4.19)\\ &= \left\langle\sum_{c\in\mathcal{C}} s_{\parallel x}\ s_{\parallel y}\ \Theta(c,x)\ \Theta(c,y)\right\rangle,\end{aligned}$$

where the characteristic function Θ is defined as

$$\Theta(c,x) := \begin{cases}1 & \text{if } x\in c,\\ 0 & \text{otherwise.}\end{cases} \qquad (4.20)$$

The spin-spin correlation of the spins at lattice site x and y is, therefore, given by

$$\langle\mathbf{s}_x\cdot\mathbf{s}_y\rangle = 2\left\langle\sum_{c\in\mathcal{C}} s_{\parallel x}\ s_{\parallel y}\ \Theta(c,x)\ \Theta(c,y)\right\rangle. \qquad (4.21)$$

Because of translation invariance of the XY model, the spin-spin correlation of the spins at lattice site x and y does not depend on the coordinates $\mathbf{x}$ and $\mathbf{y}$ of these sites, but only on the distance vector $\mathbf{d}=\mathbf{y}-\mathbf{x}$ between them. Therefore, we define the spin-spin correlation function as

$$C_s(\mathbf{d}) := \langle\mathbf{s}_r\cdot\mathbf{s}_{r+d}\rangle, \qquad (4.22)$$

where $r\in\Lambda$ is some arbitrary reference site with coordinates $\mathbf{r}$ and $r+d\in\Lambda$ is the lattice site at coordinates $\mathbf{r}+\mathbf{d}$.

In a simulation of the XY model, we may estimate the spin-spin correlation function by randomly selecting a reference site $r\in\Lambda$ in each configuration generated after equilibrium was reached. We then measure the quantity $2s_{\parallel r}\ s_{\parallel r+d}$ for every lattice site $r+d$ in the same cluster as r, since the other sites do not contribute. Because the lattice is of finite size with periodic boundary conditions, one should take the shortest distance between the lattice sites r and $r+d$ as the distance $\mathbf{d}$. If we measure the spin-spin correlation function in this way, we only take the information contained in one of the clusters into account. In order not to lose the information contained in the remaining clusters, we choose a reference site r_i in each cluster $c_i\in\mathcal{C},\ i=1,\ldots,|\mathcal{C}|$. The probability to choose a specific cluster c_i is given by $|c_i|/|\Lambda|$, so we find

$$C_s(\mathbf{d}) = 2\left\langle\sum_{i=1}^{|\mathcal{C}|}\frac{|c_i|}{|\Lambda|}\ s_{\parallel r_i}\ s_{\parallel r_i+d}\ \Theta(c_i,r+d)\right\rangle. \qquad (4.23)$$

One way of testing an improved estimator is to compare it with the (unimproved) observable. That is, we compare the estimate for the spin-spin correlation function obtained according to (4.23) with the estimate obtained by measuring the quantity $\mathbf{s}_r\cdot\mathbf{s}_{r+d}$ in every configuration. For a given distance vector $\mathbf{d}$ let us denote the (unimproved) observable by O and the improved estimator by I. The individual measurements of O and I are binned such that we have two sets of statistically independent bin average values O_i and I_i. Since

we usually have a large number of bins, the distributions in the means of the bin average values $\bar{O}$ and $\bar{I}$ can be assumed to be Gaussian. These two distributions are characterized by the (exact) average values in the means of the bin average values $\langle\bar{O}\rangle$, $\langle\bar{I}\rangle$ and the standard deviations in the means of the bin average values $\sigma_{\bar{O}}$, $\sigma_{\bar{I}}$. The distribution in the difference $\bar{O}-\bar{I}$ will then also be Gaussian and the standard deviation in the difference $\sigma_{\bar{O}-\bar{I}}$ is related to the standard deviations $\sigma_{\bar{O}}$ and $\sigma_{\bar{I}}$ by

$$\sigma_{\bar{O}-\bar{I}} = \sqrt{\sigma_{\bar{O}}^2 + \sigma_{\bar{I}}^2}, \tag{4.24}$$

which we can estimate from the (known) estimates of $\sigma_{\bar{O}}$ and $\sigma_{\bar{I}}$. In order to obtain this result for the propagation of the error, we have to assume that the two sets of bin average values O_i and I_i are independent of each other. Although the measurements of O and I are taken from the same configurations $[\mathbf{s}^j, b^j]$ generated by the algorithm, we will assume this since in general $O[\mathbf{s}^j] \neq I[\mathbf{s}^j, b^j]$. If the (unimproved) observable O and the improved estimator I indeed measure the same quantity, then the (exact) average in the difference $\langle\bar{O}-\bar{I}\rangle = \langle\bar{O}\rangle - \langle\bar{O}\rangle$ should be zero. In order to compare the improved estimator to the (unimproved) observable, we now test, for each distance vector $\mathbf{d}$, how far, in units of standard deviations estimated from $\sigma_{\bar{O}-\bar{I}}$, the difference $\bar{O}-\bar{I}$ deviates from zero. If the differences indeed satisfy the Gaussian distributions described above, then 68.2% of the differences should deviate by less than one, 27.2% between one and two, 4.2% between two and three and 0.2% by more than three standard deviations from zero. The results of this comparison are shown in Figure 4.2.

