Monte Carlo method for numerical integration

$$I = \int_V d^D x \; f(\vec{x})$$

$$= \int_V d^D x \left(N \frac{f(\vec{x})}{p(\vec{x})} \right) \frac{p(\vec{x})}{N}$$

$g(\vec{x})$

$p(\vec{x}) \geq 0$

p chosen so as to select the important region of V for the integral I.

Random sampling

$\rightarrow p(\vec{x}) = \frac{1}{V}$ $\{\vec{x}_i\}$ random points

$$I_M = \frac{1}{M} \sum_{i=1}^{M} f(\vec{x}_i) \longrightarrow I$$

$$|I_M - I| \sim O\left(\frac{1}{\sqrt{M}}\right) A$$

$$A \sim O(\exp(D))$$

$$\boxed{I = \int d^D x \ \tilde{g}(\vec{x}) \ \tilde{p}(\vec{x})} = \langle g \rangle_p$$

$$\left(\tilde{p}(\vec{x}) = \frac{e^{-\frac{E(\vec{x})}{k_B T}}}{Z} \qquad \beta = \frac{1}{k_B T} \right)$$

$$S = -\int d^D x \ \tilde{p}(\vec{x}) \log \tilde{p}(\vec{x})$$

$$= -\langle \log \tilde{p} \rangle_{\tilde{p}}$$

$$\underline{N_{points} \sim \exp(S)}$$

$\#$ of points contributing significantly to I

prob. (to get a significant point out of a random extraction) $\sim \dfrac{N_{points}}{N_{tot}}$

$$\sim e^{(S - S_{max})}$$

$$\Delta S < 0$$

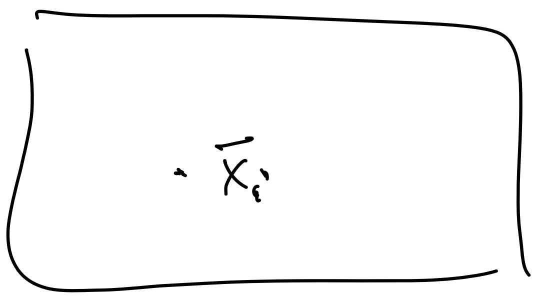
$$\underline{\Delta S = D \ \delta S}$$

$e^{-\beta \delta S}$

Rejection sampling

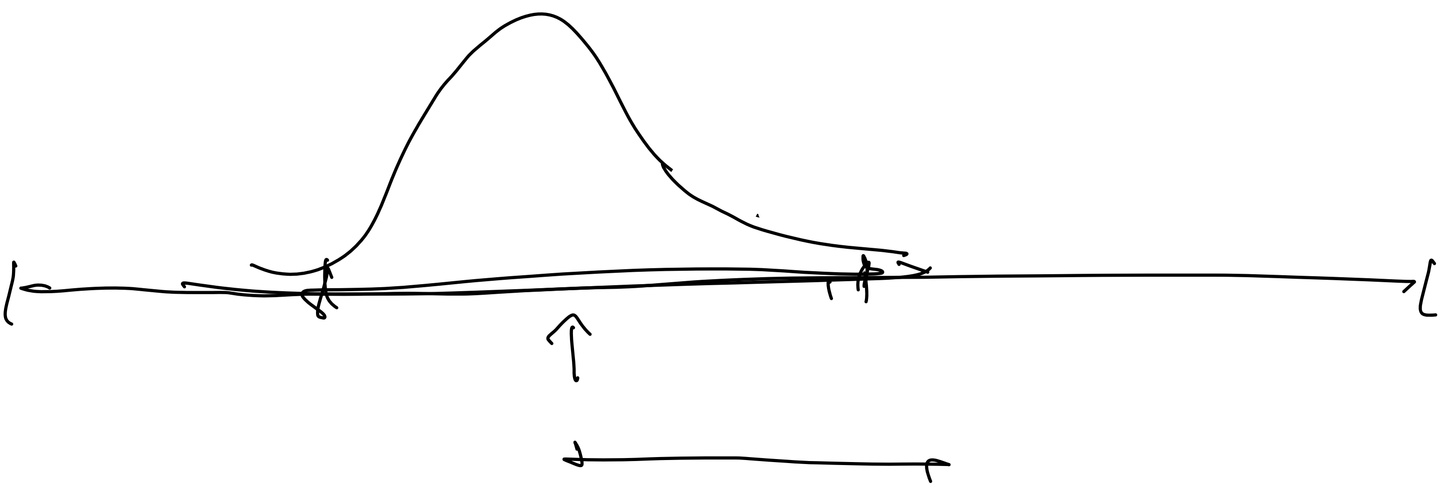
try to produce points distributed with $\tilde{p}$ but I propose them randomly ?

