Universitetet
i Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

MASTER'S THESIS

Study programme/specialisation: MA-MAT FYS Mathematics and Physics	Spring / Autumn semester, 2019 Open/Confidential
Author: Jon Erling Vembe	Jon Erling Vembe (signature of author)
Programme coordinator: Anders Tranberg, Tore Selland Kleppe Supervisor(s): Anders Tranberg	
Title of master's thesis: Lattice Monte Carlo Simulation and phase transitions in Scalar field theories	
Credits: 60	
Keywords: Phase transitions, Electroweak phase transition, Thermal field theory, Lattice QFT, Markov chain Monte-Carlo	Number of pages: 31 + supplemental material/other: ..... Stavanger, 14.06.2019 (date/year)

Lattice Monte Carlo simulation and phase transitions in scalar field theories

Jon Erling Vembe

Abstract

The cubic anisotropy model is a simple scalar field theory with a phase transition of either first- or second order depending on the values of its parameters, with similarities to the electroweak phase transition that occurred in the very early universe. Using lattice simulations, the critical mass of this phase transition is estimated for several values of the parameters, and an infinite volume estimate is produced. Further, the question of whether the phase transitions are first- or second order in the lattice model is investigated. The phase transitions are found to be of second-order when $3\lambda_2/\lambda_1 = 0, 1, 3$, and first order when $3\lambda_2/\lambda_1 = 3.003$, in line with existing theory on the matter.

Contents

1 Introduction

In cosmological models of the very early Universe, it is theorized that all fundamental forces may have been unified, gradually separating as the Universe expanded and cooled down through a series of symmetry-breaking phase transitions. Of these theories, the electroweak phase transition is particularly well-developed, modelling the separation of the electromagnetic and weak interactions. In this thesis, the aim is to test a numerical lattice simulation method of analysing equilibrium states in a simpler scalar field theory, the cubic anisotropy model. While simple compared to the standard model, it exhibits a similar transition to that in the electroweak phase transition, making it a useful toy model for methods that could be applied to more realistic models. By seeing how the equilibrium state changes as a function of the model parameters, properties of the phase transition like the order and critical point of the transition can be found.

The Monte Carlo simulations used here were implemented from scratch in C++ for this thesis, and as such a large amount of the work involved dealt with making the simulations run correctly. The payoff to doing it this way is efficiency, as the custom software allows for large simulations to be performed in short amounts of time, using only consumer-grade hardware.

The structure of this thesis is as follows: Sections 2-4 deal with the underlying physical theory, starting with phase transitions in general and how they relate to cosmological models of the early universe in section 2, then moving on to quantum field theory at non-zero finite temperatures and lattice models thereof in section 3, and giving details on the cubic anisotropy model in section 4. Section 5 deals with the Monte-Carlo methods used to perform simulations on the lattice model, while section 6 explains the parameters used to perform the simulations, as well as how final results are derived from the simulation output. Finally, those results are presented in section 7, along with some discussion of those results. Some details on how simulations are implemented are given in the appendix.

2 Phase transitions

The most familiar examples of phase transitions in everyday life are the freezing and boiling of water, in which the substance changes into a solid or gas once a critical temperature has been reached. This critical temperature varies with pressure, and as such the phase transition is said to occur on a *critical line* in the $P - T$-plane. Phase transitions are not limited to this familiar kind of solid-liquid-gas transition, however. A general phase transition is characterized by some rapid, possibly discontinuous, change in the properties of a system, in response to a change in the thermodynamic variables of the system, such as the

temperature. At the transition point(s) the free energy,

$$F = U - TS, \tag{1}$$

remains constant, but its first- or second derivatives do not. It is in these parameters that the phase change of the system can be identified. For example, in the case of water, the specific volume and internal energy of the liquid and gaseous phases are significantly different at the same temperature and pressure. This is easily seen when boiling water: The temperature remains constant while energy is transferred into the water, and only resumes increasing once all the water has vaporized, assuming energy is still being added at that point.

2.1 Order parameters and transition types

For many phase transitions, the phases can most easily be distinguished through an *order parameter* [8]. In these cases, the order parameter describes the ordering of some microscopic property, such as the alignment of the magnetic dipoles in a metal. The usefulness of an order parameter comes from unambiguously showing what phase the system is, as it is often zero in one phase, and nonzero in the other. For example, in the case of a paramagnetic metal being cooled below its Curie temperature, the magnetic dipoles strongly influence each other to align in the same direction, causing ferromagnetism. The order parameter in this case is the absolute value of the magnetization $\mathbf{M}$, it is zero in the paramagnetic phase and nonzero in the ferromagnetic case.

A good choice of order parameter, or something similar to an order parameter, is helpful to properly analyse the transition properties of a system. In the type of statistical model used here, the order parameter is derived statistically from the microscopic properties of the system. As will be shown in section 4, it is not always possible to use the actual order parameter. In those cases where it is not, a stand-in for the order parameter needs to be defined.

In the case where the discontinuity appears in a first derivative of the free energy, the transition is said to be *first-order*. As the description of boiling water above indicates, it is one such phase transition. In this case, the specific volume $v(T)$ is one of the discontinuous parameters. Boiling illustrates a common property of these phase transitions: At the critical point, the system will be a in a mixed state that occupies both phases in equilibrium, in this case both steam and liquid water. This happens because any energy added to the system is "spent" to break the inter-molecular bonds keeping the water in its liquid phase, so none is available to increase temperature. The heat "lost" to breaking these bonds is known as the *latent heat* of the phase transition.

Another common property of first-order phase transitions that appears in the water-steam phase transition is the formation of bubbles. As a small portion of the total mass changes phase, surface tension between the two phases appears. It is then most efficient for the part of the system in contact with the bubble surface to change phase first, and the bubbles grow as the phase transition progresses. Depending on the system, this causes various turbulent behaviours.

In the case of water the bubbles rise to the surface and burst, freeing the vapour within and contributing to turbulent flow on the way up.

If the discontinuity appears in a second derivative, the transition is of second order. These transitions do not involve a latent heat, and as such there is no coexistence of separate phases. In this case, the order parameter is continuous at the phase transition. The paramagnetism-ferromagnetism transition is a good example of a second-order transition, as the magnetization changes continuously from 0 to a positive value.

2.2 Spontaneous symmetry breaking

In quantum field theory, phase transitions most often appear in the context of spontaneous symmetry breaking, where the ground state of a field lacks symmetries that are present in the Lagrangian of that field. The simplest example of this occurs in the case of a massive scalar field with a ϕ^4 interaction [21]:

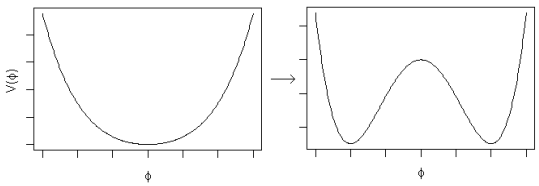

Figure 1: Scalar ϕ^4 potential with $m^2 > 0$ and $m^2 < 0$, going from a single minimum to two minima.

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4, \tag{2}$$

which has a Z_2 symmetry where $\phi \to -\phi$, and the order parameter has expectation value 0 as long as $m^2 > 0$ at tree level. In real systems, parameters like m^2 and λ are often functions of temperature, leading to the possibility of m^2 going below 0. When this happens, there are two distinct ground states $|0_+\rangle$ and $|0_-\rangle$, with tree-level expectation values at the minima of the potential:

$$\langle 0_\pm|\phi|0_\pm\rangle = \pm\sqrt{\frac{-6m^2}{\lambda}}. \tag{3}$$

In these ground states, the $\phi \to -\phi$ transformation maps $|0_+\rangle$ to $|0_-\rangle$, meaning the symmetry of the Lagrangian does not apply to the ground states. While the two ground states are equivalent, the system must fall into one or the other.

This kind of phase transition plays an important part in models of the early Universe [21, 5], where it is theorized that the fundamental forces were initially indistinguishable, but separated due to symmetry-breaking processes as the Universe expanded and the energy density decreased. The success of electroweak theory in unifying the electromagnetic and weak interactions at high energies, in particular, stands out as a key piece of evidence for this being the case. The *electroweak phase transition* of the standard model occurred around $2 \times 10^{-11}s$ after the Big Bang, separating the weak and electromagnetic interactions by breaking the $SU(2) \times U(1)$-symmetry of the electroweak gauge fields into the

$U(1)$-symmetry of electromagnetism. This happened because the *Higgs multiplet*, H, which is a complex doublet field, gains a vacuum expectation value from the potential

$$V(H) = -m^2|H|^2 + \lambda|H|^4, \tag{4}$$

when $m^2 > 0$, similar to the model in eq. 2. This vacuum expectation value is what induces the symmetry breaking through the various couplings between H and the gauge fields, resulting in the familiar weak- and electromagnetic forces. The theory of *electroweak baryogenesis*[18, 5] is a possible consequence of this phase transition, stating that if the phase transition is first-order, baryon number violation may have occurred during it, explaining the apparent prevalence of matter over antimatter in the Universe. However, in the Standard Model, a first-order phase transition is only possible if the Higgs boson has a mass lower than $m_h \lesssim 70$ GeV, which is much smaller than the LHC measurement of $m_h \approx 125$ GeV. Because of this, electroweak baryogenesis is only possible in modified versions of the standard model that make the phase transition first order.

The model considered in this thesis, the *cubic anisotropy model*, is not such a modification. It is, however, a much simpler scalar field theory with a potential similar to that in eq. 4, which is easily discretized for lattice simulations. The phase transition in the cubic anisotropy model is discussed in detail in section 4, while the equation used to estimate its order parameter at arbitrary parameter values is the main result of the next section.

3 Thermal quantum field theory on the lattice

Thermal field theory [8, 12] approaches quantum field theory at non-zero finite temperatures using many of the tools from statistical mechanics, in particular the *canonical ensemble.* The end goal of this section will be to derive the formula needed to calculate an observable A in the lattice approximation of a field with arbitrary parameters:

$$\langle \hat{A} \rangle = \frac{1}{Z} \int \prod_{\mathbf{x}} (d\phi(\mathbf{x})) \hat{A}[\phi(\mathbf{x})] e^{-S[\phi(\mathbf{x})]}. \tag{5}$$

To arrive at this result, first a quick review of statistical mechanics applied to quantum field theory at high temperatures is needed. From there, this section proceeds to the lattice model, detailing the effect of using it to approximate continuous fields, and the necessary steps to retrieve the continuum theory from lattice simulations.

