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 [1]. 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 diverse problems [2].

In this paper, we pay homage to the Ising model by applying modern computational solutions to 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—and discuss their strengths and weaknesses when appplied to predict physical phenomena.

1 Theory of the Ising Model

The prevailing physical theory explains magnetism as a consequence of the intrinsic spin all particles possess [3]. 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:

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

There also exist pairwise exchange interactions between particles due to the mutual effects of their magnetic moments. These exchange effects have strengths defined by constants J_{ij}. The potential energy contained by these pairwise interactions is:

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

For simplicity, we will ignore other, more complicated phenomena.

Now, consider a collection of particles $\vec{\mu} = (\vec{\mu_1}, \ldots, \vec{\mu_n})$ in the prescence of an external magnetic field $\vec{B}$. Note that both potential energies are minimized when the moments align with $\vec{B}$ and with each other. This corresponds to the explanation of ferromagnetism as a systematic alignment of spins. The Hamiltonian, which specifies the total energy of the system, is defined by the sum of all these

Preprint. Under review.

pairwise and unary energies:

$$E(\vec{\mu}) = -\frac{1}{2}\sum_{i}^{n}\sum_{j}^{n} J_{ij}\vec{\mu_i}\cdot\vec{\mu_j} - \sum_{i}^{n}\vec{\mu_i}\cdot\vec{B} \tag{1}$$

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

First, in the context of an atomic lattice, all sites are assumed to be identical and representable by a single fermionic spin variable. That is, each site's magnetic moment simplifies as $\vec{\mu_i} = \mu\sigma_i$, where the spin $\sigma_i \in \{-1, 1\}$.

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

Coalescing constants, the Hamiltonian reduces nicely to:

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

The Boltzmann distribution then gives the probability of some microstate of spins $\vec{\sigma} = (\sigma_1, \ldots, \sigma_k)$ 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 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.

Now, these measurables require marginal distributions in order to compute. For example, the quantities we are most interested in, the energy 2 and the magnetization $\mathbb{E}[\sigma]$, both require explicit microstates to compute. We will now examine two approaches to estimating 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. When the true distribution is not immediately available for sampling, constructing an appropriate Markov Chain is one way to approximate it.

In the context of the Ising model, the Markov transitions available from some microstate of spins $\vec{\sigma}$ are those created by flipping any spin σ_i in $\vec{\sigma}$. We can sample from these available transitions, then, by choosing a random σ_i to flip. Since we want to move towards the true, equilibrium distribution, we only transition to the new sampled 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:

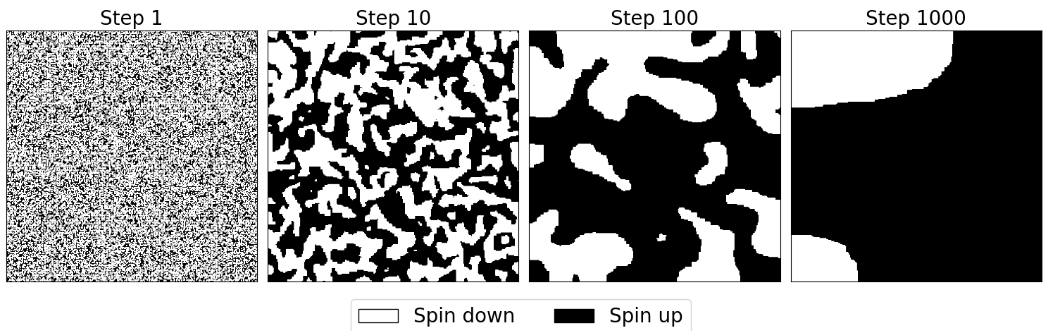
$$\Delta E = E(\vec{\sigma}') - E(\vec{\sigma}) = 2\sigma_i(J\sum_{j\in adj(i)}\sigma_j + B_i)$$

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

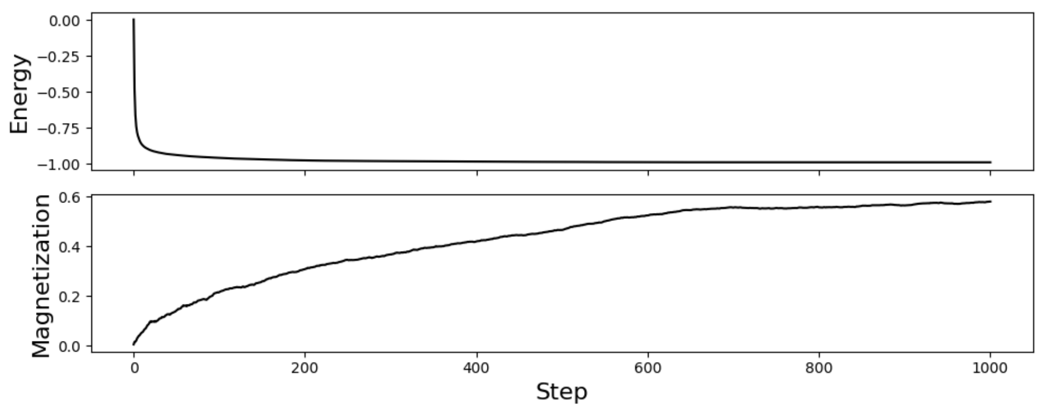
$$\frac{P(\vec{\sigma}')}{P(\vec{\sigma})} = e^{-\beta\Delta E}$$