$$I_M = \frac{1}{M}\sum_{i=1}^{M} f(\vec{x}_i)$$



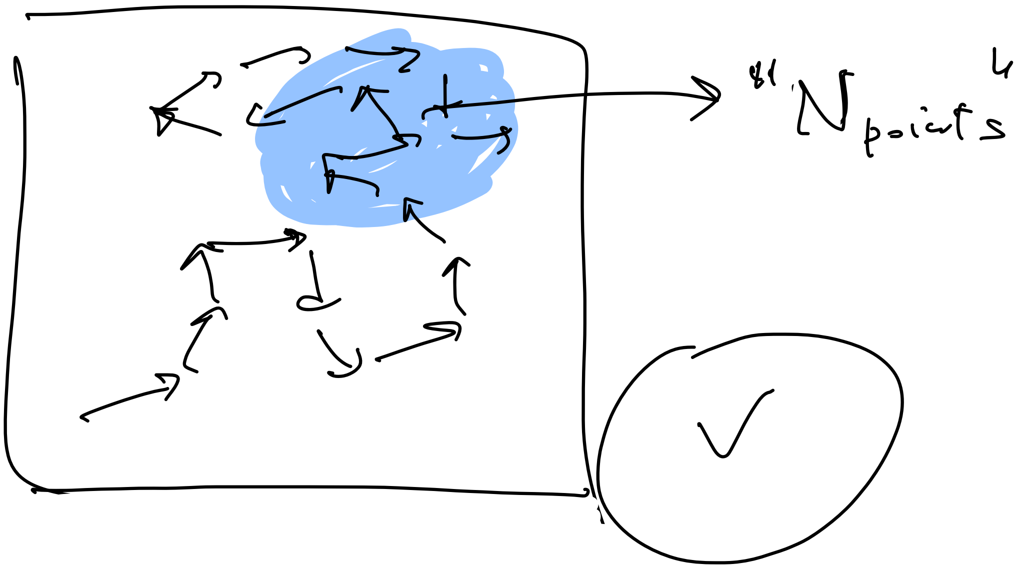
$\tilde{p}(\vec{x}_i)$: acceptance prob. for a point $\vec{x}_i$

probability of accepting a point $\sim O(\exp(-D))$



Importance sampling based Markov chains

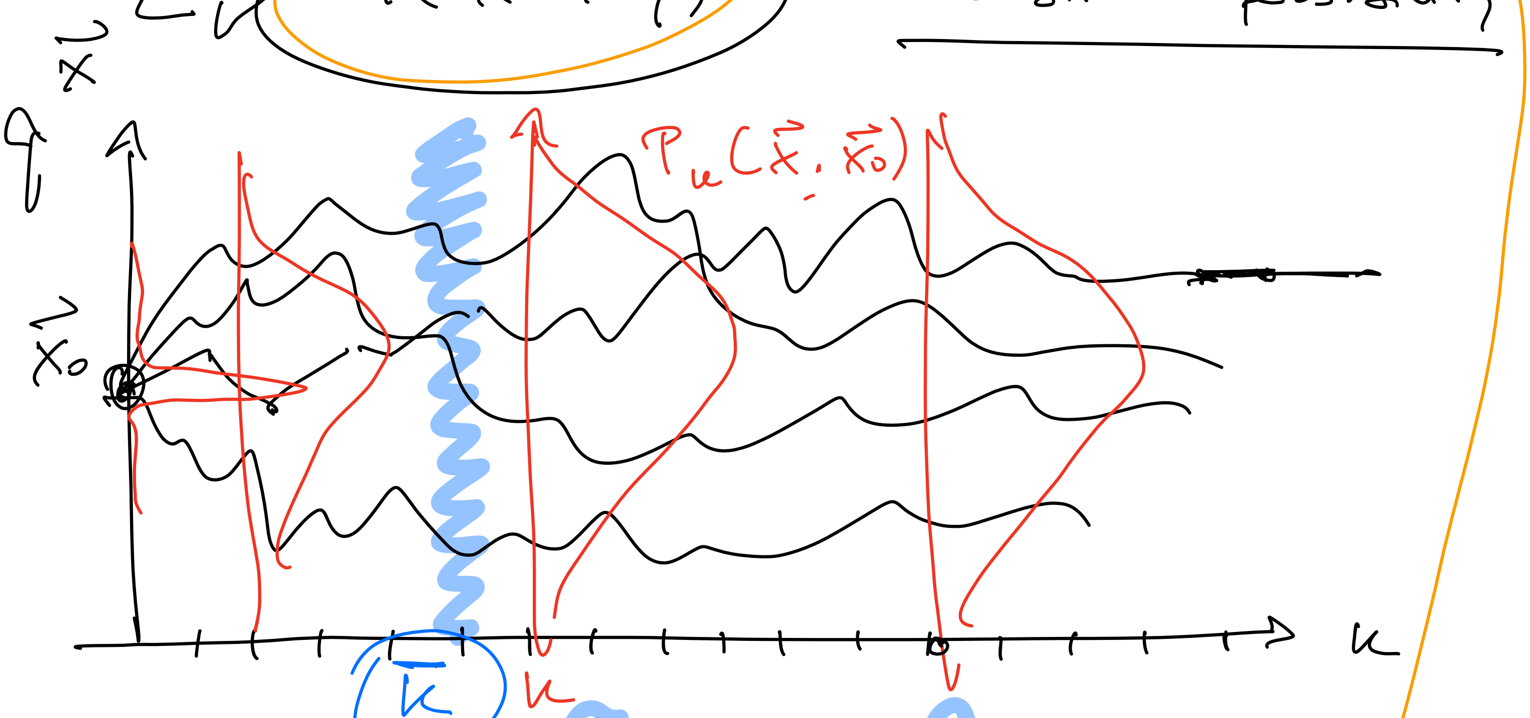
(Markov chain Monte Carlo)



