Linear Gaussian

Basic Model

$$\mathrm{x}_{t+1} = \mathrm{A}\mathrm{x}_t + \mathrm{w}_t = \mathrm{A}\mathrm{x}_t + \mathrm{w}_\bullet \qquad \mathrm{w}_\bullet \sim \mathcal{N}(0, \mathrm{Q})$$
$$\mathrm{y}_t = \mathrm{C}\mathrm{x}_t + \mathrm{v}_t = \mathrm{C}\mathrm{x}_t + \mathrm{v}_\bullet \qquad \mathrm{v}_\bullet \sim \mathcal{N}(0, \mathrm{R})$$

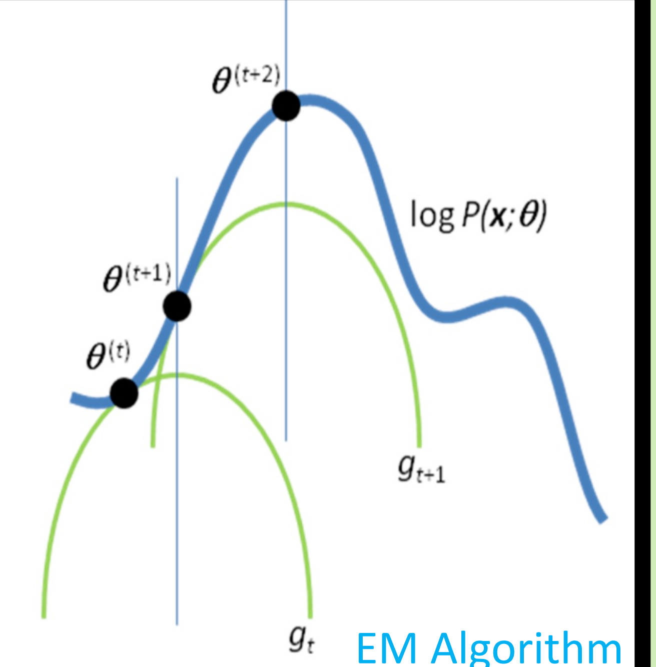
EM algorithm: Learning

increase

$$\begin{aligned}
\log \int_{\mathbf{X}} P(\mathbf{Y}, \mathbf{X} \mid \theta) d\mathbf{X} &= \log \int_{\mathbf{X}} Q(\mathbf{X}) \frac{P(\mathbf{X}, \mathbf{Y} \mid \theta)}{Q(\mathbf{X})} d\mathbf{X} \\
&\geq \int_{\mathbf{X}} Q(\mathbf{X}) \log \frac{P(\mathbf{X}, \mathbf{Y} \mid \theta)}{Q(\mathbf{X})} d\mathbf{X} \\
&= \int_{\mathbf{X}} Q(\mathbf{X}) \log P(\mathbf{X}, \mathbf{Y} \mid \theta) d\mathbf{X} - \int_{\mathbf{X}} Q(\mathbf{X}) \log Q(\mathbf{X}) d\mathbf{X} \\
&= \mathcal{F}(Q, \theta)
\end{aligned}$$

E step $\qquad Q_{k+1} \leftarrow \underset{Q}{\arg\max} \mathcal{F}(Q, \theta_k)$

M step $\qquad \theta_{k+1} \leftarrow \underset{\theta}{\arg\max} \mathcal{F}(Q_{k+1}, \theta)$



EM Algorithm

Filtering & smoothing

$$P(\mathbf{x}_t \mid \{\mathbf{y}_1, \ldots, \mathbf{y}_t\})$$
$$P(\mathbf{x}_t \mid \{\mathbf{y}_1, \ldots, \mathbf{y}_\tau\})$$

Gaussian

Basic Model

$$
\begin{aligned}
\mathrm{x}_{t+1} &= \mathrm{A}\mathrm{x}_t + \mathrm{w}_t = \mathrm{A}\mathrm{x}_t + \mathrm{w}_\bullet \qquad \mathrm{w}_\bullet \sim \mathcal{N}(0, \mathrm{Q}) \\
\mathrm{y}_t &= \mathrm{C}\mathrm{x}_t + \mathrm{v}_t = \mathrm{C}\mathrm{x}_t + \mathrm{v}_\bullet \qquad \mathrm{v}_\bullet \sim \mathcal{N}(0, \mathrm{R})
\end{aligned}
$$