3.1 Thermal field theory and statistical mechanics

The core idea of statistical mechanics is to connect the microscopic states of a system to macroscopic states described by a small number of thermodynamic parameters. In the case of statistical mechanics on a crystal lattice, for example,

the microscopic state is the position, internal energy, momentum, and other relevant properties of the atoms in the crystal. The macroscopic properties of the system, on the other hand, are the temperature, entropy, internal energy, and so on. Clearly, there are many microscopic states corresponding to each macroscopic state. To keep track of all these microscopic states, an *ensemble of states* is defined, described by the density matrix $\hat{\rho}$, [8] satisfying

$$tr[\hat{\rho}] = 1. \tag{6}$$

When a system is in the macroscopic state corresponding to $\hat{\rho}$, the expectation value of any observable is given by

$$\langle \hat{A} \rangle = tr[\hat{\rho}\hat{A}] = \sum_i \gamma^i \langle i|\hat{A}|i\rangle \,, \tag{7}$$

where γ^i is the weight of each microstate $|i\rangle$. The density matrix can be expressed purely in terms of the weighted states as

$$\hat{\rho} = \sum_i \gamma^i |i\rangle \langle i| \,. \tag{8}$$

The time evolution of $\hat{\rho}(t)$ is given by the Liouville-von Neumann equation, stating that

$$i\partial_t \hat{\rho}(t) = \hat{H}\hat{\rho}(t) + \hat{\rho}(t)\hat{H} = [\hat{H}, \hat{\rho}(t)], \tag{9}$$

where $\hat{H}$ is the standard *Hamiltonian operator* for the system, which measures the total energy of the system, i.e. the sum of the kinetic and potential energy.

In the case of thermal equilibrium, $\hat{\rho}(t)$ is constant in time, meaning that

$$i\partial_t \hat{\rho}(t) = 0. \tag{10}$$

Of the possible solutions to this equation, the one that is relevant here is the density matrix of the canonical ensemble, which is given by

$$\hat{\rho} = \frac{1}{Z} e^{-\beta \hat{H}}, \tag{11}$$

where $\beta = \frac{1}{k_b T}$ is the inverse temperature, and Z, the *partition function*, is the sum over all states weighted by energy, given by

$$Z = tr[e^{-\beta \bar{H}}]. \tag{12}$$

The canonical ensemble is the Gibbs ensemble that models a system kept at constant temperature in a heat bath [15]. This makes it particularly suited to the study of phase transitions, as the temperature can be set externally for calculations.

With this, the observable expectation value takes the form

$$\langle \hat{A} \rangle = \frac{1}{Z} tr[e^{-\beta \bar{H}} \hat{A}]. \tag{13}$$

Before moving on to the direct evaluation of Z as a *path integral*, the interpretation of the operator $\hat{\rho} = \frac{1}{Z}e^{-\beta\hat{H}}$ should be addressed. In statistical mechanics, this is a *Boltzmann distribution* [15], which gives the probability of a system to be in a state Φ with energy $H(\Phi)$ as a function of said energy. There is a second interpretation in the context of quantum field theory, however [12]. The time evolution operator $U(t_1, t_0)$, which evolves a system from its state at time t_0 to its state at t_1, is given by

$$U(t_1, t_0) = e^{i(t_1 - t_0)\hat{H}}, \tag{14}$$

and immediately bears resemblance to the $e^{-\beta\hat{H}}$-factor in the density matrix. In fact, the statistical mechanics of a 3D system at inverse temperature β can be directly interpreted as being equivalent to a 4-dimensional field theory with periodic "time" of period β. The use of an imaginary value for time is a mathematical trick that relates the thermal path integral to the usual Minkowskian path integral, with the effect of transforming the relativistic Minkowski metric by

$$ds^2 = dx^2 + dy^2 + dz^2 - dt^2 \to dx^2 + dy^2 + dz^2 + d\tau^2, \tag{15}$$

which is a 4D Euclidean metric. Thermal field theory takes place in this 4D spacetime analogue, with periodic "time" so that $\tau = \tau + \beta$.

3.2 The path integral

The next step is to reformulate Z in terms of a path integral, using the same approach to derive it as [12, 8, 21]. For a theory of a single scalar field, Z is evaluated by taking the trace over the basis of field eigenstates, $|\Phi\rangle$. Field eigenstates are defined by the eigenvalue equation

$$\hat{\phi}(\mathbf{x})\,|\Phi\rangle = \phi(\mathbf{x})\,|\Phi\rangle \tag{16}$$

Where $\hat{\phi}(x)$ is the field operator, normally used to extract field expectation values in an arbitrary state $|i\rangle$ by $\langle\phi(\mathbf{x})\rangle = \langle i|\hat{\phi}(\mathbf{x})|i\rangle$, and $\phi(\mathbf{x})$ is simply a field value at the point $\mathbf{x}$. The trace in this basis is

$$Z = \int d\Phi\,\langle\Phi|e^{-\beta\hat{H}}|\Phi\rangle\,. \tag{17}$$

With the substitution $\beta \to i(t_b - t_b)$, this becomes

$$Z = \int d\Phi\,\langle\Phi|e^{-i(t_b - t_a)\hat{H}}|\Phi\rangle = \int d\Phi\,\langle\Phi t_b|\Phi t_a\rangle\,, \tag{18}$$

where the inner product $\langle\Phi t_b|\Phi t_a\rangle$ is the *propagator* from the state $|\Phi t_a\rangle$ to $|\Phi t_b\rangle$, giving the probability amplitude of a system known to be in the state $|\Phi\rangle$ at time t_a to be in the same state at time t_b.

Defining $Z(\Phi_i, t_i; \Phi_j, t_j) = \langle\Phi_i t_i|\Phi_j t_j\rangle = \langle\Phi_i|e^{i(t_i - t_j)\hat{H}}|\Phi_j\rangle$ as the propagator from $|\Phi_i t_i\rangle$ to $|\Phi_j t_j\rangle$, the next step is to derive the path integral for Z. The

trick to doing this is to divide it into $n \to \infty$ small time steps $\delta t = t_{j+1} - t_j$ by repeatedly inserting the completeness relation

$$\int d\Phi_j \, |\Phi_j t_j\rangle \, \langle \Phi_j t_j| = 1, \tag{19}$$

resulting in

$$Z = \int d\Phi d\Phi_1 ... d\Phi_n \, \langle \Phi t_b | \Phi_n t_n \rangle \, \langle \Phi_n t_n | \Phi_{n-1} t_{n-1} \rangle \, ... \, \langle \Phi t_b | \Phi t_a \rangle \, . \tag{20}$$

Inserting time evolution back in at every step by $|\Phi_j t_j\rangle = e^{i\delta t \hat{H}} \, |\Phi_j t_{j-1}\rangle$ results in

$$Z = \int d\Phi d\Phi_1 ... d\Phi_n \, \langle \Phi t_b | e^{i\delta t \hat{H}} | \Phi_n t_b \rangle \, \langle \Phi_n t_n | e^{i\delta t \hat{H}} | \Phi_{n-1} t_{n-1} \rangle \, ... \, \langle \Phi t_b | e^{i\delta t \hat{H}} | \Phi t_a \rangle \, . \tag{21}$$

Now, each piece is evaluated separately, resulting in

$$\lim_{\delta t \to 0} \langle \Phi_i t_i | e^{i\delta t \hat{H}} | \Phi_{i-1} t_i \rangle = N e^{i\delta t \int d^3\mathbf{x} \mathcal{L}[\phi(\mathbf{x},t)_j, \partial\phi(\mathbf{x},t)_j/\partial t]}, \tag{22}$$

where $\mathcal{L}$ is the Lagrangian density of the field. The partition function is now the product of all these exponentials

$$Z = \mathcal{N} \int d\Phi d\Phi_1 ... d\Phi_n \exp\Big[\sum_j i\delta t \int d^3\mathbf{x} \mathcal{L}[\phi_j(\mathbf{x}, t_j), \partial_\mu \phi_j(\mathbf{x}, t_j)] \Big], \tag{23}$$

which, taking the limits $n \to \infty$ and $\delta t \to 0$ and dropping the normalization factor $\mathcal{N}$ results in

$$Z = \int_{\phi(\mathbf{x},t_1)=\phi(\mathbf{x},t_2)} \mathcal{D}[\phi(\mathbf{x},t)] e^{iS[\phi(\mathbf{x},t)]} \tag{24}$$

Here, the action $S[\phi(\mathbf{x},t)$ is a functional of $\phi(\mathbf{x},t)$ from time t_a to t_b, given by

$$S = \int_{t_a}^{t_b} dt \int d^3\mathbf{x} \mathcal{L}[\phi(\mathbf{x},t), \partial_\mu \phi(\mathbf{x},t)], \tag{25}$$

and the integration measure

$$\mathcal{D}\phi(\mathbf{x},t) = \mathcal{D}[\phi] = \int d\Phi d\Phi_1 ... d\Phi_n, \tag{26}$$

is understood to a functional integral over all possible 4D configurations of the field ϕ.