$$\vec{x}_0 \to \vec{x}_1 \to \vec{x}_2 \to \dots$$

$$T(\vec{x}_1 \to \vec{x}_2) =$$

$$T(\vec{x} \to \vec{y})$$ transition probability



$$\underline{\underline{P_k(\vec{x};\vec{x}_0)}}$$

equilibration time

$$P_k(\vec{x};\vec{x}_0) \underset{k \gg \bar{k}}{\simeq} P(\vec{x}) = \frac{p(\vec{x})}{N}$$

stationary regime

$$\text{"}\frac{dP_k}{dt}\text{"} = P_{k+1}(\vec{x};\vec{x}_0) - P_k(\vec{x},\vec{x}_0)$$

$$= \sum_{\vec{q}} P_k(\vec{q};\vec{x}_0)\,T(\vec{q}\to\vec{x}) - \sum_{\vec{q}} P_k(\vec{x};\vec{x}_0)\,T(\vec{x}\to\vec{q}) \underline{\underline{=0}}$$

stationary regime

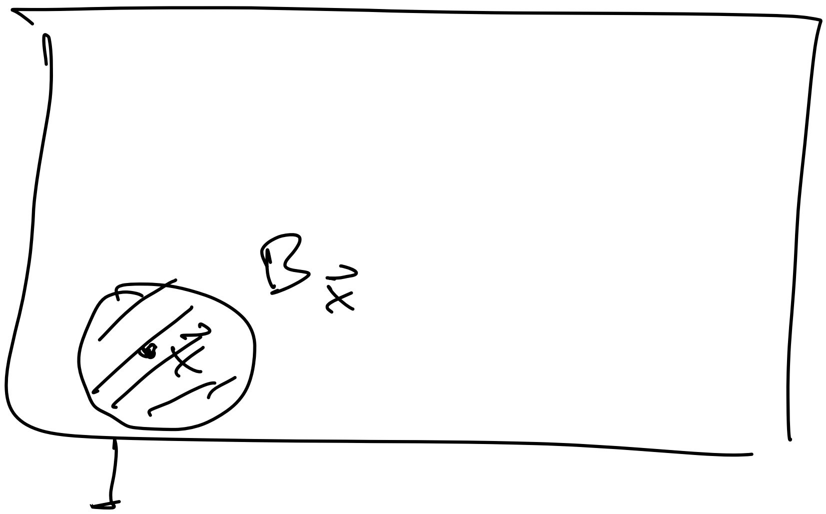
$$\boxed{\sum_{\vec{q}}\left(\frac{p(\vec{q})}{\cancel{N}}\,T(\vec{q}\to\vec{x}) - \frac{p(\vec{x})}{\cancel{N}}\,T(\vec{x}\to\vec{q})\right) = 0}$$

$$\Uparrow$$

$$\to \quad p(\vec{q})\,T(\vec{q}\to\vec{x}) = p(\vec{x})\,T(\vec{x}\to\vec{q})$$

detailed balance condition

$$T(\vec{x}\to\vec{y}) = \underbrace{T_{prop}(\vec{x}\to\vec{y})}_{\text{proposal probability}} \underbrace{A(\vec{x}\to\vec{y})}_{\text{acceptance probability}}$$



one choice

$$T_{prop}(\vec{x}\to\vec{y}) = \begin{cases} \frac{1}{B_{\vec{x}}} & \vec{y}\in B_{\vec{x}} \\ 0 & \text{otherwise} \end{cases}$$

$$T_{prop}(\vec{x}\to\vec{y}) = T_{prop}(\vec{y}\to\vec{x})$$

$$A(\vec{x}\to\vec{y}) = \frac{p(\vec{y})}{p(\vec{x})} A(\vec{y}\to\vec{x})$$

Metropolis - Hastings solution

$$\boxed{A(\vec{x} \to \vec{y}) = \min\left(1, \frac{p(\vec{y})}{p(\vec{x})}\right)}$$

practical algorithm

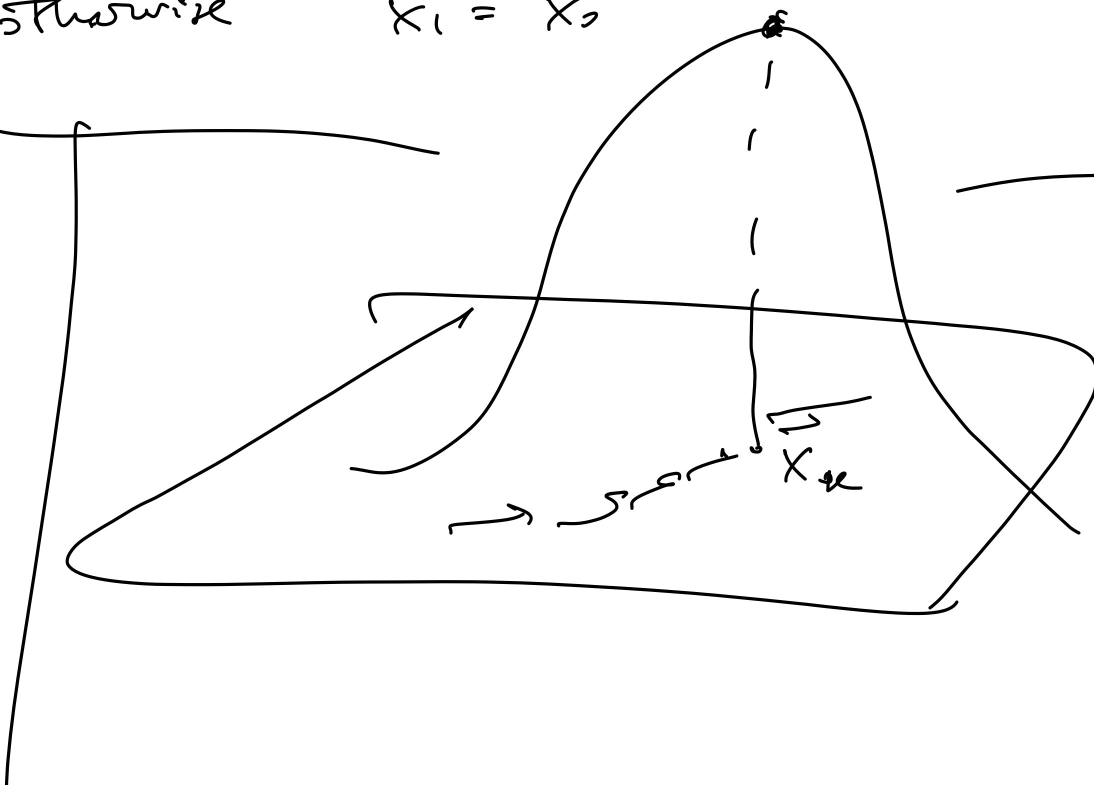
$$\underset{k}{\vec{x}_0} \longrightarrow \underset{k+1}{\vec{x}_1}$$

