Bachelor Thesis

Non-Flat Wang-Landau Algorithm

Alex Panchot

3727651

Universität Leipzig

Faculty of Physics and Earth Sciences
Institute of Theoretical Physics
Computational Quantum Field Theory (CQT)

Examiners:
Prof. Dr. Wolfhard Janke
Dr. Stefan Schnabel

April 24, 2021

Table of Contents

List of Figures

List of Tables

Section 1

Introduction

Before beginning to understand what the words "Non-Flat Wang-Landau" mean, one must begin by understanding the system or "model" in which this algorithm is designed to work with.

1.1 The Model

The Ising model was first developed by Wilhelm Lenz in 1920 and further investigated by Ernst Ising, the namesake, in his 1925 doctoral dissertation [1].

The basic idea of the model is to simply represent magnetism of a simple cubic lattice. Ferromagnetism arises when the spins, of each of the atoms in a domain, point in the same direction. The uniformity of the spins allow the microscopic magnetic moments to form a macroscopic moment. For the interests of this paper, the effects of pure ferromagnetism will not be examined, but rather of the region between complete ferro and antiferromagnetism, ferrimagnetism. Additionally, only the changes inside the sample rather than on its environment as well as ideal samples will be explored.

For the sample, each atom is represented by a value for the spin. The spin values, which represent the direction of the magnetic moment, can either point "up" or "down" representing a value of up or down spin respectively. The up spin is represented with σ_{+1} and the down spin with σ_{-1}. With a selectively placed lattice of atoms, macroscopic measurements can be made.

The volume of the lattice is defined as V, with $V = L^d$. L is the side length of the lattice and d is the dimension of the system. So for example, an 8×8 2D system has a volume of: $8^2 = 64 = V$.

1.1.1 Ising Model

For the Ising model, as described above, each spin can take a value of ± 1. Using these values for the spin, the magnetization and energy can be computed. Additional parameters such as specific heat and susceptibility can also be computed, but will not be examined in this paper.

The Hamiltonian of the atoms can be calculated with:

$$\mathcal{H} = -J \sum_{<i,j>} \sigma_i \sigma_j + \mu H \sum_i \sigma_i \tag{1.1}$$

Assuming a canonical ensemble for the lattice, the Hamiltonian is equivalent to the energy of the system. In addition, the above general formula can be simplified. The simplified version of the Hamiltonian/energy are assumed for this thesis. The value of J or the "exchange energy" is set at 1. Furthermore, the external magnetic field "H" is set at 0, such that the resulting Hamiltonian, and henceforth energy, looks like:

$$\mathcal{H} = E = -\sum_{<i,j>} \sigma_i \sigma_j \tag{1.2}$$

The energy of the system is a sum over each spin and its interaction with its nearest neighbors. For this thesis, a 2 dimensional case is treated. For the 2D case, each spin has 4 neighbors, assuming a square lattice.

While calculating the energy, the first issue appears. At the edge or corner of the lattice, these spins do not have 4 neighbors. To solve the issue of a finite lattice, the use of "periodic boundary conditions" will be used. Periodic boundary conditions are simply to set $\sigma_{n+1} = \sigma_0$, with n being the length of one side of the lattice. The use of these boundary conditions allow for a better representation of an infinite sized lattice.

1.1.2 Potts Model

Looking beyond the Ising Model, the values of each spin can be generalized beyond simply ± 1. Now, let each value of the spin exist between 1 and some value q. These values may be given a direction, which is referred to as the clock model, but for the purposes of this paper, each spin will not be given a direction. Therefore, the calculation of macroscopic magnetic properties are no longer possible. Additionally, the Hamiltonian of the system changes accordingly (equation 1.3).

$$E = -\sum_{<ij>} \delta_{\sigma_i, \sigma_j} \tag{1.3}$$

Each neighbor interaction can now only take values of 0 or 1 because of the Kronecker delta ($\delta_{\sigma_i, \sigma_j}$). This changes the total energy such that it can no longer be positive (as the sum cannot be negative).

1.1.3 Partition Function

After calculating the energy of *one state* of the spins (equation 1.3), the canonical partition function can be calculated. The partition function is a sum over *all possible states* of the spins (equation 1.4). For this and all following examples, the Potts Model is used.

$$Z = \sum_{\{\sigma\}} e^{-\beta \mathcal{H}(\{\sigma\})} = \sum_{\{\sigma\}} e^{-\beta E(\{\sigma\})} = \sum_{\{\sigma\}} e^{\beta \sum_{<jk>} \delta_{\sigma_j, \sigma_k}} \tag{1.4}$$

For this partition function, the value β_c is referred to at the inverse temperature, such that $\beta = \frac{1}{k_b T}$. Here the k_b is the Boltzmann Constant, which for simplicity's sake is set at 1 (thus $\beta = \frac{1}{T}$). Here T is the temperature of the system. In addition to the partition function, the expected energy:

$$\langle E \rangle = \frac{\sum_{\{\sigma\}} E_{\{\sigma\}} e^{-\beta E_{\{\sigma\}}}}{\sum_{\{\sigma\}} e^{-\beta E_{\{\sigma\}}}} = \frac{\sum_{\{\sigma\}} E_{\{\sigma\}} e^{-\beta E_{\{\sigma\}}}}{Z} \tag{1.5}$$

specific heat:

$$C_v = \frac{\left(\langle E^2 \rangle - \langle E \rangle^2\right)}{T^2} \tag{1.6}$$

$$\langle E^k \rangle = \frac{\sum_{\{\sigma\}} E_{\{\sigma\}}^k e^{-\beta E_{\{\sigma\}}}}{Z}, \quad k = 1, 2, 3, ...$$

and other variables can be calculated. For this particular Hamiltonian, the magnetization is not clearly defined as the spins do not represent a defined direction. A statement can be made on the most common value, but nothing can really be said concerning the "average" magnetization or susceptibility.

1.2 Algorithms

There exist various algorithms designed to solve these problems. Here, two will be discussed.

The Metropolis algorithm, which calculates variables, such as expected energy, at specific temperatures. The benefit of the metropolis algorithm is that it can estimate the expected energy much faster than computing using the true formula.

Additionally, the Wang-Landau algorithm can be used to calculate the so-called density of states. Knowing the density of states can calculate the partition function and expected energy exactly, at the expense of taking longer to calculate than the Metropolis.

1.2.1 Metropolis

The Metropolis algorithm was first developed in 1953 [5] and is a very efficient method to determine data from a probability distribution where truly random sampling is not possible. The Metropolis algorithm for the Ising Model is as follows:

1. Set up the lattice with each spin taking either random spins (hot-start) or the same spin value (cold-start).
2. Pick a random spin on the lattice, a "lattice site" and perform the following steps.
   (a) Calculate the local energy of the system, i.e calculate the energy between the lattice site and each of its nearest neighbors. For 2D, the lattice site has 4 nearest neighbors.
   (b) Flip the lattice site's spin value and calculate the new local energy.
   (c) Using the probability $\left(p = \min\left(1, e^{-\beta \Delta E}\right) \right)$ calculate if the lattice site should keep its new spin value. If a random number $[0, 1)$ is less than p, then keep the new spin value, else set the lattice site's spin to the spin in step (a).
3. Repeat step 2. for V attempts, each with a new randomly selected lattice site. A "sweep" is what completing V attempts is called.
4. After completing the sweep, the current energy and magnetization can be recorded and saved for later. These values will be used later to calculate the average energy, specific heat or magnetization. Of course, the values of energy and magnetization can be measured after each step 2., but it is more typical to measure after each sweep.
5. Repeat steps 2-4 for as many steps as "needed". There is no specific value for how many are needed as it depends on the level of accuracy needed and the V of the system.

6. Now the average energy, etc. can be calculated from each of the collected energies, but these values are only for the temperature β that the simulation was run at.

1.2.2 Density of States

Looking at the partition function, the summation of a summation means that each possible configuration of the lattice must be seen such that each possible energy is calculated. For a q-state lattice with a side length of l, the total number of configurations is q^{l^2}. For a rather small lattice of $q = 10$ and $l = 32$, the total number is 10^{1024}. This means that the exact computation of the partition function, even with the fastest computers available is not feasible.

Despite having an extremely large amount of possibilities for the energy of the system, the energy can only exist between 0 and $2 \times l \times l$. Because there are only 2048 unique energy values (for $l = 32$), many orders of magnitude of these configurations are degenerate. Calculating the number of occurrences of these degenerate values can drastically shorten the time needed to calculate the partition function or other macroscopic variables such as the expected value for the energy.

The sum in the partition function (as well as the expected energy) can use the density of states, written as $g(E)$.

$$Z = \sum_i g(E_i)\, e^{-\beta E_i}$$

$$\langle E \rangle = \frac{\sum_i E_i\, g(E_i)\, e^{-\beta E_i}}{Z}$$

Here the density of states reduces the number of calculations for the partition function to $2 \times l \times l$ or just the number of available energy states. More importantly, the density of states is independent of the temperature. This allows for the calculation of the expected energy at *any* temperature after the calculation of the density of states.

1.2.3 Wang-Landau

The Wang-Landau algorithm was first described in 2001 [4]. It aims to compute the density of states rather than to calculate for example the average energy for a specific temperature. In order to do this, it uses a changing probability that is based on the current density of states. This allows for "less-popular" energies to be visited more quickly.

The density of states ($g(E)$) are not computed directly, but rather the inverse of the density of states is calculated. Let $w(E) = \frac{1}{g(E)}$ be called the weight function. In addition to this change, each $w(E)$ is stored as the natural log in the computer as storing the full value would exceed the maximum size of a double precision floating point number ($\approx 10^{308}$). So $w(E) \to \ln w(E)$.

The Wang-Landau algorithm is as follows [2]:

1. Set up the lattice with each spin taking either random spins (so-called hot-start) or the same spin value (so-called cold-start).
2. Set each weight function to 1 such that $\ln w(E) = \ln 1 = 0$

3. Pick a random spin on the lattice, a "lattice site" and perform the following steps.

   (a) Using the probability $\left(p = \min\left(1, \frac{w(E_2)}{w(E_1)}\right) \right)$ calculate if the lattice site should flip its spin value. $w(E_1)$ is the weight function of the current energy value and $w(E_2)$ is the weight function of the energy value should the lattice site flip its spin. Of course as the weight functions are stored as the natural log, the probability is written as:

   $$p = \min\left(1, e^{\ln w(E_2) - \ln w(E_1)}\right)$$

   (b) Regardless of whether the lattice site's spin has changed on not, the weight function of the final energy changes by a so-called modification factor f. Again f is stored in the natural log so $f \to \ln f$. For the normal weight function the modification factor is division, but for the natural log it changes to be subtraction.

   $$\ln w_{new}(E) = \ln w_{old}(E) - \ln f$$

   A natural, but not necessarily best, choice for f is e, such that $\ln f = \ln e = 1$

   (c) A energy histogram is also kept and records each time the lattice is at at specific energy.

4. Repeat step 3. for V attempts

5. Repeat steps 3-4 for 10000 sweeps. Every 10000 sweeps is an arbitrary choice that appears to be frequent but not too frequent.

6. After every 10000 sweeps, calculate the so-called "flatness". Flatness can be defined in two different ways:

   (a) The definition of flatness as described by Wang and Landau is that smallest value of the histogram $H(E)$ is at least $x\%$ of the average value of the histogram $\langle H(E)\rangle$. $x\%$ is chosen based on the size and complexity of the lattice. For the Ising Model, the value of $x\%$ can be as high as 95%. As soon at the smallest value is $\geq \langle H(E)\rangle$, the lattice is flat. For example, look at the following histogram ?? (for a 8×8, $q = 2$ system):

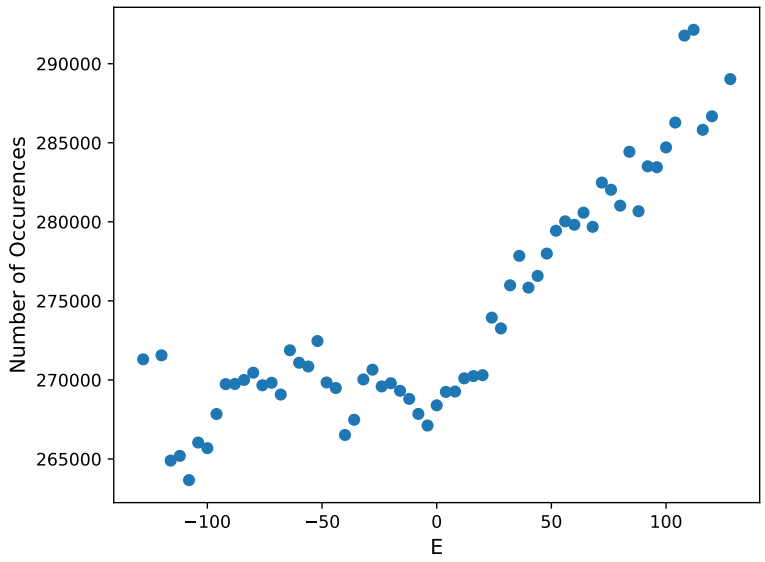

Figure 1.1: Example Histogram

   In the above histogram, assume the smallest value is 265000 and the average value is 275000. If $x\%$ is set to 50% then this histogram is considered flat as $\frac{265000}{275000} \approx 0.963 > 0.5$

   (b) An alternative definition of flatness is one where the lattice achieves flatness when each of the sites have been visited at least once. However, as this reduces the amount of sweeps required to achieve flatness, it also reduces the accuracy of the final density of states. Therefore, one may choose to change the reduction factor in order to counteract this.

