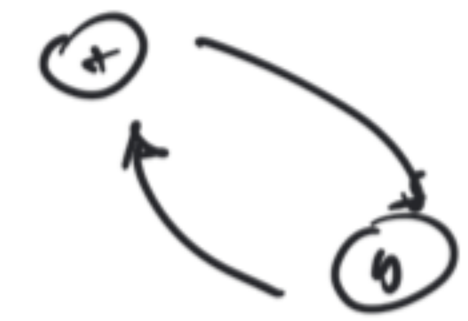
MCMC recap

$p(x)$, $x \in S$.

Detailed Balance Condition



$$\underbrace{p(x)\,\Pi(y|x) = p(y)\,\Pi(x|y)}$$

$q(y|x)$ proposal

sample $y \sim q(y|x)$

accept with prob. $A(y|x) = \min\left\{1, \frac{p(y)\,q(x|y)}{p(x)\,q(y|x)}\right\}$

GIBBS SAMPLER

$\underbrace{p(x_1, \dots, x_n)}$ knowing $p(x_j | x_{\setminus j})$

$q(y|x) \sim \mathcal{N}(x, 10^2)$

$p(x) = \text{bimodal} = \frac{1}{3}\mathcal{N}(0,1) + \frac{2}{3}\mathcal{N}(10,1)$

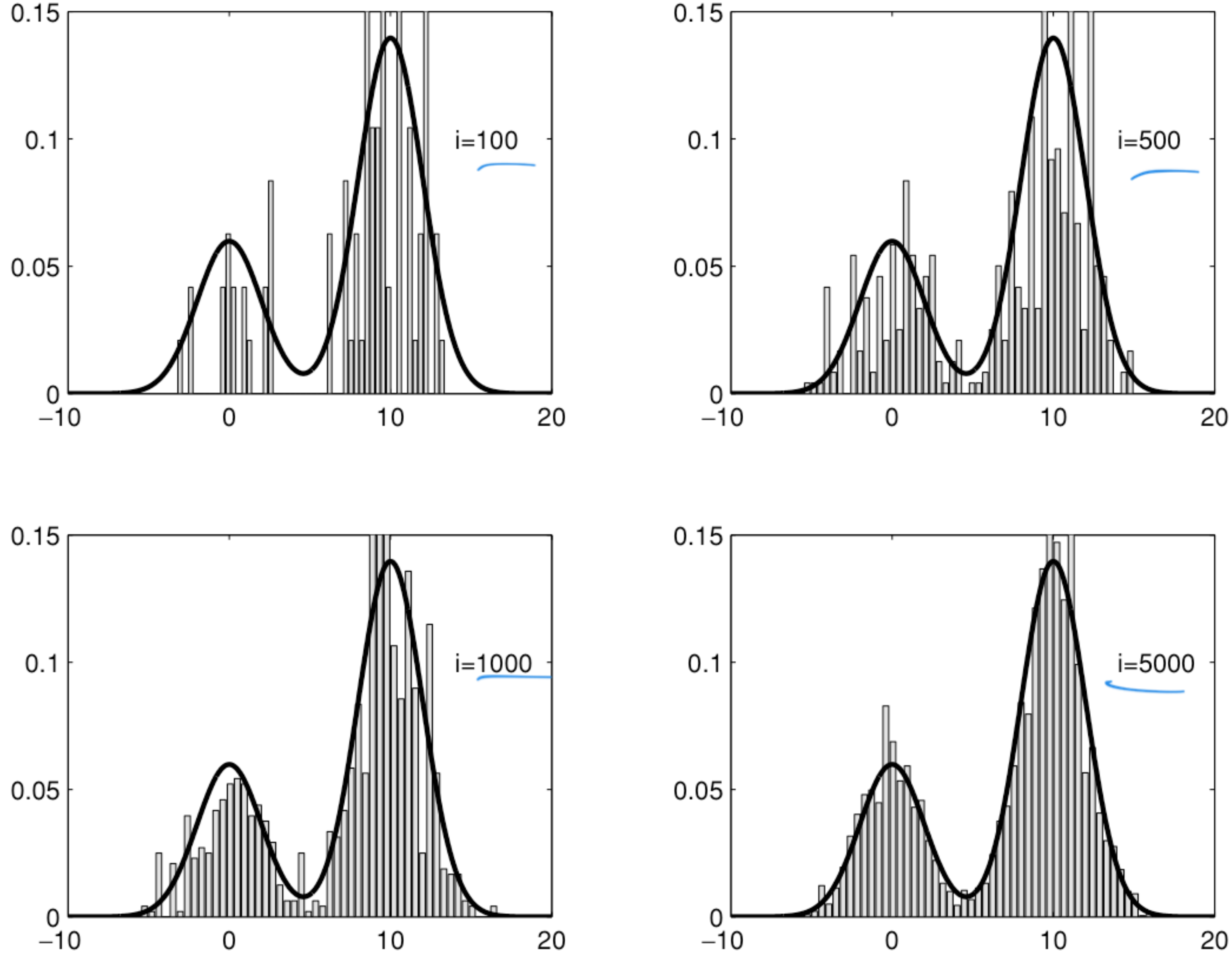


Figure 6. Target distribution and histogram of the MCMC samples at different iteration points.