Going back to imaginary time by $t_a \to 0$ and $t_b \to i\beta$, the action becomes

$$S[\phi, \beta] = \int_0^\beta d\tau \int d^3\mathbf{x} \mathcal{L}_E[\phi(\mathbf{x},\tau), \partial_\mu \phi(\mathbf{x},\tau)], \tag{27}$$

where the relativistic Minkowsky Lagrangian $\mathcal{L}$ has been replaced by the Euclidean Lagrangian, $\mathcal{L}_E$. The difference between the two is that where a Minkowsky Lagrangian for a scalar field theory generally reads

$$\mathcal{L}[\phi] = (\partial_t - \nabla^2)\phi - V[\phi], \tag{28}$$

the Euclidean Lagrangian is given by

$$\mathcal{L}[\phi] = (\partial_t + \nabla^2)\phi + V[\phi]. \tag{29}$$

With all this, the final formulation of Z as a path integral in 4D euclidean space with a periodic time boundary condition is

$$Z = \int_{\phi(\mathbf{x},0)=\phi(\mathbf{x},\beta)} \mathcal{D}[\phi].e^{-S[\beta,\phi]} \tag{30}$$

Operator expectation values are similarly expressed by the path integral

$$\langle A \rangle = \frac{1}{Z} \int_{\phi(\mathbf{x},0)=\phi(\mathbf{x},\beta)} \mathcal{D}[\phi] A e^{-S[\beta,\phi]}. \tag{31}$$

In practice, it is often necessary to use an effective field theory approximation for calculation purposes. One way to greatly simplify theories at high temperatures in equilibrium is by *dimensional reduction*[5, 7]. Making use of the periodic time boundary condition, it is possible to match the 4D theory to a 3D effective field theory when the temperature is large compared to the masses involved, by exploiting the fact that as the time dimension is compactified, the zero-mode dominates time-dependent dynamics, allowing them to be integrated out. The 3D theory then has parameters that correspond to the ones in the 4D theory as functions of temperature, and calculations done in the 3D theory can be mapped back to the full 4D theory.

There are several benefits to using the dimensionally reduced 3D theory instead of the full 4D theory, in particular that it vastly simplifies numerical calculations. To do numerical calculations, however, it is necessary to make further approximations.

3.3 Lattice regularization

In all of quantum field theory, it is common to encounter the problem of divergences at very high and very low momenta [21]. Various regularization schemes exist to combat this, but lattice regularization [4], the one used here, is particularly well-suited to numerical work, as it reduces an otherwise infinitely granular theory to a finite data set.

The idea behind lattice regularization is to model the otherwise continuous fields using a discrete lattice with spacing a, modifying the theory so that in the continuum limit, $a \to 0$, the original theory is retrieved. The introduction of a restricts all coordinates to the form

$$x_\mu = a n_\mu. \tag{32}$$

For practical calculations, it is also necessary to restrict n to a finite range, so that

$$0 < n_\mu < N, \tag{33}$$

which corresponds to only considering the system within a finite volume V. Outside this range, the lattice is assumed to be periodic, so that

$$n_\mu + N = n_\mu. \tag{34}$$

With these restrictions, each field configuration is described by N^3 points, each containing one or more scalar, vector, tensor or spinor values depending on the type of field.

The lattice serves as a UV cutoff, providing an upper bound on the frequencies that can exist on the lattice. This can be seen when taking the Fourier transform of any arbitrary function $f(x)$ on a 1-D lattice:

$$\begin{aligned}\tilde{f}(k) = \sum_n f(an)e^{i2\pi akn} &= \sum_n f(an)e^{i2\pi akn + i2\pi n} \\ &= \sum_n f(an)e^{i2\pi a(k+1/a)n} = \tilde{f}(k+1/a),\end{aligned} \tag{35}$$

meaning that frequencies separated by $1/a$ are equivalent on the lattice, effectively restricting each component of the wave vector $\mathbf{k}$ to the range

$$-1/2a < k_i \leq 1/2a. \tag{36}$$

Known as the *Brillouin zone* of the lattice [6]. Equivalently, wavelengths are restricted to

$$\lambda \geq 2a \tag{37}$$

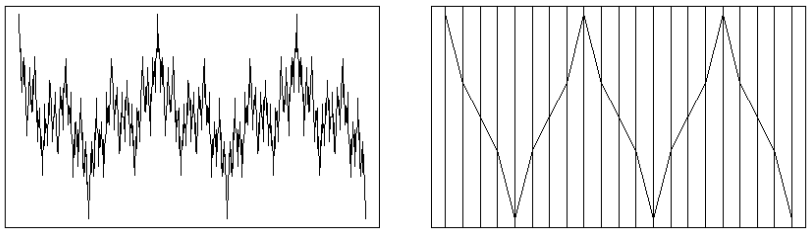

Figure 2: Lattice cutoff: Wavelengths less than twice the lattice spacing are removed by the lattice cutoff, as only the value at each discrete point is recorded

Introducing the lattice necessitates modifications to the action. A generic action integral in a 3D effective scalar field theory with a single field looks like

$$S = \int d^3\mathbf{x}\mathcal{T}(\phi(\mathbf{x}), \phi'(\mathbf{x})) + \mathcal{V}(\phi(\mathbf{x}))\big], \tag{38}$$

but with the action restricted to the lattice, it becomes a sum

$$S = a^3 \sum_{\mathbf{x}} \big[\mathcal{T}(\phi(\mathbf{x}), \phi'(\mathbf{x})) + \mathcal{V}(\phi(\mathbf{x}))\big]. \tag{39}$$

The factor of a^3 is commonly eliminated by rescaling the fields and field parameters. The details of this rescaling are given in section 4.

The kinetic term $\mathcal{T}$ can be expressed in terms of the Laplacian, ∇^2:

$$\mathcal{T}(\phi(\mathbf{x}), \phi'(\mathbf{x})) = -\phi(\mathbf{x})\nabla^2\phi(\mathbf{x}), \tag{40}$$

On the lattice, the Laplacian $\nabla^2\phi(\mathbf{x})$ needs to be defined in terms of nearest-neighbour derivatives. Defining $\mathbf{x} + i$ as the nearest-neighbour coordinate in a positive direction along the ith dimension, the simplest approximation of the Laplacian is

$$\nabla_L^2\phi(\mathbf{x}) = \frac{1}{a^2}\big[-6\phi(\mathbf{x}) + \sum_i \big(\phi(\mathbf{x}+i) + \phi(\mathbf{x}-i)\big)\big] \tag{41}$$

This only has $O(a)$ accuracy, however. This can be shown [2] by first taking the series expansion of an arbitrary function $f(a)$ around a point x:

$$f(a) = \sum_{i=0}^{\infty} \frac{f^n(x)}{n!}(a-x)^n. \tag{42}$$

Now, adding the series for $f(x+a) + f(x-a)$ results in

$$f(x+a) + f(x-a) = 2f(x) + a^2 f''(x) + \mathcal{O}(a^4). \tag{43}$$

Rearranging and dividing through by a^2 results in

$$\frac{\partial^2}{\partial x^2} f(x) = \frac{f(x-a) - 2f(x) + f(x+a)}{a^2} + \mathcal{O}(a^2), \tag{44}$$

Which leads directly to the Laplacian through $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. To eliminate the $\mathcal{O}(a^2)$ error term, an improved lattice Laplacian using next-to-nearest neighbours [17] is used instead:

$$\nabla_L^2\phi(\mathbf{x}) = \frac{1}{a^2}\Big[-\frac{15}{2}\phi(\mathbf{x}) + \frac{4}{3}\sum_i \big(\phi(\mathbf{x}+i) + \phi(\mathbf{x}-i)\big) - \frac{1}{12}\sum_i \big(\phi(\mathbf{x}+2i) + \phi(\mathbf{x}-2i)\big)\Big]. \tag{45}$$

Due to the rescaling mentioned above, the factor of $\frac{1}{a^2}$ above disappears in the Lagrangian.

Another feature of the lattice model that needs to be addressed is that of renormalization counterterms. When using perturbative methods, the need for renormalization arises when loop corrections to observables that depend on lattice parameters are found to be divergent in those parameters. In the case of

the lattice approach, a similar problem arises in that $\mathcal{O}(a)$ errors appear, which dominate results at large values of a [16].

The explanation for these errors is that they arise from effects dependent on the degrees of freedom that have been cut away by the lattice model, i.e. loop corrections to the path integral in momentum ranges outside of the lattice cutoff [16]. To compensate for this, counterterms are calculated to provide the higher-order corrections that are missing in the lattice theory.

The method to find these counterterms [16, 17] is to calculate correlation functions from which the renormalization parameters are derived in both the continuum and lattice theories, then derive counterterms that match the two, as functions of the renormalized parameters. This results in finite counterterms, which can be easily calculated and applied to produce bare parameters, which are used in the lattice calculations.

3.3.1 Lattice path integral

Recalling that the path integral was defined as a product over the infinite degrees of freedom in a continuous field, it is simple to see that it must be redefined on the lattice. Since there are only N degrees of freedom along each spatial dimension, the path integral measure returns to being a product of integration measures over each individual lattice point $\mathbf{x}$ [4],

$$\mathcal{D}[\phi]_L = \prod_{\mathbf{x}} (d\phi(\mathbf{x})) \tag{46}$$

Which leads to the partition function

$$Z = \int \prod_{\mathbf{x}} (d\phi(\mathbf{x})) e^{-S[\phi(\mathbf{x})]} \tag{47}$$

And finally, the operator expectation value formula is

$$\langle A \rangle = \frac{1}{Z} \int \prod_{\mathbf{x}} (d\phi(\mathbf{x})) \hat{A}[\phi(\mathbf{x})] e^{-S[\phi(\mathbf{x})]} \tag{48}$$

As will be shown in section 5, Monte-Carlo simulations are needed to evaluate this numerically, due to the high number of degrees of freedom in the integral.

3.3.2 Continuum and infinite volume limits

Two limits are needed to retrieve the original theory when using lattice regularization. The first is the infinite volume limit, $(aN)^3 \to \infty$, the second is the continuum limit, $a \to 0$ [4]. In practice, calculating this limit for a calculated measurement M works as follows [17]:

- Measure M by some procedure depending on observables in the lattice theory at a fixed value of a for different N
- Fit M as a function of N^{-3}, then extrapolate to $N^{-3} = 0$

- Repeat this procedure for several a, fit the infinite volume limits $M(a, 1/N^3 -> 0)$ as a function of a, and finally extract the continuum estimate of M by extrapolating to $a = 0$.

The details of how this procedure is implemented here are given in section 6.

