FYS3150 Computational Physics - Project 3
The Thermodynamics of The Ising Model

Nils Johannes Mikkelsen
(Dated: November 18, 2018)

hello

All material written for project 4 may be found at:
https://github.com/njmikkelsen/comphys2018/tree/master/Project4

I. INTRODUCTION

II. THEORY

A. Thermodynamics

The following theory is based on the FYS 3150 lectures on statistical physics [5], in addition to the online version of Harvey Gould and Jan Tobochnik's *Thermal and Statistical Physics* [3] (chapter 5 in particular).

1. Fundamentals

The theory of thermodynamics is based on the statistical notion that the macrosopic properties of a thermodynamic system is fundamentally rooted in the configurations of microsopic properties. More formally, assuming that a system may be decomposed into a strict set of degrees of freedom, a unique configuration of these degrees constitutes what is known as a *microstate*. The fundamental assumption of thermodynamics, and statistical physics in general, is that the probability of finding the system in any of its available microstates is uniform. The generalised properties of a unique microstate, say the number of up-spins in a string of spin-1/2 electrons, is known as the system's *macrostate*. Several microstates may share a common macrostate, thus leading to a statistical distribution in the system's macrostates. It therefore follows that macroscopic properties such as temperature, pressure, etc., stems from the underlying distributions of micro- and macrostates.

A fundamental property of a themodynamic system is the total number of available states Ω, whose numerical value is often ridiculously large. It's importance relates to the fundamental assumption of the uniform propbability distribution between microstates: $P_i = 1/\Omega$ (here i denotes an arbitrary microstate). This expression leads to another, arguebly more important, fundamental quantity known as entropy:

$$S = -k_B \sum_i P_i \log P_i = k_B \log \Omega \tag{1}$$

where k_B is the Boltzmann constant and the second equation applies $P_i = 1/\Omega$. The importance of entropy is most visible in its relation to the famous *Second Law of Thermodynamics* (2LT), which states that the entropy of an isolated system tends to increase:

$$\mathrm{d}S = \frac{\mathrm{d}Q}{T} \geq 0 \tag{2}$$

Here, $\mathrm{d}S$ is the infinitesimal increase in S due to an infinitesimal exchange of heat Q between a system an it's surroundings at temperature T.

2. Some selected thermodynamic properties

This project will only consider the so-called *canonical ensemble*. In this context, an ensemble, or a *statistical ensemble*, is a large collection of ideal and identical microsystems that exist is some form of statistical equilibrium. The canonical ensemble is a particular ensemble in which the system is in thermal equilibrium with its surroundings. It can be shown that such systems behave according to the Boltzmann distribution, which is a probability distribution governing the probability of finding the system with a specific energy ϵ, provided temperature T:

$$P(\epsilon) = \frac{1}{Z} e^{-\epsilon/k_B T} \tag{3}$$

Here, k_B is the Boltzmann constant and Z is the so-called partition function:

$$Z = \sum_i e^{-\epsilon_i/k_B T} \tag{4}$$

A common practice is to introduce the substitution $\beta = 1/k_B T$, simplifying both analytics and computations. One of the properties of the canonical ensemble is its drive to minimise the Helmholtz free energy:

$$F = U - TS \tag{5}$$

where $U = \langle \epsilon \rangle$ is the system's internal energy. The Helmoholtz free energy describes the eternal conflict between entropy's tendency to increase and the principle of energy minimisation.

While thermodynamics deserves a more in-depth treatment, this would only be superfluous in this report. The final parts of this section will therefore introduce dome thermodynamic properties without a strict derivation.

The first and second moments of ϵ are given by:

$$\langle \epsilon \rangle = \sum_i \epsilon_i P_i = \frac{1}{Z} \sum_i \epsilon_i e^{-\beta \epsilon_i} \tag{6a}$$

$$\langle \epsilon^2 \rangle = \sum_i \epsilon_i^2 P_i = \frac{1}{Z} \sum_i \epsilon_i^2 e^{-\beta \epsilon_i} \tag{6b}$$

such that the variance of ϵ is given by $\mathrm{Var}[\epsilon] = \langle \epsilon^2 \rangle - \langle \epsilon \rangle^2$. The energy-variance is particularly important as it is proportional to the system's heat capacity at constant volume:

$$C_V = \frac{\mathrm{Var}[\epsilon]}{k_B T^2} = \frac{1}{k_B T^2} \left(\langle \epsilon^2 \rangle - \langle \epsilon \rangle^2 \right) \tag{7}$$

Furthermore, consider a system composed of spin-1/2 particles that is subjected to an external magnetic field $\mathbf{B} = B\hat{\mathbf{z}}$ such that the energy-interaction between a particle and the field is

$$E_B = -\boldsymbol{\mu} \cdot \mathbf{B} = -\mu_z B, \quad B > 0 \tag{8}$$

where $\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z)$ is the particle's magnetic moment. To simplify notation, introduce: $\mu_z = s\mu$ where $s = \pm 1$ indicates spin-up or spin-down. The net magnetisation of the complete system is then

$$\mathcal{M} = \mu M = \mu \sum_i s_i \tag{9}$$

where M is the net spin. The first and second moments of M are given by:

$$\langle M \rangle = \sum_i M_i P_i = \frac{1}{Z} \sum_i M_i e^{-\beta \epsilon_i} \tag{10a}$$

$$\langle M^2 \rangle = \sum_i M_i^2 P_i = \frac{1}{Z} \sum_i M_i^2 e^{-\beta \epsilon_i} \tag{10b}$$

such that the variance of M is given by $\mathrm{Var}[M] = \langle M^2 \rangle - \langle M \rangle^2$. Much like how heat capacity is proportional to the energy-variance, the magnetic susceptibility χ is proportional to the variance of the net-spin M:

$$\chi = \frac{\mathrm{Var}[M]}{k_B T} = \frac{1}{k_B T} \left(\langle M^2 \rangle - \langle M \rangle^2 \right) \tag{11}$$

B. The Ising Model

1. General aspects

The Ising model is a mathematical description that attempts to explain ferromagnetism as a result of the spin-spin interactions between particles. The model asserts the (spin-1/2) particles in a magnetic material are arranged in a square lattice structure and attributed a binary spin $s = \pm 1$, where s follows the notation introduced in the previous section. The key component of the Ising model is that each spin is only allowed to interact with its neighbouring spin. Furthermore, the interaction between two spins k and l is given by $-Js_k s_l$ for all combinations of k and l. Here, $J > 0$ is a coupling constant which represents the interaction strength. The general model also allows for the presence of an external magnetic field. Seeing that the system is completely defined by the arrangement of the spins, the arrangement constitutes the microstate of the Ising model. The macrostate is uniquely determined by the system's energy level and net spin (M in equation (9)).

