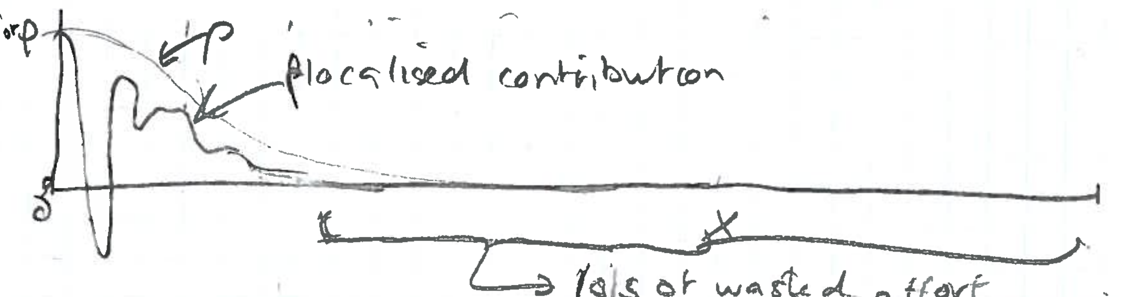
high $d \Rightarrow$ MC more efficient compared to grid

example:



suppose we know another function $p(x)$, which we can integrate, and wich is some what similar and > 0 if $f \neq 0$

$$\int dx\, f(x) = \underbrace{\int dx\, \rho(x)}\ \underbrace{\frac{f(x)}{\rho(x)}}$$

average this instead of f.

draw ρ-distributed points instead of homogeneous points concentrate where f is large

$$\int = \langle f/\rho \rangle_\rho$$

error in $\langle f \rangle$

$$\text{error}^2 = \left(\langle f^2 \rangle - \langle f \rangle^2\right) \frac{1}{N}$$

$$= \left[\frac{1}{N} \sum_i f(x_i)^2 - \left(\frac{1}{N} \sum_i f(x_i)\right)^2\right] \frac{1}{N}$$

big fluctuations in f

or error in $\langle f/\rho \rangle_\rho$

$$\text{error}^2 = \left(\left\langle \left(\frac{f}{\rho}\right)^2 \right\rangle_\rho - \left\langle \frac{f}{\rho} \right\rangle_\rho^2\right) \frac{1}{N}$$

not so big fluctuations in f/ρ

$$= \left[\frac{1}{N} \sum_i \left(\frac{f(x_i)}{\rho(x_i)}\right)^2 - \left(\frac{1}{N} \sum_i \frac{f(x_i)}{\rho(x_i)}\right)^2\right] \frac{1}{N}$$

Key is to pick good ρ
- analytically managrable
- easy to draw from

---

drawing gaussian random variables

draw X_1, X_2 between $0, 1$

$$\phi = 2\pi X_1$$

$$R = [-\log(X_2)]^{1/2} \qquad R \text{ gets } x \exp(-x^2) \text{ type dist}$$

$$Y_1 = R \cos \phi$$

$$Y_2 = R \sin \phi$$

$$P_R(R)\, dR = \overbrace{1}^{P_X(X_1)\, dX_1} \cdot \frac{dX}{dR}\, dR = \left[\frac{1}{(-\log(X_2))^{1/2}} \frac{1}{2} \frac{1}{X_2}\right]^{-1} \propto R \exp(-R^2)$$

14h45 ini.

More MC

calculating some average:

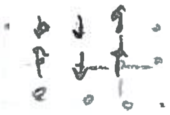
$$\langle A\rangle = \frac{\sum_{\text{states } \alpha} A(\alpha)\exp(-\beta E(\alpha))}{\sum_\alpha \exp(-\beta E(\alpha)) = Z}$$

$$\int_{\text{phasespace}} d\gamma\, A(\gamma)\exp(-\beta E(\gamma)) \Big/ \int d\gamma \exp(-\beta E(\gamma))$$

example: Ising model

$$\mathcal{H} = -J\sum_{nn} s_i s_j - h\sum_i s_i$$

coupling; field in direction



~~Hopefully~~ Supriya has discussed this system just in 2d

every spin is just up or down
⇒ should be simple enough

in 3d, 5x5x5 grid (not big)

$N = 125$ # states $2^{125} = 10^{37}$

calc Z or something
$O(1\text{ GHz})$ computer $= 10^9$ calcs/second.