Static data modelling (variants)

$$
\begin{aligned}
\mathrm{x}_\bullet &= \mathrm{w}_\bullet \qquad \mathrm{w}_\bullet \sim \mathcal{N}(\mathbf{0}, \mathrm{I}) \\
\mathrm{y}_\bullet &= \mathrm{C}\mathrm{x}_\bullet + \mathrm{v}_\bullet \qquad \mathrm{v}_\bullet \sim \mathcal{N}(\mathbf{0}, \mathrm{R})
\end{aligned}
$$

Factor Analysis

$$
R = \begin{bmatrix} r_1 & 0 & 0 & \dots & 0 \\ 0 & r_2 & 0 & \dots & 0 \\ 0 & 0 & r_\mathrm{m} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & r_\mathrm{s} \end{bmatrix}
$$

SPCA

$$\mathbf{R} = \epsilon \mathbf{I}$$

PCA

$$\mathbf{R} = \lim_{\epsilon \to 0} \epsilon \mathbf{I}$$

Time-series modelling

Kalman Filter

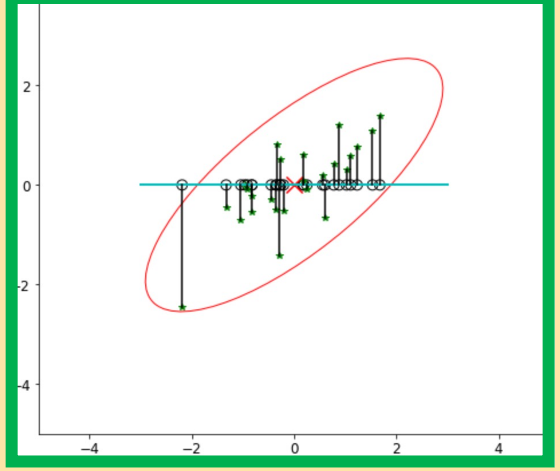
$$
\begin{aligned}
\mathbf{x}_{t+1} &= \mathbf{A}\mathbf{x}_t + \mathbf{w}_t = \mathbf{A}\mathbf{x}_t + \mathbf{w}_\bullet \qquad \mathbf{w}_\bullet \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}) \\
\mathbf{y}_t &= \mathbf{C}\mathbf{x}_t + \mathbf{v}_t = \mathbf{C}\mathbf{x}_t + \mathbf{v}_\bullet \qquad \mathbf{v}_\bullet \sim \mathcal{N}(\mathbf{0}, \mathbf{R})
\end{aligned}
$$

PCA

$$x_\bullet = w_\bullet \qquad w_\bullet \sim \mathcal{N}(\mathbf{0}, I)$$

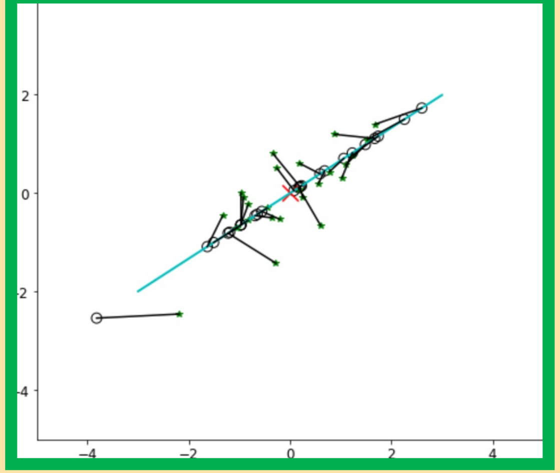
$$y_\bullet = Cx_\bullet + v_\bullet \qquad v_\bullet \sim \mathcal{N}(\mathbf{0}, R) \qquad \mathbf{R} = \lim_{\epsilon \to 0} \epsilon \mathbf{I}$$