7. Once flatness has been achieved, the histogram $H(E)$ is reset for all energy values. The modification factor f is also reduced by the square-root ($\frac{1}{2}$ in the natural log):

$$f_{new} = \sqrt{f_{old}} = (f_{old})^{\frac{1}{2}} \rightarrow \ln f_{new} = \frac{1}{2} \ln f_{old}$$

In figure 1.2, a sample Wang-Landau algorithm is shown. After each time flatness is achieved, the $\ln f$ drops by $\frac{1}{2}$ as described.

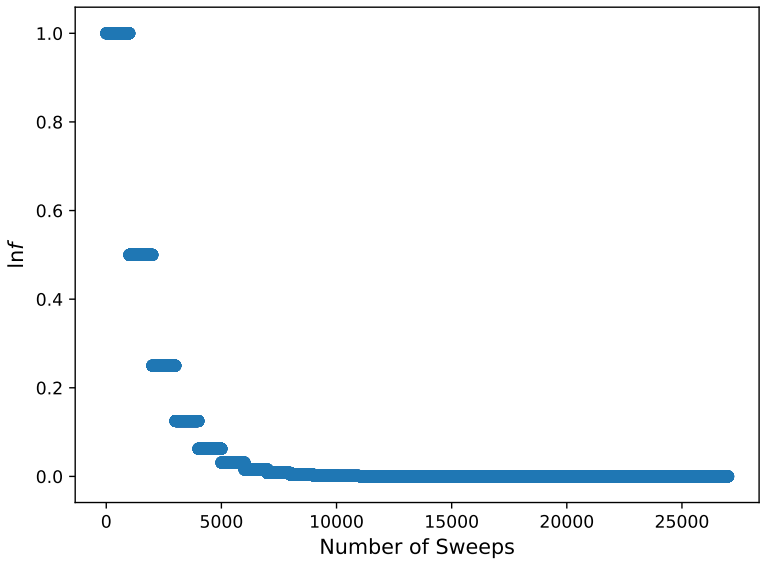

Figure 1.2: Modification Factor Example

8. Repeat steps 3-7 until $\ln f < 10^{-8}$. This takes 27 iterations.

9. After completing all 27 iterations, the density of states is calculated by simply inverting the weight function or changing the sign on the $\ln w(E)$

The calculation of the density of states is very fast, but leads to a biased error in the final results. Below is an additional step that can be taken in order to reduce these errors. All of the steps described up until this point are said to be in the 1st phase. The steps in the next sections are called the 2nd phase.

1.2.4 2nd Phase of the Wang-Landau

After calculating the final weight functions, $\ln w(E)$, the simulation is run again with as described by steps 3-5. *Important:* steps 1-2 are *not* performed as this 2nd phase should immediately follow the 1st phase.

Step 6-7 is slightly different. Here two histograms are kept rather than just one like in the 1st phase. One histogram is used to check for flatness as described in step 6. The other histogram is never reset. Both have their data created as described in step 3(c). After flatness is achieved step 7 is *not* performed. So the weight functions stay the same.

This 2nd phase is run for a certain number of iterations, with more resulting in a better final result.

After the 2nd phase is finished, the density of states can be calculated by:

$$g(E) \propto \frac{H_{phase2}(E)}{w_{phase1}(E)} = \gamma \, \frac{H_{phase2}(E)}{w_{phase1}(E)}$$

The $H_{phase2}(E)$ is the histogram which has not been reset over the entire 2nd phase. To get the real density of states, some normalization constant γ must be multiplied with the fraction. In the logarithms it looks like:

$$\ln g(E) = \ln H_{phase2}(E) - \ln w_{phase1}(E) + \ln \gamma \tag{1.7}$$

1.2.5 Normalization

After finishing with the 2nd phase, the density of states must be normalized. In equation 1.7, the logarithm of the histogram will continue to grow as the number of iterations grows longer. Thus there needs to be a corrective constant (γ) to correct this growth.

There are several different methods of normalization, but the one used in this paper will be described here.

1. Let the calculated density of states be: $\rightarrow \ln \tilde{g}(E)$
2. Now subtract the maximal value of $\ln \tilde{g}(E)$ from the calculated density of states: $\ln g_0(E) = \ln \tilde{g}(E) - \max(\ln \tilde{g}(E))$
   - The reason for subtracting the maximal value from each of the density of states is that $e^{\ln \tilde{g}(E)} = \tilde{g}(E)$ and this is too large to store in a double precision variable. So by subtracting the maximal log value from each, the largest that $\ln g_0(E)$ can be is 0. Each of the other values will be smaller than 0 but not too much smaller.
3. Now let $N_0 = \sum_i g_0(E_i) = \sum_i e^{\ln g_0(E_i)}$
   - As all of the $\ln g_0(E)$ are < 0, this is now possible for the computer.
4. The normalized density of states is now given by:

$$g(E) = \frac{g_0(E)\, q^V}{N_0} \tag{1.8}$$

   where $V = L^d$ is the volume and q is the number of spin possibilities.
   - This formula is used because it normalizes the density of states with the total number of possible states q^V. If the density of states were perfect then $N_0 = q^V$.
5. The normalized logarithmic density of states is now given by:

$$\ln g(E) = \ln g_0(E) - \ln N_0 + V \ln q \tag{1.9}$$

Section 2

The Non-Flat Profile

After the introduction of the Wang-Landau algorithm, the next step is to consider how the algorithm may be improved. For instance, how can the algorithm be sped up such that it performs a simulation with equal or lower errors in less total simulation time? For smaller system sizes, the Wang-Landau algorithm is quick enough as is. But, when larger systems are examined, the computation time needed to finish scales up considerably. For example, a 64×64, $q = 10$ system takes around an hour for an average computer to just finish the 1st phase.

Some of the perhaps simplest and easiest variables that one can change in the algorithm are the reduction factor r with $\ln f_{i+1} = r \ln f_i$ (it is set at 2 in this paper), and the number of iteration of the 2nd phase. Of course, the trade off with speed is the reduction in the final accuracy. The idea of a profile is then perhaps a bit more complex in the hope that the final result's accuracy is not affected too much.

2.1 What is a Profile?

The basic idea of a profile is to force the simulation into visiting certain regions more often. As the system contains a transition energy, visiting the transition energy more frequently will naturally help the system change between the extreme energy values more often.

During the simulation, it is quite simple to implement the profile. As explained in the introduction, the weight function is appropriately adjusted by:

$$\ln w_{new}(E) = \ln w_{old}(E) - \ln f$$

To add the profile to this, the profile itself is simply added with $p(E)$ representing the profile.

$$\ln w_{new}(E) = \ln w_{old}(E) - \frac{\ln f}{p(E)}$$

Here, it is clear that a $p(E)$ that has a large value will suppress the change in $\ln w_{new}(E)$. This means that since the weight function changes with a lower amount each time for this energy, the system can visit this energy more frequently as the probability is less unfavorable as a result of the change. Because of this, the profile needs to make the system visit the transition energy region more often, while also not forcing the system to visit unimportant energy regions as well.

2.2 The Flat Profile

While computing the standard Wang-Landau algorithm, the profile used would simply be called flat. This is because it is 1 over the entire energy region. This is in figure 2.1 for a 32×32 system.

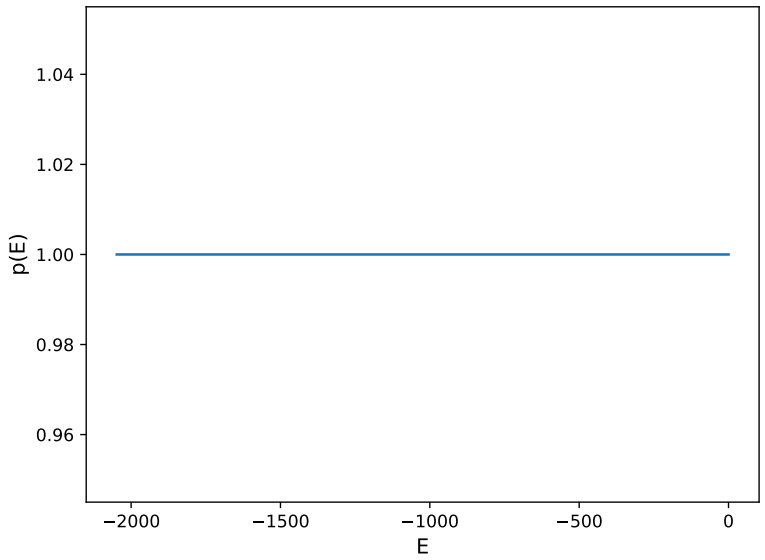

Figure 2.1: Flat Profile

As the flat profile is the default choice, it serves as the control for other profiles.

2.3 The Ideal Profile

Calculating the ideal profile would at first seem quite simple. If the logic described above is used, then it would seem easy to implement. Just increasing the profile for important regions is partly the solution, but through thorough testing, it is not the complete solution.

The important regions are at least the transition energy and the extreme energy regions, i.e the ground state and the zero energy state. The reason that these regions are considered to be important is that visiting these regions allow the system to make the energy transition more often. Simply being in that energy region more often allows for the system to cross the boundary and more quickly achieve "flatness".

Of course the big unknown here is where exactly the energy range is. The exact calculation of it by use of a timeseries is very not feasible. This is because any reduction in time by the use of the profile could be countered by the time taken to create the profile in the first place. Crucially, it is important to note that changing the lattice length or number of possible spin states may also change the transition energy.

Before looking at a calculated ideal profile, a simple profile can be created from the previous constraints. For a 32×32 system, with $q = 10$, a simple profile may look like this:

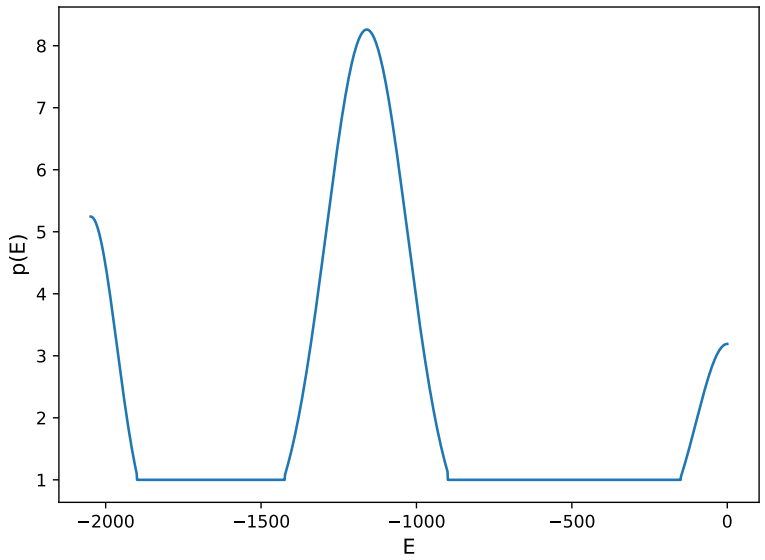

Figure 2.2: Simple Profile

This is one of many possible choices for a simple profile as this was created using Gaussian peaks. This means that the height and width of the peaks has to be carefully selected. In this profile, for instance, the extreme peaks are not at the same height. And because the region between the peaks is not considered too important, it is left at a value of 1.

And of course, the simple answer for the simple profile (figure 2.2) is that it is much slower than just the flat profile. Therefore, just throwing together a profile by purely guessing is counterproductive and a systematically calculated profile is the true answer.

Using another algorithm [3], an ideal profile can be created. However, as this algorithm [3] is somewhat complex, the algorithm will not be explained here. The results of the algorithm were also taken and used directly. Therefore the "ideal" profiles used here may not be perfect as they were created using another algorithm. The computation time of that algorithm is also not taken into consideration here, but it should not be forgotten as it part of the total time taken to calculate the final density of states.

