

Habilitation à diriger des recherches

présentée devant

L'université de Rennes 1
Institut Mathématique de Rennes

par

Fabien Campillo

QUELQUES APPLICATIONS DES PROCESSUS DE DIFFUSION FILTRAGE/STATISTIQUE - CONTRÔLE - HOMOGENÉISATION

Document de synthèse



Habilitation soutenue le 13 décembre 2004 devant le jury composé de \diamond M. JEAN DESHAYES, professeur à l'université de Rennes I, président \diamond M. JEAN JACOD, professeur à l'université Paris VI, rapporteur \diamond M. BERNARD YCART, professeur à l'université Paris VI, rapporteur \diamond M. PIERRE DEL MORAL, professeur à l'université de Nice Sophia-Antipolis, rapporteur \diamond M^{me} MICHÈLE BASSEVILLE, directeur de recherche au CNRS Rennes, examinateur \diamond M. BERNARD DELYON, professeur à l'université de Rennes I, examinateur \diamond M. FRANÇOIS LE GLAND, directeur de recherche à l'INRIA Rennes, examinateur \diamond M. ALEXANDER YU. VERETENNIKOV, professeur à l'université de Leeds, examinateur.

À ma famille, à mes amis.

REMERCIEMENTS

Je remercie vivement Jean Jacod, Bernard Ycart et Pierre Del Moral d'avoir accepté de rapporter sur cette habilitation et de participer au jury.

J'exprime également ma profonde gratitude à Michèle Basseville, Bernard Delyon, Jean Deshayes, François Le Gland et Sacha Veretennikov qui ont bien voulu faire partie de ce jury.

Je remercie spécialement Sacha qui me fait l'amitié de venir de Leeds.

Je remercie les membres de l'Institut de Recherche Mathématique de Rennes pour leur soutien dans cette démarche.

Travailler à l'Inria est un atout. La qualité des services de cet institut (administratifs, documentalistes, ingénieurs et techniciens informaticiens, services généraux...) et la valeur de ses chercheurs m'ont permis de travailler dans d'excellentes conditions. Je ne peux pas citer ici toutes les personnes que j'ai eues la chance de rencontrer à l'Inria, aussi vais-je simplement remercier les directeurs des unités de recherche qui ont su me "supporter" : Pierre Bernhard, Marc Berthod, Michel Cosnard à Sophia-Antipolis et Claude Labit à Rennes. Je souhaite également remercier les assistantes qui m'ont aidé toutes ces années : Anita Guiteau, France Limouzis, Ephie Deriche, Magali Djian, Sylvie Blanc, Huguette Bechu et enfin Stéphanie Lemaile.

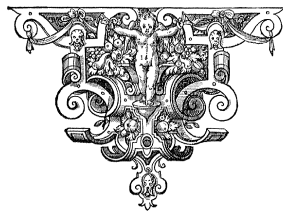


TABLE DES MATIÈRES

Avant-propos	vii
I Filtrage : statistique et approximation	1
I.1 Un modèle statistique	2
I.2 Normalité asymptotique locale	5
I.3 Convergence de la suite des processus du rapport de vraisemblance	8
I.4 Détection de rupture	9
I.4.1 Un cas linéaire	10
I.4.2 Un cas non linéaire	14
I.5 Détection de rupture : asymptotique non locale	15
I.6 Statistique et filtrage non linéaire	18
I.6.1 Fonction de log-vraisemblance	19
I.6.2 Estimation de paramètre : algorithme de gradient	19
I.6.3 Estimation de paramètre : algorithme EM	19
I.6.4 Le problème de détection	21
I.6.5 Méthodes numériques	22
Bibliographie	25
II Un problème de contrôle stochastique	35
II.1 Présentation du problème	35
II.1.1 Un exemple : un modèle de système de suspensions semi-actives	35
II.1.2 Un problème de contrôle stochastique	37
II.1.3 La mesure de probabilité invariante	37
II.2 Approche optimale	38
II.2.1 L'équation de Hamilton-Jacobi-Bellman (HJB)	38
II.2.2 Approximation	39
II.2.3 Résolution du problème approché	40
II.2.4 Convergence de la fonction coût	41
II.3 Approche sous-optimale	42
II.3.1 Algorithmes d'approximation stochastique	42
II.3.2 Discrétisation en temps	45
II.4 D'autres modèles	46
Bibliographie	48
III Homogénéisation	51
III.1 Homogénéisation d'opérateur parabolique aléatoire	51
III.1.1 Introduction	51
III.1.2 Résultats de convergence	54
III.1.3 Démonstrations	56
III.2 Méthodes de Monte Carlo pour un milieu hétérogène	58
III.2.1 Introduction	58
III.2.2 Marches aléatoires pour le calcul de coefficients effectifs	60
III.2.3 Simulation "exacte" pour des milieux fissurés	61
Bibliographie	64

AVANT-PROPOS



Après une thèse de troisième cycle sous la direction d'Étienne Pardoux à l'Université de Provence, soutenue en 1984, j'ai effectué mon service national dans le Groupe d'Étude et de Recherche en Détection Sous Marine (GERDSM) de la DCAN sous la responsabilité de Claude Jauffret. En 1986, après un court séjour post-doctoral chez Ulrich Haussmann à l'Université de Colombie Britannique, j'ai intégré le projet de recherche dirigé par Étienne Pardoux à l'INRIA Sophia-Antipolis. C'est dans le cadre de cette équipe que j'ai été amené à collaborer longuement avec François Le Gland. En 1996, j'ai monté mon projet de recherche INRIA localisé à Marseille Château-Gombert (Université de Provence). Depuis 2002 j'ai rejoint la nouvelle équipe que François Le Gland propose de former à l'IRISA Rennes.

Ce document de synthèse regroupe trois thèmes indépendants. Ceci reflète mon goût pour la diversité. L'INRIA offre également la possibilité de traiter des sujets sur un large spectre. Dans chacun des trois thèmes développés, j'ai eu la possibilité de travailler sur des aspects relativement théoriques aussi bien que sur des aspects relativement pratiques (chacun de ces trois sujets ont donné lieu à des relations industrielles contractuelles).

- Le premier est en fait le fil rouge de mes travaux et s'attache à différents problèmes de filtrage et de statistique pour des processus de diffusion partiellement observés.
- Le deuxième traite d'un problème de contrôle stochastique de type ergodique. C'est un contrat avec Renault et une collaboration avec des collègues mécaniciens qui sont à l'origine de ces travaux ; il s'agissait de calculer des lois de commande pour des amortisseurs semi-actifs.
- Le dernier enfin, traite de quelques problèmes d'homogénéisation pour des processus aléatoires en milieux hétérogène. Il s'agissait à l'origine d'un contrat avec l'IFP pour le développement de méthodes de Monte Carlo pour le calcul de coefficients effectifs.

En début de chacune de ces trois parties, je retrace le contexte de ces travaux.

CHAPITRE I

FILTRAGE : STATISTIQUE ET APPROXIMATION

Depuis ma thèse de troisième cycle, je m'intéresse aux problèmes de statistique pour des systèmes partiellement observés, ainsi qu'à leur approximation numérique. Cette thèse était constituée de deux travaux indépendants : l'analyse d'un problème de détection de changement dans un système linéaire partiellement observé et le développement d'une méthode d'approximation du filtre non linéaire.

Dans le problème de détection il s'agit de développer des algorithmes et d'en étudier également les propriétés asymptotiques. Sur le plan pratique, le problème est de développer des méthodes réalistes pour détecter un changement dans un système dynamique partiellement observé et d'identifier les caractéristiques du système après un éventuel changement. Sur le plan théorique, il s'agit d'appliquer les différentes techniques d'analyse statistique asymptotique appropriée. Enfin l'articulation de ces deux thèmes est également importante : l'analyse asymptotique permet également de proposer des procédures pratiques.

La recherche dans ce domaine est restée soutenue depuis les années 70. En 1976, Alan Willsky a proposé un des premiers articles de revue. En France, dans les années 80, Michèle Basseville et Albert Benveniste contribuent largement au développement et à la diffusion de ces méthodes aussi bien par leurs propres travaux qu'en animant la communauté scientifique. Michèle Basseville notamment a écrit un article de revue. Sur le plan mathématique, les statisticiens d'Orsay avaient investi ces domaines depuis des années : en particulier Jean Deshayes et Dominique Picard qui proposent des résultats très pointus et aboutis.

Concernant la statistique asymptotique, j'ai commencé par étudier les travaux de Ildar Ibragimov et Rafail Has'minskii ainsi que les premiers travaux de Yurii Kutoyants concernant les EDS. En 1983, ce dernier est de passage en France. Il est invité par Étienne Pardoux à donner un séminaire à l'Inria Sophia–Antipolis ; il me confie à cette occasion des épreuves de son premier livre à venir.

Yurii Kutoyants fera un plus long séjour à l'Inria Sophia–Antipolis au sein de l'équipe dirigée par Étienne Pardoux. Cela donnera une collaboration à trois avec François Le Gland. C'est essentiellement avec ce dernier que j'ai travaillé en statistique et/ou approximation en filtrage non linéaire. Cette collaboration s'est renouvelée depuis mon arrivée à Rennes.

L'autre axe de mes travaux débutés en thèse, concernait l'approximation numérique du filtre non linéaire. Suivant une idée suggérée par Donald Dawson, je proposais une méthode d'approximation de type Gauss–Galerkin. Après ma thèse j'ai passé un an au sein du

GERDSM Toulon, où j'ai travaillé en collaboration avec Claude Jauffret sur des algorithmes de filtrage non linéaire pour la trajectographie passive. Dans les années qui suivirent, à l'initiative de François Le Gland, je me suis intéressé aux méthodes d'approximation aux différences finies proposées par Harold Kushner. Toujours avec François Le Gland nous avons travaillé sur différentes applications industrielles (US Army pour un contrat de veille scientifique, le CNES, la DCAN Toulon). Le renouveau de ce thème dû aux méthodes d'approximation particulière permet de relancer de nouvelles activités.

I.1 UN MODÈLE STATISTIQUE

On considère l'espace canonique $\mathcal{C}([0, T]; \mathbb{R}^n \times \mathbb{R}^d)$ muni de sa tribu borélienne, désigné par (Ω, \mathcal{F}) , et le processus canonique $\{X_t, Y_t\}_{0 \leq t \leq T}$. On note $\mathbb{P}_\theta^\varepsilon$ la mesure de probabilité définie sur (Ω, \mathcal{F}) telle que le processus canonique admette la représentation suivante :

$$dX_t = b_\theta(X_t) dt + \varepsilon \sigma(X_t) dW_t^\theta, \quad X_0 = \varepsilon \xi, \quad (\text{I.1})$$

$$dY_t = h_\theta(X_t) dt + \varepsilon dV_t^\theta, \quad Y_0 = 0 \quad (\text{I.2})$$

pour $0 \leq t \leq T$, où W_t^θ, V_t^θ sont des processus de Wiener standard, et $W_t^\theta, V_t^\theta, \xi$ sont mutuellement indépendants.

Les processus X_t, W_t^θ, Y_t sont respectivement à valeurs dans $\mathbb{R}^n, \mathbb{R}^m, \mathbb{R}^d$. Le paramètre inconnu θ appartient à $\Theta \subset \mathbb{R}^p$. On suppose que ξ est gaussien $N(0, I_n)$ (où I_n désigne la matrice identité de taille $n \times n$).

Dans (I.1)–(I.2), X_t désigne l'état d'un processus non observé et Y_t son observation, θ est un paramètre inconnu. Le but est soit d'estimer θ soit de vérifier une hypothèse portant sur θ à partir des observations $\mathcal{Y}_t \triangleq \sigma(Y_s; 0 \leq s \leq t)$. Il est possible d'associer à chaque valeur θ du paramètre, le filtre non linéaire $\pi_t^\theta(dx) = \mathbb{P}_\theta^\varepsilon(X_t \in dx | \mathcal{Y}_t)$ et de calculer ainsi la vraisemblance de θ au vu des observations. Le coût de mise en œuvre d'un filtre pour chaque valeur du paramètre rend cette approche inutilisable en pratique. De plus, on s'intéresse à des problèmes "temps réel" : les observations doivent être traitées au fur et mesure de leur disponibilité ; le calcul de π_t^θ et le problème statistique concernant θ doivent également être traités en ligne.

Un des exemples les plus intéressants est celui de la détection de changement : supposons que le système (I.1)–(I.2) fonctionne selon un régime nominal représenté par une valeur connue θ_0 du paramètre (θ_0 est connu ou a été préalablement estimé sur un échantillon d'observations provenant du système en régime nominal), à un instant de changement inconnu τ , le paramètre passera de la valeur θ_0 à la valeur $\theta_0 + \bar{\theta}$. Le problème est double : à chaque instant t il s'agit de tester s'il y a eu changement ($\tau < t$) ou non ($\tau \geq t$) et, dans le premier cas, d'estimer la valeur du saut $\bar{\theta}$.

Avant de considérer le problème de détection de changement, nous allons nous intéresser au modèle statistique sous-jacent. L'ouvrage de référence concernant la statistique de processus de diffusion partiellement observé, reste celui de Robert Liptser et Albert Shiryaev [97, 98] réédité en 2001. Concernant le filtrage non linéaire en plus de la précédente référence, on pourra consulter les références classiques JAZWINSKI [66], KALLIANPUR [70], PARDOUX [114].

FONCTION DE VRAISEMBLANCE

La fonction de vraisemblance correspondant au paramètre θ étant donné les observations Y_t pour $0 \leq t \leq T$ est :

$$L_{\theta_0}^\varepsilon(\theta) \triangleq \mathbb{E}_{\theta_0}^\varepsilon \left[\frac{d\mathbb{P}_\theta^\varepsilon}{d\mathbb{P}_{\theta_0}^\varepsilon} \middle| \mathcal{Y}_T \right] \quad (\text{I.3})$$

où θ_0 est n'importe quelle valeur fixée du paramètre.

Pour toute valeur de θ on suppose que les coefficients $x \rightarrow b_\theta(x)$, $x \rightarrow \sigma(x)$ et $x \rightarrow h_\theta(x)$ sont lipschitziens à croissance au plus linéaire.

On pose :

$$\Phi_t^{\varepsilon, \theta} \triangleq \exp \left(\frac{1}{\varepsilon^2} \int_0^t h_\theta(X_s)^* dY_s - \frac{1}{2\varepsilon^2} \int_0^t |h_\theta(X_s)|^2 ds \right), \quad (\text{I.4})$$

et on définit la mesure de probabilité $\mathbb{P}_\theta^{\varepsilon, \circ}$ sur (Ω, \mathcal{F}) par :

$$\frac{d\mathbb{P}_\theta^\varepsilon}{d\mathbb{P}_\theta^{\varepsilon, \circ}} = \Phi_T^{\varepsilon, \theta}. \quad (\text{I.5})$$

D'après le théorème de Girsanov, sous la probabilité $\mathbb{P}_\theta^{\varepsilon, \circ}$: W_t^θ et $\frac{1}{\varepsilon} Y_t$ sont des processus de Wiener standard, et W_t^θ , Y_t , ξ sont mutuellement indépendants.

Hypothèse I.1.1 *On suppose que $m \leq n$ et que la matrice $\sigma(x)$ est de rang plein (= m) pour tout x (i.e. $\sigma(x)^* \sigma(x)$ est inversible).*

On définit la matrice de diffusion $a(x) = \sigma(x) \sigma(x)^*$. Sous l'Hypothèse I.1.1 cette matrice peut être singulière.

L'hypothèse I.1.1 n'est pas restrictive. Dans tous les cas on peut supposer que $m \leq n$ et que $\sigma(x)$ est de rang plein (= m) pour tout x . En effet, si $m > n$ i.e. si dans l'équation d'état (I.1) la dimension du processus de Wiener est plus grande que celle de X_t on peut toujours se ramener par changement de base à un processus de Wiener de dimension au plus n . On suppose donc que $m \leq n$. Maintenant si $\sigma(x)$ n'est pas de rang plein, il suffit de réduire la dimension du processus de Wiener afin que $\sigma(x)$ soit de rang plein.

