

## 1 Thermodynamic Properties of the Ising Model

If we want to calculate the thermodynamic properties of a system at temperature  $T$  (that is, in equilibrium) that has a Hamiltonian  $\mathcal{H}(\mathbf{x})$ , we have the problem of calculating a high dimensional integral with a rapidly varying integrand.

- Here “ $\mathcal{H}(\mathbf{x})$ ” is the energy of the system for a given *configuration*  $\mathbf{x}$ . A configuration is a specification of the state of the system in some relevant variables.
- For a quantum mechanical system of  $N$  spinless particles, a configuration could be the  $N$  position vectors.
- For a system of spins on a lattice, a configuration could be a specification of the spin projection at each lattice point.

We’ll use the Ising model in one and two dimensions as an example. The system consists of a set of lattice sites with spins, which take on only two possible values,  $+1$  or  $-1$ . The spins are arranged in a chain (one-dimensional Ising model) or a square lattice (two-dimensional Ising model). The interaction between spins is short-ranged, which is represented in our model by interactions only between *nearest neighbors*. That is, there is potential energy only from adjacent spins. Each spin has two nearest neighbors in one dimension and four nearest neighbors in two dimensions (diagonals don’t count here!).

- We have to specify what happens at the edge of the system. E.g., do the ends of the one-dimensional chain have only one nearest neighbor, or do we imagine the chain wrapped into a continuous band? In the latter case, we have *periodic boundary conditions*, because it is equivalent to imagining an infinite chain that repeats itself periodically.

We also allow for an external constant magnetic field,  $h$ , which interacts with each of the spins individually to contribute a Zeeman energy. (Note: the notation  $h$  for the magnetic field to avoid confusion with the Hamiltonian. In other references such as Landau and Paez,  $h \rightarrow \mu B$  is used instead.)

We choose units so that the Hamiltonian takes the simple form:

$$\mathcal{H}(\mathbf{x}) = -J \sum_{\langle i,j \rangle} S_i S_j - \sum_i h S_i , \quad (1)$$

where  $S_i = \pm 1$  is the spin.

- Here  $\mathbf{x}$  is just a shorthand for a specification of the spin  $S_i$  on each site. For a one-dimensional Ising model with  $N$  sites, you could store this as an array of length  $N$  with each element either  $+1$  or  $-1$ .
- The notation  $\langle i, j \rangle$  stands for “nearest neighbor” pairs of spins. Obviously the product  $S_i S_j$  can be either  $+1$  or  $-1$ .
- In many applications we’ll set the external field  $h$  to zero.
- The constant  $J$  (called the “exchange energy”) specifies the strength of the spin-spin interaction. For a ferromagnetic interaction,  $J > 0$ , while for an “anti-ferromagnetic interaction”  $J < 0$ . *What do you think is the physical origin of  $J$  for a ferromagnetic interaction? See below.*
- The *magnetization* of the system is the average value of the spin,  $\langle S_i \rangle$ , where the average is taken over the entire lattice. If the spins are equally likely to be up ( $+1$ ) as down ( $-1$ ), then the net magnetization is zero. In an external magnetic field, up or down will be favored energetically (if  $h > 0$  in  $\mathcal{H}(\mathbf{x})$  above, which is favored?), and there will generally be a net magnetization. In a ferromagnet at sufficiently low temperature, there will be *spontaneous* magnetization even in the absence of an external field. This may occur over regions (“domains”) or over the entire lattice.

A configuration in the Ising model is a *microstate* of the system. *If we have  $N$  sites in a one-dimensional Ising model, how many possible configurations (or microstates) are there? How many different values of the energy (with  $h = 0$ )?*