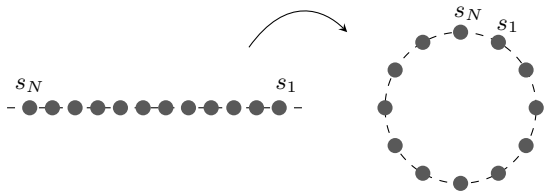


Figure 1: A visualisation of the physical interpretation of one-dimensional period boundary conditions: the single dimension is folded into a two-dimensional closed path. The bondary point s_N thus experiences being the neighbour of the complementary boundary point s_1.

For realistic systems, the model contains a number of particles on the scale of Avogadro's number ($\propto 10^{23}$), which clearly is infeasible even for the most sophisticated computers. In order to avoid considerable boundary effects, the mathematics employ period boundary conditions. For the one-dimensional case, period boundary conditions may be physically interpreted as the folding of a straight line into a closed path (see figure 1 for visuals). The two-dimensional equivalent is to fold a rectangular sheet into a torus, for example as seen in this animation [1].

By combining both the spin interactions and the external magnetisation, the energy of the system is given by

$$\epsilon = -J \sum_{\langle kl \rangle} s_k s_l - \mu B \sum_i s_i \tag{12}$$

where $\langle kl \rangle$ implies a summation over neighbouring spins. Note that (12) does not impose any particular coordinate system, in fact, the Ising model may be applied to any n-dimensional case. This project will focus on the two-dimensional case, which happens to have been solved exactly (higher-dimensional cases have not).

2. The 2×2 system

Assuming the system consists of four particles arranged in a square 2×2 lattice, there are $2^4 = 16$ number of possible arrangements of the spins s. All possible microstates are sketched out in appendix A, the macrostates are summarised in table I.

Energy $[J]$	Net Spin	Degeneracy
-8	4	1
0	2	4
0	0	4
8	0	2
0	-2	4
-8	-4	1

Table I: The macrostates of the 2×2 Ising model system: energy & net spin. The degeneracy refers to the number of microstates with the same macrostate.

Table I provides a simple opportunity to evaluate various thermodynamic properties effectively. The key observation is to identify that the summation over all iterations of a variable A_i with a degeneracy $\deg(A)$ is the same as the summation over all the values of the variable A scaled with its corresponding degeneracy:

$$\begin{aligned}\sum_i A_i &= \left(A^1 + \cdots + A^1\right) + \cdots + \left(A^n + \cdots + A^n\right) \\ &= \deg(A^1)A^1 + \cdots + \deg(A^n)A^n \\ &= \sum_A \deg(A)A\end{aligned}$$

where $A \in \{A^1, \ldots, A^n\}$. The rest of this section deals with to exact evaluation of the thermodynamic properties introduced in section II A 2 using this summation trick. The quantites are given a superscript 2×2 to indicate that they are only valid for the 2×2 Ising model system.

The partition function is equal to

$$\begin{aligned}Z^{2\times2} &= \sum_\epsilon \deg(\epsilon)e^{-\epsilon\beta} \\ &= e^{8J\beta} + 4e^0 + 4e^0 + 2e^{-8J\beta} + 4e^0 + e^{8J\beta} \\ &= 12 + 4\cosh(8J\beta) \end{aligned} \tag{13}$$

The first and second moments of the energy is therefore given by

$$\begin{aligned}Z\left\langle\epsilon\right\rangle^{2\times2} &= \sum_\epsilon \deg(\epsilon)\epsilon e^{-\epsilon\beta} \\ &= (-8J)e^{8J\beta} + 4(0)e^0 + 4(0)e^0 \\ &\quad + 2(8J)e^{-8J\beta} + 4(0)e^0 + (-8J)e^{8J\beta} \\ \left\langle\epsilon\right\rangle^{2\times2} &= -\frac{32}{Z}J\sinh(8J\beta)\end{aligned} \tag{14}$$

$$\begin{aligned}Z\left\langle\epsilon^2\right\rangle^{2\times2} &= \sum_\epsilon \deg(\epsilon)\epsilon^2 e^{-\epsilon\beta} \\ &= (-8J)^2e^{8J\beta} + 4(0)^2e^0 + 4(0)^2e^0 \\ &\quad + 2(8J)^2e^{-8J\beta} + 4(0)^2e^0 + (-8J)^2e^{8J\beta} \\ \left\langle\epsilon^2\right\rangle^{2\times2} &= \frac{128}{Z}J^2\cosh(8J\beta)\end{aligned} \tag{15}$$

which implies that the variance of the energy is

$$\begin{aligned}\mathrm{Var}[\epsilon] &= \frac{128}{Z}J^2\cosh(8J\beta) - \Big(-\frac{32}{Z}J\sinh(8J\beta)\Big)^2 \\ &= \frac{128J^2}{Z^2}\Big(Z\cosh(8J\beta) - 8\sinh^2(8J\beta)\Big)\end{aligned} \tag{16}$$

Similarly, the first and second moments of the net spin is given by

$$\begin{aligned}Z\left\langle M\right\rangle^{2\times2} &= \sum_\epsilon \deg(\epsilon)M(\epsilon)e^{-\epsilon\beta} \\ &= (4)e^{8J\beta} + 4(2)e^0 + 4(0)e^0 \\ &\quad + 2(0)e^{-8J\beta} + 4(-2)e^0 + (-4)e^{8J\beta} \\ \left\langle M\right\rangle^{2\times2} &= 0\end{aligned} \tag{17}$$

$$\begin{aligned}Z\left\langle M^2\right\rangle^{2\times2} &= \sum_\epsilon \deg(M)M(\epsilon)^2e^{-\epsilon\beta} \\ &= (4)^2e^{8J\beta} + 4(2)^2e^0 + 4(0)^2e^0 \\ &\quad + 2(0)^2e^{-8J\beta} + 4(-2)^2e^0 + (-4)^2e^{8J\beta} \\ \left\langle M^2\right\rangle^{2\times2} &= \frac{32}{Z}\left(1 + e^{8J\epsilon}\right)\end{aligned} \tag{18}$$

which implies that the variance of the net spin is

$$\mathrm{Var}[M] = \frac{32}{Z}\left(1 + e^{8J\epsilon}\right) \tag{19}$$

The heat capacity $C_V^{2\times2}$ and the magnetic susceptibility $\chi^{2\times2}$ are found using equations (7) and (11) respectively (which are essentially a rescaling of the energy and net spin variances in equations (16) and (19)). Plots of $C_V^{2\times2}$ and $\chi^{2\times2}$ are included in appendix A.

