Détection/localisation en présence de paramètres de nuisance

Igor Nikiforov

ICD/LM2S/UTT, FRE CNRS 2848

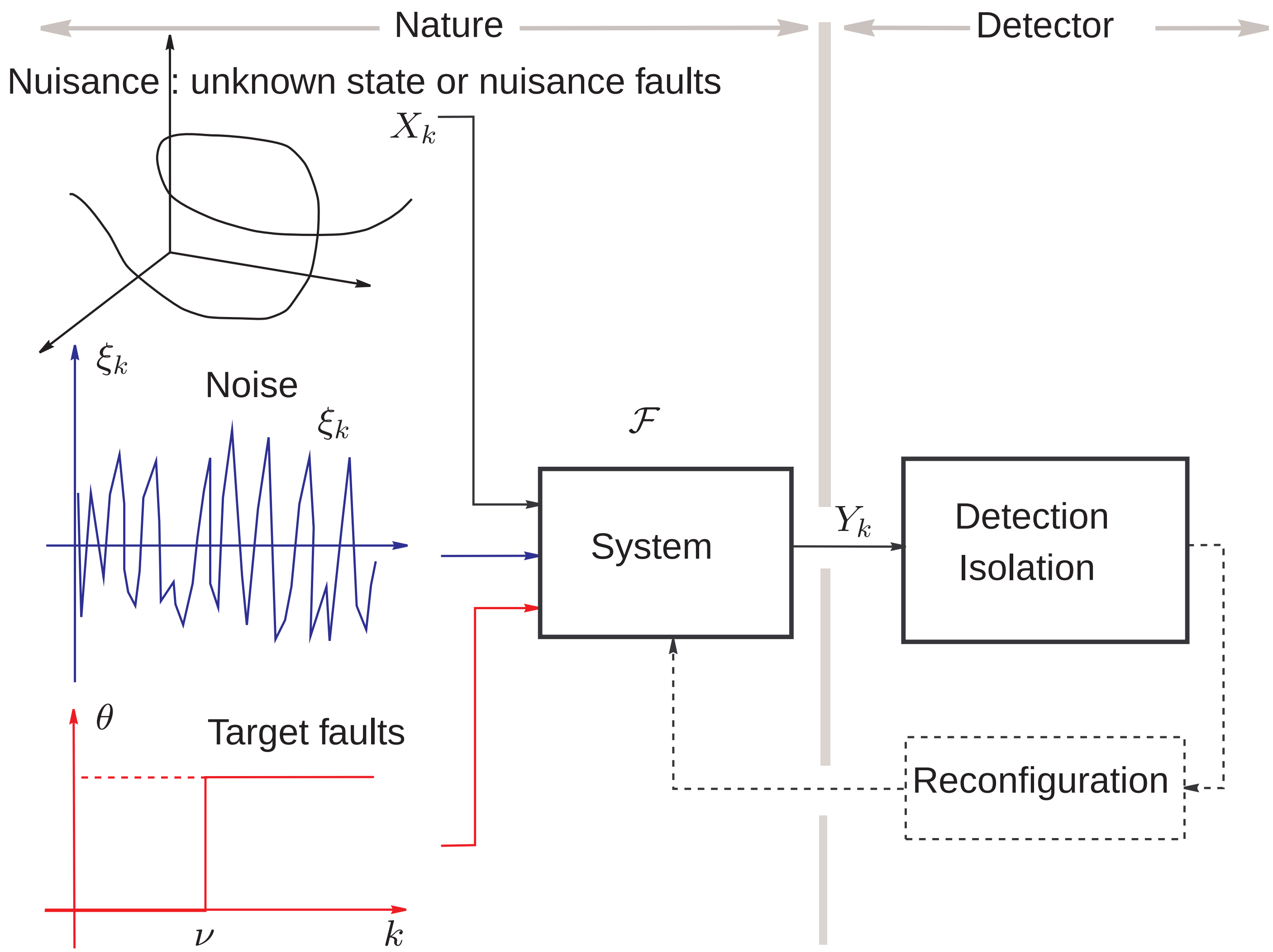
Réunion de S3 à l'ENSAM, Paris, 15 mars 2006

Plan

1. Statistical tests for monitoring : hypotheses testing and change detection/isolation
2. Hypotheses testing (with nuisance parameters)
3. Example : anomaly/target detection from a few tomographic projections
4. Change detection and detection/isolation (with nuisance parameters)
5. Examples :
   (a) SIRU integrity monitoring
   (b) GPS/Galileo navigation system integrity monitoring
6. Conclusions and perspectives

The functions of system monitoring

Model of system : $Y_k = \mathcal{F}(X_k, U_k, \theta_k, \xi_k)$



Two change situations

The two following situations are distinguished :

Hypotheses testing : the parameter vector θ is assumed to be *constant* within the entire data sample $Y_1, \ldots, Y_n$.

Change detection/isolation : the parameter θ can *change* within the data sample at an unknown instant (change point) k_0 $(1 \leq k_0 \leq n)$.

Hypotheses testing : problem statement

A (fixed) n-size sample of independent observations $Y_1, \ldots, Y_n$ is available and supposed to be generated by one among $(K+1)$ probability distributions $P_0, \ldots, P_K$. Possible known inputs $U_1, \ldots, U_n$ are assumed non random.

The hypotheses testing problem consists in deciding which distribution P_i (hypothesis $\mathcal{H}_i$) is the true one.

A *statistical test* for testing between the $\mathcal{H}_i$'s is any measurable mapping $\delta : (\mathcal{Y}, \mathcal{U}) \to \{\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_K\}$ from the observation space onto the set of hypotheses.

Hypotheses testing : criteria

The quality of a statistical test is defined with a set of *error probabilities* :

$$\alpha_i = \mathbb{P}_i\left(\delta \neq \mathcal{H}_i\right), \quad i = 0, 1, \ldots, K,$$

where $\mathbb{P}_i$ stands for observations $Y_1, \ldots, Y_n$ being generated by distribution P_i.

Binary hypotheses case : $\mathcal{H}_0$ **against** $\mathcal{H}_1$

The *power* is defined with the probability of correct decision :

$$\beta = 1 - \alpha_1 = \mathbb{P}_1\left(\delta = \mathcal{H}_1\right).$$

The pair (α_0, β) is then a sufficient performance index.

Hypotheses testing problems with nuisance parameters

Detection. Detection refers to deciding whether the monitored system is in its nominal (safe) state or not :

$$\mathcal{H}_0 : \quad \theta \in \Theta_0, \ \overbrace{X \in \mathbb{R}^q}^{\text{nuisance}} \quad \text{against} \quad \mathcal{H}_1 \ : \quad \theta \in \Theta_0^c \stackrel{\text{def.}}{=} \mathbb{R}^m \setminus \Theta_0, \ \overbrace{X \in \mathbb{R}^q}^{\text{nuisance}}$$

When some more information about is available, $\mathcal{H}_1 \ : \quad \theta \in \Theta_1 \subset \Theta_0^c$.

Isolation. In case of two fault modes or more, isolation refers to deciding which fault mode occurred.

$$\mathcal{H}_0 \quad \text{against} \quad \mathcal{H}_i \ : \quad \phi \in \Phi_i \subset \mathbb{R}^m, \ \overbrace{X \in \mathbb{R}^q}^{\text{nuisance}} \ (i = 1, \ldots, K),$$

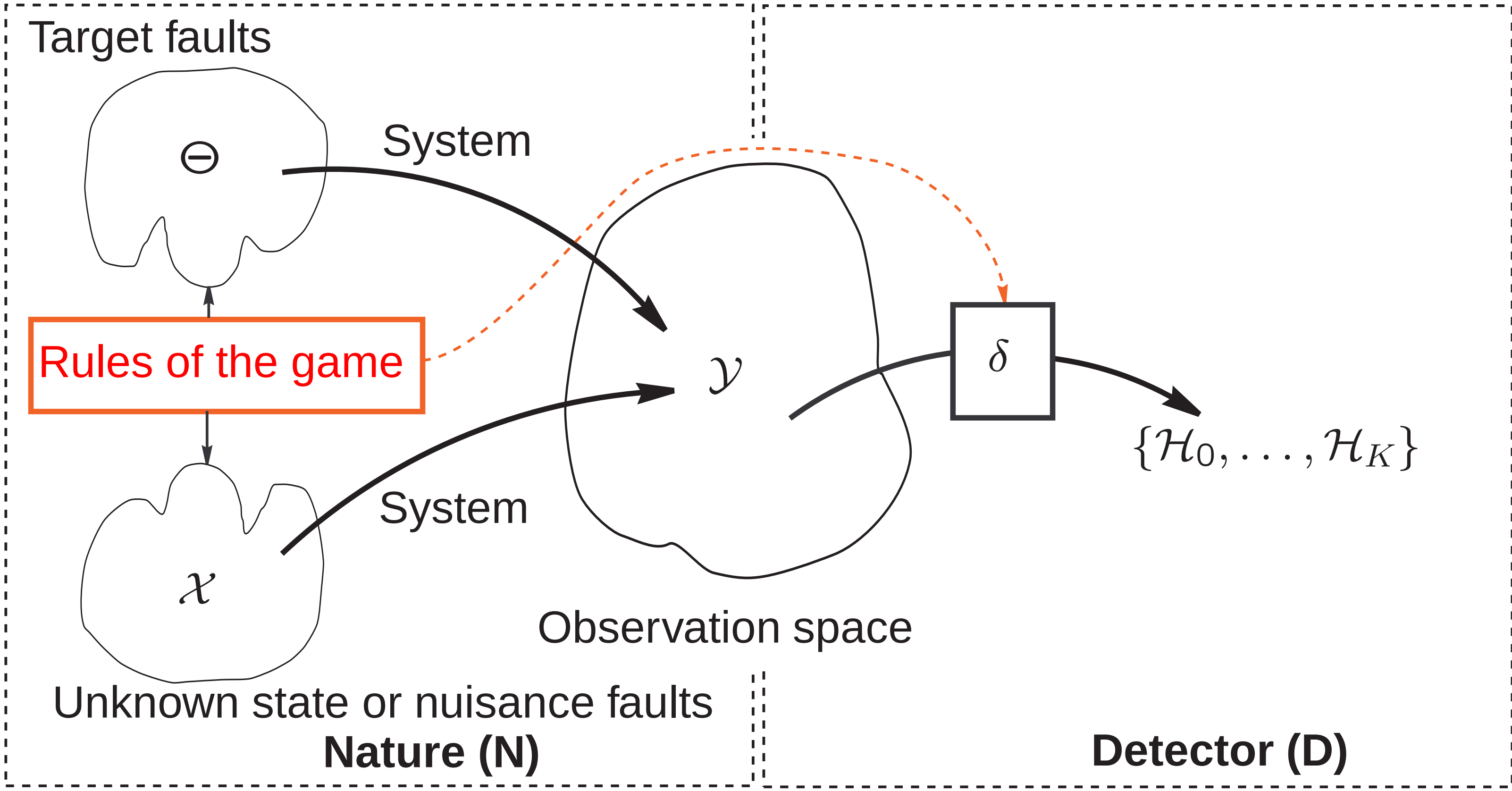
where $\Phi_i \bigcap \Phi_j = \emptyset$ for $i \neq j$.