extract $\vec{x}_1$ with $T_{prp}(\vec{x}_0 \to \vec{x}_1)$

extract $z \in [0, 1]$ random number

if $z < \min\left(1, \frac{p(\vec{x}_1)}{p(\vec{x}_0)}\right)$ ✓ next point is $\vec{x}_1$

otherwise $\vec{x}_1 = \vec{x}_0$



After repeating this process k times with $k \gg \bar{k}$

$\{\vec{x}_\ell\}$ are distributed according to $\frac{p(\vec{x})}{N}$

$\ell \gg \bar{k}$

$$\boxed{I_M = \frac{1}{M} \sum_{\ell = \cancel{\bar{k}+} n}^{M + \cancel{\bar{k}} + n} g(\vec{x}_\ell)}$$

$n \gg 1$

sum of random variables equally distributed

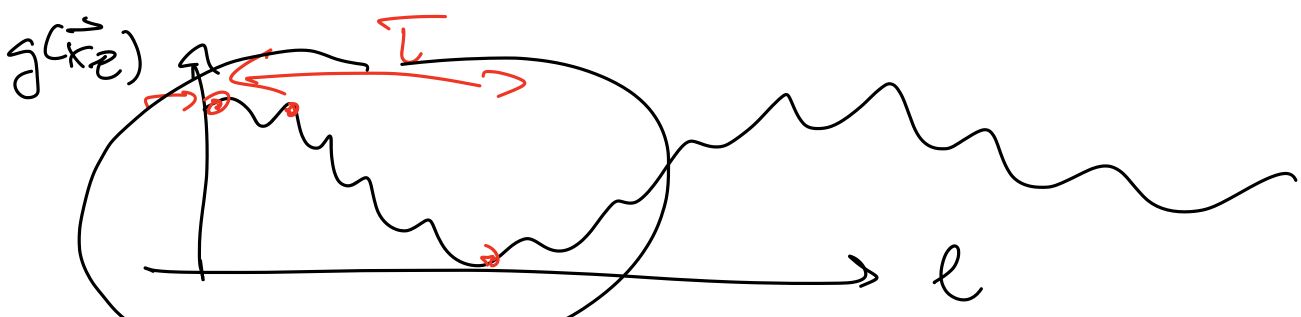
$$| I_M - I |$$

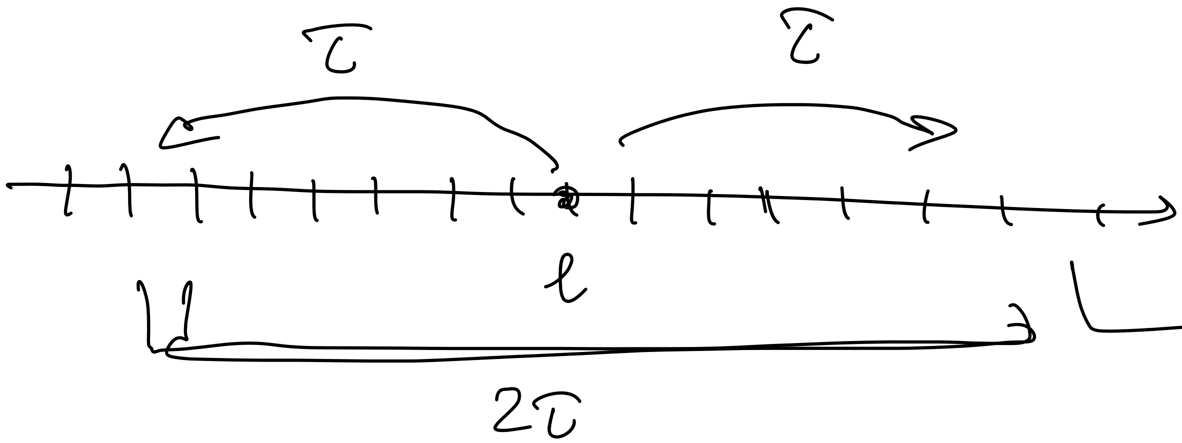
$$\sigma_g^2 = \int d^D x \, \frac{p(\vec{x})}{N} \left(g(\vec{x}) - I \right)^2$$

