

Physics 127c: Statistical Mechanics

Quantum Monte Carlo

Monte Carlo methods are good for evaluating probabilistic integrals. A key feature of quantum mechanics is that we must deal with complex amplitudes rather than real positive probabilities. Intrinsically, Monte Carlo methods are not a “natural” match with quantum problems, and can only be successfully used where the quantum problem can be made to look like some classical problem. For example, the ground state wavefunction of a single particle or Bose system is everywhere positive, and so we can evaluate integrals over this wavefunction by Monte Carlo methods. This is a useful and much used approach. On the other hand excited states of these systems, and the ground state of Fermi systems necessarily have regions of positive and negative wave function. This makes the problem *much* harder, and dealing with the “sign problem” remains an active research area.

Variational Methods

Even the familiar variational method when used in a many body situation leaves multidimensional integrals that must be evaluated, so that Monte Carlo methods are necessary. An important example of this was McMillan’s calculation [1] of the condensate fraction in zero temperature superfluid He⁴. I will only give the briefest outline here to show where the Monte Carlo comes in, and refer you to the original for the algebraic details.

For noninteracting Bosons the ground state wavefunction is simply a constant. To take into account the interactions McMillan used the trial wavefunction consisting of a product of functions of the pair separations known as a *Jastrow wavefunction*

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \prod_{i < j} f(r_{ij}) \quad \text{with} \quad f(r) = \exp[-u(r)] \quad \text{and} \quad u(r) = \left(\frac{r_0}{r}\right)^p. \quad (1)$$

The function f is to take into account the pair correlations introduced by the interactions (particularly the hard core of the atoms). The symbol $\prod_{i < j}$ is the product over all pairs, counted once. The product can be organized as $\prod_{j=1, N; i=1, j}$, hence the notation. The constants r_0 and p are to be optimized by minimizing the ground state energy

$$E_0 \leq \frac{\int \Psi H \Psi d^{3N}r}{\int \Psi^2 d^{3N}r} = \frac{\int \sum_{i < j} \left[\frac{\hbar^2}{2m} \nabla_i^2 u(r_{ij}) + V(r_{ij}) \right] \Psi^2 d^{3N}r}{\int \Psi^2 d^{3N}r} \quad (2)$$

with $u(r_{ij})$ the pair potential. Note that Ψ is symmetric in the particle coordinates as required for a Bose wavefunction, and is real. Deriving the second form of the energy expression requires a few lines of algebra (see the original paper if you have trouble with this). Since

$$\Psi^2 = \exp \left[- \sum_{i < j} 2u(r_{ij}) \right] \quad (3)$$

this *looks* like the Boltzmann average with respect to a *classical* effective Hamiltonian $\beta H_{eff} = \sum_{i < j} 2u(r_{ij})$, so that the result can be evaluated by Monte Carlo methods.

The condensate density is given as

$$n_0 = \lim_{r \rightarrow \infty} N \frac{\int \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}_1 + \mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) d^3 r_2 \dots d^3 r_N}{\int \Psi^2 d^{3N} r} \quad (4)$$

$$= n \lim_{r \rightarrow \infty} \left\langle \prod_{j>1} \frac{f(|\mathbf{r}_{1j} + \mathbf{r}|)}{f(r_{1j})} \right\rangle \simeq n \left\langle \prod_{j>1} f(r_{1j}) \right\rangle \left\langle \prod_{j>1} \frac{1}{f(r_{1j})} \right\rangle \quad (5)$$

where n is the density N/Ω and the angular bracket denotes the average with respect to Ψ^2 , which again is evaluated by Monte Carlo methods. (The last expression arises since for $r \rightarrow \infty$ the average should factorize.)

McMillan found $n_0/n \simeq 0.11$ using systems of $N = 32$ and $N = 108$ particles in periodic boundary conditions. The limitation to a *very* small number of particles is because the calculation involves N dimensional integrals. Clearly such a small number of particles is not sufficient to show the condensate exists—note the expression requires taking $r \rightarrow \infty$. However *if* we believe the condensate exists from other arguments, the McMillan calculation provides an estimate of the fraction of particles involved. This estimate of about 10% has survived more modern calculations and also experiment, although experiment too involves many difficulties in implementation and interpretation, so that the results are not completely reliable.

Path Integral or the Trotter Method

The method of the previous section suffers the limitation of all variational methods—the accuracy of the answer is limited by the ingenuity of the trial wavefunction. The methods of the present section are in principle exact, i.e. the accuracy is limited by discretization and statistical errors, that can be reduced by using more powerful computers and longer runs.

