Lissage rapide dans des modèles non linéaires et non gaussiens

Ivan Gorynin[1], Stéphane Derrode[2], Emmanuel Monfrini[1], Wojciech Pieczynski[1]

[1]Département CITI, Telecom SudParis, Institut Mines-Telecom, CNRS UMR 5157, 91011 Evry, France

[2]École Centrale de Lyon, LIRIS, CNRS UMR 5205, Écully, France.

Ivan.Gorynin@telecom-sudparis.eu, stephane.derrode@ec-lyon.fr,
Emmanuel.Monfrini@telecom-sudparis.eu, Wojciech.Pieczynski@telecom-sudparis.eu

Résumé − Nous nous intéressons à la question du filtrage et du lissage statistique dans les systèmes non linéaires et non gaussiens. La nouveauté réside dans l'approximation du système non-linéaire par un modèle à sauts dans lequel un calcul rapide et optimal du filtrage et du lissage est possible. Ces méthodes montrent leur intérêt, en particulier dans les récents modèles de volatilité stochastique asymétrique.

Abstract − We consider here the problem of statistical filtering and smoothing in nonlinear non-Gaussian systems. The novelty consists in approximating the nonlinear system by a recent switching system, in which exact fast optimal filtering and smoothing are workable. Our methods are applied to an asymetric stochastic volatility model and some experiments show their efficiency.

1 Introduction

Soient $\mathbf{X}_1^N = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$ et $\mathbf{Y}_1^N = (\mathbf{Y}_1, \ldots, \mathbf{Y}_N)$, deux suites de variables aléatoires. $\mathbf{X}_1^N$ est un processus caché et $\mathbf{Y}_1^N$ est l'observation que l'on en fait. Pour tout $n \in \{1, ..., N\}$, $\mathbf{X}_n$ et $\mathbf{Y}_n$ prennent respectivement leurs valeurs dans $\mathbb{R}^d$ et $\mathbb{R}^p$. Nous nous concentrons ici sur le calcul, pour tout $n \in \{1, ..., N\}$, de $\mathrm{E}\left[\mathbf{X}_n \left| \mathbf{y}_1^n \right.\right]$ et $\mathrm{E}\left[\mathbf{X}_n \mathbf{X}_n^\top \left| \mathbf{y}_1^n \right.\right]$ pour le filtrage ainsi que de $\mathrm{E}\left[\mathbf{X}_n \left| \mathbf{y}_1^N \right.\right]$ et $\mathrm{E}\left[\mathbf{X}_n \mathbf{X}_n^\top \left| \mathbf{y}_1^N \right.\right]$ pour le lissage, qui sont souvent les quantités d'intérêt dans ce genre de problématique. La distribution $p\left(\mathbf{x}_1^N, \mathbf{y}_1^N\right)$ du couple $(\mathbf{X}_1^N, \mathbf{Y}_1^N)$ est classiquement celle d'une chaîne de Markov cachée qui peut n'être ni linéaire, ni gaussienne. Dans ce cas, cette distribution est définie par $p\left(\mathbf{x}_1, \mathbf{y}_1\right)$ et les deux équations :

$$\mathbf{X}_{n+1} = \Phi(\mathbf{X}_n, \mathbf{U}_{n+1}); \quad (1)$$

$$\mathbf{Y}_n = \Psi(\mathbf{X}_n, \mathbf{V}_n), \quad (2)$$

où $\mathbf{U}_1, \mathbf{V}_1, \ldots, \mathbf{U}_N, \mathbf{V}_N$ sont des variables indépendantes. Dans le cas général, le filtrage et le lissage rapides et exacts ne sont pas possibles et de nombreuses méthodes d'approximation sont proposées. Parmi elles, les méthodes de Monte-Carlo séquentielles basées sur le filtrage et le lissage particulaire sont très efficaces mais rencontrent parfois des difficultés [2, 5].

Nous nous concentrons ici sur le cas où les distributions $p\left(\mathbf{x}_n^{n+1}\right)$, $p\left(\mathbf{y}_n \left| \mathbf{x}_n \right.\right)$ définies par (1)-(2) ne dépendent pas de n. La loi $p\left(\mathbf{x}_1^N, \mathbf{y}_1^N\right)$ du couple $\left(\mathbf{X}_1^N, \mathbf{Y}_1^N\right)$ est donc entièrement déterminée par la connaissance de $p\left(\mathbf{x}_1^2, \mathbf{y}_1^2\right)$.