C. Numerical Simulations

The following theory is based partially on the FYS 3150 lecture notes on Monte Carlo methods [4] and statistical physics [5], and notes on Monte Carlo integration by Dave Edwards [2].

1. *Monte Carlo integration*

In the field of numerical integration, one of the most common strategies is so-called Monte Carlo integration. Monte Carlo integration cleverly exploits the properties of stochastic variables and probability theory by rewriting the problem in terms of expected values. The basic problem is to integrate a scalar integral on the form

$$I = \int_D \mathrm{d}x\, f(x) \tag{20}$$

where $x \in D$ is not necessarily one-dimensional. Let $p(x)$ denote a particular known distribution and consider the

following change of variables:

$$\int_D \mathrm{d}x\, f(x) = \int_{x\in D} \mathrm{d}p(x)\, \frac{f(x)}{p(x)} \tag{21}$$

That is, the integral I is equal to the expected value of $f(x)/p(x)$ with respect to the probability distribution $p(x)$. The idea is to draw N_{MC} random numbers $X \sim p(x)$ and estimate the expected value with the sample expected value:

$$\left\langle \frac{f(x)}{p(x)} \right\rangle_X \approx \sum_{i=1}^{N_{\mathrm{MC}}} \frac{f(x_i)}{p(x_i)} p(x_i) = \sum_{i=1}^{N_{\mathrm{MC}}} f(x_i) \tag{22}$$

which becomes an equality when $N_{\mathrm{MC}} \to \infty$ due to the Law of Large Numbers.

Although (22) is exact in the large-N_{MC} limit, the choice of $p(x)$ may affect the lower boundary of N_{MC} for which the sample expected value is fairly accurate. Hence, choosing $p(x)$ is an important component of the design of a Monte Carlo integration. The simplest choice of $p(x)$ is the uniform distribution, which sadly is particularly unstable for $f(x)$ functions with sharp peaks.

Effective algorithms that a large number of computations of $p(x)$ have been developed, one of which is presented below. Nonetheless, the basic foundation for any Monte Carlo integration techinque is summarised by the following algorithm.

Algorithm 1 Standard Monte Carlo Integration

1: Define N_{MC}.
2: Initialise SUM $= 0$.
3: **for** $i = 1, \ldots, N_{\mathrm{MC}}$:
4: Draw x_i at random according to $p(x)$.
5: Evaluate $f(x_i)$.
6: Add contribution: SUM += $f(x_i)$
7: **end for**
8: Normalise final value: $I = \mathrm{SUM}/N_{\mathrm{MC}}$.

2. *Markov chains*

A Markov chain is a stochastic model for the discrete evolution [1] of a so-called "memoryless system". In this context, to be memoryless implies that each step is only dependent on the previous state of the system (i.e., the system does not conserve memory of earlier states). A common usage of Markov chains is to mimick time evolution, but this is just an example.

In its most general form, a (discrete) Markov chain is a sequence of random variables $X_1, X_2, \ldots$, which satisfy the Markov property. Simplified, the Markov property requires that the probability of moving from a state X_n to state X_{n+1} is only dependent on X_n. The state space may either be continuous or discrete, each of which invoke a distinct, albeit similar, formalism. Because the Boltzmann distribution is a continuous distribution, the discrete formalism will be ignored in favour of the continuous.

Say the state space for a particular Markov chain is $\mathcal{S}$, then a *stochastic kernel* on $\mathcal{S}$ is a function $p : \mathcal{S}\times\mathcal{S} \to \mathbb{R}$ that satisfies

$$p(x,y) \le 0,\ \forall x,y \in \mathcal{S} \quad \text{and} \quad \int_{\mathcal{S}} \mathrm{d}y\, p(x,y) = 1,\ \forall x \in \mathcal{S}$$

Each Markov chain is parametrised by its corresponding stochastic kernel, which in this sense is usually referred to as the transition probability density. Now for every $p(x,y)$ there exists a so-called *Markov operator* P that is such that

$$\big[(\cdot)P\big](y) = \int_{\mathcal{S}} \mathrm{d}x\, p(x,y)(\cdot)$$

Note in particular that P is a left-acting operator, as per usual in the literature.

Furthermore, say the Markov chain has continued for n steps with previous measurements $X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n$, then the next measurement $X_{n+1} = y$ is drawn according to the distribution $p^{(n+1)}(x)$ whose evolution from the current distribution $p^{(n)}(x)$ is given by the Markov operator of the Markov chain:

$$p^{(n+1)}(y) = \big[p^{(n)}(x)P\big](y) = \int_{\mathcal{S}} \mathrm{d}x\, p(x,y) p^{(n)}(x) \tag{23}$$

A special case of (23) is when the distribution is invariant of P, that is, when $p^{(n+1)} = p^{(n)}$. This behaviour is called a stationary distribution and is usually denoted by $\pi(x)$:[2]

$$\pi(y) = \int_{\mathcal{S}} \mathrm{d}x\, p(x,y)\pi(x) \tag{24}$$

Such a state is guaranteed if the Markov chain obeys the so-called *detailed balance* condition. Provided a transition probability density $p(x,y)$ and a particular distribution $\pi(x)$, then the detailed balance condition states that the Markov chain corresponding to $p(x,y)$ is *reversible* with respect to $\pi(x)$ if

$$\pi(x)p(x,y) = \pi(y)p(y,x),\ \forall x,y \in \mathcal{S} \tag{25}$$

It follows that

$$\int_{\mathcal{S}} \mathrm{d}x\, p(x,y)\pi(x) = \pi(y) \int \mathrm{d}x\, p(y,x) = \pi(y)$$

[1] A Markov chain may be generalised to continuous evolution, but this will not be used in this project.

[2] Becuase choosing π to represent something else other than π seemed like a great idea, obviously.

because $p(x, y)$ is normalised for all $x, y \in \mathcal{S}$. In conclusion, in case a Markov chain is "well-behaved", meaning it satisfies detailed balance, the so-called *limiting distribution* of the Markov chain approaches π as $n \to \infty$:

$$\lim_{n\to\infty} \left[p^{(0)}(x)P^n\right](y) = \pi(y) \tag{26}$$

where P^n implies repeated operations on $p^{(0)}$, $p^{(1)}$, etc.