The thermal average of a quantity  $A$  that depends on the configuration  $\mathbf{x}$  (examples of  $A$  are the energy per degree of freedom  $E = \langle \mathcal{H} \rangle_T / N$  or the magnetization per degree of freedom  $M = \langle \sum_i S_i \rangle_T / N$ ) is given by the *canonical ensemble* (the denominator is the *partition function*  $Z$ )

$$\langle A(\mathbf{x}) \rangle_T = \frac{\int d\mathbf{x} A(\mathbf{x}) e^{-\mathcal{H}(\mathbf{x})/kT}}{\int d\mathbf{x}' e^{-\mathcal{H}(\mathbf{x}')/kT}} = \int d\mathbf{x} A(\mathbf{x}) P_{\text{eq}}(\mathcal{H}(\mathbf{x})), \quad (2)$$

where we have pulled out the Boltzmann factors to identify the probability distribution function  $P_{\text{eq}}(\mathcal{H}(\mathbf{x}))$ :

$$P_{\text{eq}}(\mathcal{H}(\mathbf{x})) = \frac{e^{-\mathcal{H}(\mathbf{x})/kT}}{\int d\mathbf{x}' e^{-\mathcal{H}(\mathbf{x}')/kT}}. \quad (3)$$

*(Is this PDF normalized? Yes!)* In the Ising model example, the integration over  $\mathbf{x}$  is just a sum over the spin configurations. If our configurations are chosen as eigenstates of energy (which they are in the Ising model example), then we can just sum over energies rather than

every configuration:

$$\langle A \rangle_T = \frac{\sum_E (\# \text{ of states with energy } E) A(E) e^{-E/kT}}{\sum_{E'} (\# \text{ of states with energy } E') e^{-E'/kT}} = \sum_E A(E) P(E) . \quad (4)$$

If we can construct a set of  $N$  configurations  $\{\tilde{\mathbf{x}}_i\}$  (a statistical sample) that are distributed according to  $P$ , then we can apply our Monte Carlo method (i.e., we will be doing importance sampling) and  $\langle A \rangle_T = (1/N) \sum_i A(\tilde{\mathbf{x}}_i)$ . This is what the Metropolis algorithm does for us!

## 2 Exchange Integral $J$

[This summary is based on the discussion in Shang-Keng Ma's book on *Statistical Mechanics*.]

Let's look at the physical origin of the exchange integral  $J > 0$  for a ferromagnet. The short-distance exchange interaction is the strongest among electron spins, and comes from the interplay of electrostatic repulsion and the Pauli exclusion principle. Consider the interaction between electrons at two sites. If it were just a magnetic dipole interaction, then  $\uparrow\uparrow$  would not be energetically favored and we would not have ferromagnets! To find the true origin of the interaction, we must turn to quantum mechanics.

Consider the wave function of electrons 1 and 2 as a product of a spatial part  $\varphi$  and a spin part  $\chi$  (this factorization is possible if there is negligible spin-orbit coupling, which we assume):

$$\psi(1, 2) = \varphi(\mathbf{r}_1, \mathbf{r}_2) \chi(s_1, s_2) , \quad (5)$$

where the  $\mathbf{r}_i$ 's are the positions of the electrons and the  $s_i$ 's are the spins. Since electrons are identical *fermions*, the overall wavefunction must be *antisymmetric* under particle exchange (i.e.,  $\psi(2, 1) = -\psi(1, 2)$ ). That leaves two possibilities:  $\varphi$  is symmetric while  $\chi$  is antisymmetric, or *vice versa*. For example,

$\chi$	implication for $\varphi$
$\uparrow\uparrow$	$\varphi(\mathbf{r}_1, \mathbf{r}_2) = -\varphi(\mathbf{r}_2, \mathbf{r}_1)$
$\frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow)$	$\varphi(\mathbf{r}_1, \mathbf{r}_2) = +\varphi(\mathbf{r}_2, \mathbf{r}_1)$

Thus the symmetry of the spins dictates the symmetry of the spatial part; this is the key!

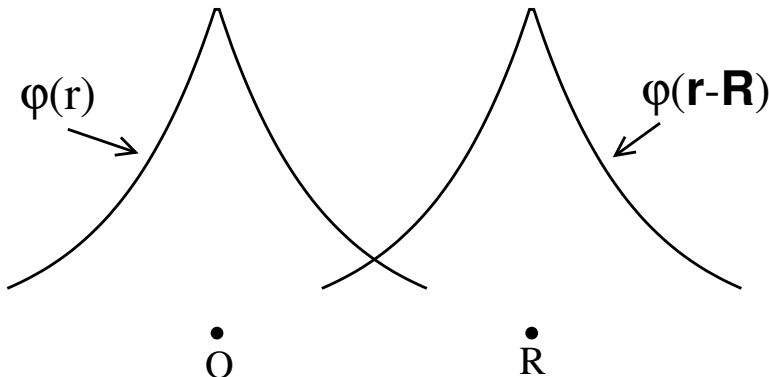
Consider first the non-interacting case and imagine the spatial wavefunctions are like hydrogen ground state wavefunctions centered at the origin  $O$  and point  $R$ . So we make the identification