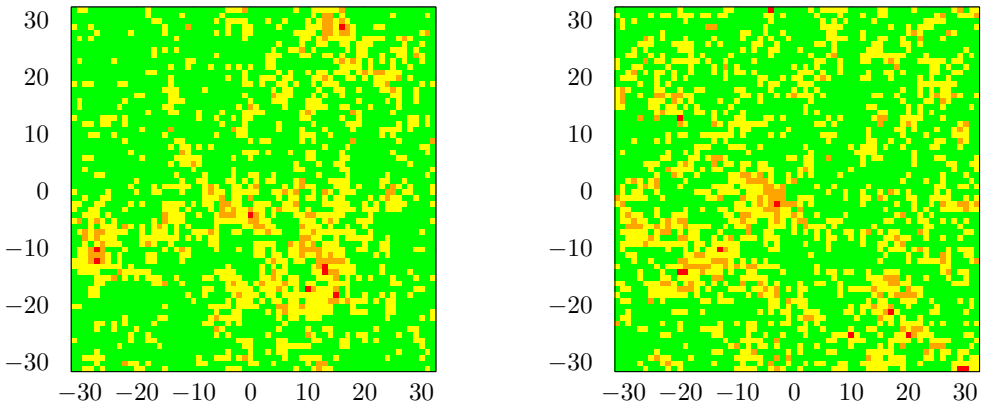


Figure 4.2: Comparison of the estimate obtained from the improved estimator with the estimate obtained from the (unimproved) observable for $\beta = 1.22$ (left) and $\beta = 0.92$ (right). A green point indicates that the difference in the mean values obtained from the improved estimator and the (unimproved) observable deviates less than one standard deviation in the difference (69.6% of the points in the left and 65.5% of the points in the right figure belong to this category.). A yellow point indicates that the difference deviates between one and two standard deviations (26.6% on the left and 29.3% on the right), an orange point that it deviates between two and three (3.5% on the left and 4.9% on the right) and a red point that it deviates by more than three standard deviations (0.2% on the left and 0.3% on the right). Note also that the points where the differences deviate the most, appear to be randomly distributed. Thus, the improved estimator and the (unimproved) observable appear to measure the same quantity. To obtain these figures, $524'288$ measurements were taken after an equilibration of $20'000$ sweeps on a lattice of 64^2 sites and a bin size of 32 measurements was used.

Figure 4.3 and 4.4 shows the spin-spin correlation function obtained according to (4.23) at two different temperatures. We can not directly compare these results with the description

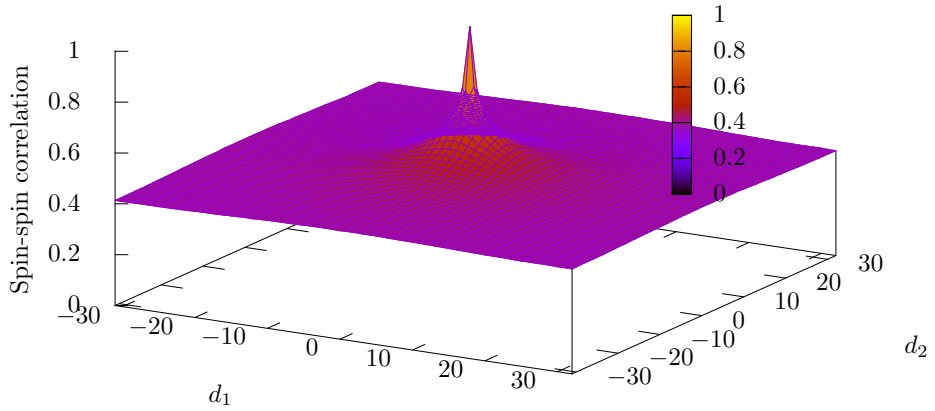


Figure 4.3: Spin-spin correlation as a function of the distance vector $\mathbf{d} = (d_1, d_2)$ obtained from 524′288 measurements, with a bin size of 32 measurements, after an equilibration of 20′000 sweeps on a lattice of 64^2 sites at $\beta = 1.22$, that is, in the spin-wave phase. The standard deviations in the plotted points range from 0.0006 to 0.0008.

of the spin-spin correlation function in Section 2.3 because this description is only valid in the limit of an infinitely large system. However, we see by comparing Figure 4.3 with Figure 4.4 that the spin-spin correlation function in the spin-wave phase falls off faster than in the disordered phase, as expected.