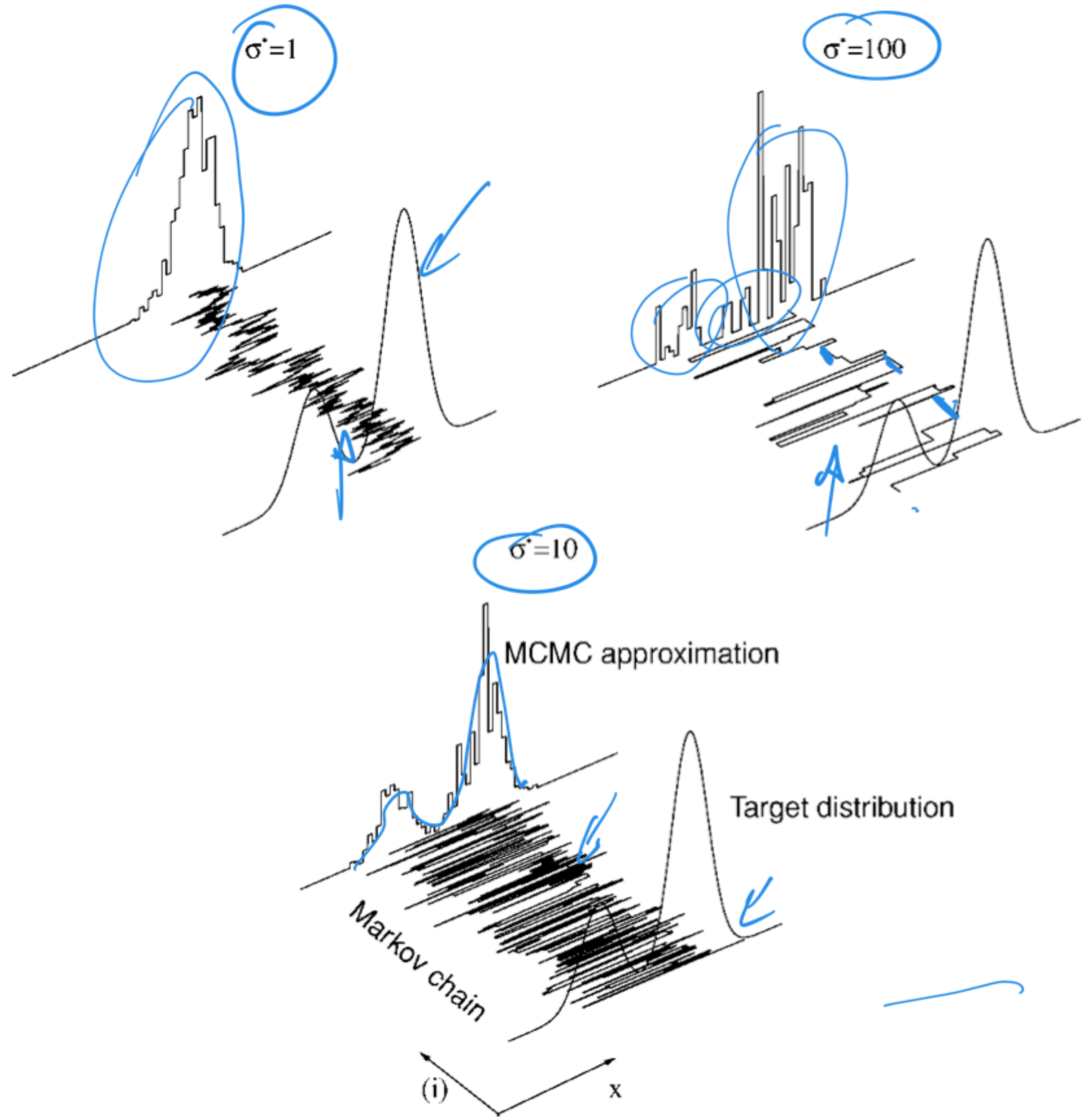


Figure 7. Approximations obtained using the MH algorithm with three Gaussian proposal distributions of different variances.