$$\varphi_1(\mathbf{r}) = \varphi(r) , \quad \varphi_2(\mathbf{r}) = \varphi(\mathbf{r} - \mathbf{R}) , \quad (6)$$

as pictured in the figure below. Then the noninteracting wave function is either the symmetric ( $\varphi_+$ ) or antisymmetric ( $\varphi_-$ ) combination:

$$\varphi_{\pm}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}[\varphi_1(\mathbf{r}_1)\varphi_2(\mathbf{r}_2) \pm \varphi_1(\mathbf{r}_2)\varphi_2(\mathbf{r}_1)] \quad (7)$$

Notice that  $\varphi_-$  vanishes when  $\mathbf{r}_1 = \mathbf{r}_2$ .



Now consider the Coulomb repulsion between the electrons:

$$U(\mathbf{r}_1 - \mathbf{r}_2) = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} . \quad (8)$$

If we calculate the change in energy  $\delta E$  in first-order perturbation theory, then

$$\delta E = \langle \varphi(1, 2) | U | \varphi(1, 2) \rangle . \quad (9)$$

The spin part is diagonal and thus contributes only in the sense of telling us whether  $\varphi_+$  or  $\varphi_-$  is the spatial wave function. We can label the two possibilities  $\delta E_{\pm}$ , with

$$\delta E_{\pm} \equiv \langle \varphi_{\pm} | U | \varphi_{\pm} \rangle = \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 |\varphi_1(\mathbf{r}_1)|^2 |\varphi_2(\mathbf{r}_2)|^2 \frac{2e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \pm \frac{J}{2} , \quad (10)$$

which defines the *exchange integral*  $J$  as

$$J \equiv \int d^3\mathbf{r}_1 d^3\mathbf{r}_2 \varphi_1(\mathbf{r}_2)\varphi_2(\mathbf{r}_1)\varphi_1(\mathbf{r}_1)\varphi_2(\mathbf{r}_2) \frac{2e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} > 0 . \quad (11)$$

So what do we learn?

- $J$  is positive definite (for our example), so  $J > 0$  means *lower* energy for  $E_-$ .
- But  $E_-$  implies  $\varphi_-$ , which means an anti-symmetric spatial wave function and therefore symmetric (e.g., parallel) spins.
- Thus the parallel alignment is preferred to lower the energy! In short, if the spins are parallel, the electrons are less likely to be near each other, and therefore will have a lower average potential energy from electrostatic repulsion.
- We can evaluate Eq. (11) using hydrogenic wave functions to get a quantitative estimate for  $J$ .

### 3 Metropolis Algorithm

The first thing to get straight is that the Metropolis algorithm has nothing to do with Superman or Fritz Lang :). It is named for the first author on the paper that described the algorithm. (A bit of trivia: another author on the paper is Edward Teller.) The Metropolis algorithm generates a *Markov chain*; this is a sequence of configurations  $\mathbf{x}_i$  that will be distributed according to the canonical distribution (i.e., the Boltzmann factor will tell us the relative probability of any configuration).

The Markov process constructs state  $\mathbf{x}_{l+1}$  from the previous state  $\mathbf{x}_l$  according to a *transition probability*  $W(\mathbf{x}_l \rightarrow \mathbf{x}_{l+1})$ . The idea is to construct  $W$  such that in the limit of a large number of configurations, the distribution of states  $\mathbf{x}_i$  approaches the equilibrium Boltzmann distribution  $P_{\text{eq}}$  from Eq. (3).

