Theory and Methods for the Ferromagnetic Ising Model

Jay Shen
Department of Physics
University of Chicago
Chicago, IL 60637
jshe@uchicago.edu

Mark Lee
Department of Statistics
University of Chicago
Chicago, IL 60637
markyl@uchicago.edu

Abstract

The Ising model is a historically important problem in statistical mechanics. Although it was originally proposed as a crude approximation of magnetic phenomena, it furnished statistical mechanics, and later probabilistic science as whole, with a new paradigm of graph-based modeling that would prove influential. Especially in the distilled, theoretical study of probabilistic graphical models (PGMs), which today is its own field, many methods developed for Ising models and spin-glasses—as well as the rich physical vocabulary of energies, entropy, and partition functions—have been repurposed and reinterpreted for application towards a wide variety of problems.

In this paper, we pay homage to the Ising model by applying modern computational solutions towards the original problem of magnetism. We discuss the theory of the Ising model from a physical perspective, and examine its correspondence with PGMs theory. We then evaluate two approaches to inference—Markov Chain Monte Carlo and belief propagation. Finally, we show the success of the Ising model at modeling some physical phenomena, and discuss its limitations.

1 Theory of the Ising Model

The prevailing physical theory explains magnetism as a consequence of the intrinsic spin of particles. This spin produces a magnetic dipole moment $\vec{\mu}$ proportional to the spin σ:

$$\vec{\mu} = \mu\sigma$$

If an external magnetic field $\vec{B}$ is present, the potential energy is given by the classical formula:

$$U_B = -\vec{\mu} \cdot \vec{B}$$

Handily, this also absorbs the splitting of atomic energy levels due to quantum mechanical phenomena.

There also exists a pairwise interaction between particles due to the mutual effects of their magnetic moments. These exchange effects have strength defined by constants J_{ij}, The potential energy of these pairwise interactions is then:

$$U_{ij} = -J_{ij}\vec{\mu_i} \cdot \vec{\mu_j}$$

Here, we ignore quantum mechanical spin-spin coupling and exclusion energies.

Note that both potential energies are minimized when the moments align with $\vec{B}$ and with each other. This corresponds to the theory of ferromagnetism as a systematic alignment of spins.

Now, consider a collection of particles $\vec{\mu_i}$ in the prescence of an external magnetic field $\vec{B}$. The Hamiltonian, which specifies the total energy of the system, is defined by the sum of all pairwise and

Preprint. Under review.

unary energies:

$$E(\vec{\mu}) = -\frac{1}{2}\sum_i \sum_j J_{ij}\vec{\mu_i}\cdot\vec{\mu_j} - \sum_i \vec{\mu_i}\cdot\vec{B} \tag{1}$$

In most cases, working with this Hamiltonian is intractable. Luckily, we can simplify it by making several strategic assumptions.

First, in the context of an atomic lattice, all relevant particles are fermionic and have spin-$\frac{1}{2}$. We assume that all particles within an atom have identical spin, so the spin of each atom as a whole is fermionic. Then the magnetic moments simplify to $\vec{\mu_i} = \mu\sigma_i$, where $\sigma_i \in \{-1, 1\}$ is the atomic spin.

Second, we make a mean-field nearest-neighbor approximation so that the strength of pairwise interactions are negligible for non-adjacent pairs. We also assume all non-negligible J_{ij}'s are equivalent, that is that the material is organized in a regular lattice.

Coalescing constants, the Hamiltonian reduces nicely to:

$$E(\vec{\sigma}) = -J\sum_i \sum_{j\in adj(i)} \sigma_i\sigma_j - \sum_i \sigma_i B_i \tag{2}$$

The Boltzmann distribution gives a probability of some microstate $\vec{\sigma}$ at inverse temperature β:

$$P(\vec{\sigma}) = \frac{1}{Z}e^{-\beta E(\vec{\sigma})} \tag{3}$$

The physical theory of the Ising model has now been developed and we can turn to the purely abstract task of inference on a PGM.

