Lecture 5 Monte-Carlo Methods and Diagrammatics

Objective for today:

- brief introduction to Monte-Carlo integration

- continous time QMC as an example method using diagrams.

· next time, will cover bold diagrammatic MC.

Disclaimer: there are many other useful MC techniques that I will not touch upon.

- Variational Monte-Carlo
- Diffusion Monte-Carlo
- ...

Monte-Carlo basics:

- generally, a method to evaluate high-dimensional integrals.
- integral is transformed into a summation over characteristic configurations that are generated from a Markov-chain process.
- expectation values in some distribution with weight / probability $p(\underline{x})$

$$\langle O\rangle_p = \frac{1}{Z} \underbrace{\int_C d\underline{x}}_{\text{integral over configuration space}} \; O(\underline{x})\; p(\underline{x})$$

with the partition function (total weight)

$$Z = \int_C d\underline{x} \; p(\underline{x})$$

A Monte-Carlo process will approximate the observable

$$\langle O\rangle_p \leftarrow \langle O\rangle_{MC} = \frac{1}{M}\sum_{i=1}^{M} O(x_i)$$

where configurations $\underline{x}_i$ are selected with probability $\frac{p(x_i)}{Z}$.

By virtue of the central limit theorem, in the limit of many samples, the error between our estimates and the exact value will scale via

$$(\Delta O)^2 = \langle (O_{MC} - O_p)^2\rangle = \frac{\operatorname{Var} O}{M}$$

i.e. $\Delta O \sim \frac{1}{\sqrt{M}}$

Generating Samples via a Markov-Chain

- transition-matrix $T_{\underline{x}y}$ between configurations, giving probability to go from $\underline{x}$ to y
- approximate the true distribution by a sequence of configurations drawn from any starting point $\underline{x} = \sigma_0$ in C.

The transition-matrix needs to satisfy

- $\sum_y T_{xy} = 1$: Normalization
- Ergodicity: any configuration $\underline{x}$ must be reachable from any other configuration y in a finite number of steps:

$$\forall x, y, \exists N < \infty \mid n \geq N : (T^n)_{x,y} \neq 0.$$

- Balance: probability to reach a certain point = probability to leave it.

$$\int_C d\underline{x} \; p(\underline{x}) \, T_{xy} = p(y)$$

(so: $p(x)$ is a left eigenvector of T_{xy})

In general, matrices satisfying balance are difficult to construct, unless requiring

- Detailed Balance:

$$\frac{T_{xy}}{T_{yx}} = \frac{p(y)}{p(x)} \quad \text{or} \quad \underbrace{p(x)\, T_{xy}}_{\text{probability to arrive at } y \text{ from } x} = \underbrace{p(y)\, T_{yx}}_{\text{prob. to go the other way.}}$$

Metropolis-Hastings algorithm

- particular scheme for constructing a set of moves satisfying detailed balance in two steps:

1) propose a given move to go from $\underline{x} \to \underline{y}$ with the a-priori probability

$$\mathcal{A}(\underline{x} \to \underline{y})$$

2) Accept or reject the move with an acceptance ~~probability~~ ~~rates~~ probability

$$\underline{B}(x \to y)$$

such that $\quad T_{xy} = \mathcal{A}(x \to y)\, \underline{B}(x \to y)$

Using detailed balance, we require:

$$p(x)\, \mathcal{A}(x \to y)\, B(x \to y) = p(y)\, \mathcal{A}(y \to x)\, B(y \to x)$$

so

$$\boxed{\frac{B(x \to y)}{B(y \to x)} = \frac{p(y)\, \mathcal{A}(y \to x)}{p(x)\, \mathcal{A}(x \to y)} \equiv R}$$

Now, Metropolis proposed ~~to~~ resolve this via

$$B(x \to y) = \min(1, R)$$

$$B(y \to x) = \min(1, 1/R).$$

i.e. R determines the probability that a given move should be accepted.

Reweighting

the exact distribution $p(x)$ may be difficult/inconvenient to calculate, and we can sample instead with a different probability distribution to estimate

$$\langle O \rangle = \frac{1}{Z} \int_C d\underline{x}\, O(\underline{x})\, p(x)$$

$$= \frac{\int_C d\underline{x}\, O(x) \frac{p(x)}{\rho(x)} \rho(x)}{\int_C dx \frac{p(x)}{\rho(x)} \rho(x)}$$

$$= \frac{\langle O \cdot \frac{p}{\rho} \rangle_\rho}{\langle p/\rho \rangle_\rho}$$

(Care needs to be taken when estimating the errors on the ratio of these averages, as they are drawn from the same samples of configurations...
$\Rightarrow$ jackknife /bootstrap procedure, see e.g. Numerical Recipes)

Quantum Field Theory with Monte Carlo:

