An introduction to particle methods in nonlinear filtering

Journées Thématiques "Filtrage particulaire"

Institut Henri Poincaré, les 2 et 3 décembre 2002

Nadia Oudjane (EDF, Clamart)

An introduction to particle methods in nonlinear filtering

1. The nonlinear filtering problem
   - Optimal filter and approximate filters
   - Some applications
2. Particle methods for nonlinear filtering
   - Weighted Monte Carlo Filter
   - Interacting Particle Filter
   - Some theoretical results
   - Improvement with robustification
   - Simulation results

General framework

▶ **State process (signal)** $(X_n)_{n\geq 0} \in E = \mathbb{R}^d$

X_n **Markov chain** $\sim (\pi_0, Q_n)$ **Ex :** $X_n = F_n(X_{n-1}, W_n)$

▶ **Observation process** $(Y_n)_{n\geq 0} \in \mathbb{R}^q$

$Y_n = h(X_n) + V_n$ $V_n \sim g_n$

▶ **Optimal filter** $(\pi_n)_{n\geq 0} \in \mathcal{P}(E)$

$\pi_n(dx) = \mathbb{P}[\, X_n \in dx \mid Y_{1:n} \,]$ **with** $Y_{1:n} = (Y_1, \cdots, Y_n)$

Prediction filter and likelihood function

- **Prediction filter** $(\pi_{n|n-1})_{n \geq 0} \in \mathcal{P}(E)$

$$\pi_{0|-1} = \pi_0 \qquad \pi_{n|n-1}(dx) = \mathbb{P}[\, X_n = dx \mid Y_{1:n-1} \,]$$

- **Likelihood function** $\Psi_n \; : E \longrightarrow \mathbb{R}^+$

$$\Psi_n(x) = g_n(Y_n - h(x))$$

Evolution of the optimal filter $\pi_{n-1} \longrightarrow \pi_n$

$$\pi_{n-1} \xrightarrow[\text{Prediction}]{(1)} \pi_{n|n-1} = Q_n \, \pi_{n-1} \xrightarrow[\text{Correction}]{(2)} \pi_n = \Psi_n \cdot \pi_{n|n-1}$$

- **Prediction** $\pi_{n-1} \longrightarrow \pi_{n|n-1}$

$$\pi_{n|n-1}(dx') = \int_E \pi_{n-1}(dx) \, Q_n(x, dx')$$

$$\pi_{n|n-1} = Q_n \, \pi_{n-1}$$

- **Correction** $\pi_{n|n-1} \longrightarrow \pi_n$

$$\pi_n(dx) = \frac{\Psi_n(x) \, \pi_{n|n-1}(dx)}{\int_E \pi_{n|n-1}(du) \, \Psi_n(u)}$$

$$\pi_n = (\Psi_n \cdot \pi_{n|n-1})$$

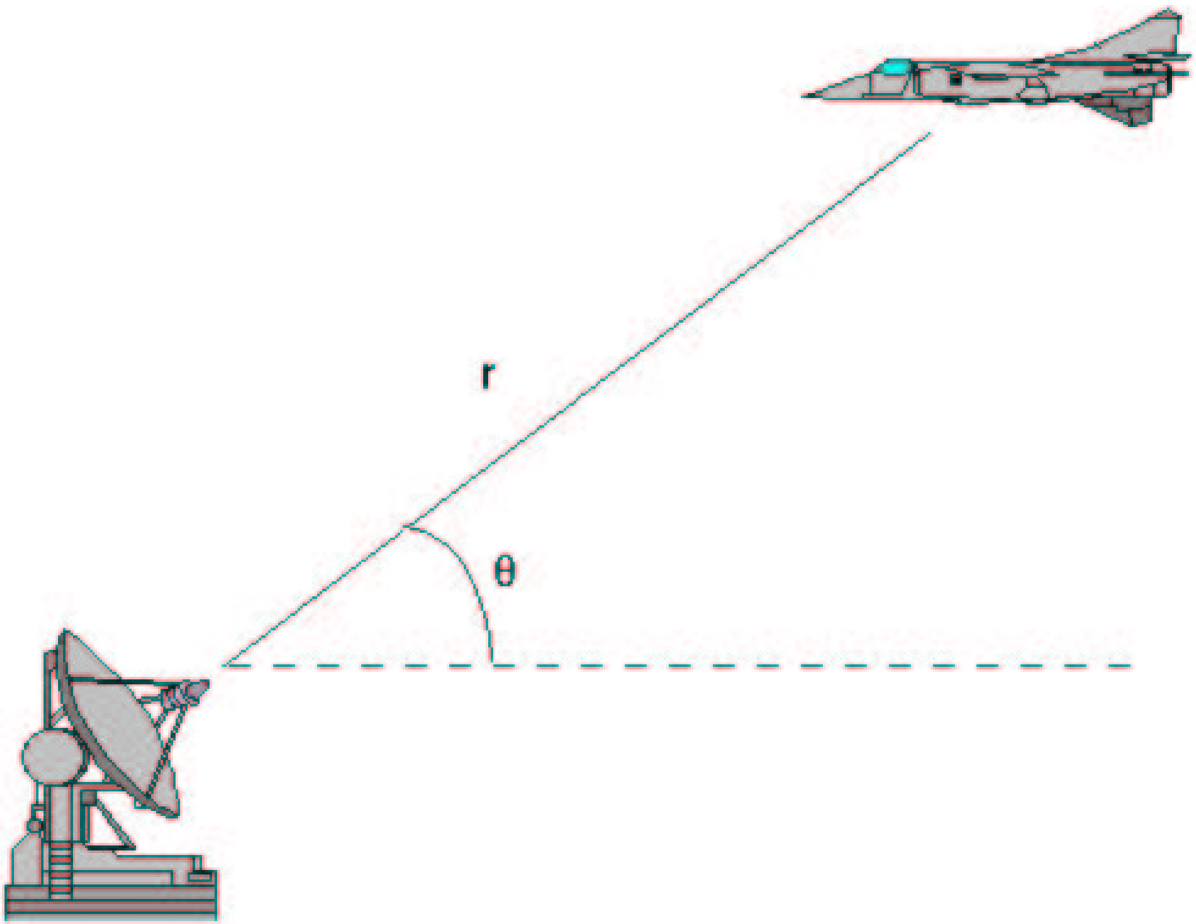
Approximate filters

- **Analytical method**
  - **Kalman Filter** — optimal in the linear Gaussian case
  - **Extended Kalman Filter** — weakly nonlinear and unimodal cases
- **Numerical methods**
  - **Fixed or adaptive grid** — Kushner, Dupuis (92) / Cai, Le Gland, Zhang (95)