D'après l'Hypothèse I.1.1, la matrice pseudo-inverse $\sigma^+(x)$ de $\sigma(x)$ est définie par :

$$\sigma^+(x) \triangleq [\sigma(x)^* \sigma(x)]^{-1} \sigma(x)^*. \quad (\text{I.6})$$

(cf. LIPTSER & SHIRYAEV [97, section 13.1]). On a alors :

$$a^+(x) = [\sigma(x) \sigma(x)^*]^+ = \sigma^+(x)^* \sigma^+(x) = \sigma(x) [\sigma(x)^* \sigma(x)]^{-2} \sigma(x)^*. \quad (\text{I.7})$$

Hypothèse I.1.2 *Pour tout $x \in \mathbb{R}^n$ et $\theta, \theta' \in \Theta$:*

$$b_\theta(x) - b_{\theta'}(x) \in \text{image}(\sigma(x)).$$

On déduit de l'Hypothèse I.1.2 que $\sigma \sigma^+ [b_\theta - b_{\theta_0}] = \sigma \sigma^+ \sigma u$, pour un certain u , mais, par définition, la matrice pseudo-inverse vérifie $\sigma \sigma^+ \sigma = \sigma$, et donc :

$$\sigma(x) \sigma^+(x) [b_\theta(x) - b_{\theta_0}(x)] = b_\theta(x) - b_{\theta_0}(x)$$

ainsi

$$\begin{aligned} dX_t &= b_\theta(X_t) dt + \varepsilon \sigma(X_t) dW_t^\theta \\ &= b_{\theta_0}(X_t) dt + [b_\theta - b_{\theta_0}](X_t) dt + \varepsilon \sigma(X_t) dW_t^\theta \\ &= b_{\theta_0}(X_t) dt + [\sigma \sigma^+ (b_\theta - b_{\theta_0})](X_t) dt + \varepsilon \sigma(X_t) dW_t^\theta \\ &= b_{\theta_0}(X_t) dt + \varepsilon \sigma(X_t) \left(\frac{1}{\varepsilon} [\sigma^+ (b_\theta - b_{\theta_0})](X_t) dt + dW_t^\theta \right) \end{aligned}$$

ce qui conduit à l'identité en loi suivante entre W_t^θ et $W_t^{\theta_0}$:

$$dW_t^{\theta_0} = \frac{1}{\varepsilon} [\sigma^+ (b_\theta - b_{\theta_0})](X_t) dt + dW_t^\theta. \quad (\text{I.8})$$

et

$$\frac{d\mathbb{P}_\theta^{\varepsilon, \circ}}{d\mathbb{P}_{\theta_0}^{\varepsilon, \circ}} = \frac{\Psi_T^{\varepsilon, \theta}}{\Psi_T^{\varepsilon, \theta_0}}. \quad (\text{I.9})$$

où

$$\Psi_t^{\varepsilon, \theta} \triangleq \exp \left(\frac{1}{\varepsilon^2} \int_0^t [a^+ b_\theta]^*(X_s) dX_s - \frac{1}{2\varepsilon^2} \int_0^t |\sigma^+ b_\theta|^2(X_s) ds \right). \quad (\text{I.10})$$

Finalement, on note que :

$$\frac{d\mathbb{P}_\theta^\varepsilon}{d\mathbb{P}_{\theta_0}^\varepsilon} = \frac{d\mathbb{P}_\theta^\varepsilon}{d\mathbb{P}_\theta^{\varepsilon, \circ}} \frac{d\mathbb{P}_\theta^{\varepsilon, \circ}}{d\mathbb{P}_{\theta_0}^{\varepsilon, \circ}} \frac{d\mathbb{P}_{\theta_0}^{\varepsilon, \circ}}{d\mathbb{P}_{\theta_0}^\varepsilon} = \frac{\Psi_T^{\varepsilon, \theta}}{\Psi_T^{\varepsilon, \theta_0}} \frac{\Phi_T^{\varepsilon, \theta}}{\Phi_T^{\varepsilon, \theta_0}}.$$

Ainsi on a prouvé que sous l'Hypothèse I.1.2, les mesures de probabilité $\mathbb{P}_\theta^\varepsilon |_{\mathcal{Y}_T}$ sont mutuellement absolument continues. Pour tout $\theta_0 \in \Theta$ fixé, la fonction de vraisemblance (I.3) admet la représentation suivante :

$$L_{\theta_0}^\varepsilon(\theta) = \mathbb{E}_{\theta_0}^\varepsilon [\Lambda_{\theta_0, \theta}^\varepsilon(T) | \mathcal{Y}_T] \quad \text{avec} \quad \Lambda_{\theta_0, \theta}^\varepsilon(T) \triangleq \frac{\Psi_T^{\varepsilon, \theta}}{\Psi_T^{\varepsilon, \theta_0}} \frac{\Phi_T^{\varepsilon, \theta}}{\Phi_T^{\varepsilon, \theta_0}}, \quad (\text{I.11})$$

où $\Phi_t^{\varepsilon, \theta}$ est défini par (I.4) et $\Psi_t^{\varepsilon, \theta}$ par (I.10), i.e.

$$\begin{aligned} \Lambda_{\theta_0, \theta}^\varepsilon(t) &= \exp \left(\frac{1}{\varepsilon^2} \int_0^t [a^+ (b_\theta - b_{\theta_0})]^*(X_s) dX_s - \frac{1}{2\varepsilon^2} \int_0^t (|\sigma^+ b_\theta|^2 - |\sigma^+ b_{\theta_0}|^2)(X_s) ds \right) \\ &\quad \times \exp \left(\frac{1}{\varepsilon^2} \int_0^t [h_\theta - h_{\theta_0}]^*(X_s) dY_s - \frac{1}{2\varepsilon^2} \int_0^t [|h_\theta|^2 - |h_{\theta_0}|^2](X_s) ds \right). \end{aligned}$$

FONCTION DE VRAISEMBLANCE ET FILTRE NON LINÉAIRE

Une autre expression de la fonction de vraisemblance est :

$$\tilde{L}^\varepsilon(\theta) \triangleq \mathbb{E}_{\theta_0}^{\varepsilon, \circ} \left[\frac{d\mathbb{P}_\theta^\varepsilon}{d\mathbb{P}_{\theta_0}^{\varepsilon, \circ}} \Big| \mathcal{Y}_T \right]. \quad (\text{I.12})$$

Notons que cette quantité ne dépend pas du choix de θ_0 , on peut donc choisir $\theta_0 = \theta$, ce qui conduit à l'expression suivante :

$$\tilde{L}^\varepsilon(\theta) = \mathbb{E}_{\theta}^{\varepsilon, \circ}[\Phi_T^{\varepsilon, \theta} | \mathcal{Y}_T]. \quad (\text{I.13})$$

Le lien avec (I.3) est :

$$L_{\theta_0}^\varepsilon(\theta) = \frac{\tilde{L}^\varepsilon(\theta)}{\tilde{L}^\varepsilon(\theta_0)}.$$

L'inconvénient de la représentation (I.3) est que la dépendance en θ de $L_{\theta_0}^\varepsilon(\theta)$ est plus explicite. Son intérêt est qu'il s'agit exactement du coefficient de normalisation du filtre non linéaire, non normalisé (cf. §I.6, p. 18).

I.2 NORMALITÉ ASYMPTOTIQUE LOCALE

Approcher un modèle statistique par un modèle statistique gaussien est une idée très classique (voir par exemple WALD [138]). En 1960, Lucien Le Cam [89] propose une telle approximation de façon "locale" ; il introduit la terminologie "normalité asymptotique locale". Il a développé cette idée (et d'autres) dans de nombreux travaux [90, 91, 92, 93, 94]. Aad Van der Vaart [135] a proposé en 2002 un survol intéressant des travaux en statistique de Lucien Le Cam (disparu en 2000).

Il existe des ouvrages aboutis comme ROUSSAS [120], IBRAGIMOV & HAS'MINSKII [62], MILLAR [105], BASAWA & SCOTT [7], LE CAM [93], VAN DER VAART [134], ainsi que le cours à l'école d'été de St-Flour DACUNHA-CASTELLE [39], et le très bon cours de troisième cycle dû à Valentine Genon-Catalot et Dominique Picard [49]. On pourra également consulter avec profit les premiers travaux publiés dans ce domaine parmi lesquels HÁJEK [56] et CHIBISOV [37]. Concernant les EDS les travaux de Yurii Kutoyants sont incontournables [85, 86, 87] ; on pourra aussi consulter ses articles précédemment publiés [82, 83, 84].

On établit maintenant la normalité asymptotique locale de la suite de modèles statistiques $\{\mathbb{P}_\theta^\varepsilon; \theta \in \Theta\}$ en tout point θ_0 lorsque $\varepsilon \downarrow 0$. On se fixe $\theta_0 \in \Theta$ et on pose :

$$\mathcal{Z}_{\theta_0}^\varepsilon(u) = L_{\theta_0}^\varepsilon(\theta_0 + \varepsilon u) = \mathbb{E}_{\theta_0}^\varepsilon \left[\frac{d\mathbb{P}_{\theta_0 + \varepsilon u}^\varepsilon}{d\mathbb{P}_{\theta_0}^\varepsilon} \middle| \mathcal{Y}_T \right]$$

ainsi :

$$\mathcal{Z}_{\theta_0}^\varepsilon(u) = \mathbb{E}_{\theta_0}^\varepsilon[\bar{\Lambda}_T^\varepsilon | \mathcal{Y}_T] \quad \text{avec} \quad \bar{\Lambda}_T^\varepsilon \triangleq \Lambda_{\theta_0, \theta_0 + \varepsilon u}^\varepsilon(T).$$

L'idée est donc développer la fonction $u \rightarrow \log \mathcal{Z}_{\theta_0}^\varepsilon(u)$ en $u = 0$, c'est-à-dire :

$$\log \mathcal{Z}_{\theta_0}^\varepsilon(u) \simeq u^* \Delta_{\theta_0}^\varepsilon - \frac{1}{2} u^* I^\varepsilon(\theta_0) u + \dots$$

(noter que $\log \mathcal{Z}_{\theta_0}^\varepsilon(0) = 0$). Le premier terme de ce développement fait apparaître le terme $\Delta_{\theta_0}^\varepsilon = (\partial_u \mathcal{Z}_{\theta_0}^\varepsilon(u)|_{u=0})^*$, i.e.

$$\Delta_{\theta_0}^\varepsilon \triangleq \left(\varepsilon \partial_\theta \log L_{\theta_0}^\varepsilon(\theta) \Big|_{\theta=\theta_0} \right)^* \quad (\text{I.14})$$

qui joue un rôle central sur le plan théorique comme pratique. Nous l'appellerons le "résidu". On détermine dans un premier temps une expression pour ce résidu.

LE RÉSIDU

On montre sans difficulté que $\Delta_{\theta_0}^\varepsilon$ admet la représentation suivante (cf. [32]) :

$$\Delta_{\theta_0}^\varepsilon = \mathbb{E}_{\theta_0, \varepsilon}(\Xi_{\theta_0}^\varepsilon(T) | \mathcal{Y}_T) \quad (\text{I.15})$$

où

$$\begin{aligned} \Xi_{\theta_0}^\varepsilon(t) &= \varepsilon (\partial_\theta \log \Psi_t^{\varepsilon, \theta})^* + \varepsilon (\partial_\theta \log \Phi_t^{\varepsilon, \theta})^* \\ &= \int_0^t [\sigma^+ \dot{b}_\theta]^*(X_s) \frac{1}{\varepsilon} (dX_s - b_\theta(X_s) ds) + \int_0^t \dot{h}_\theta(X_s)^* \frac{1}{\varepsilon} (dY_s - h_\theta(X_s) ds) \\ &= \int_0^t [\sigma^+ \dot{b}_\theta]^*(X_s) dW_s^\theta + \int_0^t \dot{h}_\theta(X_s)^* dV_s^\theta. \end{aligned} \quad (\text{I.16})$$

On adopte ici la notation suivante : Pour une fonction $(\theta, x) \rightarrow f_\theta(x)$, on pose

$$f'_\theta(x) = \partial_x f_\theta(x) \triangleq \left[\frac{\partial f_{\theta, i}(x)}{\partial x_j} \right]_{\substack{1 \leq i \leq k \\ 1 \leq j \leq n}}, \quad \dot{f}_\theta(x) = \partial_\theta f_\theta(x) \triangleq \left[\frac{\partial f_{\theta, i}(x)}{\partial \theta_j} \right]_{\substack{1 \leq i \leq k \\ 1 \leq j \leq p}}$$

(i indice de ligne, j indice de colonne) avec $f_\theta(x) = [f_{\theta, 1}(x) \cdots f_{\theta, k}(x)]^*$.

SYSTÈME RENORMALISÉ

On pose

$$\bar{X}_t^\varepsilon \triangleq \frac{1}{\varepsilon} (X_t - x_t), \quad \bar{Y}_t^\varepsilon \triangleq \frac{1}{\varepsilon} (Y_t - y_t), \quad \bar{\Xi}_t^\varepsilon \triangleq \Xi_{\theta_0}^\varepsilon(t), \quad (\text{I.17})$$

où

$$\frac{d}{dt} x_t = b_{\theta_0}(x_t), \quad x_0 = 0, \quad \frac{d}{dt} y_t = h_{\theta_0}(x_t), \quad y_0 = 0, \quad (\text{I.18})$$

et $\Xi_{\theta_0}^\varepsilon(t)$ est défini par (I.16). Notons que $\sigma(\bar{Y}_s^\varepsilon; s \leq t) = \sigma(Y_s^\varepsilon; s \leq t)$.

On s'intéresse au comportement asymptotique des processus $\bar{X}_t^\varepsilon, \bar{Y}_t^\varepsilon, \bar{\Xi}_t^\varepsilon, \bar{\Lambda}_t^\varepsilon$ lorsque $\varepsilon \downarrow 0$. D'après (I.1)–(I.2), (I.16), (I.17), (I.18) :

$$d\bar{X}_t^\varepsilon = \frac{1}{\varepsilon} [b_{\theta_0}(x_t + \varepsilon \bar{X}_t^\varepsilon) - b_{\theta_0}(x_t)] dt + \sigma(x_t + \varepsilon \bar{X}_t^\varepsilon) dW_t^{\theta_0}, \quad \bar{X}_0^\varepsilon = \xi, \quad (\text{I.19})$$

$$d\bar{Y}_t^\varepsilon = \frac{1}{\varepsilon} [h_{\theta_0}(x_t + \varepsilon \bar{X}_t^\varepsilon) - h_{\theta_0}(x_t)] dt + dV_t^{\theta_0}, \quad \bar{Y}_0^\varepsilon = 0, \quad (\text{I.20})$$

$$d\bar{\Xi}_t^\varepsilon = [\sigma^+ \dot{b}_{\theta_0}]^*(x_t + \varepsilon \bar{X}_t^\varepsilon) dW_t^{\theta_0} + \dot{h}_{\theta_0}^*(x_t + \varepsilon \bar{X}_t^\varepsilon) dV_t^{\theta_0}, \quad \bar{\Xi}_0^\varepsilon = 0, \quad (\text{I.21})$$

$$\begin{aligned} d\bar{\Lambda}_t^\varepsilon &= \bar{\Lambda}_t^\varepsilon \frac{1}{\varepsilon} [\sigma^+ (b_{\theta_0 + \varepsilon u} - b_{\theta_0})]^*(x_t + \varepsilon \bar{X}_t^\varepsilon) dW_t^{\theta_0} \\ &\quad + \bar{\Lambda}_t^\varepsilon \frac{1}{\varepsilon} [h_{\theta_0 + \varepsilon u} - h_{\theta_0}]^*(x_t + \varepsilon \bar{X}_t^\varepsilon) dV_t^{\theta_0}, \quad \bar{\Lambda}_0^\varepsilon = 1. \end{aligned} \quad (\text{I.22})$$

SYSTÈME LINÉAIRE TANGENT LIMITE

Il est clair que le système (I.19)–(I.22) converge vers le système limite suivant :

$$d\bar{X}_t^0 = F_t \bar{X}_t^0 dt + G_t dW_t^{\theta_0}, \quad \bar{X}_0^0 = \xi, \quad (\text{I.23})$$

$$d\bar{Y}_t^0 = H_t \bar{X}_t^0 dt + dV_t^{\theta_0}, \quad \bar{Y}_0^0 = 0, \quad (\text{I.24})$$

$$d\bar{\Xi}_t^0 = J_t^* dW_t^{\theta_0} + K_t^* dV_t^{\theta_0}, \quad \bar{\Xi}_0^0 = 0, \quad (\text{I.25})$$

$$d\bar{\Lambda}_t^0 = u^* \bar{\Lambda}_t^0 (J_t^* dW_t^{\theta_0} + K_t^* dV_t^{\theta_0}), \quad \bar{\Lambda}_0^0 = 1, \quad (\text{I.26})$$

avec $F_t \triangleq b'_{\theta_0}(x_t)$, $G_t \triangleq \sigma(x_t)$, $H_t \triangleq h'_{\theta_0}(x_t)$, $J_t \triangleq [\sigma^+ \dot{b}_{\theta_0}](x_t)$, $K_t \triangleq \dot{h}_{\theta_0}(x_t)$.

Le système (I.23)–(I.25) est linéaire–gaussien avec processus d'état $(\bar{X}_t^0, \bar{\Xi}_t^0)$ et processus d'observation \bar{Y}_t^0 . La solution de (I.26) est une martingale exponentielle qui s'exprime explicitement en fonction des mouvements browniens $W_t^{\theta_0}$ et $V_t^{\theta_0}$. Posons $\bar{\mathcal{Y}}_t^0 = \sigma(\bar{Y}_s^0; s \leq t)$, ainsi :

$$\text{loi} \left(\begin{bmatrix} \bar{X}_t^0 \\ \bar{\Xi}_t^0 \end{bmatrix} \middle| \bar{\mathcal{Y}}_t^0 \right) = N \left(\begin{bmatrix} \hat{X}_t^0 \\ \hat{\Xi}_t^0 \end{bmatrix}, \begin{bmatrix} R_t & r_t \\ r_t^* & Q_t \end{bmatrix} \right)$$

dont la moyenne et la covariance sont les sorties de ce filtre de Kalman–Bucy :

$$d\hat{X}_t^0 = F_t \hat{X}_t^0 dt + R_t H_t^* [d\bar{Y}_t^0 - H_t \hat{X}_t^0 dt], \quad \hat{X}_0^0 = \mathbb{E}(\xi), \quad (\text{I.27})$$

$$d\hat{\Xi}_t^0 = (K_t + H_t r_t)^* [d\bar{Y}_t^0 - H_t \hat{X}_t^0 dt], \quad \hat{\Xi}_0^0 = 0, \quad (\text{I.28})$$

$$\frac{d}{dt} R_t = F_t R_t + R_t F_t^* + G_t G_t^* - R_t H_t^* H_t R_t, \quad R_0 = \text{cov}(\xi), \quad (\text{I.29})$$

$$\frac{d}{dt} Q_t = -K_t^* H_t r_t - r_t H_t^* K_t + J_t^* J_t - r_t H_t^* H_t r_t, \quad Q_0 = 0, \quad (\text{I.30})$$

$$\frac{d}{dt} r_t = (F_t - R_t H_t^* H_t) r_t - R_t H_t^* K_t + G_t J_t, \quad r_0 = 0. \quad (\text{I.31})$$

D'après (I.26) :

$$\bar{\Lambda}_T^0 = \exp \left(u^* \bar{\Xi}_T^0 - \frac{1}{2} u^* \int_0^T (J_t^* J_t + K_t^* K_t) dt u \right)$$

et $\text{loi}(\bar{\Xi}_T^0 | \bar{\mathcal{Y}}_T^0) = N(\hat{\Xi}_T^0, Q_T)$, et d'après (I.28) et (I.30) :

$$\hat{\Xi}_T^0 = \int_0^T (K_t + H_t r_t)^* (d\bar{Y}_t^0 - H_t \hat{X}_t^0 dt)$$

$$Q_T = \int_0^T \left(-K_t^* H_t r_t - r_t H_t^* K_t + J_t^* J_t - r_t H_t^* H_t r_t \right) dt$$

On calcule alors :

$$\mathcal{Z}_{\theta_0}(u) \triangleq \mathbb{E}_{\theta_0}[\bar{\Lambda}_T^0 | \bar{\mathcal{Y}}_T^0] = \exp(u^* \Delta_{\theta_0} - \frac{1}{2} u^* I(\theta_0) u) \quad (\text{I.32})$$

avec

$$\Delta_{\theta_0} \triangleq \int_0^T (K_t + H_t r_t)^* (d\bar{Y}_t^0 - H_t \hat{X}_t^0 dt), \quad (\text{I.33})$$

$$I(\theta_0) \triangleq \int_0^T (K_t + H_t r_t)^* (K_t + H_t r_t) dt. \quad (\text{I.34})$$

CONVERGENCE

Maintenant que la limite a été identifiée, il faut démontrer la convergence. On montre de façon classique que :

$$\mathbb{P}_{\theta_0}^\varepsilon - \lim_{\varepsilon \downarrow 0} \sup_{0 \leq t \leq T} \left\{ |\bar{X}_t^\varepsilon - \bar{X}_t^0| + |\bar{Y}_t^\varepsilon - \bar{Y}_t^0| + |\bar{\Xi}_t^\varepsilon - \bar{\Xi}_t^0| + |\bar{\Lambda}_t^\varepsilon - \bar{\Lambda}_t^0| \right\} = 0. \quad (\text{I.35})$$

