PARAMETER ESTIMATION

M: MODEL

$\theta = (\theta_1, \theta_2, \dots, \theta_N)$ PARAMETERS $\qquad \theta$ IS CONTINUOUS $\qquad N$: DIMENSIONALITY

d: DATA

WANT:

POSTERIOR $\qquad$ PRIOR, $\pi(\theta)$ $\qquad$ LIKELIHOOD, $\mathcal{L}(\theta)$

$$P(\theta|d,M) = \frac{P(\theta|M)\; P(d|\theta,M)}{P(d|M)}$$

(BAYES THEOREM)

$P(d|M) \leftarrow$ EVIDENCE/MARGINAL LIKELIHOOD

$$= \frac{P(\theta|M)\; P(d|\theta,M)}{\int d\theta'\, P(\theta'|M)\, P(d|\theta',M)}$$

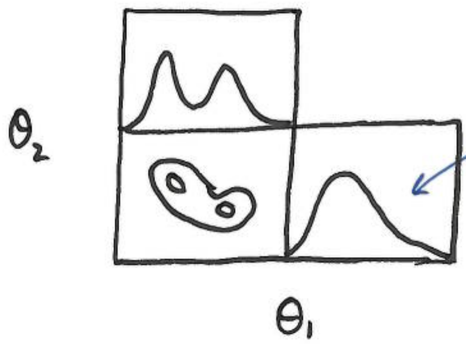
(LAW OF TOTAL PROBABILITY)

$$\propto f(\theta) \qquad f(\theta) \equiv P(\theta|M)\; P(d|\theta,M)$$

1. $f(\theta) \geq 0 \;\; \forall\, \theta$ $\qquad$ 2. $0 < \int d\theta'\, f(\theta') < \infty$

WHAT DO WE WANT FROM $P(\theta|d,M)$ OR $f(\theta)$?

- PARAMETER ESTIMATES: $\hat{\theta} \stackrel{?}{=} \int d\theta\, \theta\, P(\theta|d,M)$
- UNCERTAINTIES: $\sigma = \left[\int d\theta\, (\theta - \hat{\theta})^2\, P(\theta|d,M)\right]^{1/2}$
- CORNER PLOTS:



$$P(\theta_1|d,M) = \int d\theta_2\, d\theta_3 \dots d\theta_N\;\; P(\theta|d,M)$$

LOW-DIMENSIONAL PROBLEMS

$N \lesssim 3 \qquad \theta = (\theta_1, \theta_2, \theta_3) \qquad f(\theta) = P(\theta|M)\, P(d|\theta, M)$

EXAMPLE: $\theta_1 = F$: FRACTION OF STARS WHICH ARE BINARIES

$N=2$. $\theta_2 = m$: AVERAGE MASS OF THE STARS.

GRID-BASED METHOD.

ALGORITHM:

$0 \leq F \leq 1$ STELLAR ASTROPHYSICS

1. CHOOSE F_{min}, F_{max} $\quad m_{min} \quad m_{max}$

2. CHOOSE RESOLUTION: $I \times J \qquad I, J \simeq 100$

3. $$F_i = F_{min} + \frac{i + 1/2}{I+1} (F_{max} - F_{min})$$ OPTIONAL

   m_j similar.

4. $$f_{ij} = f(\theta_{ij}) = f(F_i, m_j)$$

5. NORMALISE:

$$I \simeq (F_{max} - F_{min})(m_{max} - m_{min}) \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} f_{ij}$$

$$p_{ij} \equiv \frac{f_{ij}}{I} \qquad p_{ij} \simeq P(\theta_{ij} | d, M) = P(F_i, m_j | d, M)$$

CALCULATE DERIVED QUANTITIES

$$P(F_i | d, M) = \sum_{j=1}^{J} p_{ij}$$

GENERAL:

$$\int d\theta\, g(\theta)\, P(\theta | d, M) \simeq \sum_{i=1}^{I} \sum_{j=1}^{J} g(\theta_{ij})\, p_{ij}$$

SAMPLING

$$f(\theta) \equiv P(\theta|M)\, P(d|\theta, M)$$

CHARACTERISING A DISTRIBUTION BY A FUNCTION

$$\theta_k \sim P(\theta|d,M) \qquad k = 1, 2, \ldots, K \qquad K \simeq 10^{3-4}$$