    Small dimension ($d \leq 3$)
- **Monte Carlo methods**
  - **Particle filter** — Del Moral, Rigal et Salut (92)
  - **Bootstrap filter or SIR** — Gordon, Salmond et Smith (93)
  - **Monte Carlo filter** — Kitagawa (96)

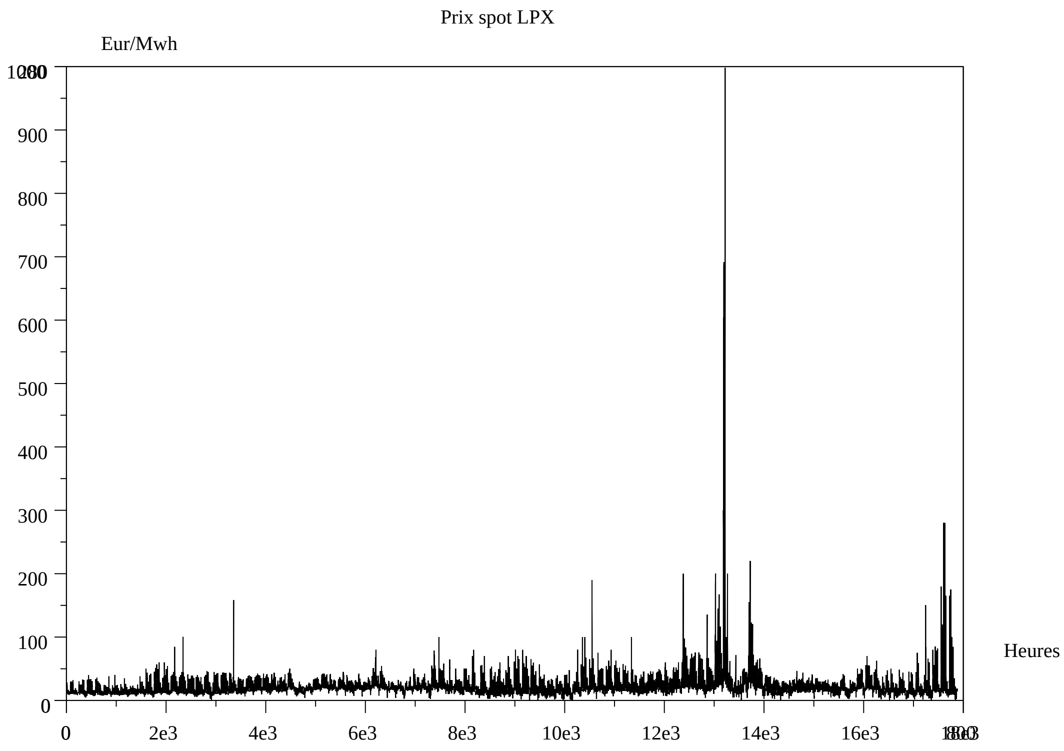
    Independent of the dimension and of the nonlinearity of the problem

The Tracking Problem



Application to mathematical finance

▶ **Aim : To fit a model on the spot price** $(S_t)_{t \geq 0} \in \mathbb{R}^+$



Description of the model

- **Multifactor model on the spot price** $(S_t)_{t\geq 0} \in \mathbb{R}^+$

$$\begin{cases} S_t = \exp(X_t) \\ dX_t = \alpha\,(Y_t - X_t)\,dt + \sigma_X\,dZ_t \\ dY_t = \mu\,dt + \sigma_Y\,dB_t \\ Y_0 = y_0 \end{cases}$$

- **Aim : To estimate** $\theta = (\alpha, \sigma_X, \mu, \sigma_Y)$

Three approaches

- **Filtering**
  - **Estimate (θ, Y) given the price observations X**
- **Filtering + Maximum likelihood (Martin Barlow)**
  - **0. Initialize $\theta = \theta_0$**
  - **1. Estimate Y given X by filtering**
  - **2. Estimate θ by Maximum likelihood given Y and go back to 1.**
- **Recursive maximum likelihood**

Principle of the Interacting Particle Filter (IPF)

$$\pi^N_{n-1} \xrightarrow[\text{Sampled Prediction}]{(1)} \pi^N_{n|n-1} = S^N(Q_n\, \pi^N_{n-1}) \xrightarrow[\text{Correction}]{(2)} \pi^N_n = \Psi_n \cdot \pi^N_{n|n-1}$$

▶ Sampled Prediction

$$\pi^N_{n|n-1} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\xi^i_{n|n-1}} \qquad \textbf{with} \qquad (\xi^1_{n|n-1}, \cdots, \xi^N_{n|n-1}) \sim Q_n\, \pi^N_{n-1}$$

▶ Correction

$$\pi^N_n = \sum_{i=1}^{N} \omega^i_n\, \delta_{\xi^i_{n|n-1}} \qquad \textbf{with} \qquad \omega^i_n = \frac{\Psi_n(\xi^i_{n|n-1})}{\sum_{j=1}^{N} \Psi_n(\xi^j_{n|n-1})}$$

Interacting Particle Filter (IPF)

$$\pi^N_{n-1} \xrightarrow{(1)} \pi^N_{n|n-1} = S^N(Q_n \, \pi^N_{n-1}) \xrightarrow[\text{Correction}]{(2)} \pi^N_n = \Psi_n \cdot \pi^N_{n|n-1}$$

Sampled

Prediction

$$\pi^N_{n-1} = \{ \omega^1_{n-1} \, \delta_{\xi^1_{n-1|n-2}} + \cdots + \omega^N_{n-1} \, \delta_{\xi^N_{n-1|n-2}} \}$$

(1.a) $\downarrow$ **Sampling** $\downarrow$

$$\{ \xi^1_{n-1}, \cdots \xi^N_{n-1} \}$$

(1.b) $\downarrow$ **Evolution** Q_n $\downarrow$

$$\pi^N_{n|n-1} = \{ \frac{1}{N} \, \delta_{\xi^1_{n|n-1}} + \cdots + \frac{1}{N} \, \delta_{\xi^N_{n|n-1}} \}$$

(2) $\downarrow$ **Correction** Y_n $\downarrow$

$$\omega^i_n \propto \Psi_n(\xi^i_{n|n-1})$$

$$\pi^N_n = \{ \omega^1_n \, \delta_{\xi^1_{n|n-1}} + \cdots + \omega^N_n \, \delta_{\xi^N_{n|n-1}} \}$$