Cette propriété est insuffisante pour assurer la convergence des espérances conditionnelles suivantes :

$$\mathbb{E}_{\theta_0, \varepsilon}(\bar{\Xi}_T^\varepsilon | \bar{\mathcal{Y}}_T^\varepsilon) \xrightarrow{\varepsilon \downarrow 0} \mathbb{E}_{\theta_0}(\bar{\Xi}_T^0 | \bar{\mathcal{Y}}_T^0), \quad \mathbb{E}_{\theta_0, \varepsilon}(\bar{\Lambda}_T^\varepsilon | \bar{\mathcal{Y}}_T^\varepsilon) \xrightarrow{\varepsilon \downarrow 0} \mathbb{E}_{\theta_0}(\bar{\Lambda}_T^0 | \bar{\mathcal{Y}}_T^0). \quad (\text{I.36})$$

D'après KLEPTSINA *et al* [72], ces dernières propriétés sont valides si, en plus de (I.35), on montre :

$$\|\bar{\mathbb{Q}}^{0, \varepsilon} - \bar{\mathbb{Q}}^0\|_{\text{VT}} \xrightarrow{\varepsilon \downarrow 0} 0$$

(norme en variation totale) où $\bar{\mathbb{Q}}^{0,\varepsilon} \triangleq \text{loi}(\bar{X}_t^0, \bar{\Xi}_t^0, \bar{Y}_t^\varepsilon; 0 \leq t \leq T)$ et $\bar{\mathbb{Q}}^0 \triangleq \text{loi}(\bar{X}_t^0, \bar{\Xi}_t^0, \bar{Y}_t^0; 0 \leq t \leq T)$. Ce qui se montre ainsi :

$$\|\bar{\mathbb{Q}}^{0,\varepsilon} - \bar{\mathbb{Q}}^0\|_{\text{TV}} = \mathbb{E}_{\theta_0}^\varepsilon \left| 1 - \frac{d\bar{\mathbb{Q}}^0}{d\bar{\mathbb{Q}}^{0,\varepsilon}} \right| = \mathbb{E}_{\theta_0}^\varepsilon |\mathbb{E}_{\theta_0}^\varepsilon [1 - \Sigma_T^\varepsilon | \mathcal{G}]|$$

où $\mathcal{G} = \sigma(\bar{X}_t^0, \bar{\Xi}_t^0, \bar{Y}_t^\varepsilon; 0 \leq t \leq T)$ et $\Sigma_t^\varepsilon = \exp(M_t^\varepsilon - \frac{1}{2} \langle M^\varepsilon \rangle_t)$ où, d'après (I.20) et (I.24) :

$$M_t^\varepsilon = \int_0^t \left\{ h'_{\theta_0}(x_s) \bar{X}_s^0 - \frac{1}{\varepsilon} [h_{\theta_0}(x_s + \varepsilon \bar{X}_s^\varepsilon) - h_{\theta_0}(x_s)] \right\}^* dV_s^{\theta_0}.$$

Il suffit alors de noter que $\mathbb{P}_{\theta_0}^\varepsilon - \lim \Sigma_T^\varepsilon \rightarrow 0$ lorsque $\varepsilon \downarrow 0$. L'idée de cette démonstration est de LE GLAND & WANG [95].

La seconde convergence de (I.36) est en fait le résultat cherché : pour tout $u \in \mathbb{R}^p$

$$\mathcal{Z}_{\theta_0}^\varepsilon(u) \xrightarrow[\varepsilon \downarrow 0]{\mathbb{P}_{\theta_0}^\varepsilon} \mathcal{Z}_{\theta_0}(u), \quad \forall u \in \mathbb{R}^p.$$

I.3 CONVERGENCE DE LA SUITE DES PROCESSUS DU RAPPORT DE VRAISEMBLANCE

Le résultat de normalité asymptotique locale de la section précédente, n'est que la convergence de la suite des marginales finies des lois des processus du rapport de vraisemblance. Des conditions suffisantes bien connues pour la convergence de la loi du processus de rapport de vraisemblance dans $\mathcal{C}_0(\mathbb{R}^p, \mathbb{R})$, sont les conditions d'Ibragimov-Has'minskii [62, §III-1] (cf. GENON-CATALOT & PICARD [49, §5.3] pour cette version) :

(i) Pour tout u

$$\mathcal{Z}_{\theta_0}^\varepsilon(u) = \exp \left\{ u^* \Delta_{\theta_0}^\varepsilon - \frac{1}{2} u^* I(\theta_0) u + O_{\theta_0}^\varepsilon(u) \right\}$$

où $I(\theta_0)$ est une matrice symétrique définie positive et

$$\text{loi}\{\Delta_{\theta_0}^\varepsilon | \mathbb{P}_{\theta_0}^\varepsilon\} \xrightarrow[\varepsilon \downarrow 0]{} N(0, I(\theta_0)), \quad \mathbb{P}_{\theta_0}^\varepsilon - \lim_{\varepsilon \downarrow 0} O_{\theta_0}^\varepsilon(u) = 0. \quad (\text{I.37})$$

(ii) Il existe des constantes $q > 0, B > 0, \beta > n, m > 0, a > 0$ telles que pour tout $R > 0$ et $\varepsilon > 0$:

$$\sup_{\substack{|u| < R \\ |v| < R \\ u, v \in U_{\theta_0}^\varepsilon}} |u - v|^{-\beta} \mathbb{E}_{\theta_0}^\varepsilon \left\{ |(Z_{\theta_0}^\varepsilon(u))^q - (Z_{\theta_0}^\varepsilon(v))^q|^m \right\} \leq B(1 + R^a) \quad (\text{I.38})$$

$$\sup_{u \in U_{\theta_0}^\varepsilon} |u|^R \mathbb{E}_{\theta_0}^\varepsilon \left\{ (Z_{\theta_0}^\varepsilon(u))^q \right\} < +\infty \quad (\text{I.39})$$

où $U_{\theta_0}^\varepsilon = \frac{1}{\varepsilon} (\Theta - \theta_0)$.

D'après IBRAGIMOV & HAS'MINSKII [62, Th. III-1.1 & III-1.3], ces conditions impliquent que l'estimateur du maximum de vraisemblance :

$$\hat{\theta}_\varepsilon \in \text{Arg max}_{\theta \in \Theta} L_{\theta_0}^\varepsilon(\theta)$$

vérifie, pour tout $K \subset \Theta$ compact, uniformément en $\theta \in K$:

(i) consistance faible :

$$\mathbb{P}_\theta^\varepsilon\text{-}\lim_{\varepsilon \downarrow 0} \hat{\theta}_\varepsilon = \theta, \quad (\text{I.40})$$

(ii) normalité asymptotique

$$\text{loi}\left\{\frac{1}{\varepsilon}(\hat{\theta}_\varepsilon - \theta) \middle| \mathbb{P}_\theta^\varepsilon\right\} \xrightarrow{\varepsilon \downarrow 0} N(0, I(\theta)^{-1/2}), \quad (\text{I.41})$$

(iii) convergence des moments

$$\lim_{\varepsilon \downarrow 0} \mathbb{E}_\theta^\varepsilon \left| \frac{1}{\varepsilon}(\hat{\theta}_\varepsilon - \theta) \right|^p = \mathbb{E}|\xi|^p \quad (\text{I.42})$$

pour tout $p \in \mathbb{N}$ où $\xi \sim N(0, I(\theta)^{-1/2})$,

(iv) efficacité asymptotique dans K

$$\lim_{\varepsilon \downarrow 0} \left\{ \inf_{\tilde{\theta}_\varepsilon} \sup_{\theta \in \mathcal{T}} \mathbb{E}_\theta^\varepsilon w_\varepsilon(\tilde{\theta}_\varepsilon - \theta) - \sup_{\theta \in \mathcal{T}} \mathbb{E}_\theta^\varepsilon w_\varepsilon(\hat{\theta}_\varepsilon - \theta) \right\} = 0 \quad (\text{I.43})$$

pour toute partie \mathcal{T} non vide de K , où $w_\varepsilon(\theta) = |\frac{1}{\varepsilon} I(\theta_0)^{-1/2} \theta|^2$ pour tout $\theta_0 \in K$.

I.4 DÉTECTION DE RUPTURE

Ce domaine est vaste. On se contentera donc de citer les articles de revue : WILLSKY [140], ZACKS [142], ISERMANN [63], BASSEVILLE [8], GERTLER [50], LAI [88]. On mentionnera également quelques monographies : PAU [115], GHOSH & SEN [51], BASSEVILLE & NIKIFOROV [14], BRODSKY & DARKHOVSKY [22], BASSEVILLE [10], des ouvrages collectifs comme BASSEVILLE & BENVENISTE [11], et d'autres ouvrages plus appliqués comme HIMMELBLAU [57]. On n'évoquera pas ici les nombreux domaines d'application.

Parmi les très nombreuses méthodes, on signalera en premier les méthodes paramétriques non bayésiennes comprenant deux techniques relativement proches : les méthodes CUSUM ("cumulative sums") et les méthodes GLRT ("generalized likelihood ratio test"). Introduites dans les années 50 par PAGE [112, 113], ces techniques ont été largement étudiées et développées : LORDEN [102] (qui montre que les procédures CUSUM de Page sont asymptotiquement minimax), HINKLEY & HINKLEY [60], HINKLEY [58, 59], BENVENISTE & BASSEVILLE [16], BASSEVILLE *et al* [13, 12], BENVENISTE *et al* [17], POLLAK & SIEGMUND [117], WILLSKY & JONES [139], NIKIFOROV [109, 110], MOUSTAKIDES [106, 107, 108], BASSEVILLE & NIKIFOROV [14], SHIRYAEV [125].

D'autres travaux traitent des aspects théoriques du sujet CHERNOFF & ZACKS [36], KANDER & ZACKS [71] (échantillon de taille fixée) SEN & SRIVASTAVA [121], KLIGENE [73, 74, 75], TELKSNYS [131, 130], KLIGENE & TELKSNYS [76], TARTAKOVSKII [128, 129].

Les liens avec le filtrage non linéaire ont été abordés par DAVIS [41], LOPARO & ROTH [100, 101] et LOPARO *et al* [99].

Pour les approches bayésiennes, on citera également les méthodes séquentielles qui remontent aux premiers travaux d'Albert Shiryaev [122, 123, 124]; voir également la monographie [127]. Cette approche a été récemment reprise par Michel Vellekoop, voir VELLEKOOP [136] et VELLEKOOP & CLARK [137].

On peut également évoquer les méthodes non paramétriques, voir BHATTACHARYA & JOHNSON [19], DARKHOVSKY [40] PETTITT [116], ZACKS [143], BRODSKY & DARKHOVSKY [22], et les méthodes par chaînes de Markov (“switching environments”), voir TUGNAIT [133, 132] et BLOM [21].

Ce domaine possède même une préhistoire, notamment dans le domaine de contrôle de qualité de processus (on n’en donnera aucune référence!).

I.4.1 UN CAS LINÉAIRE

Dans CAMPILLO [25] on a traité le problème de détection de changement dans un système linéaire gaussien.

À un instant de changement inconnu τ , le terme constant de dérive dans l’équation d’état passe de la valeur b à la valeur $b + \gamma$, où b est connu et la valeur du saut γ est inconnue, i.e.

$$dX_t = [B X_t + b + \mathbf{1}_{(\tau \leq t)} \gamma] dt + \varepsilon C dW_t, \quad X_0 = x_0, \quad (\text{I.44})$$

$$dY_t = H X_t dt + \varepsilon dV_t, \quad Y_0 = 0, \quad (\text{I.45})$$

où $0 \leq t \leq T$, W et V sont des processus de Wiener standard indépendants. Le paramètre inconnu est donc $\theta \triangleq (\tau, \gamma)$. À l’instant T on souhaite tester l’hypothèse \mathcal{H}_0 de non rupture contre l’hypothèse \mathcal{H}_1 de rupture :

$$\begin{cases} \mathcal{H}_0 : & \text{pas de rupture avant } T, \\ \mathcal{H}_1 : & \text{rupture avant } T. \end{cases} \quad (\text{I.46})$$

L’hypothèse \mathcal{H}_1 est multiple : $\mathcal{H}_1 = \{\mathcal{H}_1(\theta); \theta \triangleq (\tau, \gamma) \text{ t.q. } \tau < T, \gamma \neq 0\}$ où $\mathcal{H}_1(\theta)$ est l’hypothèse de rupture à l’instant τ de saut γ .

FONCTION DE VRAISEMBLANCE

On considère la fonction de vraisemblance $\tilde{L}^\varepsilon(\theta) = \mathbb{E}_\theta^{\varepsilon, \circ}[\Phi_T^{\varepsilon, \theta} | \mathcal{Y}_T]$ définie en (I.12)–(I.13), avec :

$$\Phi_T^{\varepsilon, \theta} = \exp\left(\frac{1}{\varepsilon} \int_0^T (H X_s)^* \frac{1}{\varepsilon} (dY_s - H X_s ds) + \frac{1}{2\varepsilon^2} \int_0^T |H X_s|^2 ds\right).$$

On obtient une expression explicite en utilisant le filtre de Kalman–Bucy. Dans le cas \mathcal{H}_0 , le filtre de Kalman–Bucy $\text{loi}(X_t | \mathcal{Y}_t) = N(\hat{X}_t, \varepsilon^2 R_t)$ est :

$$d\hat{X}_t = [B \hat{X}_t + b] dt + R_t H^* (dY_t - H \hat{X}_t dt), \quad \hat{X}_0 = x_0, \quad (\text{I.47})$$

$$\dot{R}_t = A R_t + R_t A^* + C C^* - R_t H^* H R_t^*, \quad R_0 = 0. \quad (\text{I.48})$$

On désignera par $N(\hat{X}_t^\theta, \varepsilon^2 R_t)$ le filtre de Kalman Bucy sous l’hypothèse $\mathcal{H}_1(\theta)$ (la covariance R_t ne dépend pas de θ). On introduit le processus innovation :

$$\varepsilon d\nu_t \triangleq dY_t - H \hat{X}_t dt \quad (\text{resp. } \varepsilon d\nu_t^\theta \triangleq dY_t - H \hat{X}_t^\theta dt). \quad (\text{I.49})$$

Il s’agit d’un processus de Wiener standard sous l’hypothèse \mathcal{H}_0 (resp. $\mathcal{H}_1(\theta)$).

On montre alors que :

$$\tilde{L}^\varepsilon(\theta) = \exp\left(\frac{1}{\varepsilon} \int_0^t (H \hat{X}_s^\theta)^* d\nu_s^\theta + \frac{1}{2\varepsilon^2} \int_0^t |H \hat{X}_s^\theta|^2 ds\right)$$

ainsi :

$$L^\varepsilon(\theta) \triangleq \frac{\tilde{L}^\varepsilon(\theta)}{\tilde{L}^\varepsilon(0)} = \frac{\exp\left(\frac{1}{\varepsilon} \int_0^T (H \hat{X}_t^\theta)^* d\nu_t^\theta - \frac{1}{2\varepsilon^2} \int_0^T |H \hat{X}_t^\theta|^2 dt\right)}{\exp\left(\frac{1}{\varepsilon} \int_0^T (H \hat{X}_t)^* d\nu_t - \frac{1}{2\varepsilon^2} \int_0^T |H \hat{X}_t|^2 dt\right)}$$

Notons que $\varepsilon d(\nu_t - \nu_t^\theta) = H(\hat{X}_t^\theta - \hat{X}_t) dt$, ainsi :

$$\begin{aligned} d(\hat{X}_t^\theta - \hat{X}_t) &= B(\hat{X}_t^\theta - \hat{X}_t) dt + \mathbf{1}_{(\tau \leq t)} \gamma dt - R_t H^* \varepsilon d(\nu_t - \nu_t^\theta) \\ &= B(\hat{X}_t^\theta - \hat{X}_t) dt + \mathbf{1}_{(\tau \leq t)} \gamma dt - R_t H^* H(\hat{X}_t^\theta - \hat{X}_t) dt \\ &= (B - R_t H^* H)(\hat{X}_t^\theta - \hat{X}_t) dt + \mathbf{1}_{(\tau \leq t)} \gamma dt. \end{aligned}$$

Par conséquent :

$$H(\hat{X}_t^\theta - \hat{X}_t) = S(\theta, t) = S_t^\tau \gamma \quad \text{où} \quad S_t^\tau \triangleq H \int_{t \wedge \tau}^t \Gamma(t, s) ds \quad (\text{I.50})$$

et $\Gamma(t, s)$ est la résolvante :

$$\frac{\partial}{\partial t} \Gamma(t, s) = (B - R_t H^* H) \Gamma(t, s), \quad \Gamma(s, s) = I$$

Finalement on obtient :

$$L^\varepsilon(\theta) = \exp\left(\frac{1}{\varepsilon} \int_0^T S(\theta, t)^* d\nu_t - \frac{1}{2\varepsilon^2} \int_0^T |S(\theta, t)|^2 dt\right) \quad (\text{I.51})$$

Le lien avec la fonction de vraisemblance $L_{\theta_0}^\varepsilon(\theta)$ définie en (I.11) dans le cas général de la Section I.1 est simplement $L_{\theta_0}^\varepsilon(\theta) = L^\varepsilon(\theta)/L^\varepsilon(\theta_0)$.

Le fait que dans cet exemple, le problème se ramène à un seul filtre (sous l'hypothèse \mathcal{H}_0) et au calcul d'une expression déterministe $S(\theta, t) = S_t^\tau \gamma$ est exactement le résultat obtenu par WILLISKY & JONES [139, 140] et tient à la nature additive du saut.

NORMALITÉ ASYMPTOTIQUE LOCALE

Sur le plan statistique, ce problème est en fait de la forme :

$$d\nu_t = S(\theta, t) dt + \varepsilon d\beta_t^\theta$$

où ν_t est le processus innovation obtenu en sortie du filtre de Kalman–Bucy (I.47)–(I.48) sous l'hypothèse \mathcal{H}_0 , $S(\theta, t)$ est la fonction déterministe définie par (I.50), enfin β_t^θ est un processus de Wiener standard sous $\mathcal{H}_1(\theta)$ (et β_t^0 est un processus de Wiener standard sous \mathcal{H}_0). L'étude asymptotique locale de ce problème a été traitée dans IBRAGIMOV & HAS'MINSKII [62, §II.7 et III.5] et détaillée dans cette situation dans CAMPILLO [25].