4 Cubic anisotropy model

The cubic anisotropy model [17, 1] is a simple two-field scalar field theory with a phase transition, which is well-suited to being studied in lattice models. The model is defined by the partition function

$$Z = \int \mathcal{D}\phi_1 \mathcal{D}\phi_2 e^{-S[\phi_1,\phi_2]}, \tag{49}$$

where S is the action

$$S = \int d^3x \left[\sum_{i=1,2} \left(\frac{1}{2}\phi_i \nabla^2 \phi_i + \frac{m^2}{2}\phi_i^2 + \frac{\lambda_1}{24}\phi_i^4 \right) + \frac{\lambda_2}{4}\phi_1^2\phi_2^2 \right]. \tag{50}$$

The parameters here are given in the standard particle physics units, where $\hbar = c = k_B = 1$, so that [length] = [time] = [energy]$^{-1}$ = [temperature]$^{-1}$. The dimension of the scalar fields is (length)$^{-1/2}$, so that squared fields ϕ_i^2 are given in units of (length)$^{-1}$. The coupling parameters λ have units of (length^{-1}), and m^2 is given in terms of (energy)^2 = (length)$^{-2}$ [17].

Similar to the single-field model in eq. 2, this model has a phase transition where the fields gain a nonzero expectation value when $m^2 < 0$. At tree-level, this transition is always of second order, as the stationary point at $|\phi| = 0$ is unstable. However, at $\lambda_2 > \lambda_2$, it has been shown that radiative effects cause the phase transition to become first-order [17, 1], with the latent heat increasing as a function of the ratio λ_2/λ_1. The effect this ratio has on the potential is discussed further in section 4.1.

Used as a dimensionally reduced effective field theory for a 4-dimensional high-temperature scalar field theory, the parameters of the 3D theory correspond to those of the 4D theory by [17]

$$m^2 = m_{4D}^2 + O(T^2), \phi = \phi_{4D} T^{1/2}, \lambda = \lambda_{4D} T, S = S_{4D}/T, \tag{51}$$

meaning that lowering temperature in the 4D theory rescales both the 3D parameters and the field values, as well as lowering the mass of the 3D theory.

Regardless of parameter values, the cubic anisotropy model has discrete symmetries under both $\phi_i \to -\phi_i$ and $\phi_1 \leftrightarrow \phi_2$. Like in the single field model, the ground state at $m^2 > 0$ has tree-level expectation value $\langle\phi_1\rangle = \langle\phi_2\rangle$, while at $m^2 < 0$ there are several ground states, in which the symmetries of the Lagrangian are broken, detailed in section 4.1.

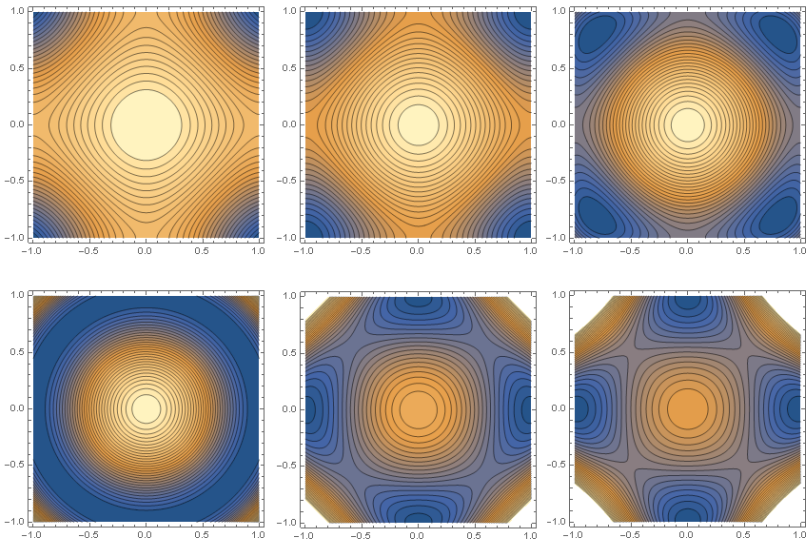

Figure 3: Contour plots of the cubic anisotropy model's potential for $x = 3\lambda_2/\lambda_1 = \{-0.5, 0, 0.5, 1.0, 2.0, 3.0\}$, showing the dependence of the shape on x.

4.1 Shape of the potential

There are several ratios of the parameters λ_1 and λ_2 of particular interest, due to the way they shape the potential $\mathcal{V}(\phi_1, \phi_2)$ in the $\phi_1\phi_2$-plane. [17] To illustrate this, it is convenient to rescale the fields by $\phi_i \to \sqrt{-\frac{\lambda_1}{6m^2}}\phi_i$. The factor $\sqrt{-\frac{\lambda_1}{6m^2}}$ comes from inverting the tree-level expectation value $v = \sqrt{\frac{-6m^2}{\lambda_1}}$ of the single-field model in eq. 2, which is a reasonable starting point to look for minima, as the λ_2-term is zero along both ϕ_i-axes. Applying the rescaling results in

$$\begin{aligned}\mathcal{V} = -\frac{3m^4}{\lambda_1}(\phi_1^2 + \phi_2^2) + \frac{3m^4}{2\lambda_1}(\phi_1^4 + \phi_2^4) + \frac{9\lambda_2 m^2}{\lambda_1}\phi_1^2\phi_2^2 \\ = X[-(\phi_1^2 + \phi_2^2) + \frac{1}{2}(\phi_1^4 + \phi_2^4) + x(\phi_1^2\phi_2^2)], \quad (52)\end{aligned}$$

where $X = \frac{3m^2}{\lambda_1}$ and $x = 3\lambda_2/\lambda_1$.

The ratio x has several special points where the potential gains additional symmetries:

- Firstly, when x=0, the potential is

$$\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2 = X\sum_i(-\phi_i^2 + \frac{1}{2}\phi_i^4) \quad (53)$$

Which corresponds to the fields being entirely separate, as there is no interaction in the kinetic term. From this, it seems reasonable to guess that the phase transition will have the same critical mass as a single-field model with the same λ_1. The minima of each ϕ_i are located at $\phi_i = \pm 1$, which corresponds to $\phi_i = \pm v$ when the coordinates are scaled back. For the combined system, then, the minima are located at $\{\phi_1, \phi_2\} = \{\pm v, \pm v\}$, and as such there are four distinct ground states to this value of x.

- The point where $x = 1$ has

$$\mathcal{V} = X[-(\phi_1^2 + \phi_2^2) + \frac{1}{2}(\phi_1^4 + \phi_2^4) + (\phi_1^2\phi_2^2)] = X[-(\phi_1^2 + \phi_2^2) + \frac{1}{2}(\phi_1^2 + \phi_2^2)^2] \tag{54}$$

Corresponding to another single-field theory under the substitution $\phi^2 = \phi_1^2 + \phi_2^2$:

$$\mathcal{V} = X[-\phi^2 + \frac{1}{2}\phi^4], \tag{55}$$

so again the phase transition must be equivalent to the one observed in the one-field theory. This time, however, there are infinitely many minima in a ring around $\phi_1 = \phi_2 = 0$, with radius 1, or v in the original coordinate system. This corresponds to an infinite number of equivalent ground states related by a rotation in the $\phi_1\phi_2$-plane, and stands out as the only value of x with this kind of continuous symmetry. The radial component is, however, completely equivalent to the same single-field model as when $x = 1$, and as such it should share its critical mass.

- At $x = -1$, the potential becomes

$$V = X[(\phi_1^2+\phi_2^2)+\frac{1}{2}(\phi_1^4+\phi_2^4)-(\phi_1^2\phi_2^2)] = X[(\phi_1^2+\phi_2^2)+\frac{1}{2}(\phi_1^2-\phi_2^2)^2]. \tag{56}$$

This potential has a serious problem: As the ϕ^4 part of the potential is zero along the diagonals $\phi_1 = \pm\phi_2$, the negative ϕ^2 term is never countered by the growth of the ϕ^4-term along this diagonal, and $\mathcal{V}$ diverges to negative infinity. As there are no minima to the potential, the ground state expectation value in this case is also divergent, and no proper analysis in the broken phase can be done.

- The final special point is when $x = 3$. In this case, $\lambda_1 = \lambda_2$, and the potential is:

$$\begin{aligned}\mathcal{V} = X[-(\phi_1^2 + \phi_2^2) + \frac{1}{2}(\phi_1^4 + \phi_2^4) + 3(\phi_1^2\phi_2^2)] \\ = X[-(\phi_1^2 + \phi_2^2) + \frac{1}{2}(\phi_1^2 + \phi_2^2)^2 + 2(\phi_1^2\phi_2^2)].\end{aligned} \tag{57}$$

While this does not look all that informative, a redefinition of the fields

by $\{\phi_1, \phi_2\} \to \frac{1}{\sqrt{2}}\{\phi'_1 + \phi'_2, \phi'_1 - \phi'_2\}$, results in

$$\mathcal{V} = X[-({\phi'}_1^2 + {\phi'}_2^2) + \frac{1}{2}({\phi'}_1^2 + {\phi'}_2^2)^2 + \frac{1}{2}({\phi'}_1^2 - {\phi'}_2^2)^2] \\ = \mathcal{V}_1 + \mathcal{V}_2 = X\sum_i[-{\phi'}_i^2 + {\phi'}_i^4], \quad (58)$$

which again corresponds to two uncoupled fields, but unlike before, the factor of $\frac{1}{2}$ preceding the ϕ^4-term is missing. This implies that the potential at $x = 3$ is not equivalent to the same one-field model as the ones at $x = 0$ and $x = 1$, and as such the critical mass could be different. Going back to the ϕ_1, ϕ_2-coordinates, the minima are easily found by noticing that the $\phi_1^2\phi_2^2$-term is zero along the orthogonal lines $\phi_1 = 0$ and $\phi_2 = 0$. From this, it's clear that the minima are located at $\phi_1, \phi_2 = \pm v, 0$ and $0, \pm v$.