4.5 Vortex-vortex correlation

Let $\square_u$ and $\square_v$ be two arbitrary plaquettes of the dual lattice. The vortex-vortex correlation $\langle \nu_{\square_u} \nu_{\square_v} \rangle$ is defined as the expectation value of the product of the vorticities of the plaquettes $\square_u$ and $\square_v$. In order to find an improved estimator for the vortex-vortex correlation, we proceed as in the previous sections and find

$$\langle \nu_{\square_u}\ \nu_{\square_v} \rangle = \left[\prod_{x \in \Lambda} \int_{S^1 | \mathbf{s}_x \cdot \mathbf{n}_W > 0} d\mathbf{s}_x \right] \sum_{[b]} \left[\prod_{x \in \Lambda} \sum_{\tilde{\mathbf{s}}_x = \mathbf{s}_x, \mathbf{s}'_x} \right] \rho[\tilde{\mathbf{s}}, b]\ \nu_{\square_u}[\tilde{\mathbf{s}}]\ \nu_{\square_v}[\tilde{\mathbf{s}}], \tag{4.25}$$

where the sum over all spin flips is restricted to flipping entire clusters. For a given bond configuration, let us consider the $2^{|C|}$ equally probable spin configurations obtained from $[\mathbf{s}, b]$ by independently flipping cluster. For a plaquette $\square$ and a cluster $c \in \mathcal{C}$ that belongs to that plaquette, we define

$$\nu_{\triangle}([\mathbf{s}], c) := \begin{cases} 0 & \text{if none of the semi-vortices of } \square \text{ belongs to } c, \\ \frac{1}{2} & \text{if the positive semi-vortex of } \square \text{ belongs to } c \text{ and} \\ -\frac{1}{2} & \text{if the negative semi-vortex of } \square \text{ belongs to } c. \end{cases} \tag{4.26}$$

With this definition, we can show that

$$\left[\prod_{x \in \Lambda} \sum_{\tilde{\mathbf{s}}_x = \mathbf{s}_x, \mathbf{s}'_x} \right] \rho[\tilde{\mathbf{s}}, b]\ \nu_{\square_u}[\tilde{\mathbf{s}}]\ \nu_{\square_v}[\tilde{\mathbf{s}}] = \left[\prod_{x \in \Lambda} \sum_{\tilde{\mathbf{s}}_x = \mathbf{s}_x, \mathbf{s}'_x} \right] \rho[\tilde{\mathbf{s}}, b] \sum_{c \in \mathcal{C}} \nu_{\triangle_u}([\mathbf{s}], c)\ \nu_{\triangle_v}([\mathbf{s}], c). \tag{4.27}$$

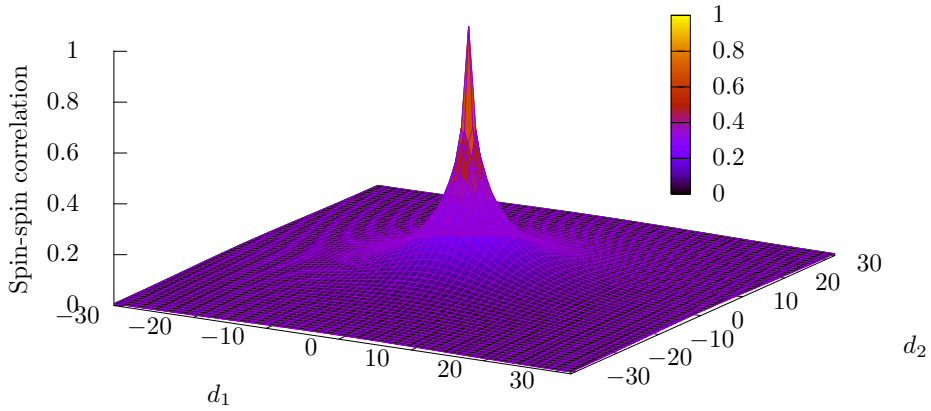


Figure 4.4: Spin-spin correlation function obtained from 524′288 measurements, with a bin size of 32 measurements, after an equilibration of 20′000 sweeps on a lattice of 64^2 sites at $\beta = 0.92$, that is, in the disordered phase. The standard deviations in the plotted points range from 0.0001 to 0.0003.

In order to do so, we distinguish the following four cases, but first, note that the right-hand side of (4.27) simplifies to $2^{|\mathcal{C}|}\rho[\mathbf{s}, b] \sum_{c\in\mathcal{C}} \nu_{\triangle_u}([\mathbf{s}], c)\, \nu_{\triangle_v}([\mathbf{s}], c)$ since $\nu_{\triangle_u}([\mathbf{s}], c)$ and $\nu_{\triangle_v}([\mathbf{s}], c)$ do not depend on the orientation of the clusters.