2 Inference on Ising Models

In many cases, the Ising model of ferromagnetism is used to produce physical measurables like energy and magnetization. For example, Ernst Ising's original inquiry asked if the Ising model was capable of demonstrating phase transitions—changes in the magnetic state due to external conditions.

Now, these measurables require marginal distributions in order to compute, for example, the unary and pairwise energies. We will now examine two approaches to computing these marginals—Markov Chain Monte Carlo and belief propagation.

2.1 Markov Chain Monte Carlo

Monte Carlo algorithms produce better and better estimates by repeated sampling from the true distribution. When the true distribution is not immediately available for sampling, employing Markov Chain Monte Carlo defines an approximate distribution that hopefully, during sampling, approaching the true distribution.

In the context of the Ising model, the Markov states proceeding from some microstate of spins $\vec{\sigma}$ are created by flipping any spin σ_i in $\vec{\sigma}$. We can sample from these adjacent microstates by choosing a random σ_i to flip. Since we want to move towards the true, equilibrium distribution, we only transition to the new microstate $\vec{\sigma}'$ if $P(\vec{\sigma}') > P(\vec{\sigma})$. Using the formula for the Boltzmann distribution, this criterion is equivalent to $E(\vec{\sigma}') < E(\vec{\sigma})$.

In practice, we will compute the change in energy $\Delta E = E(\vec{\sigma}') - E(\vec{\sigma})$, which has a nice form:

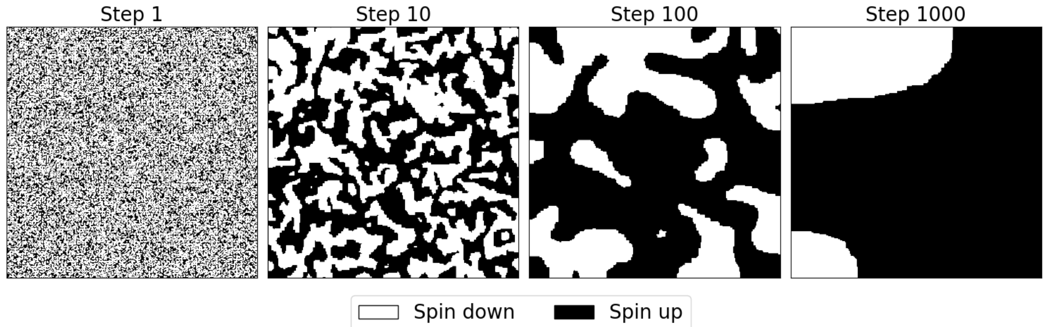
$$\Delta E = E(\vec{\sigma}') - E(\vec{\sigma}) = 2\sigma_i(J\sum_{j\in adj(i)} \sigma_j + B_i)$$

If $\Delta E < 0$, we transition states. If $\Delta E \geq 0$, we will transition microstates with probability defined by the Boltzmann factor:

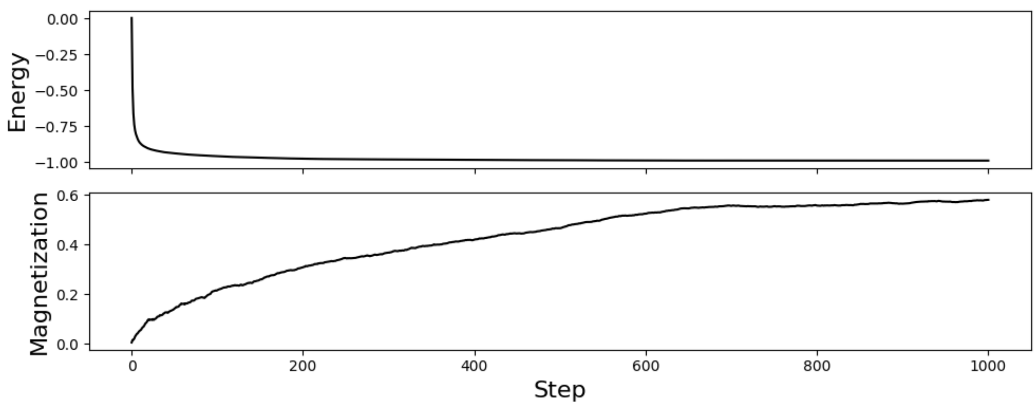
$$\frac{P(\vec{\sigma}')}{P(\vec{\sigma})} = e^{-\beta\Delta E}$$

This nicely models the phenomena of spin flips caused by field fluctuations.