A sufficient condition for this to happen is that the principle of *detailed balance* should hold. In equilibrium, the *rate* of  $\mathbf{x} \rightarrow \mathbf{x}'$  should equal the rate of  $\mathbf{x}' \rightarrow \mathbf{x}$ , or else it is not equilibrium! In calculating these rates, there are two terms to multiply:

$$\left( \frac{\text{probability}}{\text{time}} \text{of } \mathbf{x} \rightarrow \mathbf{x}' \right) \times (\# \text{ of } \mathbf{x}'\text{'s}) = \left( \frac{\text{probability}}{\text{time}} \text{of } \mathbf{x}' \rightarrow \mathbf{x} \right) \times (\# \text{ of } \mathbf{x}'\text{'s}) \quad (12)$$

or

$$W(\mathbf{x} \rightarrow \mathbf{x}') \times N P_{\text{eq}}(\mathbf{x}) = W(\mathbf{x}' \rightarrow \mathbf{x}) \times N P_{\text{eq}}(\mathbf{x}') , \quad (13)$$

which is detailed balance. This means that

$$\frac{W(\mathbf{x} \rightarrow \mathbf{x}')}{W(\mathbf{x}' \rightarrow \mathbf{x})} = \frac{P_{\text{eq}}(\mathbf{x})}{P_{\text{eq}}(\mathbf{x}')}. \quad (14)$$

But we know that the probability of configuration  $\mathbf{x}$  is

$$P_{\text{eq}}(\mathbf{x}) = \frac{1}{Z} e^{-\mathcal{H}(\mathbf{x})/kT} . \quad (15)$$

Any choice of  $W$  satisfying Eq. (14) should work, in principle.

For example, we can take

$$W(\mathbf{x}_a \rightarrow \mathbf{x}_b) = \begin{cases} \frac{1}{\tau_s} e^{-\delta E/kT} & \text{if } \delta E = E_{\mathbf{x}_b} - E_{\mathbf{x}_a} > 0 \\ \frac{1}{\tau_s} & \text{if } \delta E = E_{\mathbf{x}_b} - E_{\mathbf{x}_a} \leq 0 \end{cases} \quad (16)$$

with  $\tau_s$  arbitrary for now (set it equal to unity for convenience). We can check that it works by considering all possible outcomes, as shown in this chart:

	$\delta E_{\mathbf{x} \rightarrow \mathbf{x}'}$	$W(\mathbf{x} \rightarrow \mathbf{x}')$	$e^{-E/kT}$	$\delta E$	$W(\mathbf{x}' \rightarrow \mathbf{x})$	$e^{-E/kT}$
$E_{\mathbf{x}'} > E_{\mathbf{x}}$	$> 0$	$\frac{1}{\tau_s} e^{-(E_{\mathbf{x}'} - E_{\mathbf{x}})/kT}$	$e^{-E_{\mathbf{x}}/kT}$	$< 0$	$\frac{1}{\tau_s}$	$e^{-E_{\mathbf{x}'}/kT}$
$E_{\mathbf{x}'} < E_{\mathbf{x}}$	$< 0$	$\frac{1}{\tau_s}$	$e^{-E_{\mathbf{x}}/kT}$	$> 0$	$\frac{1}{\tau_s} e^{-(E_{\mathbf{x}} - E_{\mathbf{x}'})/kT}$	$e^{-E_{\mathbf{x}'}/kT}$

If you do the multiplications, you'll see that detailed balance is satisfied in every case.

The basic idea in implementing Metropolis for the Ising model is that from a starting configuration, one generates a candidate for a new configuration somehow (for example, by flipping one randomly selected spin). After calculating the energy change between new and old, one has the transition probability  $W$  from Eq. (16). Generate a random number between 0 and 1; if the number is less than  $\tau_s W$  keep the new configuration but otherwise flip the spin back and keep this as the “new” configuration.

Let's be extremely explicit about how we can implement this. Take  $\tau_s = 1$ . We have a starting configuration with energy  $E_a$  and a new candidate configuration with energy  $E_b$ . If  $E_b < E_a$ , then we accept the new candidate always (i.e., with probability 1). That is, if the energy is lowered, we always keep the new one. This agrees with Eq. (16). If the energy is raised, namely  $E_b > E_a$ , we keep it *some of the time*. [You might think we should always just go “downhill” in energy, but we know this is wrong: if so, we would not have a chance to follow any path through the configuration space (because we would never go uphill), so we would not have ergodicity.] Following Eq. (16), we should keep it with probability  $e^{-\delta E/kT} = e^{-(E_b - E_a)/kT}$ . This is a number between 0 and 1. Say it is 0.3. Then we generate a random number uniformly distributed between 0 and 1. If it is less than 0.3 we keep the new configuration; if not, we reject it. (To see that this is the correct procedure, imagine doing it 1000 times. You would expect roughly 300 times to get a number less than 0.3, so keeping those would give the desired probability.)