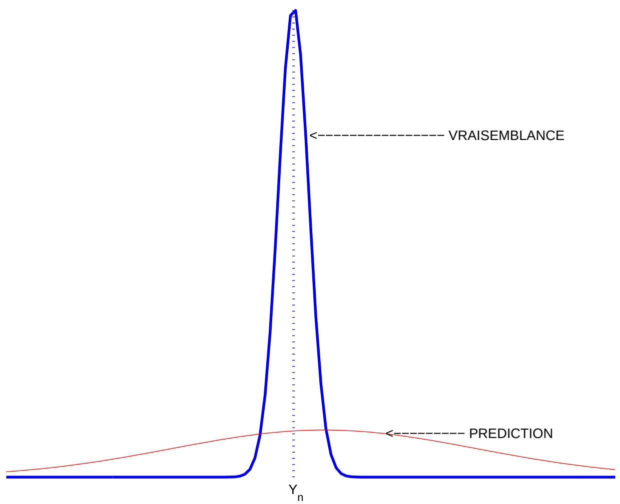
“Information of the transition”

- **Prediction** $\beta_n = \beta(Q_n) \leq 1$

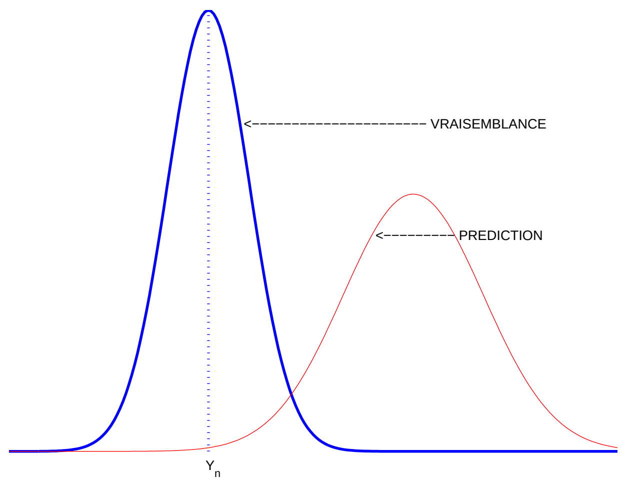
$$\|Q_n\,\mu - Q_n\,\mu'\| \leq \beta_n\,\|\mu - \mu'\|$$

Forgets the past

- **Correction** $\bar{\gamma}_n = \frac{\|\Psi_n\|}{\langle \pi_{n|n-1}, \Psi_n \rangle} \geq 1$ **Adds information**



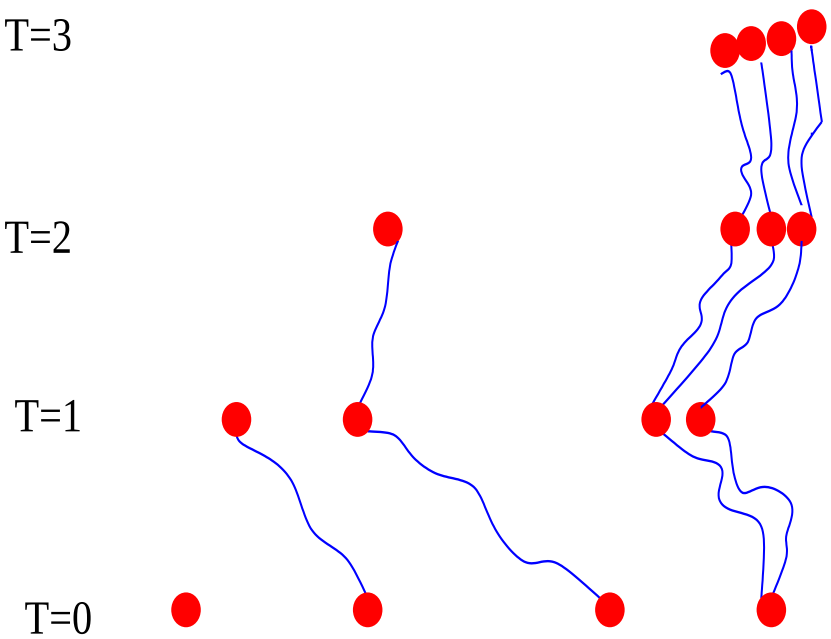
Precise Observation / $\pi_{n|n-1}$



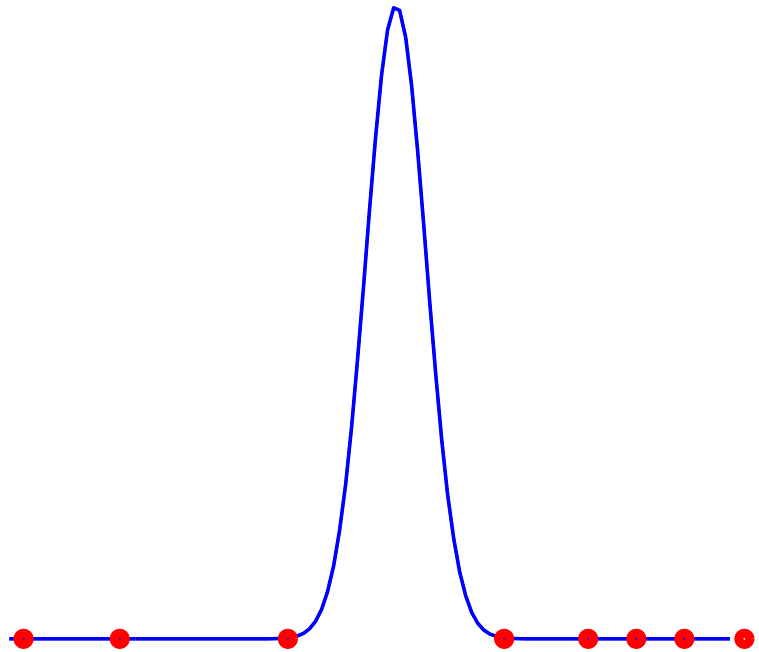
Incoherent Observation / $\pi_{n|n-1}$

Weaknesses of the IPF

- **Weak state noise** ($\beta_n \approx 1$)



- **Weak observation noise** ($\bar{\gamma}_n \ll 1$)



