Part I

Monte Carlo basics

1

Introduction

A quantum Monte Carlo method is simply a Monte Carlo method applied to a quantum problem. What distinguishes a quantum Monte Carlo method from a classical one is the initial effort necessary to represent the quantum problem in a form that is suitable for Monte Carlo simulation. It is in making this transformation that the quantum nature of the problem asserts itself not only through such obvious issues as the noncommutivity of the physical variables and the need to symmetrize or antisymmetrize the wave function, but also through less obvious issues such as the sign problem. Almost always, the transformation replaces the quantum degrees of freedom by classical ones, and it is to these classical degrees of freedom that the Monte Carlo method is actually applied. Succeeding chapters present and explain many of the quantum Monte Carlo methods being successfully used on a variety of quantum problems. In Chapters 1 and 2 we focus on discussing what the Monte Carlo method is and why it is useful.

1.1 The Monte Carlo method

The Monte Carlo method is not a specific technique but a general strategy for solving problems too complex to solve analytically or too intensive numerically to solve deterministically. Often a specific strategy incorporates several different Monte Carlo techniques. In what is likely the first journal article to use the phrase "Monte Carlo," Metropolis and Ulam (1949) discuss this strategy. To paraphrase them,

> The Monte Carlo method is an iterative stochastic procedure, consistent with a defining relation for some function, which allows an estimate of the function without completely determining it.

This is quite different from the colloquialism, "a method that uses random numbers." Let us examine the definition piece by piece. A key point will emerge.

Ulam and Metropolis were presenting the motivation and a general description of a statistical approach to the study of differential and integro-differential equations. These equations were their "defining relation for some function." The "function" was the solution of these equations. This function is of course unknown a priori. Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller (1953) a few years later would propose a statistical approach to the study of equilibrium statistical mechanics. The defining relation there was a thermodynamic average of a physical quantity over the Boltzmann distribution. The function was the physical quantity, and the unknown its average. The general description of the Monte Carlo method, given by Ulam and Metropolis and paraphrased by us, covers statistical mechanics applications plus many more applications than the original ones for differential and integro-differential equations.

The method is an "iterative stochastic procedure," meaning the procedure is applied over and over. Ostensibly, one purpose is to produce a large number of measurements on which a statistical analysis is made and validated by appealing to the law of large numbers. Certainly, this application is one use of the Monte Carlo method Ulam and Metropolis had in mind. For many other applications, the meaning of this phrase is more subtle. The movement from one step to the next depends on a stochastic procedure, creating a chain of events on which a statistical study is made. Usually this chain is a Markov chain.

Random numbers enter through the back door by saying it is a stochastic procedure. Besides being amazed at the speed with which the first computers did arithmetic, Ulam and Metropolis were also amazed that this arithmetic could generate "random numbers" uniformly distributed over the interval from zero to one, thereby enabling various schemes to sample from almost any distribution.

The defining character of the Monte Carlo strategy, however, comes from the rest of the statement: the Monte Carlo method estimates some function "without completely determining it." For Ulam and Metropolis this meant not having to obtain point-by-point values of functions obeying the differential and integro-differential equations. More broadly, presaging applications to problems in statistical mechanics, we paraphrased them by saying the Monte Carlo method makes estimates without needing to determine the function completely. Because it does not completely determine the function, the Monte Carlo method can thus provide only an estimate. As we will discuss, this estimate can have a high degree of certainty.

The Monte Carlo method is not the only method that does not need to determine a function completely. Deterministic methods also often do not need to do this. For example, a standard numerical problem is estimating the integral of a function of many variables. Deterministic algorithms use discretization schemes that estimate the integral by evaluating the integrand only at N points in each dimension. The computational effort of these methods hence scales as N^D where D is the number

of dimensions (variables). Thus the effort for high dimensional problems scales exponentially with dimension. As we will see, many Monte Carlo methods break "the curse of dimensionality" by needing to evaluate the function at a number of points that shows only a weak dependence on dimensionality. This is the key point.