The general expression needed to calculate finite temperature quantum averages is the matrix

$$\bar{\rho}_{\alpha\alpha'} = \langle \alpha | e^{-\beta H} | \alpha' \rangle \quad (6)$$

with $|\alpha\rangle$ some convenient choice of a complete set of states with a simple representation (e.g. given by defining each z -spin to be up or down in a spin- $\frac{1}{2}$ system, or position coordinates of each particle in a many particle system). The density matrix is $\bar{\rho}/\text{Tr}\bar{\rho}$ and the partition function is $Z = \text{Tr}\bar{\rho} = \sum_{\alpha} \langle \alpha | e^{-\beta H} | \alpha \rangle$. It is convenient to think of $e^{-\beta H}$ as the evolution in imaginary time.

Equation (6) is a nice formal expression, but we cannot evaluate the expression exactly because we cannot actually form the exponential of an operator, and we cannot list all the states for a large system. The trick to calculate $\bar{\rho}$ is to write this evolution as the sum of small steps

$$e^{-\beta H} = (e^{-\tau H})^n \quad \text{with} \quad \tau = \beta/n \quad (7)$$

with n large so that τ is small, and then using completeness $1 = \sum_{\beta} |\beta\rangle \langle \beta|$ to give

$$\bar{\rho}_{\alpha\alpha'} = \sum_{\beta, \gamma, \delta, \dots} \langle \alpha | e^{-\tau H} | \beta \rangle \langle \beta | e^{-\tau H} | \gamma \rangle \langle \gamma | e^{-\tau H} | \delta \rangle \dots \langle \dots | e^{-\tau H} | \alpha' \rangle. \quad (8)$$

Small τ corresponds to a large effective temperature, so that in the matrix elements $\langle \beta | e^{-\tau H} | \gamma \rangle$ etc. quantum effects will be weak. In particular if H is the sum of two noncommuting parts $H = H_1 + H_2$ then

$$e^{-\tau H} \simeq e^{-\tau H_1} e^{-\tau H_2} \quad (9)$$

since the corrections from the commutator $[H_1, H_2]$ is $O(\tau^2)$. Then also introducing the completeness relation between this product we have

$$\bar{\rho}_{\alpha\alpha'} \simeq \sum_{\beta, \beta', \gamma, \gamma', \delta, \delta', \dots} \langle \alpha | e^{-\tau H_1} | \beta \rangle \langle \beta | e^{-\tau H_2} | \beta' \rangle \langle \beta' | e^{-\tau H_1} | \gamma \rangle \dots \langle \dots | e^{-\tau H_2} | \alpha' \rangle. \quad (10)$$

Now if H_1 and H_2 separately consist of the sum of *commuting* terms, this converts the problem to one that in many cases can be reformulated as an effective *classical* problem in dimension $d + 1$ (with d the space dimension of the original problem) with the *imaginary time* τ giving the extra dimension. Averages of an operator A are given by

$$\langle A \rangle \frac{Tr(\bar{\rho}A)}{Tr(\bar{\rho})} = \frac{\sum_{\alpha, \alpha', \beta, \beta', \gamma, \gamma', \delta, \delta', \dots} \langle \alpha | e^{-\tau H_1} | \beta \rangle \langle \beta | e^{-\tau H_2} | \beta' \rangle \langle \beta' | e^{-\tau H_1} | \gamma \rangle \dots \langle \dots | e^{-\tau H_2} | \alpha' \rangle \langle \alpha' | A | \alpha \rangle}{\sum_{\alpha, \alpha', \beta, \beta', \gamma, \gamma', \delta, \delta', \dots} \langle \alpha | e^{-\tau H_1} | \beta \rangle \langle \beta | e^{-\tau H_2} | \beta' \rangle \langle \beta' | e^{-\tau H_1} | \gamma \rangle \dots \langle \dots | e^{-\tau H_2} | \alpha' \rangle \langle \alpha' | \alpha \rangle}, \quad (11)$$

which is the ratio of two averages weighted by $\bar{\rho}_{\alpha\alpha'}$. The method is to sample $\bar{\rho}_{\alpha\alpha'}$, Eq. (10) through Monte Carlo methods on its classical equivalent. This should become clearer by considering a couple of examples.

Quantum Spins

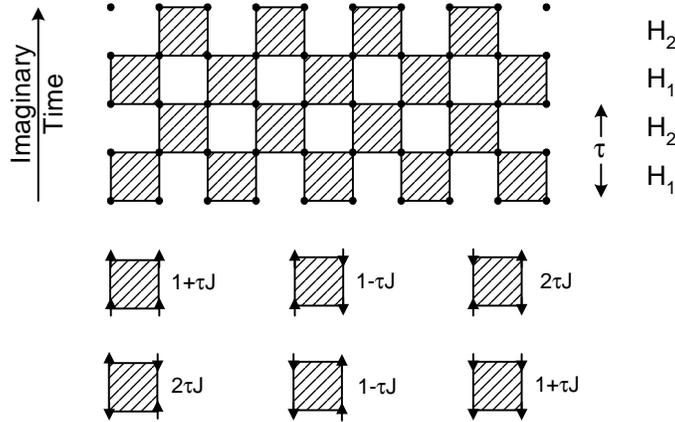


Figure 1: Shaded squares show interactions in the classical effective Hamiltonian H_{cl} . There is a spin at each solid dot, and space runs horizontally and imaginary time vertically. At successive imaginary time steps the interaction is between alternate spin pairs. The interaction squares with nonzero statistical weight $e^{-H_{cl}}$ and the value of this weight (given by the matrix elements of the quantum Hamiltonian) are shown below.

