

Filtrage non linéaire en temps discret Application du filtre particulaire avec un bruit d'observation uniforme

Matio ROBINSON¹, Yvon ANDRIANAHARISON², Falimanana
RANDIMBINDRAINIBE³
Laboratoire d'Electricité, Signaux et Automatique
Ecole Supérieure Polytechnique d'Antananarivo

Université d'ANTANANARIVO
BP 1500, Ankatso Antananarivo 101, MADAGASCAR

¹mat_robinson2000@yahoo.fr, ²yvonkyo@yahoo.fr, ³falimanana@mail.ru

Résumé :- Cet article développe la méthode de Monte Carlo séquentielle ou la méthode particulaire pour le filtrage non linéaire non nécessairement gaussien. Le filtre particulaire nécessite des hypothèses assez complexes dans son élaboration. Néanmoins, avec un bruit uniforme dans l'équation d'observation, on justifie aisément la nécessité de redistribution des particules pour avoir une approximation plus vraisemblable.

Mots-clefs :- Filtrage non linéaire, filtre particulaire, Bootstrap, approximation de Monte Carlo, chaîne de Markov.

Abstract :-This paper develops the Monte Carlo sequential method or particulate method for non linear filtering not necessarily Gaussian. The particulate filter requires hypothesis rather complex in its development. Nevertheless, with an uniform noise in the observation equation, it easily justifies the need of redistribution of particles to have a more probable approximation.

Keywords :- Nonlinear filtering, particulate filter, Bootstrap, Monte Carlo Approximation, Markov Chain.

1 Introduction

La mesure de l'état d'un système est en général impossible, pour différentes raisons entre autres, l'absence de matériels adaptés aux mesures. On fait alors des mesures observables à partir des capteurs qui sont évidemment liés à l'état du système non observé au cours du temps.

Le signal obtenu est en général perturbé par d'autres signaux d'origines inconnues supposées aléatoires.

Le filtrage [1], [2] consiste à obtenir, avec l'information des observations, une estimation de la distribution de la variable d'état non-observable. Cette estimation repose sur deux étapes classiques, la prédiction et la correction. Ces étapes permettent de construire un algorithme récursif dans la détermination du filtre en « temps réel ».

Les bruits de perturbation sont considérés comme additifs.

Le modèle du problème de filtrage [3] est un système de deux équations : une sur la dynamique de l'état $X = \{X_n; n \geq 0\}$ et une autre sur l'observation $Y = \{Y_n; n \geq 1\}$.

Notons \mathcal{Y}_n , la tribu des observations Y_1, \dots, Y_n , et les filtres prédit $\mu_n^-(dx) = \mathbb{P}_{X_n}(dx | \mathcal{Y}_{n-1})$ et corrigé $\mu_n(dx) = \mathbb{P}_{X_n}(dx | \mathcal{Y}_n)$.

Dans cet article, on établit des relations de récurrence des filtres pour un système non linéaire et non nécessairement à bruits gaussiens.

Une des particularités des filtrages à bruits d'observation uniforme est le pouvoir de construire le filtre séquentiellement sans passer par la décomposition classique de prédiction-corrrection.

Dans le cas plus général, on introduit la méthode de résolution par le filtre particulière, une méthode la plus adaptée aux filtrages non linéaires basée sur la méthode d'approximation de Monte Carlo.

2 Filtrage non linéaire

Considérons le système tel que l'état non observé $X = \{X_n; n \geq 0\}$ est une chaîne de Markov homogène à valeurs dans l'espace mesurable $(\mathbb{R}^d; \mathcal{B}(\mathbb{R}^d))$, de loi initiale μ_0 et de probabilité de transition π où pour toute fonction mesurable $f : \mathbb{R}^d \rightarrow \mathbb{R}^+$

$$\mathbb{E}[f(X_{n+1}) | X_n = y] = \int_{\mathbb{R}^d} f(x) \pi(y; dx)$$

$$\mathbb{E}[f(X_0)] = \int_{\mathbb{R}^d} f(x) \mu_0(dx)$$

L'observation est un processus tel que:

$$Y_n(x) = h(X_n) + V_n$$

h est une fonction mesurable de \mathbb{R}^d dans \mathbb{R}^k et V_n est une suite de variables c'est à dire de loi commune $\nu(dx)$.

