

F. CAMPILLO

**La méthode d'approximation de Gauss-Galerkin en filtrage non linéaire**

*M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique*, tome 20, n° 2 (1986), p. 203-223

[http://www.numdam.org/item?id=M2AN\\_1986\\_\\_20\\_2\\_203\\_0](http://www.numdam.org/item?id=M2AN_1986__20_2_203_0)

© AFCET, 1986, tous droits réservés.

L'accès aux archives de la revue « M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>



## LA MÉTHODE D'APPROXIMATION DE GAUSS-GALERKIN EN FILTRAGE NON LINÉAIRE (\*)

par F. CAMPILLO <sup>(1)</sup>

Communique par R. TEMAN

---

*Résumé — On étudie une méthode d'approximation de l'équation du filtrage non linéaire en dimension 1, avec observation en temps discret et en temps continu. Dans une première partie on présente la méthode appliquée à l'équation de Fokker-Planck. La convergence de l'approximation est démontrée. On présente enfin un exemple numérique.*

*Abstract — We study an approximation method for the one-dimensional nonlinear filtering problem, with discrete time and continuous time observation. We first present the method applied to the Fokker-Planck equation. The convergence of the approximation is established. We finally present a numerical example.*

### 1. INTRODUCTION

Les méthodes habituelles de résolution numérique des équations aux dérivées partielles font généralement appel à un nombre important de points de discrétisation en espace. De plus, sous leur forme classique, ces méthodes utilisent des grilles de discrétisation fixes dans le temps.

La méthode proposée par D. A. Dawson [4] pour la résolution numérique de l'équation de Fokker-Planck, appelée méthode de Gauss-Galerkin, combine les méthodes de quadrature de Gauss et d'approximation de Galerkin. Cette méthode, qui peut être considérée comme une méthode particulière (cf. Raviart [10]), présente le double avantage de donner des résultats acceptables même avec un petit nombre d'inconnues à calculer et une grille de discrétisation capable de s'adapter aux évolutions de la solution de l'équation aux dérivées partielles considérée. Cependant, dans sa forme actuelle, la méthode est limitée au cas d'une seule dimension d'espace.

---

(\*) Reçu en avril 1985

(1) U E R de Mathématiques, Université de Provence, 3, place V Hugo, 13331 Marseille Cedex 3

Nous allons étudier le comportement de cette méthode, appliquée au problème de filtrage non linéaire. Dans la Section 2, nous présentons la méthode d'approximation de Gauss-Galerkin appliquée à l'équation de Fokker-Planck, et nous établissons un résultat de convergence. Les résultats de cette section sont une reprise et un développement des travaux de D. A. Dawson [4].

Dans la Section 3, nous considérons dans un premier temps le problème de filtrage non linéaire avec observation en temps discret : Nous présentons l'approximation de Gauss-Galerkin et nous en démontrons la convergence. Nous considérons ensuite le problème de filtrage non linéaire avec observation en temps continu : Dans ce cas, avant d'introduire l'approximation, nous discrétisons l'équation d'observation pour nous ramener au cas précédent.

## 2. RÉOLUTION NUMÉRIQUE DE L'ÉQUATION DE FOKKER-PLANCK

### 2.1. La méthode d'approximation de Gauss-Galerkin

Introduisons tout d'abord l'équation de Fokker-Planck. Considérons l'équation différentielle stochastique :

$$(2.1) \quad dX_t = b(X_t) dt + \sigma(X_t) dW_t, \quad 0 \leq t \leq T, \quad X_0 \text{ de loi } \mu_0,$$

où  $\{X_t\}$  prend ses valeurs dans  $\mathbb{R}$ ;  $\{W_t\}$  est un processus de Wiener standard réel indépendant de  $X_0$ . Posons  $a = \sigma^2$ ,  $a' = da/dx$  et  $b' = db/dx$  on fait les hypothèses suivantes :

(H1)  $b, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ , sont des applications mesurables et bornées.

(H2)  $a' \in L^\infty(\mathbb{R})$  et il existe  $\alpha > 0$  tel que  $a(x) \geq \alpha \forall x$ .

(H3)  $b'$  est mesurable bornée, et  $a'$  est continue.

Sous les hypothèses (H1), (H2), l'équation (2.1) admet une solution unique au sens faible (cf. Stroock-Varadhan [12]). L'hypothèse (H3) sera utilisée par la suite pour démontrer la convergence de l'approximation.

Soit  $\mu_t$  la loi de  $X_t$  sur  $\mathbb{R}$  ( $\forall t$ ). Il résulte de la formule de Itô que  $\{\mu_t\}$  est solution de l'équation de Fokker-Planck (écrite sous forme faible) :

$$(2.2) \quad \langle \mu_t, \phi \rangle = \langle \mu_0, \phi \rangle + \int_0^t \langle \mu_s, L\phi \rangle ds, \quad 0 \leq t \leq T, \quad \forall \phi \in C_b^2(\mathbb{R})$$

$L$  désigne l'opérateur  $L\phi(x) := b(x)\phi'(x) + 1/2 a(x)\phi''(x)$ , et on note  $\langle \mu, \phi \rangle := \int \phi(x) \mu(dx)$ .

$N \in \mathbb{N}$  étant donné, la méthode d'approximation de Gauss-Galerkin

consiste à approcher  $\{\mu_t\}$  par une famille de mesures de probabilité  $\{\mu_t^N\}$  de la forme :

$$\mu_t^N = \sum_{i=1}^N a_t^i \delta(x_t^i) \quad (\delta(x) \text{ mesure de Dirac au point } x).$$

Les fonctions  $t \rightarrow a_t^i, x_t^i$  sont déterminées en posant :

$$(2.3) \quad \langle \mu_t^N, \phi \rangle = \langle \mu_0, \phi \rangle + \int_0^t \langle \mu_s^N, L\phi \rangle ds, \quad 0 \leq t \leq T, \quad \forall \phi \in \Pi_{2N-1}$$

( $\Pi_{2N-1}$  désigne l'espace des polynômes de degré au plus  $2N-1$ ). On remarque que  $\mu_0^N$  vérifie  $\langle \mu_0^N, \phi \rangle = \langle \mu_0, \phi \rangle \forall \phi \in \Pi_{2N-1}$ , ainsi  $\mu_0^N$  est l'approximation de Gauss-Christoffel à  $N$  points de la mesure  $\mu_0$ .

### 2.2. Convergence de l'approximation

Sous une hypothèse supplémentaire, nous allons établir un résultat de convergence de l'approximation (pour une démonstration plus détaillée cf. Campillo [3]).

On pose :

$$\begin{aligned} m_n(t) &= \langle \mu_t, x^n \rangle, & \dot{m}_n(t) &= dm_n(t)/dt, \\ m_n^N(t) &= \langle \mu_t^N, x^n \rangle, & \dot{m}_n^N(t) &= dm_n^N(t)/dt. \end{aligned}$$

( $x^n$  désigne le polynôme  $x \rightarrow x^n$ ). On fait l'hypothèse supplémentaire :

$$(H4) \quad \overline{\lim}_{n \rightarrow \infty} \{ m_{2n}(0)/(2n)! \}^{1/2n} < \infty.$$

Cette hypothèse assure en particulier l'existence de moments de tous ordres pour  $X_0$ , donc pour  $X_t (\forall t)$ . De plus, en utilisant le critère de Cauchy sur la convergence des séries, (H4) implique que la série  $\sum \frac{\theta^n}{n!} m_n(0)$  admet un rayon de convergence strictement positif. Ce résultat permet d'affirmer que  $\mu_0$  est l'unique mesure non négative sur  $\mathbb{R}$  admettant  $\{m_n(0); n \in \mathbb{N}\}$  comme moments (cf. Billingsley [1]).