From the algorithm [3], the ideal profile (figure 2.3) was made for a 32×32, $q = 10$ system:

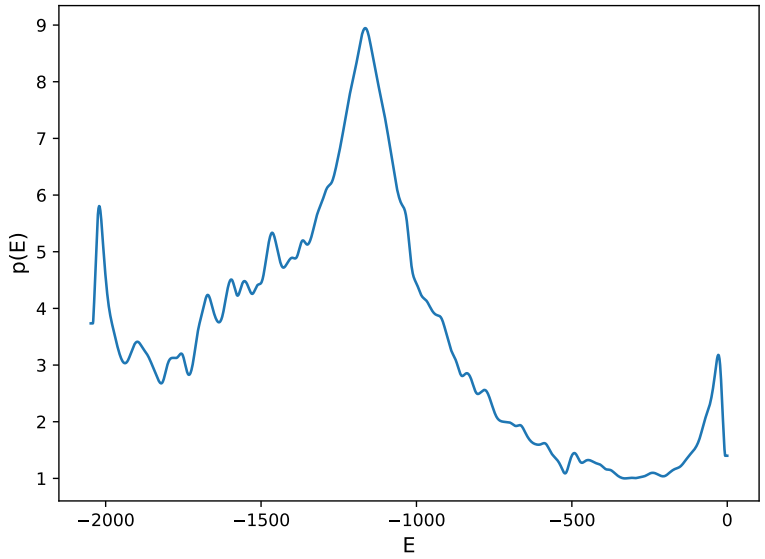

Figure 2.3: Ideal Profile, 32×32, $q = 10$ system

From the ideal profile (figure 2.3), it can easily be seen that this profile has a similar shape to the simple profile. It has a peak for the transition energy and a peak for both extreme energies. However, it also has several other peaks and the region between the transition in the center and the ground state energy is entirely larger than 1. The extreme energy peaks are also not centered at the most extreme energy values. This means that there is a transition region right before the extreme energy values.

As this profile (figure 2.3) is for a 32×32, $q = 10$ system, the profile for a different system will look different. For the 32×32, $q = 2$ and 64×64, $q = 10$ systems looked at later in the paper, the ideal profiles for these systems will be explained in their respective sections.

Section 3

$32 \times 32,\ q = 10$ System

In this section, a system's, of size 32×32 with a total number of spin possibilities of 10 ($q = 10$), density of states is calculated. The reason that this size was chosen is that it is neither the smallest nor largest system possible. Thus its density of states can be calculated rather quickly (on the order of a couple of minutes), but not so quickly that any benefit of the profile is not fully realized.

Using 10 possible spin values also allows for a first order phase transition to be seen [6] (see Appendix A). Later the $q = 2$ case will be examined as it is equivalent to the Ising model and it exhibits a second order phase transition.

Finally, both definitions of flatness will be examined here as each definition will be affected differently. For each definition of flatness, 250 different trials were run in order to get enough data for a good average. This means a unique weight function calculated after each unique 1st phase and unique density of states calculated from that weight function and a unique 2nd phase.

3.0.1 Profile

The ideal profile used in this section is the same as shown in figure 2.3.

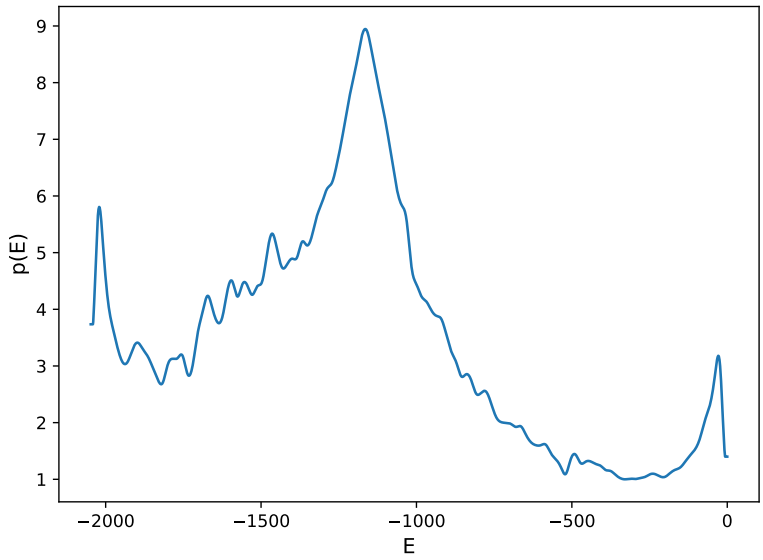

Figure 3.1: Ideal Profile for 32×32, $q = 10$ system

3.1 Alternative Definition of Flatness

First, the alternative definition of flatness is looked at. As the alternative definition simply asks that each energy value in the range of possible energy values be visited at least once, it is considerably faster than the Wang-Landau definition (which is looked at further in the paper).

3.1.1 Simulation Runtimes

First, the total runtimes of each of the 250 simulations can be looked at. The runtime is the number of sweeps that was recorded after finishing the 1st phase and after every 5 iterations of the 2nd phase up to 25. The graphs in figure 3.2 are histograms of the runtimes are shown in increasing order of iteration number. Here and henceforth "0" iterations means the number immediately after the completion of the 1st phase and without the completion of any 2nd phase.

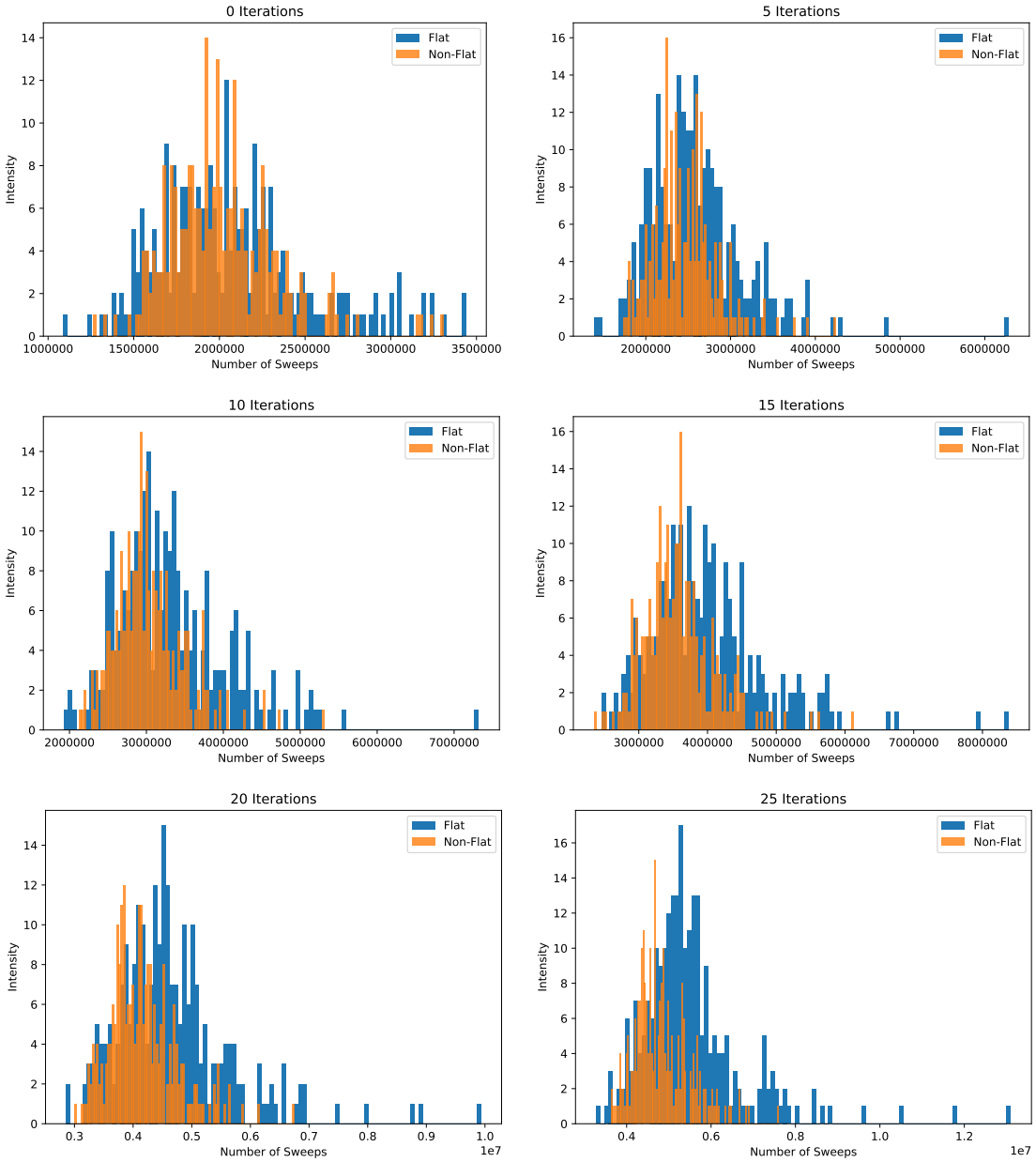

Figure 3.2: Runtime Comparisons between Flat and Non-Flat Profiles after every 5 iterations, 32×32, $q = 10$ system

From the first graph "0 Iterations" in figure 3.2, it can be seen that the variance (or width of the peaks) is quite high, but as additional iterations are performed the variance appears to go down as the peaks become thinner. In order better understand the results of the effects of the profiles, the mean runtime should be looked at. As some simulations will suffer from random error, the runtimes of those particular ones are of no direct interest. However, if the mean of many (in this case 250) simulations different for the profile, then the use of a profile is validated.

The mean of the simulations can be calculated using:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3.1}$$

$$\bar{t} = \frac{1}{250}\sum_{i=1}^{250} t_i$$

with t_i each runtime and $\bar{t}$ the average. Likewise, the standard error of the mean:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \tag{3.2}$$

$$\sigma_{\bar{t}} = \frac{\sigma}{\sqrt{250}}$$

with $\sigma_{\bar{t}}$ the error of the mean values of the runtimes and σ the standard deviation.

Flat Profile Runtimes

Using the data from figure 3.2 and the equations for the mean (equation 3.4) and error (equation 3.3), the exact data of the mean and standard error is calculated, for the flat profile, and given in table 3.1 with all values given in sweeps.

Table 3.1: Flat Profile Values, 32×32, $q = 10$ system

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	2 064 000	28 000
5	2 602 000	36 000
10	3 314 000	45 000
15	3 985 000	53 000
20	4 674 000	62 000
25	5 522 000	78 000

Again, "0 Iterations" signifies the number of sweeps after just the 1st phase. "5 Iterations" includes the results of the 1st phase and 5 2nd phase iterations, so the same results from "0 Iterations" plus 5 2nd phase iterations. "10 Iterations" includes the exact same results of "5 Iterations" plus 5 more iterations. This continues up to 25 iterations. Additionally, the values for the mean and error have been rounded to the nearest thousand as the error is large enough that the precision of the mean should not be too high.

From the exact results table 3.1, the error of the data (which is proportional to the standard deviation) is actually growing as the number of iterations increases. The shapes of the curves in the graphs of figure 3.2 appear thinner for more iterations because there are a few data points which lie far from the main peak. Looking into the raw data (the individual runtimes and not the mean values), it appears that the values that take the longest after 25 iterations also took longer after each of the other measurement points. In other words, the simulation was not slower for a single iteration but it was slower across the entire simulation, both for the 1st and 2nd phases.

Now more importantly, the results of the runtime as a result of using a profile are looked at. The same random number seeds were used and the same number of trials were performed (250) in order to make the most fair comparison.

Non-Flat Profile Runtimes

Similarly as for the flat profile, the mean and standard error for the non-flat profile is given in table 3.2. All of the data is laid out in the same manner as for the flat profile (table 3.1)

Table 3.2: Non-Flat Profile Values, 32×32, $q = 10$ system

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	2 012 000	20 000
5	2 461 000	24 000
10	3 034 000	29 000
15	3 587 000	33 000
20	4 151 000	36 000
25	4 828 000	42 000

Looking at the mean and standard error of the runtime of the non-flat profile (table 3.2), the same observations for the growing error can be made. Both the means and errors are lower for the non-flat profile however.

Comparison of Runtimes

After calculating the exact mean and errors of the runtimes for each of the profiles (tables 3.1 and 3.2), a plot can be made to better compare the data in the tables.

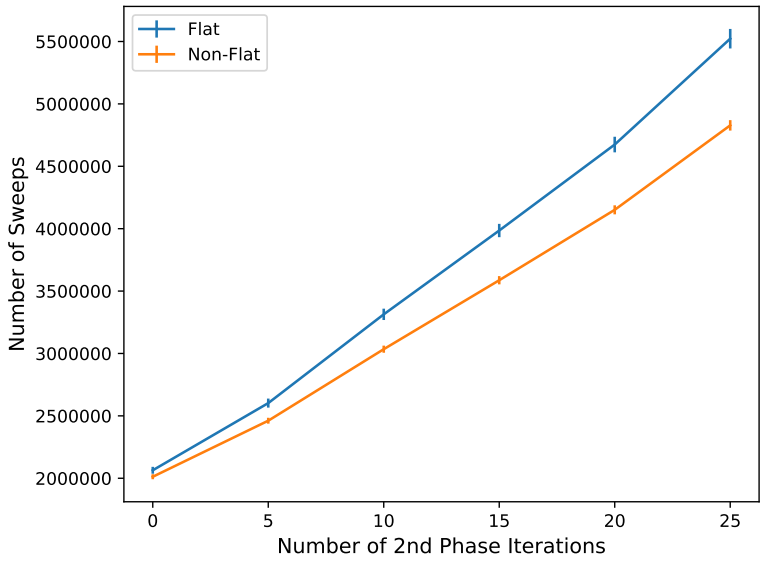

Figure 3.3: Comparison of Runtimes of Alternative Definition of Flatness, 32×32, $q = 10$ system

In figure 3.3, the total number of sweeps after is shown for each 5 iterations. Looking at 0 iterations, the non-flat profile needed fewer total sweeps. This means that the profile is better for the 1st phase as well as the 2nd phase. The figure also only has 6 data points, but it is plotted with lines connecting the points in order to make the points more visible. The error bars are also shown after each 5 iteration step. As these error bars are quite small compared to the difference in the means of the flat and non-flat, it is clear that using a profile helps for this system. Interestingly, without a 2nd phase (0 2nd phase iterations), the profile does not significantly help. But as the number of 2nd phase iterations completed increases, the effect of the profile becomes much more apparent.