On a la proposition suivante sur la prédiction d'après la formule de Bayes:

$$\mu_n^-(dx) = \int_{\mathbb{R}^d} \pi(x'; dx) \mu_{n-1}^-(dx')$$

2.1 Cas de bruits à densité

Supposons que le bruit d'observation suit la loi de densité f_V . On admet les résultats suivants qui établissent des relations de récurrence:

La loi de Y_n sachant \mathcal{Y}_{n-1} est donnée par:

$$\mathbb{P}_{Y_n}(dy | \mathcal{Y}_{n-1}) = \left[\int_{\mathbb{R}^d} f_V(y - h(x)) \mu_n^-(dx) \right] dy$$

La loi conditionnelle de (X_n, Y_n) sachant \mathcal{Y}_{n-1} est donnée par:

$$\mathbb{P}_{(Y_n, X_n)}(dx, dy | \mathcal{Y}_{n-1}) = f_V(y - h(x)) \mu_n^-(dx) dy.$$

D'où la loi corrigée est :

$\mu_n(dx) = C_n f_V(Y_n - h(x)) \mu_n^-(dx)$ où C_n est une constante de normalisation telle que

$$C_n^{-1} = \int_{\mathbb{R}^d} f_V(Y_n - h(x)) \mu_n^-(dx)$$

2.2 Cas de bruits d'observation uniforme

On suppose que la loi de la condition initiale X_0 admet une densité $p_0(x)$. On a:

Les lois conditionnelles prédites et corrigées admettent des densités :

$$\begin{aligned}\mu_n^-(dx) &= p_n^-(x)dx, \mu_n(dx) = p_n(x)dx \\ p_n^-(x) &= \int_{\mathbb{R}^d} f_W(x - f(x')) p_{n-1}(x') dx' \\ p_n(x) &= \frac{f_V(Y - h(x)) p_n^-(x)}{\int_{\mathbb{R}^d} f_V(Y - h(x)) p_n^-(x) dx}\end{aligned}$$

Si de plus, le bruit d'observation est uniforme sur $\mathcal{D} = [-a; a]^d$, $a > 0$; $f_V(x) = \frac{1}{m(\mathcal{D})} \mathbb{I}_{\mathcal{D}}(x)$ m est la mesure de Lebesgue sur \mathbb{R}^d .

$$\begin{aligned}p_n(x) &= C_n \cdot \mathbb{I}_{\mathcal{D}}(Y_n - h(x)) p_n^-(x) \\ C_n^{-1} &= \int_{\mathbb{R}^d} \mathbb{I}_{\mathcal{D}}(Y_n - h(x)) p_n^-(x) dx\end{aligned}$$

On peut donc avoir $p_n(x)$ en fonction de $p_{n-1}(x)$, sans faire la décomposition filtre prédit et filtre corrigé.

Avec une hypothèse supplémentaire où h est un difféomorphisme, on peut même simplifier cette relation.

Dans le cas général, le problème de filtrage non linéaire peut se résoudre par le filtrage de Kalman, pour les bruits gaussiens, et par le filtrage particulaire ou SIR (Sampling with Importance Resampling) pour les bruits non gaussiens.

3 Filtrage particulaire

L'idée de la simulation particulaire [5], [6] est d'approximer une suite de loi $p_k(x)$ par une suite de loi $p_k^N(x) = \sum_{i=1}^N \omega_k^i \delta_{\xi_k^i}(x)$ où δ est la mesure de Dirac; c'est la méthode de Monte Carlo séquentielle.