We implemented this process in Python on square lattices of spins. For our sampling process, we iterated through all spins in random order, at each step evaluating the transition defined by flipping the present spin. We believe this best simulates the actual equilibration process.



(a) *Spin lattice sample microstates from various points in the sampling process*



(b) *Measurables of samples during the sampling process*

Figure 1: *MCMC quenching of an initial randomly generated* 256×256 *spin lattice sample with* $J = 0.5$*,* $\beta = 10$*, and a centered Gaussian magnetic field* $B_{ij} = 0.01 \exp[-\frac{1}{1024}((i-128)^2+(j-128)^2)]$.

In the MCMC quenching simulation conducted in 1, we observe several expected phenomena. The pairwise interactions are clearly exhibited by the clusters of aligned spins which form and coalesce as the number of steps grows. The prescence of the magnetic field also seems to magnetize the entire lattice, as indicated both by the convergence on full spin alignment in 1a, and the increasing magnetization measured in 1b.

Note that computing the measurables for 1a and 1b is easy given the explicit spin lattice microstates provided by the sampling process. The magnetization is simply the average spin. The energy is computed using 2.

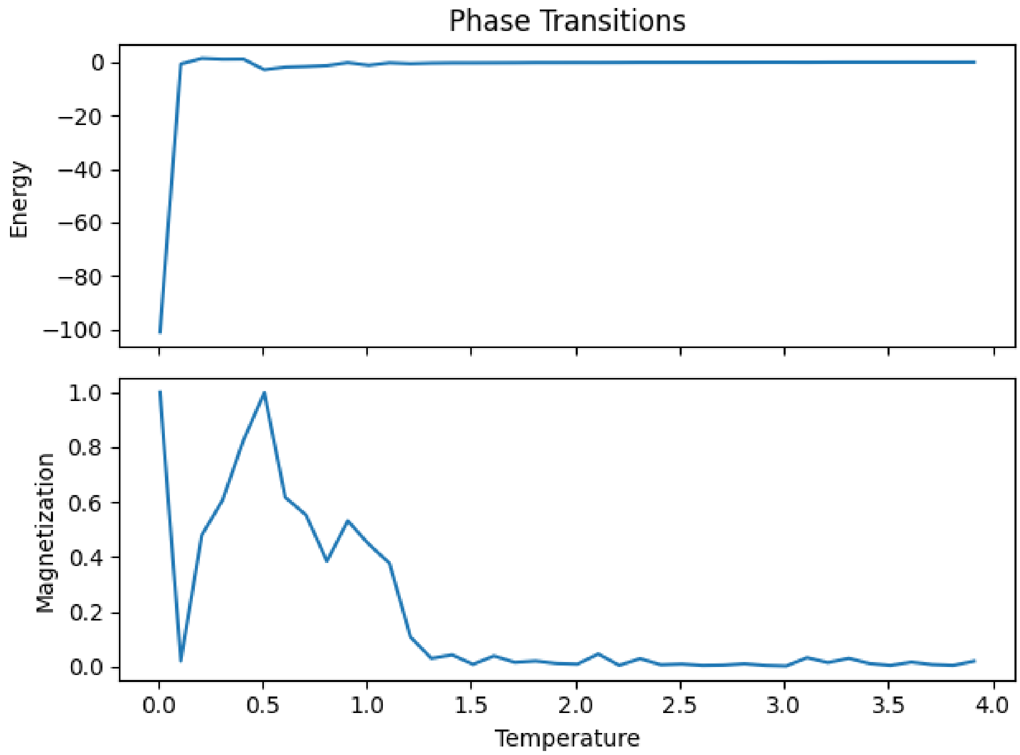
Using this MCMC inference scheme, we ran an experiment (Figure 1) simulating the phase transitions expected of the 2D Ising model.