- proposal to evaluate the partition function of a quantum problem directly! (at finite T)

$$Z = \mathrm{Tr}\left\{ e^{-\beta(\hat{H} - \mu \hat{N})} \right\}$$

which can be expressed in the interaction representation of a solvable piece $\hat{H}_0$ of the Hamiltonian

with $H = H_0 + H_1$

$$Z = \mathrm{Tr}\left\{ e^{-\beta(\hat{H}_0 - \mu N)} \; \mathcal{U}(0, \hbar\beta) \right\}$$

$$= \sum_{n=0}^{\infty} \left(-\frac{1}{\hbar}\right)^n \int_0^\beta d\tau_1 \int_{\tau_1}^\beta d\tau_2 \cdots \int_{\tau_{n-1}}^\beta d\tau_n \; \underbrace{\mathrm{Tr}\left\{ e^{-\beta(H_0 - \mu N)} H_1(\tau_n) H_1(\tau_{n-1}) \cdots H_1(\tau_1) \right\}}_{\to \text{ Feynman diagrams via Wick's theorem.}}$$

$$= \underbrace{\sum_{n=0}^{\infty}}_{\text{diagrams order}} \underbrace{\sum_{\gamma}}_{\text{topology of diagram.}} \int d\tau_1 \ldots \int d\tau_n \; W(n, \underline{\tau}, \gamma)$$

Let's call a configuration the ensemble of parameters that classify order, topology and internal vertices of a diagram.

$$\sigma = (n, \gamma, \underline{\tau}, \ldots)$$

The weight or probability of a given configuration is

$$p(\sigma) = W(\sigma) \quad d\tau_1 \, d\tau_2 \cdots d\tau_n.$$

which we assume to be positive for the moment.

Why continuous time is possible:

Integration over internal momenta proceeds as in a usual MC integration, however, we have a new class of moves which change the diagram order n.

Typically, this move will be of the following form (not thinking about phase)

$$\sigma_n = (n, \vec{\tau}) = (n, \tau_1, \dots, \tau_n)$$

$$\Updownarrow$$

$$\sigma'_{n+1} = (n+1, \vec{\tau}\,') = (n+1, \tau'_1, \dots, \tau'_{n+1})$$

We can draw the time τ'_{n+1} of the new vertex uniformly in $[0, \hbar\beta]$, so the ~~prop~~ a-priori possibility to propose the move is

$$A(\sigma_n \to \sigma'_{n+1}) = \frac{d\tau}{\hbar\beta}$$

for the reverse move of eliminating one vertex, we can just select one vertex at random

$$A(\sigma'_{n+1} \to \sigma_n) = \frac{1}{k+1}$$

Hence, the acceptance ratio becomes

$$R_{(n, \vec{\tau}; n+1, \vec{\tau}\,')} = \frac{p(\sigma'_{n+1}) \quad A(\sigma'_{n+1} \to \sigma_n)}{p(\sigma_n) \quad A(\sigma_n \to \sigma'_{n+1})}$$

$$= \frac{W(\sigma'_{n+1})\, d\tau'_1 \cdots d\tau'_{n+1}}{W(\sigma_n)\, d\tau_1 \cdots d\tau_n} \; \frac{\frac{1}{k+1}}{d\tau / \hbar\beta}$$

$$= \frac{W(\sigma'_{n+1})}{W(\sigma_n)} \; \frac{\hbar\beta}{k+1} \Rightarrow \text{all factors of the infinitesimal } d\tau \text{ cancel.}$$

$\Rightarrow$ can use continuous τ.

Note on the sign problem:

for a generic fermionic theory, the diagram weight can become negative, so we cannot use $w(\sigma)d\sigma$ as a sampling probability. Instead, we may use a re-weighting technique and take the sampling probability to be $|w(\sigma)|$

$$Z_f = \sum_{\sigma \in C} \underbrace{\frac{w(\sigma)}{|w(\sigma)|}}_{\mathrm{sgn}(w)} |w(\sigma)| = \langle \mathrm{sgn}\, w(\sigma) \rangle_{|w|}$$

and we then have to sample observables via

$$\langle O \rangle = \frac{\langle O\, \mathrm{sgn}(w) \rangle_{|w|}}{\langle \mathrm{sgn}(w) \rangle_{|w|}}$$

The fermionic partition function has a numerically much smaller value than the corresponding bosonic problem with $w_B = |W|$, $Z_B = \int_C d\sigma\, w_B(\sigma)$. The ratio is related to the free energy difference ΔF.

$$\langle \mathrm{sgn}(W) \rangle = \frac{Z_f}{Z_B} = e^{-\beta \Delta F}$$

Hence, the calculation needs to be run until the variance is of order of the expectation-value.

$$\mathrm{Var}\ \mathrm{sgn}(w) = \langle \mathrm{sgn}(w)^2 \rangle - \langle \mathrm{sgn}(w) \rangle^2$$

$$= 1 - e^{-2\beta \Delta F} \simeq 1$$

and the relative error grows with decreasing T:

$$\Delta\, \mathrm{sgn}(w) = \frac{\frac{1}{\sqrt{M}} \sqrt{\mathrm{var}\ \mathrm{sgn}(w)}}{\langle \mathrm{sgn}(w) \rangle} = \frac{e^{\beta \Delta F}}{\sqrt{M}}$$

$\Rightarrow$ computer-time grows exponentially at low T.