$\Rightarrow$ $O(10^{28}\,s) \approx O(10^{20})$ years, >> lifetime of universe.

<u>considering all states is hopeless</u>

MC says just some random subset instead α_i

$$\langle A\rangle = \frac{\frac{1}{B}\sum_{i=1}^{B} A(\alpha_i)\exp(-\beta E(\alpha_i))}{\frac{1}{B}\sum_{i=1}^{B}\exp(-\beta E(\alpha_i))} \quad \begin{matrix}\to \langle A\exp(-\beta E)\rangle_{set} \\ \to \langle \exp(-\beta E)\rangle_{set}\end{matrix}$$

Now consider low T ---- spins tend to align

most random states do not look like this.

$\underbrace{\uparrow\uparrow\uparrow\uparrow}\downarrow\downarrow\downarrow\downarrow\downarrow$

to get sensible convergence, we need these states to be sampled

$O(1)$ out of 10^{37} .... B must be very large

same trick as with MC integration: importance sampling
solution: bias the random states $\rho(\alpha)$ density

$$\langle A\rangle_{\exp(-\beta E)} = \frac{1}{B}\sum_{i=1}^{B}\xi(\alpha_i) \quad \text{with } \xi(\alpha) = \frac{A(\alpha)}{\cancel{\rho(\alpha)}}\exp(-\beta E)$$

IMPORTANCE SAMPLING

same story about errors: $(\langle\xi^2\rangle_\rho - \langle\xi^2\rangle_\rho)$

smaller than $\langle A^2\rangle_1 - \langle A\rangle_1^2$

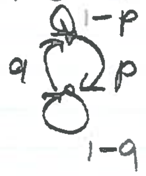
good choice would be $\rho(\alpha) \propto \exp(-\beta E(\alpha))$

Markov-chain MC

all about obtaining sampling of $\exp(-\beta E)$

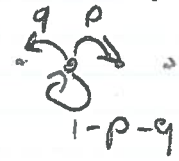
Markov chain $\begin{cases} \text{set of } n \text{ states} & \alpha_1 \; --- \; \alpha_N \\ \text{transfer probs between states} & P_{ij} \end{cases}$

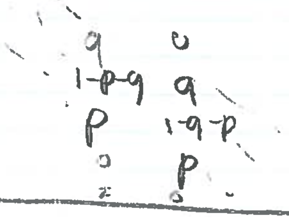
- no memory (depends only on previous state)
- used to model lots of processes; examples
  - simplistic attempt at modeling weather
    precipitation
    no precipitation
    (in a day)



$$\begin{array}{c|cc} & \text{yes} & \text{no} \\ \hline \text{yes} & 1-p & q \\ \text{no} & p & 1-q \end{array}$$

  - random walk on a line (large # states)
    state = where walker is





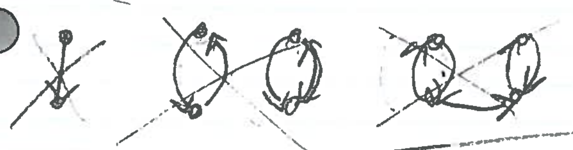
  - NS model

---

idea of metropolis type algorithms is to construct a Markov chain that moves the system to likely configurations (with the correct eq. dist)

- end up in the correct equilibrium
- accessibility: must be possible to reach all states
- (detailed) balance: there has to be an equilibrium in the Markov chain.

- accessibility:



- balance

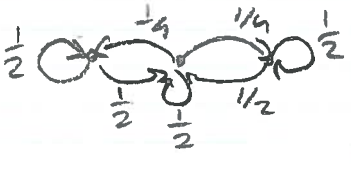
equilibrium: state α_i population a_i

$P_{ij} a_j = a_i$ eigenvalue equation for $(a_1, \ldots a_n)$ (eigenvalue 1)

example (1)

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{2}{3} \end{pmatrix} \Rightarrow a = \frac{1}{3}(1, 2)$$

example (2)



$$\Rightarrow a = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$$