Roles of informative and nuisance parameters

Statistical tests are functions of both the informative parameters ϕ and nuisance parameters X. The desirable relations between the error probabilities or the power of a test and the informative vector ϕ usually result from the application. Sometimes, the statistician must define some additional constraints (possibly artificial w.r.t. the application) resulting from the statistical nature of the problem, in order to achieve optimal properties of the test.

The main difference between ϕ and X is the following : in contrast to the informative parameter ϕ, the nuisance parameter X has no desirable impact on the performance indexes.

Bayesian approach : N is kind **[?]** $\Rightarrow$ D knows **[?]** the rules of the game

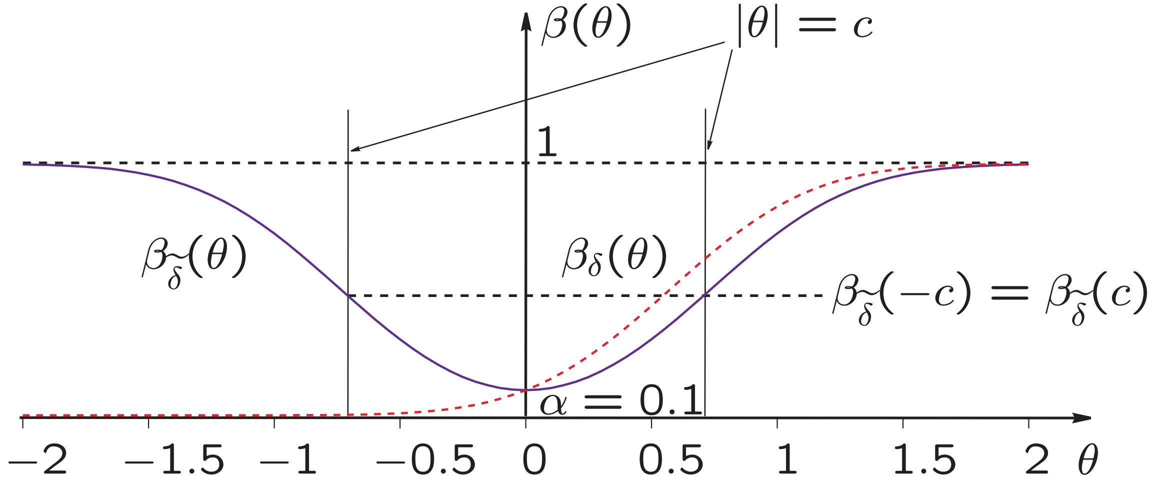


Sometimes the theory of bayesian tests is not relevant to deal with nuisance parameters, because the nuisance parameters can be intentionally chosen.

Constant power approach [Wald, 1943]

Conception of Wald : let us impose some additional constraints on the class $\mathcal{K}_\alpha = \left\{\delta : \sup_{\theta\in\Theta_0} \mathbb{P}_0(\delta \neq \mathcal{H}_0) \leq \alpha\right\}$, in order to avoid tests UMP over a subset $\overline{\Theta}_1$ of Θ_1, which are very inefficient over $\Theta_1 \setminus \overline{\Theta}_1$.

Example : It can be seen that the one-sided UMP test δ outperforms the unbiased UMP test $\widetilde{\delta}$, only for the positive alternatives ($\theta > 0$), but is very inefficient for other alternatives ($\theta < 0$).



Application to the gaussian linear model [Fouladirad and Nikiforov 2005]

A test $\delta^* \in \mathcal{K}_\alpha$ is said to be UMP with constant power over a family $\mathcal{S} = \{S_c\}$ of surfaces, if :

– $\forall \theta_1, \theta_2 \in S_c : \; \beta_{\delta^*}(\theta_1) = \beta_{\delta^*}(\theta_2)$

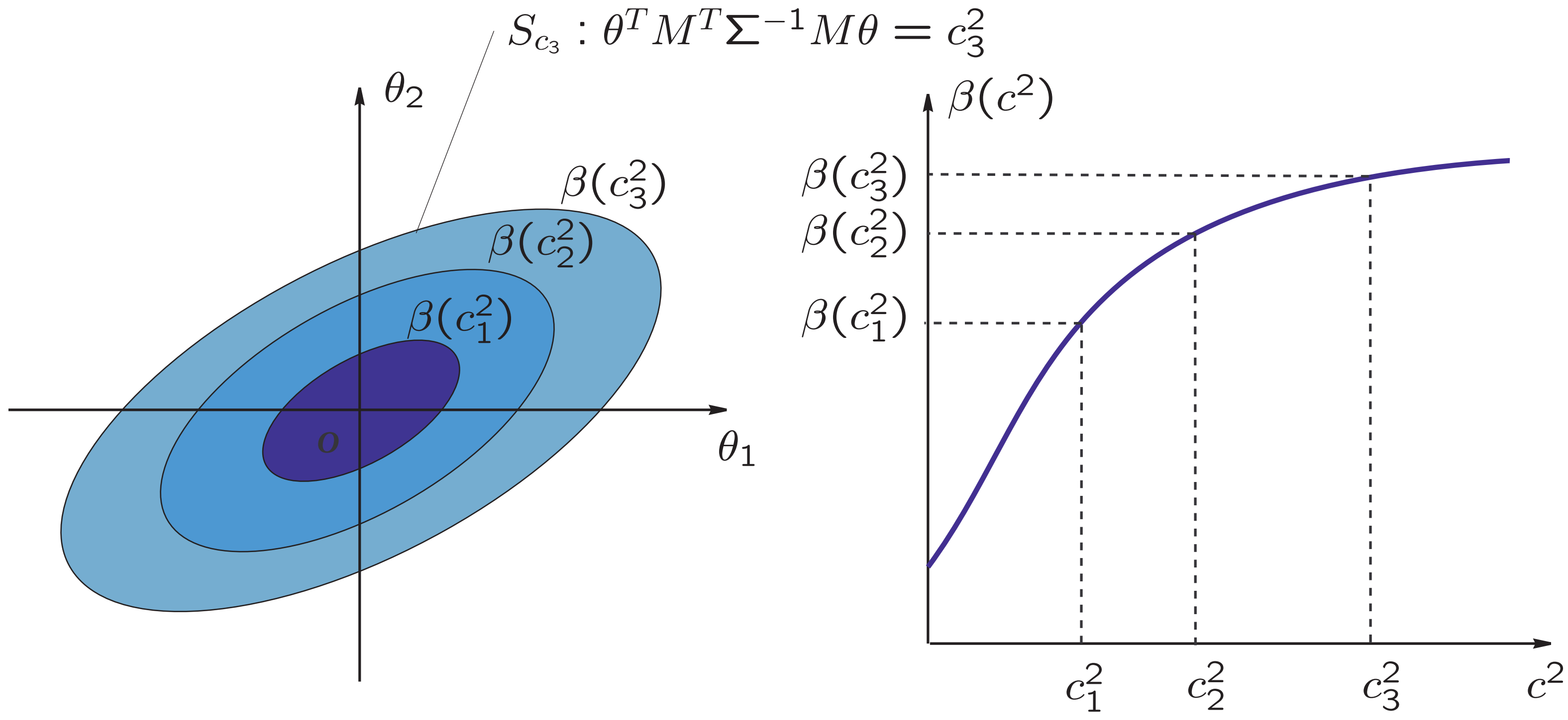
– $\forall \delta \in \mathcal{K}_\alpha$ s.t. $\forall \theta_1, \theta_2 \in S_c \; : \; \beta_\delta(\theta_1) = \beta_\delta(\theta_2)$, then : $\beta_{\delta^*}(\theta) \geq \beta_\delta(\theta)$

For the model $Y = M\theta + \xi$ with $\theta_0 = 0$, the surfaces S_c are ellipsoids :

$$S_c : \; \theta^T M^T \Sigma^{-1} M \theta = c^2 \; (c > 0)$$

and the test UMP with constant power writes as

$$\Lambda^* \stackrel{\text{def.}}{=} Y^T \Sigma^{-1} M (M^T \Sigma^{-1} M)^{-1} M^T \Sigma^{-1} Y$$ (coincides with the GLR)



$S_{c_3} : \theta^T M^T \Sigma^{-1} M\theta = c_3^2$
θ_2
θ_1
$\beta(c_3^2)$
$\beta(c_2^2)$
$\beta(c_1^2)$
o
$\beta(c^2)$
$\beta(c_3^2)$
$\beta(c_2^2)$
$\beta(c_1^2)$
c_1^2
c_2^2
c_3^2
c^2

Static and dynamic models with additive faults and nuisance parameters

The measured data Y are viewed as the output of a discrete time system [Basseville 1997]

$$\text{dynamic} \quad \left\{ \begin{array}{rcccccl} X_{k+1} & = & FX_k & + & GU_k & + & W_k \quad (+ \;\; \Gamma\, \Upsilon_x) \\ Y_k & = & HX_k & + & JU_k & + & V_k \quad (+ \;\; \Xi\, \Upsilon_y) \end{array} \right. ,$$

where $(W_k)_k \sim \mathcal{N}(0, Q_x)$, $(V_k)_k \sim \mathcal{N}(0, Q_y)$, Υ_x and Υ_y are the assumed additive faults, and the fault gains Γ and Ξ are full column rank matrices, or

$$\text{static} \quad Y = HX + \xi \; (+M\theta) \, , \quad \xi \sim \mathcal{N}(0, \Sigma)$$

where the matrix M is full column rank (f.c.r.), X is a nuisance parameter, $Y \in \mathbb{R}^r$, $\theta \in \mathbb{R}^m$, with $\text{rank}(H) + m < r$.

dynamic $\Rightarrow$ static

Method 1 : The problem is invariant to $G = \{Y \to g(Y) = Y + HX\}$

Model $Y = HX + \xi\ (+M\theta)$, $\xi \sim \mathcal{N}(0, \sigma^2 I_r)$. The fault θ detection problem consists in deciding between [Scharf and Friedlander 1994], [Nikiforov 2002], [Fouladirad and Nikiforov 2005]

$$\mathcal{H}_0 : \{\theta = 0,\ X \in \mathbb{R}^p\} \ \text{ and } \ \mathcal{H}_1 : \{\theta \neq 0,\ X \in \mathbb{R}^p\},$$

where Υ is the informative parameter and X is the nuisance parameter, while considering X as an *unknown* vector. The invariant test is based on *maximal invariant statistics*. Let us define the column space $R(H)$ of the matrix H. The standard solution is the projection of Y on the orthogonal complement $R(H)^\perp$ of the column space $R(H)$ (“parity space" in the analytical redundancy literature).