3. Markov Chain Monte Carlo: The Metropolis algorithm

As the name implies, Markov Chain Monte Carlo (MCMC) methods combine the concept of a Markov chain with Monte Carlo methods. There exists several MCMC methods, each with unique properties that may or may not be benefitial. This project will focus on the so-called *Metropolis* algorithm, which is a special case of the *Metropolis-Hastings* algorithm.

The basic idea behind the MCMC methods stems from equation (22). Ideally one would be able to draw a sufficiently large number of states x from $p(x)$ and follow algorithm 1 religiously. However, this is incredibly computation-ineffective and essentially infeasible for a standard compute. Enter Markov chains: A Markov chain with limiting distribution $\pi(x) = p(x)$ would necessarily generate states that are approximately distributed according to $p(x)$. It turns out that this may be exploited in order to avoid large computations.

The mathematics of MCMC methods are based on the detailed balance principle, consider the following rearranging of equation (25):

$$\frac{p(x, y)}{p(y, x)} = \frac{\pi(y)}{\pi(x)}$$

Now, suppose the current state of the Markov chain governed by $p(x, y)$ is X_n, i.e., the n^{th} state in the chain. The idea is to separate the transition process into two steps: an initial *proposal*, and an *acceptance/rejection* step. These steps should be independent so that probability is independent, meaning the probability density of proposing a state y, $g(y|x)$, is independent of the probability of accepting the proposal, $\alpha(x, y)$. Because they are independent, it follows that

$$p(x, y) = g(y|x)\alpha(x, y) \tag{27}$$

meaning the above criteria may be rewritten as

$$\frac{\alpha(y, x)}{\alpha(x, y)} = \frac{p(y)}{p(x)}\frac{g(x|y)}{g(y|x)} \tag{28}$$

where $p(x) = \pi(x)$ is the assertion of MCMC methods that the limiting distribution of the Markov chain governed by $p(x, y)$ is the integrand of equation (20). This expression is particularly easy to implement in case the normalisation of p and g are independent on x and y, or if they happen to cancel each other (although this is very unlikely).

The last step it to choose an acceptance probability $\alpha(y, x)$ that satisfies 28, the *Metropolis choice* is

$$\alpha(y, x) = \min\left\{1, \frac{p(y)}{p(x)}\frac{g(x|y)}{g(y|x)}\right\} \tag{29}$$

The Metropolis-Hastings algorithm may now be stated:

Algorithm 2 The Metropolis-Hastings Algorithm

```
Initialise the first state X_0.
for n = 1, ..., N_MH:
    Generate a proposal y according to g(y|x).
    Evaluate acceptance probability α(y, x).
    if α(y, x) = 1:
        Accept y.
    else
        Draw a uniformly distributed number a ∈ [0, 1].
        if a ≤ α(y, x):
            Accept y.
        else
            Reject y and set X_{n+1} = X_n.
        end if
    end if
end for
```

where N_{MH} is the number of steps in the Markov chain. Note that the algorithm above does not factor in the Monte Carlo update, this is because the central purpose of the Metropolis-Hastings algorithm is namely just to sample x values according to $p(x)$. Nonethelesss, implementations of the algorithm usually include an intermediate step inside this loop in order to avoid multiple loops.

Furthermore, although the distribution is guaranteed to converge, it does not necessarily converge within the first few steps. Accordingly, a common practice is to perform N_{prep} preparation updates to X_0 in a process known as "burn-in", before invoking the Monte Carlo steps.

As previously mentioned, this project will implement the Metropolis algorithm, which is a special case of the Metropolis-Hastings algorithm. The Metropolis assumes that g is symmetric in x and y, meaning $g(x|y) = g(y|x)$. This implies that the acceptance probability may be simplified to

$$\alpha_{\text{M}}(y, x) = \min\left\{1, \frac{p(y)}{p(x)}\right\} \tag{30}$$

III. METHOD

The main focus of this project is to study the two-dimensional Ising model, where the overaching goal is to estimate the critical temperature for the magnetic phase transition. The system is assumed to be under no influence from external magnetic fields such that the phase transition is completely dependent on unperturbed system. The time-evolution of the system is modelled using

a Markov chain, where each step represents a small increment in time. In order to study the critical temperature, the mean magnetisation, heat capacity and magnetic susceptibility will be studied as functions of temperature. The resulting integrals will be estimated using Monte Carlo simulations with Metropolis-sampled states.

Before continuing to the numerical experiments in this project, the following section will describe some general aspects of the simulation.

A. The Simulation - General Aspects

The Ising system is parametrised by temperature via $\beta = 1/k_B T$, which is assumed to be constant during the simulation. To goal is to estimate a series of expectance values as functions of temperature, meaning various integrals must be estimated using Monte Carlo methods. The energy of the ideal Ising model (infinitely large lattice) is continuous with respect to temperature, this will be approximated using periodic boundary conditions such that the continuous formalism developed above is applicable, in particular the Metropolis algorithm.

1. *The basic approach - Flip one spin*

The (micro-)state of the Ising system is the spin configuration of the lattice structure. Becuase the Ising system is canonical, the energy must necessarily behave according to the Boltzmann distribution (equation (3)). Now, assuming g is symmetric in x and y, its follows that the Metropolis acceptance probability is equal to

$$\begin{aligned}\alpha(y,x) &= \min\left\{1, \frac{\frac{1}{Z}e^{-\epsilon(y)\beta}}{\frac{1}{Z}e^{-\epsilon(x)\beta}}\right\} \\ &= \min\left\{1, e^{[\epsilon(x)-\epsilon(y)]\beta}\right\} \end{aligned} \tag{31}$$

However, the exponential may further simplified using a clever choice of g. The approach is sometimes called the *"flip one"* tactic. The idea is to assume that g is only nonzero for the states that are different from x by a single spin-permutation (i.e. a single spin flip). At its face, this imposes no significance on the exponential in $\alpha(y,x)$. However, observe that with a single spin flip, the only difference between x and y are the interactions between the flipped spin and its neighbours (see figure 2). This implies that to evaluate the change in energy and net magnetisation between x and y, only the flipped interactions have to be considered. The possible arrangements of the flipped spin and its neighbours are shown in appendix A.