Il faut noter que ce problème statistique bien qu'étant un problème de détection de rupture est régulier (i.e. $S(\theta, t)$ est régulier en θ). Ceci est dû au filtre qui "lisse" le problème.

On considère la renormalisation locale du rapport de vraisemblance :

$$Z_{\varepsilon, \theta}(U) \triangleq L_{\theta}^{\varepsilon}(\theta + \varepsilon U) \quad \text{avec} \quad U = \begin{pmatrix} u \\ v \end{pmatrix}.$$

On montre que la famille de modèles statistiques $\{P_{\theta}^{\varepsilon}; \theta \in \Theta\}_{\varepsilon > 0}$ est *localement asymptotiquement normale*, i.e. en tout $\theta \in \Theta$ on a :

$$Z_{\varepsilon, \theta}(U) = \exp \left(U^* \Delta_{\theta}^{\varepsilon} - \frac{1}{2} U^* I(\theta) U + O_{\varepsilon, \theta}(U) \right)$$

où :

- (i) $\text{loi}\{\Delta_{\theta}^{\varepsilon} | P_{\theta}^{\varepsilon}\} = N(0, I(\theta))$,
- (ii) $P_{\theta}^{\varepsilon} - \lim_{\varepsilon \downarrow 0} O_{\varepsilon, \theta}(U_{\varepsilon}) = 0$ pour toute suite bornée U_{ε} .

On calcule explicitement $\Delta_{\theta}^{\varepsilon}$ et $I(\theta)$:

$$\Delta_{\theta}^{\varepsilon} \triangleq \int_0^T \dot{S}(\theta, t)^* [d\nu_t - \frac{1}{\varepsilon} S(\theta, t) dt], \quad I(\theta) \triangleq \int_0^T \dot{S}(\theta, t)^* \dot{S}(\theta, t) dt$$

où

$$\dot{S}(\theta, t) = \left[\partial_{\tau} S(\theta, t) ; \partial_{\gamma} S(\theta, t) \right] = \left[\dot{G}_t^{\tau} \gamma ; G_t^{\tau} \right].$$

et $\dot{S}(\theta, t) \triangleq \partial_{\theta} S(\theta, t)$ et $\dot{G}_t^{\tau} \triangleq \partial_{\tau} G_t^{\tau}$.

ESTIMATEURS ET TESTS DU MAXIMUM DE VRAISEMBLANCE

L'optimisation de la fonction de vraisemblance (I.51) en γ est explicite, ce qui conduit aux estimateurs du maximum de vraisemblance (EMV) suivants :

$$\hat{\tau}^{\varepsilon} \in \text{Arg} \max_{0 \leq \tau < T} \left[\left(\int_0^T S_t^{\tau} d\nu_t \right)^* \left(\int_0^T [S_t^{\tau}]^* S_t^{\tau} dt \right)^{-1} \left(\int_0^T S_t^{\tau} d\nu_t \right) \right], \quad (\text{I.52})$$

$$\hat{\gamma}^{\varepsilon} = \varepsilon \left(\int_0^T [S_t^{\tau}]^* S_t^{\tau} dt \right)^{-1} \left(\int_0^T S_t^{\tau} d\nu_t \right) \Big|_{\tau = \hat{\tau}^{\varepsilon}} \quad (\text{I.53})$$

Le test du rapport de vraisemblance de \mathcal{H}_0 contre \mathcal{H}_1 est :

$$\begin{cases} \text{si } \hat{\ell}^{\varepsilon} \leq c \text{ on accepte } \mathcal{H}_0 \\ \text{si } \hat{\ell}^{\varepsilon} > c \text{ on accepte } \mathcal{H}_1 \end{cases} \quad (\text{I.54})$$

où

$$\hat{\ell}^{\varepsilon} \triangleq \log L^{\varepsilon}(\hat{\theta}^{\varepsilon}) = \left(\int_0^T S_t^{\tau} d\nu_t \right)^* \left(\int_0^T [S_t^{\tau}]^* S_t^{\tau} dt \right)^{-1} \left(\int_0^T S_t^{\tau} d\nu_t \right) \Big|_{\tau = \hat{\tau}^{\varepsilon}}.$$

La détermination du seuil c en fonction d'une probabilité de fausse alarme α (probabilité d'accepter \mathcal{H}_1 alors que \mathcal{H}_0 est vraie), n'est pas explicite et se fera via une analyse asymptotique.

OPTIMALITÉ ASYMPTOTIQUE LOCALE DU TEST DU RAPPORT DE VRAISEMBLANCE

Au lieu d'étudier le problème de test (I.46), on se donne $\theta_0 = \begin{pmatrix} \tau_0 \\ \gamma_0 \end{pmatrix}$ avec $\tau_0 < T$ et on teste :

$$\begin{cases} \mathcal{H}_0 : \text{pas de rupture avant } \tau_0 \text{ (i.e. } \tau \geq \tau_0), \\ \mathcal{H}_1 : \text{rupture avant } \tau_0 \text{ (i.e. } \tau < \tau_0). \end{cases} \quad (\text{I.55})$$

On pose :

$$\begin{aligned} D_{\varepsilon,K}^0 &= \left\{ U = \begin{pmatrix} u \\ v \end{pmatrix} ; \theta_0 + \varepsilon U \in \Theta, |U| \leq K, u \geq 0 \right\}, \\ D_{\varepsilon,K}^1 &= \left\{ U = \begin{pmatrix} u \\ v \end{pmatrix} ; \theta_0 + \varepsilon U \in \Theta, |U| \leq K, u < 0 \right\}. \end{aligned}$$

Un test $\Phi_\varepsilon : (\Omega, \mathcal{F}) \mapsto ([0, 1], \mathcal{B}[0, 1])$, est dit *localement asymptotiquement sans biais* (LASB) de niveau $\alpha \in]0, 1[$ si :

$$\overline{\lim}_{\varepsilon \downarrow 0} \sup_{U \in D_{\varepsilon,K}^0} \mathbb{E}_{\theta_0 + \varepsilon U}^\varepsilon [\Phi_\varepsilon] \leq \alpha \leq \underline{\lim}_{\varepsilon \downarrow 0} \inf_{U \in D_{\varepsilon,K}^1} \mathbb{E}_{\theta_0 + \varepsilon U}^\varepsilon [\Phi_\varepsilon], \forall K > 0.$$

Un test Φ_ε^* est dit *localement asymptotiquement de puissance maximum sans biais* (LAP-MSB) au niveau α s'il est LASB de niveau α et si, pour tout test Φ_ε LASB de niveau α :

$$\underline{\lim}_{\varepsilon \downarrow 0} \inf_{U \in D_{\varepsilon,K}^1} (\mathbb{E}_{\theta_0 + \varepsilon U}^\varepsilon [\Phi_\varepsilon^*] - \mathbb{E}_{\theta_0 + \varepsilon U}^\varepsilon [\Phi_\varepsilon]) \geq 0, \forall K > 0.$$

Le test du rapport de vraisemblance est :

$$\hat{\Phi}^\varepsilon = \begin{cases} 1 & \text{si } \Delta_{\theta_0}^{1,\varepsilon} \leq \rho(\Delta_{\theta_0}^{2,\varepsilon}) \\ 0 & \text{si } \Delta_{\theta_0}^{1,\varepsilon} > \rho(\Delta_{\theta_0}^{2,\varepsilon}) \end{cases} \quad \text{où } \Delta_{\theta_0}^\varepsilon = \begin{pmatrix} \Delta_{\theta_0}^{1,\varepsilon} \\ \Delta_{\theta_0}^{2,\varepsilon} \end{pmatrix} \quad (\text{I.56})$$

où la fonction $\rho(\cdot)$ est explicitement déterminée en fonction du niveau α . On s'inspire de la construction classique des tests de puissance maximale sans biais portant sur des familles de lois exponentielles avec paramètres de nuisance, voir LEHMANN [96, p. 52], que l'on retrouve dans MICHEL [104].

On déduit de la propriété de normalité asymptotique locale que le test (I.56) est localement asymptotiquement le plus puissant sans biais au niveau α .

OPTIMALITÉ ASYMPTOTIQUE DES EMV

Ici $\Theta = [0, \tilde{\tau}] \times \mathbb{R}_*^n$. On fait les hypothèses suivantes :

$$\forall \theta, \theta' \in \Theta, \theta \neq \theta' : \|S(\theta, \cdot) - S(\theta', \cdot)\|_{L^2(0,T)} > 0 \quad (\text{I.57})$$

$$\forall K \subset \Theta \text{ compact}, \exists \alpha_K, \beta_K > 0 \text{ tels que } \forall \theta \in K : \alpha_K I \leq I(\theta) \leq \beta_K I. \quad (\text{I.58})$$

Dans CAMPILLO [26] on a montré que ces hypothèses conduisent aux propriétés suivantes : pour tout $K \subset \Theta$ compact,

$$\sup_{\theta_1, \theta_2} |I(\theta_1)^{-1/2} I(\theta_2) I(\theta_1)^{-1/2}| < +\infty, \quad (\text{I.59})$$

$$t \rightarrow S(\theta, t) \text{ est continûment différentiable en } \theta \text{ dans } L^2(0, T), \quad (\text{I.60})$$

$$\sup_{\theta \in K} \sup_{\substack{U \in D_\theta^\varepsilon(K) \\ |U| < \varepsilon^{-1/2}}} |U|^{-2} \left\| \{ \dot{S}(\theta + \varepsilon I(\theta)^{-1/2} U, \cdot) - \dot{S}(\theta, \cdot) \} I(\theta)^{-1/2} U \right\|_{L^2(0,T)}^2 \xrightarrow{\varepsilon \downarrow 0} 0 \quad (\text{I.61})$$

$$\inf_{\theta \in K} \inf_{\substack{U \in D_\theta^\varepsilon(K) \\ |U| > \varepsilon^{-1/2}}} |I(\theta)^{1/2}|^{-1} \left\| S(\theta + \varepsilon I(\theta)^{-1/2} U, \cdot) - S(\theta, \cdot) \right\|_{L^2(0,T)}^2 > 0 \quad (\text{I.62})$$

où $D_\theta^\varepsilon(K) = \{U : \theta + \varepsilon U \in \Theta \cap K\}$.

D'après IBRAGIMOV & HAS'MINSKII [62, Th. III-5.1] les conditions (I.59)–(I.62) impliquent que les hypothèses d'Ibragimov–Has'minskii (I.37)–(I.39) sont alors satisfaites ainsi que les propriétés de l'EMV (I.40)–(I.43).

Une condition suffisante pour que les hypothèses (I.57) et (I.58) soient vérifiées est :

$$\exists c_H > 0 \text{ tel que } H^* H \geq c_H I .$$

I.4.2 UN CAS NON LINÉAIRE

On suppose que le système (I.1)–(I.2) fonctionne sous un mode nominal (normal) représenté par la valeur θ_0 du paramètre θ . Cette valeur nominale est connue ou estimée à partir des observations produites en fonctionnement normal.

La surveillance de ce système consiste à décider si les nouvelles observations sont produites par le modèle nominal (i.e. $\theta = \theta_0$) ou pas (i.e. $\theta \neq \theta_0$). Ceci devrait être fait seulement à l'aide du filtre non linéaire basé sur θ_0 (et aucune autre valeur du paramètre).

Une démarche bien définie (voir BASSEVILLE [8, 9], BENVENISTE *et al* [17]) consiste à calculer de façon récursive et à partir des observations courantes un *résidu* qui est un processus qui est idéalement "statistiquement nul" (i.e. un bruit blanc) lorsque le système est en bon fonctionnement et qui ne l'est plus en cas de rupture.

- D'une part le calcul de ce résidu doit être le moins coûteux possible. Dans le cas partiellement observé qui nous intéresse, cela signifie que seul le filtre non linéaire associé à θ_0 est utilisé.
- Une fois ce résidu formé, il est traité par une des techniques classiques de type Page–Hinkley.

On rappelle que le résidu introduit en (I.14) à partir de la fonction de vraisemblance (I.3) :

$$\Delta_{\theta_0}^{\varepsilon} = (\varepsilon \partial_{\theta} L_{\theta_0}^{\varepsilon}(\theta)|_{\theta=\theta_0})^* \quad \text{où} \quad L_{\theta_0}^{\varepsilon}(\theta) = \mathbb{E}_{\theta_0}^{\varepsilon} \left[\frac{d\mathbb{P}_{\theta}^{\varepsilon}}{d\mathbb{P}_{\theta_0}^{\varepsilon}} \middle| \mathcal{Y}_T \right] .$$

Ce résidu répond aux critères souhaités : il est calculé de façon récursive uniquement à partir des observations courantes, via le filtre non linéaire associé à θ_0 uniquement.

On identifie la loi asymptotique de $\Delta_{\theta_0}^{\varepsilon}$ sous l'alternative contiguë $\mathbb{P}_{\theta_0+\varepsilon u}^{\varepsilon}$: pour toute fonction test φ

$$\begin{aligned} \mathbb{E}_{\theta_0+\varepsilon u}^{\varepsilon} \varphi(\Delta_{\theta_0}^{\varepsilon}) &= \mathbb{E}_{\theta_0}^{\varepsilon} \left[\varphi(\Delta_{\theta_0}^{\varepsilon}) \frac{d\mathbb{P}_{\theta_0+\varepsilon u, \varepsilon}^{\varepsilon}}{d\mathbb{P}_{\theta_0}^{\varepsilon}} \right] \\ &= \mathbb{E}_{\theta_0}^{\varepsilon} \left[\varphi(\Delta_{\theta_0}^{\varepsilon}) \mathbb{E}_{\theta_0}^{\varepsilon} \left[\frac{d\mathbb{P}_{\theta_0+\varepsilon u, \varepsilon}^{\varepsilon}}{d\mathbb{P}_{\theta_0}^{\varepsilon}} \middle| \mathcal{Y}_T \right] \right] \\ &= \mathbb{E}_{\theta_0}^{\varepsilon} \left[\varphi(\Delta_{\theta_0}^{\varepsilon}) \mathcal{Z}_{\theta_0}^{\varepsilon}(u) \right] \xrightarrow{\varepsilon \downarrow 0} \mathbb{E}_{\theta_0} \left[\varphi(\Delta_{\theta_0}) \mathcal{Z}_{\theta_0}(u) \right] . \end{aligned}$$

Ainsi

$$\Delta_{\theta_0}^{\varepsilon} \xrightarrow{\varepsilon \downarrow 0} \begin{cases} N(I(\theta_0)u, I(\theta_0)) & \text{sous } \mathbb{P}_{\theta_0+\varepsilon u}^{\varepsilon} , \\ N(0, I(\theta_0)) & \text{sous } \mathbb{P}_{\theta_0}^{\varepsilon} . \end{cases}$$

Le test optimal du χ^2 pour détecter un changement dans la moyenne d'une variable aléatoire gaussienne Δ_{θ_0} est $\Delta_{\theta_0}^* I(\theta_0)^{-1} \Delta_{\theta_0} \gtrsim \lambda$. On utilise le même test pour le système de départ :

$$[\Delta_{\theta_0}^{\varepsilon}]^* I(\theta_0)^{-1} \Delta_{\theta_0}^{\varepsilon} \gtrsim \lambda .$$

L'intérêt de cette approche est que le filtrage particulière donne une approximation réaliste de Δ_ε ; en effet on peut prouver que :

$$\Delta_{\theta_0}^\varepsilon = \frac{\int_{\mathbb{R}^n} \dot{\eta}_T^\varepsilon(dx)}{\int_{\mathbb{R}^n} \eta_T^\varepsilon(dx)} \quad (\text{I.63})$$

où $\eta_T^\varepsilon(dx)$ est le filtre optimal non linéaire non normalisé et $\dot{\eta}_T^\varepsilon(dx)$ est sa dérivée par rapport au paramètre θ , pris dans les deux cas à la valeur nominale θ_0 du paramètre. Le filtrage particulière propose des approximations de ces deux termes mais également du rapport (I.63) (cf. CÉROU & LE GLAND [38], GUYADER *et al* [55], CAYLUS & AL [34]).

I.5 DÉTECTION DE RUPTURE : ASYMPTOTIQUE NON LOCALE

L'autre grande approche d'analyse asymptotique en statistique s'appuie sur la théorie des grandes déviations (voir BAHADUR [6], le séminaire d'Orsay [3], BIRGÉ [20]). Cette approche est dite non locale.

Le résultat présenté ici s'appuie sur les travaux de Matthew James et François Le Gland [65] qui suivaient eux-mêmes ceux de Matthew James et John Baras [64].

Pour des processus de diffusion partiellement observés dans le cadre d'une asymptotique petit bruit, Matthew James et François Le Gland ont démontré la consistance de l'estimateur du maximum de vraisemblance. Dans CAMPILLO & LE GLAND [30, 31] on applique ces résultats à un problème de détection de changement. Dans CAMPILLO *et al* [33] on reprend cette approche dans le cas totalement observé.

On considère le problème de test suivant :

- L'hypothèse nulle \mathcal{H}_0 (pas de rupture). Les observations Y_t sont issues du système :

$$\begin{aligned} dX_t &= b(X_t) dt + \varepsilon dW_t & X_0 &= x_0, \\ dY_t &= h(X_t) dt + \varepsilon dV_t & Y_0 &= 0, \end{aligned}$$

pour $0 \leq t \leq T$. On désigne par \mathbb{P}^ε la loi du processus (X_t, Y_t) .

- L'hypothèse alternative \mathcal{H}_1 (présence d'une rupture). Il existe $\tau \in [0, T']$ tel que les observations Y_t sont issues du système :

$$\begin{aligned} dX_t &= [b(X_t) + \mathbf{1}_{(t \geq \tau)} \bar{b}(X_t)] dt + \varepsilon dW_t & X_0 &= x_0, \\ dY_t &= [h(X_t) + \mathbf{1}_{(t \geq \tau)} \bar{h}(X_t)] dt + \varepsilon dV_t & Y_0 &= 0, \end{aligned}$$

pour $0 \leq t \leq T$. On désigne par $\mathcal{H}_1(\tau)$ l'hypothèse de rupture à l'instant τ et par $\mathbb{P}_\tau^\varepsilon$ la loi du processus (X_t, Y_t) sous cette hypothèse.

Ici T' est fixé tel que $T' < T$ (cette contrainte technique sera expliquée plus loin). On pose $b_\tau(t, x) = b(x) + \mathbf{1}_{(t \geq \tau)} \bar{b}(x)$ et $h_\tau(t, x) = h(x) + \mathbf{1}_{(t \geq \tau)} \bar{h}(x)$.