Evolution operator of the optimal filter $\bar{R}_n$

▶ **Integral operator** R_n

$$(R_n\,\mu)(dx') = \int_E \mu(dx)\, Q_n(x, dx')\, \Psi_n(x')$$

▶ **Normalized operator** $\bar{R}_n$

- **Definition**

$$\bar{R}_n(\mu) = \frac{R_n\,\mu}{(R_n\,\mu)(E)}$$

- **Filter evolution**

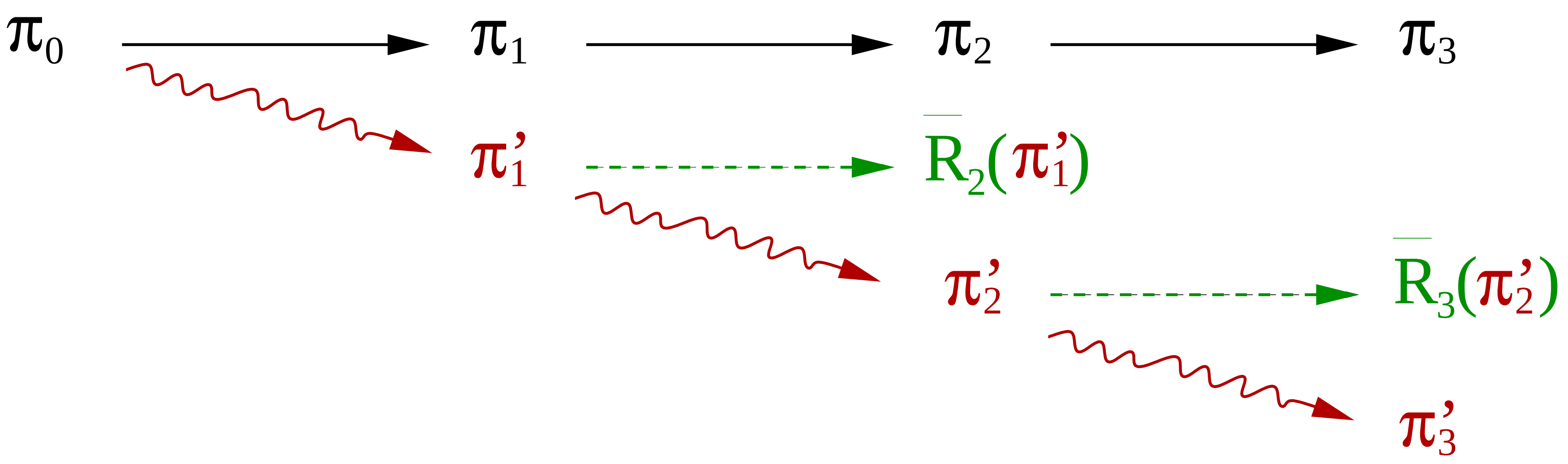
$$\pi_n = \bar{R}_n(\pi_{n-1}) = \bar{R}_n \circ \bar{R}_{n-1} \cdots \circ \bar{R}_1(\pi_0)$$

- **Notation**

$$\bar{R}_{n:1} = \bar{R}_n \circ \bar{R}_{n-1} \cdots \circ \bar{R}_1$$

Stability of the approximation

- **Optimal filter** $\pi_{k-1} \xrightarrow{\bar{R}_k} \pi_k$
- **Approximate filter** $\pi'_{k-1} \rightsquigarrow \pi'_k$
- **Decomposition of the global error** $\pi_n - \pi'_n$



$$\pi'_n - \pi_n = \sum_{k=1}^{n} [\, \bar{R}_{n:k+1}(\pi'_k) - \bar{R}_{n:k+1}(\bar{R}_k(\pi'_{k-1}))\,]$$

Convergence of the Interacting Particle Filter

- **Optimal filter** $\pi_{k-1} \xrightarrow{\bar{R}_k} \pi_k$
- **Particle filter** $\pi^N_{k-1} \rightsquigarrow \pi^N_k$
- **Result**

$$\sup_{\|\phi\|=1} \mathbb{E}[\, |\langle \pi_n - \pi^N_n , \phi \rangle| \mid Y_{1:n} \,] \leq \frac{C_n}{\sqrt{N}}$$

- **What about C_n ?**

$$C_n = \sum_{k=1}^{n} 2^{n-k} \, \beta_{n:k+1} \, \bar{\gamma}_{n:k+1}$$

- **Transport coefficients** $\beta_n \leq 1$ $\bar{\gamma}_n \geq 1$

$$\|Q_n \, \mu - Q_n \, \mu'\| \leq \beta_n \, \|\mu - \mu'\| \qquad \bar{\gamma}_n = \frac{\|\Psi_n\|}{\langle \pi_{n|n-1}, \Psi_n \rangle} \geq 1$$

Uniform convergence in time ?

▶ **Mixing assumption**

- **[R]** **with** $\varepsilon_k \geq \varepsilon > 0$ **and** $\lambda_k \in \mathcal{M}^+(E)$

$$\varepsilon_k \, \lambda_k \leq R_k(x, \cdot) \leq \frac{1}{\varepsilon_k} \, \lambda_k$$

▶ **Result**

- **[R]** **and** $\bar{\gamma}_k \leq \bar{\gamma}$ **imply**

$$\sup_{||\phi||=1} \mathbb{E}[\, | \, \langle \pi_n - \pi_n^N, \phi \rangle \, | \; | \; Y_{1:n} \,] \leq (\, 2 \, \bar{\gamma} \, / \, \varepsilon^2 \,) \, \frac{1}{\sqrt{N}}$$

How to built a robust approximate filter

- **Robustification of the transitions**
- **Introduction of numerical approximations**