Without a external magnetic field ($B = 0$), the energy difference is

$$\epsilon(y) - \epsilon(x) = -J\sum_{\langle kl\rangle} s_k(y)s_l(y) + J\sum_{\langle kl\rangle} s_k(x)s_l(x)$$

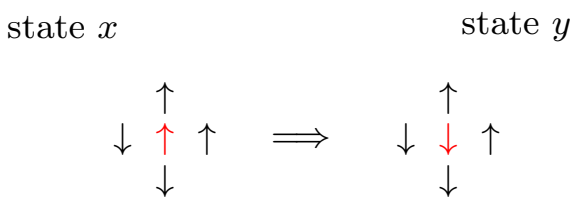


Figure 2: An example transition from state x to state y using the flip one tactic. The only difference is the flipped red spin, the rest of the spin configuration remains the same.

where $s_k(x)$ denotes the spin at label k in state x. Each of these sums may be divided into two components: one containing the flipped spin interactions and the rest:

$$\sum_{\langle kl\rangle} s_k(\cdot)s_l(\cdot) = \sum_{\langle ki\rangle} s_k(\cdot)s_i(\cdot) + \sum_{\langle kl\rangle\,\wedge\, l\neq i} s_k(\cdot)s_l(\cdot)$$

where i denotes the flipped spin. The two "rest" sums cancel and, as $s_k(x) = s_k(y) = s_k$, one finds that:

$$\epsilon(y) - \epsilon(x) = -J\sum_{\langle ki\rangle} s_k\big(s_i(y) - s_i(x)\big)$$

Furthermore, the flipped spin necessarily satisfies $s_i(y) = -s_i(x)$, meaning $s_i(y) - s_i(x) = -2s_i(x)$. In conclusion,

$$\epsilon(y) - \epsilon(x) = \Delta\epsilon = 2Js_i(x)\sum_{\langle k(i)\rangle} s_k \tag{32}$$

where $\langle k(i)\rangle$ denotes summation over the neighbours k of i, but not including i. Because there is a limited number of possible flip-one arrangements, there is thus only a limited number of possible values for $\epsilon(y) - \epsilon(x)$. This is important because it allows for the computation of the exponential in (31) outside the Metropolis loop. More on this below.

To obtain the change in net magnetisation, the same approach is followed: The total sums are divided into a flipped sum and a rest sum. However, because the net magnetisation does not loop over neighbours, the flipped sum reduces to the single flipped spin:

$$M(y) - M(x) = \sum_j s_j(y) - \sum_j s_j(x) = s_i(y) - s_i(x)$$

where j loops over all spins, while i denotes the flipped spin. Again, as $s_i(x) = -s_i(y)$ it follows that

$$M(y) - M(x) = \Delta M = 2s_i(y) \tag{33}$$

2. *The algorithm*

The algorithm is basically a special case of the combination of algorithms 1 and 2. Nonetheless, some minor details (such as evaluating the exponential outside

the Monte Carlo simulation) have been implemented to spare computation. Note that later modifications will be introduced where necessary, but the general algorithm is:

Algorithm 3 The Ising Model Metropolis Monte Carlo

Define N_{MC}, T and L.
Initialise Monte Carlo integrals.
Precompute all possible $\exp(-\Delta\epsilon\beta)$.
Initialise $L \times L$ spin state matrix.
Initialise state energy and net magnetisation.
for $n = 1, \ldots, N_{\mathrm{MC}}$:
    for $k = 1, \ldots, L$:
        for $l = 1, \ldots, L$:
            Generate random index $i \in \{1, \ldots, L\}$.
            Generate random index $j \in \{1, \ldots, L\}$.
            Propose new state y with spin s_{ij} flipped.
            Evaluate net magnetisation of (i, j) neighbours.
            Infer $\alpha = \exp\big(-\Delta\epsilon\beta\big)$.
            Draw $a \sim \mathcal{U}(0, 1)$.
            if $a \leq \alpha$:
                Accept y: Flip spin (i, j) in state matrix.
                Update state energy and net magnetisation.
            end if
        end for
    end for
    Add contribution to Monte Carlo integrals.
end for
Normalise Monte Carlo integrals according to N_{MC} and $L \times L$.

There are several components of algorithm 3 that are implemented for computational reasons, these are explained below. Observe the triple `for`-loop in 3: this is actually a double loop in disguise. The outer loop is the standard Monte Carlo integration loop from algorithm 1. The two inner loops (k and l) are actually the single Metropolis loop in algorithm 2. The idea it to loop over every spin in the lattice in order to produce an uncorrelated new state. However, flipping every spin in the lattice would just return the inverse (negative) state, thus each spin is chosen at random.

Furthermore, note that the Metropolis-Hastings algorithm (2) contains a two-step `if`-test. In particular, the algorithm computes first $\alpha = \exp(-\Delta\epsilon\beta)$ and then evaluates $\alpha(y, x)$. In case α is greater than 1, the move is directly accepted. On the other hand, if α is the smallar of the two, the move is accepted with probability α by comparing it with a random number $a \sim \mathcal{U}(0, 1)$.

B. Numerical Experiments

The following section describes the specific problems that are studied in this project. Additional details are introduced where relevant.

1. *Verifying the computations*

To begin the project, the Ising simulation will be compared to the known 2×2 system with a temperature of $T = 1$. The thermodynomic properties of interest to this experiment are: $\langle\epsilon\rangle$, $\langle M\rangle$, C_V, χ and $\langle|M|\rangle$, whose analytic solutions are described in section II B 2. The main goal of this experiment is to make a crude analysis of the Monte Carlo integrals' dependency on the number of Monte Carlo cycles N_{MC} when the Metropolis algorithm is given zero burn-in time (meaning all cycles contribute towards the integrals). The experiment will be run 40 unique times for each of the following number of Monte Carlo cycles: 10^2, 10^3, 10^4, 10^5, 10^6 and 10^7.

2. *Studying the burn-in*

The next experiment will focus on the Monte Carlo burn-in with special interest on the number of sufficient cycles and the importance of the initial state. The experiment features a 20×20 lattice with either of two initial states: the ground state (ordered) and a random state (disordered). In the ground state (lowest energy state), all spins point upwards (see equation (12)), whereas in the random state, all spins are equally distributed between down and up spins.

The jist of the experiment is to perform the simulation as before, however track the contributions to the Monte Carlo integrals through the simulation. The resulting arrays, which contain the commulative sums of the integration, are then normalised according to the number of cycles that has passed. The experiment will use an upper boundary of $N_{\mathrm{MC}} = 10^6$ for all runs. Moreover, the experiment will also be run with 3 different temperatures: $T_1 = 1$, $T_2 = 1.7$ and $T_3 = 2.4$.