4.2 Discretization on the lattice

The discretized action for the cubic anisotropy model on the lattice is

$$S = \sum_{\mathbf{x}}\left[\sum_{i=1,2}\left(-\frac{Z_\phi}{2}\phi_i\nabla^2\phi_i + \frac{Z_m(m^2+\delta m)}{2}\phi_i^2 + \frac{\lambda_1+\delta\lambda_1}{24}\phi_i^4\right) + \frac{\lambda_2+\delta\lambda_2}{4}\phi_1^2\phi_2^2\right], \quad (59)$$

where the Z- and δ-factors are multiplicative and additive renormalization coefficients as described in [17], and the spacing parameter a is baked into the units by rescaling so that

$$\phi_L = a^{1/2}\phi, m_L^2 = a^2m^2, \lambda_L = a\lambda, \quad (60)$$

defining a set of units on the lattice that are dimensionless. The symmetries described in section 4.1 are preserved by the renormalizations [17], meaning the properties of the system described there should hold on the lattice.

The order parameter used for the phase transition of this model is:

$$\phi_{av}^2 = \frac{1}{N^3}\sum_{\mathbf{x}}(\phi_1(\mathbf{x})^2 + \phi_2(\mathbf{x})^2), \quad (61)$$

which is strictly speaking not an order parameter [17], as it is nonzero in both phases. The alternative, however, would be to use

$$\phi_{av} = \sqrt{(\bar{\phi_1})^2 + (\bar{\phi_2})^2}, \bar{\phi_i} = \frac{1}{N^3}\sum_x \phi_i. \quad (62)$$

While this order parameter has a nonzero expectation value for each of the individual ground states of the system, its expectation value for a path integral

is 0, due to the symmetric nature of the potential. This makes ϕ_{av} unsuitable for this situation, and ϕ_{av}^2 must be used instead. Further, as will be shown in section 6, this choice of order parameter is helpful when reweighting.

The order parameter is also corrected by [17]

$$a\langle\phi^2\rangle = Z_m\langle\phi_a v^2\rangle - \delta\langle\phi_a v^2\rangle \tag{63}$$

From here on out, the physical volume of the lattice is defined in terms of λ_{1L}, and continuum limit approximations are taken at $\lambda_{1L} \to 0$. With this convention, measurements of the critical mass are given in terms of m_L^2/λ_{1L}^2.

5 Monte-Carlo Methods

When it comes to actually evaluating the path integral in eq. 48, the sheer dimensionality of the problem is a major detriment to numerical solutions. Treating each $\phi_i(\mathbf{x})$ as a separate degree of freedom, there are $2N^3$ such degrees of freedom, and a Monte-Carlo approach is necessary to converge to a good estimate in a reasonable time frame.

The basic application of the Monte-Carlo method [4] to this problem is to draw n field configurations $\Phi = \Phi_0, ...\Phi_n$ randomly from the distribution

$$\Phi \sim \frac{1}{Z}e^{-S}, \tag{64}$$

then, an observable's expectation value is approximated using

$$\langle\bar{A}\rangle = \frac{1}{n}\sum_{i=0}^{n} A[\Phi_i]. \tag{65}$$

Provided that a way to randomly generate *uncorrelated* samples from the target distribution can be used, the estimate $\langle\bar{A}\rangle$ will converge to the true value $\langle A\rangle$ with error $\frac{\sigma_A}{\sqrt{n}}$, where σ_A is the standard deviation of A in the sample Φ:

$$\sigma_A^2 = \frac{1}{n-1}\sum_{i=0}^{n}(A[\Phi_i] - \langle\bar{A}\rangle)^2 \tag{66}$$

The problem is that getting a sample with the distribution in eq. 64 is a non-trivial issue. A simple method to generate a sample $\mathbf{X}$ from an arbitrary distribution $P(\mathbf{x})$ is *rejection sampling* [14], in which for every iteration i, a proposed configuration $\mathbf{y} \sim g(\mathbf{y})$ is generated from an easily sampled proposal distribution g. Then, $P(\mathbf{y})$ is calculated from the target distribution, and an accept/reject step is performed, in which y is either accepted as the ith sample in $\mathbf{X}$, or rejected, with acceptance probability $p(\mathbf{y})/Mg(\mathbf{y})$. M is a constant that ensures $Mg(\mathbf{x} > p(\mathbf{x}))$ for all $\mathbf{x}$. Rejection sampling has two problems that make it unsuitable for this situation, however. The first is that the choice of an appropriate $g(x)$ is itself difficult. More importantly, however, there is

the problem that outside a very narrow range of field configurations around the stationary points of the action, $P(\phi)$ is exponentially suppressed by the action growing large. As a result, rejection sampling approaches will end up with a very high rejection rate, ending up with a large percentage of the computation time wasted on rejecting samples.

This kind of high-dimension, small-peaked problem is what *Markov chain Monte-Carlo* simulations[14, 20] are for. The fundamental idea of any MCMC strategy is to define some transition rule which results in a *Markov chain* that converges to the desired distribution. The code developed for this thesis is just such a simulation strategy, implementing a variant of the Metropolis-Hastings sampling algorithm from scratch in the C++ programming language, using the CUDA GPU programming API. The following describes the algorithm more generally, while the specifics of this implementation are found in the appendix.

Starting from some arbitrary configuration $\Phi_0 = \{\phi(\mathbf{x}_{0,0,0}), ...\phi(\mathbf{x}_{N,N,N})\}$, for every step, Φ_{i+1} is acquired by a simple procedure:

1. Propose a perturbation Φ' of Φ_i using the proposal distribution $T(\Phi; \Phi')$, which can be any easily sampled distribution as long as it is symmetric, i.e. $T(\Phi; \Phi') = T(\Phi'; \Phi)$.

2. Calculate $\Delta S = S[\Phi'] - S[\Phi_I]$

3. Generate a random number $b \sim U(0, 1)$

4. Accept or reject the sample $\boldsymbol{\phi}'$ based on b and ΔS. If $b \leq e^{-\Delta S}$, the sample is accepted, and $\boldsymbol{\phi}_{i+1} = \boldsymbol{\phi}'$. Otherwise, the sample is rejected, and $\boldsymbol{\phi}_{i+1} = \boldsymbol{\phi}_i$

A good choice of proposal distribution is critical. The first question is one of whether to sample every component in $\boldsymbol{\phi}$ all at once, or to instead focus on some small number of components for each iteration. In this case, a pseudo-Gibbs sampling scheme was used, so that at each iteration, only one lattice point $\mathbf{x}$ is resampled from a normal distribution with mean $\phi(\mathbf{x})$. This resampling affects the whole ϕ_1, ϕ_2-doublet, so that

$$\phi_i(\mathbf{x})' \sim N(\phi_i(\mathbf{x}), \sigma), \tag{67}$$

making the Markov chain a random walk, but only moving along two of the system's degrees of freedom at every step.

The standard deviation σ needs some fine-tuning for optimal results. If the variance is too large, the rejection rate will also be large, meaning the simulation spends most of its time on wasted samples. On the other hand, if the variance is too small, the simulation will just perform a random walk in a small area, with the accept/reject step having almost no impact. When no other method of fine-tuning presents itself, simply recording the rejection rate for a short test run and manually adjusting σ until the rejection rate falls within a good interval $(0.2 - 0.5)$ is sufficient.

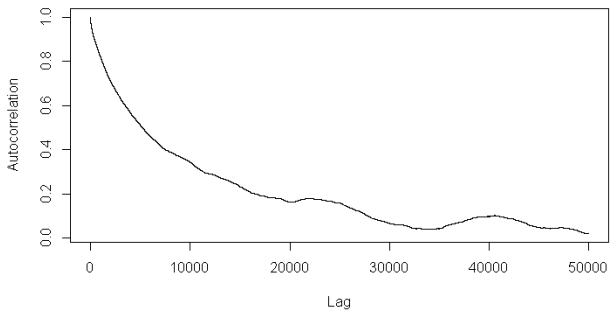

Figure 4: Autocorrelation curve of a strongly autocorrelated sample, $\tau(\phi^2) \approx 10^4$ sweeps.

Once the configurations Φ have been sampled, any number of useful observables can be calculated using eq. 65. Due to the size of Φ when the lattice is large, it is rarely practical, or even feasible, to save the raw configurations. For example, if Φ is $400,000$ samples of lattice configurations at $N = 80$, an uncompressed file containing the entire lattice at every step using double-precision floating point numbers would be at least $2 \times 80^3 \times 400000 \times 8 \approx 3.3 \times 10^12$b= 125GB in size, meaning just 16 simulation runs would be needed to fill up a mid-range hard drive with 2 TB storage space. Instead of doing this, it is much more economical to only save the observables $A[\Phi_i]$, meaning the simulation code only keeps the current configuration Φ_i in memory at any one point.

The big issue with samples acquired using this algorithm is that they are very strongly autocorrelated. Because of this, a large number of *sweeps* is needed. One sweep is defined as a number of sampling steps equal to the number of lattice points, N^3, in which the sampler has had the chance to visit every single lattice point once. As detailed in section A, a major advantage of the GPU implementation used here is that it can run several independent sampling steps in parallel. However, the parallelization makes it infeasible to record any observables during a sweep, due to the need to access every lattice point to calculate observables. Because of this, observables are at most recorded for every sweep in this implementation. To measure the autocorrelation of the samples, the integrated autocorrelation time of an observable measured by the MCMC sampler is given by

$$\tau(A) = \frac{1}{2} + \sum_{j=1}^{\infty} \rho_j \tag{68}$$

Where ρ_j, the autocorrelation at lag j, is given by

$$\rho_j = corr(\{A[\Phi_1], ...A[\Phi_{n-j}]\}, \{A[\Phi_{j+1}], ..., A[\Phi_n]\}) \tag{69}$$

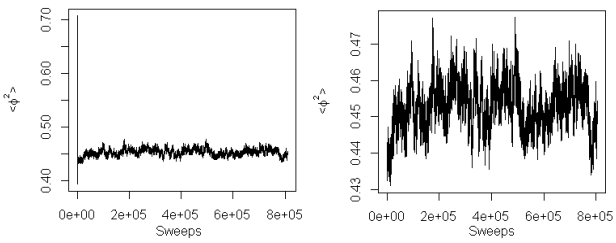

Figure 5: The first few sweeps differ significantly from the rest of the markov chain (left), but when a burn-in period of 5000 sweeps is removed, the sample is very stable. (right)

The standard error of the estimated observable is then given by

$$\sigma_{\bar{A}} = \sqrt{\frac{n}{2\tau(A)}\sigma_A^2} \tag{70}$$

The factor $n/2\tau(A)$ is known as the *effective sample size*, and as the simulation results show in section refresults, it can become very small even at large sample sizes.

