Monte Carlo method - introduction

- it is a stochastic method - we use random (or pseudorandom) numbers and quantities to simulate some physical system or to obtain average values of observables in many-particle systems

- there are various types of simulations:
  1) MC integration - a basic tool to calculate high-dimensional integrals, applicable in many areas of physics
  2) geometric MC - simulating particular systems using e.g. random walks (diffusion) (diffusion-limited aggregation, percolations ...)
  3) thermodynamic MC - e.g. Ising models etc.
  4) models of structures and there evolution
  5) kinetic MC

  etc.

- using MC method we can handle large numbers of particles $\sim 10^6$ and more

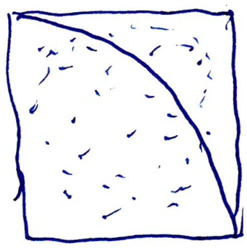
- usually it does not require large memory, only CPU time and a good random number generator

- because it is stochastic we obtain only approximations, it is not „ab initio" method

- it is relativly simple to code

- examples of MC simulations

1) approximation of π



the area of the quarter of the circle is $S = \frac{\pi}{4}$

and it can be approximated as

$$S \approx \frac{\text{number of point in } ◣}{\text{number of point in } \square}$$

if the probability of each point is the same

2) reaction $O_2^- + H_2 \to O_2^- + H_2$, $\to OH^- + H$, $\to e^- + H_2O$ at low energies

- to do full quantum calculation is very difficult

$\Rightarrow$ MC simulation - classical trajectories with random initial conditions according to the quantum initial states

$\Rightarrow$ $P_{ch}(E, b)$ - probabilities of ending in each channel depending on energy and impact parameter

$\Rightarrow$ $\sigma_{ch}(E) = 2\pi \int_0^{b_{max}} P(E,b)\, b\, db$ - cross sections for processes ending in channels

- recommended literature

- I. Nezbeda, J. Kolafa, M. Kotrla - Úvod do počítačových simulací, Praha 2003 (in Czech)

- D. Landau, K. Binder - A Guide to Monte Carlo Simulations in Statistical Physics, Cambridge 2002

- more will appear my webpage

Monte Carlo method - background

- application of MC method is usually based on determination of a certain quantity (e.g. integral) using averaging of stochastic (random) quantities X_i which are distributed according to some ~~probabitity~~ probability function
  - we will see that the error decreases with a number of averaged quantities n as $\frac{1}{\sqrt{n}}$ (at least very probably)
    - this follows from the central limit ~~theorema~~ theorem as explained below

- basic notions
  - probability function of random discrete variable X

    $P[X=x] = P(x)$ = probability that X has a value x

    clearly $P(x) \geq 0$ for all x and $\sum_x P(x) = 1$

  - cumulative distribution function of X (CDF)

    $$F(x) = P[X \leq x] = \sum_{t \leq x} P(t)$$

    or $P[x_1 \leq X \leq x_2] = F(x_2) - F(x_1)$

    it is monotonic increasing function $0 \leq F(x) \leq 1$

  - probability distribution of continuous random variable X

    described by probability density function (PDF) $\varrho(x) > 0$

    satisfying $\int_\Omega \varrho(x)\, dx = 1$ where Ω is a domain of X

    thus $P[x_1 \leq X \leq x_2] = \int_{x_1}^{x_2} \varrho(x)\, dx = F(x_2) - F(x_1)$

where $F(x)$ is again CDF of X

$$F(x) = P[X \le x] = \int_{-\infty}^{x} g(t)\,dt \quad \text{or} \quad g(x) = \frac{dF(x)}{dx}$$

Note: probability of finding an exact value x in the continous case is zero!

- <u>expected (mean) value</u> of a random variable X

$$E(X) = \sum_{i=1}^{n} x_i p_i \quad \text{(for discrete variables)}$$

$$= \int_{\Omega} x g(x)\,dx \quad \text{(for continuous variables)}$$

- <u>variance</u> (the square of the standard deviation)

$$\sigma^2(X) = D(X) = \mathrm{var}(X) =$$

$$= \sum_{i=1}^{n} [x_i - E(X)]^2 p_i = \sum_{i=1}^{n} x_i^2 p_i - E(X)^2$$

or

$$= \int_{\Omega} [x - E(X)]^2 g(x)\,dx = \int_{\Omega} x^2 g(x)\,dx - E(X)^2$$

$$= E(X^2) - E(X)^2$$

- basic relations are

$$E(X+c) = c + E(X) \qquad D(X+c) = D(X)$$

$$E(cX) = c\,E(X) \qquad D(cX) = c^2 D(X)$$

and for two random variables X and Y independent

$$E(X+Y) = E(X) + E(Y) \qquad D(X+Y) = D(X) + D(Y)$$

$$E(X \cdot Y) = E(X) \cdot E(Y)$$

e.g. $D(X+Y) = \iint [x+y-E(X+Y)]^2 g(x,y)\,dx\,dy \overset{\text{independent variables}}{=}$

