Approche non paramétrique du filtrage de système non linéaire à temps discret et à paramètres inconnus

Filtering discrete time nonlinear system with unknown parameters: a non parametric approach

Vivien Rossi [a] Jean-Pierre Vila [b]

[a] *UMR I3M, Equipe de Probabilités et Statistique, Université Montpellier II cc 51, Place Eugène Bataillon, 34095 Montpellier cedex 5. Tel (+33) 04 67 14 46 07, Fax 04 67 14 35 58*

[b] *UMR Analyse des Systèmes et Biométrie, ENSAM-INRA 2, Place Pierre Viala, 34060 Montpellier cedex 1. Tel. 04 67 99 61 22 77, Fax 04 67 52 14 27*

Abstract

Particle filters are presently among the most powerful tools for filtering discrete time non linear systems. However the presence of unknown parameters in the system equations makes their use more complex and can even impair their convergence properties. This paper shows how an on-line consistent estimation of these parameters can be obtained simultaneously to that of the state variables to be filtered. This approach relies upon a kernel-based non parametric estimation of conditional probability densities from successive Monte Carlo generations of system particles.

Résumé

Les filtres particulaires sont actuellement les outils de filtrage de système non linéaire à temps discret les plus performants. Toutefois la présence de paramètres inconnus dans les équations du système rend leur mise en œuvre délicate et compromet souvent leur convergence. Cet article montre comment une estimation en ligne convergente de ces paramètres peut être obtenue simultanément à celles des variables d'état à filtrer. Elle repose sur une méthode d'estimation non paramétrique de densités de probabilités conditionnelles par noyau de convolution, à partir de générations successives de particules du système.

Abridged English version

This paper deals with the problem of filtering non linear discrete time system of the general form

$$\begin{cases} x_{t+1} = f_t(x_t, \theta_x, \varepsilon_t) \\ y_{t+1} = h_t(x_{t+1}, \theta_y, \eta_t) \end{cases} \tag{1}$$

Email addresses: rossiv@ensam.inra.fr (Vivien Rossi), vila@ensam.inra.fr (Jean-Pierre Vila).

in which f_t, h_t are known Borelian functions, $x_t \in \mathbb{R}^d$ is the vector of state variables, $y_t \in \mathbb{R}^q$ is the vector of observed variables, $\theta = (\theta_x^T, \theta_y^T)^T \in \Theta \subset \mathbb{R}^p$ a vector of fixed unknown parameters, ε_t, and η_t are vectors of independent white noises.

The first purpose of filtering is to estimate at each time t the probability distribution of the state vector x_t conditional on the observed variables up to time t, the so-called optimal filter, or if it exists its density function $p_t(x|y_1, \ldots, y_t)$. Among the most powerful methods for filtering nonlinear systems are the particle filters which proceed through successive Monte Carlo model simulations of state particles $\tilde{x}_t^i$, allowing a convergent empirical estimation of the state conditional probability distribution (Del Moral [2]). Typically in this family of particle methods, the evolution of the particles at each step is controlled through their likelihood with respect to the observed variables. These methods are thus dependent on the availability of the analytic form of this likelihood function and on the presence of non nul observation noises. Moreover, when unknown parameters are present as in model (1) any filtering procedure has to take account of these parameters θ in the model equations and must be backed up with a simultaneous parameter estimation procedure. The estimation of θ can even be the main objective. Standard particle filter methods do not cope easily with this issue (Doucet et al. [6]).

In the following a convergent kernel-based non parametric particle procedure is proposed to perform estimation of both state and parameter vector conditional densities. It generalizes a state-dedicated non parametric conditional density estimation procedure recently developped (Rossi & Vila [13]). This new procedure relies first on the introduction into the system equations (1) of the natural model $\theta_{t+1} = \theta_t$ for the parameters whose possible values are endowed with an a priori probability density p_0^θ, the support of which is supposed to cover the true unknown parameter values θ^*. The system model (1) so completed is used to generate by simulation at each successive time step t, n particles $(\tilde{x}_t^i, \tilde{\theta}_t^i, \tilde{y}_t^i), i = 1, \ldots, n$, from which and from the current observed y_t value, kernel-based conditional density estimates of the state vector x_t and the parameter vector θ_t are built up. The recursive use of these successive density estimates in the particle generation from step to step, ensures their conditioning on the successive observations up to the one at hand.