$$\sigma_{I_M}^2 \neq \frac{\sigma_g^2}{M}$$

CLT ✗

autocorrelation time





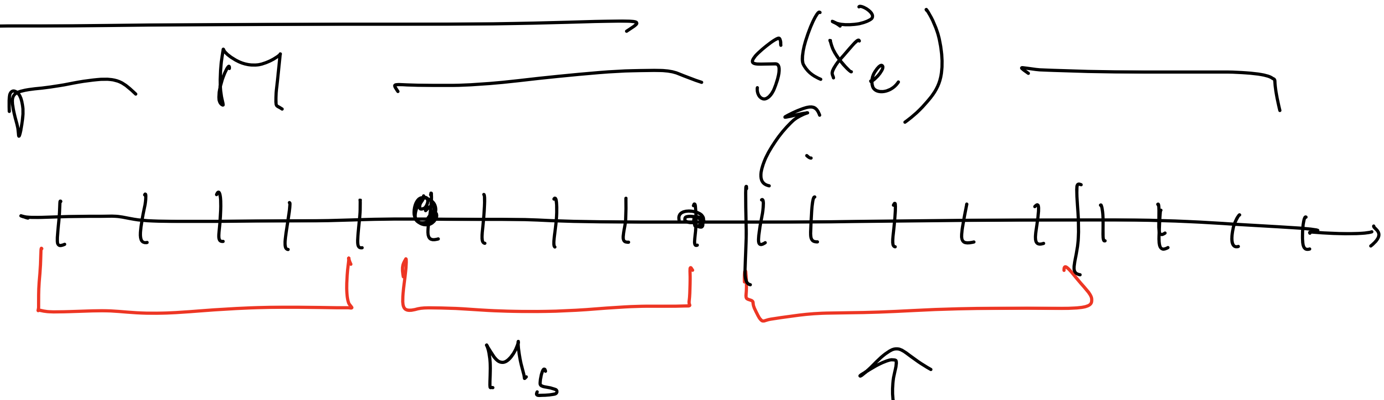
$$M \longrightarrow M_{eff} = \frac{M}{2\tau}$$

$$\sigma^2_{I_M} = \frac{\sigma_s^2}{M_{eff}} = \frac{(2\tau)}{T} \frac{\sigma_s^2}{M}$$

How to estimate τ ?

$$\overline{h}$$

Block variables



$$M = n_b M_b$$

$$I_M = \frac{1}{M} \sum_{l=1}^{M} g(\vec{x}_l)$$

$$= \frac{1}{n_b} \sum_{a=1}^{n_b} \underbrace{\frac{1}{M_b} \sum_{l=(a-1)M_b+1}^{a \cdot M_b} g(\vec{x}_l)}$$

$$= \frac{1}{n_b} \sum_{a=1}^{n_b} \bar{g}_a$$

$\bar{g}_a$ block average

$M_b \gg \tau \Longrightarrow$ two block averages are statistically independent

$$\boxed{\sigma_{I_M}^2 = \frac{\sigma_{\bar{g}}^2}{n_b} = 2\tau \frac{\sigma_g^2}{M}}$$

$$\sigma_{\bar{g}}^2 = \frac{1}{n_b} \sum_a \left(\bar{g}_a - I\right)^2$$

$$\boxed{\tau = \frac{1}{2} M_b \frac{\sigma_{\bar{g}}^2}{\sigma_g^2}}$$

Computational Cost of numerical integration using MC Monte Carlo ?