PRINCIPE DE GRANDE DÉVIATION

Sous \mathcal{H}_0 , $\{\mathbb{P}^\varepsilon\}_{\varepsilon > 0}$ satisfait à un principe de grande déviation avec fonction de taux :

$$J(\varphi) \triangleq \frac{1}{2} \int_0^T |\dot{\varphi}_t - b(\varphi_t)|^2 dt$$

si φ_t est absolument continue, $J(\varphi) = +\infty$ sinon.

Pour tout $\tau \in [0, T']$, sous $\mathcal{H}_1(\tau)$, $\{\mathbb{P}_\tau^\varepsilon\}_{\varepsilon>0}$ satisfait à un principe de grande déviation avec fonction de taux :

$$J_\tau(\varphi) \triangleq \frac{1}{2} \int_0^T |\dot{\varphi}_t - b_\tau(t, \varphi_t)|^2 dt$$

si φ_t est absolument continue, $J_\tau(\varphi) = +\infty$ sinon.

PROCESSUS DE VRAISEMBLANCE LIMITE

Le rapport de vraisemblance généralisé pour ce problème est

$$L^\varepsilon \triangleq \sup_{0 \leq \tau \leq T'} L_T^\varepsilon(\tau) = \sup_{0 \leq \tau \leq T'} \frac{\tilde{L}^\varepsilon(\tau)}{\tilde{L}^\varepsilon(T)}$$

(on reprend les définitions précédentes), de même pour les fonctions de log-vraisemblance

$$\ell^\varepsilon \triangleq \sup_{0 \leq \tau \leq T'} \ell_T^\varepsilon(\tau) \quad \text{où} \quad \ell_T^\varepsilon(\tau) \triangleq \tilde{\ell}^\varepsilon(\tau) - \tilde{\ell}^\varepsilon(T)$$

Sous \mathcal{H}_0 , i.e. en \mathbb{P}^ε probabilité, $\ell_T^\varepsilon(\tau)$ converge uniformément en $\tau \in [0, T']$ vers

$$\ell_T(\tau) \triangleq -\frac{1}{2} \inf_{\varphi: \varphi_0 = x_0} \int_0^T \{|\dot{\varphi}_t - b(\varphi_t)|^2 + |\dot{y}_t - h(\varphi_t)|^2\} dt \leq 0$$

où y_t est la sortie du système déterministe limite : $\dot{y}_t = h(x_t)$, $y_0 = 0$ et $\dot{x}_t = b(x_t)$, $x_0 = x_0$.

Sous $\mathcal{H}_1(\tau_0)$, i.e. en $\mathbb{P}_{\tau_0}^\varepsilon$ probabilité, $\ell_T^\varepsilon(\tau)$ converge uniformément en $\tau \in [0, T']$ vers

$$\begin{aligned} \ell_T^{\tau_0}(\tau) \triangleq & -\frac{1}{2} \inf_{\varphi: \varphi_0 = x_0} \int_0^T \{|\dot{\varphi}_t - b_\tau(t, \varphi_t)|^2 + |\dot{y}_t^{\tau_0} - h_\tau(t, \varphi_t)|^2\} dt \\ & + \frac{1}{2} \inf_{\varphi: \varphi_0 = x_0} \int_0^T \{|\dot{\varphi}_t - b(\varphi_t)|^2 + |\dot{y}_t^{\tau_0} - h(\varphi_t)|^2\} dt \end{aligned}$$

où $y_t^{\tau_0}$ est la sortie du système déterministe limite : $\dot{y}_t^{\tau_0} = h_{\tau_0}(t, x_t^{\tau_0})$, $y_0^{\tau_0} = 0$ et $\dot{x}_t^{\tau_0} = b_{\tau_0}(t, x_t^{\tau_0})$, $x_0^{\tau_0} = x_0$. La fonction $\tau \rightarrow \ell_T^{\tau_0}(\tau)$ atteint son maximum en $\tau = \tau_0$ et donc :

$$\tilde{\ell}_T^{\tau_0} \triangleq \max_{0 \leq \tau \leq T'} \ell_T^{\tau_0}(\tau) = \frac{1}{2} \inf_{\varphi: \varphi_0 = x_0} \int_0^T \{|\dot{\varphi}_t - b(\varphi_t)|^2 + |\dot{y}_t^{\tau_0} - h(\varphi_t)|^2\} dt$$

TEST DU RAPPORT DE VRAISEMBLANCE

Le test du rapport de vraisemblance généralisé est :

$$\Phi^\varepsilon = \mathbf{1}_{D^\varepsilon} \quad \text{où} \quad D^\varepsilon \triangleq \left\{ \sup_{0 \leq \tau \leq T'} \ell_T^\varepsilon(\tau) > c \right\} \text{ est le domaine de rejet de } \mathcal{H}_0$$

et où c est un seuil à déterminer.

Enfin on détermine les probabilités de fausse alarme et de non-détection :

$$F^\varepsilon = \mathbb{E}^\varepsilon \Phi^\varepsilon \quad \text{et} \quad N^\varepsilon = \sup_{0 \leq \tau \leq T'} \mathbb{E}^\varepsilon(1 - \Phi^\varepsilon)$$

PROBABILITÉ DE FAUSSE ALARME

De ce qui précède on obtient :

$$F^\varepsilon = \mathbb{E}^\varepsilon \mathbf{1}_{(\sup_{0 \leq \tau \leq T'} \ell_T^\varepsilon(\tau) > c)} \underset{\varepsilon \downarrow 0}{\approx} \mathbf{1}_{(\sup_{0 \leq \tau \leq T'} \ell_T(\tau) > c)}$$

on pose :

$$c_{\min} \triangleq \sup_{0 \leq \tau \leq T'} \ell_T(\tau)$$

Par programmation dynamique on peut obtenir une expression plus simple de c_{\min} (mais naturellement pas explicite), on dispose des bornes :

$$c''_{\min} \leq c_{\min} \leq c'_{\min} \leq 0$$

avec

$$c''_{\min} \triangleq -\frac{1}{2} \int_{T'}^T \{ |\bar{b}(x_t)|^2 + |\bar{h}(x_t)|^2 \} dt$$

$$c'_{\min} \triangleq -\frac{1}{2} \inf_{\varphi: \varphi_0 = x_0} \int_{T'}^T \{ |\dot{\varphi}_t - [b + \bar{b}](\varphi_t)|^2 + |\dot{y}_t - [h + \bar{h}](\varphi_t)|^2 \} dt$$

PROBABILITÉ DE NON-DÉTECTION

De ce qui précède on obtient que sous $\mathcal{H}_1(\tau_0)$:

$$\mathbb{E}_{\tau_0}^\varepsilon \mathbf{1}_{(\sup_{0 \leq \tau \leq T'} \ell_T^\varepsilon(\tau) \leq c)} \underset{\varepsilon \downarrow 0}{\approx} \mathbf{1}_{(\bar{\ell}_T^{\tau_0} \leq c)}$$

on pose :

$$c_{\max} \triangleq \inf_{0 \leq \tau_0 \leq T'} \bar{\ell}_T^{\tau_0}$$

On peut obtenir une formulation de c_{\max} par programmation dynamique. On a également les bornes :

$$0 \leq c'_{\max} \leq c_{\max} \leq c''_{\max}$$

avec

$$c''_{\max} \triangleq \frac{1}{2} \inf_{0 \leq \tau_0 \leq T'} \int_{\tau_0}^T \{ |\bar{b}(x_t^{\tau_0})|^2 + |\bar{h}(x_t^{\tau_0})|^2 \} dt$$

$$c'_{\max} \triangleq \frac{1}{2} \inf_{0 \leq \tau_0 \leq T'} \inf_{\varphi: \varphi_0 = x_0} \int_{\tau_0}^T \{ |\dot{\varphi}_t - b(\varphi_t)|^2 + |\dot{y}_t^{\tau_0} - h(\varphi_t)|^2 \} dt$$

DÉTECTABILITÉ

Le seuil c doit être choisi tel que $c_{\min} < c < c_{\max}$. Une condition de détectabilité est donc :

$$c_{\min} < c_{\max}.$$

Il est donc nécessaire que $T' < T$. On a montré que $c''_{\min} \leq c'_{\min} \leq 0 \leq c'_{\max} \leq c''_{\max}$. Une condition suffisante de détectabilité est donc $c'_{\min} < c'_{\max}$, une condition nécessaire est $c''_{\min} < c''_{\max}$.

I.6 STATISTIQUE ET FILTRAGE NON LINÉAIRE

Les statistiques de processus de diffusion avec observations partielles, sont nécessairement liées au filtrage non linéaire. Rappelons que dans les cas les plus simples, le calcul d'une statistique nécessite le plus souvent la mise en œuvre (d'une approximation) du filtre non linéaire fondé sur une valeur nominale du paramètre. Une mise en œuvre "académique" (et inaccessible) nécessiterait l'utilisation d'un filtre non linéaire par valeur du paramètre.

Une première tâche — que l'on se propose d'aborder dans cette section — est donc d'approfondir les relations entre ces statistiques et le filtre non linéaire optimal.

Reprenons le système précédent (en posant $\varepsilon = 1$) défini sur l'espace canonique $(\Omega, \mathcal{F}, \mathcal{F}_t)$:

$$dX_t = b_\theta(X_t) dt + \sigma(X_t) dW_t^\theta, \quad X_0 = \varepsilon \xi, \quad (\text{I.64})$$

$$dY_t = h_\theta(X_t) dt + dV_t^\theta, \quad Y_0 = 0 \quad (\text{I.65})$$

avec $t \geq 0$, où W_t^θ, V_t^θ sont des processus de Wiener standard, et $W_t^\theta, V_t^\theta, \xi$ sont mutuellement indépendants.

On pose :

$$\Phi_t^\theta \triangleq \exp \left(\int_0^t h_\theta(X_s)^* dY_s - \frac{1}{2} \int_0^t |h_\theta(X_s)|^2 ds \right) \quad (\text{I.66})$$

et on définit la mesure de probabilité \mathbb{P}_θ° sur (Ω, \mathcal{F}_T) par :

$$\left. \frac{d\mathbb{P}_\theta}{d\mathbb{P}_\theta^\circ} \right|_{\mathcal{F}_T} = \Phi_T^\theta. \quad (\text{I.67})$$

On définit le filtre non linéaire non normalisé :

$$\int \varphi(x) p_t^\theta(x) dx \triangleq \mathbb{E}_\theta^{\circ, \varepsilon} [\varphi(X_t) \Phi_t^{\varepsilon, \theta} | \mathcal{Y}_t]$$

La densité "non normalisée" $p_t^\theta(x)$ est solution de l'équation de Zakai (voir PARDOUX [114] pour une présentation rigoureuse) :

$$dp_t^\theta(x) = (\mathcal{L}_\theta)^* p_t^\theta(x) dt + h_\theta(x)^* p_t^\theta(x) dY_t \quad (\text{I.68})$$

où \mathcal{L}_θ est le générateur infinitésimal associé au processus de Markov X_t solution de (I.64), i.e.

$$\mathcal{L}_\theta \varphi(x) \triangleq \frac{1}{2} \sum_{i,j=1}^n a_{i,j}(x) \frac{\partial^2 \varphi(x)}{\partial x_i \partial x_j} + \sum_{i=1}^n b_{i,\theta}(x) \frac{\partial \varphi(x)}{\partial x_i}. \quad (\text{I.69})$$

Le lien avec le filtre non linéaire est évidemment :

$$\mathbb{E}_\theta^\varepsilon [\varphi(X_t) | \mathcal{Y}_t] = \frac{(p_t^\theta, \varphi)}{(p_t^\theta, 1)} \quad (\text{I.70})$$

avec la notation $(p_t^\theta, \varphi) = \int \varphi(x) p_t^\theta(x) dx$.

I.6.1 FONCTION DE LOG-VRAISEMBLANCE

Une des expressions $\theta \rightarrow \tilde{\ell}_T(\theta)$ de la fonction de log-vraisemblance s'écrit à l'aide du numérateur de l'expression (I.70), en effet :

$$\tilde{\ell}_T(\theta) = \log \mathbb{E}_\theta^\circ[\Phi_T^\theta | \mathcal{Y}_T] = \log(p_t^\theta, 1). \quad (\text{I.71})$$

À première vue cette expression ne semble pas d'un grand intérêt pratique. En fait les bonnes méthodes d'approximation du filtre non linéaire, notamment les approximations particulières, font apparaître cette expression de façon naturelle (ce n'est pas le cas du filtre de Kalman étendu). En d'autres termes, lors de la résolution numérique du filtre on disposera d'une approximation de $\tilde{\ell}_T(\theta)$.

I.6.2 ESTIMATION DE PARAMÈTRE : ALGORITHME DE GRADIENT

On a montré que :

$$\dot{\tilde{\ell}}(\theta) = \frac{(\dot{p}_T^\theta, 1)}{(p_T^\theta, 1)} \quad (\text{I.72})$$

où $p_t^\theta(x)$ est la solution de l'équation de Zakai (I.68) et $\dot{p}_t^\theta(x)$ est la solution de l'équation dérivée suivante :

$$\begin{aligned} dp_t^\theta(x) &= \mathcal{L}_\theta^* p_t^\theta(x) dt + h_\theta(x)^* \dot{p}_t^\theta(x) dY_t \\ &\quad - \text{div}(\dot{b}_\theta(x) p_t^\theta(x)) dt + \dot{h}_\theta(x)^* p_t^\theta(x) dY_t, \quad \dot{p}_0^\theta(x) = 0 \end{aligned} \quad (\text{I.73})$$

qui est formellement la dérivée de l'équation de Zakai (I.68) par rapport au paramètre θ , ainsi $\dot{p}_t^\theta(x) = (\partial/\partial\theta)p_t^\theta(x)$.

I.6.3 ESTIMATION DE PARAMÈTRE : ALGORITHME EM

On reprend quelques résultats obtenus dans CAMPILLO & LE GLAND [29].

L'algorithme EM repose sur la remarque suivante, qui est une application de l'inégalité de Jensen :

$$\tilde{\ell}_T(\theta) - \tilde{\ell}_T(\theta') = \log \mathbb{E}_{\theta'} \left[\frac{dP_\theta}{dP_{\theta'}} \middle| \mathcal{Y}_T \right] \geq \mathbb{E}_{\theta'} \left[\log \frac{dP_\theta}{dP_{\theta'}} \middle| \mathcal{Y}_T \right] \triangleq Q_T(\theta, \theta'). \quad (\text{I.74})$$

À T fixé, l'algorithme EM consiste à se donner $\hat{\theta}_0$, à poser $p = 0$ et à effectuer les itérations :

- (i) poser $p \leftarrow p + 1$ et $\theta' \leftarrow \hat{\theta}_{p-1}$,
- (ii) (étape E) calculer la fonction $\theta \rightarrow Q_T(\theta, \theta')$,
- (iii) (étape M) déterminer $\hat{\theta}_p$ qui maximise $\theta \rightarrow Q_T(\theta, \theta')$.

La propriété la plus puissante de cet algorithme est que la suite $\{\hat{\theta}_p\}_{p \geq 0}$ est maximisante, i.e. $\tilde{\ell}(\hat{\theta}_{p+1}) > \tilde{\ell}(\hat{\theta}_p)$ à moins que $\hat{\theta}_{p+1} = \hat{\theta}_p$. Les propriétés de convergence des suites $\{\tilde{\ell}(\hat{\theta}_p)\}_{p \geq 0}$ et $\{\hat{\theta}_p\}_{p \geq 0}$ ont été étudiées notamment dans WU [141] et DEMBO & ZEITOUNI [45]; ces derniers l'appliquent dans le même cadre de processus de diffusion partiellement observés. Par ailleurs cet algorithme n'a d'intérêt que si les étapes (E) et (M) peuvent se faire relativement simplement.

Il est donc nécessaire de préciser le calcul de $\theta \rightarrow Q_T(\theta, \theta')$ pour l'étape (E) et le calcul de $\theta \rightarrow \partial_{1,0}Q_T(\theta, \theta') = \frac{\partial}{\partial \theta}Q_T(\theta, \theta')$ pour l'étape (M).