**THÉORÈME 2.1 :** *Sous les hypothèses (H1)-(H4), l'approximation de Gauss-Galerkin est convergente :  $\mu_t^N \xrightarrow[N \rightarrow \infty]{} \mu_t, t \geq 0$ .  $\square$*

Pour démontrer ce théorème, on utilise plusieurs lemmes. Dans la suite nous allons raisonner pour  $t \in [0, T]$ ; tous les résultats seront vrais pour tout  $T \geq 0$ .

LEMME 2.2 : Il existe des réels  $K_n, K'_n$  ne dépendant pas de  $N$  tels que :

$$(i) \quad |m_n^N(t)| \leq K_n \forall n, N, 2N \geq n+1, 0 \leq t \leq T;$$

$$(ii) \quad |\dot{m}_n^N(t)| \leq K'_n \forall n, N, 2N \geq n+1, 0 \leq t \leq T;$$

$$(iii) \quad \text{La série } \sum_{n \in \mathbb{N}} \frac{\theta^n}{n!} K_n \text{ admet un rayon de convergence } > 0. \quad \square$$

*Preuve* : Nous montrons qu'il existe  $K_n$  et  $K'_n$  tels que :

$$(2.4) \quad |m_n(t)| \leq K_n, \forall n, 0 \leq t \leq T.$$

$$(2.5) \quad |\dot{m}_n(t)| \leq K'_n, \forall n, 0 \leq t \leq T.$$

$$(2.6) \quad \text{La série } \sum_{n \in \mathbb{N}} \frac{\theta^n}{n!} K_n \text{ admet un rayon de convergence } > 0.$$

Supposons que  $|m_{2n-2}(t)| \leq K_{2n-2}$  ( $0 \leq t \leq T$ ), en prenant  $\phi(x) = x^{2n}$  dans (2.2), on obtient après majoration :

$$m_{2n}(t) \leq m_{2n}(0) + c \int_0^t \{ m_{2n}(s) + n^2 K_{2n-2} \} ds.$$

Puis en utilisant le lemme de Gronwall :

$$m_{2n}(t) \leq c \{ m_{2n}(0) + n^2 K_{2n-2} \}$$

où  $c$  désigne une constante qui dépend de  $T$ , mais pas de  $n$ . Soit  $K_0$  telle que  $|m_0(t)| \leq K_0, 0 \leq t \leq T$ , on définit par récurrence :

$$(2.7) \quad K_{2n} = c \{ m_{2n}(0) + n^2 K_{2n-2} \}$$

alors  $|m_{2n}(t)| \leq K_{2n}$  ( $\forall n, 0 \leq t \leq T$ ). Par ailleurs

$$|x|^{2n} \leq 1/2 \left\{ \frac{x^{2n}}{2n} + 2nx^{2n-2} \right\},$$

on peut donc choisir :  $K_{2n-1} := 1/2 \left\{ \frac{K_{2n}}{2n} + 2nK_{2n-2} \right\}$ . Ainsi (2.4) est démontré. En écrivant explicitement  $K_{2n}$  à partir de (2.7) on peut montrer (2.6). (2.5) se vérifie sans difficultés. Pour établir le lemme il suffit de remarquer que le raisonnement ci-dessus reste valable pour les moments  $m_n^N(t)$ , avec les mêmes constantes  $K_n$  et  $K'_n$ .  $\square$ .

**LEMME 2.3 :** *Il existe une famille de lois  $\{ \nu_t ; 0 \leq t \leq T \}$ , et une sous-suite  $\{ \nu_t^N ; 0 \leq t \leq T \}_{N \in \mathbb{N}}$  extraite de  $\{ \mu_t^N ; 0 \leq t \leq T \}_{N \in \mathbb{N}}$ , telles que :  $\nu_t^N \xrightarrow{N \rightarrow \infty} \nu_t$ ,  $0 \leq t \leq T$ .  $\square$*

*Preuve :* D'après le lemme 2.2, (i), (ii),  $\forall n$  fixé, la famille  $\{ m_n^N(\cdot) ; N > 1/2(n + 1) \}$  est bornée et équicontinue dans  $C[0, T]$ , donc relativement compacte. Par un procédé de diagonalisation de Cantor on montre qu'il existe une suite croissante d'entiers  $\{ N_p ; p \in \mathbb{N} \}$  et des fonctions  $m_n^* \in C[0, T]$ , telles que :

$$(2.8) \quad m_n^{N_p}(\cdot) \xrightarrow{p \rightarrow \infty} m_n^*(\cdot) \quad \text{dans } C[0, T], \quad \forall n.$$

Par ailleurs on considère le résultat suivant (cf. Shohat-Tamarkin [11]) :

Étant donnée une suite  $(m_n)$  de réels, une condition nécessaire et suffisante pour qu'il existe une mesure non négative qui admette  $m_n$  pour moments  $(\forall n)$ , est que :

$$\left\{ \begin{array}{l} \forall k \in \mathbb{N}, C_0, C_1, \dots, C_k \in \overline{\mathbb{R}} \\ \left[ \sum_{i=0}^k C_i x^i \geq 0, \forall x \in \mathbb{R} \right] \Rightarrow \left[ \sum_{i=0}^k C_i m_i \geq 0 \right]. \end{array} \right.$$

Cette dernière propriété est satisfaite par  $\{ m_n^{N_p}(t) \}_{n \in \mathbb{N}}$ , donc se conserve par passage à la limite lorsque  $p \rightarrow \infty$ . D'après (2.8), il existe donc une mesure non négative  $\nu_t$  qui admet  $m_n^*(t)$  pour moments  $(\forall n, 0 \leq t \leq T)$ . Ainsi pour  $t \leq T$  :

$$(2.9) \quad \int x^n \mu_t^{N_p}(dx) \xrightarrow{N \rightarrow \infty} \int x^n \nu_t(dx).$$

D'après le lemme 2.2, la série  $\sum \frac{\theta^n}{n!} m_n^*(t)$  admet un rayon de convergence strictement positif, donc (cf. Billingsley [1])  $\nu_t$  est l'unique loi sur  $\mathbb{R}$  qui vérifie (2.9), ce qui permet d'affirmer (cf. Breiman [2], p. 181) que  $\mu_t^{N_p} \Rightarrow \nu_t$  quand  $p \rightarrow \infty$ .

**LEMME 2.4 :** *Sous les hypothèses (H1)-(H3), l'équation de Fokker-Planck (2.2), admet une unique solution  $t \rightarrow \mu_t$ , fonction à valeurs dans l'espace des mesures non négatives sur  $\mathbb{R}$ .  $\square$*

*Preuve :* En utilisant la formule de Itô on vérifie aisément que la loi de  $X_t$  résout l'équation (2.2), d'où l'existence d'une solution. Soit  $\phi(\cdot, \cdot) \in$

$C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R})$  et  $\tilde{\mu}$  une solution de (2.2) à valeurs dans l'espace des mesures signées sur  $\mathbb{R}$ . Alors,

$$(2.10) \quad \langle \tilde{\mu}_t, \phi(t) \rangle = \langle \mu_0, \phi(0) \rangle + \int_0^t \langle \tilde{\mu}_s, \phi'(s) + L\phi(s) \rangle ds.$$

Par ailleurs, on considère l'équation aux dérivées partielles rétrograde,

$$(2.11) \quad \frac{\partial v(s)}{\partial s} + Lv(s) = 0, \quad s < t, \quad v(t) = \bar{v}.$$

D'après les hypothèses faites, et en utilisant des théorèmes de régularité des solutions des EDP paraboliques (cf. Ladyzenskaja *et al.* [7]), on a :  $\forall \bar{v} \in \mathbb{D}(\mathbb{R})$ , (2.11) admet une solution  $v \in C_b^{1,2}([0, t] \times \mathbb{R})$ , où  $\mathbb{D}(\mathbb{R})$  désigne l'espace des applications  $\mathbb{R} \rightarrow \mathbb{R}$  de classe  $C^\infty$  à support compact. Après différence entre deux solutions, pour démontrer l'unicité il suffit de vérifier que :  $\mu_0 = 0 \Rightarrow \{ \tilde{\mu}_t = 0, 0 \leq t \}$ .