$$T_\epsilon \sim \frac{A(\tau)}{\epsilon^2}$$

$$|I_M - I| \sim \sqrt{\frac{M\tau}{M}} \sim \frac{1}{\sqrt{M_{eff}}}$$

independent of D

generically

$$\tau \simeq a \cdot D L^z \qquad z < \infty$$

dynamical exponent

L : length of each integration dimension

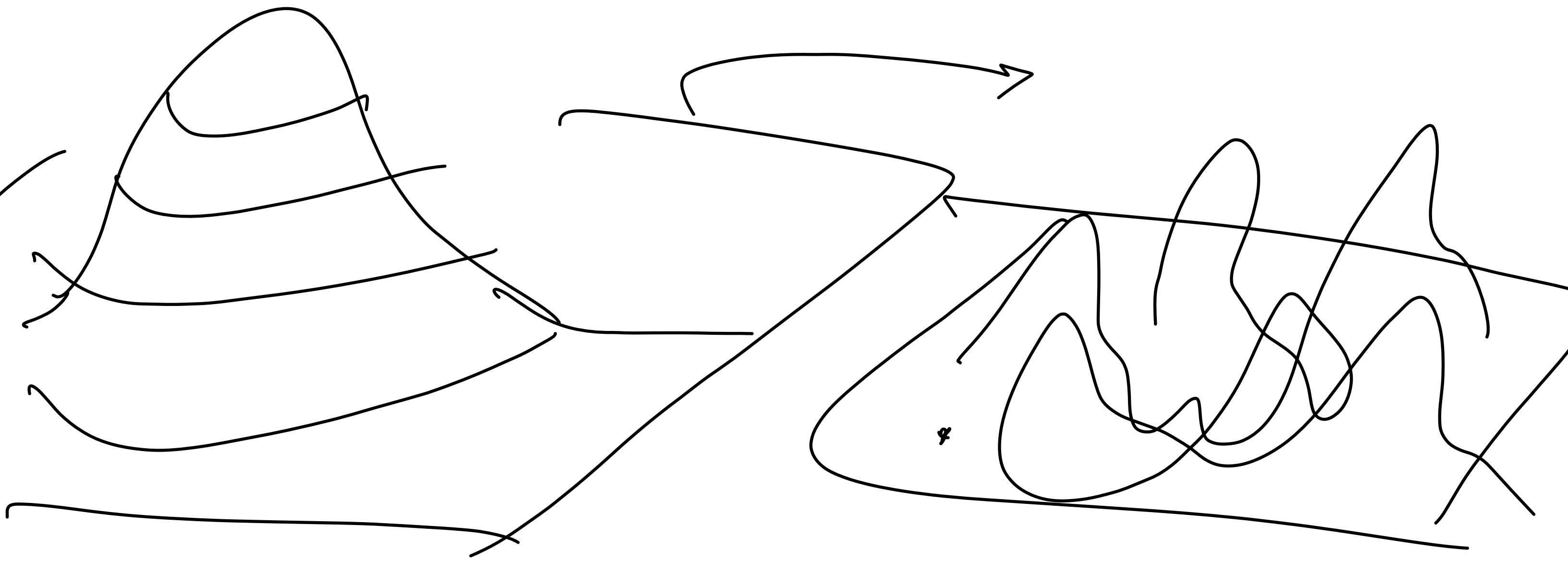
$$V \approx L^D$$

generically $z \simeq 2$ ✓

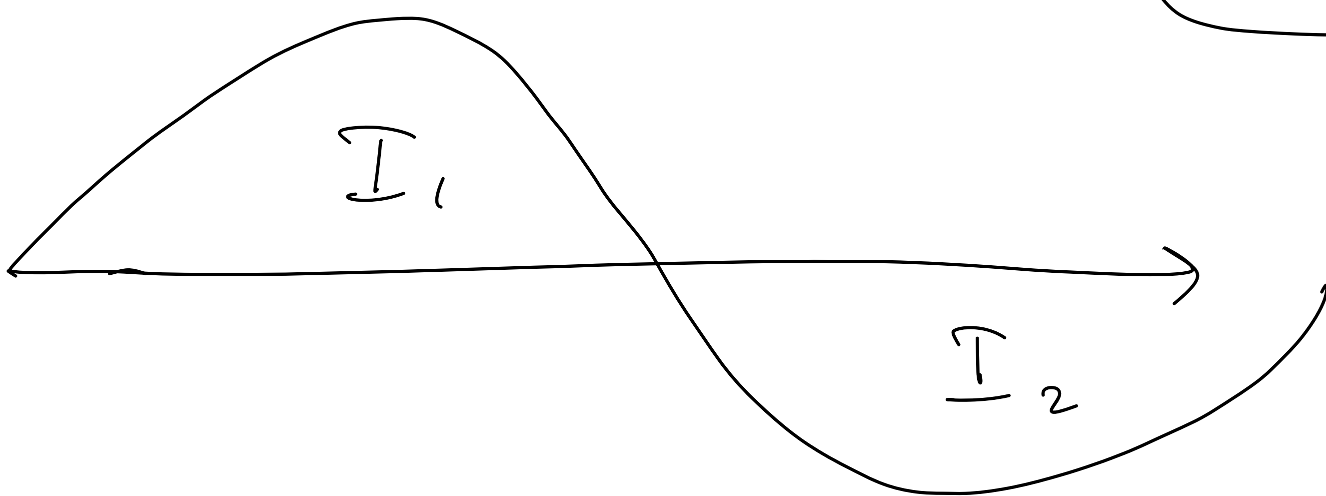
Bad situations : $z = \infty$

$\tau \sim O(\exp(D))$
(L)

1) $\frac{p(\vec{x})}{N}$ has a very complex landscape



2) sign problem $I = \int d^D x \, g(\vec{x}) \frac{P(\vec{x})}{N}$



$$I = I_1 + I_2 = \text{poly}(D)$$

$I_1 \to \mathcal{O}(\exp(D))$ $\qquad$ $I_2 \to \mathcal{O}(\exp(D))$

Minimization of a multivariable function

$$\boxed{\begin{array}{l} f(\vec{x}) \qquad \vec{x} \in \mathbb{R}^N \\ \vec{x}^{\,*} : \quad \min\limits_{\vec{x} \in A (\subseteq \mathbb{R}^N)} f(\vec{x}) = f(\vec{x}^{\,*}) \end{array}}$$

Examples (from physics & beyond)

1) $\min H(\{\vec{x}_i, \vec{p}_i\})$

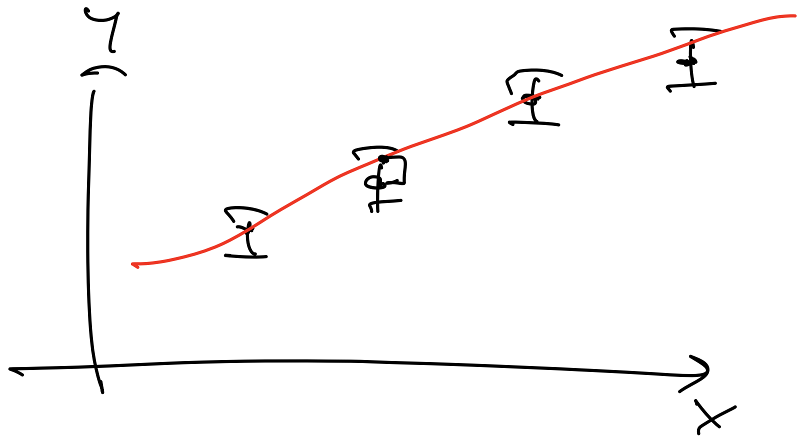
2) $\langle \underline{\Psi}(\vec{\alpha}) \rangle$ $\qquad$ $\vec{\alpha} = \alpha_1, \dots, \alpha_M \to \vec{x}$