Consider for simplicity the one dimensional nearest neighbor spin- $\frac{1}{2}$ Heisenberg model with Hamiltonian

$$H = -J \sum_i \vec{\sigma}_i \cdot \vec{\sigma}_{i+1} = -J \sum_i (\sigma_i^z \sigma_{i+1}^z + 2\sigma_i^+ \sigma_{i+1}^- + 2\sigma_i^- \sigma_{i+1}^+) \quad (12)$$

with $\vec{\sigma}$ the Pauli spin matrices, and $\sigma^\pm = \frac{1}{2}(\sigma_x \pm i\sigma_y)$ the spin raising and lowering operators ($\sigma^+ |\downarrow\rangle = |\uparrow\rangle$, etc.). The complete set of states $|\alpha\rangle$ can be chosen as the product states given by the spin on each site being up \uparrow or down \downarrow . We write H as the sum over odd (H_1) and even (H_2) terms

$$H = -J \sum_{i \text{ odd}} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1} - J \sum_{i \text{ even}} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1}. \quad (13)$$

Each sum consists of commuting terms ($H_1 = -J(\vec{\sigma}_1 \cdot \vec{\sigma}_2 + \vec{\sigma}_3 \cdot \vec{\sigma}_4 + \dots)$ etc.) that can be treated separately. Thus $\langle \alpha | e^{-\tau H_1} | \beta \rangle$ is the product of pairwise terms involving the matrix elements of (expand the exponential for small τ)

$$\exp(\tau J \vec{\sigma}_1 \cdot \vec{\sigma}_2) \simeq \begin{bmatrix} 1 + \tau J & 0 & 0 & 0 \\ 0 & 1 - \tau J & 2\tau J & 0 \\ 0 & 2\tau J & 1 - \tau J & 0 \\ 0 & 0 & 0 & 1 + \tau J \end{bmatrix} \quad (14)$$

with respect to the basis states $s_1 s_2 = \uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow$. The contribution to Eq. (11) of a pair of spins over some imaginary time interval τ is interpreted as the Boltzmann factor $e^{-H_{cl}}$ for the plaquette given by the two spins in a $d + 1$ dimensional classical system at fictitious temperature 1 and with Hamiltonian H_{cl} where $H_{cl}(s_1, s_2, s'_1, s'_2)$ depends on the spin configurations at the beginning and the end of the time interval and is given by the corresponding matrix element of Eq. (14)

$$\exp[-H_{cl}(s_1, s_2, s'_1, s'_2)] = [\exp(\tau J \vec{\sigma}_1 \cdot \vec{\sigma}_2)]_{s_1 s_2, s'_1 s'_2}. \quad (15)$$

For example, $H_{cl}(\uparrow\downarrow\downarrow\uparrow) = -\ln(2\tau J)$. This gives us the *quantum* properties in terms of an effective classical problem in $1 + 1$ dimension, for example

$$Z = \text{Tr} \bar{\rho} = \sum_{\substack{\text{states} \\ \{s_{i,n}\}}} e^{-H_{cl}(\{s_{i,n}\})} \quad (16)$$

(see Fig. 1). The distribution $e^{-H_{cl}(\{s_{i,n}\})}$ is sampled by standard Monte Carlo methods.

Notice that the success of the method requires the matrix elements in Eq. (14) to be positive or zero, since they become the statistical weights of the classical problem. Thus not all quantum spin problems can be treated in this way. In fact the antiferromagnet would give negative weights for the off-diagonal elements, and so the method would apparently break down¹.

The expansions in small τ introduce errors into the method: these can be systematically reduced by increasing n (and the size of the computer!). There will also be statistical errors that can be reduced by sampling more configurations.