To see this differently, figure 3.4 shows the number of sweeps taken for the non-flat for a specific number of iterations divided by the flat at that same number of iterations. Thus, the flat profile is always 1 and the non-flat profile is a percentage of that. For example the first data point (for 0 iterations) is ≈ 0.975. This means that the non-flat took $\approx 97.5\%$ as many sweeps to complete the same thing as the flat profile.

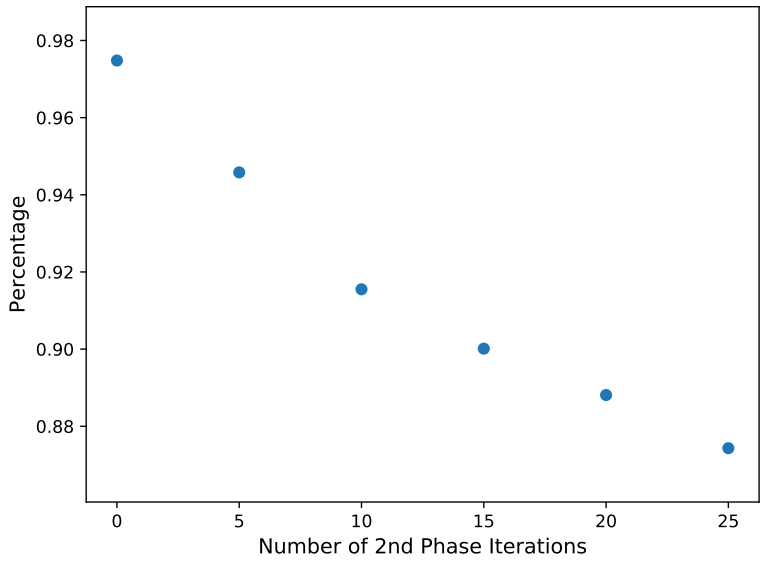

Figure 3.4: Non-Flat Runtime as a percentage of Flat Runtime, 32×32, $q = 10$ system

From figures 3.3 and 3.4, a simple conclusion can be made. Only after the use of a 2nd phase will the effects of a profile become especially apparent (for this particular system at least). The curves also show that performing more iterations in the 2nd phase will increase the benefits of using the profile. Looking at figure 3.4, the slope between each of the points appears to be getting smaller, this would mean that while increasing the number of iterations does result in a lower percentage, the change in percentage is less. Therefore, there could be an assumption that there exists a point at which the number of iterations does not change the percentage gap between flat and non-flat profiles.

3.1.2 Comparing Calculated Density of States

As critical as the runtime is, the whole purpose of the Wang-Landau algorithm is to calculate the density of states. So if the density of states is not properly computed under the use of a profile (or calculated with significant random errors), then the use of a profile in the first place is made useless.

From each of the 250 simulations performed, a slightly different density of states is produced. Each of these is slightly different so it is important to normalize them in such a way that a meaningful comparison can be made.

Importantly, when the term density of states is used here, it refers to the logarithm of the density of states. So all of the comparisons concern the logarithm and not the true density of states.

Standard Error

After normalizing each of simulation's density of states (by the method described in the introduction), the standard error (equation 3.3) of the density of states for each energy can be analyzed. This measures the fluctuations of a theoretically identical value (the density of state of a specific energy) between different trials. The error measures the error in the density of states from a single profile.

Here the error of the mean is calculated using:

$$\sigma_{\ln \bar{g}(E)} = \sqrt{\frac{\mathrm{Var}(\ln g(E))}{250}} \tag{3.3}$$

with E a single energy value.

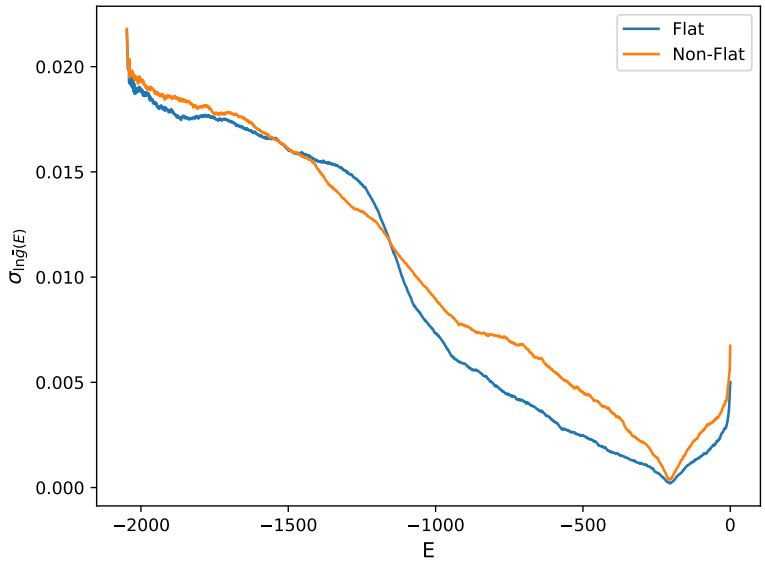

Figure 3.5: Comparison of Standard Error $\sigma_{\ln \bar{g}(E)}$, 32×32, $q = 10$ system

Looking at figure 3.5, the curves are not too interesting. The difference between the profiles shows that perhaps the non-flat has a slightly higher standard error, but even if it is not due to random error, the difference is very small. What is important to see is that the standard error ($\sigma_{\ln \bar{g}(E)}$) is the lowest (near zero) for the energy at which the density of states is the highest. The inverse is true for the ground state as it has the highest standard error and has the lowest density of state. The reason for this is simply due to the normalization method as described in the introduction. As each of the density of states were normalized at nearly the same point, the error for that point should be nearly zero.

Difference of Normalized Density of States

In addition to examining the error of the logarithmic density of states from a single profile, the comparison between the flat and non-flat profiles is also important. To compare them, the ratio of the mean density of states is used. In the logarithm this turns into a difference. To get the mean of the density of states, the mean formula is used:

$$\ln \bar{g}(E) = \frac{1}{250} \sum_{i=1}^{250} \ln g_i(E) \tag{3.4}$$

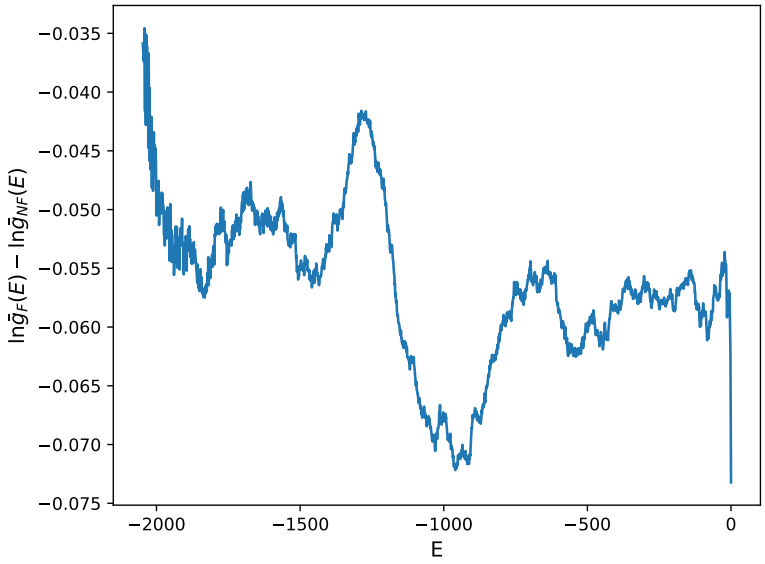

Figure 3.6: Difference of Calculated DoS for Flat and Non-Flat Profiles, 32×32, $q = 10$ system

In figure 3.6, the $\ln \bar{g}_F(E)$ represents the flat profile and $\ln \bar{g}_{NF}(E)$ the non-flat. It is symmetric but only if each half is flipped across both the horizontal and vertical axes. Also, the entire shape is in the negative region, which means that the average for the flat profile is smaller than the non-flat profile. As there is not an exact value for the true density of states, one cannot say which is closer to the true value.

3.2 Wang-Landau Definition of Flatness

Now, the Wang-Landau definition of flatness is tested and analyzed. However, the first thing that must be looked at is how the flatness criteria is defined. In the original paper [4], the definition of flatness is not explicitly clear. The paper states that as soon as the minimum is $x\%$ of the average, it is considered flat. But the exact value of x is not stated. This is simply due to the fact that x varies depending on the system and its parameters. In order to fully examine the effect of a profile on this definition, the values of x are: $x = 0.2$, $x = 0.5$, $x = 0.8$. This represents a low, middle and high amount of flatness.

3.2.1 Simulation Runtimes

First, the simulation runtimes are looked at in the same way as they were for the alternative method of flatness. As tables 3.1 and 3.2 from the alternative method are not particularly useful on their own, they are included in Appendix B and only the resulting figures are included in the main analysis.

$x = 0.2$

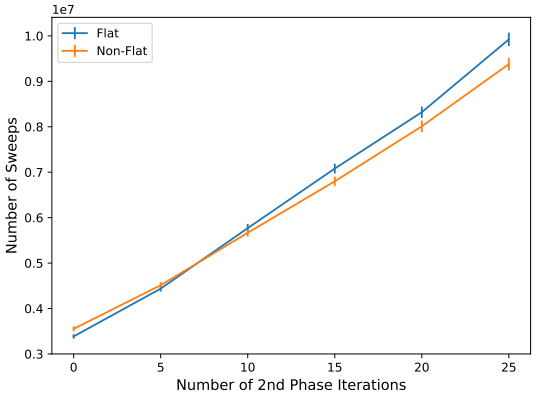

Figure 3.7: Comparison of Runtimes of Wang-Landau Definition of Flatness, 32×32, $q = 10$ system, $x = 0.2$

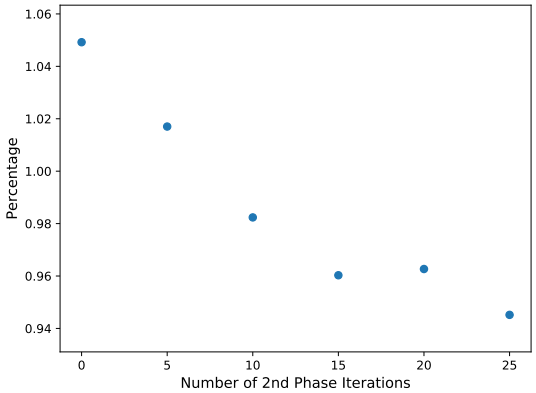

Figure 3.8: Non-Flat Runtime as a Percentage of Flat Runtime, 32×32, $q = 10$ system, $x = 0.2$

$x = 0.5$

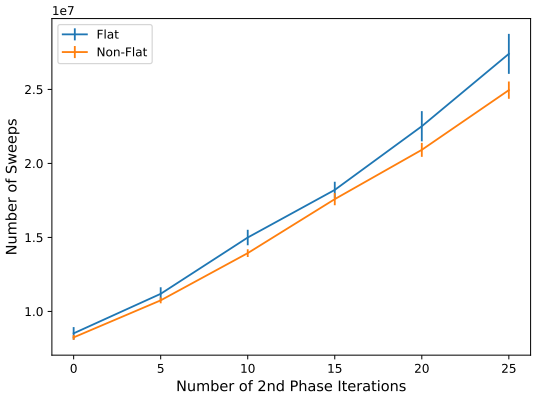

Figure 3.9: Comparison of Runtimes of Wang-Landau Definition of Flatness, 32×32, $q = 10$ system, $x = 0.5$

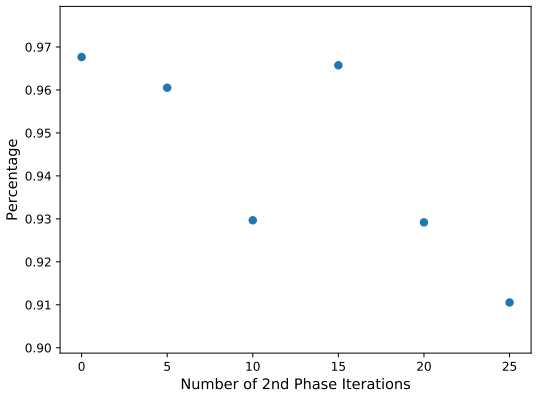

Figure 3.10: Non-Flat Runtime as a Percentage of Flat Runtime, 32×32, $q = 10$ system, $x = 0.5$

Interestingly, for figure 3.10, there is a significant outlier. This may be due to random error but it is interesting as all of the other cases ($x = 0.2$, $x = 0.8$. or alternative method) show a clean decaying behavior.