ξ_k^i sont appelées des particules, et on a $\xi_k^i \sim p_k(x)$ et $p_k(\xi_k^i) = \omega_k^i$ est le poids de la particule ξ_k^i .

Il se peut dans la simulation que les poids ω_k^i soient presque nuls et ne contribuent plus à l'approximation. On dit qu'il y a dégénérescence des particules. Ces particules de poids négligeables devront être remplacées par les particules de poids plus importantes non négligeables. On a donc un regroupement de parti-

cules dans une région d'état pour ralentir la dégénérescence.

Dans ce cas, on redistribue les particules de poids plus importantes pour compléter le nombre total des particules. Cette redistribution ne doit pas se faire sur un petit nombre de particules sinon on perd l'idée de convergence de la méthode.

On peut considérer un rééchantillonnage multinomial. A partir des particules ξ_k^1, \dots, ξ_k^N avec les probabilités $\omega_k^1, \dots, \omega_k^N$, on fait un échantillonnage de N particules de poids $1/N$. Cela revient à multiplier les particules les plus importantes.

Le filtrage particulaire est une méthode numérique qui est une approche adaptée au cas de modèle d'observation non linéaire avec des bruits non-gaussiens.

Considérons une variable aléatoire X à valeurs dans \mathbb{R}^d et de densité p_X , i.e

$$\mathbb{E}\phi(X) = \int_{\mathbb{R}^d} \phi(x) p_X(x) dx, \quad \forall \phi \in \mathcal{C}_b(\mathbb{R}^d)$$

De même $p_{X,Y}$, la densité conjointe de X et Y est définie par :

$$\begin{aligned}\mathbb{E}\phi(X, Y) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(x) p_{X,Y}(x, y) dx dy, \\ \forall \phi &\in \mathcal{C}_b(\mathbb{R}^d \times \mathbb{R}^d)\end{aligned}$$

On définit la densité conditionnelle de X sachant $Y = y$, notée $p_{X|Y=y}$:

$$p_{X|Y=y}(x) := \frac{p_{X,Y}(x, y)}{p_Y(y)} = \frac{p_{X,Y}(x, y)}{\int_{\mathbb{R}^d} p_{X,Y}(x, y) dx}$$

L'espérance conditionnelle associée à cette densité conditionnelle est:

$$\mathbb{E}\phi(X|Y = y) = \frac{\int_{\mathbb{R}^d} \phi(x) p_{X,Y}(x, y) dx}{\int_{\mathbb{R}^d} p_{X,Y}(x, y) dy}$$

3.1 Echantillonnage d'importance

Considérons une loi $p(x)$ absolument continue par rapport à une autre loi $\tilde{p}(x)$:

$$p(x) = 0 \Rightarrow \tilde{p}(x) = 0.$$

$$\mathbb{E}\phi(X) = \int_{\mathbb{R}^d} \phi(x) p(x) dx$$

$$= \int_{\mathbb{R}^d} \phi(x) \frac{p(x)}{\tilde{p}(x)} \tilde{p}(x) dx, \quad \forall \phi \in \mathcal{C}_b(\mathbb{R}^d)$$

Si on simule un N échantillons $\tilde{\xi}_1, \dots, \tilde{\xi}_N$ à partir de la loi $\tilde{p}(x)$, alors on peut avoir l'approximation par la méthode de Monte Carlo.

$$\mathbb{E}\phi(X) \simeq \sum_{i=1}^N \omega^i \phi(\tilde{\xi}_i)$$

Où $\tilde{\xi}_1 \sim \tilde{p}(x)$ et $\omega^i = p(\tilde{\xi}_i)/\tilde{p}(\tilde{\xi}_i)$

3.2 Filtre SIR ou bootstrap

La base du filtrage particulaire [4] est de déterminer des approximations du filtre prédit et du filtre corrigé sous la forme:

$$\mu_n^-(dx) \simeq \sum_{i=1}^N \omega_{n-}^i \delta_{\xi_{n-}^i}(x) dx$$

$$\mu_n(dx) \simeq \sum_{i=1}^N \omega_n^i \delta_{\xi_n^i}(x) dx$$

A chaque instant k , on détermine des suites $(\omega_{n-}^i, \xi_{n-}^i)$ et (ω_n^i, ξ_n^i) , $i=1, \dots, N$.