For a given time step t, almost sure L_1-convergence results of these conditional density estimates to their true counterparts, as the number n of generated particles grows to infinity, are established with their rate of convergence. These L_1-convergence results have no equivalent in the standard particle filter methods which can only proceed with discrete approximations of conditional probability measures. Most remarkably, these convergence results are obtained under simple assumptions as $\lim_{n\to\infty} nh_n^{d+p+q}/\log n = \infty$ and $\lim_{n\to\infty} h_n = 0$ with h_n the kernel bandwidth parameter, which are independent of the time step t.

Moreover, almost sure convergence of conditional expectation estimates $\widehat{x}_t^n$ and $\widehat{\theta}_t^n$ of the state vector x_t and the parameter vector θ as n grows to infinity, are also provided.

Finally, a direct application of a theorem of Schervish ([14]) ensures that the conditional θ_t probability density function concentrates around the true unknown parameter value θ^* as t grows to infinity and then the same is true for its filter estimate when in addition n tends also to infinity. In a similar way the direct application of a theorem of Schwartz ([15]) ensures that the conditional θ_t expectation $\mathbb{E}[\theta_t|y_{1:t}]$, converges to θ^* almost surely as t tends to infinity and the same is true for its filter estimate $\widehat{\theta}_t^n$ when in addition n grows to infinity.

1. Introduction

Les équations d'état et d'observation d'un système dynamique contiennent souvent des paramètres fixes mais inconnus. Un tel contexte est formalisé par le système (1). L'idée d'utiliser des méthodes de filtrage particulaire pour estimer à la fois les variables d'état x_t non observées du système et les paramètres

inconnus constants θ du modèle n'est pas nouvelle (voir Doucet et al. ([6]) ou Liu & West ([12]) pour un état de l'art). Elle repose sur l'introduction d'une loi a priori $\pi_0(\theta)$ pour les paramètres et une extension du vecteur d'état qui devient $z_t = (x_t, \theta_t)$, et consiste à réestimer cette loi par approximation discrète à chaque pas de temps. Elle se heurte généralement au problème d'un modèle d'évolution à introduire pour ces paramètres. Pour les méthodes particulaires classiques (Doucet [6]) l'introduction du modèle le plus naturel $\theta_{t+1} = \theta_t$ restreint l'exploration de l'espace des paramètres à la phase d'initialisation de l'algorithme, empéchant sa convergence. Différentes approches ont été proposées pour pallier cette difficulté : introduction d'une dynamique d'évolution artificielle de type simple $\theta_{t+1} = \theta_t + \zeta_t$ ou plus complexe (Kitagawa [10], Higuchi [8]), au risque de dénaturer la dynamique réelle du système; introduction d'un modèle de Markov caché (Gilks & Berzuini [7]) et d'une étape de type chaîne de Markov de Monte-Carlo pour accroître la diversité particulaire, sans conduire toutefois à l'estimation proprement dite des paramètres. Pour éviter ces additions une approche par régularisation de l'estimation discrète à chaque pas de la loi a posteriori des paramètres, source de difficultés pour les méthodes particulaires classiques, a également été proposée (West [16], Liu & West [12]), au prix d'une complexification supplémentaire des algorithmes. Plus généralement, les techniques de régularisation sont utilisées sur les filtres particulaires pour éviter leur dégénérescence. Le cas le plus fréquent, la régularisation du modèle d'état est présentée et étudiée dans Del Moral & Miclo([4]) et LeGland & Oudjane([11]). Cependant cette approche souffre encore de certaines des restrictions des méthodes classiques (nécessité de non nullité des variables de bruit et de la disponibilité analytique de leurs lois). Cette restriction a été levée dans Del Moral & al.([3]) par la régularisation du modèle d'observation mais le modèle d'état étant non régularisé, l'approche demeure sensible au problème de dégénérescence des filtres particulaires. Afin de contourner simultanément ces deux problèmes, Rossi & Vila ([13]) ont régularisé conjointement le modèle d'état et le modèle d'observation. Leur démarche peut s'interpréter comme une généralisation des approches précédentes grâce à une extension du concept de particule incluant les variables d'état et les variables d'observation. Toutefois, la construction et l'étude théorique des filtres correspondants sont différentes car elles s'appuient sur l'estimation non paramétrique, par noyau de convolution, des densités conditionnelles.
L'approche présentée ici pour l'estimation des paramètres inconnus du système (1), s'inspire des résultats de Rossi & Vila ([13]). Comme pour les approches traditionnelles évoquées ci-dessus, le vecteur d'état est étendu au paramètre : $z_t = (x_t, \theta_t)$. Cependant, le fait de travailler avec un modèle d'état régularisé permet de respecter formellement le modèle naturel $\theta_{t+1} = \theta_t$, pour les paramètres inconnus, sans rencontrer les problèmes habituels de dégénérescence.