E step



Init: C = [1,0]
Generate X

M step



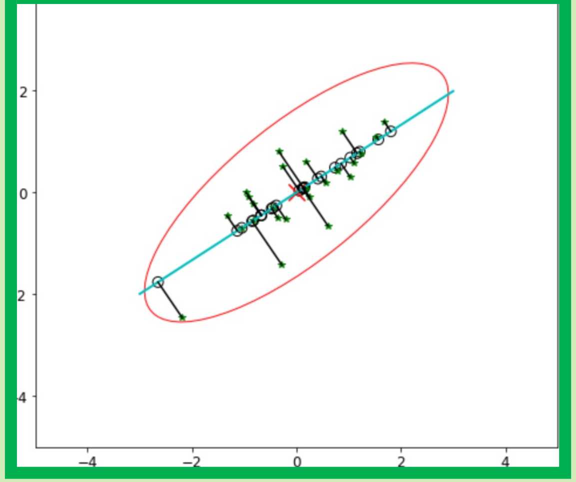
M: update C

PCA Simulation

E: $\mathbf{X} = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{Y}$

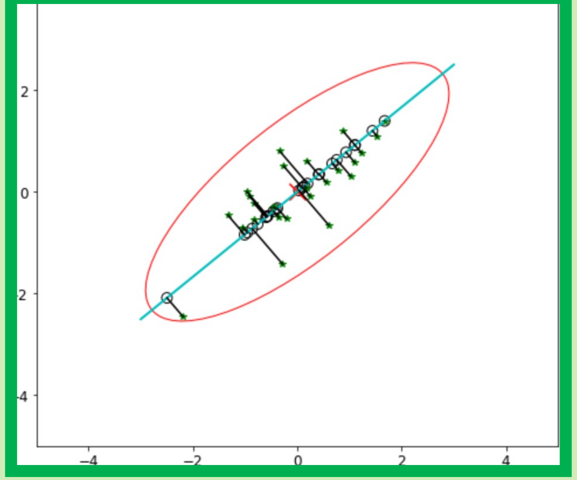
M: $\mathbf{C}^{\text{new}} = \mathbf{Y}\mathbf{X}^T(\mathbf{X}\mathbf{X}^T)^{-1}$