2.2 Belief Propagation

Message-passing belief propagation algorithms are another way to estimate true marginals for general graphical models. They are derived from the theory of variable elimination, but are applicable to general graphs which may include cycles.

It turns out that belief propagation is especially effective and simple to implement on Ising models. This follows from the Boltzmann distribution that defines the joint distribution of the Ising model:

$$\tilde{P}(\vec{\sigma}) = \exp\Big[\beta J \sum_i \sum_{j \in adj(i)} \sigma_i \sigma_j + \beta \sum_i \sigma_i B_i\Big] \tag{4}$$



It factorizes nicely over the Ising lattice:

$$\tilde{P}(\vec{\sigma}) = \prod_i \exp\Big[\beta \sigma_i B_i\Big] \prod_{j \in adj(i)} \exp\Big[\beta J \sigma_i \sigma_j\Big] \tag{5}$$

This defines a factor graph of node and edge potentials representing the unary and pairwise potentials, respectively:

$$\phi_i = \exp\Big[\beta \sigma_i B_i\Big] \qquad\qquad \phi_{ij} = \exp\Big[\beta J \sigma_i \sigma_j\Big]$$

We then run a belief propagation routine on the factor graph according to the update scheme in Algorithm 1.

In Figure 2, we repeatedly run Algorithm 1 on the same system from Figure 1, until the maximum change in the messages falls below some ϵ. We also compute some measurables by taking samples from the marginal beliefs at each step and computing their mean energy and magnetization.

After calibration, the approximate marginals visualized in Figure 2a reflect what we would expect, full system spin alignment by the magnetic field. The trend in the magnetization is also observable in the plots of sample measurables which also reflects the appropriate minimization of energy.

2.3 Comparing MCMC and belief propagation

There are a number of considerations to keep in mind when deciding how to do inference on Ising models.

The benefits of MCMC inference lie in the easy access to valid microstate samples throughout the estimation process, which makes the computation of meaningful measurables easy. This is a direct consequence of the simulatory nature of MCMC which emulates, as closely as possible, the physical processes that underlie our desired marginal distributions.

Unfortunately, MCMC is slow, and very difficult to parallelize. This slowness is compounded by its tendency to follow unoptimal branches in the search tree and get stuck in local minima.

Algorithm 1 Belief propagation calibration step on factor graph

Input

Factor graph: $G(V = \{v_i\}_i \cup \{v_{ij}\}_{i,j}, E)$
Unary Factors: $\{\phi_i(\sigma_i)\}_i$
Pairwise Factors: $\{\phi_{ij}(\sigma_i, \sigma_j)\}_{i,j}$

Output:

Unary beliefs: $\{\beta_i(\sigma_i)\}_i$
Pairwise beliefs: $\{\beta_{ij}(\sigma_i, \sigma_j)\}_{i,j}$

// Pass messages from unary nodes
for unary $v_i \in V$ **do**
  for pairwise $v_{ij} \in Nb(v_i)$ **do**
    $\delta[v_i \to v_{ij}] = \phi_i \prod_{k \neq j} \delta[v_{ik} \to v_i]$ // Nothing to sum!
  end for
end for
// Pass messages from pairwise nodes
for pairwise $v_{ij} \in V$ **do**
  for $\{v_i, v_j\} = Nb(v_{ij})$ **do**
    $\delta[v_{ij} \to v_i] = \sum_{v_j} \phi_{ij} \delta[v_j \to v_{ij}]$
    $\delta[v_{ij} \to v_j] = \sum_{v_i} \phi_{ij} \delta[v_i \to v_{ij}]$
  end for
end for
// Compute beliefs
for unary $v_i \in V$ **do**
  $\beta_i = \phi_i \prod_j \delta[v_{ij} \to v_i]$
end for
for pairwise $v_{ij} \in V$ **do**
  $\beta_{ij} = \phi_{ij} \delta[v_i \to v_{ij}] \delta[v_j \to v_{ij}]$
end for
return $\{\beta_i\}_i, \{\beta_{ij}\}_{ij}$