$x = 0.8$

For $x = 0.8$, the results came out slightly differently. The simulation was allowed to run for 48 hours, and while $x = 0.2$ and $x = 0.5$ each finished 250 different trials (in the 48 hours), the $x = 0.8$ only managed to finish 47 and 77 for flat and non-flat respectively (in the 48 hours). This means that already the non-flat must be quicker as it finished over 50% more trials than the flat profile.

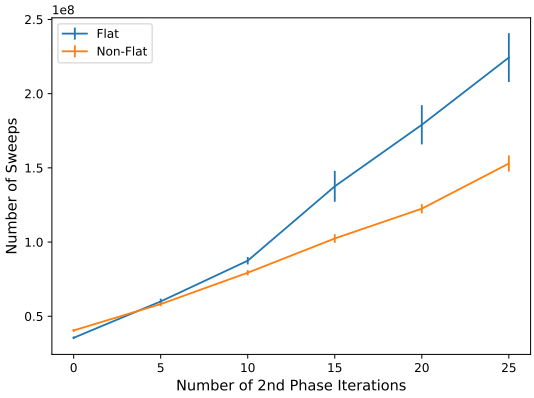

Figure 3.11: Comparison of Runtimes of Wang-Landau Definition of Flatness, 32×32, $q = 10$ system, $x = 0.8$

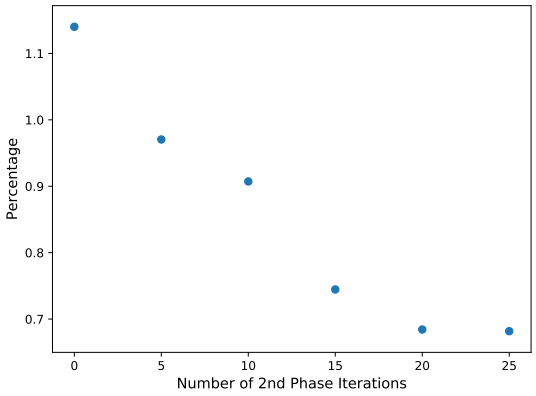

Figure 3.12: Non-Flat Runtime as a Percentage of Flat Runtime, 32×32, $q = 10$ system, $x = 0.8$

Conclusion

After looking at each of the figures, the general consensus is that the non-flat profile is still faster for the Wang-Landau definition. It also does not matter what the minimum $x\%$ is; however, as the scaling factor grows larger ($0.2 \rightarrow 0.8$) the difference after 25 iterations also grows. What does not change significantly is the effect on the 1st phase. While the alternative method and $x = 0.5$ both have the non-flat profile help reduce the total sweeps in the 1st phase, for $x = 0.2$ and $x = 0.8$ the non-flat actually increases the total number of sweeps in the 1st phase slightly.

All of the Wang-Landau flatness simulations also took much longer to finish than the alternative definition. Changing the x value also did nothing significant for the accuracy of the density of states, but it did significantly increase the number of sweeps needed to finish. This can be seen in figure 3.13 for the flat profile and figure 3.14 for the non-flat profile. From these figures, it is very clear that while $x = 0.2$ and $x = 0.5$ are not far apart, $x = 0.8$ takes an order of magnitude longer.

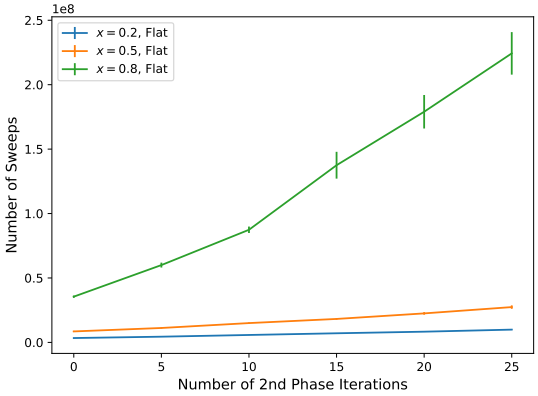

Figure 3.13: Comparison of all of the Runtimes of Wang-Landau Definition of Flatness of a Flat Profile, 32×32, $q = 10$ system

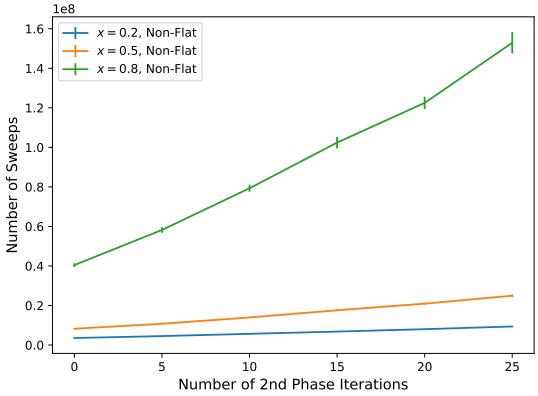

Figure 3.14: Comparison of all of the Runtimes of Wang-Landau Definition of Flatness of a Non-Flat Profile, 32×32, $q = 10$ system

3.2.2 Comparing Calculated Density of States

Again (as with the alternative definition of flatness), the density of states must be equal for both the flat and non-flat profiles in order to verify the success of the non-flat profile. All of the figures are created using the same definition and method as described in the alternative method.

$x = 0.2$

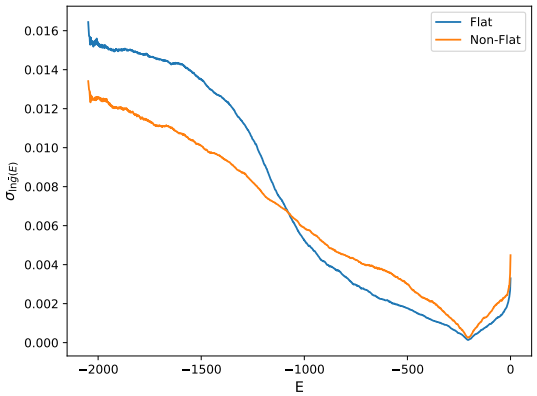

Figure 3.15: Comparison of Standard Errors, 32×32, $q = 10$ system, $x = 0.2$

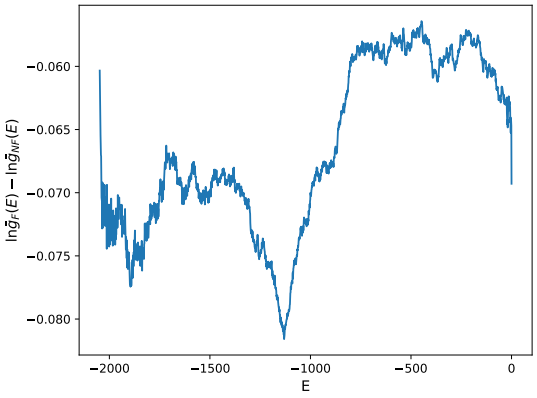

Figure 3.16: Difference of Calculated DoS for Flat and Non-Flat Profiles, 32×32, $q = 10$ system, $x = 0.2$

$x = 0.5$

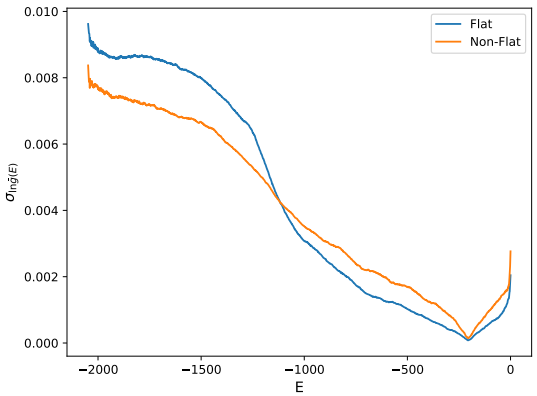

Figure 3.17: Comparison of Standard Errors, 32×32, $q = 10$ system, $x = 0.5$

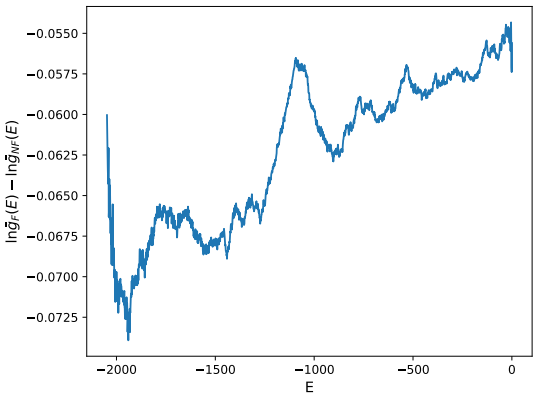

Figure 3.18: Difference of Calculated DoS for Flat and Non-Flat Profiles, 32×32, $q = 10$ system, $x = 0.5$

$x = 0.8$

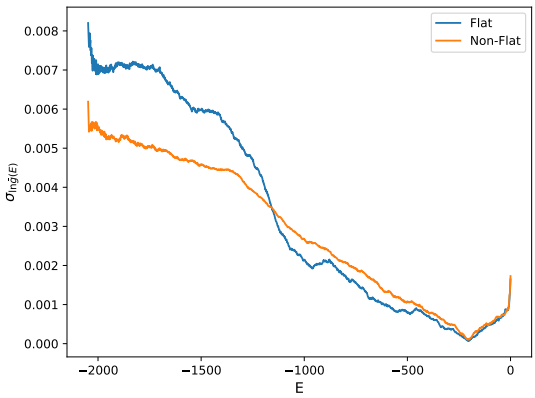

Figure 3.19: Comparison of Standard Errors, 32×32, $q = 10$ system, $x = 0.8$

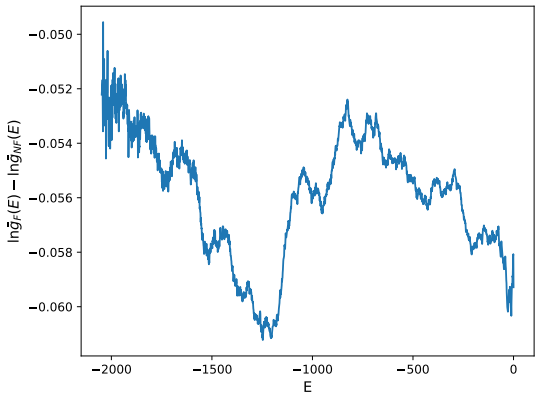

Figure 3.20: Difference of Calculated DoS for Flat and Non-Flat Profiles, 32×32, $q = 10$ system, $x = 0.8$

Conclusion

Looking at the standard errors of the alternative definition (figure 3.5) and the Wang-Landau definition (figures 3.15, 3.17, and 3.19) of flatness, a few general statements can be made. As more sweeps are performed (this corresponds to a higher value of x) the absolute error of both flat and non-flat reduce. This should be expected as a longer runtime will of course reduce random errors. However each of the figures show a consistent trend. On the lower half of the energy (towards the ground state), the flat profile produces a density of state with a higher error, but on the higher half of the energy, the non-flat has a higher error. This may be due to some systematic error if each of the cases produce the same result.

Looking at the difference of the logarithmic density of states for alternative (figure 3.6) and Wang-Landau (figures 3.16, 3.18, and 3.20) shows a different behavior. For these figures, none of their shapes look similar. The only similarity that they do share is that they all are negative, meaning that the density of states for the non-flat is slightly larger than for the flat. As the shapes are all different, the cause is most likely a random error rather than a systematic one that appears to be in the standard errors.

Section 4

$32 \times 32,\ q = 2$ System

In this section a similarly sized system (32×32), with a different amount of spins, is analyzed. By using $q = 2$, the system becomes equivalent to the Ising Model. Using this number of spins will also allow for the observation of the effects of a profile on a second order phase transition.

Here only the alternative definition of flatness was used as it was the definition of flatness (from the $q = 10$ system) that performed the most consistently and quickly. Similarly as for $q = 10$, 250 trials were performed. However, as a single simulation for $q = 2$ completes much more quickly, the number of 2nd phase iterations was allowed to go until 100 instead of 25, with a spacing of 20 instead of 5.

4.0.1 Profile

Dissimilar to the ideal profile for the $q = 10$ system, finding the ideal profile for $q = 2$ is not as straight forward. Using the same algorithm [3] as before leads to the following profile (figure 4.1).