1. If one of the plaquettes has zero vorticity in all the $2^{|\mathcal{C}|}$ spin configurations, then also $\nu_{\square_u}\nu_{\square_v} = 0$ in all of them. Since then one of the plaquettes has no semi-vortices, $\sum_{c\in\mathcal{C}} \nu_{\triangle_u}([\mathbf{s}], c)\, \nu_{\triangle_v}([\mathbf{s}], c) = 0$ and (4.27) is satisfied.

2. Let us assume that both plaquettes have non-zero vorticity in one of the $2^{|\mathcal{C}|}$ spin configurations, but that none of the semi-vortices of one plaquette belongs to the same cluster as one of the semi-vortices of the other plaquette. We will then find the value $\nu_{\square_u}\nu_{\square_v} = 1$ just as many times as the value -1, thus, $\nu_{\square_u}\nu_{\square_v}$ averages to zero among the $2^{|\mathcal{C}|}$ spin configurations. In this case also $\sum_{c\in\mathcal{C}} \nu_{\triangle_u}([\mathbf{s}], c)\, \nu_{\triangle_v}([\mathbf{s}], c) = 0$ and, hence, (4.27) is satisfied.

3. Let us now assume that one of the semi-vortices of each plaquette belong to the same cluster but that the remaining two do not. If the semi-vortices that belong to the same cluster are both either positive or negative, then we find, since the vorticities of the two plaquettes are determined by three clusters, $2 \cdot 2^{|\mathcal{C}|-3} = 2^{|\mathcal{C}|-2}$ spin configurations in which $\nu_{\square_u}\nu_{\square_v} = 1$. If, on the other hand, one of the semi-vortices that belong to the same cluster is positive and the other negative, then we find $2^{|\mathcal{C}|-2}$ configurations in which $\nu_{\square_u}\nu_{\square_v} = -1$. In the remaining configurations of both cases $\nu_{\square_u}\nu_{\square_v} = 0$ and the left-hand side of (4.27) either evaluates to $2^{|\mathcal{C}|-2}\rho[\mathbf{s}, b]$ or $-2^{|\mathcal{C}|-2}\rho[\mathbf{s}, b]$. Since there is only one cluster $c \in \mathcal{C}$ for which $\nu_{\triangle_u}([\mathbf{s}], c)$ and $\nu_{\triangle_v}([\mathbf{s}], c)$ are non-zero, the right hand side of (4.27) either evaluates to $2^{|\mathcal{C}|}\rho[\mathbf{s}, b]\frac{1}{4} = 2^{|\mathcal{C}|-2}\rho[\mathbf{s}, b]$ or $-2^{|\mathcal{C}|-2}\rho[\mathbf{s}, b]$. Hence, (4.27) is satisfied.

4. In the last case, both of the semi-vortices of one of the plaquettes belong to the same clusters as the semi-vortices of the other plaquette. Much in the same way as in the

previous case, if both positive semi-vortices belong to one cluster and both negative semi-vortices to the other, then the left-hand side of (4.27) evaluates to $2^{|\mathcal{C}|-1}\rho[\mathbf{s},b]$. If, on the other hand, one positive and one negative semi-vortex belong to the same cluster and the remaining two semi-vortices to the other cluster, then the left-hand side of (4.27) evaluates to $-2^{|\mathcal{C}|-1}\rho[\mathbf{s},b]$. Since there are now two cluster $c \in \mathcal{C}$ for which $\nu_{\triangle_u}([\mathbf{s}],c)$ and $\nu_{\triangle_v}([\mathbf{s}],c)$ are non-zero, the right-hand side of (4.27) either evaluates to $2^{|\mathcal{C}|}\rho[\mathbf{s},b]2\frac{1}{4} = 2^{|\mathcal{C}|-1}\rho[\mathbf{s},b]$ or $-2^{|\mathcal{C}|-1}\rho[\mathbf{s},b]$. Hence, (4.27) is satisfied.

In all cases (4.27) is satisfied and thus, the vortex-vortex correlation of the plaquettes $\square_u$ and $\square_v$ is given by

$$\langle \nu_{\square_u}\, \nu_{\square_v} \rangle = \left\langle \sum_{c\in\mathcal{C}} \nu_{\triangle_u}(c)\, \nu_{\triangle_v}(c) \right\rangle . \tag{4.28}$$

Because of translation invariance, the vortex-vortex correlation of the plaquettes $\square_u$ and $\square_v$ does not depend on the coordinates of the two plaquettes on the dual lattice but only on the distance vector $\mathbf{d} = \mathbf{v} - \mathbf{u}$ between them. Therefore, the vortex-vortex correlation function is defined by

$$C_v(\mathbf{d}) := \langle \nu_{\square_r}\, \nu_{\square_{r+d}} \rangle, \tag{4.29}$$

where r is some arbitrary reference site on the dual lattice with coordinates $\mathbf{r}$ and $r + d$ is the site with coordinates $\mathbf{r} + \mathbf{d}$.