The parity vector $Z = WY$ is a maximal invariant to the group G.

$$WH = 0, \; W^T W = P_H = I_r - H(H^T H)^- H^T, \; WW^T = I_{r-\mathrm{rank}(H)}.$$

Transformation by W removes the interference of the nuisance parameter X.

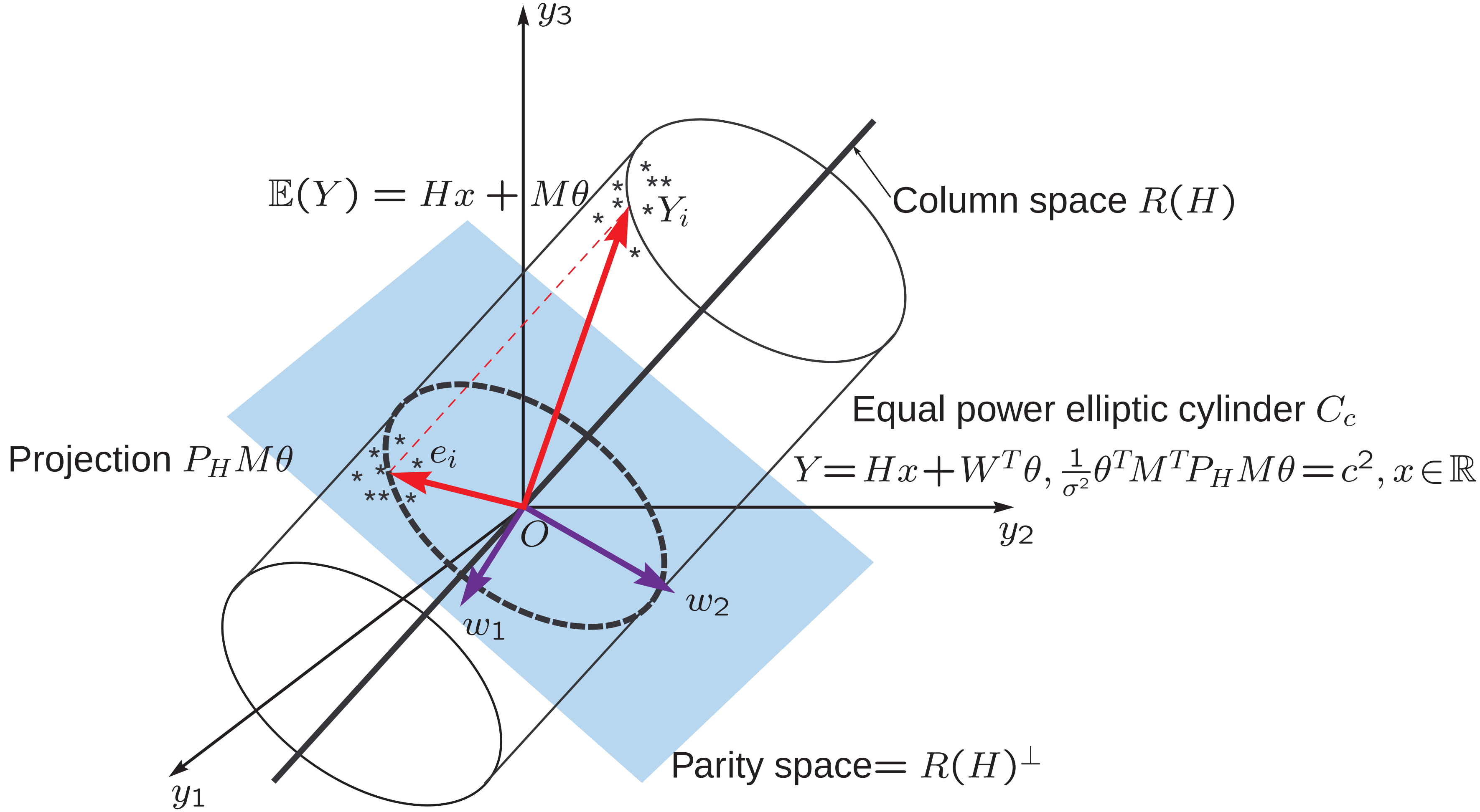
$$Z = WY = W\xi \; (+WM\theta)$$

For the following family of surfaces S_c [Fouladirad and Nikiforov 2005] :

$$S_c : \; \frac{1}{\sigma^2}\theta^T M^T P_H M\theta = c^2 \; (c > 0)$$

the test *UMP with constant power* writes as

$$\Lambda_1(Y) = \frac{1}{\sigma^2} Y^T P_H M (M^T P_H M)^{-1} M^T P_H Y$$



y_3
$\mathbb{E}(Y) = Hx + M\theta$
Y_i
Column space $R(H)$
Equal power elliptic cylinder C_c
Projection $P_H M\theta$
e_i
$Y = Hx + W^T\theta, \frac{1}{\sigma^2}\theta^T M^T P_H M\theta = c^2, x \in \mathbb{R}$
O
y_2
w_1
w_2
Parity space$= R(H)^{\perp}$
y_1

Method 2 : Reparameterization – thanks to Wald

The model can be re-written as :

$$Y = \widetilde{H}\theta + \xi,\ \widetilde{H} = (M \quad H),\ \theta = \begin{pmatrix} \Upsilon \\ X \end{pmatrix}$$

where the $r \times (m+p)$-dimensional matrix $\widetilde{H}$ is assumed to be f.c.r. Using the reparameterization approach to dealing with the nuisance X, we define $\eta_i(\theta) = \theta_i (i = 1, \ldots, m+p)$. The hypotheses can be re-written as :

$$\mathcal{H}_0 : \{\theta \in \omega : \eta_i(\theta) = \theta_i = 0;\ i = 1, \ldots, m\} \text{ and } \mathcal{H}_1 : \{\theta \in \mathbb{R}^{m+p} \setminus \omega\}.$$

The Jacobian matrix writes $J = I_{m+p}$ and matrix $J\mathcal{F}^{-1}J$ is equal to $\mathcal{F}^{-1} = \sigma^2(\widetilde{H}^T\widetilde{H})^{-1}$, $\widehat{\theta} = (\widetilde{H}^T\widetilde{H})^{-1}\widetilde{H}^T Y$. Let $\mathcal{F}_m$ be the $m \times m$-dimensional upper-left block of matrix $\mathcal{F}^{-1}$ and $\widehat{\theta}_m$ contain the first m components of $\widehat{\theta}$:

$$\Lambda_2(Y) = \eta(\widehat{\theta})^T\, \widetilde{\mathcal{F}}_m^{-1}(\widehat{\theta})\, \eta(\widehat{\theta}) = \widehat{\theta}_m^T\, \widetilde{\mathcal{F}}_m^{-1}\, \widehat{\theta}_m.$$

Method 3 : GLR

$$\begin{aligned}
\Lambda_3(Y) = 2\widehat{\Lambda}(Y) &= 2\log \frac{\sup_{\Upsilon,X} f_{HX+M\Upsilon,\sigma^2 I_r}(Y)}{\sup_X f_{HX,\sigma^2 I_r}(Y)} \\
&= -\frac{1}{\sigma^2}\|Y - M\widehat{\theta} - H\widehat{X}\|_2^2 + \frac{1}{\sigma^2}\|Y - H\widehat{X}\|_2^2 \\
&= \frac{1}{\sigma^2} Y^T \left(P_H - P_{M,H}\right) Y
\end{aligned}$$

where $P_{M,H} = I_r - \widetilde{H}\left(\widetilde{H}^T\widetilde{H}\right)^{-1}\widetilde{H}^T$.

Conclusion : if the matrix $(M \;\; H)$ is f.c.r. [Nikiforov 2002]

$$\Lambda_1(Y) = \Lambda_2(Y) = \Lambda_3(Y)$$

These tests are minimax and UMP with constant power.

Fault isolation : Two approaches

Again a linear gaussian model :

$$Y = HX + \xi \; (+M\Upsilon) \, , \quad \xi \sim \mathcal{N}(0, \sigma^2 I_r)$$

First approach : several simultaneous faults can be considered.

The approach consists in partitioning Υ in ϕ, and deciding in favor of the fault mode ϕ, while considering the other fault modes collected in ψ as nuisance information. The model writes with $\widetilde{H} \stackrel{\text{def.}}{=} (M_\phi \;\; M_\psi \;\; H)$:

$$Y = HX \; (+M_\psi \Upsilon_\psi) + \xi \; (+M_\phi \Upsilon_\phi)$$

The $r \times (m_\phi + m_\psi + p)$-dimensional matrix $\widetilde{H}$ is assumed to be f.c.r. Fault isolation can thus be seen as fault detection in the presence of a nuisance.

Second approach : a single fault at a time is assumed.