2. Un algorithme non paramétrique de filtrage particulaire et d'estimation de paramètres

Dans sa composante filtrage, cette présentation est restreinte à l'estimation de la densité conditionnelle du vecteur d'état x_t à l'instant t, $p_t(x|y_1, \cdots, y_t)$, mais l'approche présentée s'adapte immédiatement à l'estimation de la densité conditionnelle de toute la trajectoire des états jusqu'à l'instant t soit $p_t(x^1, \cdots, x^t|y_1, \cdots, y_t)$, ou à celle de la densité prédictive de l'état à l pas en avant $p_{t+l}(x|y_1, \cdots, y_t)$.

2.1. *Rappel et notations*

- Soit $X_1, \cdots, X_n$ des variables aléatoires de $\mathbb{R}^d$ indépendantes et identiquement distribuées. L'estimateur g_n de leur densité commune g, associé à un noyau donné K est ([1])

$$g_n(x) = \frac{1}{nh_n^d} \sum_{i=1}^{n} K(\frac{x - X_i}{h_n}) = (K_{h_n} * \mu_n)(x) \qquad x \in \mathbb{R}^d$$

où h_n, réel lié à n, est le paramètre de fenêtre, et $\mu_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}$ est la mesure empirique associée aux $X_1, \cdots, X_n$. En d'autres termes, g_n est la densité obtenue en régularisant la mesure empirique des $X_1, \cdots, X_n$, par convolution avec $\frac{1}{h_n}K(\frac{\cdot}{h_n})$. On notera par commodité $K_{h_n}(x) = \frac{1}{h_n^d}K(\frac{x}{h_n})$ $x \in \mathbb{R}^d$.

- Soit $W^{s,r}(\Omega) = \{f \in L_r(\Omega) | D^\alpha f \in L_r(\Omega), \forall\alpha : |\alpha| \leq s\}$ espace de Sobolev standard.
- On notera $p_t(x|y_{1:t})$ et $p_t(\theta|y_{1:t})$ les densités de x_t et de θ_t conditionnellement aux observations $y_1, \ldots, y_t$ et par $p_t^n(x|y_{1:t})$, $p_t^n(\theta|y_{1:t})$ leurs estimations par noyau avec n particules.
- Les densités conditionnelles $p_t(x|y_{1:t})$ et $p_t(\theta|y_{1:t})$ étant définies $P_{Y_{1:t}}$ presque partout, nous utilisons les conventions $p_t(x|y_{1:t}) = p_t(\theta|y_{1:t}) \doteq 0$ si $p_t(y_{1:t}) = 0$ et $p_t^n(x|y_{1:t}) = p_t^n(\theta|y_{1:t}) \doteq 0$ si $p_t^n(y_{1:t}) = 0$, pour leurs estimations par noyaux avec n particules.

2.2. *Hypothèses*

- Soit p_0^x la densité de probabilité du vecteur d'état à l'instant $t = 0$, supposée connue.
- Soit p_0^θ une densité de probabilité représentant une connaissance a priori sur les valeurs possibles des paramètres $\theta \in \Theta$, non nulle en les vraies valeurs θ^* de ces paramètres inconnus.
- Le modèle d'état f_t et le modèle d'observation h_t sont supposés connus à chaque instant t et structurellement exacts pour le processus réel étudié.
- Les lois de probabilités $\mathcal{L}_{\varepsilon_t}$ et $\mathcal{L}_{\eta_t}$ caractérisant respectivement les bruits ε_t et η_t éventuellement nuls, sont simulables sinon connues analytiquement, pour tout t.

Remarque : Les filtres particulaires classiques requièrent la non dégénérescence et la connaissance analytique de la densité $p(y_t|x_t)$ qui intervient dans la pondération des particules générées.

2.3. *Algorithme*

Soit $K_{h_n}^x, K_{h_n}^\theta$ et $K_{h_n}^y$ des noyaux de Parzen-Rosenblatt ([1]) de dimension d, p et q respectivement.
Soit $n > 0$ donné. L'algorithme procède par simulation de n particules à chaque pas de temps.