In a simulation, we may estimate the vortex-vortex correlation function by randomly selecting a plaquette r in each of the measured configurations. If we do so, then we will often pick a plaquette with zero vorticity in all the $2^{|\mathcal{C}|}$ configurations, and we lose all of the information about the vortex-vortex correlation function contained in the configuration. Preferably, we would like to pick one plaquette, with non-zero vorticity in one of the $2^{|\mathcal{C}|}$ configurations, in each cluster, but then we would consider the same pair of vortices more than once. The problem is that the plaquettes, with non-zero vorticity in one of the $2^{|\mathcal{C}|}$ configurations, lie between clusters. However, the semi-vortices only belong to one cluster. Thus, if we, instead of selecting a plaquette in each cluster, randomly pick a semi-vortex, then we find that

$$\langle \nu_{\square_u}\, \nu_{\square_v} \rangle = \left\langle \sum_{i=1}^{|\mathcal{C}|} \frac{|\triangle(c_i)|}{|\Lambda|} \nu_{\triangle_{r_i}}(c_i)\, \nu_{\triangle_{r_i+d}}(c_i) \right\rangle , \tag{4.30}$$

where $|\triangle(c_i)|$ denotes the number of semi-vortices in the cluster c_i, and we used that the number of plaquettes of the lattice equals the number of lattice sites.

The improved estimator for the vortex-vortex correlation function can be tested with the same method we used for the spin-spin correlation function, that is, by comparing with an (unimproved) estimator. One important feature of such an unimproved estimator is simplicity. The implementation of an unimproved estimator should be simple enough such that we can be sure that it works correctly. This is certainly satisfied by the (unimproved) observable but, as in most configurations a randomly selected plaquette has zero vorticity in all $2^{|\mathcal{C}|}$ configurations, the standard deviation in the mean might be quite large. Thus, we might not realize by comparing with this observable, that the improved estimator produces wrong results. We, therefore, test the improved estimator for the vortex-vortex correlation function not just against the (unimproved) observable but also against a semi-improved estimator. The only difference to the observable is, that the semi-improved estimator chooses a reference plaquette among the plaquettes with non-zero vorticity in one of the $2^{|\mathcal{C}|}$ configurations. The results are shown in Figure 4.5.

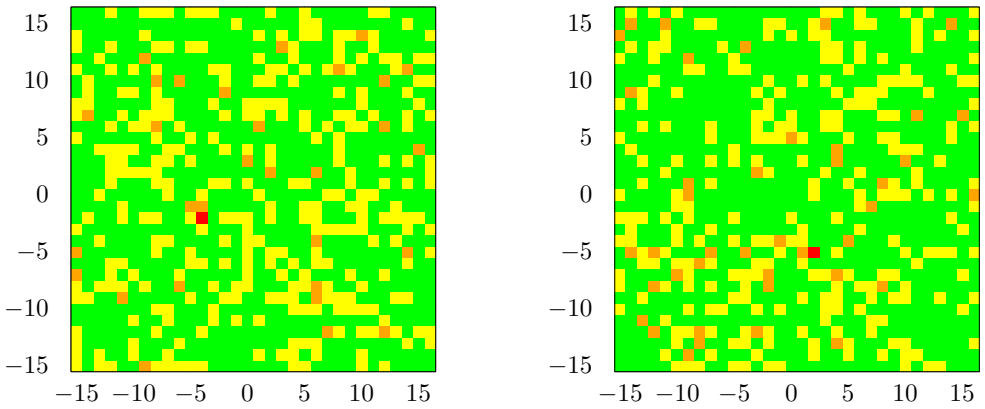


Figure 4.5: Comparison of the estimate obtained from the improved estimator with the estimate obtained from the observable (left) and the semi-improved estimator (right) for $\beta = 1.22$. Points where the difference in the mean values obtained from the improved estimator and the observable or of the semi-improved estimator deviate less than one standard deviation in the difference, are indicated by a green point (67.4% in the left and 70.4% in the right figure). Those which deviate between one and two standard deviations are indicated by a yellow point (29.6% in the left and 25.5% in the right figure), those that deviate between two and three standard deviations are indicated by an orange point (2.9% in the left and 4% in the right figure) and those that deviate by more than tree standard deviations are indicated by a red point (0.1% in the left 0.1% in the right figure). Thus, the improved estimator and the observable and the semi-improved estimator appear to measure the same quantity. To obtain these figures, $5240'288$ measurements were taken after an equilibration of $20'0000$ sweeps on a lattice of 32^2 sites and a bin size of 32 measurements was used.