Soit  $t \geq 0$  et  $\bar{v} \in \mathbb{D}(\mathbb{R})$ , par (2.11) on associe à  $\bar{v}$  une application  $v \in C_b^{1,2}([0, t] \times \mathbb{R})$ . D'après (2.10) (avec  $\mu_0 = 0$ ) et (2.11) :

$$\langle \tilde{\mu}_t, v(t) \rangle = \int_0^t \langle \tilde{\mu}_s, v'(s) + Lv(s) \rangle ds = 0$$

donc  $\langle \tilde{\mu}_t, v(t) \rangle = \langle \tilde{\mu}_t, \bar{v} \rangle = 0$ , ceci pour tout  $\bar{v} \in \mathbb{D}(\mathbb{R})$ , d'où  $\tilde{\mu}_t = 0$ .  $\square$

*Preuve du théorème 2.1* : Si on établit que :

$$(2.12) \quad \langle v_t, \phi \rangle = \langle \mu_0, \phi \rangle + \int_0^t \langle v_s, L\phi \rangle ds, \quad t \leq T, \quad \forall \phi \in C_b^2(\mathbb{R})$$

où  $\{ v_t \}$  est la limite d'une sous-suite dont l'existence est assurée par le lemme 2.3, alors d'après le lemme 2.4 on a  $v_t = \mu_t$ ,  $0 \leq t \leq T$ . On en déduit donc qu'une sous-suite de  $\mu_t^N$  converge vers  $\mu_t$ . Mais en reprenant la démonstration, par unicité de la limite on montre qu'il y a convergence de toute la suite. Il s'agit donc de montrer que (2.12) est vérifiée pour tout  $\phi \in C_b^2(\mathbb{R})$ . Nous allons considérer plusieurs étapes.

*Étape 1* :  $\phi$  polynôme. Pour tout  $N \geq (d+1)/2$ ,  $d = \text{degré } \phi$ , on a :

$$\langle v_t^N, \phi \rangle = \langle \mu_0^N, \phi \rangle + \int_0^t \langle v_s^N, L\phi \rangle ds$$

donc (2.12) s'obtient par convergence dominée lorsque  $N \rightarrow \infty$ .

*Étape 2* :  $\phi(x) = e^{i\theta x} p(x)$ ,  $\theta \in \mathbb{R}$ ,  $p$  polynôme. Prenons d'abord  $\phi(x) = e^{i\theta x}$ ,  $|\theta| \leq \theta_1$ , où  $\theta_1$  est le rayon de convergence donné par le lemme 2.2, (iii).

On pose :  $\phi_n(x) = \sum_{k=0}^n (i\theta x)^k / k !$

$\phi_n$  vérifie (2.12), donc quand  $n \rightarrow \infty$  on obtient :  $\Phi(\theta) = 0, \forall |\theta| \leq \theta_1$ , où  $\Phi(\theta) := \langle \nu_t, \phi \rangle - \langle \mu_0, \phi \rangle - \int_0^t \langle \nu_s, L\phi \rangle ds$ .

Ainsi  $\forall p \geq 1, \Phi^{(p)}(\theta) = 0, |\theta| \leq \theta_1$ , on en déduit que (2.12) est vraie pour tout  $\phi$  de la forme  $e^{i\theta x} p(x), |\theta| \leq \theta_1, p$  polynôme. En utilisant l'inégalité :

$$\left| \exp \{ i(\theta + \theta_1) x \} + \exp \{ i\theta_1 x \} \sum_{k=0}^n \frac{(i\theta x)^k}{k !} \right| \leq c \frac{|\theta|^{n+1}}{(n + 1)!} |x|^{n+1}$$

on montre par le même raisonnement que (2.12) est vérifiée pour tout  $\phi$  de la forme  $e^{i\theta x} p(x), |\theta| \leq 2\theta, p$  polynôme. Ainsi de suite on démontre l'étape 2.

*Étape 3* :  $\phi \in C^2(\mathbb{R})$  à support compact. Il existe  $\phi_n$  de la forme :

$$\phi_n(x) = \sum_{k=-n}^n a_k^n \exp \{ i b_k^n x \} \quad \text{telle que} \quad \|\phi^{(j)} - \phi_n^{(j)}\|_\infty \xrightarrow{n \rightarrow \infty} 0 \quad (j = 0, 1, 2).$$

(2.12) est vérifiée pour  $\phi_n(\forall n)$ , donc pour  $\phi$  par passage à la limite.

*Étape 4* :  $\phi \in C_b^2(\mathbb{R})$ . Cette étape se démontre à partir de l'étape précédente par convergence dominée.  $\square$

*Remarque 2.5* : On a démontré la conservation du critère de Cauchy : Si  $\overline{\lim}_{n \rightarrow \infty} \{ m_{2n}(0) / (2n)! \}^{1/2n} < \infty$ , alors  $\overline{\lim}_{n \rightarrow \infty} \{ m_{2n}(t) / (2n)! \}^{1/2n} < \infty$  pour tout  $t \in [0, T]$ . Ce résultat sera utilisé pour la démonstration de la convergence dans le cas du filtrage non linéaire.  $\square$

### 3. RÉSOLUTION NUMÉRIQUE DE L'ÉQUATION DE ZAKAI

#### 3.1. Filtrage avec observation en temps discret

On considère le système :

$$(3.1) \quad \begin{cases} dX_t = b(X_t) dt + \sigma(X_t) dW_t, & X_0 \sim \mu_0, \\ y_k = h(X_{t_k}) + v_k, \end{cases}$$



où  $0 \leq t \leq T$ ,  $0 < t_1 < \dots < t_K = T$  est une suite d'instants donnés, pour simplifier nous prendrons  $t_k = k\Delta$ , avec  $\Delta = T/K$  ( $K \in \mathbb{N}$ ).

$\{X_t\}$ ,  $\{W_t\}$ ,  $\{y_k\}$  et  $\{v_k\}$  sont des processus à valeurs dans  $\mathbb{R}$ ;  $\{v_k\}$  est une suite de variables gaussiennes indépendantes,  $v_k \sim N(0, R)$ ;  $\{W_t\}$  est un processus de Wiener standard indépendant de  $\{v_k\}$ ;  $X_0$  est indépendant de  $W$  et  $\{v_k\}$ . Supposons satisfaites les hypothèses (H1)-(H4) de la section précédente, ainsi que l'hypothèse :

(H5)  $h : \mathbb{R} \rightarrow \mathbb{R}$  est une application mesurable et bornée.

$\{X_t\}$  décrit l'évolution d'un système physique,  $\{y_k\}$  son observation en temps discret. Le problème de filtrage consiste à déterminer  $\eta_t$ , la loi conditionnelle de  $X_t$  sachant  $\{y_k; t_k \leq t\} = \{y_1, \dots, y_{[t/\Delta]}\}$  ( $[t/\Delta]$  désigne la partie entière de  $t/\Delta$ ).