Avec le modèle $\begin{cases} Y_n = h(X_n) + V_n \\ X_{n+1} = f(X_n) + W_n \end{cases}$

$X_0 \sim \mu_0(dx)$, $W_n \sim p_W(w)dw$ et $V_n \sim p_V(v)dv$

$$p_n^-(x) \simeq \sum_{i=1}^N \omega_{n-}^i \delta_{\xi_{n-}^i}(x)$$

$$p_n(x) \simeq \sum_{i=1}^N \omega_n^i \delta_{\xi_n^i}(x)$$

Considérons le filtre SIS ou Sequential Importance Sampling et le filtre SIR.

3.2.1 Filtre SIS

Dans l'étape de prédiction avec la formule de Bayes [5],

$\xi_n^i = f(\xi_{n-1}^i) + W_n^i$ avec $W_n^i \sim p_W(w)dw$.

Et dans l'étape de correction,

$$\omega_n^i = \frac{\omega_{n-1}^i p_V(Y_n - h(\xi_n^i))}{\sum_{j=1}^N \omega_{n-1}^j p_V(Y_n - h(\xi_n^j))}$$

Par contre, le filtre SIR utilise le rééchantillonnage.

3.2.2 Filtre SIR

On commence par un rééchantillonnage [4], c'est l'étape de sélection, c'est à dire, on choisit $\hat{\xi}_{n-1}^i$ parmi $(\xi_{n-1}^1, \dots, \xi_{n-1}^N)$ en fonction des poids $(\omega_{n-1}^1, \dots, \omega_{n-1}^N)$. On peut vérifier la valeur du paramètre

$$N_n^{eff} = \frac{1}{\sum_{i=1}^N \omega_n^i} \in [1, N]$$

qui représente le nombre efficace des particules.

Quand N_n^{eff} est proche de N , alors les particules sont d'égale importance.

En général, ce n'est pas le cas et on doit redistribuer les particules pour avoir des poids très proches de $1/N$.

La redistribution se fait selon la loi multinomiale qui élimine les particules de poids négligeables en faveur de celles qui apportent plus de contributions dans l'approximation.

La prédiction est une étape de mutation, $\xi_n^i = f(\hat{\xi}_{n-1}^i) + W_n^i$ avec $W_n^i p_W(w)dw$

La correction est une étape de pondération, telle que

$$\omega_n^i = \frac{p_V(Y_n - h(\xi_n^i))}{\sum_{j=1}^N p_V(Y_n - h(\xi_n^j))}$$

La dégénérescence des poids peut être résolue par une redistribution de Kitagawa, consistant à construire une suite arithmétique de raison $1/N$ de N nombres telle que le premier terme est simulé selon la loi uniforme $\mathcal{U}\left[0, \frac{1}{N}\right]$. Cet algorithme a une complexité de l'ordre de N mais n'est pas toujours efficace.

4 Applications

4.1 Cas de bruit uniforme

Supposons que

$$\begin{cases} X_{n+1} = X_n + W_n & (1) \\ Y_n = X_n + V_n & (2) \end{cases}$$

W_n et V_n sont des processus c'est à dire de densités f_W et f_V . W_n et V_n sont indépendantes entre elles et indépendantes de la condition initiale X_0 .

Supposons que $d = 1$, le bruit de mesure est gaussien centré de matrice de variance Q et le bruit d'observation est uniforme sur un domaine borné : $\mathcal{D} = [-a, a]$.

$$f_W(w) = g_Q(w), f_V(v) = \frac{1}{m(\mathcal{D})} \mathbb{I}_{\mathcal{D}}(v)$$

où g_Q est la densité gaussienne $\mathcal{N}(0, Q)$ sur \mathbb{R} avec $Q > 0$.