The vortex-vortex correlation function is shown in Figure 4.6. The form of the curve is almost the same above and below the critical temperature, except that the peak in the middle is much less high and slightly less spread in the quasi-ordered phase. Both of these differences can best be seen from Figure 4.7 which shows the vortex-vortex correlation function along one of the two dual-lattice directions that passes through the reference point, i.e. the restriction of the vortex-vortex correlation function to distance vectors $\mathbf{d} = (d_1, d_2)$ with either $d_1 = 0$ or $d_2 = 0$. From the difference in height of the peak we may conclude that there are, on average, fewer vortices and anti-vortices in the quasi-ordered phase than in the disordered phase. This was to be expected as the system becomes more turbulent with increasing temperature. The negative values around the peak of the vortex-vortex correlation function indicate that in both phases, vortices and anti-vortices form closely bound pairs. In the quasi-ordered phase, however, the mean distance between a vortex and its anti-vortex partner appears to be slightly larger. Note that this does not exclude that there are free vortices in the disordered phase, it merely states that a relatively large part of the vortices and anti-vortices still form bound pairs. There appears to be no contribution to the vortex-vortex correlation function, other than that they contribute to the peak at the center, that can be assigned to free vortices and anti-vortices. There is also no significant difference between the correlation function in the two phases that could be attributed to the occurrence of free vortices in the disordered phase. Hence, it appears that the vortex-vortex correlation is not sensitive to the BKT-transition.

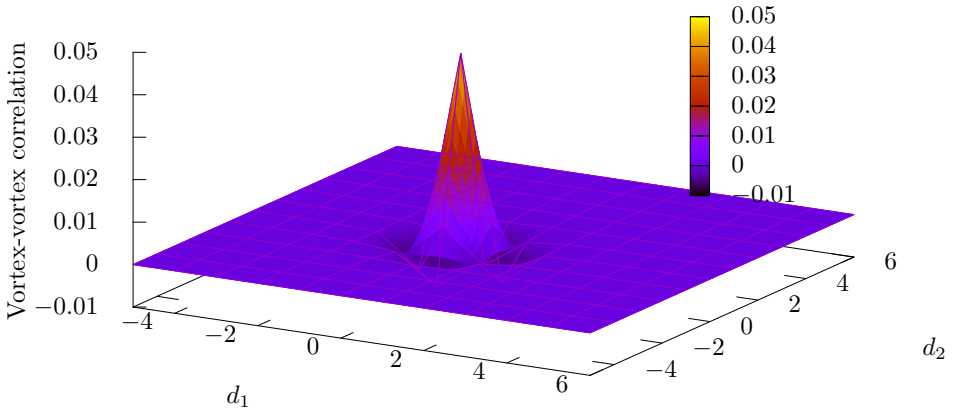


Figure 4.6: Vortex-vortex correlation as a function of the distance vector $\mathbf{d} = (d_1, d_2)$ obtained from 524′288 measurements of the improved estimator, with a bin size of 32 measurements, after an equilibration of 20′000 seeps on a lattice of 32^2 sites at an inverse temperature of $\beta = 0.92$. The standard deviations in the vortex-vortex correlations range from $5 \cdot 10^{-6}$ to $4 \cdot 10^{-5}$.

