Multivariate Self-Starting Shiryaev (M3S)

Konstantinos Bourazas[†] and Panagiotis Tsiamyrtzis[*] [†]

[†]Dept. of Statistics, Athens University of Economics and Business, [*]Dept. of Mechanical Engineering, Politecnico di Milano

kbourazas@aueb.gr, panagiotis.tsiamyrtzis@polimi.it

A brief description

- **3S** is a Bayesian on-line change point detection scheme, generalizing Shiryaev's procedure (1963) and providing posterior inference about the parameters of interest.
- **M3S** is the multivariate extension of 3S, aiming in detecting persistent shifts in the mean vector or the covariance matrix of multivariate Normal data.

3S scheme

- $\boldsymbol{x} = (x_1, ..., x_n)$ is a random sample.
- $\boldsymbol{\theta}$ are the **In Control (IC)** parameters.
- $\boldsymbol{\phi}$ are the **Out Of Control (OOC)** parameters.
- $g(\boldsymbol{\theta}, \boldsymbol{\phi})$ is the link function of the OOC scenario.
- τ is the unknown change point.

The likelihood is:

$$f(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{\phi},\tau) = \begin{cases} f(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{\phi},\tau \le n) = \prod_{i=1}^{\tau-1} f(x_i|\boldsymbol{\theta}) \prod_{i=\tau}^{n} f(x_i|g(\boldsymbol{\theta},\boldsymbol{\phi})) & \text{if } \tau \le n \\ f(\boldsymbol{x}|\boldsymbol{\theta},\tau > n) = \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta}) & \text{if } \tau > n \end{cases}$$

The stopping time is (constant decision limit):

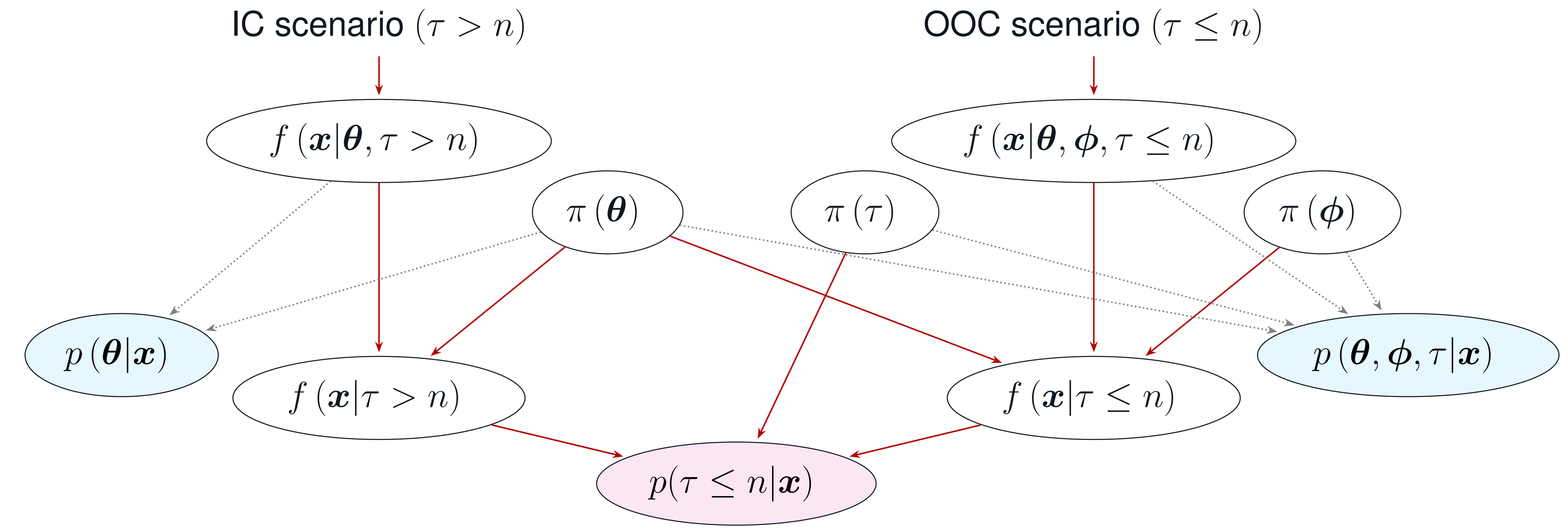
$$T(p^*) = inf\{n \ge 1 : p(\tau \le n|\boldsymbol{x}) \ge p^*\}, \text{ where}$$