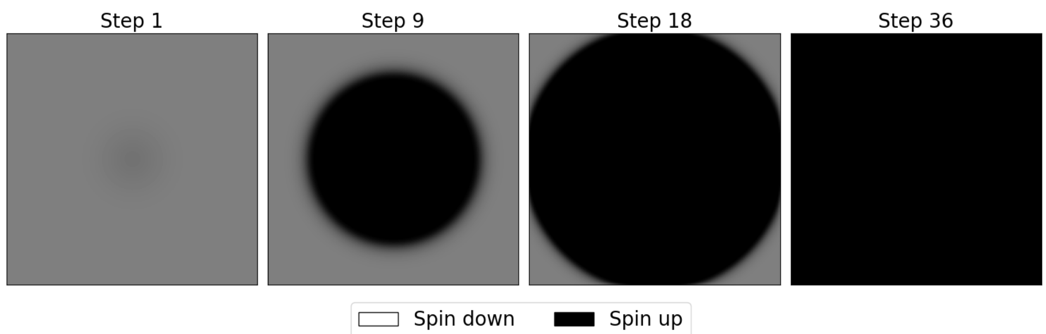
Belief propagation has its upsides in the immediate access to highly-accurate marginal beliefs upon calibration. Belief propagation is also fast. It is highly parallelizable, especially on factor graphs, and inherently bypasses many of the search tree inefficiencies that Monte Carlo methods struggle with.

The downsides of belief propagation lie in the poor interpretability of intermediate steps. Since the marginal beliefs do not necessarily agree with each other before calibration, they do not correspond to physically meaningful microstates upon sampling. As a consequence, it is difficult to take measurements of the system without sampling. Computing the partition function derivatives may be one solution, but we leave that problem to future studies.

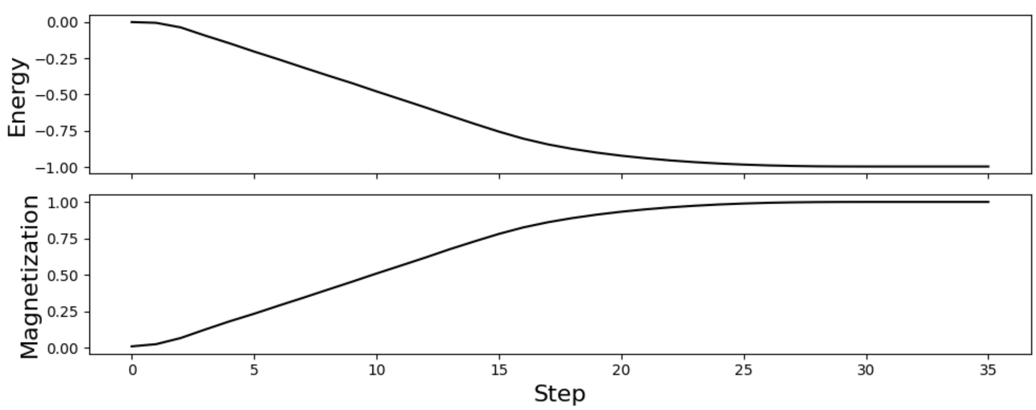
For most physical use cases, we conclude that MCMC is the better approach to inference. The interpretability and ease of measurements when dealing with explicit microstates is extremely valuable, especially in the context of magnetism. As we showed, simulating and analyzing physical phenomena such as phase transitions are straightforward and can be very precise.

3 Conclusions

In this paper we have



(a) *Spin lattice marginal beliefs at various points in the calibration process.*



(b) *Mean value of measurables taken from samples from marginals*

Figure 2: *Belief propagation to estimate marginal distributions of spin lattice with* $J = 0.5$, $\beta = 10$, *and a centered Gaussian magnetic field* $B_{ij} = 0.01 \exp[-\frac{1}{1024}((i-128)^2 + (j-128)^2)]$. *We run belief propagation until the difference between messages falls below* $\epsilon = 0.001$.