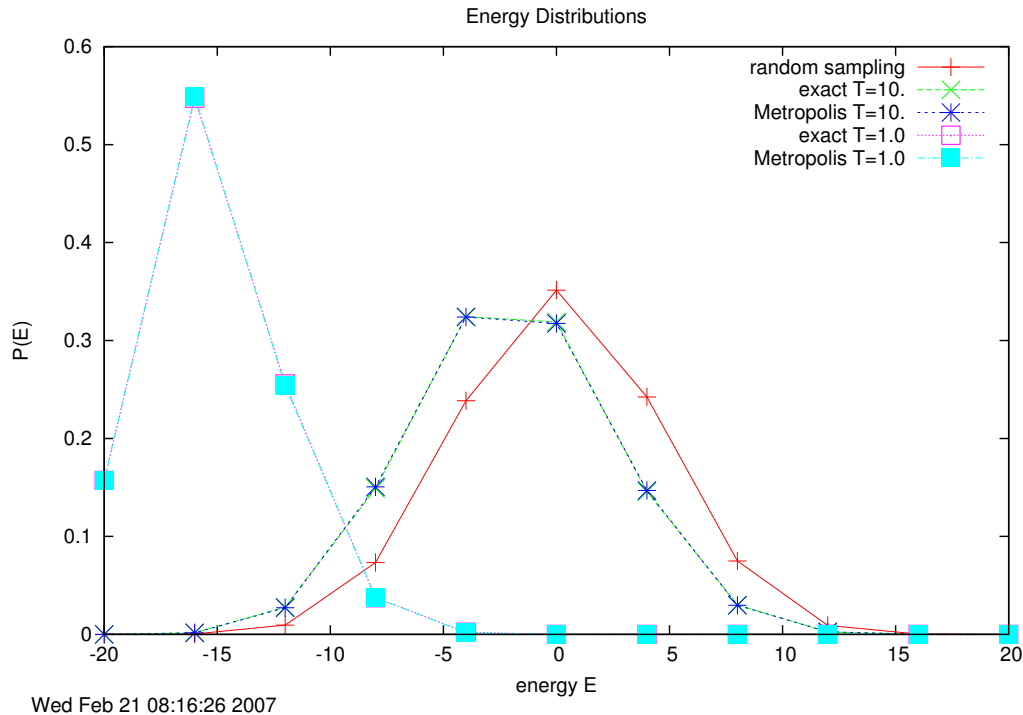
There are various implementation issues. These include:

- **Equilibration.** If we are going to use configurations generated by Metropolis to calculate thermal averages, we want to make sure we have reached “equilibrium.” How long does this take and what is the signature?
- **Autocorrelation.** When evaluating an average over Monte Carlo configurations, we want to skip the first  $n_1$  steps and then use data taken every  $n_0$  Monte Carlo steps. How do we determine how large to take  $n_0$  and  $n_1$ ? The *autocorrelation function* can be used to determine a reasonable choice for  $n_0$ .

## 4 Follow-up notes

- Recap of energy distributions figure.

Here is a typical figure you might have obtained from the `sampling_test` program:



Some observations:

- The “random sampling” selects configurations (recall that a configuration is one particular specification of spins in the chain) at random, so each one is equally likely to be chosen. Then the energy is calculated and histogrammed. Since there are only two out of about one million configurations that have the lowest energy of  $-20$ , the curve will be almost zero there. The peak in the middle means that there are more ways to have zero energy than any other energy. The curve is symmetric because there are equal numbers of configurations with a given positive energy as with the corresponding negative energy.
- The exact  $P(E)$  at temperature  $T$  combines two factors: the Boltzmann factor  $e^{-E/kT}/Z$  and the number of states with energy  $E$ . For  $T = 1$  (we use units where  $k = 1$ ), the first factor is largest for  $E = -20$  but the second is smallest there. The product is a moderate probability. For  $E = -16$ , the first factor is smaller but the second is sufficiently larger that the net result is the maximum probability. Even though the second factor is largest at  $E = 0$ , the Boltzmann factor is very small (remember that  $Z$  is large), so there is only a very small probability  $P(0)$ .