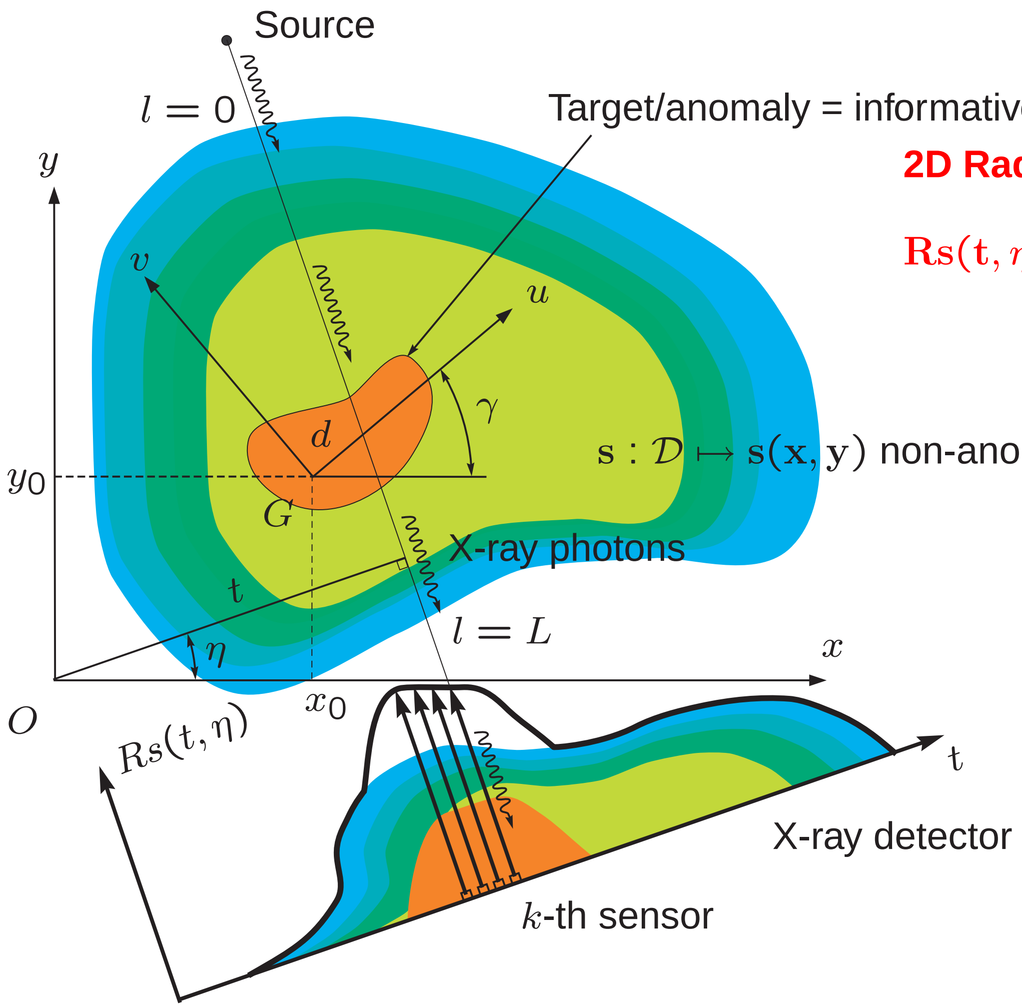
The second approach is multiple hypotheses testing. The problem is to test the null hypothesis $\mathcal{H}_0 : \{Y \sim \mathcal{N}(HX, \sigma^2 I_r);\ X \in \mathbb{R}^p\}$ against the alternative hypotheses $\mathcal{H}_j : \{Y \sim \mathcal{N}(HX + M_j \Upsilon_j, \sigma^2 I_r);\ X \in \mathbb{R}^p, \theta_j \neq 0\}$ $(j = 1, \ldots, K)$. Unfortunately, this hypotheses testing problem is not invariant under the group of permutations of the $\mathcal{H}_j$'s.

$$\widehat{\delta}(Y) = \begin{cases} \mathcal{H}_0 \text{ if } & \max_{i=1:K} \frac{f_{WM_i\widehat{\Upsilon}_i}(Z)}{f_0(Z)} < h \\ \mathcal{H}_j \text{ if } j = \arg\max_{i=1:K} & \frac{f_{WM_i\widehat{\Upsilon}_i}(Z)}{f_0(Z)} \geq h \end{cases} \quad Z = WY$$

where $f_\theta(Z)$ is the density of $\mathcal{N}(\theta, \sigma^2 I_{r-p})$ and

$$\widehat{\Upsilon}_j = \arg\min \|Z - WM_j\Upsilon_j\|^2 = \frac{M_j^T W^T Z}{M_j^T W^T W M_j}.$$

Example : target/anomaly detection [Fillatre and Nikiforov 2005]



2D Radon transform of the function $\mathbf{s}$ **:**

$$\mathbf{Rs}(\mathbf{t}, \eta) = \int_0^{\mathbf{L}} \mathbf{s}(\mathbf{t} \cos \eta - \mathbf{l} \sin \eta, \mathbf{t} \sin \eta + \mathbf{l} \cos \eta) \mathbf{dl}$$

$\mathbf{s} : \mathcal{D} \mapsto \mathbf{s}(\mathbf{x}, \mathbf{y})$ non-anomalous environment = nuisance parameters

Change detection - retrospective

Bayesian approach [Kolmogorov and Shiryaev 1960, Shiryaev 1961]

Consider the following *continuous* model of change :

$$dx_t = \nu \mathbf{1}_{t \geq t_0} dt + \sigma d\omega_t, \;\; P_\pi(t_0 < t) = 1 - e^{-\lambda t}$$

where $(\omega_t)_t$ is a normalized Brownian motion. The criteria :

$$\overline{\tau} \overset{\text{def.}}{=} E_\pi(N - t_0 | N > t_0) \rightarrow \min$$

$$E_\pi(N | N \leq t_0) \geq \overline{T}$$

The main result (lower bound) :

$$\overline{\tau} = \frac{1}{\rho_{1,0}} [\log \overline{T} + \log \rho_{1,0} - 1 - C + O(\rho_{1,0})] \text{ as } \lambda \rightarrow 0, \; \overline{T} \rightarrow \infty, \rho_{1,0} = \frac{\nu^2}{2\sigma^2}.$$

Non-Bayesian approach [Lorden 1971]

Let $(Y_k)_{k\geq 1}$ be an independent random sequence observed *sequentially* :

$$\mathcal{L}(Y_k) = \begin{cases} P_0 & \text{if} \quad k \leq k_0 - 1 \\ P_1 & \text{if} \quad k \geq k_0 \end{cases}, \quad k_0 = 1, 2, \ldots$$

The change time k_0 is an *unknown nonrandom* value. The problem is to *detect* the change in $P_l, l = 0, 1$ as soon as possible. The criteria is :

$$\overline{\mathbb{E}}^*(N) \stackrel{\text{def.}}{=} \sup_{k\geq 0} \operatorname{esssup} E_1((N - k_0 + 1)^+ | \mathcal{Y}_1^{k_0-1}) \to \min$$

over the classe $\mathcal{K}_\gamma = \{N : E_0(N) \geq \gamma\}$. The main result (lower bound) :

$$\overline{\mathbb{E}}^*(N) \sim \frac{\log \gamma}{\rho_{1,0}} \;\text{ as }\; \gamma \to \infty, \quad \rho_{1,0} = \int f_1 \ln \frac{f_1}{f_0} d\mu$$

where $\rho_{1,0}$ is the Kullback-Leibler information.

CUSUM Algorithm [Page 1954]

Consider for some $j : 1 \leq j \leq k$ the hypotheses

$$\begin{aligned} \mathcal{H}_j &= (Y_1, \ldots, Y_{j-1}) \sim P_0 \text{ and } (Y_j, \ldots, Y_k) \sim P_1 \\ \mathcal{H}_0 &= (Y_1, \ldots, Y_k) \sim P_0 \end{aligned}$$

The log-likelihood ratio for testing $\mathcal{H}_j$ against $\mathcal{H}_0$ is

$$S_j^k = \log \frac{f_j(Y_1, \ldots, Y_k)}{f_0(Y_1, \ldots, Y_k)} = \sum_{i=j}^{k} \log \frac{f_1(Y_i)}{f_0(Y_i)}$$

Maximum likelihood principle and the recursive form of the CUSUM algorithm :

$$N = \min\{k \geq 1 : g_k \geq h\}, \quad g_k = \max_{1 \leq j \leq k} S_j^k = \left(g_{k-1} + \log \frac{f_1(Y_k)}{f_0(Y_k)}\right)^+, \; g_0 = 0$$

Change detection/isolation

For introducing the second situation, the nuisance-free case $X = 0$ is considered. The sequence of observations $(Y_n)_{n\geq 1}$ is generated by the distribution $\mathcal{P}_{\theta(k)}$. Until time $k_0 - 1$, the parameter is $\theta(k) = \theta_0$ and, from k_0 onwards, it becomes $\theta(k) = \theta_l$ for some l, $1 \leq l \leq K$. The fault onset time k_0 and fault index l are assumed unknown and non random*.

The change detection/isolation algorithm should compute a *pair* (N, ν) based on the observations $(Y_k)_{k\geq 1}$, where ν, $1 \leq \nu \leq K$, is the *final decision* and N is the *alarm time* at which a ν-type change is detected.

*The results of a Bayesian approach can be found in [Speyer 1999, Lai 2000].

Independent observations :

A finite family of distributions $\mathcal{P} = \{P_i,\ i = 0,\dots,K\}$ with densities $\{f_i,\ i = 0,\dots,K\}$ is considered. In the parametric case, it is assumed that $\mathcal{P} = \{P_\theta,\ \theta \in \Theta\}$, where $\theta \in \mathbb{R}^r$, $\Theta = \bigcup_{i=0}^{K}\{\theta_i\}$ and the density function of this family is denoted by $f_\theta(Y)$.

Static regression model with redundancy :

$$Y_k = HX_k + \xi_k + G_l(k, k_0), \quad \xi_k \sim \mathcal{N}(0, \sigma^2 I_r)$$

where H is a full rank matrix of size $r \times p$ with $r > p$ and $G_l(k, k_0)$ is the l-type change occurring at time k_0, namely :

$$G_l(k, k_0) = \begin{cases} 0 & \text{if} \quad k < k_0 \\ G_l & \text{if} \quad k \geq k_0 \end{cases}, \quad 1 \leq l \leq K.$$

Dependent observations :

It is assumed that under P_0, the conditional density function of Y_k given $\mathcal{Y}_1^{k-1} = Y_1, \ldots, Y_{k-1}$ is $f_0(Y_k|\mathcal{Y}_1^{k-1})$ for every $k \geq 1$ and under $P_{k_0}^l$, the conditional density function is $f_0(Y_k|\mathcal{Y}_1^{k-1})$ for $k < k_0$ and is $f_l(Y_k|\mathcal{Y}_1^{k-1})$ for every $k \geq k_0$, $1 \leq l \leq K$.