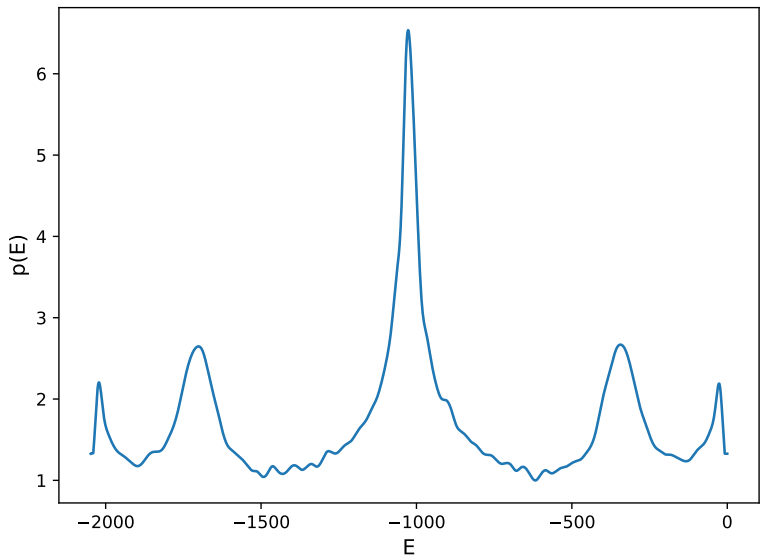

Figure 4.1: Calculated Ideal Profile for 32×32, $q = 2$ system

However, this profile (figure 4.1) does not return better results than a flat profile. This profile (figure 4.1) is called the "calculated profile". Interestingly, this can be corrected by simply deleting the center peak like in figure 4.2. This profile (figure 4.2) is called the "corrected profile". As the algorithm [3] worked for $q = 10$, it is somewhat worrying that the algorithm fails for $q = 2$. If the algorithm works for a certain q, but not another, the profiles produced by the algorithm must be checked against a flat profile. This defeats the purpose of the profile in the first place.

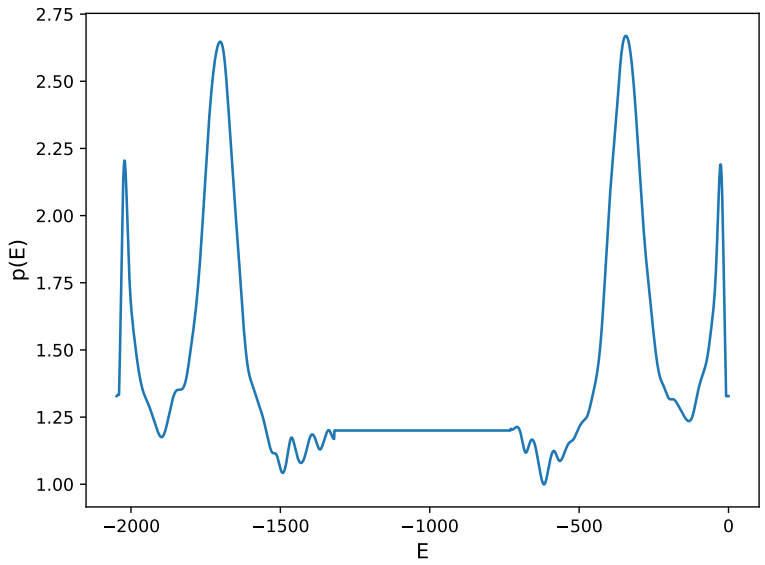

Figure 4.2: Corrected Ideal Profile for 32×32, $q = 2$ system

As the cutting of the center peak is done with little precision (and replaces all of the values with a constant value), the question is whether this is the best profile for $q = 2$? And the answer is probably not. However, the corrective cutting is good enough to make it faster than the flat profile.

4.1 Alternative Definition of Flatness

As mentioned earlier, only the alternative definition of flatness is looked at here.

4.1.1 Simulation Runtimes

Comparison of Runtimes

Looking at the data (tables B.9, B.10, B.11) leads to interesting observations that are dissimilar to the ones made for the $q = 10$ system.

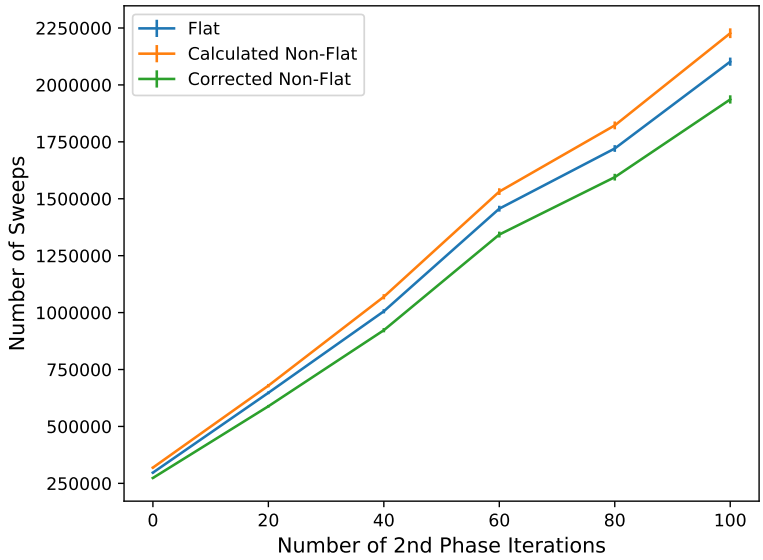

Figure 4.3: Comparison of Runtimes, 32×32, $q = 2$ system

Looking at figure 4.3, the calculated non-flat profile is worse and the corrected non-flat is better as expected (at reducing the number of total sweeps needed). The error bars in figure 4.3 are also too small to see for just a few 2nd phase iterations and just barely visible for the higher number of iterations performed. More interesting is to see the percentages, which also show how the non-flat profiles move closer or further from the flat profile between measurement points.

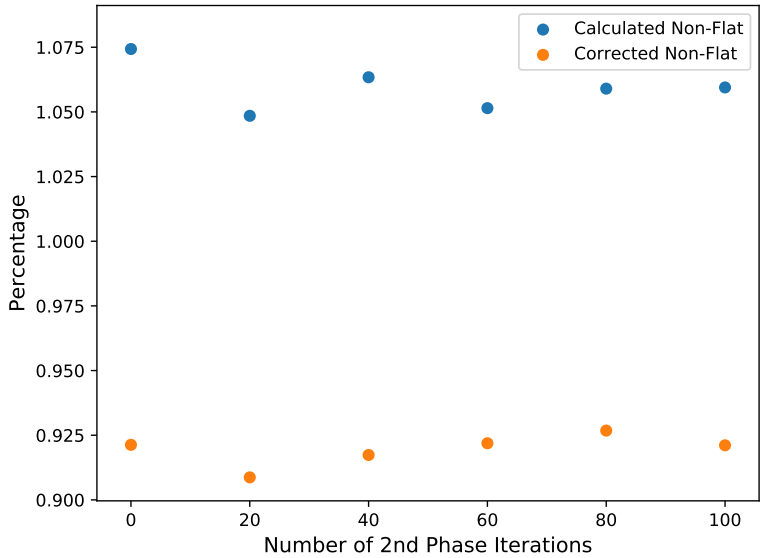

Figure 4.4: Non-Flat Runtime as a Percentage of Flat Runtime, 32×32, $q = 2$ system

Interestingly, in figure 4.4, the percentage does not change over the number of iterations. This means that while the absolute difference between the flat and non-flat grows, the relative change does not. Therefore, the non-flat profile performs equally well in the 1st and 2nd phase. With the $q = 10$ system, this is not true. In the $q = 10$ system, the non-flat profile performs better overall when the 2nd phase is longer. This is an interesting observation that does not have a clear answer.

4.1.2 Comparing Calculated Density of States

As before, the 250 different trials performed each produce a different density of states. As the calculated profile is shown to be worse than the flat profile, only the corrected profile will be looked at here.

Standard Error

The standard error (formula 3.3) for the density of states for each energy is shown in figure 4.5.

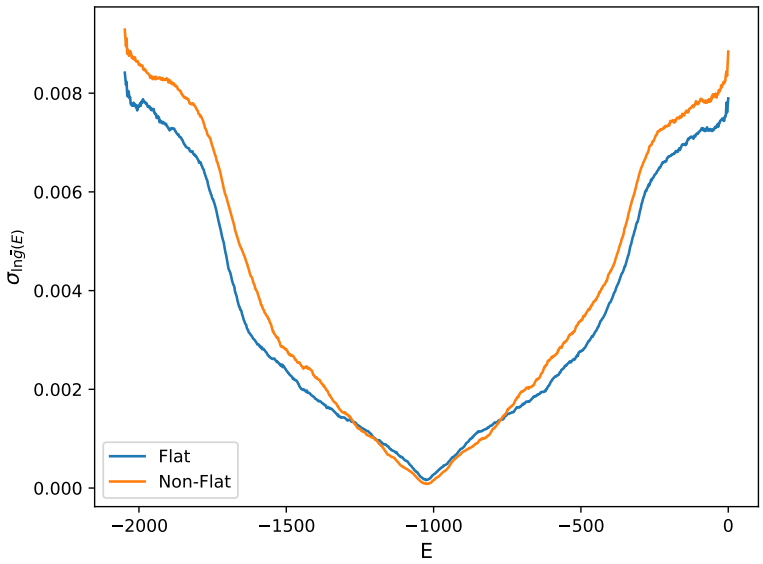

Figure 4.5: Comparison of Std Errors, 32×32, $q = 2$ system

Somewhat as expected, figure 4.5 for $q = 2$ is similar to figure 3.5 for $q = 10$ with the respect to the standard errors. Both feature near zero error at the maximal density of states with the error growing further from this point. This is simply due to the normalization method used.

Difference of Normalized Density of States

The difference of the mean of the logarithmic density of states (equation 3.4) is shown in figure 4.6. The difference is calculated similarly to figure 3.6.

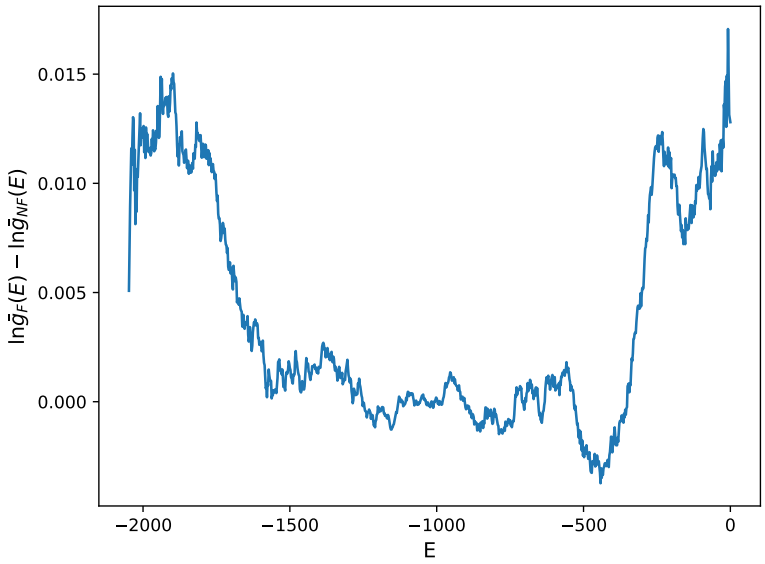

Figure 4.6: Difference of Calculated DoS for Flat and Non-Flat Profiles, 32×32, $q = 2$ system

Conclusion

For $q = 2$ the standard errors (figure 4.5) are appropriately shaped. One difference between the $q = 2$ and $q = 10$ is that the $q = 2$ has the non-flat profile with a slightly higher error over the entire energy range. The difference in the mean of the logarithmic density of states is similar to the $q = 10$ figures except that is positive instead of negative. Therefore the flat logarithmic density of states are larger and not smaller than the non-flat ones. The arbitrary shape in figure 4.6 is most likely due to random errors.

Section 5

$64 \times 64,\ q = 10$ System

By now, it is clear that using the correct profile can reduce the time taken to calculate the density of states for a certain system. However, a major problem is that calculating the profile itself takes more time than simply using a flat profile from the beginning. Therefore, the ideal solution is to use the calculated profile from a smaller system and scale it up for a larger system. Calculating for a smaller system takes much less time and can be more accurate.

5.0.1 Profile

In order to scale up the profile for a 64×64, $q = 10$ system the calculated ideal profile from a 16×16, $q = 10$ and 32×32, $q = 10$ system are were used. These two smaller systems' profiles are quite similar, as shown in figure 5.1. As smaller systems do not have the same energy ranges, the smaller system's profile is stretched such that it fits the larger systems energy range.