The Monte Carlo method is a powerful but relatively simple idea. The seed of the idea came to Ulam over his frustration with the card game of Solitaire. At any intermediate point in the game, he was frustrated by the difficulty in calculating whether he would win or lose, but he realized that if he played the game often enough, he could reliably estimate the probability of winning. The speed with which the first electronic computer could do calculations prompted him to grow this idea into a new way to solve mathematical problems. Sharing his thoughts with von Neumann led them to develop a stochastic method to simulate the transport of neutrons through fissile material. This transport was naturally stochastic: defined by mean-free paths, the neutron has a specific probability of traveling a certain distance without a collision, and on collision, a scattering cross-section specifies probabilities for the collision to be elastic, inelastic, fissile, or absorptive. The first Monte Carlo method was a stochastic procedure, expressible on a computer, that mimicked what occurs in nature. It iteratively and stochasticly solved a Boltzmann-like transport problem without ever directly appealing to a Boltzmann equation.

The Metropolis-Ulam paper was written a couple of years after the first application of the method. It noted the general utility of stochastic procedures for solving problems, even those without stochastic analogs. Metropolis takes credit for the name "Monte Carlo" (Metropolis, 1985, 1987). It was not chosen to connect the method with the games of chance in this famous casino area of Monaco, but was Metropolis's way of ribbing Ulam, who had an uncle who occasionally had a "need" to go to the casino. The Monte Carlo method was seen as a procedure to which theorists would occasionally need to resort. It was perhaps unforeseen how frequently we have to resort to it.

1.2 Quantum Monte Carlo

Metropolis and Ulam (1949) credit Fermi with foreseeing that the Monte Carlo method would be useful for solving quantum problems. As they relate, Fermi noted that this new way of computing allows us to obtain the ground state of the time-independent Schrödinger equation

$$-\nabla^2 \psi\,(x, y, z) = [E - V(x, y, z)]\,\psi\,(x, y, z)$$

by introducing an imaginary-time dependence via the transformation

$$u\,(x, y, z, \tau) = \psi\,(x, y, z)\,e^{-E\tau},$$

so $u(x, y, z, \tau)$ obeys the diffusion-like equation

$$\frac{\partial u}{\partial \tau} = \nabla^2 u - Vu.$$

Already known was that this equation had a Monte Carlo expression as a collection of weighted particles, in which each independently performs a random walk while at the same time having its weight subjected to multiplication by the value of V at the point (x, y, z). In this type of Monte Carlo simulation the total weight of particles decays exponentially in (imaginary) time at a rate controlled by the eigenvalue E, and the spatial distribution of particles provides an estimate of $\psi(x, y, z)$, the solution of the time-independent Schrödinger equation. As we will show later, this eigenpair corresponds to the ground state solution. Indeed, Fermi's transformation is the starting point for all ground state quantum Monte Carlo methods. Computers, however, had to advance before this application of the Monte Carlo method became feasible for nontrivial quantum problems.

Fermi's observation is equivalent to transforming the time-dependent Schrödinger equation

$$i\frac{\partial \psi}{\partial t} = -\nabla^2 \psi + V\psi \tag{1.1}$$

from real time to imaginary time by the change of variables $it = \tau$:

$$\frac{\partial \psi}{\partial \tau} = \nabla^2 \psi - V\psi. \tag{1.2}$$

As Feynman later noted (Feynman and Hibbs, 1965), the same change of variables (with periodic boundary conditions in imaginary time) transforms the real-time path-integral formulation of quantum mechanics into finite-temperature quantum statistical mechanics. Finite-temperature quantum Monte Carlo simulations are typically performed in imaginary time. We will flesh out this observation later. We have a more modest, immediate task – fleshing out the Monte Carlo method and the need to resort to it.

1.3 Classical Monte Carlo

Because it is a convenient model to explain why we sometimes need Monte Carlo simulations, we start by discussing the one-dimensional Ising model in an external magnetic field. It is not a quantum model, but it is likely the simplest possible model with many interacting degrees of freedom.