$$p(\tau \le n|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\tau \le n)\pi(\tau \le n)}{f(\boldsymbol{x}|\tau \le n)\pi(\tau \le n) + f(\boldsymbol{x}|\tau > n)\pi(\tau > n)}$$

The marginal distributions involved in the computation are:

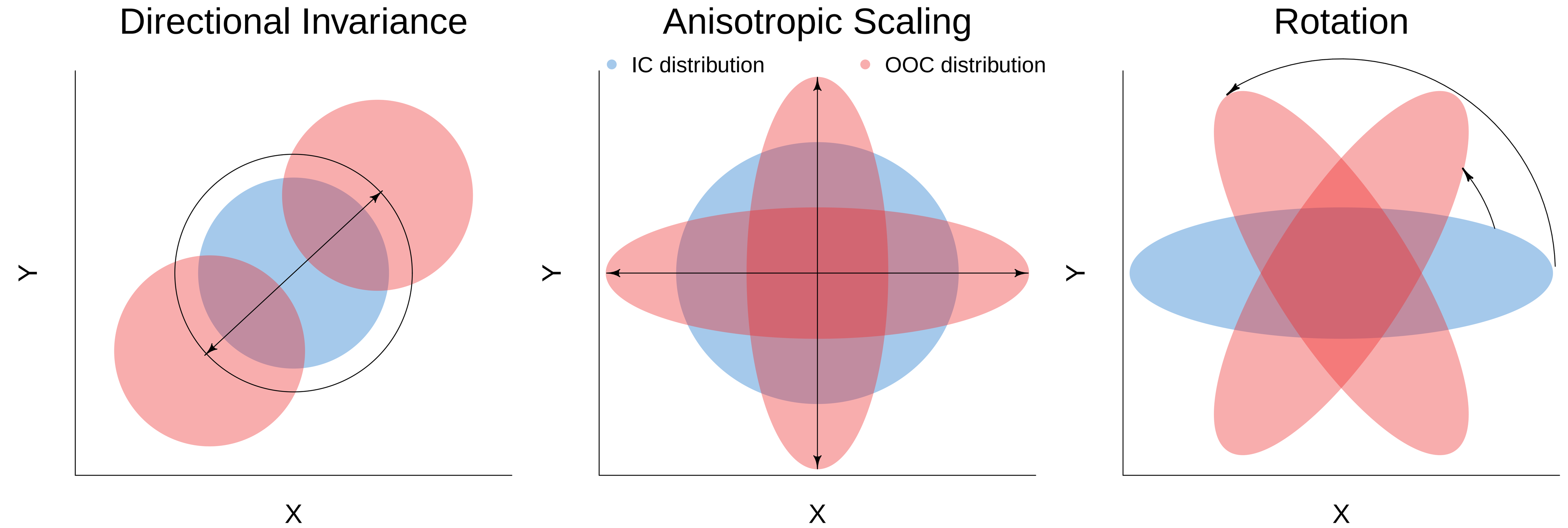
$$f(\boldsymbol{x}|\tau > n) = \int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta},\tau > n)\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$$

$$f(\boldsymbol{x}|\tau \le n) = \int_{\Phi}\int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{\phi},\tau \le n)\pi(\boldsymbol{\theta})\pi(\boldsymbol{\phi})d\boldsymbol{\theta}d\boldsymbol{\phi}$$



Desired properties for multivariate detection models

- **Directional invariance**: the detection of changes in any direction.
- **Anisotropic scaling**: the changes are weighted by the covariance matrix.
- **Rotation**: detection of changes in the correlation between the variables.