- For larger temperatures, the Boltzmann factor is more uniform, and so the factor of the number of states with energy  $E$  dominates. Thus, the random sampling distribution corresponds to very high temperature.
- The Metropolis distribution uses only a representative subset of the full set of configurations to estimate  $P(E)$ . But it does a very good job of reproducing the exact distribution at a given temperature.

- **Thermal averages.** Recall the basic idea for evaluating thermal averages. We generate a set of “configurations,” which we’ll label  $\mathbf{x}_0, \mathbf{x}_1, \dots$  or collectively as  $\{\mathbf{x}_i\}$ , that are distributed according to an appropriate probability distribution function (PDF). One of the most physical PDF’s is the Boltzmann distribution. In this case, the thermal average of  $A(\mathbf{x})$  is approximated by

$$\langle A \rangle_T \equiv \frac{\int d\mathbf{x} A(\mathbf{x}) e^{-\mathcal{H}(\mathbf{x})/kT}}{\int d\mathbf{x} e^{-\mathcal{H}(\mathbf{x})/kT}} \approx \sum_i A(\mathbf{x}_i) . \quad (17)$$

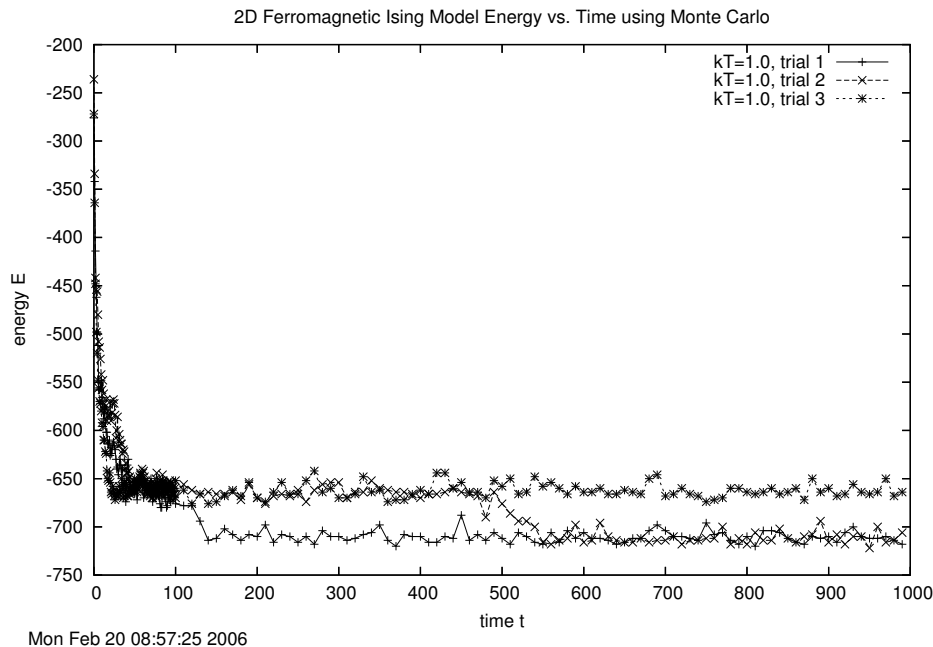
Thus, we find the average simply by summing the quantity  $A$  (e.g., the magnetization) over our representative set of  $\mathbf{x}_i$ ’s (e.g., spin configurations). At high temperatures, there will be a large spread of configurations contributing while at low temperature, states close to the ground state will increasingly dominate (i.e., appear more times in the set of  $\mathbf{x}_i$ ’s).