To summarise, the experiment is run with $N_{\mathrm{MC}} = 10^6$ for 3 different temperatures from 2 dfferent initial states, making a total 6 different runs. In addition to the Monte Carlo integrals, this experiment will also track how the number of accepted proposals evolves with increasing number of Monte Carlo cycles.

3. *The energy-distribution: A frequentist approach*

Having found a sufficient burn-in equilibriation time, the next step is to study the established equilibrium as it continues to evolve. To do so, the focus falls on the energy-distribution of the spin-states produced by the Metroplis algorithm (after converging), which will be determined using a frequentist approach: namely measuring the frequency of each energy as the system evolves.

Similar to the previous experiment, this experiment will also use a lattice with features 20×20 and be run with temperatures T_1, T_2 and T_3. The system is initialised from the ground state and given a burn-in equal

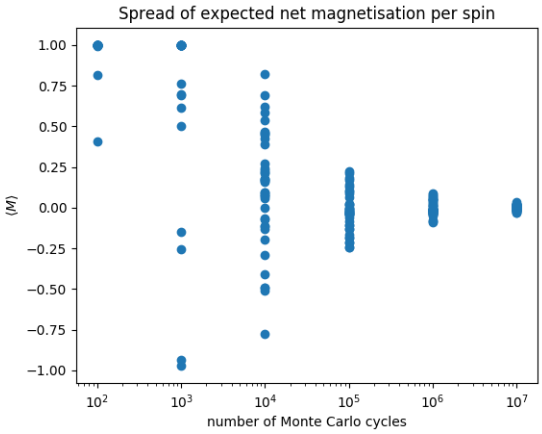


Figure 3: Estimates of the expected net magnetisation for the 2×2 Ising system with $T = 1$. The analytic value is 0, which the graph shows a clear tendency towards.

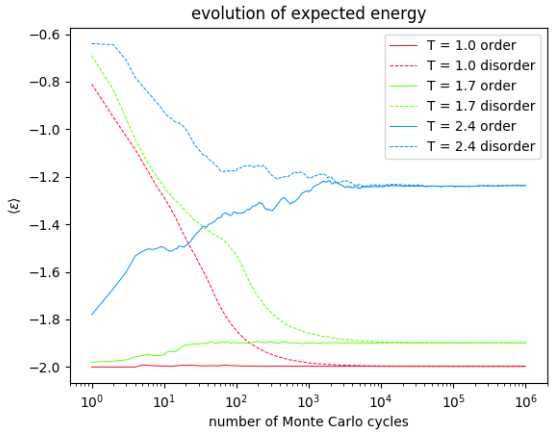


Figure 4: Energy

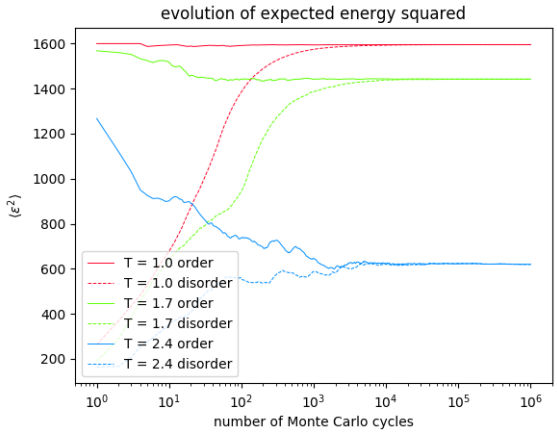


Figure 5: Energy

to the sufficient burn-in found in the previous experiment. Starting from the equilibrium state, the system is simulated over 10^6 steps in order to provide a decent sample size. Particular interest is placed on how the spread in the energy (during evolution) behaves with respect to the temperature of the system.

The Boltzmann distribution is a gaussian with respect to energy, meaning the resulting distributions should ideally be bell-curves.

4. Studying the phase transition

IV. RESULTS

A. Experiment 1

The experiments showed a clear tendency for inaccurate estimates when the number of Monte Carlo cycles dropped below $\approx 10^6$. Figure 3 shows the experiment-estimates for the expected net magnetisation as a function of the number of Monte Carlo cycles. The mean estimate clearly approach 0 (the analytical solution) with increasing N_{MC}, which is an indication of the algorithm successfully simulating the Ising model. This is further supported by the decreasing spread of the estimates. The integral estimates for $N_{\text{MC}} \geq 10^4$ are summarised in table V using the mean estimates. Overall, the experiment has shown that the simulation provides reliable results for $N_{\text{MC}} \gtrapprox 10^7$.

B. Experiment 2

The results from the experiment are shown in figures 4-9. Interestingly, figure 9 reveals that the rate of the number of accepted proposals converges on 1 for all temperatures, which is indicative of a state of equilibrium. Furthermore, the figure also exposes the temperature dependence of the equilibration time (number of sufficient Monte Carlo cycles for reaching equilibrium), in particular, systems with greater temperature converge faster than systems with less temperature. All systems show clear equilibrium behaviour with a burn-in of 10^5 Monte Carlo cycles.

C. Experiment 3

The evolution of the energy of the equilibrium systems are shown in figure 10, the corresponding energy frequency-distributions are shown in figure 11. A striking problem with this experiment is the strong cut-off of the $T = 1$ energy spectrum. A likely reason for this cut-off is the limited dimensionality of the Ising lattice. This is supported by the observation that the $T = 1.7$ and $T = 2.4$ spectra are much more gaussian in their ap-