$m(\mathcal{D})$ est la mesure de Lebesgue de \mathcal{D} , on suppose également que $\int_{\mathcal{D}} v dv = 0$.

On obtient:

$$p_n^-(x) = \int_{\mathbb{R}} g_Q(x - x') p_{n-1}^-(x') dx'$$

$$p_n(x) = C_n \mathbb{I}_{\mathcal{D}+Y_n}(x) p_n^-(x)$$

$$C_n^{-1} = \int_{\mathbb{R}} \mathbb{I}_{\mathcal{D}+Y_n}(x) p_n^-(x) dx$$

$$\mathcal{D} + Y_n = \{y \in \mathbb{R}; y - Y_n \in \mathcal{D}\}$$

C_n est la constante de normalisation.

La densité $p_n(x)$ est portée par le domaine $\mathcal{D} + Y_n$ pour tout $n \geq 1$ et posons $q_n(x) = p_n(Y_n - x)$ et on a sur \mathcal{D} $q_n(x) =$

$$= C_n \int_{\mathbb{R}} g_Q(Y_n - Y_{n-1} - x + x') q_{n-1}(x') dx'$$

D'où l'algorithme de calcul du filtre :

Algo Filtre (N, N_1 : entiers ; p_0 : densité; g fonctions ; Y_1, \dots, Y_N : observations)

Filtre q_n

Début

Initialisation de q_0 à p_0

Pour $n=1$ à N faire

Pour $k=1$ à N_1

Simuler X_k par q_{n-1}

Calculer $g_Q(Y_n - Y_{n-1} - x + X_k)$

Fin pour k

$$q_n(x) = \frac{C_n}{N_1} \sum_{k=1}^{N_1} g_Q(Y_n - Y_{n-1} - x + X_k)$$

Fin pour n

Fin

Les signaux mesuré et observé sont représentés à la figure suivante (figure 1)

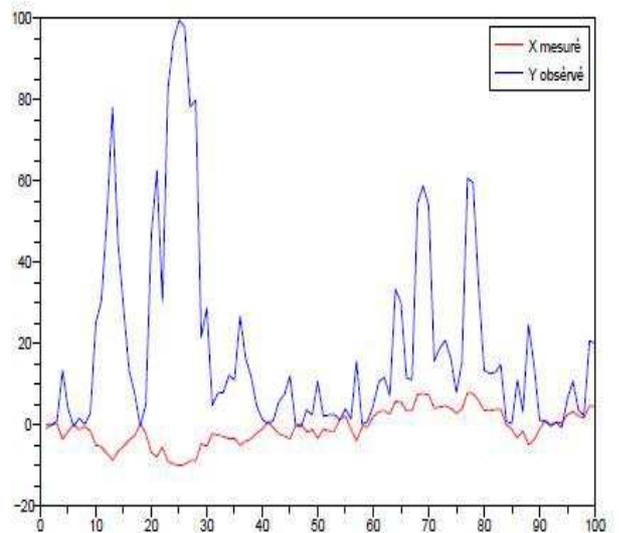


Figure 1: Signal mesuré et signal observé

Il en est de même pour les filtres p_n ; $n \leq 30$ (figure 2)

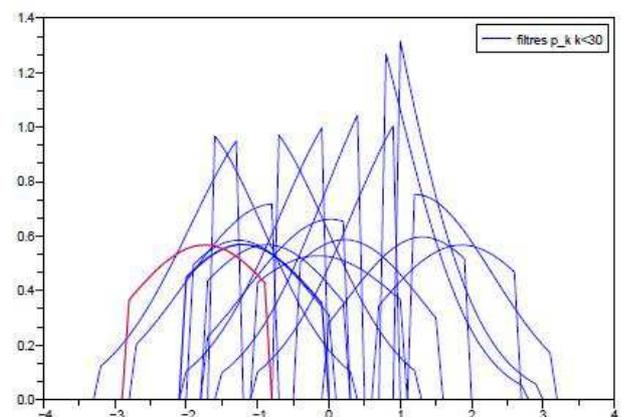


Figure 2 : filtres p_n ; $n \leq 30$

4.2 Filtrage particulière, suivi de mobile

Considérons un mobile qui se déplace dans un plan où sa position est le vecteur d'état à l'instant n , $X_n = \begin{pmatrix} X_n^1 \\ X_n^2 \end{pmatrix}$.