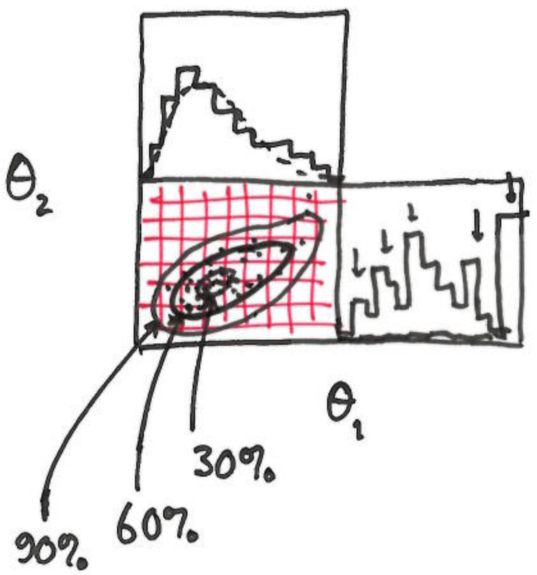
MONTE CARLO INTEGRATION

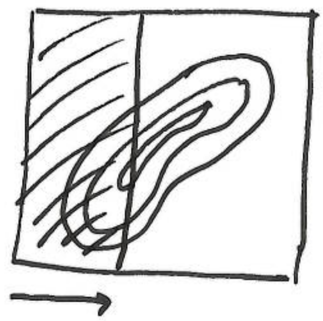
$$\int d\theta\, P(\theta|d,M)\, g(\theta) \simeq \frac{1}{K}\sum_{k=1}^{K} g(\theta_k) \quad \text{IF } \theta_k \sim P(\theta|d,M)$$

UNCERTAINTY IN ESTIMATE OF INTEGRAL GOES AS $K^{-1/2}$.

- NORMALISATION: $g(\theta) = 1 \qquad \frac{1}{K}\sum_{k=1}^{K} 1 = 1.$
- MEAN: $g(\theta) = \theta \qquad \frac{1}{K}\sum_{k=1}^{K} \theta_k = \hat{\theta}$
- COVARIANCE MATRIX: $g(\theta) = (\theta - \hat{\theta})(\theta - \hat{\theta})^T$

$$\Sigma = \frac{1}{K}\sum_{k=1}^{K} (\theta_k - \hat{\theta})(\theta_k - \hat{\theta})^T.$$





METROPOLIS ALGORITHM

HAVE $f(\theta) = P(\theta|M)\,P(d|\theta, M)$

ALGORITHM:

CHOOSE $K. \simeq 10^{3-4}$
CHOOSE θ_0 ; EVALUATE $f_0 = f(\theta_0)$ } INITIALISE

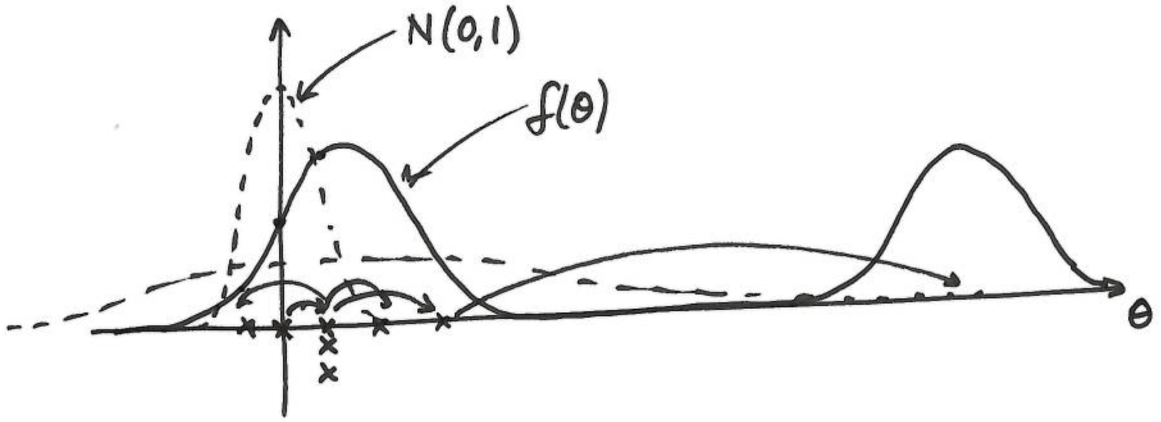
FOR $k = 1, 2, \dots K$:
{
$\theta_* \sim P(\theta_* | \theta_{k-1})$; EVALUATE $f_* = f(\theta_*)$ } PROPOSE NEW POINT