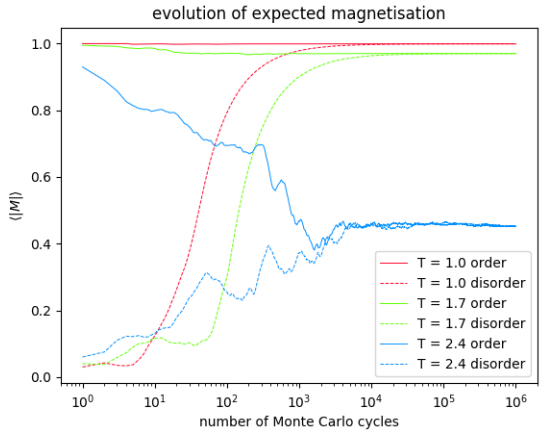


Figure 6: Energy

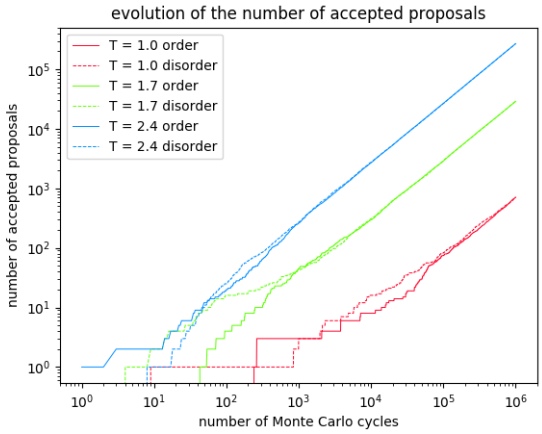


Figure 9: Energy

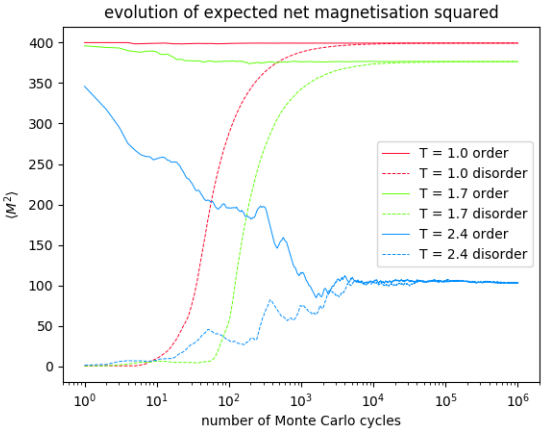


Figure 7: Energy

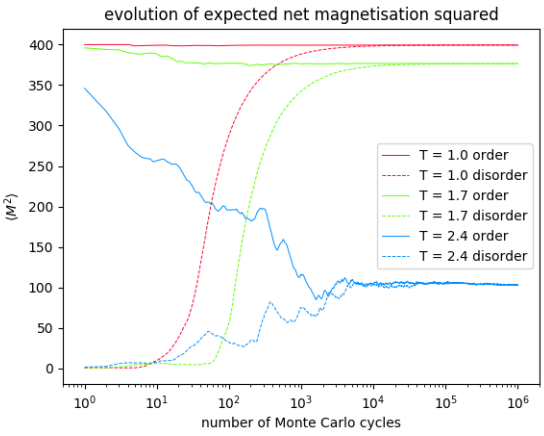


Figure 8: Energy

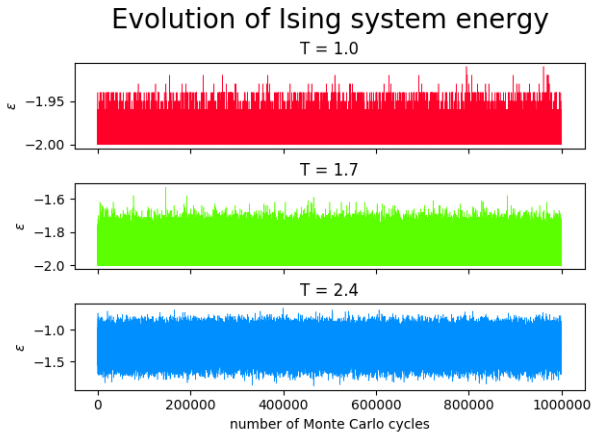


Figure 10: Energy

pearance. Moreover, note that the $T = 1.7$ spectrum is slightly cut-off from below in addition to displaying some discrepancies in comparison to standard bell-curve (thes splitting of the central peak in particular). The $T = 2.4$ spectrum however, is centered about a sufficiently large energy so that the distribution is much more gaussian-like.

The spread (standard deviation) in the energy evolution are summarised in table II. The pattern is clear: increasing the temperature increases the spread in energy. This is reassuring because it aligns with the expected behavior the relationship between temperature and energy from our understanding of thermodynamics. Increasing the temperature leads to increased probability of the system occupying a certain energy, meaning a greater number of large energy-leaps between steps is more likely to occur. However, it is important that because of the cut-off seen in figure 10, the results in table II are subjected to a slight regularization of the true spread. Nonetheless, the trend is still a valid observation.

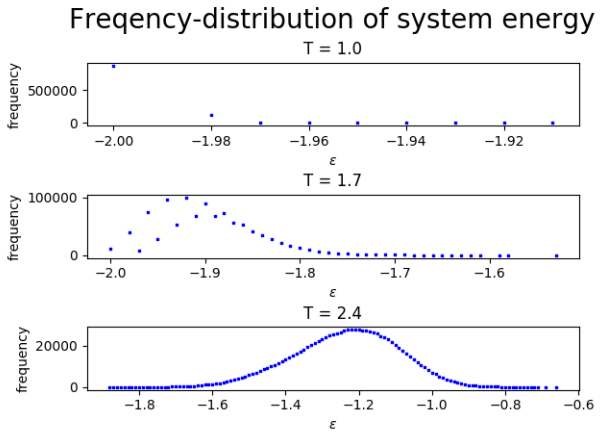


Figure 11: Energy

T	σ_ϵ
1.0	$7.9148 \cdot 10^{-3}$
1.7	$4.9684 \cdot 10^{-2}$
2.4	$1.4220 \cdot 10^{-1}$

Table II: Spread in energy during evolution from equilibrium of the 20×20 system.

V. DISCUSSION

VI. CONCLUSION

[1] Animation of a sheet being folded into a torus. https://mathematica.stackexchange.com/questions/42493/morphing-a-sheet-of-paper-into-a-torus.

[2] Dave Edwards. Monte carlo integration. http://www.cs.utah.edu/~edwards/research/mcIntegration.pdf.

[3] Harvey Gould and Jan Tobochnik. *Thermal and Statistical Physics*. 2011. publisher: Princeton Unversity Press, online version: http://stp.clarku.edu/notes/.

[4] Morten Hjorth-Jensen. Fys 3150 lecture notes: Introduction to monte carlo methods. http://compphysics.github.io/ComputationalPhysics/doc/pub/mcint/html/mcint.html.

[5] Morten Hjorth-Jensen. Fys 3150 lecture notes: Statistical physics. http://compphysics.github.io/ComputationalPhysics/doc/pub/statphys/html/statphys.html.