On a le système d'équations d'état:

$$\begin{cases} X_{n+1}^1 = X_n^1 + W_n^1 \\ X_{n+1}^2 = X_n^2 + W_n^2 \end{cases}$$

W^1 et W^2 sont des bruits blancs gaussiens.

La loi de transition est une loi gaussienne centrée.

On dispose de k -capteurs pour les observations. A chaque instant n , on dispose de la mesure $Y_n^i = h_i(X_n) + V_n^i$ pour chaque capteur i .

h_i^1 mesure la distance euclidienne entre le capteur i et le mobile h_i^2 est l'angle polaire entre l'horizontal et le segment capteur-signal. V_n^i est l'erreur de mesure suivant une loi uniforme.

On a le système d'équations d'observations:

$$\begin{cases} Z_{i,n}^1 = \sqrt{(X_n^1)^2 + (X_n^2)^2} + V_{i,n}^1 \\ Z_{i,n}^2 = \arctan\left(\frac{X_n^2}{X_n^1}\right) + V_{i,n}^2 \end{cases}$$

$V_{i,n}^1$ et $V_{i,n}^2$ sont des bruits d'observations uniformes du capteur i .

On reproduit le système d'observation par le système:

$$\begin{cases} Y_{i,n}^1 = Z_{i,n}^1 \cdot \cos(Z_{i,n}^2) \\ Y_{i,n}^2 = Z_{i,n}^1 \cdot \sin(Z_{i,n}^2) \end{cases}$$

L'algorithme se présente comme ceci :

Algo Particulaire (p_{X_0} , h , f : fonctions ; $\sigma_W > 0$, p_V densité ; N_1, N : entiers.

$$\text{particules } (\hat{\xi}_N^1, \dots, \hat{\xi}_N^N)$$

Début

Initialisation des particules : $\hat{\xi}_1^0, \dots, \hat{\xi}_N^0$

$$\hat{\xi}_n^0 \sim p_{X_0}, \omega_n^0 = \frac{1}{N}$$

Pour $i=1$ à N faire

Pour $n=1$ à N faire

Mada-ENELSA Vol 1, 2013

Mutation

$$\xi_n^i = f(\hat{\xi}_{n-1}^i) + W_n^i$$

Pondération

$$\omega_n^i = \frac{p_V(Y_n - h(\xi_n^i))}{\sum_{j=1}^N p_V(Y_n - h(\xi_n^j))}$$