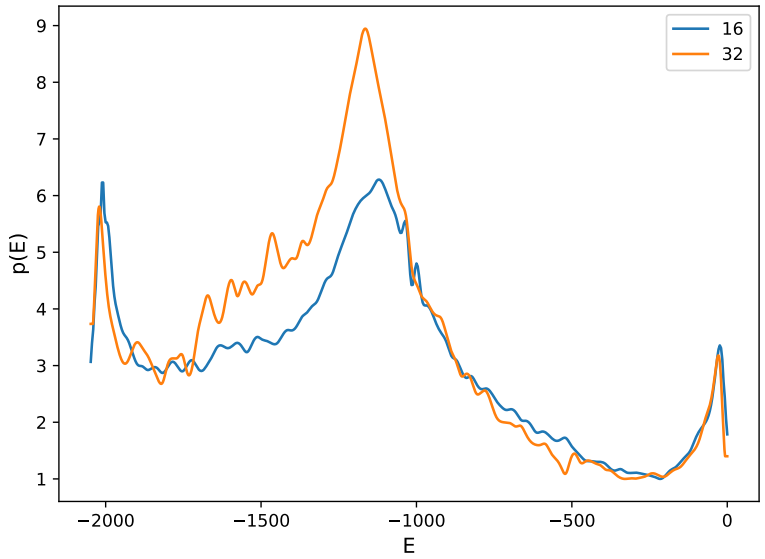

Figure 5.1: Calculated Ideal Profiles for 16×16, $q = 10$ and 32×32, $q = 10$ system

Looking at figure 5.1, the two profiles are nearly the same for most of the energy range. The only difference is at the center peak and for the region on the left of it. In order to make the profile for the 64×64 system, the absolute difference between the 32 and 16 sized systems were added to the 32 sized system.

$$p_{64}(E) = p_{32}(E) + |p_{32}(E) - p_{16}(E)| \tag{5.1}$$

This results in a profile as depicted in figure 5.2.

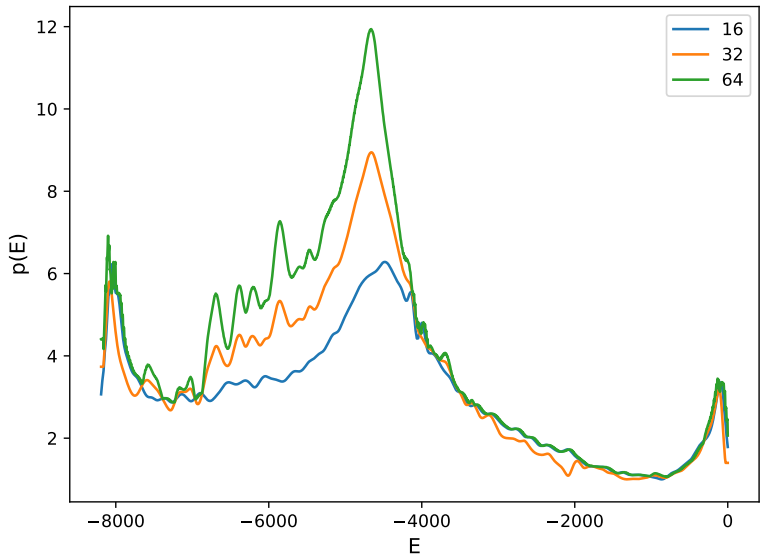

Figure 5.2: Calculated Ideal Profiles for 16×16, $q = 10$, 32×32, $q = 10$, and 64×64, $q = 10$ system

While there was an attempt to produce a profile as per the already used algorithm [3], it did not produce satisfactory results as after a long computation time, only something that still was very noisy was produced. Therefore, there is not a "perfect" profile to exactly compare the profile in figure 5.2 to. However, the algorithmic profile [3] does look similar enough to the profile in figure 5.2 such that the profile in figure 5.2 is not a truly unfounded guess.

5.1 Alternative Definition of Flatness

Once again, only the alternative definition of flatness is used for this system. For this larger 64 sized system, only 50 simulations were run. This is because each one was expected to take much longer than its 32 sized counterpart and therefore fewer were run due to the computer's 48 hour computation time limit.

5.1.1 Simulation Runtimes

Using the data (tables ?? and ??), the same types of observations can be made for this 64×64 system as for the 32×32 system.

Comparison of Runtimes

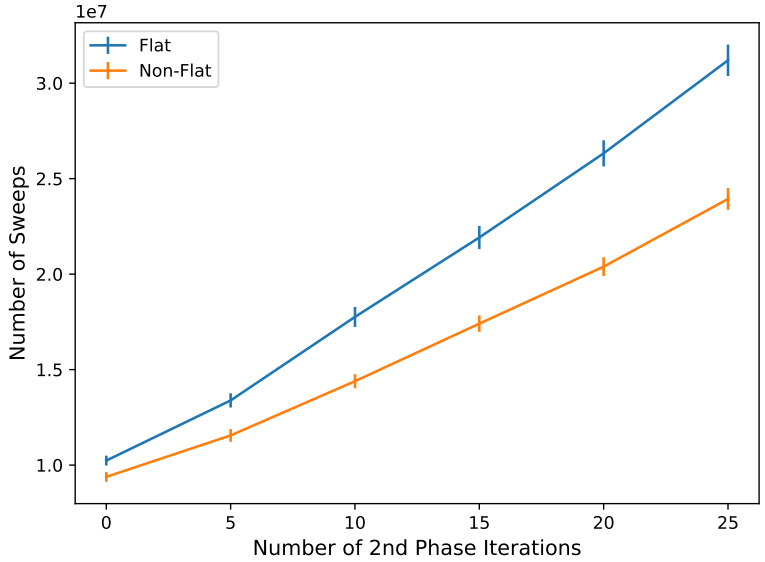

Figure 5.3: Comparison of Runtimes of Alternative Definition of Flatness, 64×64, $q = 10$ system

Figure 5.3 shows another clear example of the non-flat profile helping reduce the number of sweeps needed. In this case, the difference between flat and non-flat looks even better than for the 32 sized system (figure 3.3). Also interesting to note is that the number of sweeps is in the 10^7 range. While this is around 10 times larger than for the 32 sized system, it is a similarly sized value to the Wang-Landau definition ($x = 0.2$ and $x = 0.5$). When compared to the Wang-Landau definition ($x = 0.8$), calculating the density of states for a 64×64 system is actually quicker than for a smaller 32×32 system!

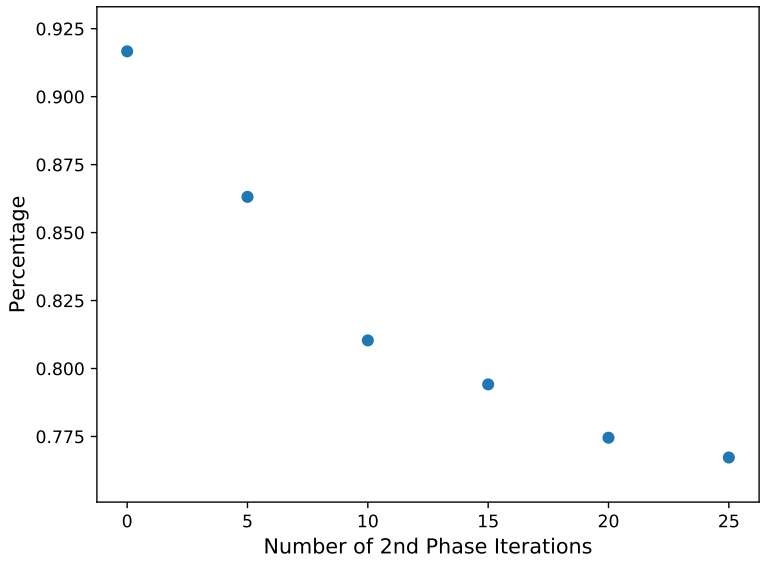

Figure 5.4: Non-Flat Runtime as a Percentage of Flat Runtime, 64×64, $q = 10$ system

Looking at figure 5.4, the percentage change looks the same as for the all of the other $q = 10$ ones. The 1st phase does not benefit significantly from the profile while the 2nd phase does.

5.1.2 Comparing Calculated Density of States

Standard Error

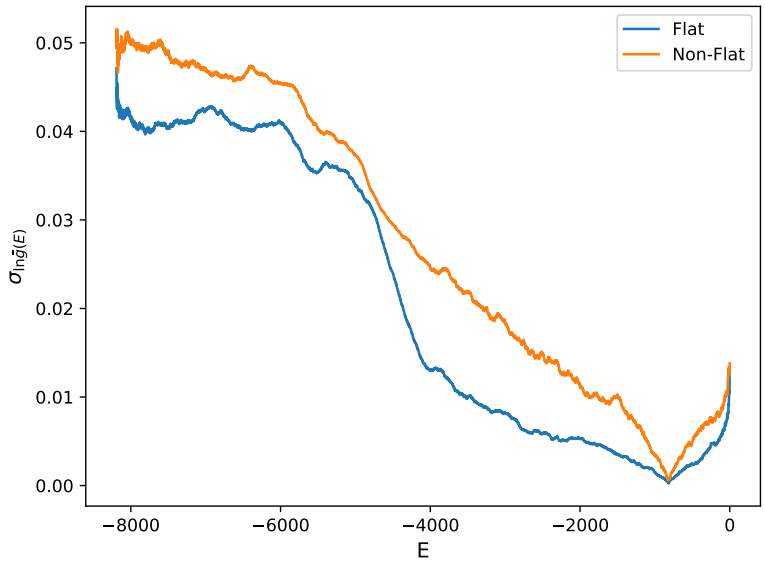

Figure 5.5: Comparison of Standard Errors, 64×64, $q = 10$ system

Compared to the 32×32 systems, this system's errors (figure 5.5) are higher for the non-flat profile across the entire energy range. For the 32×32 system, the error was smaller for the non-flat towards the ground state and become larger towards the 0 energy. The errors are also larger in magnitude, but this should be expected as a larger system will have larger random errors as well as the fact that only 50 iterations were performed instead of the 250 performed for the smaller systems.

Difference of Normalized Density of States

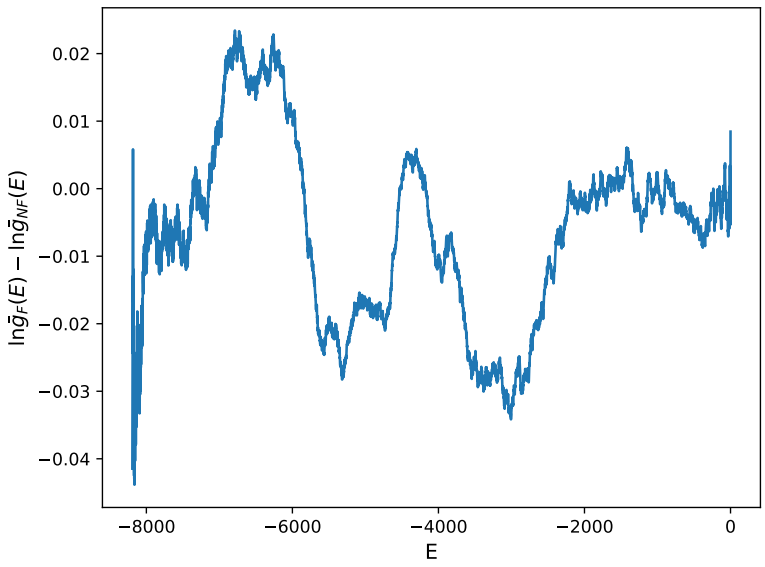

Figure 5.6: Difference of Calculated DoS for Flat and Non-Flat Profiles, 64×64, $q = 10$ system

The difference in the mean of the logarithmic density of states (figure 5.6) "looks" similar to the 32×32 system. This is good as it means that the profile does not systematically affect the calculation of the density of states.

Conclusion

Overall, the use of profile of a smaller system that is scaled up for a larger system does not appear to have any significant drawbacks. However as non-flat has errors in the mean (figure 5.5) and a difference (figure 5.6) from the flat profile, care must be taken in order to assure that the calculated (logarithmic) density of states are correct. The flat profile cannot be used as the real (logarithmic) density of states as it too has errors.

Section 6

Conclusion

While at first the implementation of a profile seems like a rather simple task, it could not be further from the truth. While many many profiles were tested, only after the implementation of another algorithm [3] did a successful profile finally appear to work with the Wang-Landau algorithm [4]. However, even this profile does not come without a drawback. While it is true that the profile itself did not introduce any significant error on the final results (density of states in this case), the profile itself did require significant calculations in order to perfect it. Though as shown in Section 5, this profile calculation time can be mitigated by the successful extrapolation to other lattice sizes. Unfortunately, while the profile can be scaled for lattice size, it is unique for each number of total spins (q) in the system.

While looking at the results of the runtimes of the two different total spin possibilities, ($q = 2$) and ($q = 10$), an interesting conclusion on the profile's effectiveness can be made. For the $q = 2$ case, the profile works equally and consistently well for both the 1st and 2nd phases. For the $q = 10$ case, the profile does not appear to help the 1st phase (and in some cases it actually is worse than the flat profile). The profile only begins to show a clear improvement after a considerable amount of sweeps in the 2nd phase.