Example application:

Single-impurity Anderson model: $H = H_{bath} + H_{hyb} + H_{U} + H_{imp}$.

$$H_{bath} = \sum_{k,\sigma} \epsilon_{k,\sigma} \hat{c}^{\dagger}_{k\sigma} \hat{c}_{k\sigma}$$

$$H_{imp} = \sum_{\sigma} E^{\sigma\sigma'} \hat{d}^{\dagger}_{\sigma} \hat{d}_{\sigma'} \quad ; \quad \text{here: } E^{\sigma\sigma'} = \varepsilon_0 \delta_{\sigma\sigma'}$$

$$H_{hyb} = \sum_{k\sigma} \left(V_k c^{\dagger}_{k\sigma} d_{\sigma} + h.c. \right)$$

$$H_U = U \hat{n}_{\uparrow} \hat{n}_{\downarrow} \qquad \hat{n}_{\sigma} = \hat{d}^{\dagger}_{\sigma} \hat{d}_{\sigma}$$

there is only one interaction-term H_U while all other parts are free.

The task is to calculate the impurity Green's function

$$G_d(\tau) = \langle T_\tau \, d(\tau) \, d^{\dagger}(0) \rangle$$

$$= \frac{1}{Z} \underset{\uparrow}{\mathrm{Tr}} \left\{ e^{-(\beta-\tau)H} d \, e^{-\tau H} d^{\dagger} \right\}$$

trace over bath + impurity D.O.F.

The bath can be integrated out analytically, given that fermions $\hat{c}^{(\dagger)}$ obey a free theory: This yields a renormalized propagator for the $\hat{d}$'s $\rightarrow$ $d=0$ problem!

$$S_0^{eff} = \int d\tau d\tau' \sum_{\sigma\sigma'} d^{\dagger}_{\sigma}(\tau) \Big[(\partial_\tau + E^{\sigma\sigma'}) \delta(\tau-\tau') + \Delta^{\sigma\sigma'}(\tau-\tau') \Big] d_{\sigma}(\tau')$$

where the hybridization-function derives from the coupling:

$$\Delta^{\sigma\sigma'}(i\omega_n) = \sum_{k\mu} V_k^{*\sigma\mu} (i\omega_n - \varepsilon_{k\mu})^{-1} V_k^{\mu\sigma'}$$

We thus have a bare Green's function for d-el is:

$$\left(\mathcal{G}_d^0\right)_{\sigma\sigma'} = -\left(\frac{1}{i\omega_n + \underline{E} + \underline{\Delta}}\right)_{\sigma\sigma'} \overset{\text{here}}{=} \frac{1}{i\omega_n - \epsilon_0 - \Delta_\sigma}\,\delta_{\sigma\sigma'}$$

In addition, we have the interaction term with the local Hamiltonian H_U, which we treat perturbatively:

$$H_{int}(\tau) = H_U(\tau) = U\,\hat{n}_\uparrow(\tau)\,\hat{n}_\downarrow(\tau)$$

At n^{th} order, we have to consider diagrams arising from

$$\mathrm{Tr}\left\{ e^{-\beta H_0}\; H_{int}(\tau_n) \cdots H_{int}(\tau_1) \right\}$$

$$= U^n \left\langle \left(\hat{n}_\uparrow(\tau_n)\,\hat{n}_\downarrow(\tau_n)\right)\left(\hat{n}_\uparrow(\tau_{n-1})\,\hat{n}_\downarrow(\tau_{n-1})\right) \cdots \left(\hat{n}_\uparrow(\tau_1)\,\hat{n}_\downarrow(\tau_1)\right)\right\rangle_0$$

$$= \left\langle d_\uparrow^\dagger(\tau_n)\,d_\uparrow(\tau_n)\;d_\downarrow^\dagger(\tau_n)\,d_\downarrow(\tau_n) \cdots d_\uparrow^\dagger(\tau_1)\,d_\uparrow(\tau_1)\,d_\downarrow^\dagger(\tau_1)\,d_\downarrow(\tau_1)\right\rangle_0$$

Wick's th.