$\Longrightarrow$ **Robust approximate filter**

- **EX : Regularized Particle Filter**

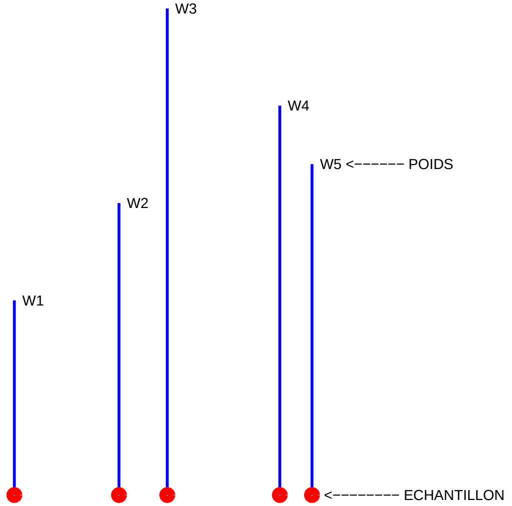
Regularization

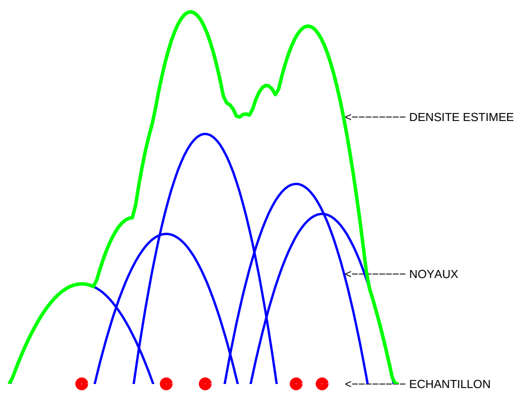
- **Regularization kernel (of order 2)** K

$$K \geq 0 \qquad \int K = 1 \qquad \int x_i \, K = 0 \qquad \int |x_i \, x_j| \, K < \infty$$

- **Scaled kernel** K_h $\qquad K_h(x) = \frac{1}{h^d} \, K(\frac{x}{h})$

- $$\mu^N = \sum \omega^i \, \delta_{\xi^i} \xrightarrow[K_h * \cdot]{\text{Regularization}} \mu^{N,h} = \sum \omega^i \, K_h(\cdot - \xi^i)$$





Regularization and robustification

▶ Optimal filter

$$\pi_{n-1} \xrightarrow[\text{Prediction}]{(1)} \pi_{n|n-1} = Q_n\, \pi_{n-1} \xrightarrow[\text{Correction}]{(2)} \pi_n = \Psi_n \cdot \pi_{n|n-1}$$

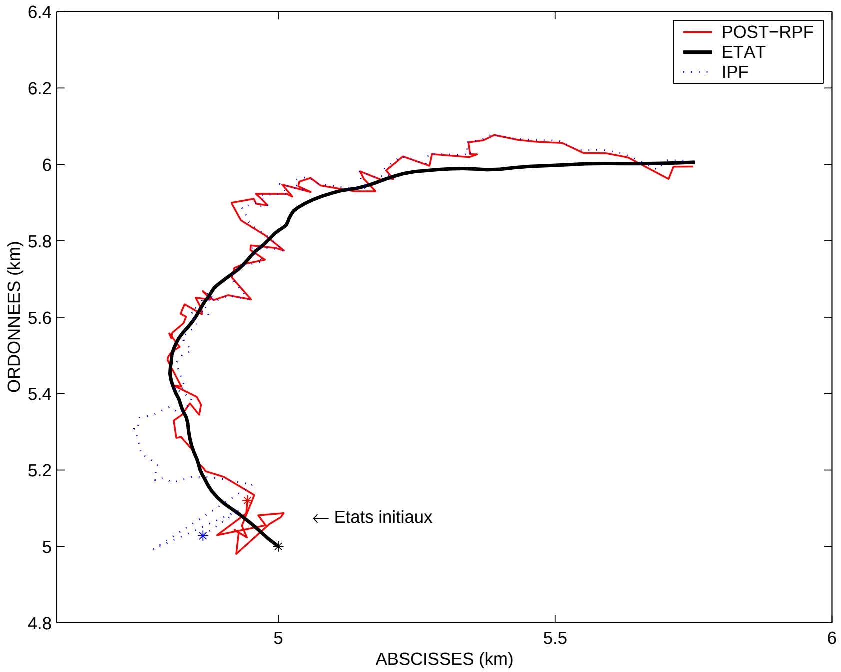
▶ Regularized filter

$$\pi^h_{n-1} \xrightarrow[\text{Modified Prediction}]{(1)} \pi^h_{n|n-1} = K_h * (Q_n\, \pi^h_{n-1}) \xrightarrow[\text{Correction}]{(2)} \pi^h_n = \Psi_n \cdot \pi^h_{n|n-1}$$

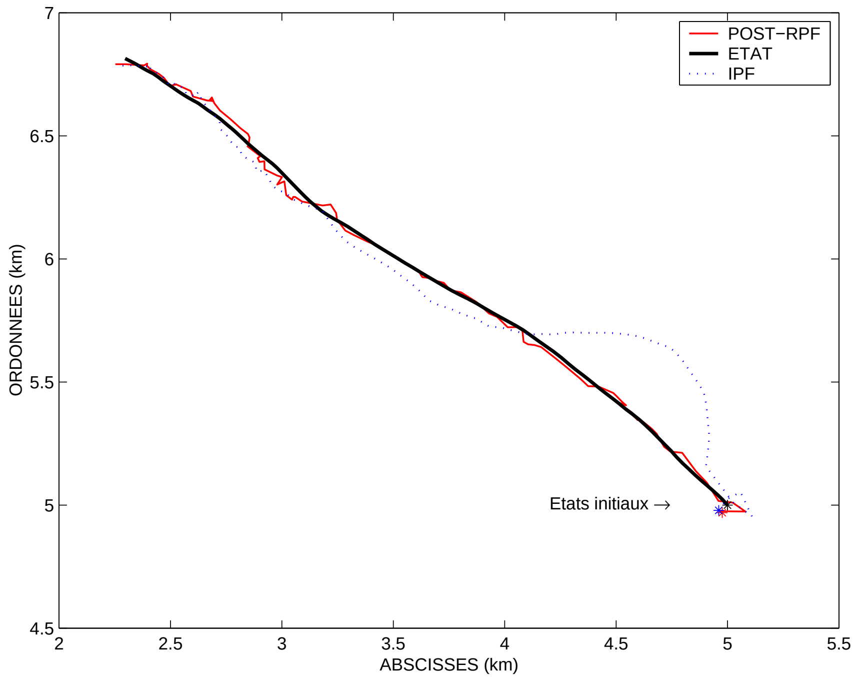
▶ Regularized Particle filter

$$\pi^{N,h}_{n-1} \xrightarrow[\text{Modified and Sampled Prediction}]{(1)} \nu^{N,h}_{n|n-1} = K_h * S^N (Q_n\, \pi^{N,h}_{n-1}) \xrightarrow[\text{Correction}]{(2)} \pi^{N,h}_n = \Psi_n \cdot \nu^{N,h}_{n|n-1}$$

Influence of the dynamical noise (1)