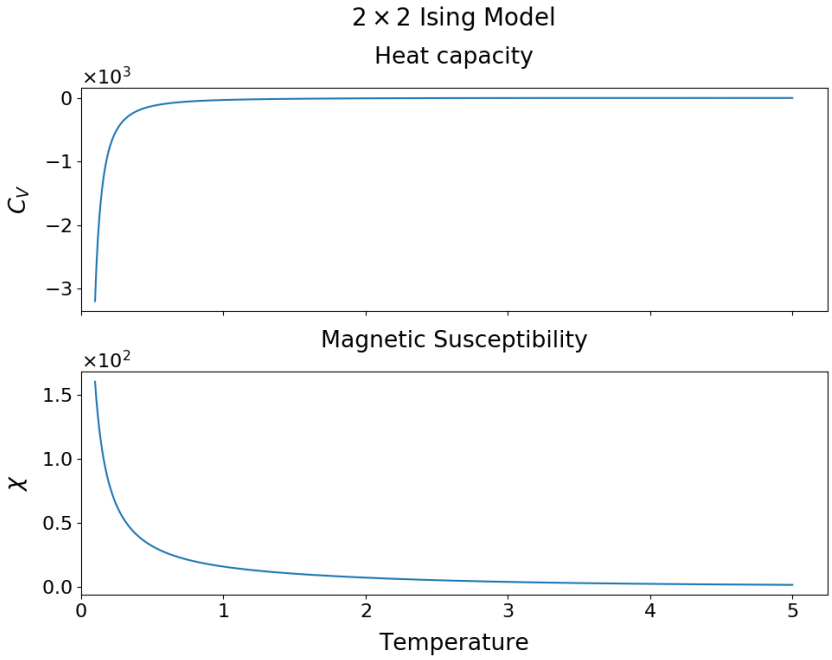
Appendix A: Miscellaneous Material

1. All Possible Spin Configurations of The 2×2 Ising Model

Table III: A display of all possible arrangements of the 2×2 system. Each spin is allowed to occupy a spin-up state ($\uparrow$) or a spin-down state ($\downarrow$).

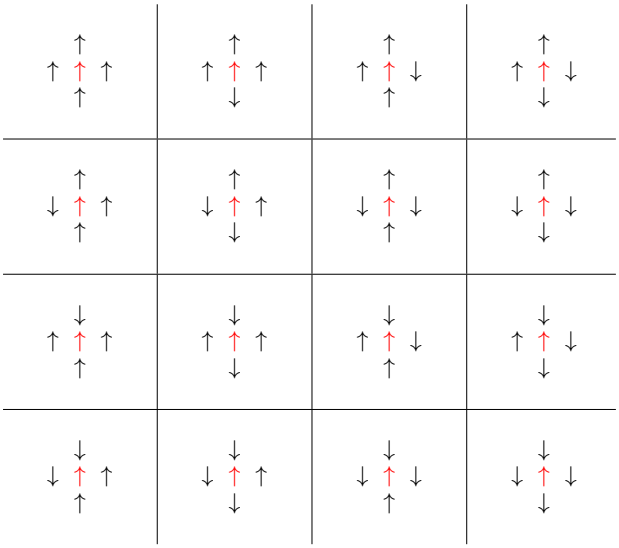
$\downarrow\downarrow$ $\downarrow\downarrow$	$\downarrow\downarrow$ $\downarrow\uparrow$	$\downarrow\downarrow$ $\uparrow\downarrow$	$\downarrow\downarrow$ $\uparrow\uparrow$
$\downarrow\uparrow$ $\downarrow\downarrow$	$\downarrow\uparrow$ $\downarrow\uparrow$	$\downarrow\uparrow$ $\uparrow\downarrow$	$\downarrow\uparrow$ $\uparrow\uparrow$
$\uparrow\downarrow$ $\downarrow\downarrow$	$\uparrow\downarrow$ $\downarrow\uparrow$	$\uparrow\downarrow$ $\uparrow\downarrow$	$\uparrow\downarrow$ $\uparrow\uparrow$
$\uparrow\uparrow$ $\downarrow\downarrow$	$\uparrow\uparrow$ $\downarrow\uparrow$	$\uparrow\uparrow$ $\uparrow\downarrow$	$\uparrow\uparrow$ $\uparrow\uparrow$

2. Heat Capacity and Magnetic Susceptibility of The 2×2 Ising Model



3. All Possible Spin Configurations In The Flip One Tactic

Table IV: A display of half of the possible arrangements of the spins involved in the flip one tactic. Each spin is allowed to occupy a spin-up state (↑) or a spin-down state (↓). The central red spin is the spin to be flipped, the other half of the possible arrangements are the resulting flipped states.



Appendix B: Additional Results

1. Numerical Results From Experiment 1

Estimates	10^4 cycles	10^5 cycles	10^6 cycles	10^7 cycles
$\langle\epsilon\rangle$	$-1.9961 \pm 0.072\%$	$-1.9959 \pm 0.026\%$	$-1.9960 \pm 0.009\%$	$-1.9960 \pm 0.003\%$
$\langle\epsilon^2\rangle$ $[10^1]$	$1.5969 \pm 0.072\%$	$1.5967 \pm 0.026\%$	$1.5968 \pm 0.009\%$	$1.5968 \pm 0.003\%$
$\langle\lvert M\rvert\rangle$ $[10^{-1}]$	$9.9871 \pm 0.048\%$	$9.9863 \pm 0.019\%$	$9.9867 \pm 0.006\%$	$9.9866 \pm 0.002\%$
$\langle M\rangle$ $[10^{-4}]$	$1178.0 \pm 297.5\%$	$-235.7 \pm 516.0\%$	$-19.0 \pm 2268.9\%$	$9.4 \pm 1654.9\%$
$\langle M^2\rangle$	$3.9935 \pm 0.060\%$	$3.9931 \pm 0.022\%$	$3.9933 \pm 0.007\%$	$3.9933 \pm 0.002\%$
algorithm run time [s]	$0.031 \pm 2.873\%$	$0.305 \pm 3.306\%$	$3.067 \pm 3.110\%$	30.812 ± 2.799

Table V: The average estimates of the various thermodynamic properties for the 2×2 Ising system. The averages are evaluated from 40 unique runs without burn-in. The confidence intervals are given by the standard deviation of the estimates relative to the averages. Note that the large error boundaries of $\langle M\rangle$ is actually a poor representation of the quality of the estimate, because the analytical value is 0. Hence, any significant deviation from 0 would necessarily bring a large spread in average estimate, regardless of the accuracy of the estimate.