State space model with additive changes :

$$\begin{aligned} X_{k+1} &= F_k X_k + \zeta_k + U_l(k, k_0) \\ Y_k &= H_k X_k + \xi_k + G_l(k, k_0), \end{aligned}$$

where $U_l(k, k_0)$ is defined exactly as $G_l(k, k_0)$, ζ_k and ξ_k are zero mean Gaussian white noises.

Several criteria to evaluate a change detection/isolation algorithm

Worst case conditional detection/isolation delay [Nikiforov 1995] and [Lai 2000]

$$\overline{\mathbb{E}}^*(N) \stackrel{\text{def.}}{=} \sup_{k_0 \geq 1, 1 \leq l \leq K} \text{esssup}\, \mathbb{E}^l_{k_0} \left((N - k_0 + 1)^+ | \mathcal{Y}_1^{k_0 - 1} \right)$$

where $x^+ = \max(0, x)$, is required to be *as small as possible* for a given minimum γ of the mean times before false alarm or false isolation :

$$\mathcal{K}_\gamma = \left\{ (N, \nu) \colon \mathbb{E}_0 \left(\inf_{r \geq 1} \{N_r : \nu_r = j\} \right) \geq \gamma, \mathbb{E}^l_1 \left(\inf_{r \geq 1} \{N_r : \nu_r = j\} \right) \geq \gamma \right\} \qquad (1)$$

for $1 \leq l, j \neq l \leq K$. For a more tractable performance index, the isolation constraint in (1) has been replaced by the probability of false isolation :

$$\mathbb{P}^l_1(\nu = j \neq l) \leq \beta \sim \gamma^{-1} \text{ as } \gamma \to \infty.$$

An asymptotic lower bound for the worst case delay, which extends the result of Lorden (1971) to multi-hypotheses case :

$$\overline{\mathbb{E}}^*(N;\gamma) \gtrsim \frac{\log \gamma}{\rho^*} \text{ as } \gamma \to \infty,$$

$$\text{where} \quad \rho^* \stackrel{\text{def.}}{=} \min_{1 \le l \le K} \min_{0 \le j \ne l \le K} \rho_{l,j}$$

$$\text{and} \quad 0 < \rho_{l,j} \stackrel{\text{def.}}{=} \mathbb{E}_1^l \left(\log \frac{f_l(Y_i)}{f_j(Y_i)} \right) < \infty$$

is the K-L information, which definition, in the general case of dependent observations, is more complicated [Lai, 2000].

Bayesian approach [Lai 2000]

Assume a prior distribution $Q = \{q_1, \ldots, q_{K-1}\}$ on the change-type l and an independent prior distribution π on the change time k_0. Lai proposes to minimize the following mean delay for detection/isolation for every $1 \leq l \leq K-1$

$$\sum_{k_0=1}^{\infty} \pi(k_0) \mathbb{E}^{l}_{k_0}(N - k_0 + 1)^+ \geq \frac{|\log \alpha|}{\min_{0 \leq j \neq l \leq K-1} \rho_{l,j}} (1 + o(1))$$

as $\alpha \to 0$, for a given error probability α :

$$\sum_{l=1}^{K-1} \sum_{k=0}^{\infty} q_l \pi(k+1) \mathrm{Pr}^{l}_{k+1}(k < N < \infty \cap \nu \neq l) + \sum_{k=0}^{\infty} \pi(k+1) \mathrm{Pr}_0(N \leq k) \leq \alpha$$

Uniformly constrained conditional probability of false isolation

The drawback of the previous criterion lies in that the probability of false isolation is constrained only if the change time is $k_0 = 1$. A more tractable criterion consists in minimizing the maximum detection/isolation delay [Nikiforov 2003] :

$$\overline{\mathbb{E}}(N) \stackrel{\text{def.}}{=} \sup_{k_0 \geq 1, 1 \leq l \leq K} \mathbb{E}^l_{k_0}(N - k_0 + 1 \mid N \geq k_0)$$

subject to : $\mathcal{K}_{\gamma,\beta} = \left\{(N,\nu) : \mathbb{E}_0(N) \geq \gamma,\ \sup_{k_0 \geq 1} \mathbb{P}^l_{k_0}(\nu = j \neq l | N \geq k_0) \leq \beta\right\}$, for $1 \leq l, j \neq l \leq K$. An asymptotic lower bound is given by :

$$\overline{\mathbb{E}}(N; \gamma, \beta) \gtrsim \max\left\{\frac{\log \gamma}{\rho^*_{\mathrm{d}}}, \frac{\log \beta^{-1}}{\rho^*_{\mathrm{i}}}\right\}$$

as $\min\{\gamma, \beta^{-1}\} \to \infty$, where $\rho^*_{\mathrm{d}} = \min_{1 \leq j \leq K} \rho_{j,0}$ and $\rho^*_{\mathrm{i}} = \min_{1 \leq l \leq K} \min_{1 \leq j \neq l \leq K} \rho_{l,j}$.

Recursive algorithm [Nikiforov 2000]

The above mentioned criterion can be realized by using the following recursive change detection/isolation algorithm :

$$N_r = \min_{1 \le l \le K} \{N_r(l)\}, \quad \nu_r = \arg\min_{1 \le l \le K} \{N_r(l)\},$$

where $N_r(l) = \inf \left\{ n \ge 1 : \min_{0 \le j \ne l \le K} \left[S_n(l,j) - h_{l,j} \right] \ge 0 \right\}$,

$$S_n(l,j) = g_n(l,0) - g_n(j,0), \; g_n(l,0) = \left(g_{n-1}(l,0) + Z_n(l,0) \right)^+ ,$$

with $Z_n(l,0) = \log f_l(Y_n)/f_0(Y_n)$, $g_0(l,0) = 0$ for every $1 \le l \le K$ and $g_n(0,0) \equiv 0$,

$$h_{l,j} = \begin{cases} h_d & \text{if} \quad 1 \le l \le K \quad \text{and} \quad j = 0 \\ h_i & \text{if} \quad 1 \le j, l \le K \quad \text{and} \quad j \ne l \end{cases} .$$

Uniformly constrained probabilities of false alarm and false isolation within a time window

For some safety-critical applications, it is necessary to warrant that the false alarm and false isolation probabilities within a time window with size m_α are lower than a prescribed upper bound [Lai 2000] :

$$\mathbb{E}^l_{k_0}(N - k_0 + 1)^+ \text{ is required to be } \textit{as small as possible}$$

subject to the following constraints :

$$\mathcal{K}_\alpha = \Big\{ (N,\nu) \; : \sup_{k \geq 1} \mathbb{P}_0(k \leq N < k + m_\alpha) \leq \alpha\, m_\alpha$$
$$\sup_{k_0 \geq 1} \mathbb{P}^l_{k_0}(k_0 \leq N < k_0 + m_\alpha \cap \nu \neq l) \leq \alpha\, m_\alpha \Big\}$$

on the false alarm and false isolation probabilities.

For every $1 \leq l \leq K$, an asymptotic lower bound for $\mathbb{E}^{l}_{k_0}(N-k_0+1)^+$ over the class $\mathcal{K}_\alpha$, which holds uniformly in k_0 when $\alpha \to 0$ is given by :

$$\mathbb{E}^{l}_{k_0}\,(N-k_0+1)^+ \;\geq\; \frac{\mathbb{P}_0(N \geq k_0)\,|\log \alpha|}{\rho_l + o(1)}$$

where $\rho_l \overset{\text{def.}}{=} \min_{j \neq l}\ \rho(\theta_l, \theta_j)$.

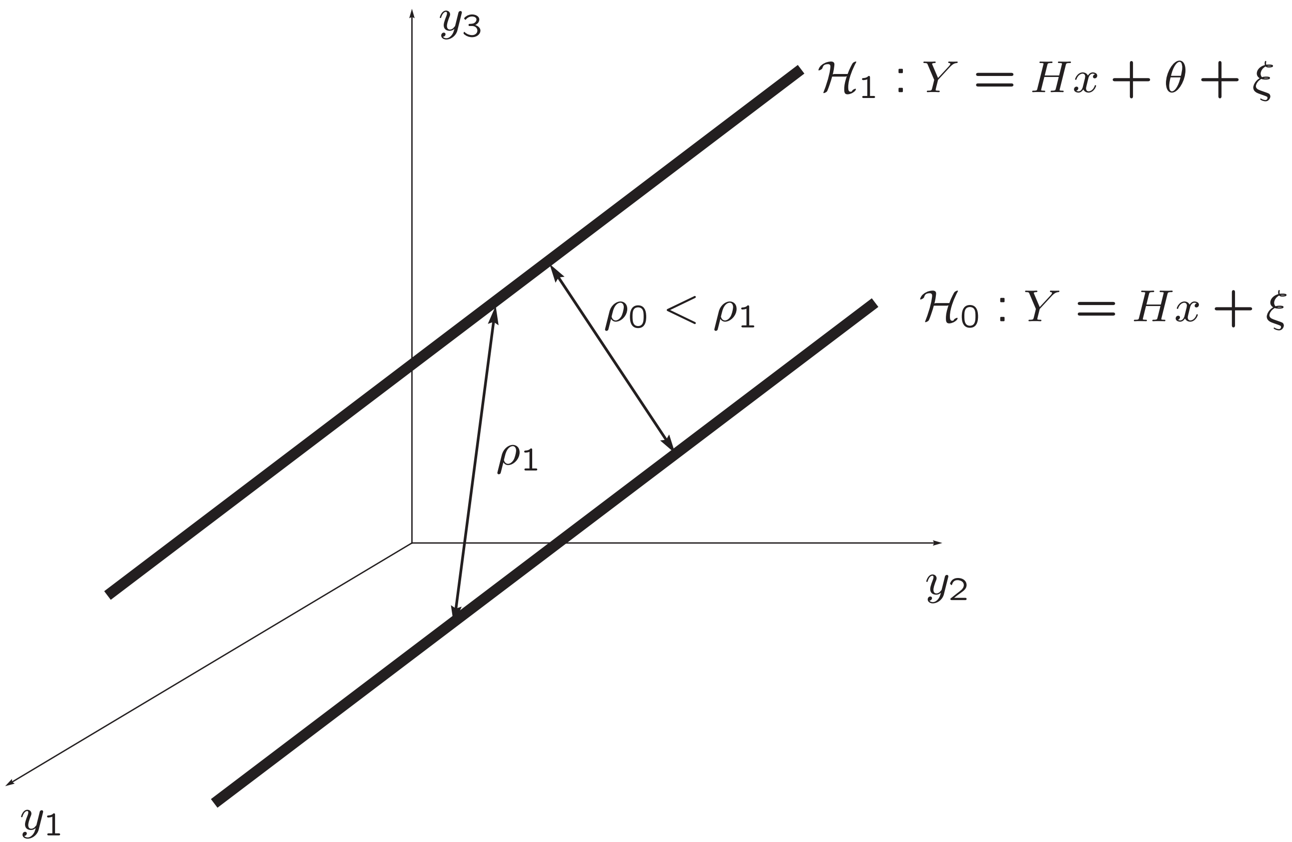
What is the moral of the story ?