$N(\theta_*; \theta_{k-1}, \mathbb{1})$

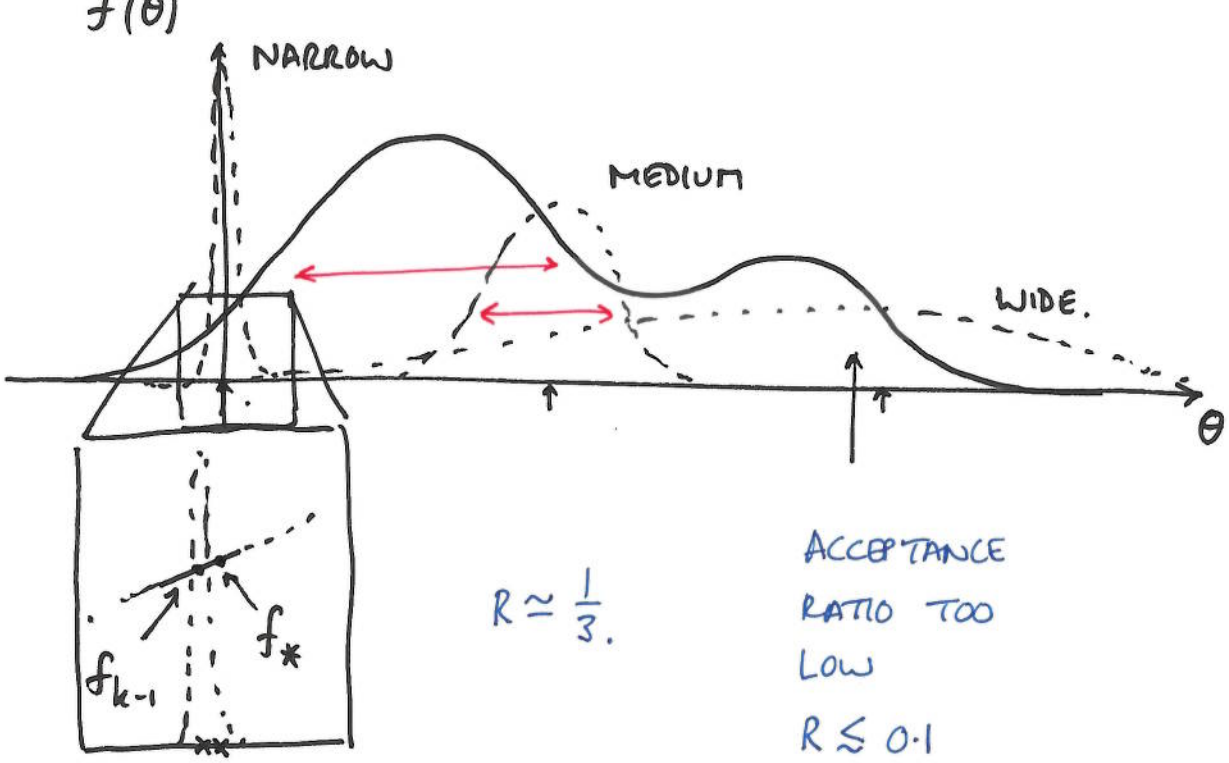
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$p \sim U(0,1)$
if $p \leq f_* / f_{k-1}$
  $\theta_k = \theta_*$; $f_k = f_*$
else
  $\theta_k = \theta_{k-1}$
}