The energy for the model is given by

$$E = -\sum_i \left(Js_i s_{i+1} + Hs_i\right), \tag{1.3}$$

where $J > 0$ is the (ferromagnetic) exchange integral, i identifies a lattice site, the on-site Ising variable s_i takes the values ± 1, and H is the external magnetic field. The lattice has N sites, and we assume periodic boundary conditions; that is, we assume $s_i = s_{i+N}$.

Two questions are typically asked of an interacting many-body problem: as N approaches the thermodynamic limit, what are its zero temperature properties? And what are its finite temperature properties? For the Ising model in an external magnetic field, the zero temperature state, that is, ground state, has an energy of $-N(J \pm H)$ with all Ising variables being $+1$ or -1, depending on the sign of H. In zero field, the ground state is doubly degenerate with the energy $-NJ$ being the lowest for spins all positive and all negative. Knowing the energy and the state is sufficient to determine all ground state properties. We thus focus on the finite temperature properties.

For finite temperatures, the partition function Z for the one-dimensional Ising model is known (Thompson, 1972):

$$Z = \lambda_+^N + \lambda_-^N, \tag{1.4}$$

where

$$\lambda_\pm = e^{J/kT}\cosh(H/kT) \pm \left[e^{2J/kT}\sinh^2(H/kT) + e^{-2J/kT}\right]^{1/2}, \tag{1.5}$$

and hence so is the free energy, $F = -kT \ln Z$. From the free energy all the equilibrium thermodynamics of the model follows by differentiation.

Typically, the finite-temperature property of interest is the equation of state, with emphasis on the behavior of the system near phase transitions. For a model of interacting spins, the appearance of magnetic long-range order is the candidate transition. Statistical mechanics teaches us that to study phase transitions we need to extrapolate the system to the thermodynamic limit and look for singular behavior in the free energy or in derived thermodynamic functions.

It is easy to take the thermodynamic limit of the present model. From (1.4) and (1.5), we find that the free energy per site is

$$F/N = -kT \ln \lambda_+,$$

and the magnetization per site is

$$M/N = -\frac{\partial}{\partial H}(F/N) = \sinh(H/kT)\left[\sinh^2(H/kT) + e^{-4J/kT}\right]^{-1/2}.$$

In zero field, we see that the magnetization is zero for all nonzero temperatures. This precludes the model from spontaneously magnetizing, a property that distinguishes the one-dimensional Ising model from Ising models in higher dimensions.

Being able to write an exact expression for the partition function is equivalent to saying the model is solved. Equations (1.4) and (1.5) were obtained by the transfer matrix technique (Thompson, 1972). While this technique also yields the exact solution of other models, for example, the zero-field two-dimensional Ising model, statistical mechanics is most often tasked with producing good approximations or good numerical simulations. We now address the basic challenges for the efficient Monte Carlo simulation of an Ising model.

For Monte Carlo methods, the most convenient approach to thermodynamics is the computation of averages from the Boltzmann probabilities for the possible configurations of the microscopic variables. If X represents a physical quantity of interest that is a function of the Ising variables, we can compute the temperature-dependent average (thermal expectation value) of X from

$$\langle X(T)\rangle = \frac{\sum_C X(C)e^{-E(C)/kT}}{\sum_C e^{-E(C)/kT}} = \sum_C X(C)\frac{e^{-E(C)/kT}}{Z}. \tag{1.6}$$

Here, C represents a configuration of the Ising variables; that is, $C = (s_1, s_2, \ldots, s_N)$, $X(C)$ and $E(C)$ are the values of X and the energy E for this configuration, and the summation is over the set of all possible configurations. The exponential $e^{-E(C)/kT}$ is the *Boltzmann factor* for configuration C, and the partition function $Z = \sum_C e^{-E(C)/kT}$ is its normalization constant. We are able thus to rewrite (1.6) as

$$\langle X(T)\rangle = \sum_C X(C)p(C), \tag{1.7}$$

where

$$p(C) = \frac{e^{-E(C)/kT}}{Z} \tag{1.8}$$

is the probability of configuration C and $\sum_C p(C) = 1$.