Physically, 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 that spin. We believe this is a good trade off between performance and an accurate simulation of the physical equilibration.



(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 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.

Using this MCMC inference scheme, we ran an experiment (Figure 2) simulating the phase transitions expected of the 2D Ising model [4].

As seen, the energy and magnetization of the lattice both degenerate to zero as the temperature increases. The physical interpretation of this is that, at high temperatures, the field fluctuations which cause spin flips become more and more likely, and thus the alignment of spins is constantly being disrupted. It has been shown analytically that the Ising model in 2D exhibits this phenomenon of phase transitions. This simulation reinforces that claim. We can also identify what is called the Curie temperature, the point at which the phase transition occurs.

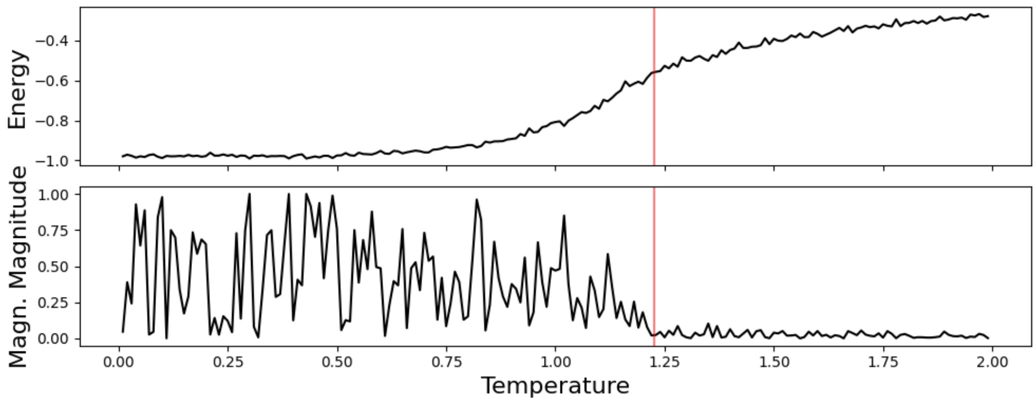


Figure 2: 100×100 *spin lattice run with* $J = 0.5$, $\vec{B} = 0$ *and* $T \in \{0.01, 0.02, \ldots, 2\}$. *Note that* $\beta = \frac{1}{T}$. *An hand-estimated Curie temperature is marked in red.*

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.

Belief propagation is especially simple to implement on Ising models. This follows from the Boltzmann distribution that defines the joint distribution:

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

It factorizes nicely over the graph representing 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:

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

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

We tested this on the same system from 1. Algorithm 1 is run repeatedly until the maximum change in the messages falls below some ϵ. We also take pseudo-measurements by sampling from the marginal beliefs at each step and computing their mean energy and magnetization.

After calibration, the approximate marginals visualized in Figure 3a 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 pseudo-measurables, which reflect the appropriate minimization of energy.

2.3 Comparing MCMC and belief propagation

When deciding which paradigm of inference to use for the Ising model, there are a number of considerations with respect to speed, output, interpretability, etc. Here, we will discuss these factors with respect to MCMC and belief propagation.

MCMC is useful because it provides easy access to valid microstate samples throughout the estimation process. Importantly, this makes the computation of meaningful measurables easy and intuitive. This is a direct consequence of the simulatory nature of MCMC, which emulates, as closely as possible, the underlying physical processes.

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 $\{\beta_i\}_i, \{\beta_{ij}\}_{ij}$

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.

Belief propagation, on the other hand, is useful because of 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.