- Etape $t = 0$: Pour $i = 1, \ldots, n$: $\bar{x}_0^i \sim p_0^x$, $\bar{\theta}_0^i \sim p_0^\theta$, $\tilde{\varepsilon}_0^i \sim \mathcal{L}_{\varepsilon_0}$, $\tilde{\eta}_1^i \sim \mathcal{L}_{\eta_1}$. $t = t + 1$.
- Etape $t > 0$: ∘ Génération de n particules, pour $i = 1, \ldots, n$
  - si $t = 1$: $\tilde{x}_1^i = f_1(\bar{x}_0^i, \bar{\theta}_0^i, \tilde{\varepsilon}_0^i)$, $\tilde{\theta}_1^i = \bar{\theta}_0^i$, $\tilde{y}_1^i = h_1(\tilde{x}_1^i, \tilde{\theta}_1^i, \tilde{\eta}_1^i)$.
  - si $t > 1$: $(\bar{x}_{t-1}^i, \bar{\theta}_{t-1}^i) \sim p_{t-1}^n(x, \theta|y_{1:t-1})$, $\tilde{\varepsilon}_{t-1}^i \sim \mathcal{L}_{\varepsilon_{t-1}}$, $\tilde{\eta}_t^i \sim \mathcal{L}_{\eta_t}$
    $\tilde{x}_t^i = f_t(\bar{x}_{t-1}^i, \bar{\theta}_{t-1}^i, \tilde{\varepsilon}_{t-1}^i)$, $\tilde{\theta}_t^i = \bar{\theta}_{t-1}^i$, $\tilde{y}_t^i = h_t(\tilde{x}_t^i, \tilde{\theta}_t^i, \tilde{\eta}_t^i)$

  ∘ Estimation des densités conditionnelles

$$
\begin{aligned}
p_t^n(x, \theta|y_{1:t}) &= \Big(\sum_{i=1}^{n} K_{h_n}^y(\tilde{y}_t^i - y_t) \times K_{h_n}^\theta(\tilde{\theta}_t^i - \theta) \times K_{h_n}^x(\tilde{x}_t^i - x)\Big) / \sum_{i=1}^{n} K_{h_n}^y(\tilde{y}_t^i - y_t) \\
p_t^n(\theta|y_{1:t}) &= \Big(\sum_{i=1}^{n} K_{h_n}^y(\tilde{y}_t^i - y_t) \times K_{h_n}^\theta(\tilde{\theta}_t^i - \theta)\Big) / \sum_{i=1}^{n} K_{h_n}^y(\tilde{y}_t^i - y_t) \\
p_t^n(x|y_{1:t}) &= \Big(\sum_{i=1}^{n} K_{h_n}^y(\tilde{y}_t^i - y_t) \times K_{h_n}^x(\tilde{x}_t^i - x)\Big) / \sum_{i=1}^{n} K_{h_n}^y(\tilde{y}_t^i - y_t)
\end{aligned}
$$

  ∘ $t = t + 1$, retour à l'étape t.

Remarque : le conditionnement par $y_1, \ldots, y_{t-1}$, non explicite dans la construction de ces estimateurs de densités conditionnelles, est pris en compte au travers de la loi $p_{t-1}^n(x, \theta|y_{1:t-1})$ qui sert à l'étape t à générer les n particules.

3. Propriétés de convergence du filtre

3.1. *Convergence des estimations des densités conditionnelles*

Théorème 3.1 (convergence L_1 p.s.) *Si $K^x_{h_n}, K^\theta_{h_n}$ et $K^y_{h_n}$ sont des noyaux de Parzen-Rosenblatt, positifs et bornés, si $p(y|y_{1:t-1})$ est continue et strictement positive en y_t pour tout t, alors*

$$\begin{cases} \lim\limits_{n\to\infty} \dfrac{nh_n^{q+d+p}}{\log n} = \infty \\ \lim\limits_{n\to\infty} h_n = 0 \end{cases} \Longrightarrow \begin{array}{l} \lim\limits_{n\to\infty} \|p_t^n(x,\theta|y_{1:t}) - p_t(x,\theta|y_{1:t})\|_{L1} = 0 \quad p.s. \\ \lim\limits_{n\to\infty} \|p_t^n(x|y_{1:t}) - p_t(x|y_{1:t})\|_{L1} = 0 \quad p.s. \\ \lim\limits_{n\to\infty} \|p_t^n(\theta|y_{1:t}) - p_t(\theta|y_{1:t})\|_{L1} = 0 \quad p.s. \end{array}$$

Preuve : la démonstration est une généralisation simple de la convergence L_1 p.s. du filtre à convolution de particules avec rééchantillonnage étudié par Rossi & Vila ([13]) dédié à l'estimation de la densité conditionnelle du vecteur d'état seul. Elle s'obtient par récurrence sur t et s'inspire d'une démonstration de Devroye([5]) assurant la robustesse des estimateurs à noyaux.