Finally, the calculated density of states do not appear to have any considerable statistical errors that come directly from the non-flat profile. As there is not a true density of states to compare the data to, the best method is to see how little the non-flat fluctuates from the flat. As the final mean of the density of states are close to each other, any statistical error must come from the algorithm and not the profile itself. Additionally, as each of the difference of the mean of the logarithms figures show a different shape, it is assured that the shapes come from random rather statistical error.

Beyond just the Ising or Potts model (or even this algorithm), the success of these profiles show that the same or a similar idea can be extended to other calculations and algorithms. In fact, could the algorithm [3] used to create these profiles itself benefit from the use of a profile?

All in all, the shortest path between two points is *not* always a straight line.

Appendix A

Calculation of Macroscopic Variables With the Density of States

After the calculation of the density of states, the analysis of the macroscopic variables can begin to take place (which is why the density of states was created in the first place). As the calculation of the density of states, in the Wang-Landau algorithm, used the natural logarithm of the density of states, the same issue (of the magnitude of the actual density of states) will arise when trying the calculate the expected energy or specific heat.

Using clever logarithm tricks, the problem can be overcome. Below are the steps for calculating the sum of these "impossibly" large values. The partition function is written as:

$$Z = \sum_i g(E_i)\ e^{-\beta E_i} = g(E_0)\ e^{-\beta E_0} + g(E_i)\ e^{-\beta E_i} + ...$$

While Z is too large to be stored, $\log(Z) = z$ is possible:

$$\log(Z) = \log\Big(\sum_i g(E_i)\ e^{-\beta E_i}\Big) = \log\Big(g(E_0)\ e^{-\beta E_0} + g(E_i)\ e^{-\beta E_i} + ...\Big)$$

Unfortunately, the logarithm cannot be broken up easily. However, by way of comparisons the value of z can be calculated:

1. Let $z = -\inf$
2. Let $a = \log\Big(g(E_0)\ e^{-\beta E_0}\Big) = \log\Big(g(E_0)\Big) - \beta E_0$
3. Redefine $z = \max(a, z) + \log\Big(1 + e^{\ \min(a,s)\ -\ \max(a,s)}\Big)$
4. Let $a = \log\Big(g(E_1)\ e^{-\beta E_1}\Big) = \log\Big(g(E_1)\Big) - \beta E_1$
5. Repeat step 3 using the value of a from step 4 and the existing value of z from step 3
6. Repeat steps 4 and 5 until all of the energy values have been represented.

After calculating all of the values, $z = \log(Z)$. Unfortunately, Z is too large to be displayed on its own. However, the calculation of the expected energy allows for the issue to be solved. If:

$$\langle E\rangle = \frac{\sum_i E_i\ g(E_i)\ e^{-\beta E_i}}{Z} = \frac{S}{Z} = \frac{e^{\log S}}{e^{\log Z}} = e^{\log S - \log Z} = e^{s-z}$$

The difference between s and z is small enough such that taking the exponential is not an issue. The calculation of s is identical to z, except for the definition of a. For s:

$$a = \log\Big(E_i\ g(E_i)\ e^{-\beta E_i}\Big) = \log\Big(E_i\Big) + \log\Big(g(E_i)\Big) - \beta E_i$$

As E_i is negative, this is a problem for $\log\Big(E_i\Big)$. However, the solution is to simply add a constant, thus making the value inside the log positive. Thus, for the calculation of s :

$$a = \log\Big(-E_i\Big) + \log\Big(g(E_i)\Big) - \beta E_i$$

For a system of length 60, Wang and Landau [4] present the critical temperature as $T_c = 0.70171$ or $\beta_c = 1.42509$. Using the above formula for the calculation of $\langle E \rangle$ and plotting as a function of β_c, the correct transition temperature is seen, validating this way of calculating these values. In addition, the specific heat as well as free energy $F = -k_b\ T\ \log(Z) = -k_b\ T\ z$ and entropy $S(T)$ can be calculated. These two variables can only be calculated with the density of states as they depend on the exact value of the partition function.

Below are the graphs for a system of length 64 with the critical temperature ($\beta_c = \ln(1+\sqrt{q})$) represented by the vertical line.

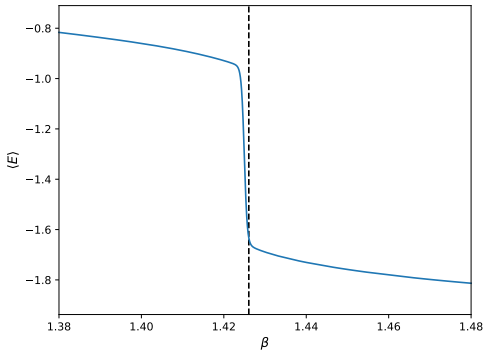

Figure A.1: Expected Energy ($\langle E \rangle$)

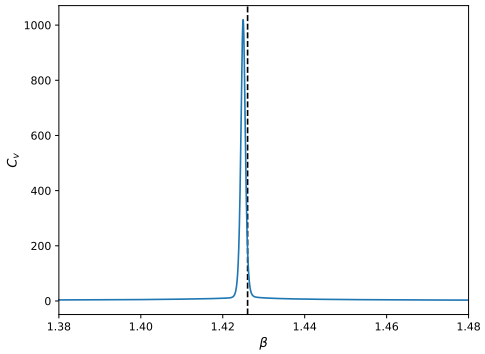

Figure A.2: Specific Heat (C_v)

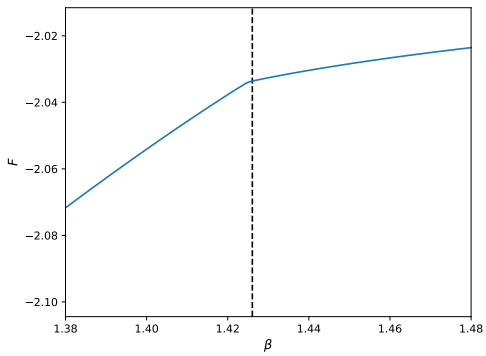

Figure A.3: Free Energy (F)

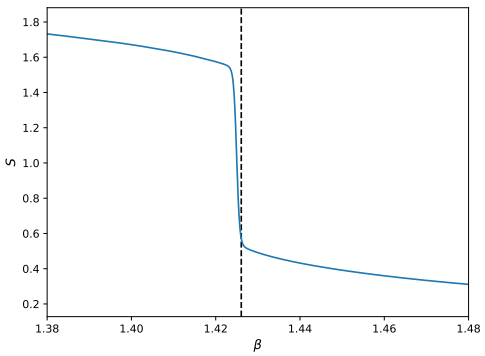

Figure A.4: Entropy (S)

Once the density of states is determined, the calculation of all four variables takes a matter of seconds even for hundreds of data points.

Appendix B

Runtime Data

In this appendix all of the runtime tables are included.

B.1 32×32, $q = 10$

B.1.1 Alternative Definition of Flatness

Table B.1: Flat Profile Values, 32×32, $q = 10$, Alternative

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	2 064 000	28 000
5	2 602 000	36 000
10	3 314 000	45 000
15	3 985 000	53 000
20	4 674 000	62 000
25	5 522 000	78 000

Table B.2: Non-Flat Profile Values, 32×32, $q = 10$, Alternative

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	2 012 000	20 000
5	2 461 000	24 000
10	3 034 000	29 000
15	3 587 000	33 000
20	4 151 000	36 000
25	4 828 000	42 000

B.1.2 $x = 0.2$

Table B.3: Flat Profile Values, 32×32, $q = 10$, $x = 0.2$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	3 390 000	50 000
5	4 430 000	70 000
10	5 770 000	90 000
15	7 080 000	110 000
20	8 320 000	130 000
25	9 920 000	150 000

Table B.4: Non-Flat Profile Values, 32×32, $q = 10$, $x = 0.2$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	3 550 000	50 000
5	4 510 000	70 000
10	5 670 000	80 000
15	6 800 000	100 000
20	8 010 000	130 000
25	9 380 000	140 000

B.1.3 $x = 0.5$

Table B.5: Flat Profile Values, 32×32, $q = 10$, $x = 0.5$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	8 520 000	420 000
5	11 200 000	440 000
10	14 990 000	520 000
15	18 210 000	550 000
20	22 510 000	1 020 000
25	27 400 000	1 350 000

Table B.6: Non-Flat Profile Values, 32×32, $q = 10$, $x = 0.5$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	8 250 000	170 000
5	10 750 000	200 000
10	13 940 000	260 000
15	17 590 000	410 000
20	20 920 000	470 000
25	24 950 000	580 000

B.1.4 $x = 0.8$

Table B.7: Flat Profile Values, 32×32, $q = 10$, $x = 0.8$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	35 500 000	870 000
5	60 000 000	1 900 000
10	87 400 000	2 500 000
15	137 500 000	10 400 000
20	179 000 000	13 000 000
25	224 300 000	16 500 000

Table B.8: Non-Flat Profile Values, 32×32, $q = 10$, $x = 0.8$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	40 440 000	930 000
5	58 200 000	1 400 000
10	79 300 000	1 800 000
15	102 400 000	2 900 000
20	122 500 000	3 100 000
25	152 900 000	5 400 000

B.2 $32 \times 32,\ q = 2$

Table B.9: Flat Profile Values, $32 \times 32,\ q = 2$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	297 000	3 000
20	648 000	6 000
40	1 006 000	9 000
60	1 457 000	13 000
80	1 721 000	15 000
100	2 102 000	18 000

Table B.10: Calculated Non-Flat Profile Values, $32 \times 32,\ q = 2$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	319 000	4 000
20	679 000	7 000
40	1 070 000	11 000
60	1 531 000	15 000
80	1 822 000	17 000
100	2 227 000	21 000

Table B.11: Corrected Non-Flat Profile Values, $32 \times 32,\ q = 2$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	274 000	3 000
20	589 000	6 000
40	923 000	9 000
60	1 343 000	13 000
80	1 595 000	15 000
100	1 936 000	18 000

B.3 $64 \times 64,\ q = 10$

Table B.12: Flat Profile Values, $64 \times 64,\ q = 10$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	10 240 000	260 000
5	13 390 000	370 000
10	17 760 000	520 000
15	21 920 000	610 000
20	26 320 000	680 000
25	31 200 000	820 000

Table B.13: Non-Flat Profile Values, $64 \times 64,\ q = 10$

Number of 2nd Phase Iterations	$\ln \bar{g}(E)$	$\sigma_{\ln \bar{g}(E)}$
0	9 380 000	260 000
5	11 550 000	330 000
10	14 390 000	370 000
15	17 410 000	430 000
20	20 390 000	490 000
25	23 940 000	570 000

References

[1] Richard Fitzpatrick. *The Ising model*. Last Visited 10/05/2019. Mar. 2006. URL: http://farside.ph.utexas.edu/teaching/329/lectures/node110.html.

[2] D. P. Landau, Shan-Ho Tsai, and M. Exler. "A new approach to Monte Carlo simulations in statistical physics: Wang-Landau sampling". In: *American Journal of Physics* 72.10 (Oct. 2004), pp. 1294–1302. DOI: 10.1119/1.1707017. eprint: https://doi.org/10.1119/1.1707017. URL: https://doi.org/10.1119/1.1707017.

[3] Simon Trebst, David A. Huse, and Matthias Troyer. "Optimizing the ensemble for equilibration in broad-histogram Monte Carlo simulations". In: *Phys. Rev. E* 70 (4 Oct. 2004), p. 046701. DOI: 10.1103/PhysRevE.70.046701. URL: https://link.aps.org/doi/10.1103/PhysRevE.70.046701.

[4] Fugao Wang and D. P. Landau. "Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram". In: *Phys. Rev. E* 64 (5 Oct. 2001), p. 056101. DOI: 10.1103/PhysRevE.64.056101. URL: https://link.aps.org/doi/10.1103/PhysRevE.64.056101.

[5] Wikipedia. *Metropolis–Hastings algorithm*. Last Visited 10/05/2019. 2019. URL: https://en.wikipedia.org/wiki/Metropolis%E2%80%93Hastings_algorithm.

[6] Wikipedia contributors. *Potts model — Wikipedia, The Free Encyclopedia*. [Online; accessed 31-May-2019]. 2019. URL: https://en.wikipedia.org/w/index.php?title=Potts_model&oldid=895015463.

Acknowledgements

I would like to first and foremost thank Dr. Stefan Schnabel for his many hours in helping me develop the ideas on how to better understand both the physical system and computer code that I was working with. I would especially like to thank him for his corrections of my numerous naive assumptions. Finally, I would like to thank him for this implementation of the algorithm [3] that computed the ideal profile. Without his help in this aspect, I could not have progressed successfully. I would also like to thank Prof. Dr. Wolfhard Janke for his general help with both the thesis and the interesting $1/t$ extension as a project.

Declaration of Originality

I hereby confirm, that I have written this thesis independently and without illicit help of others. No other sources, but those given in the bibliography, were used. All included figures and tables were generated independently by means of *LaTeX*, *python* and *matplotlib*. Calculations for the numerical data was created using *C++* and analyzed using *python*. This work was not submitted in same or similar form yet.

Alex Panchot