ACCEPT/REJECT; SATISFIES DETAILED BALANCE



PROPOSAL DISTRIBUTION



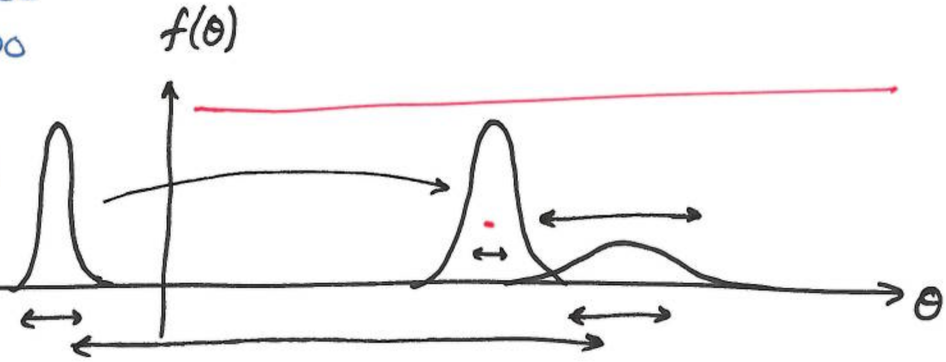
$R \simeq \frac{1}{3}.$

ACCEPTANCE RATIO TOO LOW

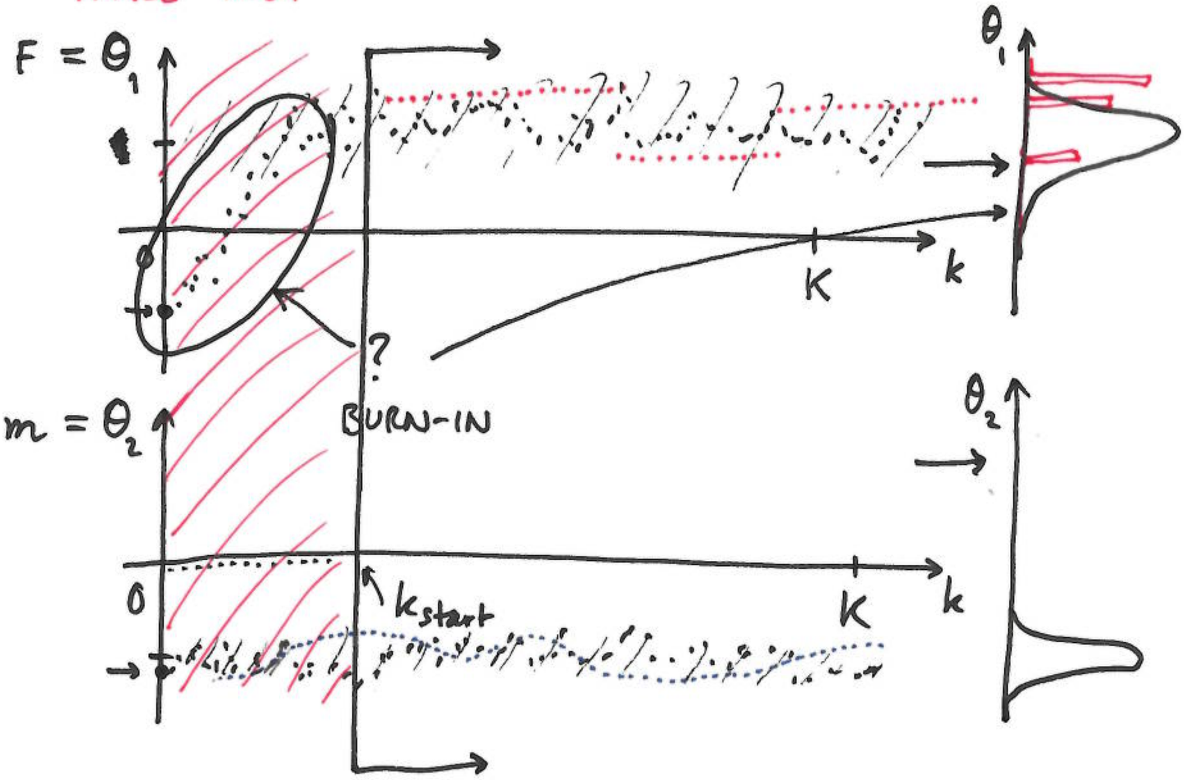
$R \lesssim 0.1$

ACCEPTANCE RATIO TOO HIGH.

$R \gtrsim 0.9$



TRACE PLOT



$\theta_1 \; \theta_2 \; \theta_3 \; \dots \; \theta_K$

$f_1 \; f_2 \; f_3 \; \dots \; f_k.$

k_{start} : FIRST k SUCH THAT $f_{k_{start}} \gtrsim 0.1 \times \max(\{f_k\})$

- AUTO-CORRELATION FUNCTION OF THE CHAIN
- EMPIRICAL COVARIANCE MATRIX OF THE CHAIN : $\hat{\underline{\underline{\Sigma}}}$
  ↳ PROPOSAL DISTRIBUTION $P(\theta_* | \theta_{k-1}) = N(\mu = \theta_{k-1}, \hat{\underline{\underline{\Sigma}}})$