The performance indexes of the discussed change detection/isolation algorithms are functions of K-L "distances" ρ_l, as it is the case for the hypotheses testing algorithms :

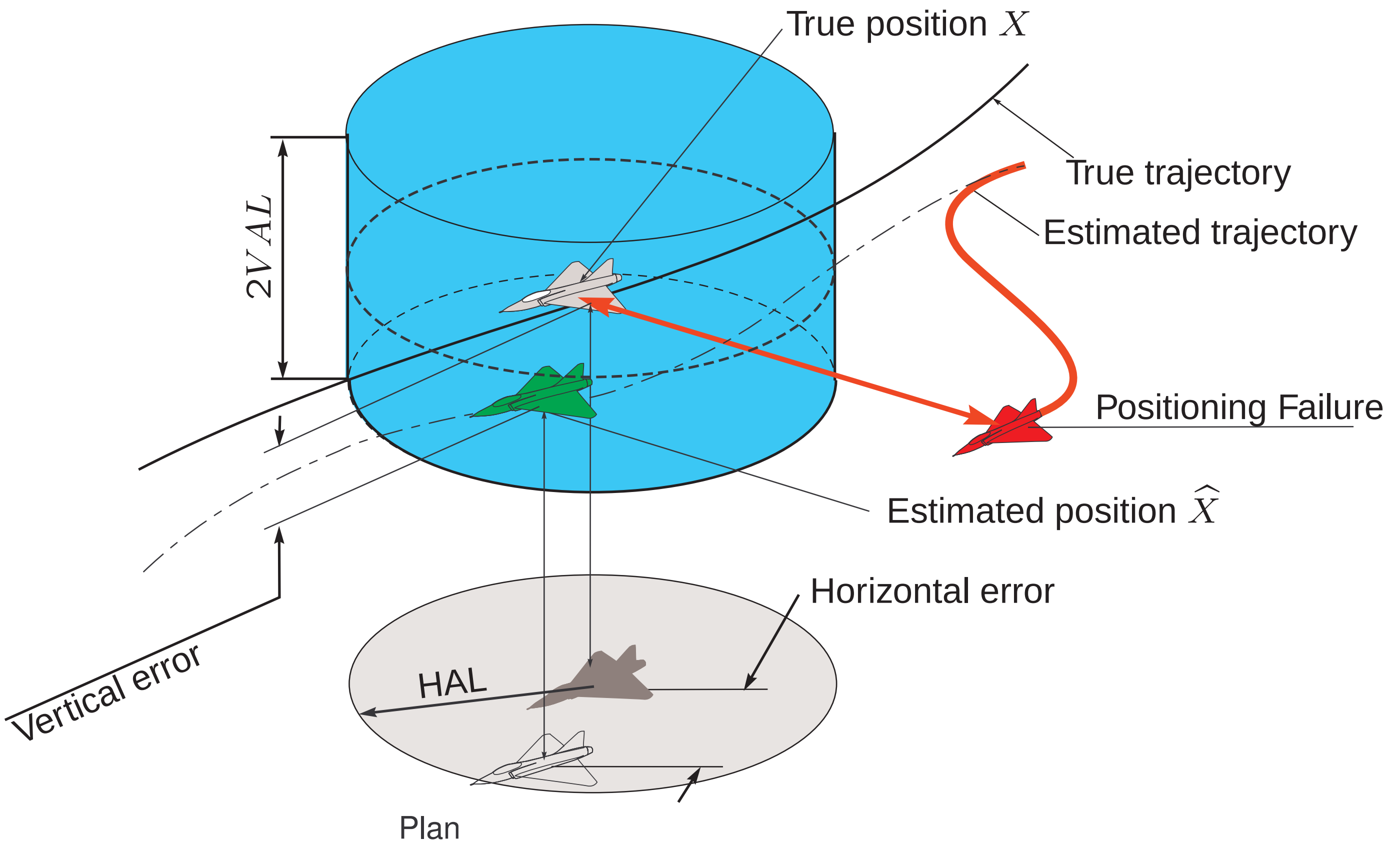
$$\text{Typical criterion} \;\gtrsim\; \frac{\log(\text{something})}{\rho_l} \qquad \text{(asymptotically !)}$$

Change detection/isolation : dealing with nuisance parameters

Geometric interpretation of the invariant sequential detection/isolation



Example (a) : SIRU integrity monitoring



Problem statement : SIRU integrity monitoring

For many safety-critical applications, a major problem of the existing navigation systems consists in its lack of integrity. The goal of the integrity monitoring is to detect and isolate faults so that they can be removed from the navigation solution before they sufficiently contaminate the output.

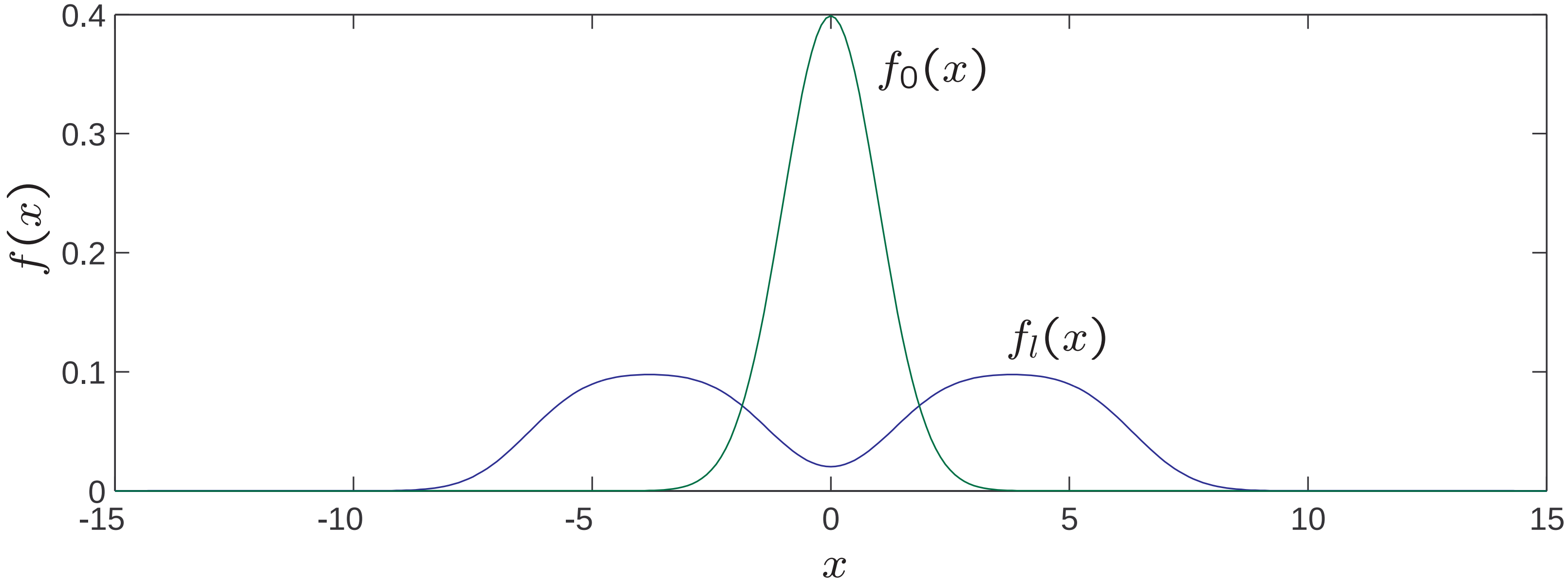
$$Y_k = H\mathcal{A}_k + \zeta_k, \; \zeta_k = \zeta_{k-1} + \xi_k + G_l(k, k_0),$$