Entre deux instants d'observation (i.e.  $t_{k-1} < t < t_k$ ), l'évolution de  $\eta_t$  est décrite par l'équation de Fokker-Planck :

$$\frac{d}{dt} \langle \eta_t, \phi \rangle = \langle \eta_t, L\phi \rangle, \quad \forall \phi \in C_b^2(\mathbb{R}).$$

A l'instant d'observation  $t = t_k$ , il est facile de vérifier en utilisant la formule de Bayes, que :

$$\langle \eta_{t_k}, \phi \rangle = \frac{\langle \eta_{t_k^-}, f(\cdot, y_k) \phi \rangle}{\langle \eta_{t_k^-}, f(\cdot, y_k) \rangle}$$

$$\text{où :} \quad * \quad \langle \eta_{t_k^-}, \phi \rangle := \lim_{t \rightarrow t_k, t < t_k} \langle \eta_t, \phi \rangle;$$

$$* \quad f(x, y) := \exp \left\{ \frac{1}{R} h(x) y - \frac{1}{2R} h(x)^2 \right\}.$$

Ainsi  $\{\eta_t\}$  est solution de l'équation :

$$\left\{ \begin{array}{l} \langle \eta_t, \phi \rangle = \langle \mu_0, \phi \rangle + \int_0^t \langle \eta_s, L\phi \rangle ds + \sum_{k=1}^{[t/\Delta]} \frac{\langle \eta_{t_k^-}, \{f(\cdot, y_k) - 1\} \phi \rangle}{\langle \eta_{t_k^-}, f(\cdot, y_k) \rangle} \\ \forall \phi \in C_b^2(\mathbb{R}). \end{array} \right. \quad (3.2)$$

On se propose d'approcher  $\eta_t$  par une mesure de probabilité  $\eta_t^N$ , de la forme :

$$\eta_t^N = \sum_{i=1}^N a_i^t \delta(x_i^t), \text{ où les processus stochastiques } \{a_i^t\} \text{ et } \{x_i^t\} \text{ sont déterminés}$$

en posant :

$$\left\{ \begin{aligned} \langle \eta_t^N, \phi \rangle &= \langle \mu_0, \phi \rangle + \int_0^t \langle \eta_s^N, L\phi \rangle ds + \sum_{k=1}^{[t/\Delta]} \frac{\langle \eta_{t_k^-}^N, \{ f(\cdot, y_k) - 1 \} \phi \rangle}{\langle \eta_{t_k^-}^N, f(\cdot, y_k) \rangle} \\ \forall \phi &\in \Pi_{2N-1}. \end{aligned} \right.$$

(3.3)

**THÉORÈME 3.1 :** *Sous les hypothèses (H1)-(H5), pour toute trajectoire  $\{y_1, \dots, y_K\}$  donnée, l'approximation de Gauss-Galerkin est convergente :  $\eta_t^N \xrightarrow[N \rightarrow \infty]{} \eta_t, 0 \leq t \leq T.$  □*

*Preuve :* On note  $m_n(t) = \int x^n \eta_t(dx)$ ; supposons que

$$(3.4) \quad \eta_t^N \xrightarrow[N \rightarrow \infty]{} \eta_t,$$

$$(3.5) \quad \overline{\lim}_{n \rightarrow \infty} \{ m_{2n}(t)/(2n)! \}^{1/2n} < \infty,$$

soient vérifiées pour  $t = t_{k-1}$ ; nous allons montrer que (3.4) (3.5) sont vérifiées pour  $t \in [t_{k-1}, t_k]$ . Pour démontrer le théorème il nous suffira donc d'établir (3.4) (3.5) pour  $t = 0$ .

Pour  $t \in ]t_{k-1}, t_k[$ , l'évolution de  $\eta_t$  est décrite par l'équation de Fokker-Planck, on déduit du théorème 2.1 (et de (3.5) en  $t = t_{k-1}$ ) que (3.4) est vérifiée pour tout  $t \in ]t_{k-1}, t_k[$ . Comme :

$$\langle \eta_{t_k}^N, \phi \rangle = \frac{\langle \eta_{t_k^-}^N, f(\cdot, y_k) \phi \rangle}{\langle \eta_{t_k^-}^N, f(\cdot, y_k) \rangle},$$

on en déduit que (3.4) est aussi vraie pour  $t = t_k$ . Par ailleurs,

$$\begin{aligned} \overline{\lim}_{n \rightarrow \infty} \left\{ \frac{m_{2n}(t_k)}{(2n)!} \right\}^{1/2n} &= \overline{\lim}_{n \rightarrow \infty} \left\{ \frac{1}{(2n)!} \frac{\int f(x, y_k) x^{2n} \eta_{t_k^-}(dx)}{\int f(x, y_k) \eta_{t_k^-}(dx)} \right\}^{1/2n} \\ &\leq \overline{\lim}_{n \rightarrow \infty} \{ m_{2n}(t_k^-)/(2n)! \}^{1/2n}. \end{aligned}$$

En utilisant (3.4) (pour  $t = t_{k-1}$ ) et la remarque 2.5, on montre que cette dernière expression est finie. On en déduit donc que (3.5) est vraie pour  $t = t_k$ .

Pour terminer la démonstration, il suffit de vérifier (3.4) (3.5) pour  $t = 0$ . D'après l'équation (3.3),  $\eta_0^N$  est l'approximation de Gauss-Christoffel de

$\eta_0 = \mu_0$ , et la convergence  $\eta_0^N \xrightarrow{N \rightarrow \infty} \eta_0$  se déduit donc du théorème 2.1. Par ailleurs (3.5) en  $t = 0$  est exactement l'hypothèse (H4).  $\square$

**3.2. Filtrage avec observation en temps continu**

On considère le système non linéaire :

$$(3.6) \quad \begin{cases} dX_t = b(X_t) dt + \sigma(X_t) dW_t, & X_0 \sim \mu_0; \\ dY_t = h(X_t) dt + dV_t, & Y_0 = 0. \end{cases}$$

$0 \leq t \leq T$ . On suppose les hypothèses des sections précédentes satisfaites. L'observation  $\{ Y_t \}$  à valeurs dans  $\mathbb{R}$ , est ici en temps continu,  $\{ V_t \}$  est un processus de Wiener standard indépendant de  $X_0$  et  $\{ W_t \}$ .

Le problème de filtrage consiste à déterminer  $v_t$  la loi conditionnelle de  $X_t$  sachant  $\mathcal{F}_t := \sigma \{ Y_s; s \leq t \}$ . Pour caractériser  $v$ , on peut utiliser la méthode de la probabilité de référence. Soit  $\hat{P}$  la loi déterminée par :

$$\frac{d\hat{P}}{dP} = \{ Z_T \}^{-1}, \quad \text{avec} \quad Z_t = \exp \int_0^t \left\{ h(X_s) dY_s - \frac{1}{2} h(X_s)^2 ds \right\}.$$

Le calcul de la loi conditionnelle de  $X_t$  sachant  $\mathcal{F}_t$  sous  $P$ , est lié à une expression calculée sous la probabilité  $\hat{P}$ , par la formule de Kallianpur-Striebel :

$$E[\phi(X_t) | \mathcal{F}_t] = \hat{E}[\phi(X_t) Z_t | \mathcal{F}_t] / \hat{E}[Z_t | \mathcal{F}_t].$$

On définit  $\tilde{v}$  la loi conditionnelle non normalisée de  $X_t$  sachant  $\mathcal{F}_t$ , en posant :  $\langle \tilde{v}_t, \phi \rangle := \hat{E}[\phi(X_t) Z_t | \mathcal{F}_t]$ .  $\tilde{v}$  est solution de l'équation de Zakaï (écrite sous forme faible) :

$$(3.7) \quad \langle \tilde{v}_t, \phi \rangle = \langle \mu_0, \phi \rangle + \int_0^t \langle \tilde{v}_s, L\phi \rangle ds + \int_0^t \langle \tilde{v}_s, h\phi \rangle dY_s, \quad \forall \phi \in C_b^2(\mathbb{R}).$$

On peut déterminer  $\tilde{v}_t^N$  — l'approximation de Gauss-Galerkin de  $\tilde{v}_t$  —, mais après discrétisation en temps, l'équation de  $\tilde{v}_t^N$  ne fait intervenir qu'une observation discrète en temps. Il est donc préférable de discrétiser directement l'équation d'observation dans (3.6) :

$$y_k = h(X_{t_k}) + v_k$$

( $t_k = k\Delta$ ), où  $v_k := (V_{t_{k+1}} - V_{t_k})/\Delta$  et  $y_k$  est l'approximation de  $\{ Y_{t_{k+1}} - Y_{t_k} \}/\Delta$ .