$$= \det \mathcal{D}_n^\uparrow \times \det \mathcal{D}_n^\downarrow$$

using $\quad -\langle T\, d(\tau_i)\, d^\dagger(\tau_j)\rangle = \mathcal{G}_d^0(\tau_i - \tau_j)$

$$\Rightarrow \quad \left(\mathcal{D}_n^\sigma\right)_{ij} = \mathcal{G}_\sigma^0(\tau_i - \tau_j).$$

and the partition function has the form:

$$Z = Z_0 \sum_{n=0}^{\infty} \frac{(-U)^n}{\hbar^n n!} \int_0^\beta d\tau_1 \cdots \int_0^\beta d\tau_n \prod_\sigma \left(\det \mathcal{D}_n^\sigma\right).$$

and the weight-function is $\quad w(n, \vec{\tau}) = \dfrac{(-U)^n}{\hbar^n n!} \prod_\sigma \left(\det \mathcal{D}_n^\sigma\right)$

Note: the trivial sign-problem for $U>0$ can be managed by redefinitions.

E.g.
$$\left\{\begin{array}{l} d^\dagger_\downarrow \to \tilde{d}_\downarrow \\ d_\downarrow \to \tilde{d}^\dagger_\downarrow \end{array}\right\} \quad \text{p-h. conjugate } \downarrow\text{-spins.}$$

$$\Rightarrow \left\{\begin{array}{ll} \varepsilon_{0\downarrow} \to -\varepsilon_{0\downarrow} & \Delta_\downarrow(\tau) \to -\Delta_\downarrow(-\tau) \\ \varepsilon_{0\uparrow} \to \varepsilon_{0\uparrow} + U & \underline{\underline{U \to -U}} \end{array}\right.$$

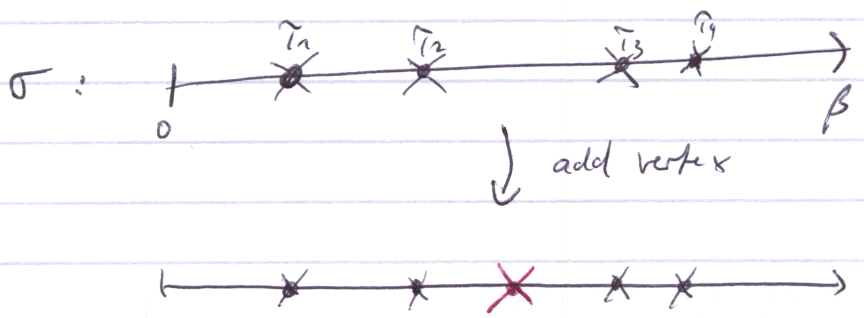
(Better approach to do this symmetrically)

Anyway, we can get a positive weight-function $\omega(n, \vec{\tau})$.

$\Rightarrow$ Updates can proceed as discussed previously

by inserting/removing a single vertex at τ_i

(topology does not occur, as we summed over all contractions of the correlator in form of a determinant!)
$\rightsquigarrow (d=0)$



with acceptance ratio $R = \frac{\beta U}{(k+1)} \prod_\sigma \frac{\det \mathcal{D}^\sigma_{n+1}}{\det \mathcal{D}^\sigma_n}$

$\uparrow$

can evaluate ratio efficiently!

Measurements:

ultimately, we want to measure observables, such as the propagator

$$G_\sigma(\tau-\tau') = -\frac{Z_0}{Z}\sum_{n=0}^{\infty}\frac{(-U)^n}{\hbar^n n!}\int d\tau_1\cdots d\tau_n \underbrace{\left\langle T_\tau\, d_\sigma(\tau)\, d_\sigma^\dagger(\tau')\, H_1(\tau_1)\cdots H_1(\tau_n)\right\rangle}_{g(n,\tau,\tau';\vec{\tau})}$$

while we sample a series (with the same summations and integrations) with probability / weight $w(n,\vec{\tau})$

the suitable estimator is

$$G_\sigma(\tau-\tau') \leftarrow \left\langle \frac{\langle g(n,\tau,\tau';\vec{\tau})\rangle_0}{\underbrace{\langle w(n,\vec{\tau})\rangle_0}_{\Sigma\text{ all Wick-contractions in the free theory}}}\right\rangle_{MC.}$$

MC: ~~sum~~ average over MC samples.

as previously for $w(n,\vec{\tau})$, the expression for $g(n,\tau,\tau';\vec{\tau})$ expands to a determinant of Green's fct$^{\text{s}}$.

Summary / Outlook:

- we've seen some of the basic ingredients for constructing continuous time / determinant Monte Carlo algorithms.
- there is a considerable body of work on these and similar algorithms. See
  - E. Gull et al. Rev. Mod. Phys. 83 (2011)
  - Philipp Werner: Continuous time impurity solvers. Notes (2011)