Fin pour n

$$N_n^{eff} = \frac{1}{\sum_{i=1}^N \omega_n^i}$$

Test si $\frac{N_n^{eff}}{N} \leq 0.75$, alors

Sélection : redistribution des particules

$$(\xi_{n-1}^1, \dots, \xi_{n-1}^N)$$

Si non continuer

$$(\hat{\xi}_N^1, \dots, \hat{\xi}_N^N) = (\xi_N^1, \dots, \xi_N^N)$$

Fin pour i

Fin

Les signaux mesuré et observé par un capteur sont représentés à la figure 3

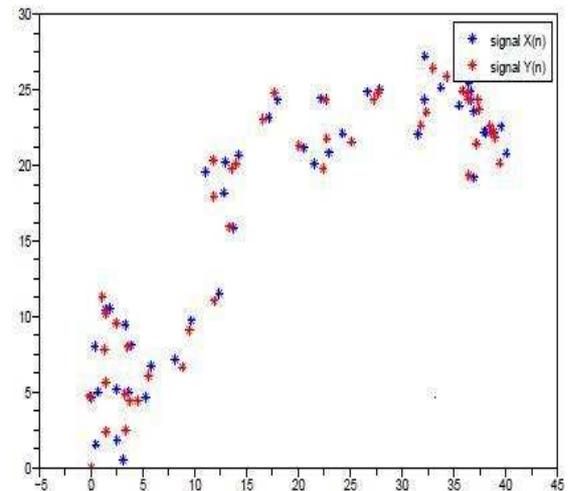


Figure 3: Signaux mesurés et observés par un capteur

Dans cette application: $(p_V(x) = \mathbb{I}_{[0;a]}(x)$ avec $a > 0$.