Interacting Particles

(See Ceperley and Pollock [2] for more details.) Here we use $|\alpha\rangle \rightarrow |\vec{x}\rangle$ where $\vec{x} = \vec{x}_1, \vec{x}_2, \dots, \vec{x}_N$ is the vector representing the position of the N particles, and

$$\bar{\rho}(\vec{x}, \vec{x}'; \beta) = \langle \vec{x} | e^{-\beta H} | \vec{x}' \rangle. \quad (17)$$

The Trotter decomposition tells us to look at $\bar{\rho}(\vec{x}, \vec{x}'; \tau) = \langle \vec{x} | e^{-\tau H} | \vec{x}' \rangle \simeq \langle \vec{x} | e^{-\tau H_1} e^{-\tau H_2} | \vec{x}' \rangle$ with $\tau = \beta/n$ small, and we split $H = H_1 + H_2$ with H_1 the kinetic energy, and H_2 the potential energy, and consider the small τ evolution under H_1 and H_2 separately.

The Hamiltonian H_1 is just that of free particles, and the imaginary time evolution is given by the modified Schrodinger equation

$$-\frac{\partial}{\partial \tau} \bar{\rho}(\vec{x}, \vec{x}'; \tau) = -\frac{\hbar^2}{2m} \nabla^2 \bar{\rho}(\vec{x}, \vec{x}'; \tau) \quad (18)$$

¹In fact on a bipartite lattice such as a simple cubic lattice a simple transformation in which the spins are rotated through π about the x-axis on one of the component lattices can be used: in the new basis the matrix elements are all positive. However for other lattices, such as a triangular lattice, this cannot be done.

with ∇^2 the N -particle Laplacian. This is just the *diffusion* equation, with solution

$$\bar{\rho}(\vec{x}, \vec{x}'; \tau) = \left(\frac{m}{2\pi\hbar^2\tau} \right)^{3N/2} \exp \left[-\frac{m}{2\hbar^2\tau} (\vec{x} - \vec{x}')^2 \right]. \quad (19)$$

The potential term (evolution under H_2) gives

$$\bar{\rho}(\vec{x}, \vec{x}'; \tau) = \langle \vec{x} | e^{-\tau V} | \vec{x}' \rangle = e^{-\tau V(\vec{x})} \delta(\vec{x} - \vec{x}'). \quad (20)$$

Note that at small τ the evolution under H_1 gives nonzero only for \vec{x} close to \vec{x}' . This means it does not matter which order we write H_1 and H_2 in the splitting of H , the basic idea of splitting the evolution into small τ steps.

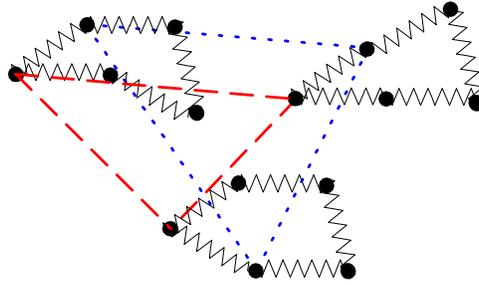


Figure 2: Classical system giving $\text{Tr} \bar{\rho}(\vec{x}, \vec{x}; \beta)$ for 3 quantum particles using $n = 5$. The dashed and dotted lines signify the interaction $V(\{\vec{x}^{(i)}\})$ depending on the position of the particles at “times” τ and 2τ (other interactions not shown).

Now we have

$$\begin{aligned} \bar{\rho}(\vec{x}, \vec{x}'; \beta) = & \text{const} \int \dots \int d^{3N} x^{(1)} d^{3N} x^{(2)} \dots d^{3N} x^{(n-1)} \\ & \times \exp \left(-\frac{m}{2\hbar^2\tau} (\vec{x}^{(1)} - \vec{x})^2 \right) \exp(-\tau V(\vec{x}^{(1)})) \exp \left(-\frac{m}{2\hbar^2\tau} (\vec{x}^{(2)} - \vec{x}^{(1)})^2 \right) \exp(-\tau V(\vec{x}^{(2)})) \\ & \times \dots \exp \left(-\frac{m}{2\hbar^2\tau} (\vec{x}' - \vec{x}^{(n-1)})^2 \right) \exp(-\tau V(\vec{x}')). \end{aligned} \quad (21)$$

We can again interpret this as the partition function of a classical system: N chains of n springs of spring constant $m/\hbar^2\tau^2$ with a potential $V(\vec{x}^{(j)})$ tethered between the end points \vec{x} and \vec{x}' at temperature τ .

For $\text{Tr} \bar{\rho}(\vec{x}, \vec{x}; \beta)$ the final position $\vec{x}' = \vec{x}$ and we integrate over all \vec{x} as well. This example is sketched in Fig. 2 for 3 quantum particles using a β -splitting of $n = 5$. Equation (21) is a discrete version of the *path integral* formulation of the statistical mechanics of interacting many body systems.

References

- [1] W. McMillan, Phys. Rev. A **138**, 442 (1965) (available [online](#))
- [2] D. M. Ceperley and E. L. Pollock, Phys. Rev. Lett. **56**, 351 (1986) (available [online](#))