Une expression pour $Q_T(\theta, \theta')$ est donnée par :

$$Q_T(\theta, \theta') = \frac{(p_T^{\theta, \theta'}, 1)}{(p_T^{\theta'}, 1)} \quad (\text{I.75})$$

où $p_t^{\theta'}(x)$ est solution de (I.68) et $\dot{p}_t^{\theta, \theta'}(x)$ est solution de :

$$\begin{aligned} d\dot{p}_t^{\theta, \theta'}(x) &= \mathcal{L}_{\theta'}^* \dot{p}_t^{\theta, \theta'}(x) dt + h_{\theta'}(x) * \dot{p}_t^{\theta, \theta'}(x) dY_t \\ &\quad - \operatorname{div}\{[b_\theta - b_{\theta'}](x) p_t^{\theta'}(x)\} dt + [h_\theta - h_{\theta'}](x) p_t^{\theta'}(x) dY_t \\ &\quad - \frac{1}{2}|a^{-1/2}(b_\theta - b_{\theta}')|^2(x) p_t^{\theta'}(x) dt - \frac{1}{2}|h_\theta - h_{\theta'}|^2(x) p_t^{\theta'}(x) dt, \\ &\quad \text{avec } \dot{p}_0^{\theta, \theta'}(x) \equiv 0. \end{aligned} \quad (\text{I.76})$$

Une expression pour $\partial_{1,0}Q_T(\theta, \theta')$ est donnée par :

$$\partial_{1,0}Q_T(\theta, \theta') = \frac{(\dot{q}_T^{\theta, \theta'}, 1)}{(p_T^{\theta'}, 1)} \quad (\text{I.77})$$

où $p_t^{\theta'}(x)$ est solution de (I.68) et $\dot{q}_t^{\theta, \theta'}(x)$ est solution de :

$$\begin{aligned} d\dot{q}_t^{\theta, \theta'}(x) &= \mathcal{L}_{\theta'}^* \dot{q}_t^{\theta, \theta'}(x) dt + h_{\theta'}(x) * \dot{q}_t^{\theta, \theta'}(x) dY_t \\ &\quad - \operatorname{div}\{\dot{b}_\theta(x) p_t^{\theta'}(x)\} dt + \dot{h}_\theta(x) p_t^{\theta'}(x) dY_t \\ &\quad - \frac{1}{2}[\dot{b}_\theta a^{-1}(b_\theta - b_{\theta}')](x) p_t^{\theta'}(x) dt - \frac{1}{2}[\dot{h}_\theta^*(h_\theta - h_{\theta}')](x) p_t^{\theta'}(x) dt, \\ &\quad \text{avec } \dot{q}_0^{\theta, \theta'}(x) \equiv 0. \end{aligned} \quad (\text{I.78})$$

On peut obtenir d'autres expressions à l'aide du lisseur. Pour cela on introduit l'équation de Zakai rétrograde :

$$dv_t^\theta(x) + \mathcal{L}_\theta v_t^\theta(x) dt + h_\theta(x) * v_t^\theta(x) dY_t = 0, \quad v_T^\theta(x) \equiv 1 \quad (\text{I.79})$$

alors en posant $r_t^\theta(x) = p_t^\theta(x) v_t^\theta(x)$ on a

$$(r_t^\theta, \varphi) = \mathbb{E}_\theta^\circ[\phi(X_t) \Phi_T^\theta | \mathcal{Y}_T].$$

On introduit le filtre et le lisseur normalisés

$$\pi_t^\theta(x) = \frac{p_t^\theta(x)}{(p_t^\theta, 1)}, \quad (\text{I.80})$$

$$\rho_t^\theta(x) = \frac{r_t^\theta(x)}{(r_t^\theta, 1)} = \frac{p_t^\theta(x) v_t^\theta(x)}{(p_T^\theta, 1)} \quad (\text{I.81})$$

Ce qui conduit aux expressions suivantes :

$$\begin{aligned}
Q_T(\theta, \theta') &= \int_0^T (\pi_t^{\theta'}, [b_\theta - b_{\theta'}]^* \nabla(\frac{\rho_t^{\theta'}}{\pi_t^{\theta'}})) dt - \frac{1}{2} \int_0^T (\rho_t^{\theta'}, |a^{-1/2} [b_\theta - b_{\theta'}]|^2) dt \\
&\quad + \int_0^T (\rho_t^{\theta'}, [h_\theta - h_{\theta'}]^*) \circ dY_t + \int_0^T (\rho_t^{\theta'}, [h_\theta - h_{\theta'}]^* h_{\theta'}) dt \\
&\quad - \frac{1}{2} \int_0^T (\rho_t^{\theta'}, |h_\theta - h_{\theta'}|^2) dt, \tag{I.82}
\end{aligned}$$

$$\begin{aligned}
\partial_{1,0} Q_T(\theta, \theta') &= \int_0^T (\pi_t^{\theta'}, \dot{b}_\theta^* \nabla(\frac{\rho_t^{\theta'}}{\pi_t^{\theta'}})) dt - \frac{1}{2} \int_0^T (\rho_t^{\theta'}, \dot{b}_\theta^* a^{-1} [b_\theta - b_{\theta'}]) dt \\
&\quad + \int_0^T (\rho_t^{\theta'}, \dot{h}_\theta^*) \circ dY_t + \int_0^T (\rho_t^{\theta'}, \dot{h}_\theta^* h_{\theta'}) dt \\
&\quad - \frac{1}{2} \int_0^T (\rho_t^{\theta'}, \dot{h}_\theta^* [h_\theta - h_{\theta'}]) dt. \tag{I.83}
\end{aligned}$$

On obtient la formulation suivante des étapes (E)/(M) :

- (E) on calcule le lisseur $\{\rho_t^{\theta'}\}_{0 \leq t \leq T}$ et donc le filtre $\{\pi_t^{\theta'}\}_{0 \leq t \leq T}$;
- (M) on maximise $\theta \rightarrow Q_T(\theta, \theta')$, à chaque valeur de θ , on calcule à chaque t les produits scalaires dans (I.82) puis on intègre en t les résultats obtenus.

L'intérêt de ces expressions via le lissage est que la dépendance en θ est plus explicite. Mais surtout, lorsque b_θ et h_θ sont linéaires en θ , on voit que l'étape (M) est explicite.

Dans CAMPILLO & LE GLAND [29] on considère également la discrétisation en temps de cet algorithme et on retrouve les algorithmes connus dans le domaine des modèles de Markov cachés.

I.6.4 LE PROBLÈME DE DÉTECTION

On considère le problème de détection de changement présenté au §I.5 p. 15 (toujours avec $\varepsilon = 1$) :

$$\begin{aligned}
dX_t &= b_\tau(t, X_t) dt + \sigma(X_t) dW_t & X_0 &= x_0, \\
dY_t &= h_\tau(t, X_t) dt + dV_t & Y_0 &= 0,
\end{aligned}$$

avec $b_\tau(t, x) = b(x) + \mathbf{1}_{(t \geq \tau)} \bar{b}(x)$ et $h_\tau(t, x) = h(x) + \mathbf{1}_{(t \geq \tau)} \bar{h}(x)$.

L'équation de Zakai s'écrit :

$$dp_t^\tau(x) = \mathcal{L}_2^* p_t^\tau(x) dt - \text{div}[b_\tau(t, x) p_t^\tau(x)] dt + h_\tau(t, x) p_t^\tau(x) dY_t$$

pour $0 \leq t \leq T$ avec la condition initiale $p_0^\tau = \delta_{x_0}$. Ici \mathcal{L}_2 désigne les termes d'ordre 2 du générateur infinitésimal (qui ne dépendent pas de τ), i.e.

$$\mathcal{L}_2 \varphi(x) \triangleq \sum_{i,j=1}^n \frac{\partial^2 a_{i,j}(x)}{\partial x_i \partial x_j} \varphi(x).$$

On obtient la formulation suivante de la log-vraisemblance : $\tilde{\ell}(\tau) = \log(p_T^\tau, 1)$. Le log-rapport de vraisemblance est donc :

$$\max_{0 \leq \tau \leq T'} \tilde{\ell}(\tau) - \tilde{\ell}(T) = \max_{0 \leq \tau \leq T'} (p_T^\tau, 1) - (p_T^T, 1).$$

Il semble donc nécessaire de résoudre autant d'équations de Zakai qu'il y a de paramètres τ à tester. On peut faire beaucoup mieux ; considérons l'équation de Zakai rétrograde :

$$dv_t^\tau(x) + \mathcal{L}_2 v_t^\tau(x) dt + b_\tau(t, x)^* \nabla v_t^\tau(x) dt + h_\tau(t, x) v_t^\tau(x) dY_t = 0$$

pour $0 \leq t \leq T$ avec la condition finale $v_T^\tau \equiv 1$. Un résultat connu (cf. PARDOUX [114]) est que $(p_t^\tau, 1) = (p_t^\tau, v_t^\tau)$ pour tout $t \in [0, T]$ i.e. cette quantité ne dépend pas de t . On peut donc prendre $t = \tau$ et considérer $\tilde{L}(\tau) = (p_\tau^\tau, v_\tau^\tau)$.

Ainsi, au lieu de considérer les deux familles d'EDPS précédentes indicées par $\tau \in [0, T]$, il suffit de considérer les deux EDPS, la première progressive, la seconde rétrograde

$$dp_t(x) = \mathcal{L}_2^* p_t(x) dt - \operatorname{div}[b(x) p_t(x)] dt + h(x) p_t(x) dY_t, \quad p_0 = \delta_{x_0} \quad (\text{I.84})$$

$$d\bar{v}_t(x) + \mathcal{L}_2 \bar{v}_t(x) dt + [b + \bar{b}](x)^* \nabla \bar{v}_t(x) dt + [h + \bar{h}](x) \bar{v}_t(x) dY_t = 0, \quad \bar{v}_T \equiv 1 \quad (\text{I.85})$$

et de poser

$$\tilde{\ell}(\tau) - \tilde{\ell}(T) = \log(p_\tau, \bar{v}_\tau) - \log(p_T, 1). \quad (\text{I.86})$$

Cette méthode n'est pas récursive : l'équation (I.84) est récursive mais le calcul du log-rapport de vraisemblance (I.86) nécessite la résolution (rétrograde) de (I.85), ce calcul n'est donc pas récursif. Mais, au lieu de calculer le log-rapport de vraisemblance généralisé $\max_{0 \leq \tau \leq T} \tilde{\ell}(\tau) - \tilde{\ell}(T)$, on peut se limiter à $\max_{T-\delta_1 \leq \tau \leq T-\delta_2} \tilde{\ell}(\tau) - \tilde{\ell}(T)$ avec $\delta_1 > \delta_2 > 0$ fixés. En conclusion le calcul de ce log-rapport de vraisemblance généralisé, limité à une fenêtre glissante, nécessite la résolution du filtre optimal (I.84) qui se fait de façon récursive et le calcul rétrograde sur une fenêtre glissante de taille fixe du terme correctif (I.85).

I.6.5 MÉTHODES NUMÉRIQUES

On considère un système différentiel stochastique (non observé) :

$$dX_t = b(X_t) dt + \sigma(X_t) dW_t, \quad X_0 \sim \mu_0, \quad t \geq 0 \quad (\text{I.87})$$

et des observations à temps discrets

$$y_k = h(X_{t_k}) + v_k, \quad t_k = k \Delta, \quad k \geq 1 \quad (\text{I.88})$$

où W_t est un processus de Wiener standard, v_k est un bruit blanc gaussien $N(0, I)$. Ces bruits ainsi que X_0 sont supposés mutuellement indépendants. On définit le filtre non linéaire

$$\pi_t(dx) = \mathbb{P}(X_t \in dx | y_k, t_k \leq t)$$

et le filtre prédit $\pi_{t-}(dx) = \mathbb{P}(X_t \in dx | y_k, t_k < t)$. L'évolution de ce filtre est régie par les deux étapes classiques :

prédiction (équation de Fokker-Planck)

$$\frac{d}{dt} \pi_t = \mathcal{L}^* \pi_t, \quad t \in [t_{k-1}, t_k[. \quad (\text{I.89})$$

La valeur finale de cette équation est $\pi_{t_k^-}$. \mathcal{L} est le générateur infinitésimal du processus (I.87).

correction (formule de Bayes)

$$\pi_{t_k}(dx) = \frac{\Psi_k(x) \pi_{t_k}^-(dx)}{\langle \pi_{t_k}^-, \Psi_k \rangle}. \quad (\text{I.90})$$

où $\Psi_k(x)$ est la vraisemblance locale $\Psi_k(x) \propto \exp(-\frac{1}{2} \|y_k - h(x)\|^2)$.

Les équations “optimales” (I.89)–(I.90) ne présentent en elles-mêmes aucun intérêt en pratique. Il est parfaitement connu que dans le cas d’un système linéaire-gaussien, la loi π_t est gaussienne et se calcule explicitement.

Dans les années 80, à la suite de Václav Beneš [15], beaucoup de travaux ont porté sur l’existence de filtres de dimension finie. Il a été finalement démontré que pour la plupart des problèmes de filtrage non linéaire il n’existe pas de filtre de dimension finie (cf. CHALEYAT-MAUREL & MICHEL [35], OCONE & PARDOUX [111]). Jean Lévine [103] avait bien exploité les possibilités numériques des filtres de dimension finie. Une des carences de cette approche est que l’existence d’un filtre fini-dimensionnel n’assure en rien que ce filtre possède de bonnes qualités numériques (contrairement au cas linéaire gaussien). Une petite expérience dans le domaine me laisse penser que les équations de ces filtres peuvent être numériquement très mal conditionnées. Ceci en fait, à mon avis, une “mauvaise” méthode. Une “bonne” méthode doit :

- (i) être simple à mettre en œuvre (et d’une certaine manière simple à comprendre pour faciliter sa diffusion),
- (ii) donner de bons résultats dans les cas standard,
- (iii) offrir de nombreuses extensions permettant de s’adapter à de nombreuses situations, d’améliorer les performances etc.

Pour le moment, en dehors du cas linéaire gaussien, seul le filtrage particulaire présente ces caractéristiques. Le filtre de Kalman étendu est également un outil puissant, toutefois il existe de nombreux cas où il ne peut s’appliquer simplement ; de plus il peut présenter des problèmes de stabilité. Mais lorsqu’il fonctionne correctement il est toujours moins coûteux que le filtre particulaire.

MÉTHODES DE MOMENTS

La première idée que j’ai explorée au tout début de ma thèse consistait à proposer une approximation du filtre non linéaire sous forme d’une somme pondérée de masses de Dirac [26]. L’idée provient de Donald Dawson [42] qui l’avait appliquée à l’équation de Fokker-Planck.

Le filtre non linéaire est donc approché de la façon suivante :

$$\pi_t(dx) \simeq \pi_t^N(dx) = \sum_{i=1}^N \omega_t^i \delta_{x_t^i}(dx) \quad (\text{I.91})$$

où, à N fixé, l’ensemble des poids/points est calculé en posant que les $2N - 1$ premiers moments de π_t et π_t^N coïncident (en fait on utilise des moments “modifiés” afin d’obtenir un problème bien conditionné).

Cette méthode est très efficace en dimension 1 mais uniquement en dimension 1 ! Toutefois il faut retenir un point important, l’approximation (I.91) de forme “particulaire” facilite grandement le calcul des fonctions de la loi conditionnelle :

$$\mathbb{P}(X_t \in B | Y_s; s \leq t) \simeq \sum_{i=1}^N \omega_t^i \mathbf{1}_B(x_t^i) \quad \mathbb{E}(\varphi(X_t) | Y_s; s \leq t) \simeq \sum_{i=1}^N \omega_t^i \varphi(x_t^i).$$

On retient donc qu'une approximation comme (I.91) résout de façon très spectaculaire ce type de calcul qui en théorie nécessite l'évaluation d'une intégrale.

Cette idée se retrouve d'une certaine manière (mais de façon plus efficace) dans le filtre "unscented" de Simon Julier et Jeffrey Uhlmann [67, 68].

MÉTHODES AUX DIFFÉRENCES FINIES

Cette approche a été proposée par Harold Kushner (voir KUSHNER [78, 79, 80] ou KUSHNER & DUPUIS [81]) et reprise par François Le Gland dans sa thèse [52].

Dit de façon simple, l'équation de Fokker-Planck (I.89) est approchée par des schémas aux différences finies. Plus précisément, le processus de Markov X_t est approché par une chaîne de Markov à valeurs dans un espace fini ; le problème de filtrage non linéaire associé est alors exactement résolu.

Cette technique présentent quelques sévères défauts :

- Cette approche implique que l'on travaille, concernant la variable d'état, sur un domaine fini (en utilisant des conditions aux bords ad hoc). Ainsi il est nécessaire que la loi conditionnelle que l'on souhaite approcher, "vive" dans une région bornée de l'espace d'état.
- Un autre problème est que les efforts de calcul sont uniformément répartis sur ce domaine quel que soit le problème traité. Il serait souhaitable que les efforts de calcul se portent sur les zones d'intérêt (les zones de forte probabilité) en traitant les zones de peu d'intérêt de façon grossière. François Le Gland a développé des techniques multigrilles pour palier cette difficulté (cf. CAI *et al* [23]).
- Dans tous les cas cette approche est très coûteuse en espace mémoire et en temps de calcul. Elle a pu trouver un regain d'intérêt à l'apparition des super-calculateurs.

En revanche, cette approche permet de calculer une approximation fine de la densité conditionnelle. Il n'existe pas à ma connaissance d'autres méthodes permettant d'offrir une telle approximation. L'approche particulière n'a pas pour but d'approcher la densité conditionnelle en tant que fonction.

VERS LE PARTICULAIRE

Lorsque l'équation d'état n'est pas bruitée, i.e. lorsque $\sigma \equiv 0$ dans (I.87), le générateur infinitésimal \mathcal{L} ainsi que l'EDP (I.89) sont du premier ordre. Il est alors possible de proposer une approximation de type particulière (i.e. de la forme (I.91)) ; voir CAMPILLO & LE GLAND [28]. Cette idée de François Le Gland s'inspirait des méthodes particulières proposées par Pierre-Arnaud Raviart [118]. Les points x_t^i évoluent de façon déterministe selon l'équation d'état (I.87), la mise à jour des poids ω_t^i selon la formule de Bayes en utilisant (I.88).

Cette méthode très simple présente de bonnes possibilités de parallélisation (cf. BERNIER *et al* [18], CAMPILLO *et al* [27]). Toutefois cette approche n'est pas par essence adaptative : les particules ne se concentrent pas nécessairement dans les zones d'intérêt de l'espace d'état. Il faut également noter que cette méthode n'utilise aucune technique de Monte Carlo.

MÉTHODES PARTICULAIRES

Nous étions pourtant encore loin de ce qui s'appelle maintenant "approximation particulière". Deux arguments essentiels manquaient à notre approche : l'échantillonnage i.e. une simulation de type Monte Carlo et le rééchantillonnage i.e. favoriser les particules qui sont dans les zones d'intérêt. L'approximation particulière offre une solution élégante et surtout efficace à ces deux problèmes.

En 1993, Neil Gordon, David Salmond et Adrian Smith [53] en ont proposé une des premières versions, le filtre particulière de type "bootstrap". À la même époque, Pierre Del Moral développait des algorithmes de même nature.

On trouvera dans DOUCET *et al* [47] et DJURIC & GOODSILL [46] une présentation complète du sujet. ARULAMPALAM [2] est un bon article d'introduction. Enfin, Pierre Del Moral a profondément fait progresser les outils d'analyse de ces méthodes, voir DEL MORAL & MICLO [44] et DEL MORAL [43]. Ces outils s'étendent à d'autres domaines que le filtrage.

Le succès du filtrage particulière est essentiellement dû aux bons résultats obtenus dans plusieurs domaines d'applications et notamment en poursuite à partir de divers types de mesures (cf. GUSTAFSSON *et al* [54] ou la monographie RISTIC *et al* [119]). Enfin, le filtrage particulière commence à être appliqué en détection de changement, voir ANDRIEU *et al* [1], AZIMI-SADJADI & KRISHNAPRASAD [4, 5], KADIRKAMANATHAN *et al* [69], DRIESSEN & BOERS [48] qui reprennent les travaux de Henk Blom [21] en "non particulière".

BIBLIOGRAPHIE

- [1] C. ANDRIEU, A. DOUCET, S. S. SINGH, AND V. B. TADIC. Particle methods for change detection, identification and control. *Proceedings of the IEEE*, 92(3) :423–438, 2004.
- [2] M. S. ARULAMPALAM, S. MASKELL, N. J. GORDON, AND T. CLAPP. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, SP-50(2 (Special issue on Monte Carlo Methods for Statistical Signal Processing)) :174–188, February 2002.
- [3] ASTÉRISQUE. *Grandes Déviations et Applications Statistiques, Séminaire Orsay 1977–1978*, volume 68 of *Astérisque*. Société Mathématique de France, Paris, 1979.
- [4] B. AZIMI-SADJADI AND P. S. KRISHNAPRASAD. A particle filtering approach to change detection for nonlinear systems. *EURASIP Journal on Applied Signal Processing*, 15 :2295–2305, 2004. <http://www.hindawi.co.uk/open-access/asp/volume-2004/S1110865704408051.pdf>.
- [5] B. AZIMI-SADJADI AND P. S. KRISHNAPRASAD. Change detection for nonlinear systems : A particle filtering approach. In *Proceedings of 2002 American Control Conference, ACC2002*, 2002. http://www.ecse.rpi.edu/~babak/ACC_CAMERA_READY.pdf.
- [6] R. R. BAHADUR. On the asymptotic efficiency of tests and estimators. *Sankhyā*, 22 :229–252, 1960.
- [7] I. V. BASAWA AND D. J. SCOTT. *Asymptotic optimal inference for nonergodic models*, volume 17 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1983.