On définit  $\mathcal{F}_t^\Delta := \sigma \{ y_1, \dots, y_{[t/\Delta]} \}$  et  ${}^\Delta v_t$  la loi conditionnelle de  $X_t$  sachant  $\mathcal{F}_t^\Delta$ . Comme nous l'avons vu dans la Section 3.1, l'évolution de  $\{ {}^\Delta v_t \}$  est décrite par l'équation :

$$(3.8) \left\{ \begin{aligned} \langle {}^\Delta v_t, \phi \rangle &= \langle \mu_0, \phi \rangle + \int_0^t \langle {}^\Delta v_s, L\phi \rangle ds + \\ &+ \sum_{k=1}^{[t/\Delta]} \frac{\langle {}^\Delta v_{t_k^-}, \{ f_\Delta(\cdot, y_k) - 1 \} \phi \rangle}{\langle {}^\Delta v_{t_k^-}, f_\Delta(\cdot, y_k) \rangle} \\ \forall \phi &\in C_b^2(\mathbb{R}), \end{aligned} \right.$$

avec  $f_\Delta(x, y) := \exp \{ h(x) y \Delta - 1/2 h(x)^2 \Delta \}$ . On a le résultat suivant (la démonstration de ce théorème se trouve dans Campillo [3]) :

**THÉORÈME 3.2 :** *Outre les hypothèses (H1)-(H5), supposons que  $h \in C_b^2(\mathbb{R})$ , alors pour toute trajectoire observée  $\{ Y_s; s \leq t \} : {}^\Delta v_t \xrightarrow{\Delta \rightarrow 0} v_t, 0 \leq t \leq T$  (à condition de définir  $v_t$  « sous forme robuste », cf. par exemple Pardoux [9]).  $\square$*

On utilise la méthode de Gauss-Galerkin afin d'approcher  ${}^\Delta v_t$  par une mesure de probabilité  ${}^\Delta v_t^N$  de la forme :  ${}^\Delta v_t^N = \sum_{i=1}^N a_i^t \delta(x_t^i)$ . Les processus stochastiques  $\{ a_i^t \}$  et  $\{ x_t^i \}$  — qui dépendent de  $\Delta$  et  $N$  — sont déterminés en posant :

$$(3.9) \left\{ \begin{aligned} \langle {}^\Delta v_t^N, \phi \rangle &= \langle \mu_0, \phi \rangle + \int_0^t \langle {}^\Delta v_s^N, L\phi \rangle ds + \\ &+ \sum_{k=1}^{[t/\Delta]} \frac{\langle {}^\Delta v_{t_k^-}^N, \{ f_\Delta(\cdot, y_k) - 1 \} \phi \rangle}{\langle {}^\Delta v_{t_k^-}^N, f_\Delta(\cdot, y_k) \rangle} \\ \forall \phi &\in \Pi_{2N-1}. \end{aligned} \right.$$

D'après le théorème 3.1, pour tout  $\Delta$ , on a la convergence suivante :

$$(3.10) \quad {}^\Delta v_t^N(\omega) \xrightarrow{N \rightarrow \infty} {}^\Delta v_t(\omega), \quad \forall \omega \text{ p.s.}, \quad 0 \leq t \leq T.$$

On introduit  $\mathbb{M}_+$  l'ensemble des mesures de probabilités sur  $\mathbb{R}$ , et  $C_u(\mathbb{R})$  l'ensemble des fonctions bornées, uniformément continues. Soit  $\{ f_p; p \geq 0 \}$  une suite dense dans  $C_u(\mathbb{R})$ . On définit :

$$d(\mu, \lambda) := \sum_{p \in \mathbb{N}} \frac{1}{2^p} \frac{|\langle \mu, f_p \rangle - \langle \lambda, f_p \rangle|}{\| f_p \|}$$

( $\|f\| = \sup \{|f(x)|, x \in \mathbb{R}\}$ )  $d(\cdot, \cdot)$  définit une métrique sur  $\mathbb{M}_+$ , qui induit une topologie équivalente à celle induite par la convergence étroite des mesures (cf Stroock-Varadhan [12]) Ainsi (3 10) et le théorème (3 2) impliquent

$$(3 11) \quad \left\{ \begin{array}{l} \text{A tout } \Delta > 0 \text{ on peut associer } N(\Delta) \in \mathbb{N} \text{ tel que} \\ \Delta v_t^{N(\Delta)}(\omega) \xrightarrow{\Delta \rightarrow 0} v_t(\omega), \quad \forall \omega \text{ p s}, \quad 0 \leq t \leq T \end{array} \right.$$

Ce dernier résultat de convergence n'est pas entièrement satisfaisant (on ne sait pas choisir explicitement  $N(\Delta)$ ) Mais comme en pratique l'équation d'observation est toujours en temps discret, pour un pas de discrétisation  $\Delta$  donné la convergence (3 10) est satisfaisante

On aurait pu obtenir une convergence en moyenne quadratique dans le cas des observations en temps continu, ce qui n'est pas non plus d'un grand intérêt

Pour obtenir une convergence pour chaque trajectoire observée, on pourrait penser utiliser la « forme robuste » de l'équation de Zakai, cela n'a pas été réalisable, car la multiplication par  $\exp\{-h(x) Y_t\}$  fait sortir de l'espace des polynômes de degré au plus égal à  $2N - 1$

## 4. ÉTUDE NUMÉRIQUE

### 4.1. Présentation de l'algorithme

#### 4 1 a Équation de Fokker-Planck

Nous considérons d'abord l'algorithme d'approximation de l'équation de Fokker-Planck La mise en œuvre pratique de celui-ci nécessite une discrétisation en temps de l'équation

$$(4 1) \quad \langle \mu_t^N, \tilde{\pi}_p \rangle = \langle \mu_0, \tilde{\pi}_p \rangle + \int_0^t \langle \mu_s^N, \tilde{\pi}_p \rangle ds, \quad 0 \leq t \leq T, \quad p = 0 \quad 2N - 1$$

où  $\{\tilde{\pi}_p, p = 0 \quad 2N - 1\}$  désigne une base de  $\Pi_{2N-1}$  Tous les schémas de discrétisation peuvent être envisagés, mais afin de simplifier l'exposé nous allons utiliser le schéma d'Euler L'équation (4 1) discrétisée s'écrit donc

$$(4.2) \quad \begin{cases} \langle \mu_0^N, \tilde{\pi}_p \rangle = \langle \mu_0, \tilde{\pi}_p \rangle & (p = 0 \dots 2N - 1) \\ \langle \mu_k^N, \tilde{\pi}_p \rangle = \langle \mu_{k-1}^N, \tilde{\pi}_p + (t_k - t_{k-1}) L\tilde{\pi}_p \rangle & (p = 0 \dots 2N - 1), \quad k = 1 \dots K \end{cases}$$

où  $0 = t_0 < t_1 < \dots < t_K = T$  et  $\mu_k^N = \sum_{i=1}^N a_k^i \delta(x_k^i)$  désigne l'approximation de  $\mu_t^N$  à l'instant  $t = t_k$ .

A partir de  $\{ a_{k-1}^i, x_{k-1}^i; i = 1 \dots N \}$ , (4.2) nous permet de calculer les moments modifiés de  $\mu_k^N$  relativement à la base  $\{ \tilde{\pi}_p \}$ , soit :

$$\tilde{m}_k^p := \langle \mu_k^N, \tilde{\pi}_p \rangle = \sum_{i=1}^N a_{k-1}^i \{ \tilde{\pi}_p + (t_k - t_{k-1}) L\tilde{\pi}_p \} (x_{k-1}^i).$$

Étant donné  $\{ \tilde{m}_k^p; p = 0 \dots 2N - 1 \}$ , on veut maintenant calculer  $\{ a_k^i, x_k^i; i = 1 \dots N \}$  tel que :

$$(4.3) \quad \sum_{i=1}^N a_k^i \tilde{\pi}_p(x_k^i) = \tilde{m}_k^p, \quad p = 0 \dots 2N - 1.$$

Pour résoudre ce problème on utilise la méthode de quadrature de Gauss-Christoffel (tous les résultats énoncés dans la suite proviennent de Gautschi [6] et Wheeler [13]).