Corollaire 3.2 (vitesse de convergence L_1 intégrée) *Supposons que $p_Y = p(y_t|y_{1:t-1}) \in W^{s,1}(\mathbb{R}^q)$, que $p_{X\theta Y} = p(x_t, \theta_t, y_t|y_{1:t-1}) \in W^{s,1}(\mathbb{R}^{d+p+q})$ et que les noyaux $K^y_{h_n} \in L_1(\mathbb{R}^q)$ et $K^{x\theta y}_{h_n} = K^x_{h_n} K^\theta_{h_n} K^y_{h_n} \in L_1(\mathbb{R}^{d+p+q})$, sont de classe $s \geq 1$. Supposons de plus que pour $\varepsilon > 0$, pour $(K_{h_n}, f, \delta) \in \{(K^y_{h_n}, p_Y, q), (K^{x\theta y}_{h_n}, p_{X\theta Y}, d+p+q)\}$ on ait $\int \|u\|^{\delta+\epsilon} K_{h_n}(u)^2 \mathrm{d}x < \infty$ et $\int (1 + \|u\|^{\delta+\epsilon}) f(u) \mathrm{d}u < \infty$. Alors*

$$\mathbb{E}[\|p_t^n(x,\theta|y_{1:t}) - p_t(x,\theta|y_{1:t})\|_{L_1}] \leq u_t[O_t(h_n^s) + O_t(\frac{1}{\sqrt{nh_n^{d+q+p}}})] \quad \textit{avec } u_t = 2^t - 1 \tag{2}$$

où $O_t(.)$ est tel que $\lim_{v_n \to 0} O_t(v_n)/v_n \leq A_t$ avec $A_t > 0$ une constante liée au modèle (1) considéré. L'inégalité (2) est aussi vérifiée pour $\mathbb{E}[\|p_t^n(x|y_{1:t}) - p_t(x|y_{1:t})\|_{L_1}]$ et $\mathbb{E}[\|p_t^n(\theta|y_{1:t}) - p_t(\theta|y_{1:t})\|_{L_1}]$. Ces espérances sont définies par rapport à toutes les variables simulées $(\tilde{x}_t^i, \tilde{\theta}_t^i, \tilde{y}_t^i)$ et les variables d'observation $y_1, \ldots, y_t$.

Preuve : Ces vitesses de convergence s'obtiennent par généralisation du calcul de la vitesse de convergence du filtre à convolution de particules ([13]). La démonstration repose sur une décomposition de chacun des premiers membres de ces inégalités et des majorations utilisant les résultats de Holmström & Klemelä([9]).

Remarques :

- La condition $\lim_{n\to\infty} nh_n^{d+p+q}/\log n = \infty$ et $\lim_{n\to\infty} h_n = 0$ assurant la convergence du filtre est indépendante de l'instant t.
- Ces résultats de convergence L_1 n'ont pas d'équivalent pour les filtres particulaires standard qui procèdent par approximation discrète des mesures de probabilités conditionnelles du vecteur d'état.

3.2. *Estimations ponctuelles de l'état et des paramètres*

Soit $(\bar{x}_t^i, \bar{\theta}_t^i), i = 1, \ldots, n$, obtenus par tirage selon la loi de densité $p_t^n(x, \theta|y_{1:t})$ estimée après l'observation y_t. Soit $\widehat{x}_t^n = \frac{1}{n}\sum_{i=1}^n \bar{x}_t^i$ et $\widehat{\theta}_t^n = \frac{1}{n}\sum_{i=1}^n \bar{\theta}_t^i$.

Théorème 3.3 (convergence ponctuelle p.s.) *Si $Var[x_t, \theta_t|y_{1:t}]$ existe et est finie*

$$\begin{cases} \lim\limits_{n\to\infty} \dfrac{nh_n^{q+d+p}}{\log n} = \infty \\ \lim\limits_{n\to\infty} h_n = 0 \end{cases} \Longrightarrow \begin{array}{l} \lim\limits_{n\to\infty} |\widehat{x}_t^n - \mathbb{E}[x_t|y_{1:t}]| = 0 \quad p.s. \\ \lim\limits_{n\to\infty} |\widehat{\theta}_t^n - \mathbb{E}[\theta_t|y_{1:t}]| = 0 \quad p.s. \end{array}$$

Ces espérances sont relatives aux lois $p_t(x|y_{1:t})$ et $p_t(\theta|y_{1:t})$ respectivement.