$$= \iint (x - E(X) + y - E(y))^2 g_X(x)\, g_Y(y)\,dx\,dy =$$

$$= \iint [(x-E(X))^2 + 2\underbrace{(x-E(x))(y-E(y))}_{0} + (y-E(Y))^2]\, g_X\, g_Y\,dx\,dy =$$

$$= D(X) + D(Y)$$

- central limit theorem (CLT)

  - let us have n independent random variables

    $$X_1, X_2, \dots X_n$$

    with the same probability distribution

    i.e. all expected values $E(X_i) = m$ for all i

    and variance $D(X_i) = b^2$, for all i

    are the same

    then a new random varible

    $$S_n = X_1 + X_2 + \dots + X_n$$

    satisfies $E(S_n) = nm$ and $D(S_n) = nb^2$

  - moreover let us consider a normal distribution

    $N(\mu, \sigma^2)$ with $\mu = nm$ and $\sigma^2 = nb^2$ or $\sigma = \sqrt{n}\, b$

    with the PDF $g_n(x)$ (probable error of $S_n \sim \sqrt{n}$)

  - them the central limit theorem gives

    $$P[s_1 \leq S_n \leq s_2] \simeq \int_{s_1}^{s_2} g_n(x)\, dx$$

    for arbitrary s_1 and s_2 and sufficiently large n

  in other words: probability distribution of sums S_n of a large number of independent random variables is approximately normal

Note: there are generalizations for much weaker conditions - $X_1, \dots, X_n$ of various distributions and not necessarily independent

but no single X_i can be dominant

that is why normal distribution appears in nature if several random factors influences the system

- expected (mean) value of a function of a random variable

if $Y=f(X)$ is a random variable as a function of X

then $E(Y) = E(f(X)) = \int_{\Omega} f(x)\, g(x)\, dx \quad \left(\text{or } = \sum_{i=1}^{n} f(x_i)\, p_i \right)$

and generaly $E(f(X)) \neq f(E(X))$

- examples of probability distributions

1) uniform distribution on $\langle a, b \rangle$

PDF: $g(x) = \begin{cases} \frac{1}{b-a} & \text{on } \langle a,b \rangle \\ 0 & \text{elsewhere} \end{cases}$, CDF: $F(X) = \frac{x-a}{b-a}$ on $\langle a,b \rangle$

$$E(x) = \frac{a+b}{2} \quad , D(X) = \sigma^2(X) = \frac{(b-a)^2}{12}$$

2) normal distribution $N(\mu, \sigma^2)$ on $\langle -\infty, \infty \rangle$

PDF: $g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $E(X) = \mu$, $D(X) = \sigma^2$

CDF: $F(X) = \int_{-\infty}^{x} g(t)\, dt = \frac{1}{2}\left[1 + \mathrm{Erf}\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right)\right]$

where the error function is defined as

$$\mathrm{Erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

3σ rule - random „measurement" of X lies almost always in $P[\mu - 3\sigma \leq X \leq \mu + 3\sigma] = \mathrm{Erf}\left(\frac{3}{\sqrt{2}}\right) = 0{,}9973$

- probable error of measurement of a random variable X is defined using the ralation

$$P[\mu - r \leq X \leq \mu + r] = 0{,}5$$

and it is approximately $r \approx 0{,}67\,\sigma$

- error estimation of the MC method

- in the MC method we calculate averages of $X_1, \dots, X_n$, i.e.

$$A_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

where again $E(X_i) = m$ and $D(X_i) = b^2$ for all i

- according to the CLT the probability distribution of A_n will be approximately normal with

$$E(A_n) = \sum_{i=1}^{n} \frac{E(X_i)}{n} = m$$

and $$D(A_n) = \sum_{i=1}^{n} \frac{D(X_i)}{n^2} = \frac{b^2}{n} \quad \text{or} \quad \sigma = \frac{b}{\sqrt{n}}$$

- thus $$P\left[m - \frac{3b}{\sqrt{n}} \le A_n \le m + \frac{3b}{\sqrt{n}}\right] \simeq 0{,}9973$$

or $$P\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m\right| < \frac{3b}{\sqrt{n}}\right] \approx 0{,}9973$$

and the probable error of A_n is $\boxed{r = 0{,}67 \frac{b}{\sqrt{n}}}$

! the word probable is important, actual error can be larger but with quite small probability