On introduit  $\{ \pi_p; p = 0 \dots N \}$ , la famille des polynômes orthogonaux relativement à la mesure  $\mu_k^N$ , i.e.  $\pi_p$  est de degré  $p$  et  $\langle \mu_k^N, \pi_p \pi_q \rangle = 0$  si  $p \neq q$ . Ces polynômes sont définis à une constante multiplicative près et on peut décider par exemple que le coefficient du monôme de plus haut degré dans  $\pi_p$  est 1 ( $p = 0 \dots N$ ). Dans ce cas la famille  $\{ \pi_p \}$  satisfait à une relation de la forme :

$$(4.4) \quad \begin{cases} \pi_{p+1}(x) = (x - \alpha_p) \pi_p(x) - \beta_p \pi_{p-1}(x), & p = 0 \dots N - 1 \\ \pi_0(x) = 1, \quad \pi_{-1}(x) = 0, & \beta_p > 0 \text{ pour } p \geq 1, \quad \beta_0 = 0. \end{cases}$$

Le calcul de  $\{ a_k^i, x_k^i \}$  solution de (4.3), se ramène au calcul des coefficients  $\{ \alpha_p, \beta_p; p = 0 \dots N - 1 \}$  de la manière suivante : posons

$$(4.5) \quad J_N = \begin{bmatrix} \alpha_0 & \sqrt{\beta_0} & & & \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & (0) \\ & & & & \\ (0) & \sqrt{\beta_{N-2}} & \alpha_{N-2} & \sqrt{\beta_{N-1}} & \\ & & \sqrt{\beta_{N-1}} & \alpha_{N-1} & \end{bmatrix}.$$

$J_N$  admet  $N$  valeurs propres réelles, deux à deux distinctes  $\xi_1 \dots \xi_N$ ; soit  $V_1 \dots V_N$  les vecteurs propres orthonormés respectivement associés. On a le résultat suivant :

$$(4.6) \quad \begin{cases} a_k^i = (V_{1i})^2, & i = 1 \dots N \\ x_k^i = \xi_i, & i = 1 \dots N \end{cases}$$

où  $V_{1i}$  désigne la première composante du vecteur  $V_i$ .

Il reste à donner un algorithme permettant de calculer  $\{\alpha_p, \beta_p; p = 0 \dots N - 1\}$  — les coefficients définissant  $\{\pi_p\}$  — à partir de  $\{\tilde{m}_k^p; p = 0 \dots 2N - 1\}$  les moments modifiés de  $\mu_k^N$  relativement à la base  $\{\tilde{\pi}_p\}$ . Numériquement, il est préférable d'utiliser une base  $\{\tilde{\pi}_p\}$  formée de vecteurs orthogonaux. Nous allons utiliser la base définie par :

$$(4.7) \quad \begin{cases} \tilde{\pi}_{p+1}(x) = (x - \tilde{\alpha}_p) \tilde{\pi}_p(x) - \tilde{\beta}_p \tilde{\pi}_{p-1}(x), & p = 0 \dots 2N - 2 \\ \tilde{\pi}_0(x) = 1, \quad \tilde{\pi}_{-1}(x) = 0, \quad \tilde{\beta}_p > 0 \text{ pour } p \geq 1, \quad \tilde{\beta}_0 = 0. \end{cases}$$

où les coefficients  $\{\tilde{\alpha}_p, \tilde{\beta}_p; p = 0 \dots 2N - 2\}$  sont donnés.

Le calcul de  $\{\alpha_p, \beta_p; p = 0 \dots N - 1\}$  à partir de  $\{\tilde{m}_k^p; p = 0 \dots 2N - 1\}$  et de  $\{\tilde{\alpha}_p, \tilde{\beta}_p; p = 0 \dots 2N - 2\}$ , s'effectue à l'aide de l'algorithme de Chebyshev modifié (cf. [13]) :

*Initialisation :*

$$(4.8) \quad \begin{cases} \sigma(-1, 0) \leftarrow 0 \\ \sigma(0, q) \leftarrow \tilde{m}_k^q, \quad q = 0 \dots 2N - 1 \\ \alpha_0 \leftarrow \tilde{\alpha}_0 + \tilde{m}_k^1 / \tilde{m}_k^0 \\ \beta_0 \leftarrow 0. \end{cases}$$

*Incrémentations :*

pour  $k = 1 \dots N - 1$  faire :

$$(4.8'') \quad \begin{cases} \left. \begin{aligned} \sigma(p, q) &\leftarrow \sigma(p - 1, q + 1) - (\alpha_{p-1} - \tilde{\alpha}_q) \sigma(p - 1, q) \\ &\quad - \beta_{p-1} \sigma(p - 2, q) + \tilde{\beta}_q \sigma(p - 1, q - 1) \end{aligned} \right\} q = p \dots 2N - p + 1 \\ \alpha_p &\leftarrow \tilde{\alpha}_p + \frac{\sigma(p - 1, p)}{\sigma(p - 1, p - 1)} + \frac{\sigma(p, p + 1)}{\sigma(p, p)} \\ \beta_p &\leftarrow \sigma(p, p) / \sigma(p - 1, p - 1). \end{cases}$$

L'algorithme d'approximation de Gauss-Galerkin, pour l'équation de Fokker-Planck, s'écrit donc :

Données :

$$\begin{aligned}
 & * \tilde{\alpha}_p, \tilde{\beta}_p \quad (p = 0 \dots 2N - 2) \\
 & * \tilde{m}_0^p := \langle \mu_0, \tilde{\pi}_p \rangle \quad (p = 0 \dots 2N - 1).
 \end{aligned}$$

Incrémentations :

pour  $k = 0 \dots K$  faire :

$$(4.9) \quad \left\{ \begin{array}{l}
 1. \text{ Calcul de } \{ \alpha_p, \beta_p; p = 0 \dots N - 1 \} \text{ à partir de } \{ \tilde{\alpha}_p, \tilde{\beta}_p; p = 0 \dots 2N - 2 \} \text{ et } \{ \tilde{m}_k^p; p = 0 \dots 2N - 1 \}, \text{ avec l'algorithme (4.8).} \\
 2. \text{ Calcul des valeurs propres et vecteurs propres orthonormés de } J_N \text{ définie en (4.5).} \\
 3. \text{ Calcul de } \{ a_k^i, x_k^i; i = 1 \dots N \} \text{ par (4.6).} \\
 4. \tilde{m}_{k+1}^p \leftarrow \sum_{i=1}^N a_k^i \{ \tilde{\pi}_p + (t_{k+1} - t_k) L \tilde{\pi}_p \} (x_k^i) \quad (p = 0 \dots 2N - 1).
 \end{array} \right.$$

4.1. b. *Équation du filtrage non linéaire*

Pour le problème de filtrage non linéaire, on doit résoudre numériquement une équation de la forme (cf. (3.3) et (3.7)) :

$$(4.10) \quad \langle v_t^N, \tilde{\pi}_p \rangle = \langle \mu_0, \tilde{\pi}_p \rangle + \int_0^t \langle v_s^N, \tilde{\pi}_p \rangle ds + \sum_{q=1}^{[t/\Delta]} \frac{\langle v_{r_q}^N, (f(\cdot, y_q) - 1) \tilde{\pi}_p \rangle}{\langle v_{r_q}^N, f(\cdot, y_q) \rangle}$$

où  $r_q := q\Delta$  ( $q = 1 \dots Q$ ) désignent les instants d'observation,  $\Delta := T/Q$  pour  $Q \in \mathbb{N}$  donné.