Batch Means

$\int p(x) f(x)\, dx$

MCMC $\quad p(x)$

$x_1, \dots, x_N$ sample $\rightsquigarrow \underbrace{I_N(f) = \frac{1}{N}\sum_i f(x_i)} \to E[f(x)] = I(f)$

$\Downarrow$

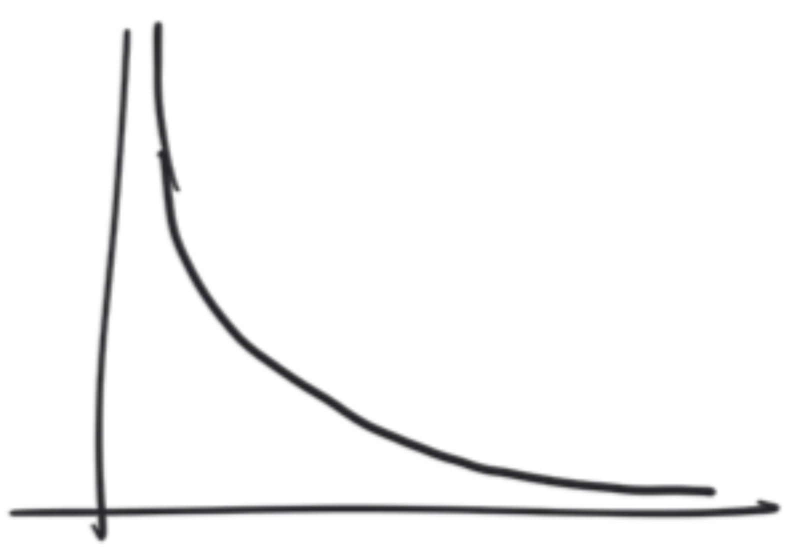
CORRELATED $\quad \sigma^2/N$

$$\sigma^2 = VAR[f(x)] + \boxed{\sum_{k\geq 1} \gamma_k}$$

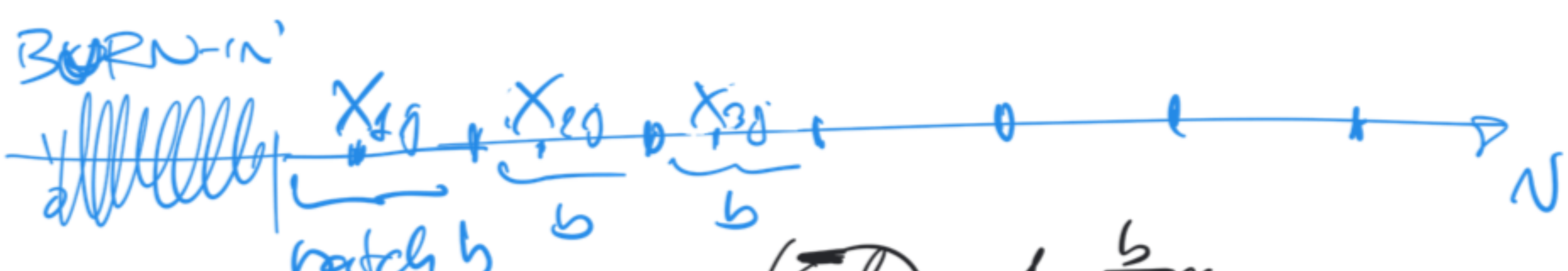
LAG-K AUTOCOVARIANCE

$$\gamma_k = CORR[f(x_i), f(x_{i+k})]$$

$$\hat{\gamma}_k = \frac{1}{N}\sum_{i=1}^{N-k} [f(x_i) - I_N(f)][f(x_{i+k}) - I_N(f)]$$



$\gamma_k \sim a^k$, $a < 1$, ergo $\sum_{k\geq 1} \gamma_k < \infty$



b - size of the batch

X_{bj} ≡ jth element in batch b.

mean of batch b $\leftarrow$ $\bar{X}_b = \frac{1}{b}\sum_{j=1}^{b} f(X_{bj})$

Idea: $\bar{X}^f_h$ are "almost" uncorrelated. Sample $\bar{K}$ batch of size b

Then I estimate $I_N(f)$ as $(N = b \cdot \bar{K})$

$$I^b_N(f) = \frac{1}{K} \sum_{j=1}^{K} \bar{X}^f_j$$

$$\hat{\sigma}^2_b = \frac{1}{K-1} \sum_{k=1}^{K} \left(\bar{X}^f_k - I^b_N(f) \right)^2$$

estimates the variance of batch samples, which is $\frac{\sigma^2}{b}$

I compute the confidence interval using $I^b_N(f) \pm 1{,}96 \cdot \sqrt{\frac{\hat{\sigma}^2_b}{K}}$

what about b and K (typically $b \approx 500$/1000 or more)

$\hat{}$ estimate from a trial run using $\hat{\tau}^*$

$K \geq 30$

Simulated Annealing