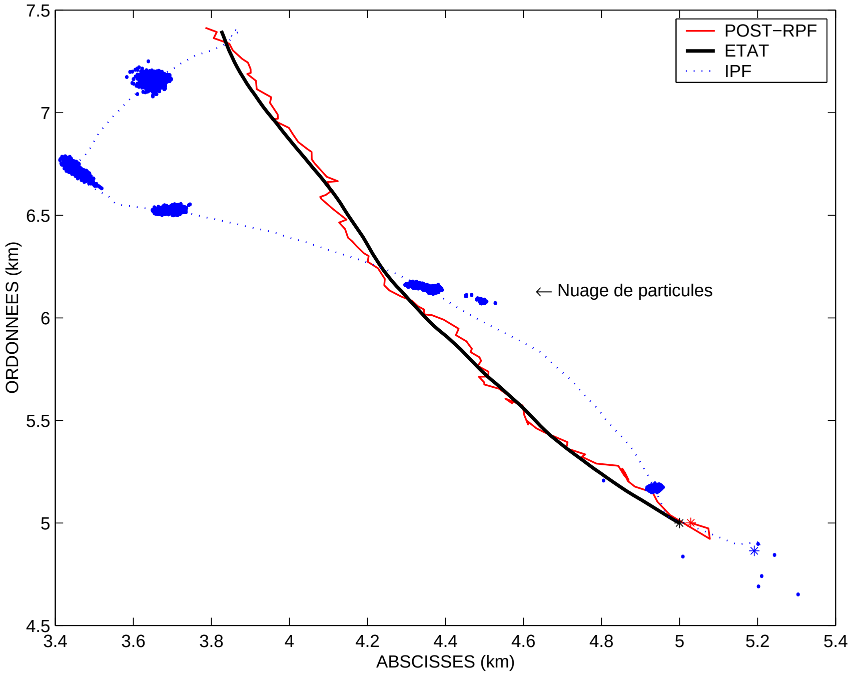


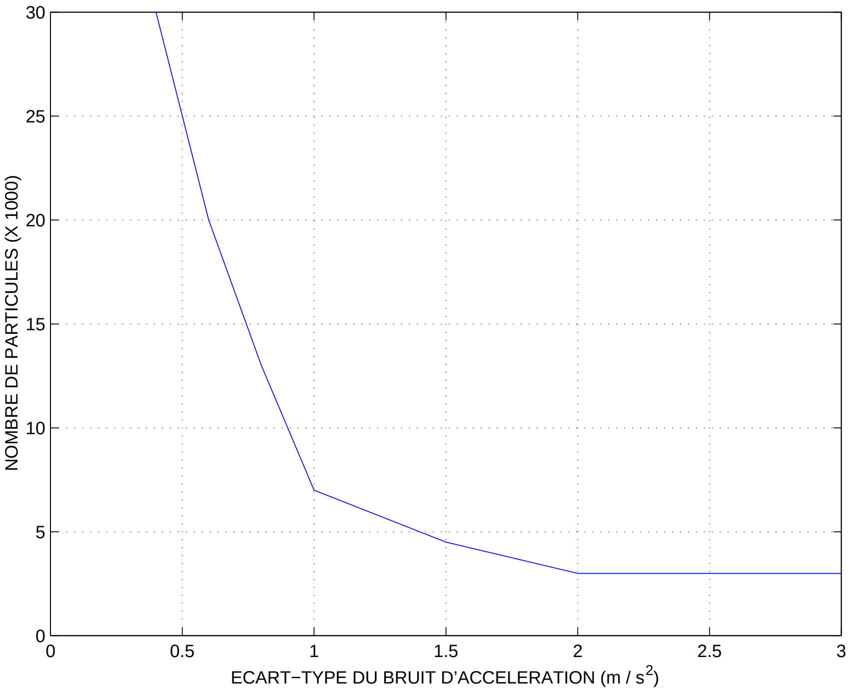
Dynamical noise c=2 m/s^2



Dynamical noise c=1.5 m/s^2

Influence of the dynamical noise (2)