Nous nous proposons d'approcher cette distribution par celle d'un mélange de K^2 gaussiennes

$$p\left(\mathbf{x}_1^2, \mathbf{y}_1^2\right) \approx \sum_{1 \le i,j, \le K} \alpha_{ij}\, p_{ij}(\mathbf{x}_1^2, \mathbf{y}_1^2), \quad (3)$$

ce que l'on sait pouvoir faire, sous des hypothèses peu coûteuses, avec une précision arbitrairement grande en augmentant le nombre K^2 de composantes. On peut alors considérer α_{ij} comme la loi d'un couple (R_1, R_2) de variables aléatoires discrètes $\alpha_{i,j} = P(R_1 = i, R_2 = j)$ où R_1 et R_2 sont à valeurs dans l'ensemble $\Omega = \{1, \ldots, K\}$. Ainsi, $p\left(\mathbf{x}_1^2, \mathbf{y}_1^2\right) = \sum_{1 \le r_1, r_2, \le K} p\left(\mathbf{r}_1^2\right) p\left(\mathbf{x}_1^2, \mathbf{y}_1^2 \left| \mathbf{r}_1^2 \right.\right)$ peut être identifié à (3) puisque $p\left(\mathbf{x}_1^2, \mathbf{y}_1^2\right)$ est une marginale de

$$p\left(\mathbf{x}_1^2, \mathbf{r}_1^2, \mathbf{y}_1^2\right) = p\left(\mathbf{r}_1^2\right) p\left(\mathbf{x}_1^2, \mathbf{y}_1^2 \left| \mathbf{r}_1^2 \right.\right), \quad (4)$$

Soit alors le processus discret $\mathbf{R}_1^N = (R_1, \ldots, R_N)$ à valeurs dans l'ensemble des classes $\Omega = \{1, \ldots, K\}$. En supposant que $p\left(r_1\right) = p\left(r_2\right)$, la loi (4) définit entièrement la distribution stationnaire du triplet $\mathbf{T}_1^N = (\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$. Comme nous l'exposons dans les parties 2 et 3, le filtrage (cf. [4]) et le lissage rapides et exacts sont possibles dans certains modèles de ce type. Dans la partie 4, nous appliquons la méthode proposée à des modèles de volatilité stochastique.

2 Modèles cachés conditionnellement linéaires à sauts markoviens

Considérons $\mathbf{X}_1^N$ et $\mathbf{Y}_1^N$ les processus réels définis précédemment ainsi que le processus $\mathbf{R}_1^N$ pour lequel R_n est

à valeurs dans un ensemble de classes $\Omega = \{1, \ldots, K\}$.

Définition 1

Le triplet $\mathbf{T}_1^N = (\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$ est dit "modèle caché conditionnellement linéaire à sauts markoviens" (MCCLSM) si $\mathbf{T}_1^N$ est de Markov avec les propriétés suivantes :

$$p\left(r_{n+1}, \mathbf{y}_{n+1} \,|\mathbf{x}_n, r_n, \mathbf{y}_n\right) = p\left(r_{n+1}, \mathbf{y}_{n+1} \,|r_n, \mathbf{y}_n\right); \quad (5)$$

$$\begin{aligned}\mathbf{X}_{n+1} = \mathbf{F}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1})\mathbf{X}_n + \\ \mathbf{G}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1})\mathbf{W}_{n+1} + \mathbf{H}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1}),\end{aligned} \quad (6)$$

où $\mathbf{F}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})$ et $\mathbf{G}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})$ sont des matrices, $(\mathbf{W}_{n+1})$ est un processus de bruit blanc gaussien et $\mathbf{H}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})$ sont des vecteurs appropriés. On peut alors établir le résultat suivant (cf. [14])

Proposition 1 (Filtrage)