La discrétisation de (4.10) se réduit donc à la discrétisation de l'intégrale  $\int \langle v_s^N, \tilde{\pi}_p \rangle ds$ . Soit  $0 = t_0 < t_1 < \dots < t_K = T$  une grille de discrétisation en temps contenant les instants d'observation  $\{ r_q \}$ . L'équation (4.10), après discrétisation à l'aide du schéma d'Euler, s'écrit :

$$\langle v_0^N, \tilde{\pi}_p \rangle = \langle \mu_0, \tilde{\pi}_p \rangle \quad (p = 0 \dots 2N - 1)$$



pour  $k = 1 \dots K$  :

$$(4.11) \left\{ \begin{array}{l} \langle v_{k^-}^N, \tilde{\pi}_p \rangle = \langle v_{k-1}^N, \tilde{\pi}_p + (t_k - t_{k-1}) L \tilde{\pi}_p \rangle \quad (p = 0 \dots 2N - 1) \\ \text{si } t_k \notin \{r_q; q = 1 \dots Q\} : v_k^N = v_{k^-}^N \\ \text{si } \exists q \in \{1, \dots, Q\} \text{ tel que } t_k = r_q : \\ \langle v_k^N, \tilde{\pi}_p \rangle = \frac{\langle v_{k^-}^N, f(\cdot, y_q) \tilde{\pi}_p \rangle}{\langle v_{k^-}^N, f(\cdot, y_q) \rangle} \quad (p = 0 \dots 2N - 1). \end{array} \right.$$

L'algorithme complet est alors équivalent à celui présenté pour l'équation de Fokker-Planck.

En pratique, la base  $\{\tilde{\pi}_p\}$  de  $\Pi_{2N-1}$  utilisée est celle des polynômes d'Hermite. Pour le calcul des valeurs propres de  $J_N$  nous avons utilisé une variante de l'algorithme *QL* pour les matrices symétriques et tridiagonales; le programme complet se trouve dans Garbow *et al.* [5].  $\square$

## 4.2. Exemple

Dans cette section nous présentons un exemple d'application de la méthode de Gauss-Galerkin en filtrage non linéaire. Les calculs ont été faits sur un ordinateur VAX 730 en FORTRAN 77 double précision.

La méthode d'approximation appliquée à l'équation de Fokker-Planck sur des exemples non linéaires a donné de bons résultats jusqu'à  $N = 10$  ( $N$  : nombre de points de Gauss). Au-delà on se heurte à des problèmes de mauvais conditionnement.

Pour le problème de filtrage, nous avons d'abord testé la méthode sur des exemples linéaires. Nous avons comparé les résultats obtenus à ceux donnés par un filtre de Kalman-Bucy. Là aussi nous avons obtenu de bons résultats, même avec très peu de points de Gauss ( $N = 3$  ou  $4$ ).

Nous présentons maintenant un exemple numérique; considérons le problème de filtrage non linéaire :

$$(4.12) \quad \begin{cases} dX_t = -\beta X_t dt + \sqrt{2\beta} \sigma dW_t, & X_0 \sim N(0, \sigma^2) \\ dY_t = b \exp(jX_t) dt + r dV_t, & Y_0 = 0 \end{cases}$$

$0 \leq t \leq T$ . Le processus de Wiener standard  $\{V_t\}$  et l'observation  $\{Y_t\}$  sont à valeurs dans le plan complexe ( $j^2 = -1$ );  $\{W_t\}$  est un processus de Wiener standard réel indépendant de  $\{V_t\}$ ;  $X_0$  est indépendant de  $\{W_t\}$  et  $\{V_t\}$ .

Soit  $v_t$  la loi conditionnelle de  $X_t$  sachant  $\mathcal{F}_t = \sigma\{Y_s; s \leq t\}$ . Nous avons mis en œuvre trois méthodes d'approximation de  $v_t$  :

- (AGG) *Approximation de Gauss-Galerkin* :  $v_t$  est approchée par une loi  $v_t^1$  de la forme  $v_t^1 = \sum_{i=1}^N a_i \delta(x_i)$  ( $N$  : nombre de points de Gauss).  
Le calcul de  $v_t^1$  a été présenté dans la Section 4.1.b.  $\square$
- (DF) *Différences Finies* : nous utilisons un schéma aux différences finies en espace, afin de résoudre numériquement l'équation de Zakaï pour la densité conditionnelle non normalisée de  $v_t$ .  $v_t$  est ainsi approchée par une loi  $v_t^2$  de la forme  $v_t^2(dx) = p(t, x) dx$  (pour les détails de cette méthode cf. Le Gland [8]).  $\square$
- (FKE) *Filtre de Kalman Étendu* :  $v_t$  est approchée par une loi gaussienne  $v_t^3 = N(\hat{X}_t^3, Q^3(t))$ ,  $\hat{X}_t^3$  et  $Q^3(t)$  sont les sorties du filtre de Kalman étendu associé à (4.12).  $\square$

*Remarques 4.1 :*

1. Ces trois méthodes sont en fait mises en œuvre après discrétisation en temps du système (4.12).
2. La condition initiale  $X_0$ , ainsi que les processus de Wiener (en temps discret : les bruits blancs gaussiens)  $\{W_t\}$  et  $\{V_t\}$  ont été simulés sur ordinateur.
3. La méthode (DF) est utilisée comme méthode de référence : nous allons comparer les moments conditionnels calculés par (AGG) avec ceux calculés par (DF). Cependant (DF) présente l'inconvénient de ne pouvoir être appliquée de manière simple dans le cas où le support de la loi  $v_t$  ne reste pas (lorsque  $t$  varie) dans un domaine borné et fixe de  $\mathbb{R}$ . En effet dans (DF) la densité conditionnelle  $p(t, x)$  est calculée sur un domaine  $[-M, M]$  fixé à l'avance.  $\square$

Au vu des exemples numériques (dont deux sont présentés en fin de section) on peut faire plusieurs constatations :

- (i) Les estimateurs  $\hat{X}_t^i := \langle v_t^i, x \rangle$  ( $i = 1, 2, 3$ ) de  $X_t$ , donnés par les trois méthodes, sont équivalents (cf. exemple 4.2, fig. 1). En revanche, contrairement à (AGG), (FKE) donne une mauvaise estimation de la variance conditionnelle  $Q(t) := \langle v_t, x^2 \rangle - \langle v_t, x \rangle^2$  (cf. exemple 4.2, fig. 2).
- (ii) (AGG) suit correctement l'évolution des moments conditionnels (pour l'exemple 4.2, les 14 premiers moments sont estimés de manière satisfaisante).
- (iii) Même pour un faible nombre de points de Gauss ( $N = 2$  dans l'exemple 4.3), (AGG) donne des résultats significatifs (cf. exemple 4.2, fig. 5).  $\square$

Exemple 4.2 : Nombre de points de Gauss :  $N = 10$ .

Instant final :  $T = 10$ .

Pas d'incrémentation :  $\Delta t = 0,01$ .

Coefficients :  $\beta = \sigma = b = 1, r = 0,5$ .

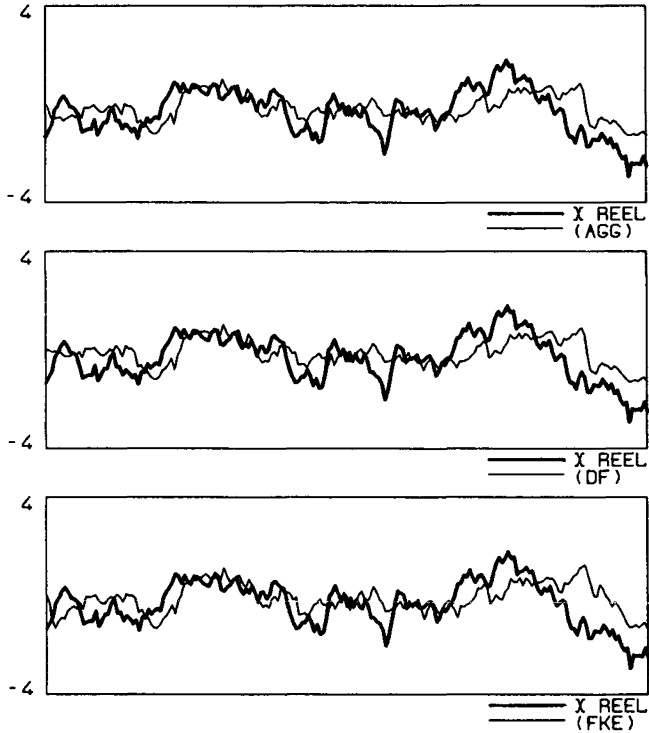


Figure 1. — Comparaison de  $X_i$ (X REEL)  
et des estimateurs  $\hat{X}_i := \langle \psi^i, x \rangle$  ( $i = 1, 2, 3$ ).

