FYS4411/9411, MARCH 23, 2023

each experiment $-\alpha-$

$$\mu_\alpha = \frac{1}{n}\sum_{k=1}^{n} x_{\alpha,k} \quad \left(\neq \mu_{exact} = \int_{x\in D} p(x)\,x\,dx\right)$$

$$\sigma_\alpha^2 = \frac{1}{n}\sum_{k=1}^{n} (x_{\alpha,k}-\mu_\alpha)^2$$

- repeat m-times

$$\mu_m = \frac{1}{m}\sum_{\alpha=1}^{m} \mu_\alpha = \frac{1}{mn}\sum_{\alpha k} x_{\alpha,k}$$

if each $x_{\alpha,k}$ are i.i.d, then in the limit $m \to \infty$, we approach $p(x) \sim e^{-(\mu-x)^2/2\sigma_m^2}$

Total variance

$$\sigma_m^2 = \frac{1}{mn^2}\sum_{\alpha=1}^{m}\sum_{kl=1}^{n} (x_{\alpha,k}-\mu_m) \times (x_{\alpha,l}-\mu_m)$$

sample variance of all mn experiments

$$\sigma^2 = \frac{1}{mn}\sum_{\alpha k} (x_{\alpha,k}-\mu_m)^2$$

$$\sigma_m^2 = \frac{1}{mn^2} \sum_{\alpha k} (x_{\alpha k} - \mu_m)^2$$

$$+ \frac{2}{mn^2} \sum_{\alpha=1}^{m} \sum_{k<l}^{n} (x_{\alpha,k} - \mu_m) \times (x_{\alpha,l} - \mu_m)$$

$$= \frac{\sigma^2}{n} + \operatorname{cov}(x) \geq 0$$

standard deviation is

$$\mathrm{STD} = \sigma / \sqrt{n} \quad \text{if } \operatorname{cov}(x) = 0$$

$$f_d = \frac{1}{nm} \sum_{\alpha=1}^{n} \sum_{k=1}^{n-d} (x_{\alpha,k} - \mu_m) \times (x_{\alpha,k+d} - \mu_m)$$

$$d = |k - l|$$

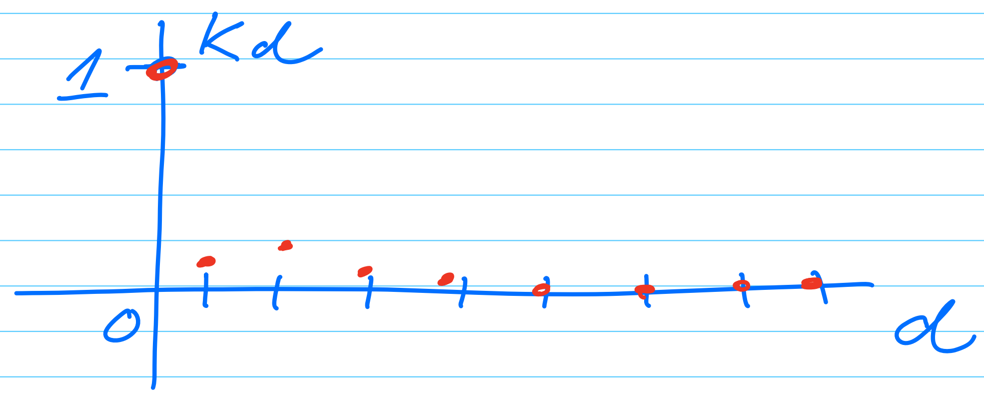
correlation function

$$K_d = \frac{f_d}{\sigma^2}$$

$$f_0 = \frac{1}{m} \frac{1}{n} \sum_{\alpha=1}^{m} \sum_{k=1}^{n} (x_{\alpha k} - \mu_m)^2$$