Considérons le MCCLSM $\mathbf{T}_1^N = (\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$. Alors, pour tout $n \in \{1, \ldots, N\}$,

$$\mathrm{E}\left[\mathbf{X}_n \,|\mathbf{y}_1^n\right] = \sum_{r_n} p\left(r_n \,|\mathbf{y}_1^n\right) \mathrm{E}\left[\mathbf{X}_n \,|r_n, \mathbf{y}_1^n\right] \quad \text{et,} \quad (7)$$

$$\mathrm{E}\left[\mathbf{X}_n\mathbf{X}_n^\top \,|\mathbf{y}_1^n\right] = \sum_{r_n} p\left(r_n \,|\mathbf{y}_1^n\right) \mathrm{E}\left[\mathbf{X}_n\mathbf{X}_n^\top \,|r_n, \mathbf{y}_1^n\right]. \quad (8)$$

peuvent être calculées récursivement avec une complexité linéaire en n par :

$$\begin{aligned}&\mathrm{E}\left[\mathbf{X}_{n+1} \,|r_{n+1}, \mathbf{y}_1^{n+1}\right] = \sum_{r_n} p\left(r_n \,|r_{n+1}, \mathbf{y}_1^{n+1}\right) \\ &\times \left(\mathbf{F}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})\mathrm{E}\left[\mathbf{X}_n \,|r_n, \mathbf{y}_1^n\right] + \mathbf{H}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})\right)\end{aligned} \quad (9)$$

$$\begin{aligned}&\text{et } \mathrm{E}\left[\mathbf{X}_{n+1}\mathbf{X}_{n+1}^\top \,|\mathbf{r}_n^{n+1}, \mathbf{y}_1^{n+1}\right] = \sum_{r_n} p\left(r_n \,|r_{n+1}, \mathbf{y}_1^{n+1}\right) \times \\ &\Big(\mathbf{F}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})\mathrm{E}\left[\mathbf{X}_n\mathbf{X}_n^\top \,|\mathbf{r}_n^{n+1}, \mathbf{y}_1^n\right] \mathbf{F}_{n+1}^\top(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}) \\ &+ \mathbf{F}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})\mathrm{E}\left[\mathbf{X}_n \,|\mathbf{r}_n^{n+1}, \mathbf{y}_1^n\right] \mathbf{H}_{n+1}^\top(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}) \\ &+ \mathbf{H}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})\mathrm{E}\left[\mathbf{X}_n^\top \,|\mathbf{r}_n^{n+1}, \mathbf{y}_1^n\right] \mathbf{F}_{n+1}^\top(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}) \\ &+ \mathbf{G}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})\mathbf{G}_{n+1}^\top(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}) \\ &+ \mathbf{H}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})\mathbf{H}_{n+1}^\top(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})\Big),\end{aligned} \quad (10)$$

En effet, d'après (5), on a $p\left(\mathbf{x}_n \,|r_n, \mathbf{y}_n^{n+1}\right) = p\left(\mathbf{x}_n \,|r_n, \mathbf{y}_n\right)$. Plus généralement, on a $p\left(\mathbf{x}_n \,|r_n, \mathbf{y}_n^N\right) = p\left(\mathbf{x}_n \,|r_n, \mathbf{y}_n\right)$ et $p\left(\mathbf{x}_n \,|\mathbf{y}_1^N\right) = \sum_{r_n} p\left(r_n \,|\mathbf{y}_1^N\right) p\left(\mathbf{x}_n \,|r_n, \mathbf{y}_1^N\right)$ ce qui permet d'établir (11) et (12).

Proposition 2 (Lissage)

Considérons le MCCLSM $\mathbf{T}_1^N = (\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$. Alors, pour tout $n \in \{1, \ldots, N\}$,

$$\mathrm{E}\left[\mathbf{X}_n \,|\mathbf{y}_1^N\right] = \sum_{r_n} p\left(r_n \,|\mathbf{y}_1^N\right) \mathrm{E}\left[\mathbf{X}_n \,|r_n, \mathbf{y}_1^n\right] \quad \text{et,} \quad (11)$$

$$\mathrm{E}\left[\mathbf{X}_n\mathbf{X}_n^\top \,|\mathbf{y}_1^N\right] = \sum_{r_n} p\left(r_n \,|\mathbf{y}_1^N\right) \mathrm{E}\left[\mathbf{X}_n\mathbf{X}_n^\top \,|r_n, \mathbf{y}_1^n\right]. \quad (12)$$

peuvent être calculées avec une complexité linéaire en N.