- **Ensuring representative configurations.** When carrying out the sum over configurations, we want to use ones that are actually representative. Three things we need to do:
  1. Throw out configurations from the beginning until the system has “equilibrated.” How might we tell when this has happened?
  2. Don’t use each configuration resulting from a MCS (“Monte Carlo Step”) but skip every  $n$  configurations to avoid correlations. Basically,  $n$  should be enough MCS’s so that the next configuration used has lost its memory of the previous one used. You can experimentally determine an appropriate “ $n$ ” by looking at an autocorrelation function. We’ll see how this works in Session 14.
  3. Adjust the “step size” between successive  $\mathbf{x}$ ’s so that the Metropolis success rate is reasonable. A rate around 50% is the canonical trade-off between computational efficiency and exploring enough of the configuration space.
- **Phase transition.** The optimized Ising model code with a ferromagnetic interaction has a phase transition in two dimensions. There is a “Curie temperature” below which the magnetization is nonzero. Can you see evidence of this phase transition from



running `ising_opt`? In the “Phase Transition” section of Session 13, you are asked to calculate the absolute value of the magnetization, which is most usefully defined for this problem as the average spin *per site* (i.e.,  $M = \frac{1}{N} \sum_{i=0}^{N-1} S_i$  with  $S_i = \pm 1$  and  $N$  is the number of lattice sites). To study the phase transition, you should look at the time *average* of this quantity, which means you average the result over all the Monte Carlo steps (after equilibration). If you are well below the phase transition, the spins will be either all up or all down, so you should find  $M = \pm 1$ . The reason to take the absolute value is that, for small systems (like the ones with  $L \leq 20$ ), in the course of a trial you could go from  $M = +1$  to  $M = -1$ . In this case, the average would be less than 1, which does not correctly reflect the physics. This problem is fixed by taking the absolute value. As a function of temperature, you should see the transition from disorder (small average magnetization) to order ( $\langle |M| \rangle = 1$ ) as the temperature is lowered.

- **More Explanation of Metropolis, etc.** The excerpt from the Binder/Heermann book (and the rest of the book) has good suggestions for the practical implementation of Monte Carlo approaches. You are strongly encouraged to look at the notes by Morten Hjorth-Jensen, which are linked on the 6810 webpage under “Supplementary Readings.”
  - Chapter 9 is “Random walks and the Metropolis algorithm.” Hjorth-Jensen explains how a Markov process (which is the basis of the Metropolis algorithm) is actually a discretized version of the diffusion equation. There is a discussion of entropy and equilibrium and ergodicity in Sect. 9.4 and a clear and concise discussion of the Metropolis algorithm in Sect. 9.5.
  - Chapter 10 is “Monte Carlo methods in statistical physics,” which has a review of the important thermodynamics as well as an explicit and detailed discussion of the Ising model. (See Fig. 10.2 for a graph of the absolute value of the average magnetization per spin vs. temperature for different size lattices.)
  - Chapter 11 is “Quantum Monte Carlo methods,” which includes a thorough discussion of variational Monte Carlo.
- **Equilibration and Cooling.** Consider the 2D Ising model with ferromagnetic  $J$ . On the next page is an example of what you might find from running the `ising_model.cpp` code with  $kT = 1$  three times. We can see a rapid decrease in energy at small times  $t$ , which reaches a plateau by  $t = 50$  about which successive configurations fluctuate. This is the equilibration or thermalization time (so these are the configurations we should throw out). But we see later that in one trial the energy drops to a lower plateau starting around  $t = 100$  and another drops to the same lower plateau starting



around  $t = 500$ . We might imagine the third trial will eventually drop down to the lower energy, but we can't predict when. This problem of equilibrating to a higher energy (and there may be many other possibilities if we run for a long time) is *not* fixed by skipping a larger number of configurations at the beginning. (What would you pick? Up to  $t = 100, 500, 2000$ ?) Rather, we give the system a chance to settle into the lowest energy by *cooling*. That is, we start the simulation at a higher temperature with a random configuration and equilibrate at that temperature. Then we use the final configuration as the starting configuration (rather than starting again with a random configuration) for a somewhat lower temperature until it equilibrates. And so on, until we get to the desired temperature. How much to lower the temperature at each stage depends on the problem. (We'll experiment with this rate in a slightly different context in Session 15 with simulated annealing.)