- [8] M. BASSEVILLE. Detecting changes in signals and systems—a survey. *Automatica*, 24(3) :309–326, 1988.
- [9] M. BASSEVILLE. Information criteria for residual generation and fault detection and isolation. *Automatica*, 33(5) :783–803, May 1997.
- [10] M. BASSEVILLE. On-board component fault detection and isolation using the statistical local approach. *Automatica*, 34(11) :1391–1416, Nov. 1998.
- [11] M. BASSEVILLE AND A. BENVENISTE, editors. *Detection of abrupt changes in signal and dynamical systems*, volume 77 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1986.
- [12] M. BASSEVILLE, A. BENVENISTE, G. V. MOUSTAKIDES, AND A. ROUGÉE. Detection and diagnosis of changes in the eigenstructure of nonstationary multivariable systems. *Automatica*, 23(4) :479–489, 1987.
- [13] M. BASSEVILLE, A. BENVENISTE, AND G. V. MOUSTAKIDES. Detection and diagnosis of abrupt changes in modal characteristics of nonstationary digital signals. *IEEE Transactions on Information Theory*, 32(3) :412–417, 1986.
- [14] M. BASSEVILLE AND I. V. NIKIFOROV. *Detection of abrupt changes : theory and application*. Prentice Hall Information and System Sciences Series. Prentice Hall Inc., Englewood Cliffs, NJ, 1993. <http://www.irisa.fr/sisthem/kniga/>.
- [15] V. E. BENEŠ. Exact finite-dimensional filters for certain diffusions with nonlinear drift. *Stochastics*, 5(1+2) :65–92, 1981.
- [16] A. BENVENISTE AND M. BASSEVILLE. Detection of abrupt changes in signals and dynamical systems : some statistical aspects. In *Analysis and optimization of systems, Part 1 (Nice, 1984)*, volume 62 of *Lecture Notes in Control and Information Sciences*, pages 145–155. Springer-Verlag, Berlin, 1984.
- [17] A. BENVENISTE, M. BASSEVILLE, AND G. V. MOUSTAKIDES. The asymptotic local approach to change detection and model validation. *IEEE Transactions on Automatic Control*, 32(7) :583–592, 1987.
- [18] H. BERNIER, F. CAMPILLO, F. CÉROU, AND F. LE GLAND. Parallélisme de données et filtrage non linéaire : Analyse de performance. Rapport Technique RT-167, INRIA, November 1994. <http://www.inria.fr/rrrt/rt-0167.html>.
- [19] G. K. BHATTACHARYA AND R. A. JOHNSON. Non-parametric tests for shift at an unknown time point. *The Annals of Mathematical Statistics*, 39 :1731–1743, 1968.
- [20] L. BIRGÉ. *Approximation dans les Espaces Métriques et Théorie de l’Estimation. Inégalité de Cramér-Chernoff et Théorie Asymptotique des Tests*. Thèse de doctorat d’état, Université Paris VII, 1980.
- [21] H. A. P. BLOM. An efficient filter for abruptly changing systems. In *Proceedings of the 33rd Conference on Decision and Control, Las Vegas, NV*, pages 656–658. IEEE-CSS, December 1984.
- [22] B. E. BRODSKY AND B. S. DARKHOVSKY. *Nonparametric Methods in Change-Point Problems*, volume 243 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 1993.
- [23] Z. CAI, F. LE GLAND, AND H. ZHANG. An adaptive local grid refinement method for nonlinear filtering. Rapport de Recherche 2679, INRIA, October 1994. <http://www.inria.fr/rrrt/rr-2679.html>.
- [24] F. CAMPILLO. *Filtrage et Détection de Ruptures de Processus Partiellement Observées*. Thèse de Troisième Cycle, Université de Provence, Marseille, 1984.

- [25] F. CAMPILLO. La méthode d'approximation de Gauss–Galerkin en filtrage non linéaire. *RAIRO Modélisation Mathématique et Analyse Numérique (M2AN)*, 20(2) :203–223, 1986.
- [26] F. CAMPILLO. Testing for a change–point in linear systems with incomplete observation. In *Asymptotic theory for non i.i.d. processes (Luminy–Marseille, 1984)*, volume 8 of *Travaux Rech.*, pages 225–242. Publications des Facultés Universitaires Saint–Louis, Brussels, 1986.
- [27] F. CAMPILLO, F. CÉROU, F. LE GLAND, AND R. RAKOTOZAFY. Particle and cell approximations for nonlinear filtering. Rapport de Recherche RR-2567, INRIA, June 1995. <http://www.inria.fr/rrrt/rr-2567.html>.
- [28] F. CAMPILLO AND F. LE GLAND. Application du filtrage non-linéaire en trajectographie passive. In *12ème Colloque GRETSI sur le Traitement du Signal, Juan les Pins, June 12-16, 1989*.
- [29] F. CAMPILLO AND F. LE GLAND. MLE for partially observed diffusions : direct maximization vs. the EM algorithm. *Stochastic Processes and their Applications*, 33(2) :245–274, 1989.
- [30] F. CAMPILLO AND F. LE GLAND. Likelihood-based statistics for partially observed diffusion processes. In *Proceedings of the 1st European Control Conference, Grenoble, July 2-5, 1991*.
- [31] F. CAMPILLO AND F. LE GLAND. Threshold selection in the GLR test for change detection in partially observed diffusion processes. In *Proceedings of the 2nd European Control Conference, Groningen, June 28-July 1, 1993*.
- [32] F. CAMPILLO AND F. LE GLAND. Local asymptotic normality of partially observed small noise diffusions, 2004. En préparation.
- [33] F. CAMPILLO, Y. KUTOYANTS, AND F. LE GLAND. Small noise asymptotics of the GLR test for off-line change detection in misspecified diffusion processes. *Stochastics and Stochastics Reports*, 70(1–2) :109–129, 2000.
- [34] N. CAYLUS, A. GUYADER, F. LE GLAND, AND N. OUDJANE. Application du filtrage particulière à l'inférence statistique des HMM. In *Actes des XXXVIèmes Journées de Statistique (SFDS'04)*, 2004.
- [35] M. CHALEYAT-MAUREL AND D. MICHEL. Des résultats de non existence de filtre de dimension finie. *Stochastics*, 13(1-2) :83–102, 1984.
- [36] H. CHERNOFF AND S. ZACKS. Estimating the current mean of a normal distribution which is subjected to changes in time. *Annals of Mathematical Statistics*, 35 :999–1018, 1964.
- [37] D. M. CHIBISOV. Asymptotic expansions for Neyman's $C(\alpha)$ tests. In *Proceedings of the Second Japan-USSR Symposium on Probability Theory (Kyoto, 1972)*, volume 330 of *Lecture Notes in Mathematics*, pages 16–45, Springer-Verlag, Berlin, 1973.
- [38] F. CÉROU AND F. LE GLAND. Efficient particle filters for residual generation in partially observed SDE's. In *Proceedings of the 39th Conference on Decision and Control, Sydney 2000*, pages 1200–1205, December 2000.
- [39] D. DACUNHA–CASTELLE. Vitesse de convergence pour certains problèmes statistiques. In Paul-Louis Hennequin, editor, *École d'Été de Probabilités de Saint-Flour VII, 1977*, volume 678 of *Lecture Notes in Mathematics*, pages 1–172. Springer-Verlag, Berlin, 1978.
- [40] B. S. DARKHOVSKY. A non–parametric method for the a posteriori detection of the disorder time of a sequence of independent random variables. *Theory of Probability and its Applications*, 21(1), 1976.

- [41] M. H. A. DAVIS. The application of nonlinear filtering to fault detection in linear systems. *IEEE Transactions on Automatic Control*, AC-20(2) :257–259, April 1975.
- [42] D. A. DAWSON. Galerkin approximation of nonlinear Markov processes. In *Statistics and related topics (Ottawa, Ont., 1980)*, pages 317–339. North-Holland, Amsterdam, 1981.
- [43] P. DEL MORAL. *Feynman-Kac formulae*. Springer-Verlag, New York, 2004.
- [44] P. DEL MORAL AND L. MICLO. Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to nonlinear filtering. In J. Azéma, M. Émery, M. Ledoux, and M. Yor, editors, *Séminaire de Probabilités XXXIV*, volume 1729 of *Lecture Notes in Mathematics*, pages 1–145. Springer-Verlag, Berlin, 2000.
- [45] A. DEMBO AND O. ZEITOUNI. Parameter estimation of partially observed continuous time stochastic processes via the EM algorithm. *Stochastic Processes and their Applications*, 23(1) :91–113, 1986.
- [46] P. M. DJURIC AND S. J. GOODSILL, editors. *Special issue on Monte Carlo methods for statistical signal processing*, volume 50. IEEE Transactions on Acoustics, Speech, and Signal Processing, February 2002.
- [47] A. DOUCET, N. DE FREITAS, AND N. GORDON, editors. *Sequential Monte Carlo Methods in Practice*. Statistics for Engineering and Information Science. Springer-Verlag, New York, 2001.
- [48] H. DRIESSEN AND Y. BOERS. An efficient particle filter for nonlinear jump Markov systems. To appear. http://www.iee.org/OnComms/PN/controlauto/Paper3-Driessen_H.pdf.
- [49] V. GENON-CATALOT AND D. PICARD. *Éléments de Statistique Asymptotique*, volume 11 of *Mathématiques & Applications*. Springer-Verlag, Paris, 1993.
- [50] J. GERTLER. Survey of model-based failure detection and isolation in complex plants. *IEEE Control Systems Magazine*, 8(6) :3–11, 1988.
- [51] B. K. GHOSH AND P. K. SEN, editors. *Handbook of sequential analysis*. Marcel Dekker, New York, 1991.
- [52] F. LE GLAND. *Estimation de Paramètres dans les Processus Stochastiques, en Observation Incomplète – Applications à un Problème de Radio-Astronomie*. Thèse de Docteur-Ingénieur, Université de Paris IX – Dauphine, 1981.
- [53] N. J. GORDON, D. J. SALMOND, AND A. F. M. SMITH. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings, Part F*, 140(2) :107–113, April 1993.
- [54] F. GUSTAFSSON, F. GUNNARSSON, N. BERGMAN, U. FORSELL, J. JANSSON, R. KARLSSON, AND P.-J. NORDLUND. Particle filters for positioning, navigation and tracking. *IEEE Transactions on Signal Processing*, 50(2), 2002. <http://www.control.isy.liu.se/~fredrik/reports/01SPpf4pos.pdf>.
- [55] A. GUYADER, F. LE GLAND, AND N. OUDJANE. A particle implementation of the recursive MLE for partially observed diffusions. In *Proceedings of the 13th IFAC/IFORS Symposium on System Identification*, pages 1305–1310, Rotterdam, August 27-29 2003.
- [56] J. HÁJEK. Local asymptotic minimax and admissibility in estimation. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, 1970/1971)*, Vol. I : *Theory of statistics*, pages 175–194, Berkeley, Calif., 1972. University of California Press.
- [57] D. M. HIMMELBLAU. *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*. Elsevier, 1978.

- [58] D. V. HINKLEY. Inference about the change-point in a sequence of random variables. *Biometrika*, 57 :1–17, 1970.
- [59] D. V. HINKLEY. Inference about the change-point from cumulative sum tests. *Biometrika*, 58 :509–523, 1971.
- [60] D. V. HINKLEY AND E. A. HINKLEY. Inference about the change-point in a sequence of binomial variables. *Biometrika*, 57 :477–488, 1970.
- [61] I. A. IBRAGIMOV AND R. Z. HAS'MINSKII. Asymptotic behavior of statistical estimators in the smooth case 1 : study of the likelihood ratio. *Theory of Probability and its Applications*, 17(3) :445–462, 1972.
- [62] I. A. IBRAGIMOV AND R. Z. HAS'MINSKII. *Statistical Estimation, Asymptotic Theory*, volume 16 of *Applications of Mathematics*. Springer–Verlag, New York, 1981.
- [63] R. ISERMANN. Process fault detection based on modeling and estimation method - A survey. *Automatica*, 20(4) :387–404, 1984.
- [64] M. R. JAMES AND J. S. BARAS. Nonlinear filtering and large deviations : a PDE-control theoretic approach. *Stochastics*, 23(3) :391–412, 1988.
- [65] M. R. JAMES AND F. LE GLAND. Consistent parameter estimation for partially observed diffusions with small noise. *Applied Mathematics & Optimization*, 32(1) :47–72, 1995.
- [66] A. H. JAZWINSKII. *Stochastic processes and filtering Theory*. Academic Press, New York, 1970.
- [67] S. J. JULIER. The scaled unscented transformation. In *Proceedings of the IEEE American Control Conference*, pages 4555–4559, 2002.
- [68] S. J. JULIER AND J. K. UHLMANN. Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations. In *Proceedings of the IEEE American Control Conference*, pages 887–892, 2002.
- [69] V. KADIRKAMANATHAN, P. LI, M. H. JAWARD, AND S. G. FABRI. Particle filtering-based fault detection in non-linear stochastic systems. *Internat. J. Systems Sci.*, 33(4) :259–265, 2002.
- [70] G. KALLIANPUR. *Stochastic Filtering Theory*. Application of Mathematics. Springer–Verlag, New York, 1980.
- [71] Z. KANDER AND S. ZACKS. Test procedures for possible changes in parameters of statistical distributions occurring at unknown time points. *The Annals of Mathematical Statistics*, 37 :1196–1210, 1966.
- [72] M. L. KLEPTSINA, ROBERT S. LIPTSER, AND A. P. SEREBROVSKI. Nonlinear filtering problem with contamination. *The Annals of Applied Probability*, 7(4) :917–934, 1997.
- [73] N. KLIGENE. On the asymptotic behaviour in detecting change in a time series. In *Transactions of the ninth Prague conference on information theory, statistical decision functions, random processes, Vol. B (Prague, 1982)*, pages 23–28. Reidel, Dordrecht, 1983.
- [74] N. KLIGENE. The probabilities of errors in likelihood estimation of a change-point in statistical models. In *Proceedings of the Fourth Prague Symposium on Asymptotic Statistics (Prague, 1988)*, pages 321–328, Prague, 1989. Charles University.
- [75] N. KLIGENE. Likelihood inference about a change point in switching auto-regression. *Informatika*, 2(1) :53–65, 137, 144, 1991.
- [76] N. KLIGENE AND L. TELKSNYS. Methods of detecting instants of change of random process properties. *Avtomatika i Telemekhanika*, 10 :5–56, 1983.

- [77] F. KONECNY. Maximum likelihood estimation of a drift parameter from a partially observed diffusion in the case of small measurement noise. *Statist. Decisions*, 8(2) :115–130, 1990.
- [78] H. J. KUSHNER. Probability limit theorems and the convergence of finite difference approximations of partial differential equations. *J. Math. Anal. Appl.*, 32 :77–103 (erratum and addendum, *ibid.* 34, 1971), 1970.
- [79] H. J. KUSHNER. A survey of some applications of probability and stochastic control theory to finite difference methods for degenerate elliptic and parabolic equations. *SIAM Review*, 18(4) :545–577, 1976.
- [80] H. J. KUSHNER. *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, volume 129 of *Mathematics in Science and Engineering*. Academic Press, New York, 1977.
- [81] H. J. KUSHNER AND P. DUPUIS. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of Mathematics*. Springer-Verlag, New York, second edition, 2001.
- [82] YU. A. KUTOYANTS. On a problem of testing hypotheses and asymptotic normality of stochastic integrals. *Theory of Probability and its Applications*, 20 :376–383, 1975.
- [83] YU. A. KUTOYANTS. Estimation of a trend parameter of a diffusion process in the smooth case. *Theory of Probability and its Applications*, 22 :399–405, 1977.
- [84] YU. A. KUTOYANTS. Estimation of a parameter of a diffusion process. *Theory of Probability and its Applications*, 23 :641–649, 1978.
- [85] YU. A. KUTOYANTS. *Parameter estimation for stochastic processes*, volume 6 of *Research and Exposition in Mathematics*. Heldermann Verlag, Berlin, 1984. First edition in Russian 1980.
- [86] YU. A. KUTOYANTS. *Identification of dynamical systems with small noise*, volume 300 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1994.
- [87] YU. A. KUTOYANTS. *Statistical Inference for Ergodic Diffusion Processes*. Springer Series in Statistics. Springer-Verlag, New York, 2003.
- [88] T. L. LAI. Sequential changepoint detection in quality control and dynamical systems (with discussion). *Journal of the Royal Statistical Society, Series B*, 57(4) :613–658, 1995.
- [89] L. LE CAM. Locally asymptotically normal families of distributions. Certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses. *University of California Publications in Statistics*, 3 :37–98, 1960.
- [90] L. LE CAM. *Théorie asymptotique de la décision statistique*. Séminaire de Mathématiques Supérieures, No. 33 (Été, 1968). Les Presses de l'Université de Montréal, 1969.
- [91] L. LE CAM. Limits of experiments. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, 1970/1971)*, Vol. I : *Theory of statistics*, pages 245–261, Berkeley, 1972. University of California Press.
- [92] L. LE CAM. Sur l'approximation de familles de mesures par des familles gaussiennes. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 21(3) :225–287, 1985.
- [93] L. LE CAM. *Asymptotic methods in statistical decision theory*. Springer Series in Statistics. Springer-Verlag, New York, 1986.