Remarque 1

Les probabilités $p\left(r_n \,|\mathbf{y}_1^n\right)$ et $p\left(r_n \,|\mathbf{y}_1^N\right)$ sont calculées récursivement avec une complexité linéaire grâce aux probabilités *forward* et *backward* dans la chaîne de Markov couple $(\mathbf{R}_1^N, \mathbf{Y}_1^N)$, la markovianité de cette chaîne étant impliquée par (5). En posant $\boldsymbol{v}_n = (r_n, \mathbf{y}_n)$, les probabilités $\alpha_n(r_n) = p\left(r_n, \mathbf{y}_1^n\right)$ et $\beta_n(r_n) = p\left(\mathbf{y}_{n+1}^N \,|\boldsymbol{v}_n\right)$ sont calculables récursivement par :

$$\begin{aligned}&\alpha_1(r_1) = p\left(\boldsymbol{v}_1\right); \\ &\alpha_{n+1}(r_{n+1}) = \sum_{r_n \in \Omega} \alpha_n(r_n) p\left(\boldsymbol{v}_{n+1} \,|\boldsymbol{v}_n\right); \quad (13) \\ &\beta_N(r_N) = 1; \\ &\beta_n(r_n) = \sum_{r_{n+1} \in \Omega} \beta_{n+1}(r_{n+1}) p\left(\boldsymbol{v}_{n+1} \,|\boldsymbol{v}_n\right). \quad (14)\end{aligned}$$

On a alors $p\left(r_n \,|\mathbf{y}_1^n\right) = \alpha_n(r_n) / \sum_{r_n^* \in \Omega} \alpha_n(r_n^*)$ et

$$p\left(r_n \,|\mathbf{y}_1^N\right) = \frac{\alpha_n(r_n)\beta_n(r_n)}{\sum_{r_n^* \in \Omega} \alpha_n(r_n^*)\beta_n(r_n^*)}. \quad (15)$$

3 MCCLSM gaussiens

Notons $\mathbf{Z}_n = (\mathbf{X}_n^\top, \mathbf{Y}_n^\top)^\top$, z_n sa réalisation, et considérons le cas stationnaire où la distribution de $\mathbf{T}_1^N = (\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$ est entièrement définie par la connaissance de $p\left(\mathbf{r}_1^2\right)$ et des distributions gaussiennes $p_{\mathbf{r}_1^2}(\mathbf{x}_1^2, \mathbf{y}_1^2)$, de moyennes

$$\boldsymbol{\Upsilon}(\mathbf{r}_1^2) = \begin{bmatrix} \mathbf{M}(r_1) \\ \mathbf{M}(r_2) \end{bmatrix} = \begin{bmatrix} \mathrm{E}\left[\mathbf{Z}_1 \,|r_1\right] \\ \mathrm{E}\left[\mathbf{Z}_2 \,|r_2\right] \end{bmatrix} \quad (16)$$

et de matrices de covariance

$$\boldsymbol{\Xi}(\mathbf{r}_1^2) = \begin{bmatrix} \boldsymbol{S}(r_1) & \boldsymbol{\Sigma}(\mathbf{r}_1^2) \\ \boldsymbol{\Sigma}^\top(\mathbf{r}_1^2) & \boldsymbol{S}(r_2) \end{bmatrix} \quad (17)$$

Si on pose

$$\mathbf{A}(\mathbf{r}_1^2) = \boldsymbol{\Sigma}^\top(\mathbf{r}_1^2)\, \boldsymbol{S}^{-1}(r_1), \quad (18)$$

et

$$\mathbf{B}(\mathbf{r}_1^2)\mathbf{B}^\top(\mathbf{r}_1^2) = \boldsymbol{S}(r_2) - \boldsymbol{\Sigma}^\top(\mathbf{r}_1^2)\boldsymbol{S}^{-1}(r_1)\boldsymbol{\Sigma}(\mathbf{r}_1^2), \quad (19)$$

le triplet $\mathbf{T}_1^N$ est défini par :

$$\mathbf{R}_1^N \text{ est une chaîne de Markov;} \quad (20)$$

$$\mathbf{Z}_{n+1} = \mathbf{A}(\mathbf{R}_n^{n+1})(\mathbf{Z}_n - \mathbf{M}(R_n)) + \mathbf{B}(\mathbf{R}_n^{n+1})\mathbf{W}_{n+1} + \mathbf{M}(R_{n+1}), \quad (21)$$