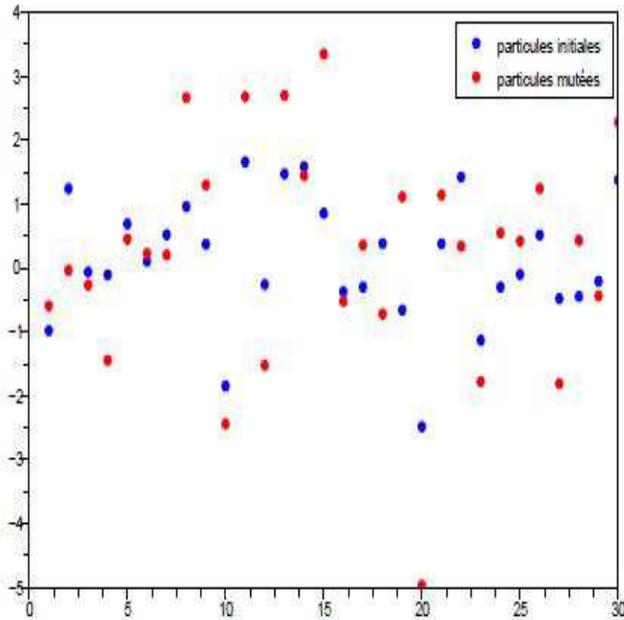


Figure 4: Processus de mutation

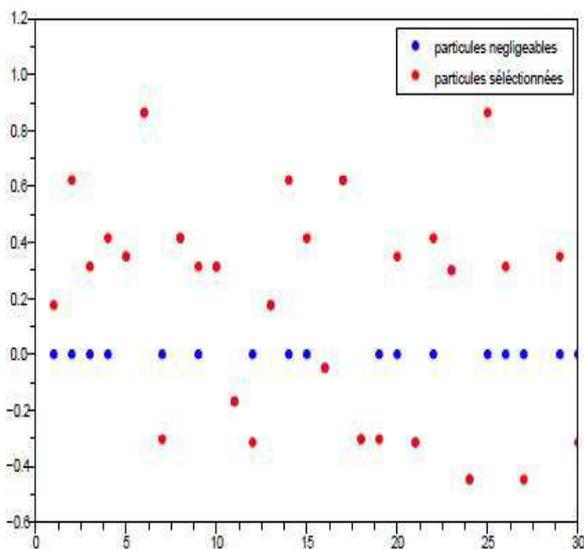


Figure 5: Processus de sélection

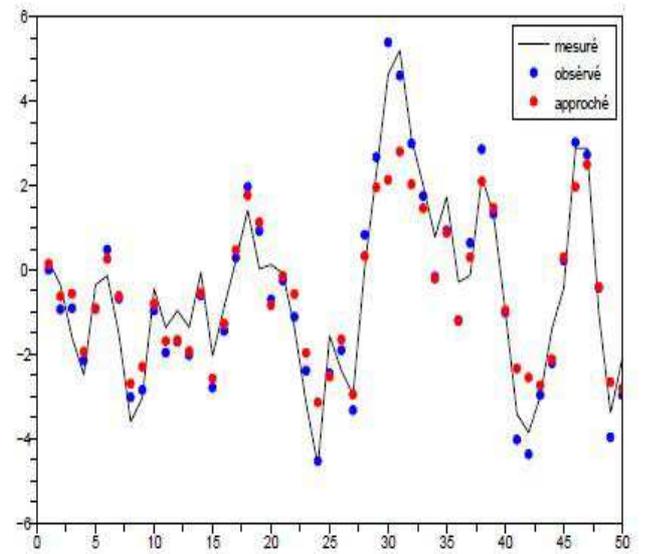


Figure 6: Filtrage particulaire

4.3 Interprétations

La particularité des densités des filtres est d'avoir des supports compacts bornés [8] dus aux bruits d'observation uniforme. Figure 1, Figure 2 et Figure 3.

La méthode d'approximation des intégrales de Monte Carlo est nécessaire. La solution théorique est assez compliquée voire difficile à mettre en œuvre dans la simulation quand la dimension est supérieure à 2. Par contre, le filtre particulaire est plus adapté dans la généralisation.

La mutation des particules est assez simple à élaborer dans la Figure 4. La vraisemblance est une forme de distribution uniforme et la normalisation entraîne une distribution de particules équiprobables.

Les particules de poids nuls sont éliminées et remplacées par une redistribution des particules de poids non nuls. Figure 5.

Le filtre est représenté par son espérance mathématique à chaque instant. Ces moyennes

approchent à chaque instant d'observation la trajectoire réelle. Figure 6.

5 Conclusion

Le filtrage particulaire [6] est une des méthodes adaptées à la résolution d'un problème de filtrage non linéaire. Sa mise en œuvre est conditionnée par la connaissance de la loi de distribution initiale, de la loi de transition de la chaîne de Markov, de la loi d'émission et d'un programme de redistribution. Pour un système à bruit d'observation uniforme, les particules sont distribuées de façon équiprobable. Une sélection parmi les particules importantes est nécessaire pour avoir le nombre suffisant pour approximer le filtre. La convergence est justifiée, comme dans toutes les méthodes de Monte Carlo par le théorème central limite et la loi des grands nombres.

La complexité des algorithmes est assez raisonnable et la stabilité du filtre par rapport à la mesure initiale peut être étudiée.

Un des inconvénients du filtrage particulaire se trouve dans le rééchantillonnage des particules si le bruit d'observation est trop faible. Dans ce cas, on peut considérer la méthode MCMC pour générer les particules [7].

On étend l'étude pour un système où l'équation d'évolution est une équation différentielle stochastique c'est à dire une équation de la dynamique dans le temps continu.

Références

[1] E.PARDOUX. Mathématiques pour l'ingénieur. DESS Université de PROVENCE 1991.

[2] F-E. RACICOT R.THEORET, Quelques applications du filtre de Kalman en Finance: Estimation et Prévision de la volatilité Stochastique et du rapport cours-bénéfices. Université de Montréal, 2005.

[3] E.PARDOUX. Filtrage non linéaire et équations aux dérivées partielles stochastiques associées. In P.L Hennequin, editor, Ecole d'été

de Probabilités de Saint-Flour XIX-1989, volume 1464 of Lecture Notes in Mathematics, pages 67-163. Springer Verlag, Berlin, 1991.

[4] Fabien CAMPILLO. Filtrage particulaire et modèles de Markov cachés. Master 2 Université de Toulon, 2006.

[5] François Le GLAND. Filtrage Bayésien et Approximation Particulaire. Cours Ecole Nationale Supérieure de Techniques Avancées, module Commande des Systèmes, 2011.

[6] Nicolas PAUL. Filtrage Particulaire. Séminaire de Statistique, Conservatoire National des Arts et Métiers, 2006.

[7] Nicolas WIDYNSKI. Filtre particulaire: suivi multi-objets dans les séquences multimodales. Lip6-Télécom Paris Tech, 2008.

[8] M.JEANBLANC, T.SIMON. Éléments de calculs stochastique. IRBID Septembre 2006.