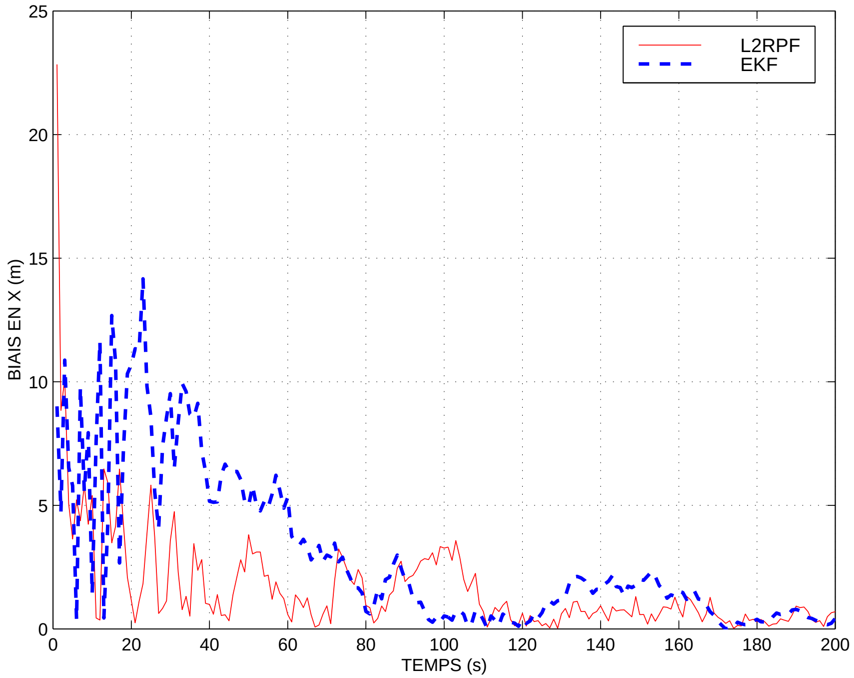




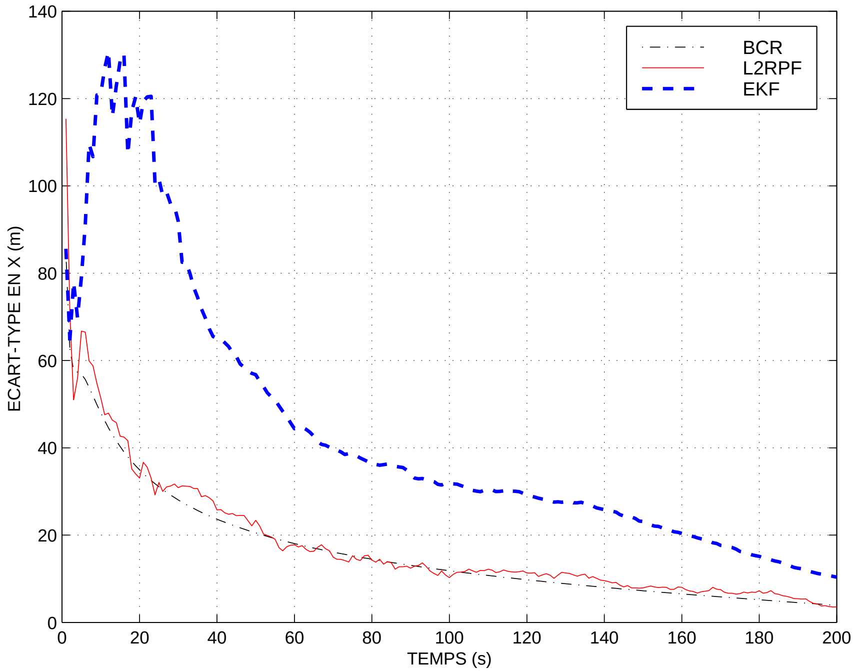
Dynamical noise c=1 m/s^2

N of particles / Dynamical noise

RPF / EKF : Range and Bearing Tracking

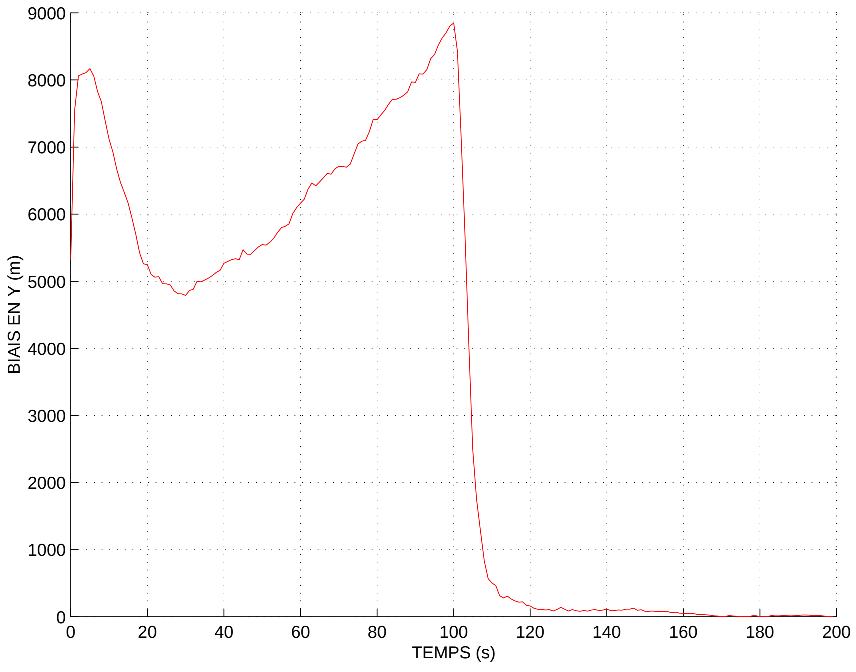


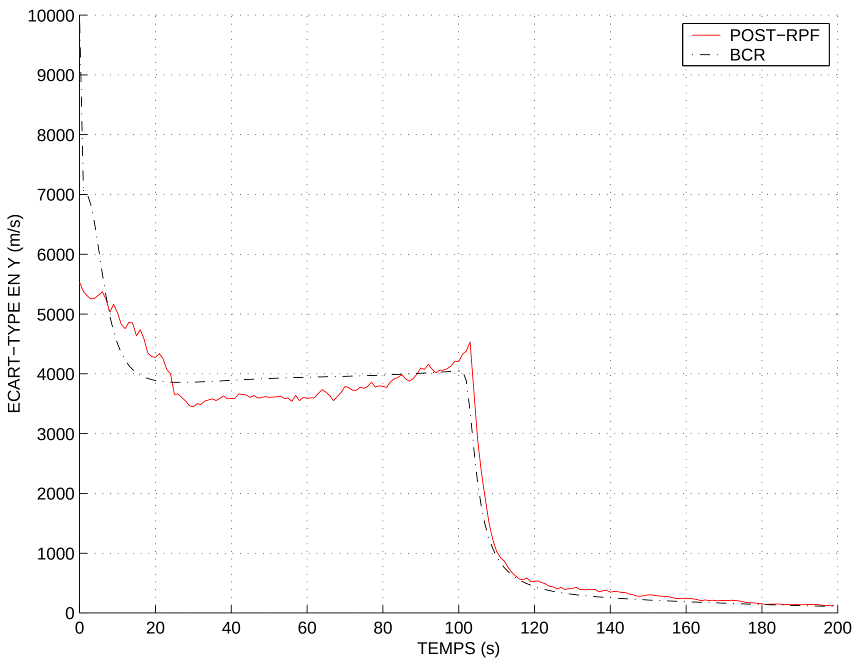
Bias on X



Standard deviation on X

RPF / EKF : Bearing only tracking

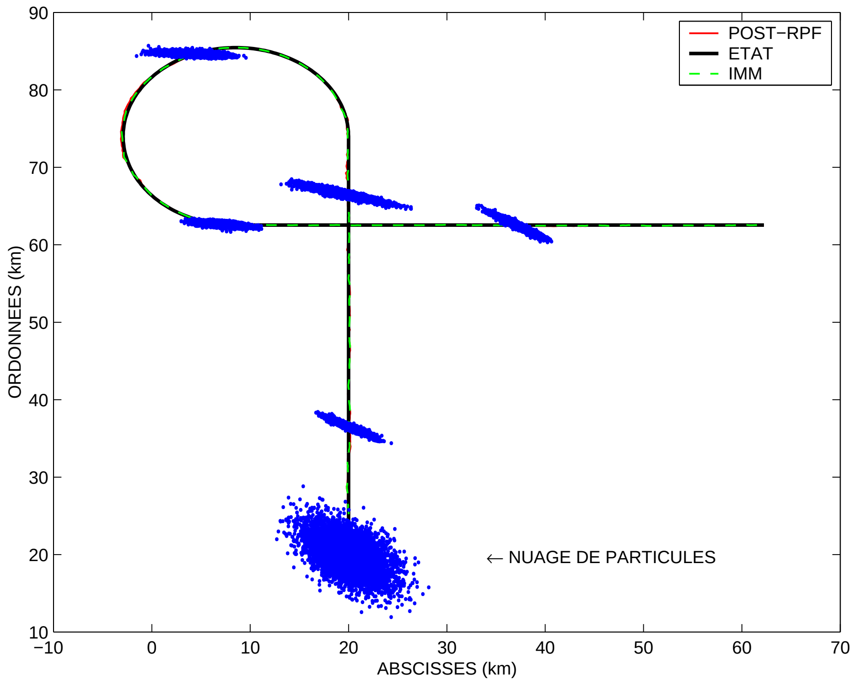