où, comme précédemment, $(\mathbf{W}_{n+1})$ est un processus de bruit blanc gaussien. Nous savons déjà que ces modèles peuvent être très proches des modèles à sauts linéaires et gaussiens qui n'offrent pas la possibilité d'un filtrage rapide [3, 13]. Considérons maintenant $\mathbf{A}(\mathbf{r}_1^2)$ de la forme

$$\mathbf{A}(\mathbf{r}_1^2) = \begin{bmatrix} \mathbf{A}_1(\mathbf{r}_1^2) & \mathbf{A}_2(\mathbf{r}_1^2) \\ \mathbf{0} & \mathbf{A}_4(\mathbf{r}_1^2) \end{bmatrix}. \quad (22)$$

Nous pouvons alors énoncer [1] :

Proposition 3

Le MCCLSM $\mathbf{T}_1^N = (\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$ défini par (20)-(22) est un MCCLSM caractérisé par $\mathbf{F}_{n+1}$, $\mathbf{G}_{n+1}$ et $\mathbf{H}_{n+1}$ indépendants de n définis par (23)-(25).

$$\mathbf{F}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}) = \mathbf{A}_1(\mathbf{r}_n^{n+1}), \tag{23}$$

$$\mathbf{H}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}) = \mathbf{A}_2(\mathbf{r}_n^{n+1})\mathbf{y}_n + \mathbf{N}_1(\mathbf{r}_n^{n+1}) + \boldsymbol{Q}_2(\mathbf{r}_n^{n+1})\boldsymbol{Q}_4^{-1}(\mathbf{r}_n^{n+1})(\mathbf{y}_{n+1} - \mathbf{A}_4(\mathbf{r}_n^{n+1})\mathbf{y}_n - \mathbf{N}_2(\mathbf{r}_n^{n+1})), \tag{24}$$

$$\mathbf{G}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})\mathbf{G}_{n+1}^T(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}) = \boldsymbol{Q}_1(\mathbf{r}_n^{n+1}) - \boldsymbol{Q}_2(\mathbf{r}_n^{n+1})\boldsymbol{Q}_4^{-1}(\mathbf{r}_n^{n+1})\boldsymbol{Q}_3(\mathbf{r}_n^{n+1}). \tag{25}$$

avec

$$\mathbf{Q}(\mathbf{r}_1^2) = \begin{bmatrix} \boldsymbol{Q}_1(\mathbf{r}_1^2) & \boldsymbol{Q}_2(\mathbf{r}_1^2) \\ \boldsymbol{Q}_3(\mathbf{r}_1^2) & \boldsymbol{Q}_4(\mathbf{r}_1^2) \end{bmatrix} = \mathbf{B}(\mathbf{r}_1^2)\mathbf{B}^\top(\mathbf{r}_1^2),$$

$$\mathbf{N}_1(\mathbf{r}_n^{n+1}) = \mathbf{M}_1(r_{n+1}) - \mathbf{A}_1(\mathbf{r}_n^{n+1})\mathbf{M}_1(r_n) - \mathbf{A}_2(\mathbf{r}_n^{n+1})\mathbf{M}_2(r_n),$$

$$\mathbf{N}_2(\mathbf{r}_n^{n+1}) = \mathbf{M}_2(r_{n+1}) - \mathbf{A}_4(\mathbf{r}_n^{n+1})\mathbf{M}_2(r_n).$$

où $\mathbf{M}_1(r_n) = \mathrm{E}[\mathbf{X}_n | r_n]$ et $\mathbf{M}_2(r_n) = \mathrm{E}[\mathbf{Y}_n | r_n]$.