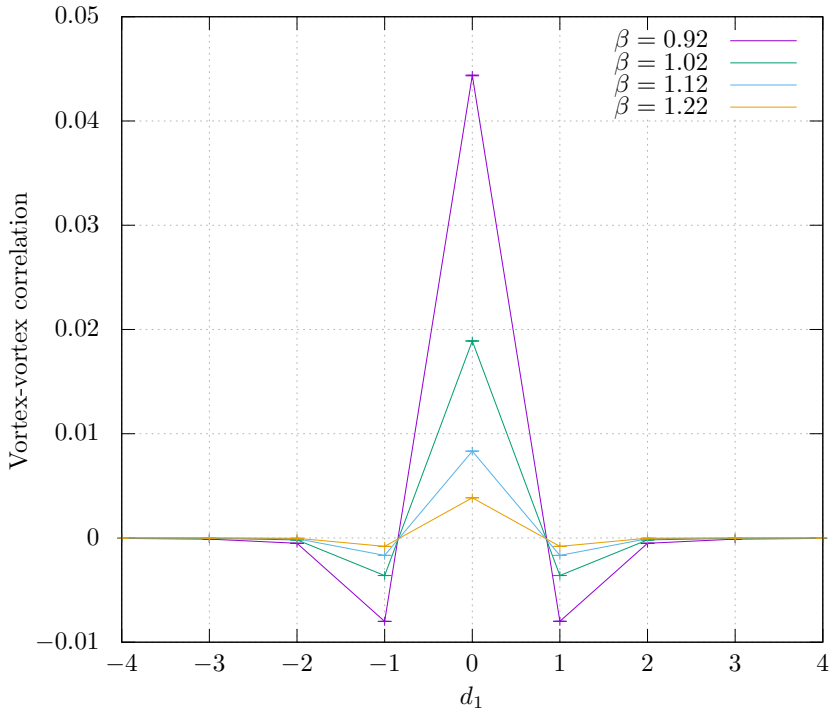


Figure 4.7: Restriction of the vortex-vortex correlation function to distance vectors $\mathbf{d} = (d_1, d_2)$ with either $d_1 = 0$ or $d_2 = 0$. This plot was obtained from 524′288 measurements of the improved estimator for the vortex-vortex correlation function, with a bin size of 32 measurements, after an equilibration of 20′000 seeps on a lattice of 32^2 sites.

5 Conclusion and Outlook

We have implemented the multi-cluster variant of the Wolff cluster algorithm for the XY model in order to study the vortices that form on the lattice. The fact that the vorticity of each vortex is determined by exactly two of the clusters to which the vortex belongs, allows to introduce semi-vortices. In difference to vortices, semi-vortices are static in the sense that they maintain their vorticity among the $2^{|\mathcal{C}|}$ configurations obtained by flipping clusters. It is these semi-vortices from which we hoped to gain further insights into the BKT-transition. Although the semi-vortices allowed us to derive an improved estimator for the vortex-vortex correlation function, it turned out that this function is not sensitive to the BKT-transition.

The vortex-vortex correlation function is not the key variable to gain further insights into the BKT-transition as it fails to capture the occurrence of free vortices in the disordered phase. However, when considering single configurations generated during a simulation in the disordered phase, there are occasionally such free vortices and anti-vortices (see Figure 5.1). These are not present in the quasi-ordered phase shown in Figure 5.2. In order to proceed,

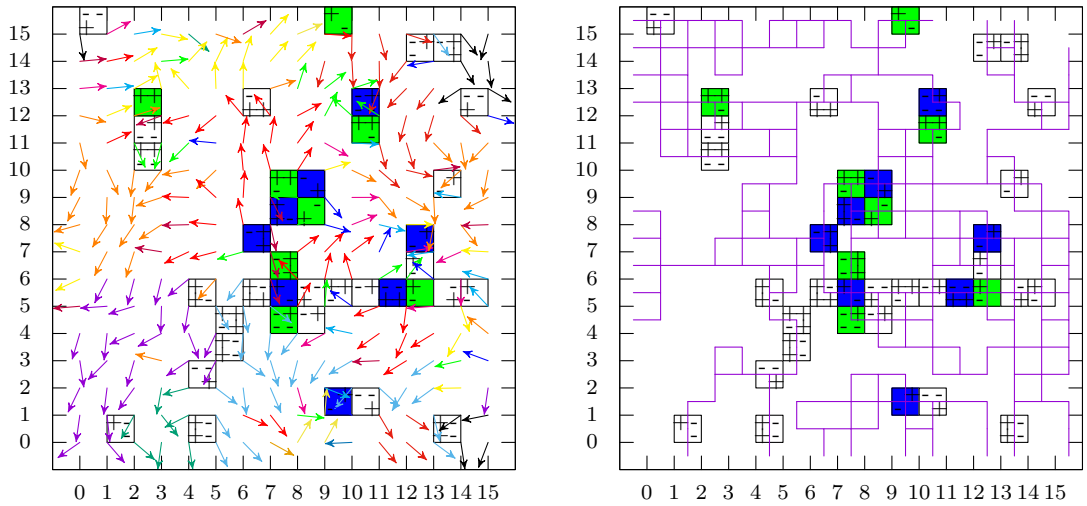


Figure 5.1: A configuration generated during a simulation at $\beta = 0.92$, i.e. in the disordered phase. On the left-hand side, the spins of the configuration are shown where the colors indicate the clusters to which they belong. The violet lines shown on the right-hand side separate the different clusters of the configuration. The black squares indicate plaquettes which are, in one of the $2^{|\mathcal{C}|}$ configurations, a vortex. Those filled are vortices (blue) or anti-vortices (green) in the depicted configuration.

one would first require a definition for free vortices and anti-vortices which is not just based on distance to the next vortex or anti-vortex. Such a definition could then allow to find an observable which is sensitive to the BKT-transition.