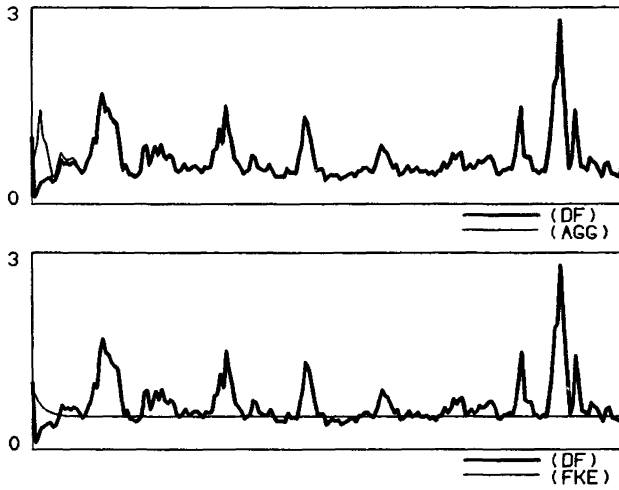


Figure 2. — Variances conditionnelles  $Q^i(t) := \langle v_t^i, x^2 \rangle - \langle v_t^i, x \rangle^2$  ( $i = 1, 2, 3$ ). Comparaisons (DF)-(AGG) et (DF)-(FKE).

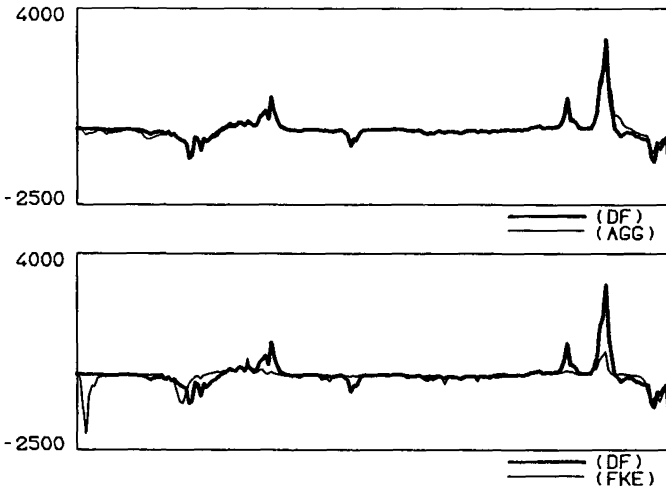


Figure 3. — Moments conditionnelles d'ordre 9 :  $\langle v_t^i, x^9 \rangle$  ( $i = 1, 2, 3$ ). Comparaisons (DF)-(AGG) et (DF)-(FKE).

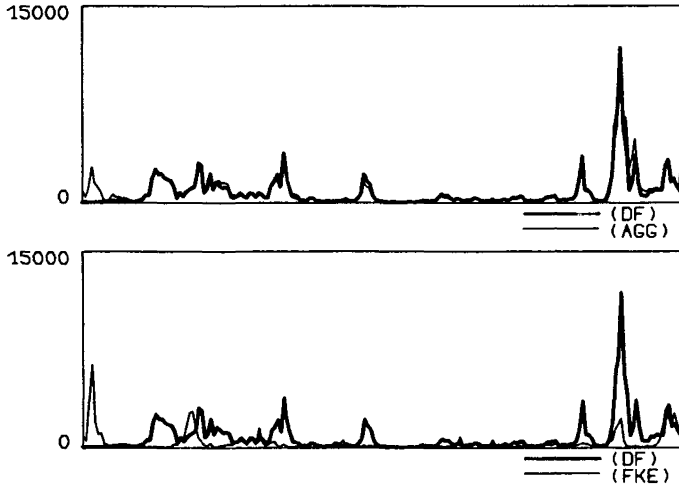


Figure 4. — Moments conditionnels d'ordre 10 :  $\langle v_i^i, x^{10} \rangle$  ( $i = 1, 2, 3$ ). Comparaisons (DF)-(AGG) et (DF)-(FKE).  $\square$

Exemple 4.3 : Nombre de points de Gauss :  $N = 2$ .

Instant final :  $T = 10$ .

Pas d'incrémentation :  $\Delta t = 0,01$ .

Coefficients :  $\beta = \sigma = b = r = 1$ .

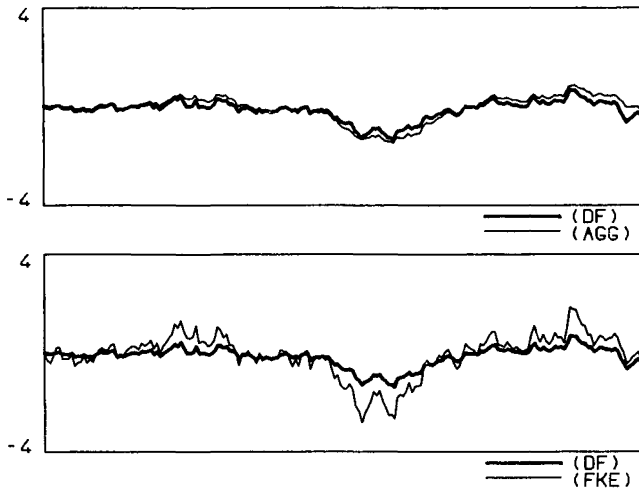


Figure 5. — Moments conditionnels d'ordre 3 :  $\langle v_i^i, x^3 \rangle$  ( $i = 1, 2, 3$ ). Comparaisons (DF)-(AGG) et (DF)-(FKE).

## BIBLIOGRAPHIE

- [1] P BILLINGSLEY, *Probability and Measure* J Wiley (1979)
- [2] L BREIMAN, *Probability* Addison-Wesley (1968)
- [3] F CAMPILLO, *Filtrage et Detection de Ruptures de Processus Partiellement Observe*, These de Troisieme Cycle, Universite de Provence (1984)
- [4] D A DAWSON, *Galerkin approximation of nonlinear Markov processes*, in Statistics & Related Topics, M Csorgo, D A Dawson, J N K Rao, A K Md Saleh (eds) North-Holland (1981)
- [5] B S GARBOW, J M BOYLE, J J DONGARA & C B BOYLER, *Matrix Eigen System Routines — EISPACK — Guide Extension*, Lecture Notes in Computer Science (1977)
- [6] W GAUTSCHI, *On generating orthogonal polynomials* SIAM J Sci Stat Comp, vol 3, n° 3 (1982), 289-317
- [7] O A LADYZENSKAJA, V A SOLONNIKOV, N N URAL'CEVA, *Linear and Quasi-linear Equations of Parabolic Type* Amer Math Socie (1968)
- [8] F LE GLAND, *Estimation de Parametres dans les Processus Stochastiques en Observation Incomplete Application a un Probleme de Radio-Astronomie* These de Docteur-Ingenieur, Universite Paris 9 (1981)
- [9] E PARDOUX, *Equations du filtrage non-lineaire de la prediction et du lissage* Stochastics, 6 (1982), 193-231
- [10] P A RAVIART, *An analysis of particle methods* CEME Course in Numerical Methods in Fluids Dynamics, Como, July 83, Lecture Notes in Mathematics Springer Verlag
- [11] J A SHOHAT & J D TAMARKIN, *The Problem of Moments* Amer Math Socie (1950)
- [12] D W STROOCK, S R S VARADHAN, *Multidimensional Diffusion Processes* Springer Verlag (1979)
- [13] J C WHEELER, *Modified moments and gaussian quadrature* Rocky Montain J Math, 4 (1974), 287-296