E step



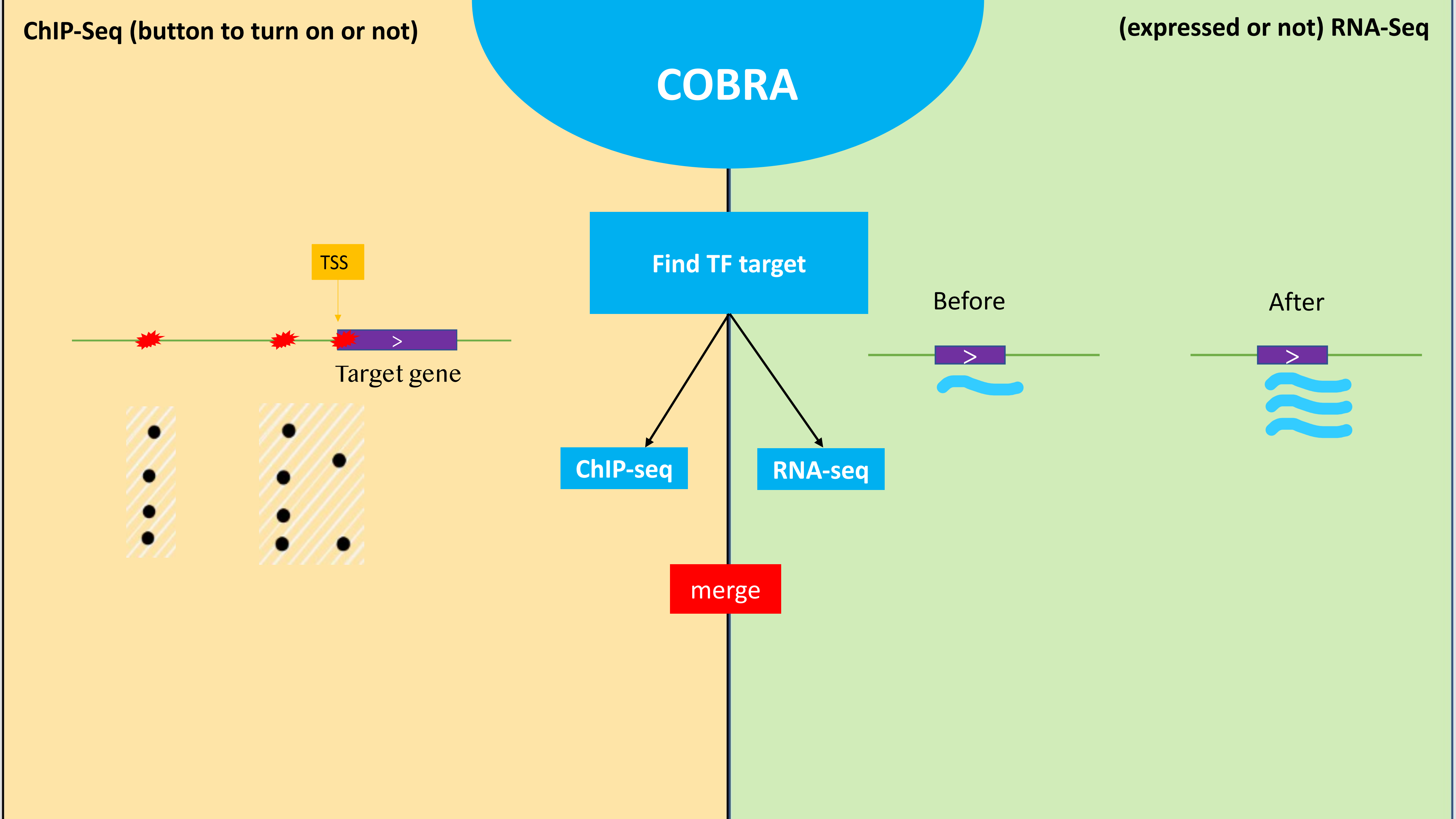
E: update X

M step

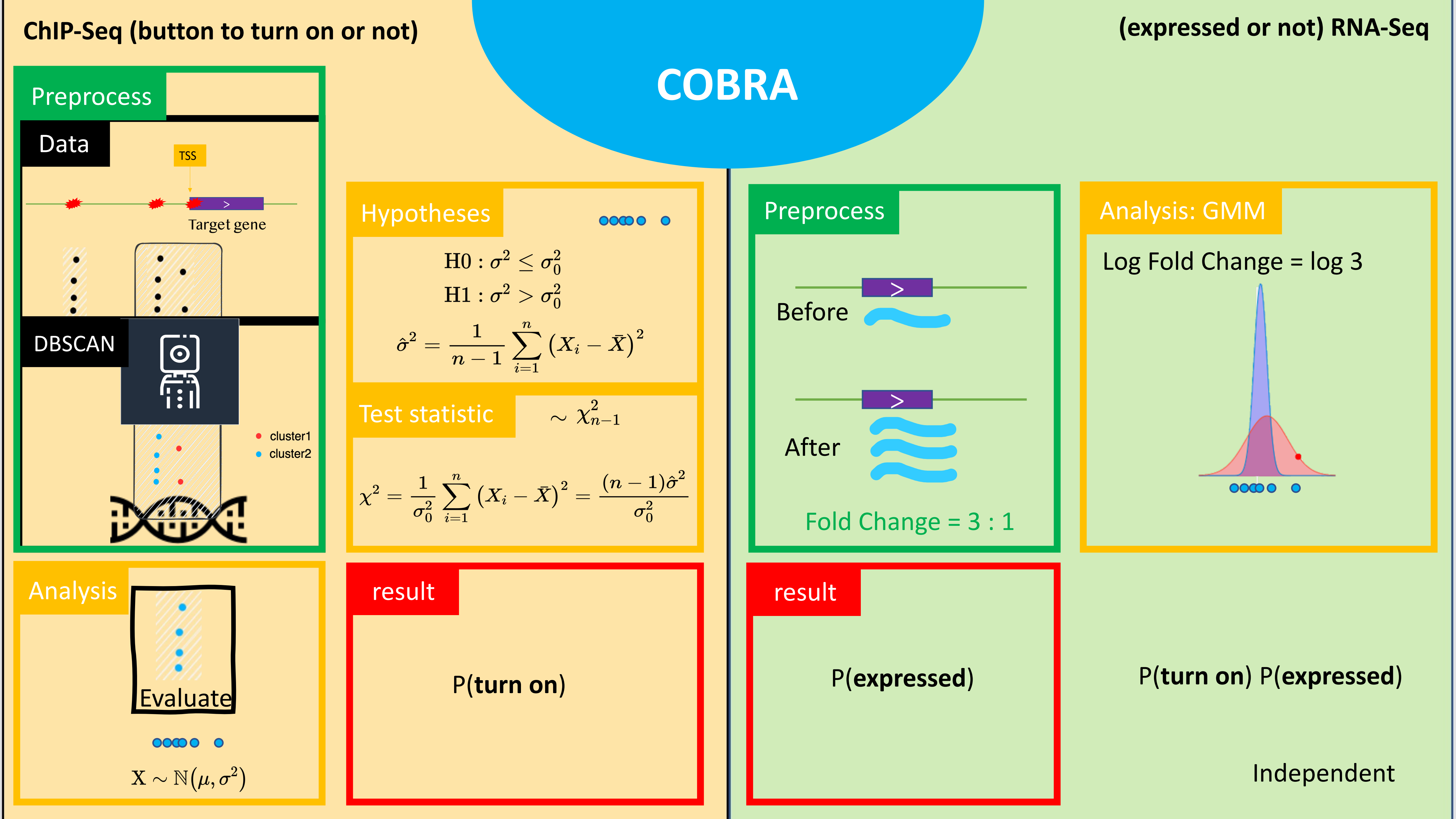


M: update C

code

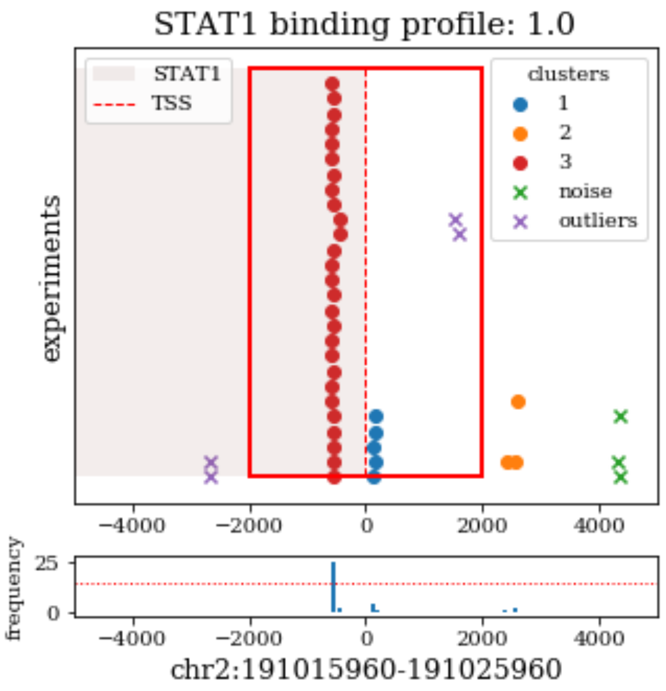


ChIP-Seq (button to turn on or not)
COBRA
(expressed or not) RNA-Seq
TSS
Find TF target
Before
After
Target gene
ChIP-seq
RNA-seq
merge

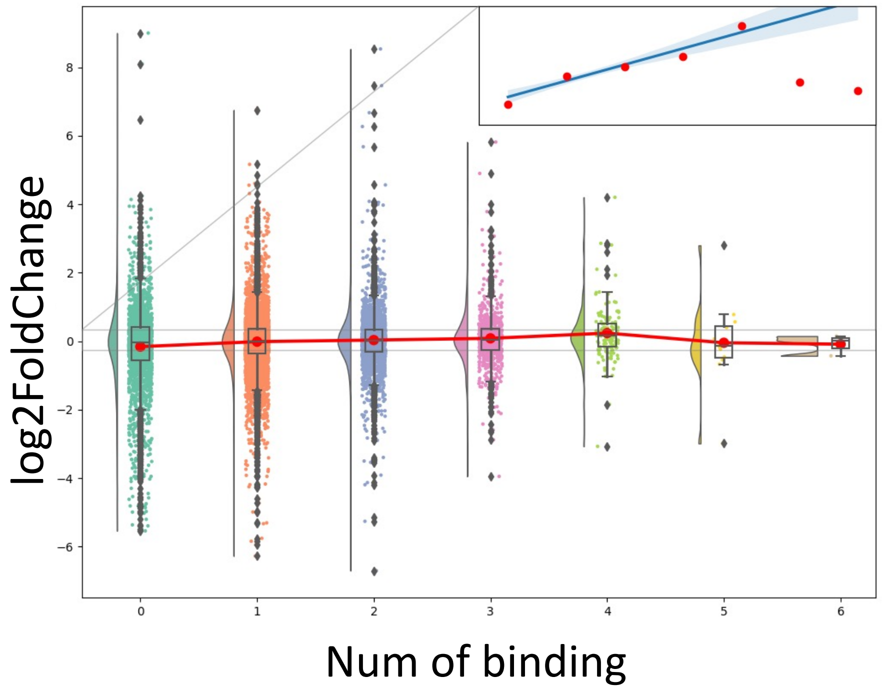

ChIP-Seq (button to turn on or not)
COBRA
(expressed or not) RNA-Seq
Preprocess
Data
TSS
Target gene
DBSCAN
cluster1
cluster2
Hypotheses
$H0 : \sigma^2 \leq \sigma_0^2$
$H1 : \sigma^2 > \sigma_0^2$
$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$
Test statistic
$\sim \chi^2_{n-1}$
$\chi^2 = \frac{1}{\sigma_0^2} \sum_{i=1}^{n} (X_i - \bar{X})^2 = \frac{(n-1)\hat{\sigma}^2}{\sigma_0^2}$
Preprocess
Before
After
Fold Change = 3 : 1
Analysis: GMM
Log Fold Change = log 3
Analysis
Evaluate
$X \sim \mathbb{N}(\mu, \sigma^2)$
result
P(**turn on**)
result
P(**expressed**)
P(**turn on**) P(**expressed**)
Independent

COBRA

ChIP-Seq



ChIP-Seq + RNA-Seq



Comparison