## 5 The Autocorrelation Function

[This discussion is based on Pang, *Computational Physics*.]

When we apply the Metropolis algorithm to generate a set of configurations that we wish to use to evaluate a physical quantity  $A$  (e.g., the magnetization of a spin system), we do not use all of them. We skip some at the beginning to eliminate the effects of the initial configuration (“thermalization” or “equilibration”) but we also don't want to use consecutive

configurations after that because they are likely to be highly correlated. So we skip some configurations before keeping another one. But how many should we skip? To find out, we analyze the *autocorrelation function* of  $A$ .

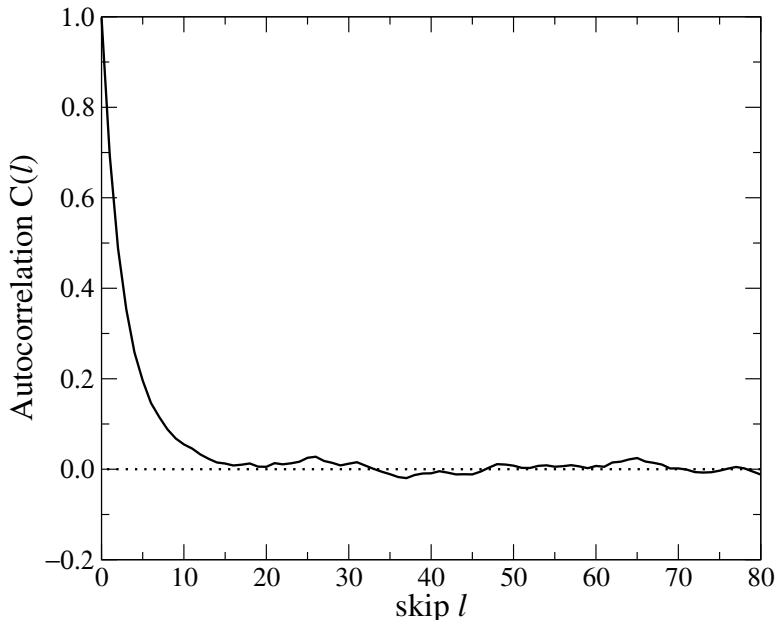
The autocorrelation function  $C(l)$  is

$$C(l) = \frac{\langle A_{n+l}A_n \rangle - \langle A_n \rangle^2}{\langle A_n^2 \rangle - \langle A_n \rangle^2} \quad (18)$$

with

$$\langle A_n \rangle = \frac{1}{N_{\text{spins}}} \sum_{n=1}^{N_{\text{spins}}} A_n \quad \text{and} \quad \langle A_{n+l}A_n \rangle = \frac{1}{N_{\text{spins}}} \sum_{n=1}^{N_{\text{spins}}} A_{n+l}A_n, \quad (19)$$

where  $A_n$  is the value of  $A$  in the  $n^{\text{th}}$  configuration. (Note that  $\langle A_n^2 \rangle$  is just  $\langle A_{n+l}A_n \rangle$  with  $l = 0$ .) Because  $n$  labels the Monte Carlo step, this correlation is effectively in time (rather than space, for example). Thus, we can use  $C(l)$  to see whether the value at  $n+l$  is related to the value at  $n$ , averaged over  $n = 1$  to  $N_{\text{spins}}$  (strictly speaking, to  $N_{\text{spins}} - l$ ). It is clear that  $C(0) = 1$ , but what do we expect as  $n$  increases? If quantities  $A$  and  $B$  are *uncorrelated*, then the average of their product is the product of their averages:  $\langle AB \rangle = \langle A \rangle \langle B \rangle$ . So if the configurations become uncorrelated, then  $\langle A_{n+l}A_n \rangle \approx \langle A_{n+l} \rangle \langle A_n \rangle \approx \langle A_n \rangle^2$  (since it shouldn't matter whether we average starting at 0 or  $l$ ) and  $C(l) \rightarrow 0$ . Note that it will not go exactly to zero, but fluctuate with each Monte Carlo step about zero (see the figure below). The correlation time (how many configurations to skip when averaging) is roughly how long it takes to go from 1 to the fluctuation region.



In the figure, the correlation time is about 15.