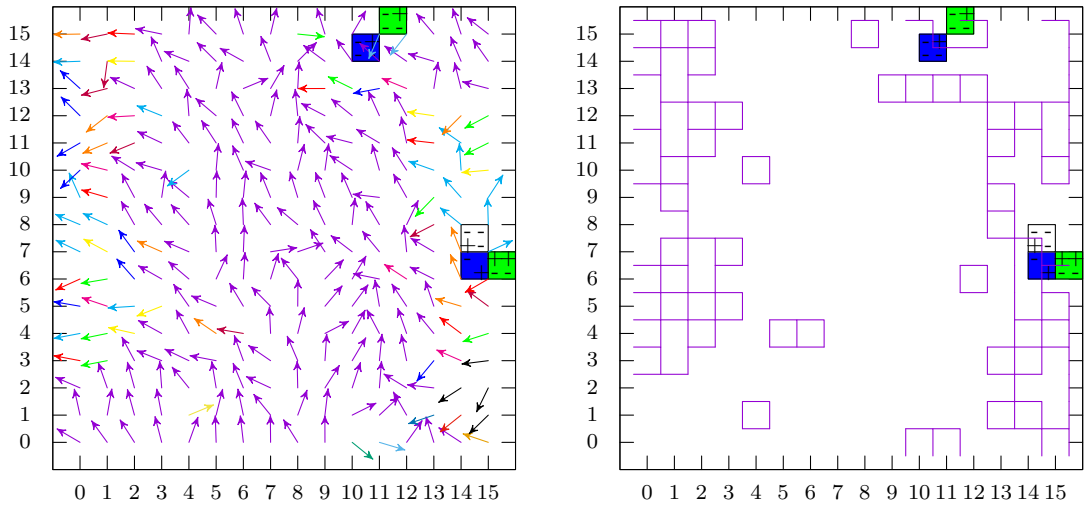


Figure 5.2: A configuration generated during a simulation at $\beta = 1.18$, i.e. in the quasi-ordered phase.

Acknowledgments

Most of all, I would like to express my gratitude to my advisor Uwe-Jens Wiese for his patience and advice throughout the last year. Further, I would like to thank Manes Hornung and João C. Pinto Barros for their help and great discussions. Finally, I thank Nathalie Aeschlimann for her everlasting support and continuous encouragement.

References

[1] V. L. Berezinskii, Destruction of long range order in one-dimensional and two-dimensional systems having a continuous symmetry group. I. Classical systems, *Sov. Phys. JETP* **32**, 493 (1971).

[2] W. Bietenholz, J. C. P. Barros, S. Caspar, M. Hornung, U.-J. Wiese, Meron- and Semi-Vortex-Clusters as Physical Carriers of Topological Charge and Vorticity, arXiv: `1912.04112 [hep-lat]` (2019).

[3] T. H. Hansson, D. Haviland, G. Ingelman, Strange phenomena in matter's flatlands, *The Royal Swedish Academy of Sciences* (2016).

[4] J. M. Kosterlitz, The critical properties of the two-dimensional xy model, *J. Phys. C* **7**, 1046 (1974).

[5] J. M. Kosterlitz, D. J. Thouless, Ordering, metastability and phase transitions in two-dimensional systems, *J. Phys. C* **6**, 1181 (1973).

[6] D. P. Landau, K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, ed. 3, 2009).

[7] U.-J. Wiese, *Statistical Mechanics,* Lecture notes (Institute for Theoretical Physics, University of Bern, 2010), (`http://www.wiese.itp.unibe.ch/lectures/thermodyn.pdf`).

[8] U. Wolff, Asymptotic freedom and mass generation in the O(3) nonlinear σ-model, *Nucl. Phys. B* **334**, 581 (1990).

[9] U. Wolff, Collective Monte Carlo Updating for Spin Systems, *Phys. Rev. Lett.* **62**, 361 (1989).

[10] U. Wolff, Collective Monte Carlo updating in a high precision study of the x-y model, *Nucl. Phys. B* **322**, 759 (1989).

Erklärung

gemäss Art. 30 RSL Phil.-nat. 18

Name/Vorname: Descombes, Marius

Matrikelnummer: 08-631-707

Studiengang: Theoretische Physik

Bachelor ☐ Master ☑ Dissertation ☐

Titel der Arbeit: Semi-Vortex Cluster Algorithm for the two dimensional XY Model

LeiterIn der Arbeit: Prof. Dr. Uwe-Jens Wiese

Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe r des Gesetzes vom 5. September 1996 über die Universität zum Entzug des auf Grund dieser Arbeit verliehenen Titels berechtigt ist.
Für die Zwecke der Begutachtung und der Überprüfung der Einhaltung der Selbständigkeitserklärung bzw. der Reglemente betreffend Plagiate erteile ich der Universität Bern das Recht, die dazu erforderlichen Personendaten zu bearbeiten und Nutzungshandlungen vorzunehmen, insbesondere die schriftliche Arbeit zu vervielfältigen und dauerhaft in einer Datenbank zu speichern sowie diese zur Überprüfung von Arbeiten Dritter zu verwenden oder hierzu zur Verfügung zu stellen.

Bern, 18.12.2020

Ort/Datum

Unterschrift