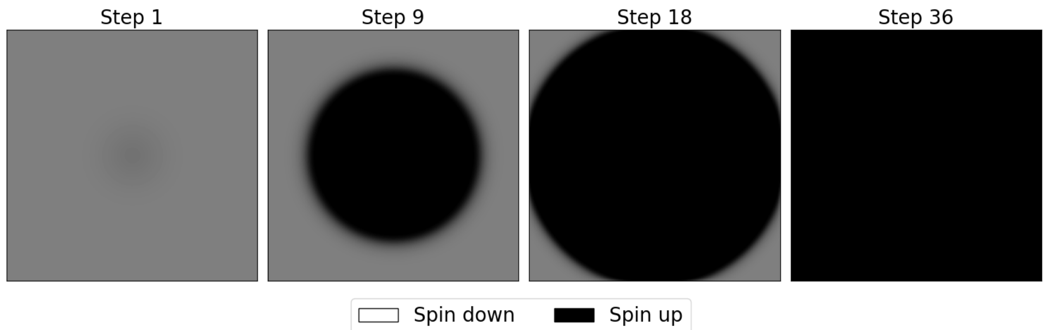
The downsides of belief propagation lie in the poor interpretability of intermediate steps. Since the marginal beliefs do not necessarily agree with each other until they are fully calibrated, they do not correspond to physically meaningful microstates when sampled. As a consequence, it is difficult to take measurements of the system. Deriving macrostate properties from the partition function 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 measurement is extremely valuable, and lends itself to the precise simulation and analysis of physical phenomena such as phase transition.

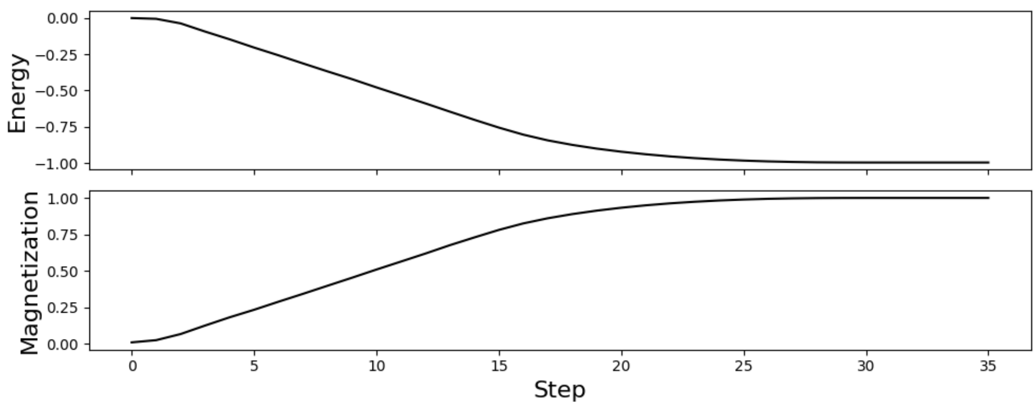
3 Conclusions

In this paper, we revisited the Ising model, an important probabilistic model of ferromagnetism, to pay homage and exercise modern computational methods. We discuss the physical theory and used it to motivate the PGMs methods applied. We evaluated two approaches, Markov Chain Monte Carlo and belief propagation, for marginal inference on the Ising model, and discussed their respective strengths and weaknesses.

We concluded MCMC was the better approach in the context of the Ising model because of its strong interpretability and ease of measurement. These strengths enables us to demonstrate the



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



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

Figure 3: *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* $\epsilon = 0.001$*.*

phenomenon of magnetic phase transitions in a highly emulative way. However, for faster inference that is interested in the true equilibrium beliefs of the system, belief propagation may prove more effective.

For the future, there is much to explore both with respect to the problem of magnetism as well as the Ising model. For example, deriving beliefs about macrostate measurables, which might involve the partition function, is an interesting and challenging task that might lessen the reliance on MCMC. There is also the problem of better simulation and analysis of phase transition phenomena, which might necessitate the adaptation of renormalization group theory to machine learning.

We conclude by affirming the power and usefulness of the Ising model in the context of physical as well as abstract probabilistic modeling. Within its rich theory, we can intuit much about modern probabilistic reasoning and its connection with the study of physical ensembles. Additionally, in the rich history of the model, in the story leading from ferromagnetism to deep neural networks, we can find solace in the fact that today's crude toy models might inspire tommorow's transhuman super-intelligence.

4 Supplementary Material

Our implementations of the above algorithms, using Python and the NumPy package, can be found at `https://github.com/jshe2304/ttic31180`.

References

[1] Wilhelm Lenz. Beiträge zum Verständnis der magnetischen Eigenschaften in festen Körpern. *Physikalische Zeitschrift*, 21:613–615, 1920.

[2] Daphne Koller. *Probabilistic Graphical Models*. The MIT Press, 2009.

[3] David Griffiths. *Introduction to Quantum Mechanics*. Cambridge University Press, 3 edition, 2018.

[4] Rudolf Peierls. On Ising's model of ferromagnetism. *Mathematical Proceedings of the Cambridge Philosophical Society*, 32, 1936.