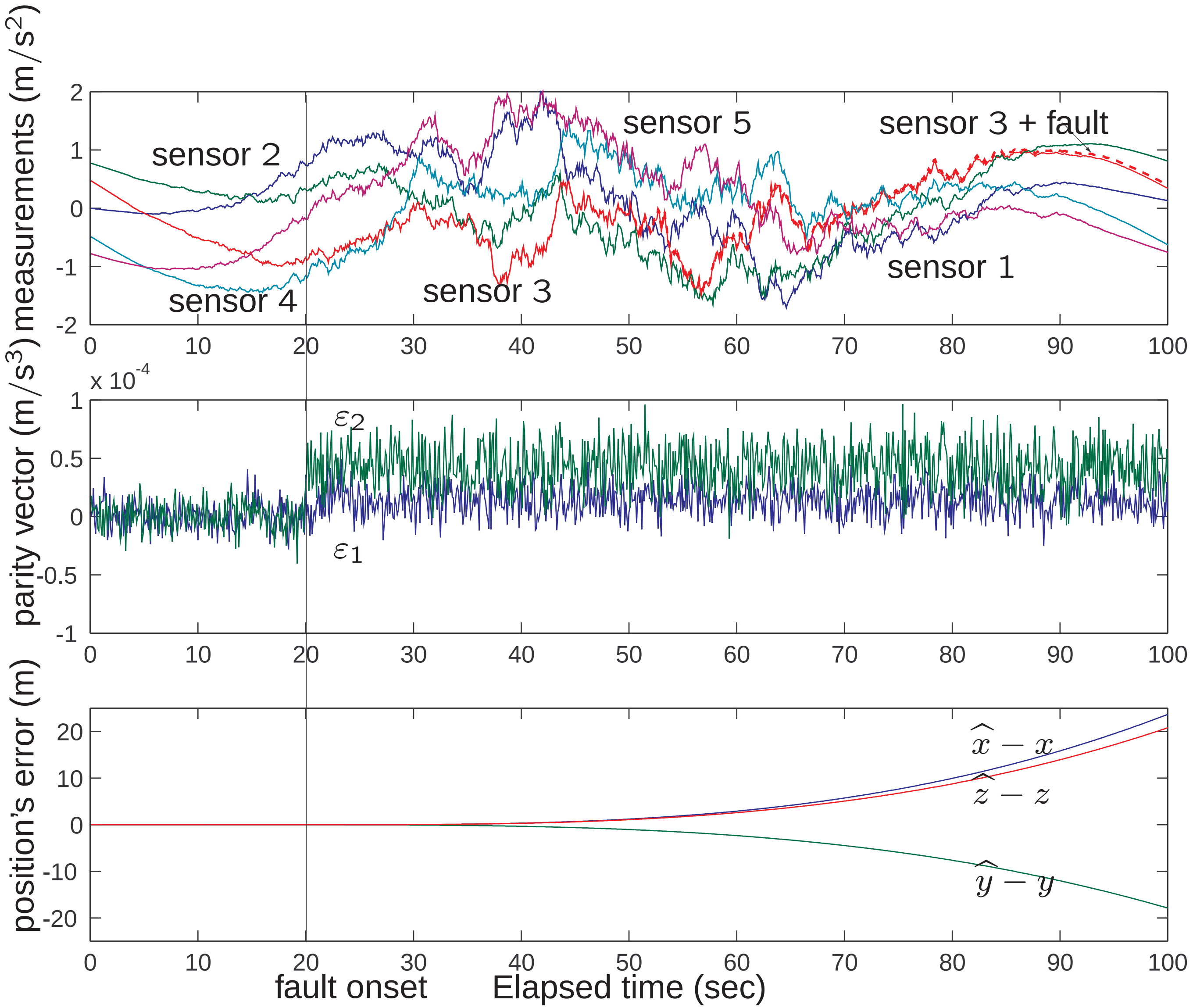
$$\nabla Y_k = H \nabla \mathcal{A}_k + \xi_k + G_l(k, k_0), \; \nabla(.)_k \stackrel{\text{def.}}{=} (.)_k - (.)_{k-1}.$$

The parity vector sequence $(e_t)_{t \geq 1}$ can be modeled as

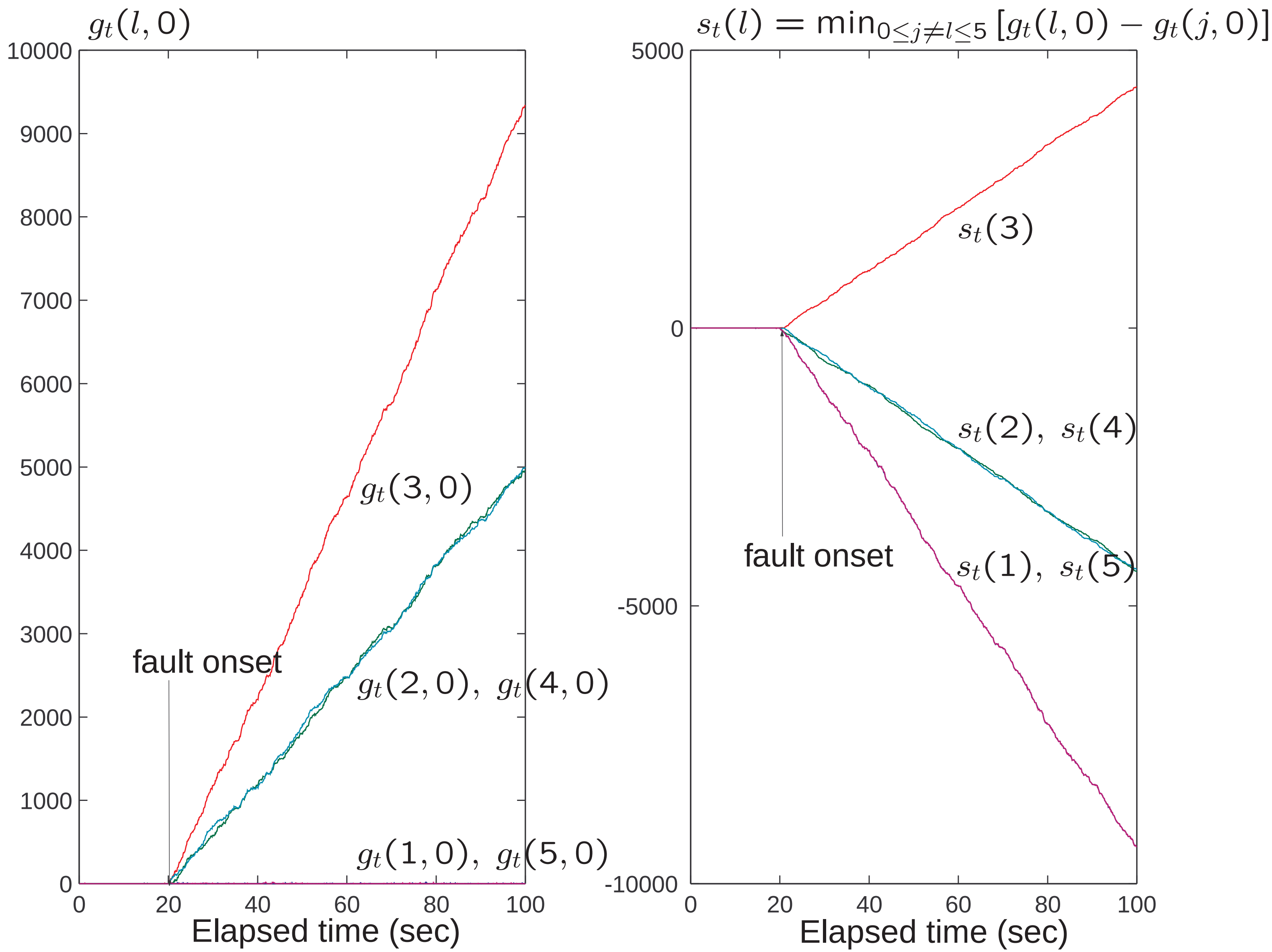
$$e_t = W\nabla Y_t = W\xi_t + WG_l(t,k), \quad l = 1, \ldots, s.$$

where the projection matrix W satisfies the following conditions $WH = 0, \quad W^T W = P_H, \quad WW^T = I_{s-3}$.