Si l'on revient alors au cas général du modèle de Markov couple, non linéaire et non gaussien, il nous suffit de pouvoir simuler des réalisations $(\mathbf{x}_1'^M, \mathbf{y}_1'^M)$ du processus défini par $p(\mathbf{x}_1, \mathbf{y}_1)$, (1) et (2) pour pouvoir estimer les paramètres du MCCLSM qui sont, pour $1 \leq i, j \leq K$, $p_{ij} = p(r_1 = i, r_2 = j)$, $\boldsymbol{\Upsilon}_{i,j} = \boldsymbol{\Upsilon}(i,j)$, et $\boldsymbol{\Xi}_{ij} = \boldsymbol{\Xi}(i,j)$. Après une initialisation obtenue classiquement grâce aux estimateurs empiriques utilisés sur une segmentation K-means, ces paramètres sont ré-estimés dans une procédure itérative de type EM grâce aux estimateurs empiriques pondérés (à l'itération q) :

$$p_{ij}^{(q+1)} = \frac{1}{M-1}\sum_{n=1}^{M-1}\psi_n^{(q)}(i,j), \quad \boldsymbol{\Upsilon}_{i,j}^{(q+1)} = \frac{\sum_{n=1}^{M-1}\mathbf{z}_n'^{n+1}\psi_n^{(q)}(i,j)}{\sum_{n=1}^{M-1}\psi_n^{(q)}(i,j)},$$

$$\boldsymbol{\Xi}_{ij}{}^{(q+1)} = \frac{\sum_{n=1}^{M-1}\left(\mathbf{z}_n'^{n+1} - \boldsymbol{\Upsilon}_{i,j}^{(q+1)}\right)\left(\mathbf{z}_n'^{n+1} - \boldsymbol{\Upsilon}_{i,j}^{(q+1)}\right)^\top\psi_n^{(q)}(i,j)}{\sum_{n=1}^{M-1}\psi_n^{(q)}(i,j)},$$

avec $\psi_n^{(q)}(i,j) = p\left(r_n = i, r_{n+1} = j \left| \mathbf{x}_1'^M, \mathbf{y}_1'^M\right.\right)$ utilisant les paramètres courants $\left(p_{ij}^{(q)}, \boldsymbol{\Upsilon}_{i,j}^{(q)}, \boldsymbol{\Xi}_{ij}{}^{(q)}\right)$ et obtenues par une procédure forward-backward dans le triplet $\mathbf{T}_1'^M$ où $\alpha_n(r_n) = p(r_n, \mathbf{z}_1'^n)$ et $\beta_n(r_n) = p\left(\mathbf{z}_{n+1}'^M | \mathbf{t}_n'\right)$ sont donc calculées récursivement pour $1 \leq n < M$, par

$$\psi_n^{(q)}(\mathbf{r}_n^{n+1}) = \frac{\alpha_n(r_n)p\left(\mathbf{t}_{n+1}' | \mathbf{t}_n'\right)\beta_{n+1}(r_{n+1})}{\sum_{r_n^*, r_{n+1}^*}\alpha_n(r_n^*)p\left(\mathbf{t}_{n+1}^* | \mathbf{t}_n^*\right)\beta_{n+1}(r_{n+1}^*)}, \tag{26}$$

avec $\boldsymbol{t}_i^* = (\mathbf{x}_i', r_i^*, \mathbf{y}_i')$, $(\mathbf{x}_1'^M, \mathbf{y}_1'^M)$ étant fixé.

4 Applications à des modèles de volatilité stochastique

Afin de mettre en évidence l'intérêt des méthodes de filtrage et de lissage présentées ici, nous proposons deux séries d'expériences. Pour la première, nous nous intéressons au cas du modèle classique de volatilité stochastique (MVS) [8, 9] défini par

$$\mathbf{X}_1 = \mu + \mathbf{U}_1; \tag{27}$$

$$\mathbf{X}_{n+1} = \mu + \phi(\mathbf{X}_n - \mu) + \sigma\mathbf{U}_{n+1}; \tag{28}$$

$$\mathbf{Y}_n = \beta\exp(\mathbf{X}_n/2)\mathbf{V}_n, \tag{29}$$

où μ est un paramètre fixé et $\mathbf{U}_1, \mathbf{V}_1, \ldots, \mathbf{U}_N, \mathbf{V}_N$ sont des gaussiennes centrées, réduites et indépendantes. Ce modèle est bien un modèle de Markov caché défini par les équations (1)-(2).

TABLE 1 – EQM pour le lissage dans le MVS lorsque $\mu = 0.5$, $\beta = 0.5$, et 4 valeurs de σ^2 et ϕ.