In an application, X might represent the energy

$$\langle E(T)\rangle = \sum_C E(C)p(C),$$

the magnetization

$$\langle M(T)\rangle = \sum_C \left(\sum_i s_i(C)\right)p(C),$$

or the spin-spin correlation function

$$\langle s_i s_j(T)\rangle = \sum_C s_i(C)s_j(C)p(C).$$

The critical tasks in computing an average this way are generating all configurations C and computing for each the probability $p(C)$. To do the latter, we have to evaluate the partition function, which requires the computation of all the Boltzmann factors. For our N-site model, the number of possible configurations is 2^N, so the complexity of this task grows exponentially fast, quickly making the computation of the partition function not practical. A way of providing a reliable solution to a problem without completely solving the problem is needed. We need the Monte Carlo method.

A Monte Carlo method for problems like the Ising model does not generate all configurations but instead only selects M representative configurations with the correct probability. It replaces (1.7) by

$$\langle X\rangle_{\mathrm{MC}} = \frac{1}{M}\sum_{i=1}^{M} X(C_i). \tag{1.9}$$

If all the measurements of $X(C_i)$ are statistically independent, then the Monte Carlo method also provides an estimate of the statistical error σ_X associated with this average,

$$\sigma_X \approx \sqrt{\frac{1}{M-1}\left(\frac{1}{M}\sum_{i=1}^{M} X^2(C_i) - \left(\frac{1}{M}\sum_{i=1}^{M} X(C_i)\right)^2\right)}. \tag{1.10}$$

Ideally, $M \ll 2^N$; that is, the average is estimated without completely solving the problem.

We now begin the process of defining specific Monte Carlo techniques that allow us to generate the configurations C_i appearing in (1.9). One basic question we need to answer is, how does the Monte Carlo method select the configurations with the correct probability if we cannot in a reasonable amount of computer time generate all the configurations necessary to compute the partition function?[1] As we shall see there are ways to select the configurations with the proper probability from an unnormalized probability density. The most famous way is called the *Metropolis algorithm*, the work by Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller we mentioned earlier. We discuss this algorithm in Section 2.5.1. After discussing this and multiple other Monte Carlo topics, we return to the Ising model in Chapters 4 and 5. In the latter chapter, we quantize it, and then develop our first quantum Monte Carlo algorithms for simulating quantum spin models.

[1] Several modern Monte Carlo methods developed for finite-temperature classical statistical mechanics enable the estimation of the logarithm of the partition function. The value of this estimate, however, is a product of the simulation and is not something needed to execute the simulation.

Overall, we focus on Markov chain Monte Carlo algorithms for quantum many-body models on a lattice. Not all Monte Carlo algorithms are Markov chain algorithms. Some just throw darts in a blindfolded way. The concept of a Markov chain is defined and discussed in Chapter 2. It produces the stochastic chain of events on which we perform a statistical analysis. Lattice models are models such as the Heisenberg model for interacting quantum spins and the Hubbard model for interacting electrons where spins and electrons exist only on lattice sites. Zero-temperature algorithms for interacting electrons and Bosons on a lattice are very similar to zero-temperature algorithms used for interacting electrons and Bosons in the continuum. For lattice models a greater variety of finite-temperature algorithms exists.

Suggested reading

N. Metropolis and S. Ulam, "The Monte Carlo method," *J. Am. Stat. Assoc.* **44**, 335 (1949).
N. Metropolis, "The beginning of the Monte Carlo method," *Los Alamos Science*, Special Issue, 1987.
R. Eckhardt, "Stan Ulam, John von Neumann, and the Monte Carlo method," *Los Alamos Science*, Special Edition, 1987.
J. E. Gubernatis, "The heritage," in *The Monte Carlo Method in the Physical Sciences: Celebrating the 50th Anniversary of the Metropolis Algorithm*, American Institute of Physics Conference Proceedings 690, ed. J. E. Gubernatis (Melville, NY: American Institute of Physics, 2003).