$$= \sigma^2$$

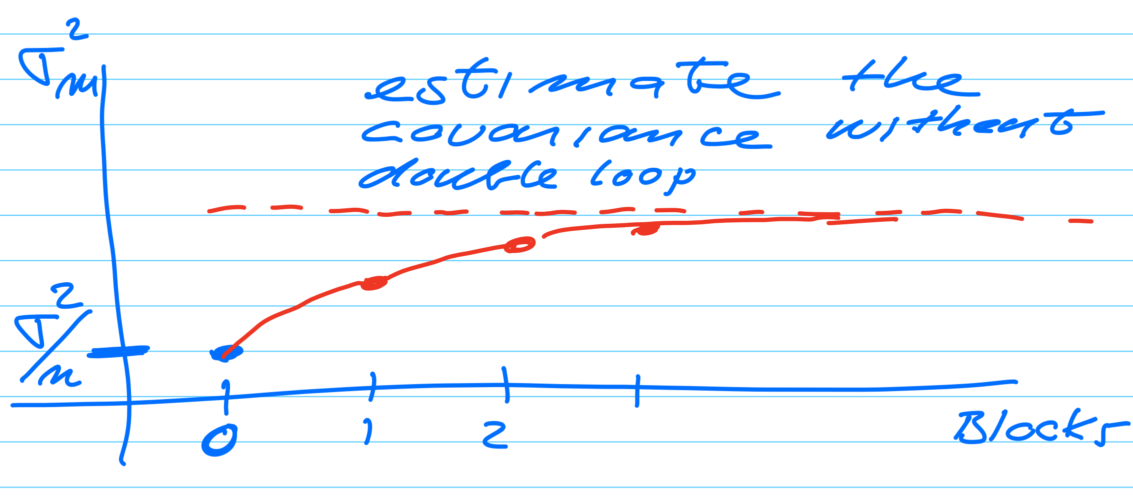
$K_0 = 1 \qquad K_d = 0 \quad d \neq 0$



$$\sigma_m^2 = \frac{\sigma^2}{n} + \frac{2}{n} \sum_{d=1}^{n-1} f_d$$

$$= \frac{\sigma^2}{n} \left[1 + 2 \sum_{d=1}^{n-1} K_d \right]$$

Flyvbjerg-Petersen, J chem Phys 91, 461 (1989)



Bootstrap is normally used with smaller samples

$$X = \{x_1\, x_2, \ldots x_n\}$$, new sample

$$X' = \{x_1'\, x_2' \ldots x_n'\}$$

repeat this B-times. With large samples, it becomes expensive (FLOPS) and time consuming.

Blocking algorithm

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i'$$

$$\sigma^2(\mu) = \mathbb{E}[x^2] - \mu^2$$

$m = 1$

$$\sigma_m^2 = \sigma_1^2 = \frac{\sigma^2}{n} + \frac{2}{n^2} \sum_{k<l} (x_k - \mu) \times (x_l - \mu)$$

$$= \sigma^2(\mu)$$

$$\gamma_{ij} = \mathbb{E}[x_i x_j] - \underbrace{(\mathbb{E}[\mu])^2}_{\mu^2}$$

$$= \gamma_t \qquad t = |i-j|$$

$$\gamma_0 = \sigma^2$$

$$\sigma^2(\mu) = \frac{1}{n^2} \sum_{i,j} \gamma_{i,j}$$

$$= \frac{1}{n}\Big[\underbrace{\gamma_0}_{\sigma^2} + \underbrace{2\sum_{t=1}^{n-1}\left(1-\frac{t}{n}\right)\gamma_t}_{\sum_{d=1}^{n-1} f_d} \Big]$$

algorithm:

Transform data set

$$X = \{x_1, x_2, \ldots x_n\}$$

into half as large data sets

$$\{x_1', x_2', x_3', \ldots x_n'\}$$

$$n' = \frac{1}{2} n$$

$$x_i' = \frac{1}{2}\left[x_{2i-1} + x_{2i}\right]$$

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$$

$$= \{1, 2, 3, 4, 5, 6\}$$

$$\mu = \frac{1}{6}\sum_{i=1}^{6} x_i = 7/2$$

$$X' = \left\{ \frac{x_1+x_2}{2}, \frac{x_3+x_4}{2}, \frac{x_5+x_6}{2} \right\}$$

$$\mu' = \frac{1}{3}\sum_{i=1}^{3} x_i' = 7/2$$

$$\sigma^2(\mu) = \frac{1}{6}\sum_{i=1}^{6} (x_i)^2 - \mu^2 = 8/3$$

$$\sigma'^2(\mu') = \frac{1}{3}\sum_{i=1}^{3} \frac{(x_{2i-1} + x_{2i})^2}{2^2} - \mu'^2$$

$$= \frac{1}{3}\left(\frac{(x_1+x_2)^2}{4} + \frac{(x_3+x_4)^2}{4} + \frac{(x_5+x_6)^2}{4} \right)$$