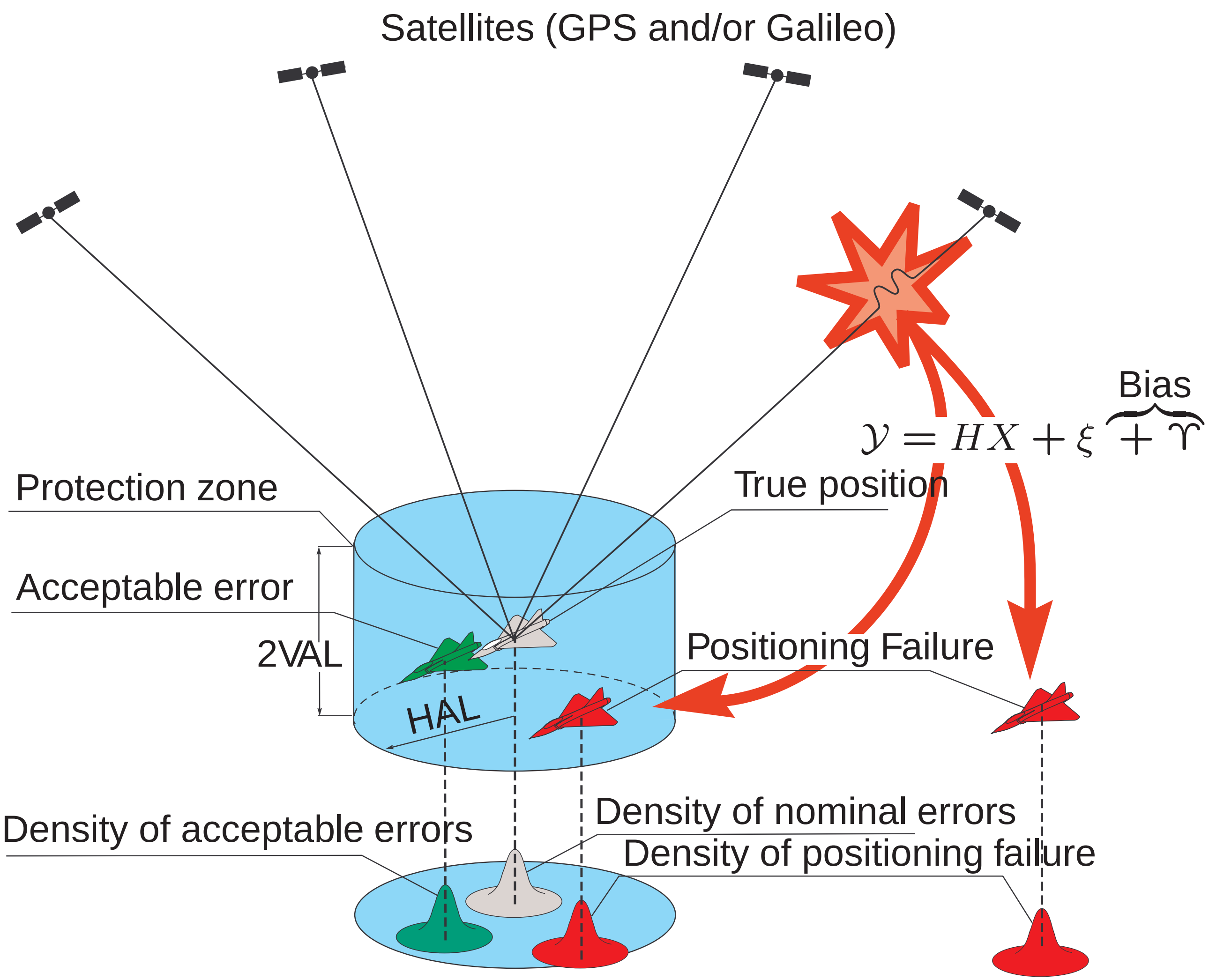




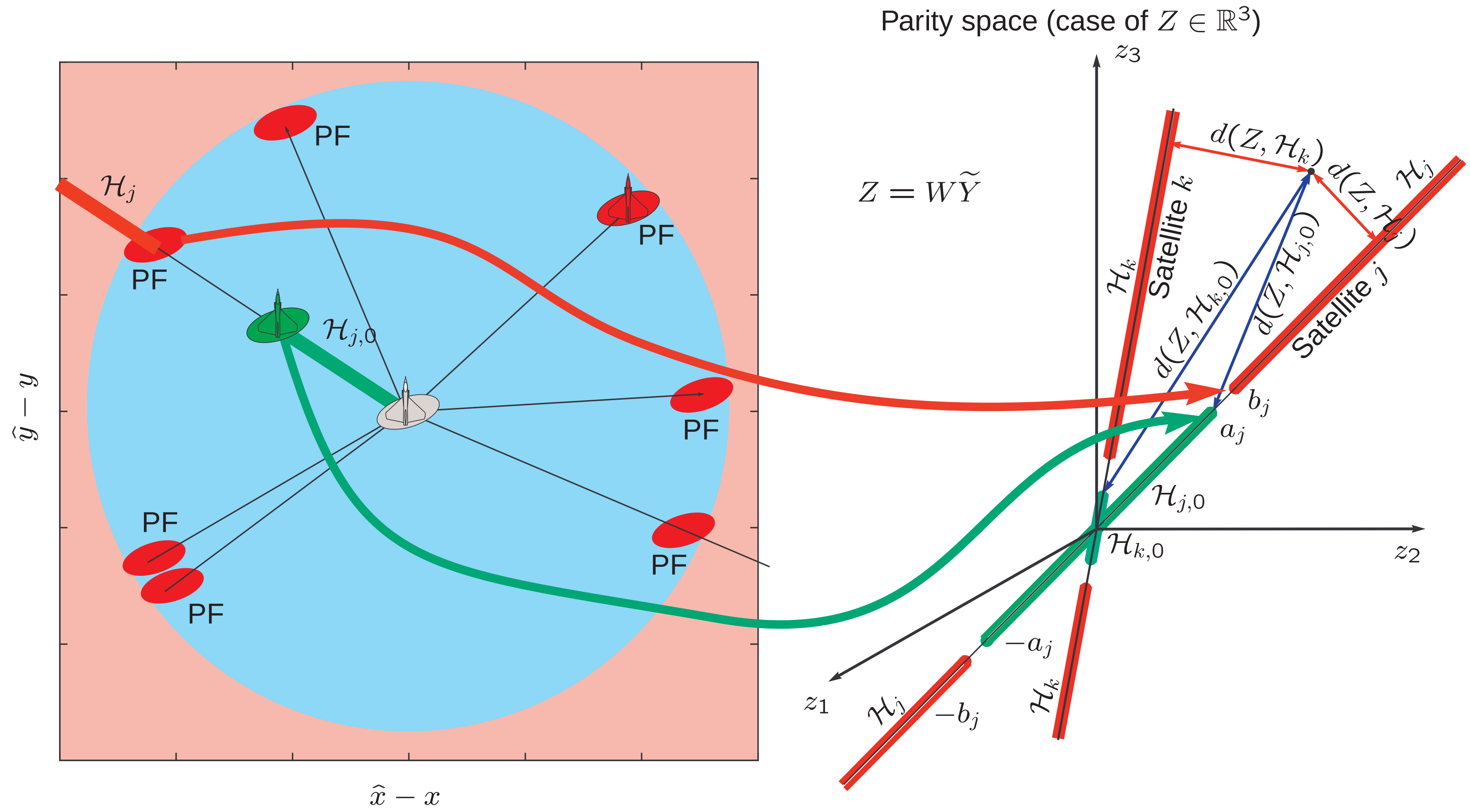
measurements (m/s^2)
sensor 2
sensor 5
sensor 3 + fault
sensor 4
sensor 3
sensor 1
parity vector (m/s^3)
x 10^-4
ε2
ε1
position's error (m)
x̂ − x
ẑ − z
ŷ − y
fault onset
Elapsed time (sec)


$g_t(l,0)$
fault onset
$g_t(3,0)$
$g_t(2,0),\ g_t(4,0)$
$g_t(1,0),\ g_t(5,0)$
Elapsed time (sec)
$s_t(l) = \min_{0\le j\ne l\le 5}\,[g_t(l,0) - g_t(j,0)]$
$s_t(3)$
$s_t(2),\ s_t(4)$
fault onset
$s_t(1),\ s_t(5)$
Elapsed time (sec)

Example (b) : GPS/Galileo integrity monitoring



Principle of advanced RAIM FDE algorithms

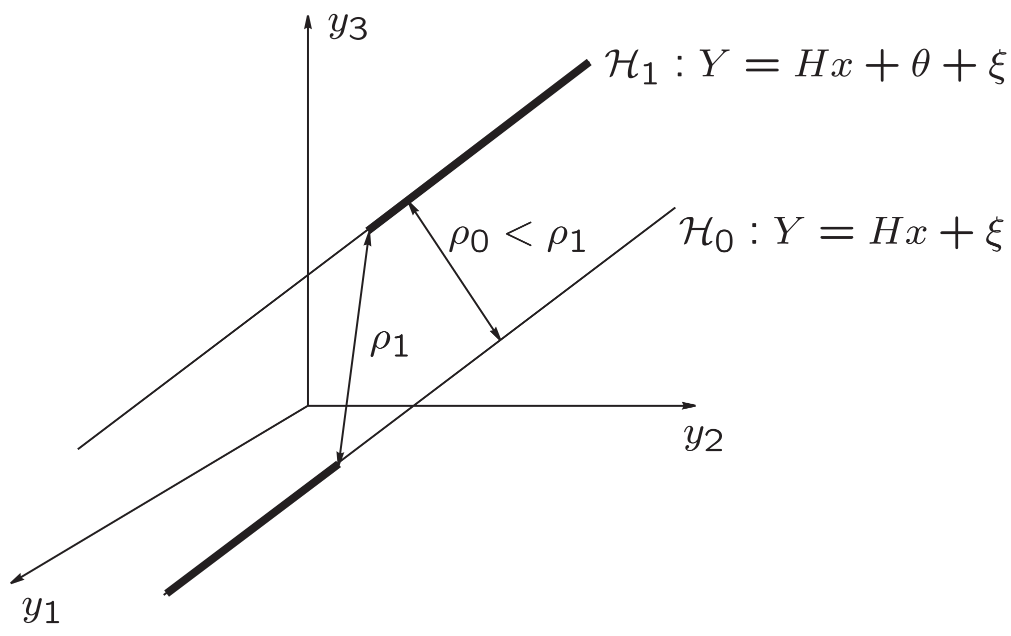


Conclusions :

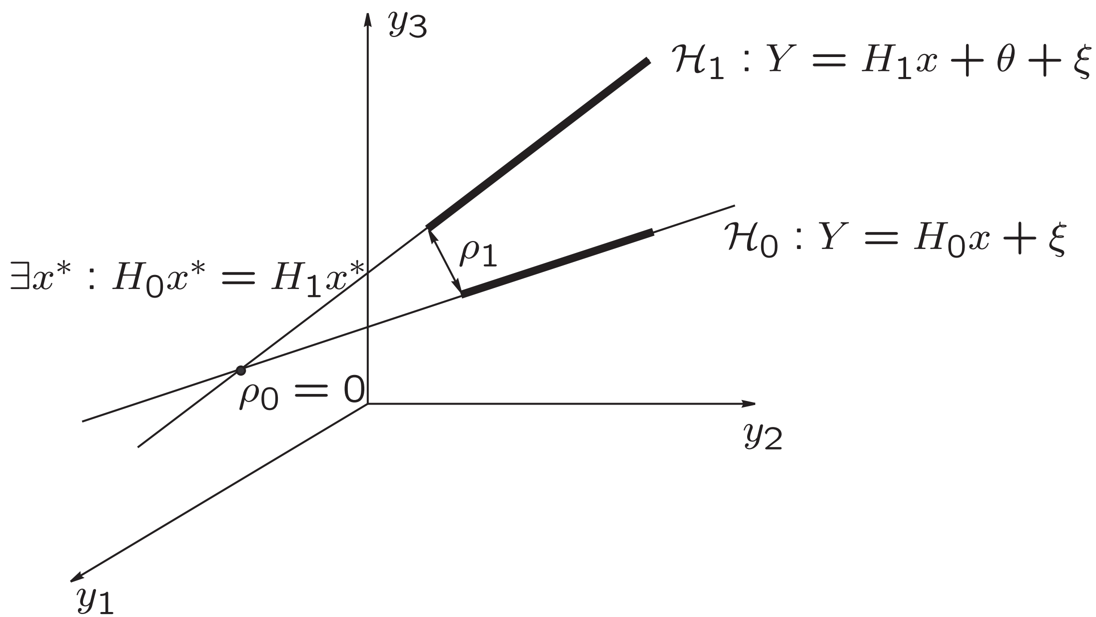
- Theoretical tools : minimax, UMP with constant power, statistical approach - analytical redundance, change detection/isolation
- Optimality criterion : tradeoff between the practical needs and theoretical results
- Nuisance parameters : Bayesian or non Bayesian, minimax, invariance

Perspectives :

- Nonlinear systems or nonlinear hypotheses
- Nonlinear nuisance : nonlinear estimation or nonlinear rejection ?
  GLR is not stable !
- Bounded nuisance



y_3
$\mathcal{H}_1 : Y = Hx + \theta + \xi$
$\mathcal{H}_0 : Y = Hx + \xi$
$\rho_0 < \rho_1$
ρ_1
y_2
y_1



y_3
$\mathcal{H}_1 : Y = H_1x + \theta + \xi$
$\mathcal{H}_0 : Y = H_0x + \xi$
ρ_1
$\exists x^* : H_0x^* = H_1x^*$
$\rho_0 = 0$
y_2
y_1