Another factor to consider in these simulations is the *burn-in time*. For a number of iterations depending on Φ_0, Φ does not follow the target distribution, but instead moves towards the nearest maximum. To ensure that the burn-in period does not affect results, these samples are discarded.

Finally, to help the simulation converge to the desired distribution more efficiently, an *overrelaxation* update [11] can be performed every m sweeps for some small number m. In the overrelaxation update, Φ_i is transformed under one of the spontaneously broken symmetries of the Lagrangian before the next Metropolis-Hastings update can take place. This way, the sampler is "forced" to visit all the peaks of the distribution related by these symmetry transformations. While the simulation should in principle be able to do this without the overrelaxation update, the probability of moving past the energy barriers between wells is often low enough that the minor additional computational effort to add in overrelaxation updates is justified.

6 The simulations

With all the necessary tools and theory in place, the simulations that were performed for this thesis can be discussed. This section describes the purpose and parameters of the simulations, as well as the processing of the output necessary to get the final measurements, which are presented in section 7.

The simulations are performed with two simple goals:

- Determine the order of the phase transitions at several $x = 3\lambda_2/\lambda_1$, and check if this matches existing theory.

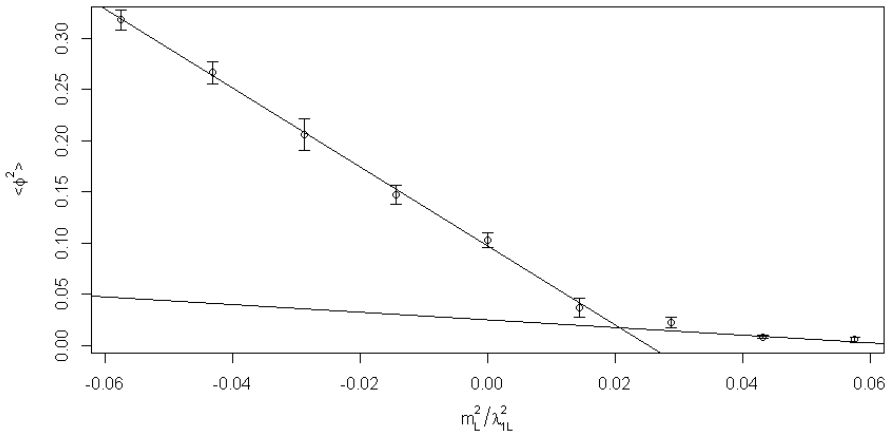

Figure 6: Large-scale estimation of the critical mass at $a\lambda_1 = 0.5$, $N = 48$ by linear model intersection, providing the estimate $m_L^2/\lambda_{L1}^2 = 0.0205 \pm 0.0057$ at 95% confidence.

- Determine the critical mass of the phase transitions at those x.

To help validate the simulations in the initial search, as well as checking convergence and adjusting the proposal distribution, several observables are recorded every sweep:

- The action S, a useful indicator of algorithm convergence, as it should stabilize after some number of iterations.
- The order parameter ϕ_{av}^2.
- The means of each component field, $\phi_{i,av} = N^{-3} \sum_{\mathbf{x}} \phi_i(\mathbf{x})$, which are helpful to see the effect x has on the shape of the potential.
- The rejection rate for each sweep, which is not used for analysing the final output, but rather to indicate if the proposal distribution is good.

For both of these goals, an initial estimate of the critical mass is needed. As long as the transition is second-order or weakly first-order, the first estimate for each x can be acquired by doing short ($10^4 < n < 10^5$) simulation runs, keeping λ_{1L} and N constant while varying m in some small range around 0, $|m^2| < \lambda_1^2$. Then, calculating $\langle \bar{\phi}^2 \rangle$ for each run, along with an error estimate $\sigma_{\langle \bar{\phi}^2 \rangle}$, the critical mass is estimated using a piecewise linear model fit:

$$\langle \bar{\phi}^2 \rangle (m^2) = \begin{cases} am^2 + b & m^2 < m_{break}^2 \\ cm^2 + d & m^2 \geq m_{break}^2 \end{cases} \tag{71}$$

The breakpoint m^2_{break} is simply chosen by iterating over the simulation values of m^2, and picking the one that yields the best combined coefficient of determination, R^2 for the piecewise function. Using these fit parameters, it is easy to estimate m_c^2 as the intersection of the two linear models:

$$\bar{m_c^2} = \frac{d-b}{a-c}. \tag{72}$$

By the usual formula for propagation of uncertainty

$$\sigma_f = \sqrt{\sum_{i=1}^{n} \sigma_{x_i}^2 \Big(\frac{\partial f}{\partial x_i}\Big)^2}, \tag{73}$$

the standard deviation of this estimate is given by

$$\sigma_{\bar{m^2}} = \sqrt{(\sigma_a^2 + \sigma_c^2)\frac{(d-b)^2}{(a-c)^4} + (\sigma_b^2 + \sigma_d^2)\frac{1}{(a-c)^2}}. \tag{74}$$

This method only determines a very rough initial estimate, however. Near the critical mass, lattice estimates of the order parameter do not behave as a piecewise linear model, regardless of the order of the phase transition, making the breakpoint of a piecewise linear model an insufficient estimate for proceeding.

If the transition is of first order, the discontinuity in $\langle\phi^2\rangle$ ensures that a simple intersection of two linear models will be off-target by a significant amount as the strength of the phase transition increases. Instead, as in [17], the critical mass is estimated by making use of the fact that in first-order transitions the two phases coexist at the critical mass, defining it as the mass where $P(\phi^2_{av})_{m^2}$ is a two-peaked distribution with peaks of equal volume.

In the case of second-order transitions, there is only one peak to the distribution, meaning the critical mass cannot be estimated by equalizing the peak sizes. Instead, the curve intersection method used earlier is applied with some modifications. Observing that near the critical mass, the measured values of $\langle\phi^2_{av}\rangle$ curve upwards before transitioning into the broken phase and resuming linear development, a nonlinear term is added to the piecewise model fit:

$$\bar{\langle\phi^2\rangle}(m^2) = \begin{cases} am^2 + b & m^2 < m^2_{break} \\ cm^2 + d + \frac{1}{fx+g} & m^2 \geq m^2_{break} \end{cases} \tag{75}$$

And the critical mass estimate is found from the intersection of the two pieces, which is found numerically

To avoid re-running simulations indefinitely when trying to find the best estimate of the critical mass, once a simulation has been run to a good amount of precision, the expectation value of the order parameter at different masses

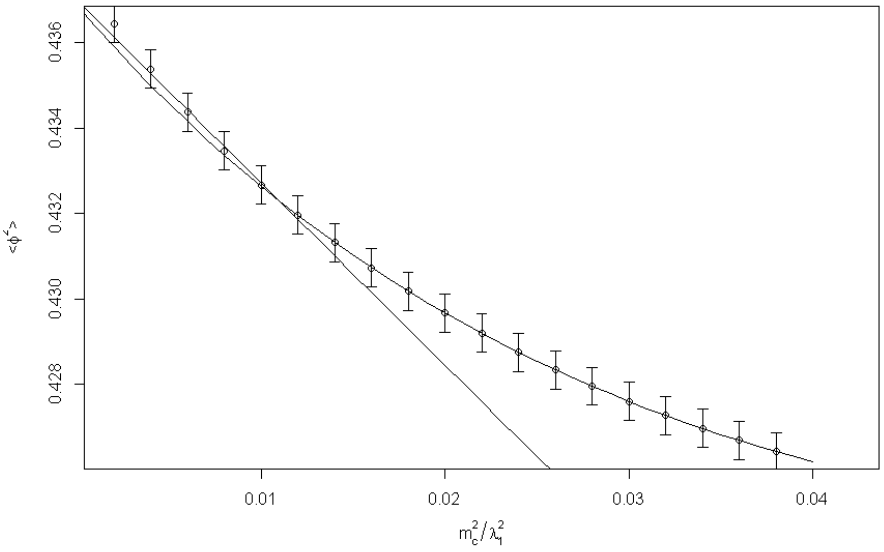

Figure 7: Estimating the critical mass of a phase transition by curve fitting, taking the intersection of the two curves as the critical mass.

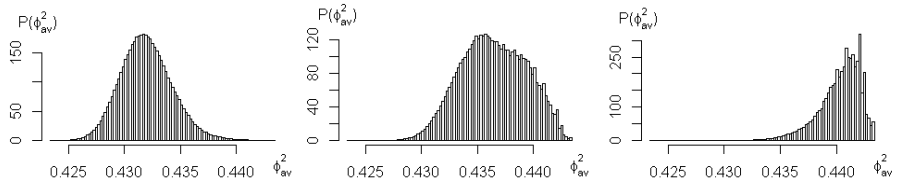

Figure 8: Reweighting a histogram from $m^2/\lambda_1^2 = 0.012$ (left) to $m^2/\lambda_1^2 = 0.022$ (centre) and $m^2/\lambda_1^2 = 0.032$ (right). Estimates from the rightmost histogram are unreliable, and reweighting much further in the same direction fails entirely.

can be acquired by *reweighting*. The idea behind this reweighting is that in the probability distribution

$$P_{m^2}(\phi_{av}^2) = \int d\phi' e^{-S_m(\phi')} \delta(\phi_{av}'^2 - \phi_{av}^2), \tag{76}$$

ϕ_{av}^2 appears as the sum over all the ϕ^2-terms in the action divided by $\frac{m^2}{2}N^3$, which means that the distribution at any other value of m^2 can be acquired by

$$P_{m_2^2}(\phi_{av}^2) \propto exp\left[\frac{N^3}{2}(m_1^2 - m_2^2)\phi_{av}^2\right] P_{m_1^2}(\phi_{av}^2). \tag{77}$$

Using the reweighted probability distribution, $\langle\phi^2\rangle$ is measured at the new m^2 by

$$\langle\phi^2\rangle_{m^2} = \int d(\phi_{av}^2) P_{m^2}(\phi_{av}^2)\phi_{av}^2 \tag{78}$$

The range of masses that this can be reliably applied to in practice is limited, however. If, for example, the peak of $P_{m^2}(\phi_{av}^2)$ is located outside of the region that has actually been sampled by the simulation, no amount of reweighting on that sample will give the right distribution. More rigorously, the relative error for a particular value of ϕ_{av}^2, $\delta P_{m^2}(\phi_{av}^2)/P_{m^2}(\phi_{av}^2)$, remains constant under reweighting [17], which makes the absolute error of histogram bins grow exponentially as tail values are amplified by the reweighting. Because of this, the simulations need to get close enough to the critical mass that both peaks of the distribution are well-sampled before reweighting can handle the rest. This is also why the initial estimation runs at a large m^2-range had to be performed separately, as the ranges of ϕ_{av} in samples taken at very different m^2 easily end up with zero overlap.