parameters

$$\min_{\vec{\alpha}} E(\vec{\alpha}) = \min_{\vec{\alpha}} \frac{\langle \Psi(\vec{\alpha}) | \hat{H} | \Psi(\vec{\alpha}) \rangle}{\langle \Psi(\vec{\alpha}) | \Psi(\vec{\alpha}) \rangle}$$

3) Fit experimental data to a function

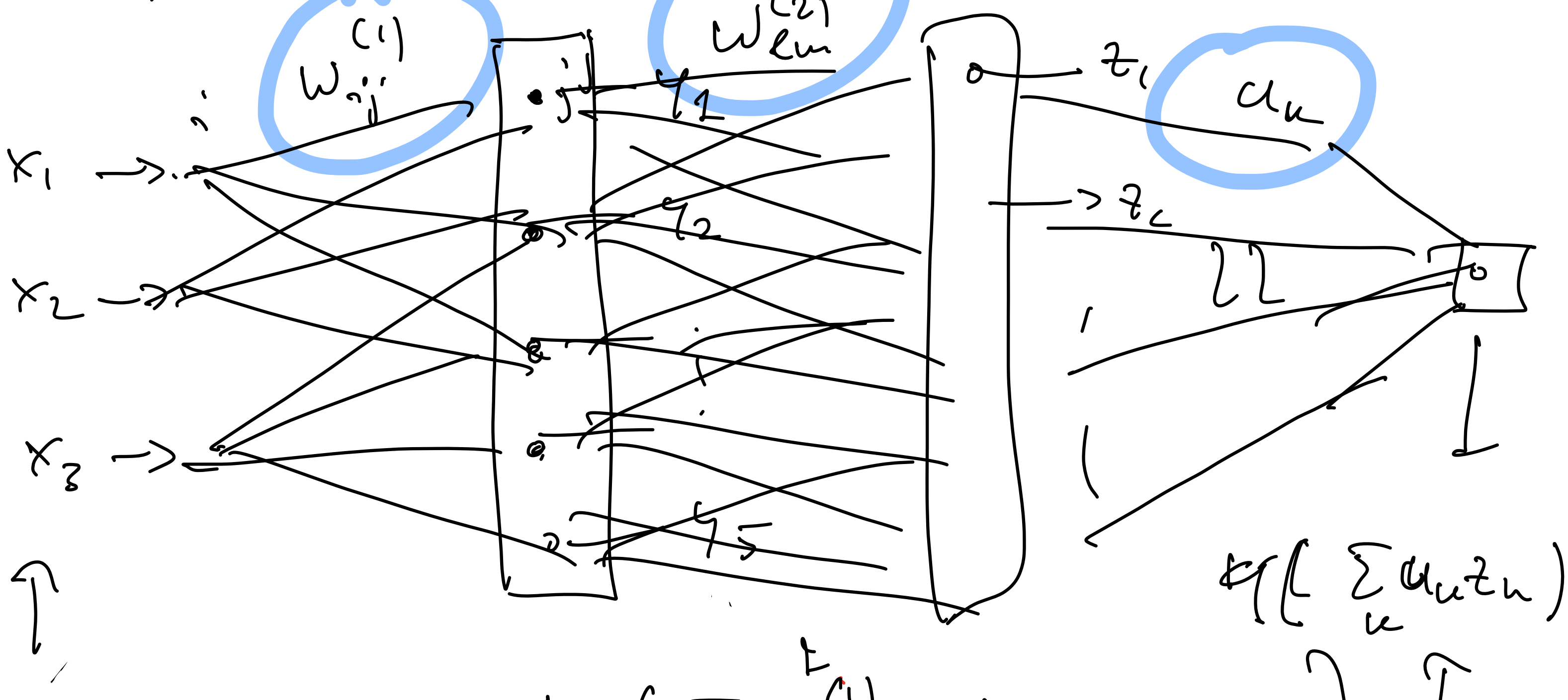
$\{x_i, y_i, \sigma_i\}$



$$\boxed{g(x, \vec{\alpha})}$$

$$\min_{\vec{\alpha}} \chi^2(\vec{\alpha}) = \min_{\vec{\alpha}} \sum_i \frac{(g(x_i; \vec{\alpha}) - y_i)^2}{\sigma_i^2}$$