M3S - Mean vector model

Assume $\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_n)$, where $\boldsymbol{X}_i|(\boldsymbol{\mu},\boldsymbol{\Sigma}) \overset{iid}{\sim} N_D(\boldsymbol{\mu},\boldsymbol{\Sigma})$

$\pi(\boldsymbol{\mu},\boldsymbol{\Sigma}) \propto L(\boldsymbol{\mu},\boldsymbol{\Sigma}|\boldsymbol{Y})^{\alpha_0}\pi_0(\boldsymbol{\mu},\boldsymbol{\Sigma})$, where
$\boldsymbol{Y} = (y_1, ..., y_{n_0})$ are the historical data
$0 \le \alpha_0 \le 1$ controls the influence of $\boldsymbol{Y}$ of the power prior,
$\pi_0(\boldsymbol{\mu},\boldsymbol{\Sigma}) = NIW(\boldsymbol{\mu}_0,\lambda,\nu,\boldsymbol{\Psi})$ is the initial prior.

$\tau \sim DiscreteW(p, b)$, where
τ is the location of a potential change point,
p is the probability for an observation to be OOC,
b represents the wear out effect.

$g(\boldsymbol{\mu},\boldsymbol{\Sigma}) = \boldsymbol{\mu} + r\boldsymbol{L}^{1/2}\boldsymbol{T}_\theta$, where

$$\boldsymbol{L} = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_D^2 \end{pmatrix}, \boldsymbol{T}_\theta = \begin{pmatrix} cos\theta_1 \\ \vdots \\ cos\theta_{D-1}\prod_{i<D-1} sin\theta_i \\ \prod_{i=1}^{D-1} sin\theta_i \end{pmatrix}$$

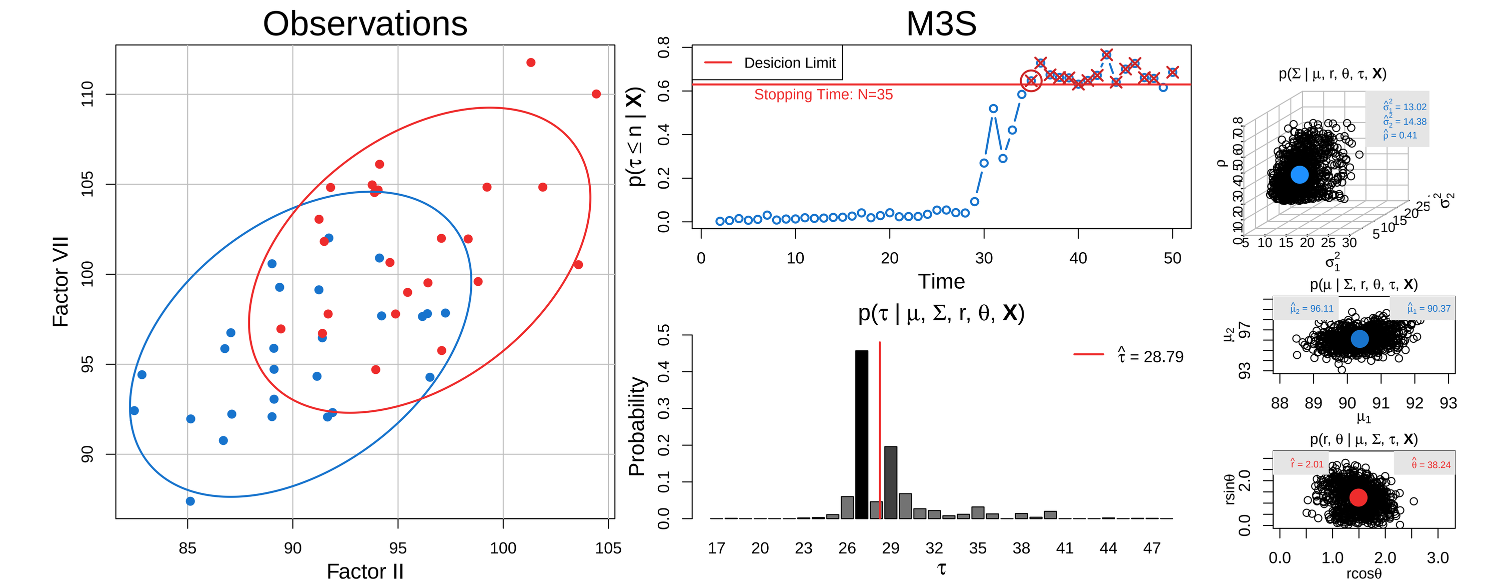
$r \sim NC_{\chi_D}(d)$ is the radius of a potential shift,
$\boldsymbol{T}_\theta \sim vMF(\boldsymbol{\mu}_\theta, \kappa)$ represents the angles of a potential shift.