Several parameter combinations were used to perform the simulations. As discussed in section 4, x has a large effect on the potential, and as such the primary focus is to compare the behaviour at different values of it in the simulations. The three values $x = 0, 1, 3$ were all simulated on to verify the expected behaviour that the transition is second-order, that $x = 0$ and $x = 1$ give equivalent results for the critical mass, and that $x = 3$ gives a different result to

the two. For first-order transitions, $x = 3.003$ was used. The reason for this rather low value is that for $x > 3$, the transition is expected to have an increasing latent heat of the phase transition in x. This presents a problem for the MCMC algorithm, as the region between the two peaks is exponentially suppressed.[17] While *multicanonical* simulations could fix this, there was not enough time to implement and debug this method in the simulation program, and as such $x = 3.003$ was chosen so that the phase transition is only weakly first order, with a small energy barrier to tunnel through. This allows the algorithm to get a representative sample of the distribution in a reasonable number of iterations, without having to use multicanonical simulations.

Setting the lattice scale by $\lambda_{1L} = 0.5$, the critical mass for each of these four cases was determined through the methods described above, at several lattice volumes. For the second-order transitions, the simulations were run at $N = 70, 76, 80, 90$, while for the first-order transition, $N = 80, 90, 100$ were used. Performance issues were the primary reason not to use $N = 100$ for the second-order transitions, as simulation runs as long as the one performed at $N = 100$ were not feasible to perform at all x. Using the measurements at these varying N, the infinite volume limit was extracted by fitting m_c^2/λ_1^2 as a function of $1/N^3$ as a linear model, and extrapolating to $1/N^3 \to 0$. Again due to time constraints, the continuum limit $\lambda_{1L} \to 0$ was not evaluated in the final simulation set.

7 Results and discussion

As discussed in section 4, the potential's minima vary in the $\phi_1\phi_2$-plane depending on x. By comparing scatter plots of the recorded observables $\phi_{1,av}, \phi_{2,av}$ from simulations in the broken phase to contour plots of the potential at the same parameters (Figure 9), it is easy to see that the correct distributions are being sampled from. Similarly, for the first order phase transitions, it can be verified that the symmetric and broken phases have been sampled from by inspecting scatter plots and seeing

While initial runs with the linear model intersection method at low lattice sizes ($N \leq 60$) were unsuccessful at finding the first-order phase transition at $x = 3.003$, increasing the lattice size to $N = 70$ and up resulted in visibly bimodal $P(\phi_{av}^2)$-distributions, from which the critical mass could be measured.

m_c^2/λ_1^2	N = 70	N=76	N = 80	N = 90	N = 100	$N \to \infty$
$x = 0$	0.00258	0.00509	0.00749	0.00786		0.0130 ± 0.0018
$x = 1$	0.000151	0.00406	0.00664	0.00794		0.0150 ± 0.0036
$x = 3$	0.00484	0.00517	0.00950	0.00989		0.0152 ± 0.0062
$x = 3.003$			0.00716	0.0090	0.0096	0.0124 ± 0.0012

Table 1: Table of m^2/λ_1^2 for each simulation at $\lambda_{1L} = 0.5$, as well as continuum limits

The estimated infinite volume, continuum-limit critical masses are given in

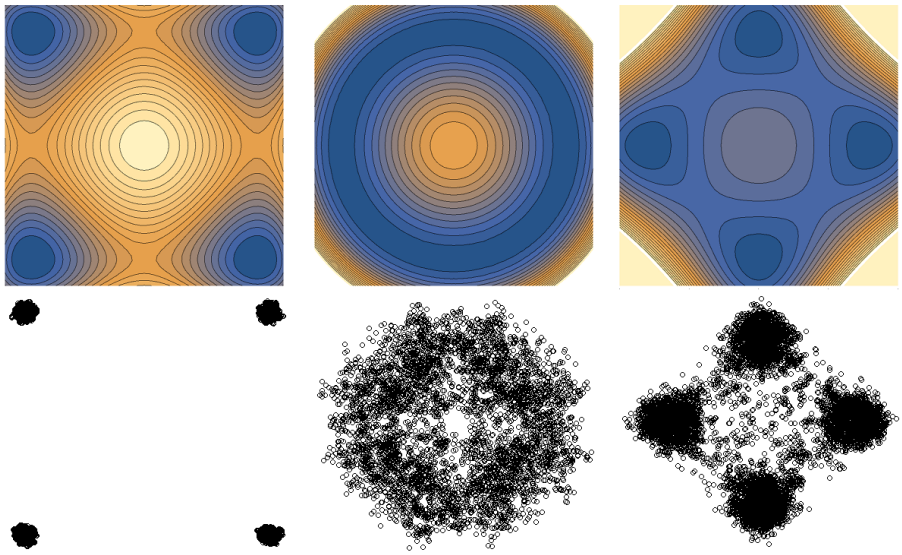

Figure 9: Correspondence between contour plots of the potential as a function of ϕ_1 and ϕ_2, and scatter plots of $phi_{1,av}$, $\phi_{2,av}$ in simulation results at $x = 0, 1, 3$.

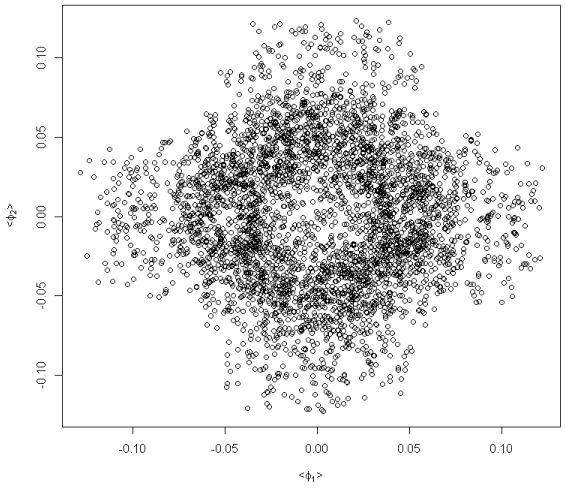

Figure 10: Scatter plot at x=3.003, $m^2/\lambda^2 = 0.0146$, indicating that simulation has sampled from both broken and symmetric phases. (N=80)

table 1. The $x = 0$ and $x = 1$ simulation runs display a close similarity, as predicted in section 4. While $x = 3$ was speculated to have some level of deviation, the infinite volume estimate does not support this to any significant degree. As predicted, however, no bimodal distributions occurred at any of these x. With no evidence of a first-order transition, these values of x appear to be second order, as expected.

The case where $x = 3.003$ proved to be troublesome, in that early quick estimates were far off target, needing iterative simulations to get close enough for successful reweighting from one phase to another. Long autocorrelation times corroborated this issue, for example the simulation at $N = 100$ had an integrated autocorrelation time $\tau = 2052$ sweeps. In the end, however, the approach of reweighting to a bimodal distribution with peaks of equal size succeeded in producing estimates of the critical mass.

There are several possible sources of systematic error that could apply to these measurements. The most obvious is the possibility of undiscovered bugs in the simulation code. While runs to validate that the code behaves correctly accounted for the majority of simulation time for this thesis, the possibility that bugs still remain cannot be dismissed. Beyond bugs in the implementation, the method used to estimate the critical mass at second order may need refinement. Reweighted estimates of $\langle \phi^2 \rangle$ for multiple m^2 sometimes gave ambiguous results with a barely visible break point, leading to a small difference between the fitted models and high error margin for the fit parameters. The reason for this appears to be that simulations at masses that are too high lack the statistics to reweight into the broken phase. As such, the easiest way to rectify this issue would simply be to run more simulations at lower masses, as well as considering improvements to the piecewise fits used in the estimation process.

The simulation code is built to be easily extended, and as such may be applied to different projects in the future. Several possible improvements to the code could be made, such as a full implementation of multicanonical sampling, further optimization for performance, and even the possibility of running multiple instances of the program at the same time. Another important improvement is the ability to recover simulations in the case of interruptions. As has been mentioned, time constraints reduced the amount of data used for final estimates, and a major factor in this was unexpected interruptions to the simulations, such as due to overheating.

Acknowledgements

I would like to thank my supervisor, Anders Tranberg, for being an excellent source of guidance and direction throughout the work on this thesis, without which it would not have been possible to complete. I would also like to thank my family and friends for their constant support throughout the project.

A Implementation details

As has been mentioned, writing simulation code accounted for the majority of the work done for this thesis. In this appendix, the details of that implementation are discussed, including some notes on necessary optimizations to make the simulations run fast. The focus on speed was motivated primarily by the hardware used. As the author of this text owns a rather powerful computer, it would be a waste to not utilize the available resources properly. In particular, the graphics card is a very powerful resource, easily capable of millions of parallel computations[1]. As an example of the benefit to doing this, consider the following: In the early stages of development, the first batch of 16 mass-varying simulations at $N = 32$ took 19 hours, with a meagre sample size of only 1000 sweeps. Towards the end of the project, a single simulation at $N = 90$ with a sample size of 810000 sweeps could be performed in the same timeframe.