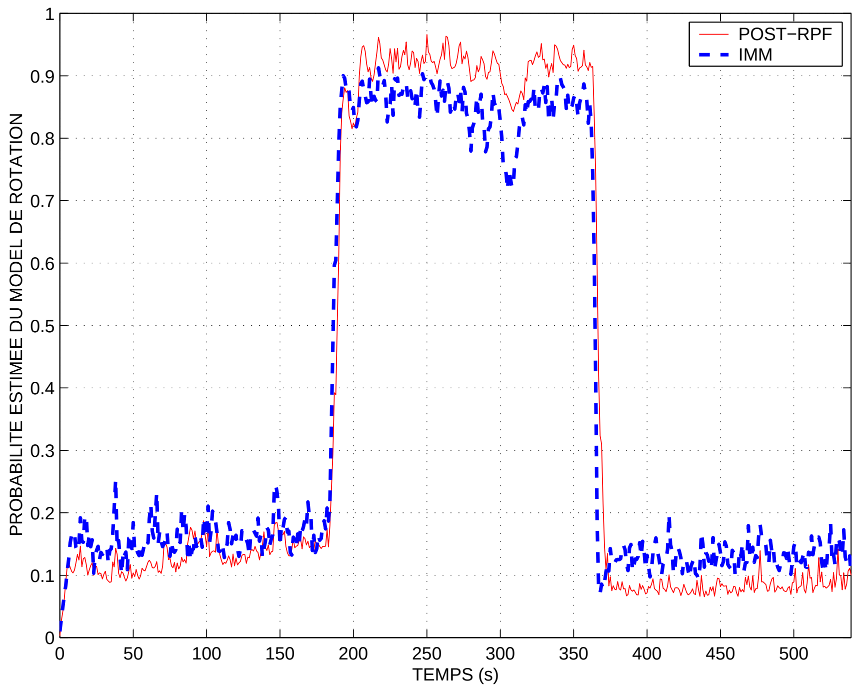


Bias on Y

Standard deviation on Y

RPF / IMM : manœuvering target (1)

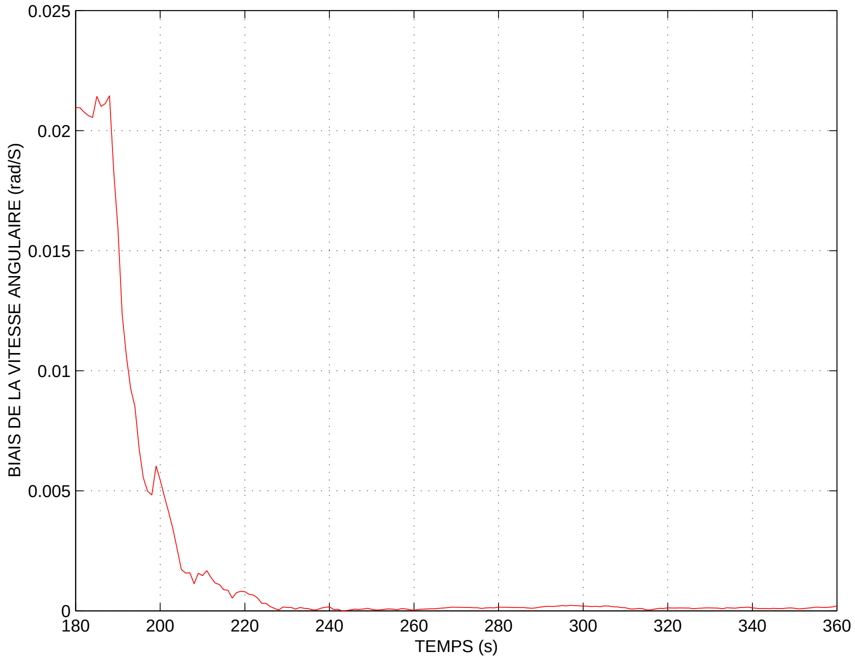


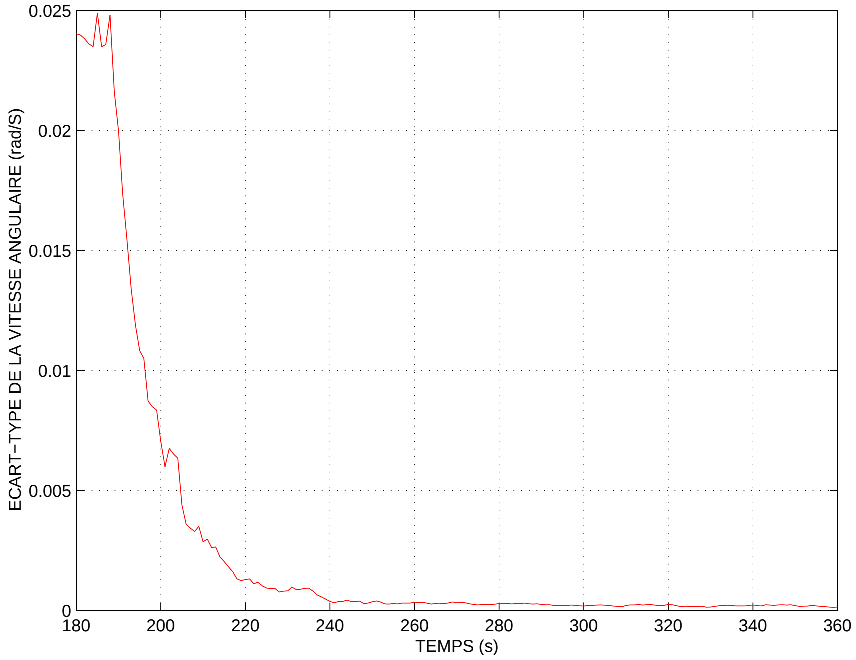


Real path and estimated path

Probability of the rotation model

RPF / IMM : manœuvering target (2)





Bias on the angular rate **Standard deviation on the angular rate**