			K				
	ϕ	σ^2	2	3	5	7	LP
1	0.99	0.02	0,39	0,23	0,14	0,12	0.12
2	0.90	0.19	0,48	0,39	0,35	0,34	0.33
3	0.80	0.36	0,56	0,50	0,47	0,47	0.46
4	0.50	0.75	0,70	0,67	0,66	0,66	0.66

TABLE 2 – EQM pour le lissage dans le MVSA lorsque $\mu = 0.5$, $\beta = 0.5$, $\phi = 0.5$ et 5 valeurs de λ^2 et ρ.

			K				
	ρ	λ^2	2	3	5	7	LP
1	-0.90	0.19	0,23	0,21	0,20	0,20	0.19
2	-0.80	0.36	0,36	0,34	0,32	0,32	0.32
3	-0.50	0.75	0,57	0,55	0,55	0,55	0.54
4	-0.30	0.91	0,65	0,63	0,62	0,62	0.62
5	-0.00	1.00	0,70	0,67	0,66	0,66	0.66

Dans la seconde série d'expériences, nous nous intéressons à une extension récente du MVS introduite dans [11, 12, 15] sous le nom de modèle de volatilité stochastique asymétrique (MVSA). Ce modèle est tel que :

$$\mathbf{X}_1 = \mu + \mathbf{U}_1; \tag{30}$$

$$\mathbf{X}_{n+1} = \mu + \phi(\mathbf{X}_n - \mu) + \sigma\left(\rho\mathbf{V}_n + \lambda\mathbf{U}_{n+1}\right); \tag{31}$$

$$\mathbf{Y}_n = \beta\exp(\mathbf{X}_n/2)\mathbf{V}_n. \tag{32}$$

et, comme dans le MVS, $\mathbf{U}_1, \mathbf{V}_1, \ldots, \mathbf{U}_N, \mathbf{V}_N$ sont des gaussiennes centrées, réduites et indépendantes. Comme le montre (31), le MVSA n'est pas nécessairement un modèle de Markov caché " à bruit indépendant" classique puisque $\mathbf{X}_{n+1}$ est fonction de $\mathbf{X}_n$ et $\mathbf{Y}_n$. Cependant, $\left(\mathbf{X}_1^N, \mathbf{Y}_1^N\right)$ reste markovien ce qui fait du MVSA un modèle de Markov caché "à bruit correlé". Comme il est aisément possible de simuler des réalisations du MVSA, les méthodes de filtrage et de lissage que nous proposons restent valables.

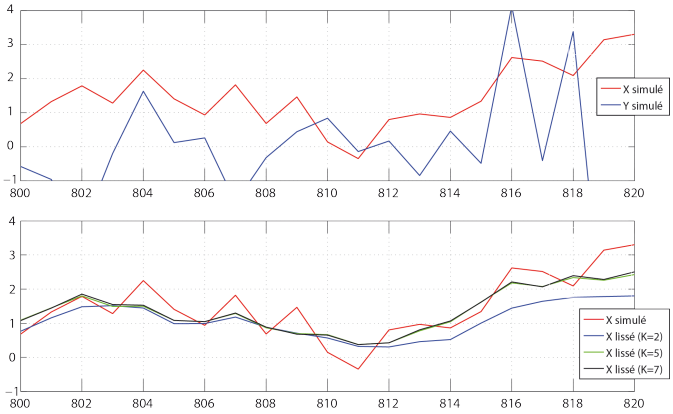

FIGURE 1 – Trajectoire simulée pour un MVS ($\phi = 0.5$) filtrée et lissée pour $K = 2$, 5 et 7.

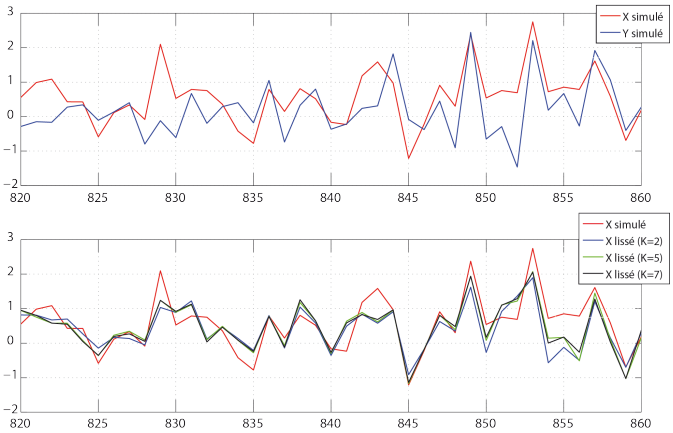

FIGURE 2 – Trajectoire simulée pour un MVSA ($\phi = 0.5$ et $\rho = -0.8$) filtrée et lissée pour $K = 2$, 5 et 7.