The need to run N^3 sampling steps per sweep is a major drain on resources when the sampling algorithm is run linearly. As N increases, the time to compute a single sweep increases increase as to $t \propto N^3$. In an effort to combat this problem, the simulation code was written to run in parallel on the computer's graphics card. GPUs are powerful parallel processing devices, easily capable of running calculations on every single lattice point simultaneously. A simple way that this benefits simulation speed is in evaluating any observable that is a sum of some quantity over the whole lattice by *parallel reduction*[19], in which a sum over a large array A of size n is evaluated by $n/2$ parallel threads. The basic algorithm works as follows:

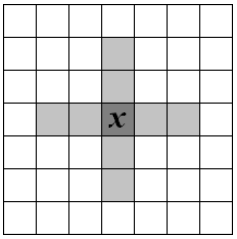

Figure 11: Nearest- and next-to-nearest neighbours of the lattice point $\mathbf{x}$ in two dimensions

- Allocate an array B which all threads have access to
- Each thread i computes $B_i = A_i + A_{i+n/2}$
- Every jth step after the first, each thread calculates $B_i = B_i + B_{i+n/(2(j+1))}$, effectively summing one half of the array with the other every step
- This is repeated until all values in A have been added together in B_0, which is the final result.

While the specifics of the implementation vary, the benefit of this algorithm is a theoretical $\mathcal{O}(log(n))$ computation time, a significant improvement over the $\mathcal{O}(n)$ time to sum over an array linearly.

[1] The theoretical limit of the GPU in question is $65535^3 * 1024$ [19], but no simulations done here came close to that number

Parallelization has its own issues, however. If every single point was sampled at the same time, calculating ΔS and performing an accept-reject step for each stage, the possibility of a *race condition* would emerge, where multiple parallel threads try to read or write at the same memory address. The trick to avoiding this, as well as massively reducing the computation work to calculate ΔS, is to realize that for each sampling step, ΔS does not depend on the entire lattice. Instead, it only depends on the region of nearest- and next-to-nearest neighbours of the lattice site being considered when using the $\mathcal{O}(a^2)$ Laplacian. For a sampling step taking place at a lattice position $\mathbf{x}'$, in which ϕ is the sampling configuration of the previous step and θ is the proposed field configuration where $\theta(\mathbf{x}')$ is the single altered field value, ΔS is given by

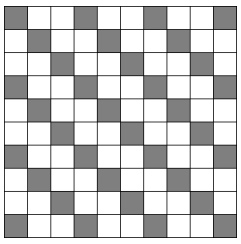

Figure 12: Independent diagonals in the 2D plane.

$$\Delta S = S[\theta] - S[\phi] = \sum_{\mathbf{x}} \big[\mathcal{L}(\theta(\mathbf{x})) - \mathcal{L}(\phi(\mathbf{x}))\big]. \tag{79}$$

However, since $\Theta = \Phi_i$ for all points except $\mathbf{x}$, the vast majority of terms in the sum cancel, leaving only the ones that depend on $\delta\phi = \theta - \phi(x)$. Dividing the Lagrangian into kinetic and potential terms

$$\mathcal{L}(\phi(\mathbf{x}), \nabla^2\phi(\mathbf{x})) = \mathcal{T}(\phi(\mathbf{x}), \nabla^2\phi(\mathbf{x})) + \mathcal{V}(\phi(\mathbf{x})), \tag{80}$$

it is apparent that all $\mathcal{V}$-terms except the ones at $\mathbf{x}$ cancel, as well as all $\mathcal{T}$ except at $\mathbf{x}$, nearest neighbours of $\mathbf{x}$, and next-to-nearest neighbours, through the Laplacian in each of these terms relying on $\phi\mathbf{x}'$. This expression

$$\begin{aligned}\Delta S = \Delta\mathcal{T}(\mathbf{x}'), \nabla^2(\theta(\mathbf{x}') \\ + \sum_{j=\pm\{1,2\}} \sum_{\mathbf{i}} \big[\Delta\mathcal{T}(\mathbf{x}' + j\mathbf{i}), \nabla^2(\theta(\mathbf{x}' + j\mathbf{i}))\big] + \mathcal{V}(\theta) - \mathcal{V}(\phi(\mathbf{x})),\end{aligned} \tag{81}$$

is not without issue however, as it still contains many terms that cancel. This is seen by examining $\Delta\mathcal{T}(\mathbf{x})$ more closely, noting that its form changes depends on whether it is evaluated at $\mathbf{x}'$ or its neighbours. With $\Delta\phi(\mathbf{x}') = \theta(\mathbf{x}') - \phi(\mathbf{x}')$ and $\Delta\phi(\mathbf{x}')^2 = \theta(\mathbf{x}')^2 - \phi(\mathbf{x}')^2$, the parts of the $\mathcal{T}$-terms that do not cancel are

$$\begin{aligned}\Delta\mathcal{T}(\mathbf{x}') = -\frac{15}{2}\Delta\phi(\mathbf{x}')^2 \\ + \Delta\phi(\mathbf{x}') \sum_i \big[\frac{4}{3}(\phi(\mathbf{x}+i) + \phi(\mathbf{x}-i)) - \frac{1}{12}(\phi(\mathbf{x}+2i) + \phi(\mathbf{x}-2i))\big],\end{aligned} \tag{82}$$

$$\Delta\mathcal{T}(\mathbf{x}' \pm \mathbf{j}) = \frac{4}{3}\phi(\mathbf{x}' \pm \mathbf{j})\Delta\phi(\mathbf{x}'), \tag{83}$$

$$\Delta\mathcal{T}(\mathbf{x}' \pm 2\mathbf{j}) = \frac{1}{12}\phi(\mathbf{x}' \pm 2\mathbf{j})\Delta\phi(\mathbf{x}'). \tag{84}$$

Collecting these terms into a single "Effective $\Delta\mathcal{T}$" results in

$$\Delta\mathcal{T}'(\mathbf{x}') = -\frac{15}{2}\Delta\phi(\mathbf{x}')^2 + \Delta\phi(\mathbf{x}')\sum_i \big[\frac{8}{3}(\phi(\mathbf{x}+i)+\phi(\mathbf{x}-i)) - \frac{1}{6}(\phi(\mathbf{x}+2i)+\phi(\mathbf{x}-2i))\big], \quad (85)$$

and

$$\Delta S(\mathbf{x}') = \mathcal{T}'(\mathbf{x}') + \mathcal{V}(\theta(\mathbf{x}')) - \mathcal{V}(\phi(\mathbf{x}')). \quad (86)$$

This expression achieves the goal of expressing ΔS only in terms of lattice values at $\mathbf{x}'$ and its next-to-nearest neighbourhood, and is easily generalized to two fields. This is what allows sampling steps to be performed in parallel, as long as they deal with proposed changes to lattice points a sufficient distance apart. Figure 11 shows the dependency zone of a sampling step. As long as other sampling steps only depend on the light grey areas, in which the lattice sites are only read from instead of written to, they can be performed in parallel. The full set of points that are independent in this way is simply all parallel diagonal planes in the lattice with a separation of two lattice sites in between them. This is exactly what the simulation kernel does, shifting the planes by one lattice site per cycle, achieving a full sweep in just three calls to the parallel sampling kernel.

References

[1] Peter Brockway Arnold and Laurence G. Yaffe. Epsilon expansion analysis of very weak first order transitions in the cubic anisotropy model: Part 1. *Phys. Rev.*, D55:7760–7775, 1997.

[2] E. Ward Cheney and David R. Kincaid. *Numerical Mathematics and Computing*. Brooks/Cole Publishing Co., Pacific Grove, CA, USA, 2nd edition, 1985.

[3] Andrew G. Cohen, D. B. Kaplan, and A. E. Nelson. Progress in electroweak baryogenesis. *Ann. Rev. Nucl. Part. Sci.*, 43:27–70, 1993.

[4] Michael Creutz. Quarks, gluons, and lattices. 01 1983.

[5] Michela D'Onofrio. *Topics of particle physics in the very early universe*. PhD thesis, Helsinki U., 2014.

[6] Martin Dove. *Introduction to Lattice Dynamics*. 11 1993.

[7] K. Kajantie, M. Laine, K. Rummukainen, and M. Shaposhnikov. Generic rules for high temperature dimensional reduction and their application to the standard model. 1995.

[8] Faqir C Khanna, Adolfo P C Malbouisson, Jorge M C Malbouisson, and Ademir E Santana. *Thermal Quantum Field Theory : Algebraic Aspects and Applications*. WORLD SCIENTIFIC, 2009.

[9] Ulrike Kraemmer and Anton Rebhan. Advances in perturbative thermal field theory. *Rept. Prog. Phys.*, 67:351, 2004.

[10] M. Laine and A. Rajantie. Lattice continuum relations for 3-D SU(N) + Higgs theories. *Nucl. Phys.*, B513:471–489, 1998.

[11] M. Laine and K. Rummukainen. The MSSM electroweak phase transition on the lattice. *Nucl. Phys.*, B535:423–457, 1998.

[12] Mikko Laine and Aleksi Vuorinen. Basics of Thermal Field Theory. *Lect. Notes Phys.*, 925:pp.1–281, 2016.

[13] G. Peter Lepage and Paul B. Mackenzie. Viability of lattice perturbation theory. *Phys. Rev. D*, 48:2250–2264, Sep 1993.

[14] Jun Liu. *Monte Carlo Strategies in Scientic Computing*. 02 2009.

[15] F. Mandl. *Statistical Physics*. Manchester physics series. Wiley, 1971.

[16] Guy D. Moore. Curing O(a) errors in 3-D lattice SU(2) x (1) Higgs theory. *Nucl. Phys.*, B493:439–474, 1997.

[17] Guy D. Moore, Kari Rummukainen, and Anders Tranberg. Nonperturbative computation of the bubble nucleation rate in the cubic anisotropy model, 2001.

[18] David E. Morrissey and Michael J. Ramsey-Musolf. Electroweak baryogenesis. *New J. Phys.*, 14:125003, 2012.

[19] NVIDIA. Cuda c programming guide, 2018.

[20] Kari Rummukainen. Monte carlo simulation methods, lecture notes. `https://www.mv.helsinki.fi/home/rummukai/lectures/montecarlo_oulu/`.

[21] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.