Application to Real Data

We have $n = 50$ medical data (blood-clotting related) of **Factor II** (Thrombin) and **Factor VII** (Proconvertin).

$$(\boldsymbol{\mu},\boldsymbol{\Sigma})|(\boldsymbol{\mu}_0,\lambda,\nu,\boldsymbol{\Psi},\boldsymbol{Y},\alpha_0) \sim NIW\left(\begin{pmatrix}90.1\\94.9\end{pmatrix}, 13, 13, \begin{pmatrix}187.6 & 80.4\\80.4 & 177.4\end{pmatrix}\right)$$

$r \sim NC_{\chi_D}(\sqrt{2})$, $\theta \sim U(0, 2\pi)$, $\tau \sim DW(10^{-3}, 1.1)$, $PFA \le 5\%$



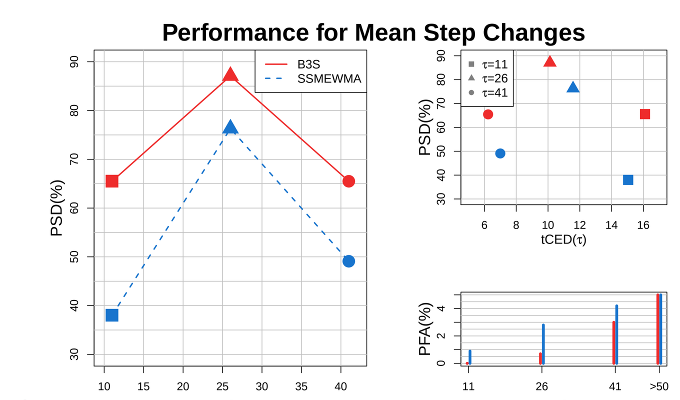
Simulation Study

Competing methods: SSMEWMA (2007), M3S

IC data: $\boldsymbol{X}_i|(\boldsymbol{\mu},\boldsymbol{\Sigma}) \sim N_2\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right)$

OOC data: $\boldsymbol{X}_i|(\boldsymbol{\mu}',\boldsymbol{\Sigma}) \sim N_2\left(\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right)$

change point: $\{11, 26, 41\}$



Performance Metrics: $PFA = P(T \le n|\tau > n) = 0.05$, $PSD(\tau) = P(\tau \le T \le n)$, $tCED(\tau) = E_\tau(T - \tau + 1|\tau \le T \le n)$

Conclusions

- M3S is a **multivariate generalization** of Shiryaev procedure, allowing prior flexibility, providing posterior inference and achieving the desired detection properties.
- It achieves greater PSD and it is more resistant in absorbing a shift.

References

Hawkins D. & Maboudou-Tchao E. (2007). "Self-Starting Multivariate Exponentially Weighted Moving Average Control Charting" *Technometrics*, Vol 49, pp 199-209

Shiryaev A. (1963). "On optimum methods in quickest detection problems", *Theory of Probability & Its Applications*, Vol 8, No 1, pp 22-46