Preuve : ce résultat s'obtient par majoration des premiers membres des égalités précédentes et établissement de la nullité asymptotique des termes majorants, notamment par une application du lemme de Borel-Cantelli.

3.3. *Convergence des estimations bayésiennes des paramètres*

S'il existe un estimateur convergent des vraies valeurs $\theta^* \in \Theta$ des paramètres inconnus, le théorème de Schervish ([14]) et le théorème de Schwartz ([15]) assurent respectivement que lorsque $t \to \infty$

a) $\forall A \in \Theta \ \lim_{t\to\infty} P(A|y_{1:t}) = \mathrm{I}_A(\theta^*) \quad p.s.$ avec I_A fonction indicatrice de A.

b) $\lim_{t\to\infty} \mathbb{E}[\theta|y_{1:t}] = \theta^* \quad p.s.$

Le théorème 3.1 associé à a) garantit donc la concentration progressive de la loi conditionnelle du vecteur θ_t autour des vraies valeurs θ^* des paramètres lorsque n et t tendent vers l'infini.

De même, le théorème 3.3 associé à b) garantit la convergence presque sûre de θ_t vers θ^* dans les mêmes conditions.

Références

[1] D. Bosq, J.P. Lecoutre, Théorie de l'estimation fonctionnelle. Economica, Paris (1987).

[2] P. Del Moral, Feynman-Kac Formulae. Genealogical and Interacting Particle Systems with Applications. Springer-Verlag New York (2004).

[3] P. Del Moral, J. Jacod, Ph. Protter, The Monte Carlo Method for filtering with discrete-time observations. Prob. Theory Relat. Fields, vol. 120, p346-368 (2001).

[4] P. Del Moral, L. Miclo, Branching and Interacting Particle Systems Approximations of Feynman-Kac Formulae with Applications to Nonlinear Filtering. Séminaire de Probabilité XXXIV, Ed. J. Azéma, M. Emery, M. Ledoux, M. Yor, Lecture Notes in Mathematics, Springer (2000).

[5] L. Devroye, A Course in Density Estimation. Birkhäuser, Boston (1987).

[6] A. Doucet, N. de Freitas, N. Gordon, Sequential Monte Carlo Methods in Practice. Statistics for Engeering and Information Science, Springer (2001).

[7] W.R. Gilks, C. Berzuini, Following a moving target-Monte Carlo inference for dynamic Baysesian models. J. R. Statist. Soc. B, 63(1), 127-146 (2001).

[8] T. Higuchi, Monte Carlo filter using the genetic algorithm operators. J. Stat. Comput. Simulation, 59(1)1, 1-23 (1997).

[9] L. Holmström, J. Klemelä, Asymptotic Bounds for the Expected L^1 Error of a multivariate Kernel Density Estimator. Journal of Multivariate Analysis, 42, p245-266 (1992).

[10] G. Kitagawa, A self-organizing state-space model". Journal of the American Statistical Association, 93(443), 1203-1215 (1998).

[11] F. LeGland, N. Oudjane, A robustification approach to stability and to uniform particle approximation of nonlinear filters : the example of pseudo-mixing signals. Stochastic Process. Appl. 106, no. 2, p279-316.

[12] J. Liu, M. West, Combined Parameter and State Estimation in Simulation-Based Filtering, in : A. Doucet, N. de Freitas, N. Gordon (Eds), Sequential Monte Carlo Methods Practice, Springer-Verlag, New York, 197-223 (2001).

[13] V. Rossi, J-P. Vila, Nonlinear filtering in discrete time : A particle convolution approach. Rapport technique 04-03, groupe biostatisque ENSAM/INRA/UM2 Montpellier, octobre (2004) sousmis à Stat. Inf. Stoc. Proc.

[14] M.J. Schervish, Theory of statistics, Springer-Verlag, New York (1995).

[15] L. Schwartz, On Bayes Procedures, Z. Wahrscheinlichkeitstheorie 4, 10-26 (1965).

[16] M. West, Approximating posterior distribution by mixture, Journal of the Royal Statistical Society (ser :B), vol. 55, 409-422 (1993).