La méthode proposée est testée pour $K = 2$, 3, 5, et 7 et comparées au lissage particulaire (LP) classique (1500 particules) (cf. [6, 7, 10]). On se restreint au cas où $d = p = 1$ et à une taille d'échantillon $N = 1000$. L'échantillon d'apprentissage est de taille $M = 20000$. Le critère de comparaison utilisé est l'erreur quadratique moyenne (EQM) pour 100 expériences indépendantes et les résultats sont présentés dans les tableaux 1 et 2. Les Figures 1 et 2 présentent les résultats obtenus pour des MVS et MVSA respectivement.

Nous montrons ainsi que pour $K = 7$ les performances sont quasi optimales et peu éloignées de celles observées pour $K = 5$. L'apprentissage des paramètres du mélange de gaussiennes est gourmand en temps de calcul. En revanche, lorsque l'apprentissage n'est pas à faire, le coût de calcul est équivalent à celui d'un lissage particulaire basé sur K^2 particules.

Références

[1] N. Abbassi, D. Benboudjema, and W. Pieczynski. Kalman filtering approximations in triplet Markov Gaussian switching models. In *Proc. IEEE Int. Workshop SSP'11*, pages 290–294, Nice, France, June 2011.

[2] O. Cappé, E. Moulines, and T. Rydén. *Inference in Hidden Markov Models*. Springer-Verlag, 2005.

[3] S. Derrode and W. Pieczynski. Exact fast computation of optimal filter in Gaussian switching linear systems. *IEEE Signal Processing Letters*, 20(7) :701–704, July 2013.

[4] S. Derrode and W. Pieczynski. Fast filter in nonlinear systems with application to stochastic volatility model. In *Proc. of the EUSIPCO'14*, pages 290–294, Lisbon, Portugal, September 2014.

[5] A. Doucet and A. Johansen. *A tutorial on particle filtering and smoothing : Fifteen years later*. Eds. London, U.K., Oxford Univ. Press, 2011.

[6] D. Duffie, J. Pan, and K. J. Singleton. Transform analysis and asset pricing for affine jump diffusions. *Econometrica*, 68(6) :1343–1376, 2000.

[7] B. Eraker. Do stock prices and volatility jump ? Reconciling evidence from spot and option prices. *The Journal of Finance*, 59 :1367–1404, 2004.

[8] E. Ghysels, A. Harvey, and E. Renault. Stochastic volatility. *Handbook of Statistics*, 14 :119–192, 1995.

[9] E. Jacquier, N. G. Polson, and P. Rossi. Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics*, 12(4) :371–389, 1994.

[10] C.-J. Kim and C. R. Nelson. *State-space models with regime switching*. MIT Press, 1999.

[11] N. Y. Nikolaev, E. Smirnov, and L. M. de Menezes. Nonlinear filtering of asymmetric stochastic volatility models and value-at-risk estimation. In *Proc. of the IEEE Conf. Computational Intelligence for Financial Engineering and Economics (CIFEr-2014)*, pages 310–317, London, U.K., 2014.

[12] Y. Omori and T. Watanabe. Block sampler and posterior mode estimation for asymmetric stochastic volatility models. *Computational Statistics and Data Analysis*, 52(6) :2892–2910, 2008.

[13] Y. Petetin and F. Desbouvries. A class of fast exact Bayesian filters in dynamical models with jumps. *IEEE Trans. on Signal Processing*, 62(14) :3643–3653, 2014.

[14] W. Pieczynski. Exact filtering in conditionally Markov switching hidden linear models. *Comptes Rendus Mathématique*, 349(9-10) :587–590, 2011.

[15] M. Takashi, Y. Omori, and T. Watanabe. Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics and Data Analysis*, 53(6) :2404–2426, 2009.