4) Train a neural network



$$y_j = h_j\left(\sum_i W_{ij}^{(1)} x_i\right)$$

non-linear function

$g(\vec{x}; \{W_{ij}^{(k)}\})$

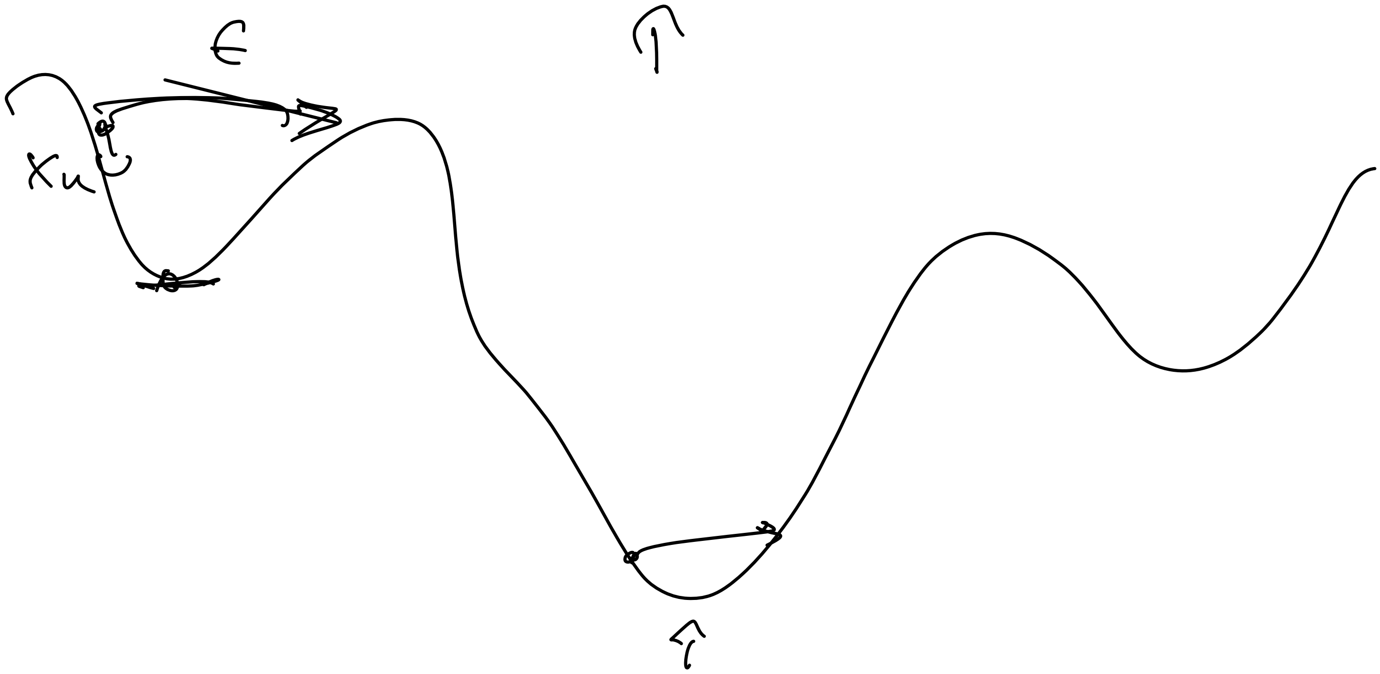
target response

Only certain solution to minimization: enumeration

$\vec{x} \in \mathbb{R}^N \longrightarrow \vec{x} \in \mathbb{C}^N$

Heuristic solutions :

$\vec{x}'$: $\boxed{\vec{\nabla} f(\vec{x}) = 0}$



Gradient descent

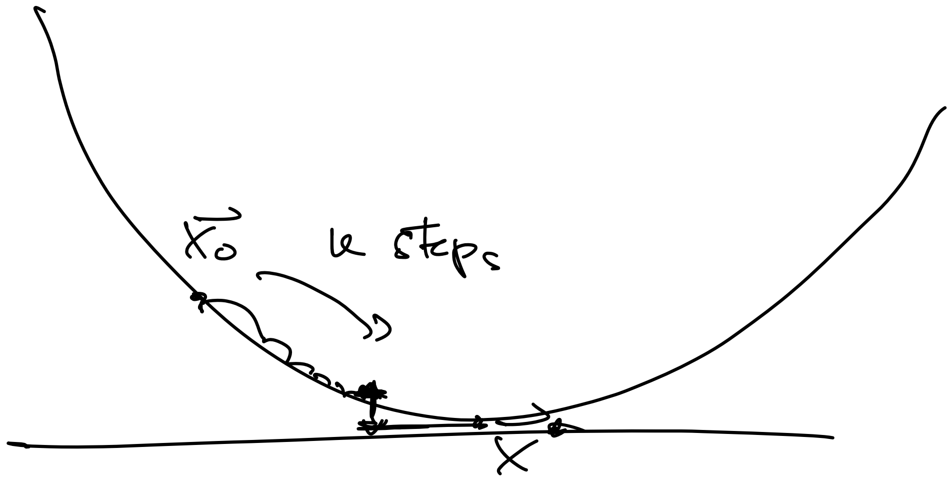
$$\vec{x}_k \to \vec{x}_{k+1} = \vec{x}_k - \epsilon \vec{\nabla} f(\vec{x}_k)$$

ϵ = "small" step

$$f(\vec{x}_{k+1}) - f(\vec{x}_k) =$$

$$f(\vec{x}_k - \epsilon \vec{\nabla} f) - f(\vec{x}_k)$$

$$= -\epsilon \|\vec{\nabla} f\|^2 + \frac{\epsilon^2}{2} \vec{\nabla} f_k^T H \vec{\nabla} f_k + o(\epsilon^3)$$



$\|\epsilon \vec{\nabla} f\|$

prove : $f(x_k) - f(\vec{x}_{min}) \sim O\left(\frac{1}{k}\right)$

worst-case scenario

$\sim \exp\left(-\frac{k}{k_0}\right)$