$$- \mu'^2 = 8/3$$

$$\boxed{\mu = \mu'} \quad \boxed{\sigma^2(\mu) = \sigma'^2(\mu')}$$

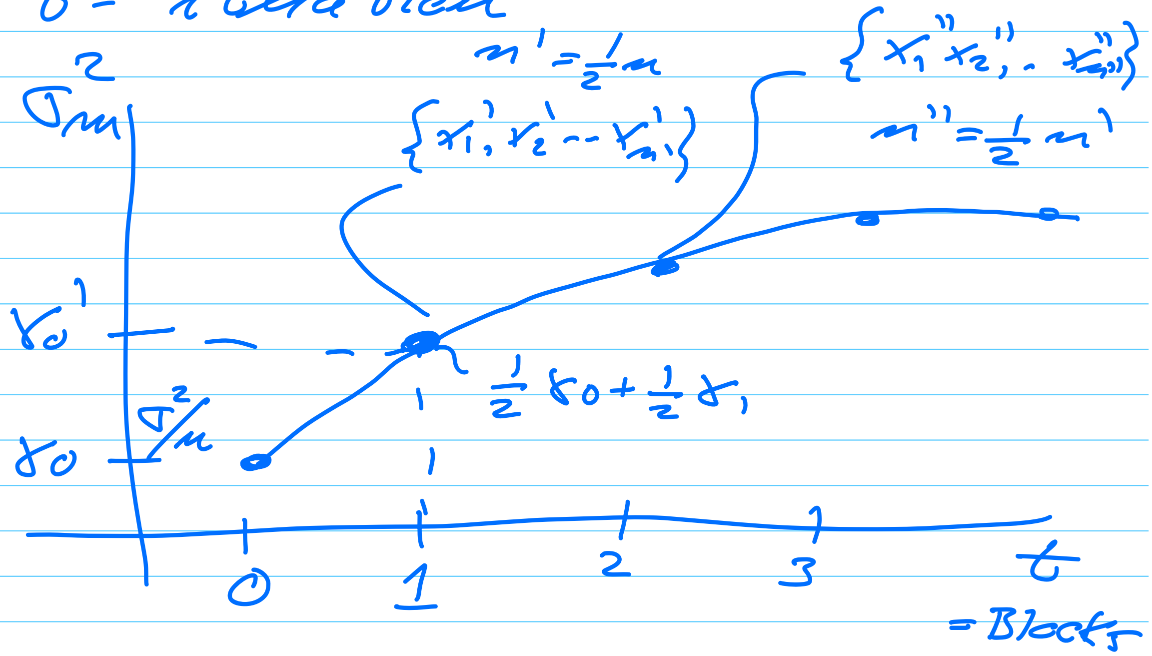
$$\# \text{ MCS} = 2^P$$

$x_i' \; ; \; \gamma_{ij}' = \gamma_0'$

$x_i' \; ; \; \gamma_{ij}' = \gamma_t \; , \; \gamma_0 = \sigma^2$

$$\gamma_t' = \begin{cases} \frac{1}{2}\gamma_0 + \frac{1}{2}\gamma_1 & t = 0 \\ \frac{1}{4}\gamma_{2t-1} + \frac{1}{2}\gamma_{t\cdot 2} + \frac{1}{4}\gamma_{2t+1} & t > 0 \end{cases}$$

$\gamma_0' = \frac{1}{2}\gamma_0 + \frac{1}{2}\gamma_1$, after we have calculated $\gamma_0 = \sigma^2$ continue to split data in halves till γ_t' reduces to a constant value after t- iteration



Onebody densities

unravel information about correlations

$$\rho(\vec{r}_1) = \int d\vec{r}_2 \, d\vec{r}_3 \dots d\vec{r}_N \times \left| \psi_T(\vec{r}_1, \vec{r}_2, \dots \vec{r}_N; \vec{\alpha}) \right|^2$$

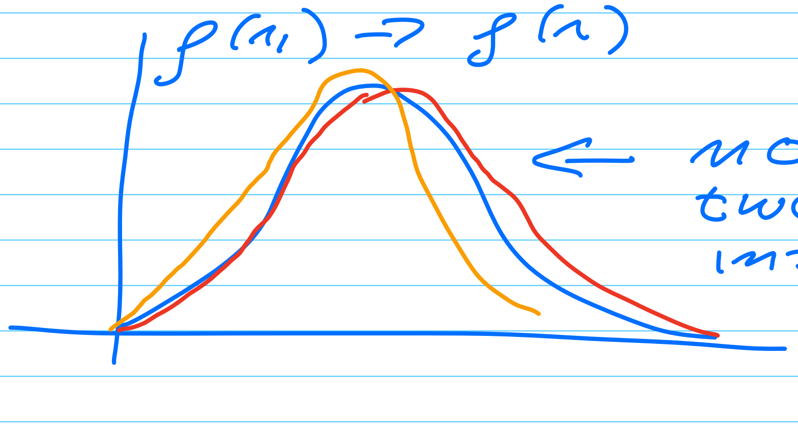
(after optimization)

2-dim

$\vec{r}_1 \Rightarrow r_1 = \sqrt{x_1^2 + y_1^2}$ (no angle ϕ-dependence)

make a table

$\rho(x_1, y_1)$ gives $\rho(r_1)$



- repulsive (pushed out)
- attractive (pulled in)