- [94] L. LE CAM AND G. L. YANG. *Asymptotics in statistics, Some Basic Concepts*. Springer Series in Statistics. Springer-Verlag, New York, second edition, 2000.
- [95] F. LE GLAND AND B. WANG. Asymptotic normality in partially observed diffusions with small noise : application to FDI. In Boženna Pasik-Duncan, editor, *Workshop on Stochastic Theory and Control, University of Kansas 2001. In honor of Tyrone E. Duncan on the occasion of his 60th birthday*, number 280 in Lecture Notes in Control and Information Sciences, pages 267–282. Springer-Verlag, Berlin, 2002.
- [96] E. L. LEHMANN. *Testing statistical hypotheses*. Springer Texts in Statistics. Springer-Verlag, New York, second edition, 1997.
- [97] R. SH. LIPTSER AND A. N. SHIRYAEV. *Statistics of Random Processes I. General Theory*, volume 5 of *Applications of Mathematics*. Springer-Verlag, second edition, 2001.
- [98] R. SH. LIPTSER AND A. N. SHIRYAEV. *Statistics of Random Processes II. Applications*, volume 6 of *Applications of Mathematics*. Springer-Verlag, second edition, 2001.
- [99] K. A. LOPARO, Z. ROTH, AND S. J. ECKERT. Nonlinear filtering for systems with random structure. *IEEE Transactions on Automatic Control*, 31(11) :1064–1068, 1986.
- [100] K. L. LOPARO AND Z. ROTH. Nonlinear filtering for process diagnostics. In *Proceedings of the 1981 Joint Automatic Control Conference*, June 1981.
- [101] K. L. LOPARO AND Z. ROTH. On the solution of the Zakai equation for the process diagnostics problem. In *Proceedings of the 33rd Conference on Decision and Control, Las Vegas, NV*, pages 636–637. IEEE-CSS, December 1984.
- [102] G. LORDEN. Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics*, 42 :1897–1908, 1971.
- [103] J. LÉVINE. Finite dimensional realizations of stochastic PDE’s and applications to filtering. *Stochastics and Stochastics Reports*, 37(1+2) :75–103, 1991.
- [104] R. MICHEL. On the asymptotic efficiency of conditional tests for exponential families. *The Annals of Statistics*, 7(6) :1256–1263, 1979.
- [105] P. W. MILLAR. The minimax principle in asymptotic statistical theory. In *Eleventh Saint Flour probability summer school—1981 (Saint Flour, 1981)*, volume 976 of *Lecture Notes in Mathematics*, pages 75–265. Springer-Verlag, Berlin, 1983.
- [106] G. V. MOUSTAKIDES. Robust detection of signals : a large deviations approach. *IEEE Transactions on Information Theory*, 31(6) :822–825, 1985.
- [107] G. V. MOUSTAKIDES. Optimal stopping times for detecting changes in distributions. *The Annals of Statistics*, 14(4) :1379–1387, 1986.
- [108] G. V. MOUSTAKIDES. Optimality of the CUSUM procedure in continuous time. *The Annals of Statistics*, 32(1) :302–315, 2004.
- [109] I. V. NIKIFOROV. Cumulative sums for detection of changes in random process characteristics. *Automation and Remote Control*, 2 :48–58, 1979.
- [110] I. V. NIKIFOROV. Sequential detection of changes in stochastic processes. In *Identification and system parameter estimation 1991, Vol. 1,2 (Budapest, 1991)*, volume 3 of *IFAC Sympos. Ser.*, pages 11–19. Pergamon, Oxford, 1992.
- [111] D. L. OCONE AND É. PARDOUX. A Lie algebraic criterion for non-existence of finite dimensionally computable filters. In G. Da Prato and L. Tubaro, editors, *Stochastic PDE’s and Applications II, Trento 1988*, volume 1390 of *Lecture Notes in Mathematics*, pages 197–204. Springer-Verlag, Berlin, 1989.

- [112] E. S. PAGE. Continuous inspection schemes. *Biometrika*, 41 :100–115, 1954.
- [113] E. S. PAGE. A test for a change in a parameter occurring at an unknown point. *Biometrika*, 42 :523–527, 1955.
- [114] É. PARDOUX. Filtrage non linéaire et équations aux dérivées partielles stochastiques associées. In P.–L. Hennequin, editor, *École d'Été de Probabilités de Saint-Flour XIX, 1989*, volume 1464 of *Lecture Notes in Mathematics*, pages 67–163. Springer-Verlag, Berlin, 1991.
- [115] L. F. PAU. *Failure diagnosis and performance monitoring*. Marcel Dekker, New York, 1981.
- [116] A. N. PETTITT. A nonparametric approach to the change-point problem. *Journal of the Royal Statistical Society, Series C*, 28(2) :126–135, 1979.
- [117] M. POLLAK AND D. SIEGMUND. A diffusion process and its applications to detecting a change in the drift of a Brownian motion. *Biometrika*, 72 :267–280, 1985.
- [118] P.–A. RAVIART. Particle approximation of first order systems. *Journal of Computational Mathematics*, 4(1) :50–61, 1986.
- [119] B. RISTIC, S. ARULAMPALAM, AND N. GORDON. *Beyond the Kalman Filter : Particle Filters for Tracking Applications*. Artech House, 2004.
- [120] G. G. ROUSSAS. *Contiguity of probability measures : some applications in statistics*, volume 63 of *Cambridge Tracts in Mathematics and Mathematical Physics*. Cambridge University Press, London, 1972.
- [121] A. SEN AND M. S. SRIVASTAVA. On tests for detecting change in mean. *The Annals of Statistics*, 3 :98–108, 1975.
- [122] A. N. SHIRYAEV. The problem of quickest detection of a violation of stationary behavior. *Dokl. Akad. Nauk SSSR*, 138 :1039–1042, 1961.
- [123] A. N. SHIRYAEV. On optimum methods in quickest detection problems. *Theory of Probability and its Applications*, VIII(1) :22–46, 1963.
- [124] A. N. SHIRYAEV. Some exact formulas in a disorder problem. *Theory of Probability and its Applications*, X(3) :348–354, 1965.
- [125] A. N. SHIRYAEV. Minimax optimality of the method of cumulative sums (CUSUM) in the case of continuous time. *Russian Mathematical Surveys*, 51 :750–751, 1996.
- [126] A. N. SHIRYAEV AND V. SPOKOINY. *Statistical experiments and decisions : asymptotic theory*. World Scientific, Singapore, 2000.
- [127] A. N. SHIRYAYEV. *Optimal Stopping Rules*, volume 8 of *Applications of Mathematics*. Springer-Verlag, New York, 1978.
- [128] A. G. TARTAKOVSKII. Optimal detection of changing properties of random sequences, I : Nonsequential detection. *Automation and Remote Control*, 47(7) :918–926, 1987.
- [129] A. G. TARTAKOVSKII. Optimal detection of changing properties of random sequences, II : Sequential detection. *Automation and Remote Control*, 48(8) :1059–1066, 1988.
- [130] L. TELKSNYS. Determination of the most probable changes in the properties of multivariate dynamical systems with unknown parameters. *Statist. Problemy Upravlenija—Trudy Sem. Processy Optimal. Upravlenija V Sekcija, Vyp.* 24 :9–26, 91, 1977.
- [131] L. TELKSNYS, editor. *Detection of changes in random processes*. Translation Series in Mathematics and Engineering. Optimization Software Inc. Publications Division, New York, 1986.

- [132] J. K. TUGNAIT. Adaptive estimation and identification for discrete systems with Markov jump parameters. *IEEE Transactions on Automatic Control*, 27(5) :1054–1065, 1982.
- [133] J. K. TUGNAIT AND A. H. HADDAD. A detection-estimation scheme for state estimation in switching environments. *Automatica*, 15(4) :477–481, 1979.
- [134] A. W. VAN DER VAART. *Asymptotic statistics*. Cambridge University Press, 1998.
- [135] A. W. VAN DER VAART. The statistical work of Lucien Le Cam. *The Annals of Statistics*, 30(3) :631–682, June 2002. Issue dedicated to the memory of Lucien Le Cam.
- [136] M. H. VELLEKOOP. *Rapid Detection and Estimation of Abrupt Changes by Non-linear Filtering*. PhD thesis, Imperial College, 1998.
- [137] M. H. VELLEKOOP AND J. M. C. CLARK. A nonlinear filtering approach to changepoint detection problems : direct and differential-geometric methods. *SIAM Journal on Control and Optimization*, 42(2) :469–494 (electronic), 2003.
- [138] A. WALD. Tests of statistical hypothesis concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3) :426–482, 1943.
- [139] A. S. WILLSKY AND H. L. JONES. A generalized likelihood ratio approach to state estimation in linear systems subject to abrupt changes. In *Proceedings of the 23th Conference on Decision and Control*, 1974.
- [140] A. S. WILLSKY AND H. L. JONES. A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic Control*, AC–21(1) :108–112, 1976.
- [141] C.-F. J. WU. On the convergence properties of the EM algorithm. *The Annals of Statistics*, 11(1) :95–103, 1983.
- [142] S. ZACKS. Survey of classical and Bayesian approaches to the change-point problem : fixed sample and sequential procedures of testing and estimation. In *Recent advances in statistics*, pages 245–269. Academic Press, New York, 1983.
- [143] S. ZACKS. Detection and change-point problems. In B. K. Ghosh and P. K. Sen, editors, *Handbook of sequential analysis*, pages 531–562. Marcel Dekker, New York, 1991.

CHAPITRE II

UN PROBLÈME DE CONTRÔLE STOCHASTIQUE

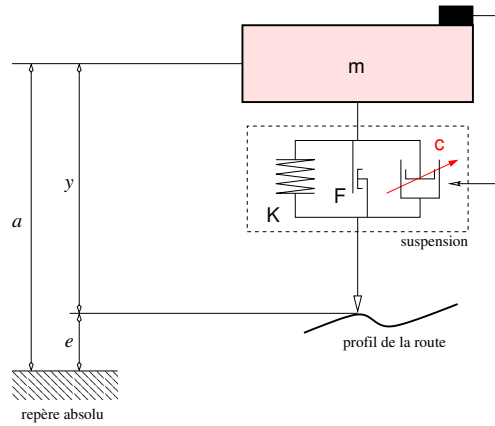
 mon arrivée à l'Inria Sophia–Antipolis, Étienne Pardoux me propose de travailler avec nos “alter ego mécaniciens”, Sergio Bellizzi et Robert Bouc du Laboratoire de Mécanique et d'Acoustique (CNRS Marseille), sur un contrat proposé par Renault. Le but était de développer des lois de commande pour amortisseurs semi–actifs. Nous nous sommes intéressés aux lois de commande optimale qu'il a été possible de développer sur les modèles les plus simples. Sergio Bellizzi et Robert Bouc ont étudié des modèles plus réalistes sur lesquels il n'est pas possible d'établir de résultat d'optimalité. Pour l'approche “optimale” nous avons été conseillés par Jean–Pierre Quadrat, un des spécialistes du domaine.

C'est une étude exemplaire à plus d'un titre. D'une part, cette collaboration avec des mécaniciens nous a permis de travailler sur un contrat de longue haleine (cinq années) et de couvrir un spectre théorie/application très large : Renault a finalement développé des prototypes testés sur des tables vibrantes. D'autre part, même pour un modèle simple du point de vue de la mécanique, l'équation différentielle stochastique considérée est non uniformément elliptique (une seule composante est bruitée) et à coefficients discontinus mais présentant une discontinuité simple. Cela en fait un problème particulièrement intéressant, nécessitant quelques outils évolués de calcul stochastique.

II.1 PRÉSENTATION DU PROBLÈME

II.1.1 UN EXEMPLE : UN MODÈLE DE SYSTÈME DE SUSPENSIONS SEMI–ACTIVES

Dans cette section, on décrit une technique de contrôle d'amortissement d'un système de suspension de véhicule, comprenant un ressort, un amortisseur, une masse et tenant compte des effets de frottement sec (cf. Figure II.1). Le but est d'améliorer le confort de roulement. Parmi les différentes alternatives aux systèmes classiques de suspension (systèmes passifs), on distingue les techniques actives des techniques semi–actives. Les procédés actifs génèrent des forces externes modulées selon une certaine loi de commande. Ces systèmes sont en général coûteux et complexes ; voir les survols GOODDALL & KORTUM [13] et HROVAT [16]. Les systèmes semi–actifs font appel à des technologies nettement moins coûteuses et plus



a_t	déplacement absolu de la masse m
y_t	déplacement relatif ($y = a - e$)
e_t	entrée aléatoire (profil de la route)
m	masse amortie
c	coefficient d'amortissement (contrôlé)
K	coefficient de raideur
F	constante de frottement sec

FIG. II.1 – Modèle à 1 degré de liberté.

simples. Ils agissent uniquement sur les lois d'amortissement ou de rappel. Ils ne peuvent donc que dissiper ou emmagasiner l'énergie. On présente ici un procédé de contrôle de la loi d'amortissement permettant de moduler le diamètre de l'orifice par lequel le fluide de l'amortisseur s'écoule ; voir ALANOLY & SAKAR [1] et TAKAHASHI *et al* [21]. On traite le cas à 1 degré de liberté. Ce modèle a été repris par HEINRICHER & F. MARTINS [14] et MENDIONDO & STOCKBRIDGE [19].

L'équation du mouvement pour un modèle à 1 degré de liberté est (cf. figure II.1 pour la définition exacte de chacun des termes) :

$$m \ddot{y}_t + c \dot{y}_t + K y_t + F \text{signe}(\dot{y}_t) = -m \ddot{e}_t. \quad (\text{II.1})$$

Dans ce modèle, la force de frottement sec est de type frottement pur de Coulomb (voir BOUC [3]).

L'entrée aléatoire \ddot{e}_t modélise l'accélération du profil, non mesurée, de la surface de la route. La force de rappel $K y_t + F \text{signe}(\dot{y}_t)$ est composée d'une première partie linéaire $K y_t$ et d'une seconde non linéaire $F \text{signe}(\dot{y}_t)$, décrivant la force de frottement sec. La force d'amortissement $c \dot{y}_t$ où $c > 0$ est le coefficient d'amortissement instantané. On supposera que le contrôle agit sur ce dernier terme.

Le but est de calculer une loi de commande en boucle fermée $c = c(y_t, \dot{y}_t) \in [c_{\min}, c_{\max}]$ (c_{\min} et c_{\max} correspondent à l'ouverture minimale et maximale de l'orifice réglant le débit d'huile) telle que la solution du système (II.1) minimise un critère lié au confort dynamique :

$$\overline{|\ddot{a}|^2} = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \int_0^T |\ddot{a}_t|^2 dt,$$

notons que $\ddot{a}_t = \ddot{y}_t + \ddot{e}_t$. \ddot{e}_t est un bruit blanc gaussien, i.e. $\ddot{e}_t = -\sigma \frac{d}{dt} W_t$ où W_t est un processus de Wiener standard.

Posons :

$$u \triangleq \frac{c}{m}, \alpha \triangleq \frac{K}{m}, \beta \triangleq \frac{F}{m} \text{ et } X_t \triangleq \begin{pmatrix} y_t \\ \dot{y}_t \end{pmatrix}.$$

avec $\alpha > 0, \beta > 0$ et u est à valeurs dans $U = [u_{\min}, u_{\max}] \subset]0, +\infty[$.

II.1.2 UN PROBLÈME DE CONTRÔLE STOCHASTIQUE

Le processus $X_t = \begin{pmatrix} y_t \\ \dot{y}_t \end{pmatrix}$, à valeurs dans \mathbb{R}^2 , est une solution de l'équation différentielle stochastique suivante :

$$dX_t = b(u(X_t), X_t) dt + \begin{pmatrix} 0 \\ \sigma \end{pmatrix} dW_t \quad (\text{II.2})$$

où

$$b(u, x) \triangleq \begin{pmatrix} b_1(u, x) \\ b_2(u, x) \end{pmatrix} = \begin{pmatrix} x_2 \\ -u x_2 - \alpha x_1 - \beta \text{signe}(x_2) \end{pmatrix}$$

et W_t est un processus de Wiener standard ; $x \mapsto u(x)$ est un contrôle en boucle fermée à valeurs dans U .

On considère une fonctionnelle de coût de type ergodique :

$$J(u) \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \int_0^T f(u(X_t), X_t) dt \quad (\text{II.3})$$

avec $f(u, x) \triangleq [b_2(u, x)]^2$. On adopte également la notation $b^u(x) \triangleq b(u(x), x)$, $f^u(x) \triangleq f(u(x), x)$...

On montrera plus loin que l'EDS (II.2) admet une mesure invariante et qu'elle est unique. La fonction coût (II.3) est donc bien définie.

Le problème de contrôle stochastique non linéaire consiste à déterminer, parmi une classe \mathcal{U} de lois de commande en boucle fermée, une loi \hat{u} qui minimise la fonction coût (II.3) sous les contraintes (II.2).

Une analyse superficielle de ce problème, ainsi que l'approximation numérique obtenue en pratique, conduit à considérer cette classe de lois de contrôle :

$$\mathcal{U} \triangleq \{u : \mathbb{R}^2 \mapsto [u_{\min}, u_{\max}] \text{ continue sur } \mathbb{R} \times \mathbb{R}_*\}$$

où $\mathbb{R}_* = \mathbb{R} \setminus \{0\}$.

II.1.3 LA MESURE DE PROBABILITÉ INVARIANTE

La fonction coût (II.3) s'écrit :

$$J(u) = \langle f^u, \mu_u \rangle = \int_{\mathbb{R}^2} f^u(x) \mu_u(dx), \quad \forall u \in \mathcal{U}, \quad (\text{II.4})$$

où μ_u est la mesure de probabilité invariante du système (II.2).

Dans CAMPILLO *et al* [8] on montre l'existence et l'unicité de la mesure invariante :

Existence : D'une part on montre que $\mathbb{E}|X_t|^2$ est borné uniformément en $t \geq 0$ et $u \in \mathcal{U}$. D'autre part on montre que le processus X_t est fellérien, i.e. l'application $x \rightarrow \mathbb{E}\phi(X_t^x)$ est continue pour tout $u \in \mathcal{U}$, $t \geq 0$ et $\phi \in C_b(\mathbb{R}^2)$; X_t^x désigne la solution du système (II.2) issue de x à l'instant $t = 0$.

Cette dernière propriété se démontre par une technique de changement de probabilité consistant à "supprimer" les termes de dérive discontinus (contenant la fonction signe et le contrôle u). Après changement de probabilité le système est linéaire gaussien.

Le contrôle uniforme du moment d'ordre 2 et la propriété de Feller impliquent que le système (II.2) admet une mesure invariante (cf. ETHIER & KURTZ [11, th. 9.3 ch. 4]).

Unicité : Soit μ une mesure de probabilité invariante du système et X_t la solution de ce système avec loi initiale μ .

On fait appel à la même technique de changement de loi : sous cette loi X_t admet une densité strictement positive sur tout \mathbb{R}^2 (et non pas p.p.). Cette propriété se conserve en revenant à la loi initiale. Ainsi μ admet une densité $p(x)$ strictement positive pour tout x p.p.

Ainsi, s'il existe deux mesures de probabilité invariantes, elles doivent être nécessairement équivalentes. Il existe donc au plus une mesure de probabilité invariante extrémale, ce qui implique l'unicité.

La mesure invariante est la solution de l'équation de Fokker–Planck stationnaire suivante écrite au sens faible :

$$\langle \mathcal{L}_u \phi, \mu_u \rangle = 0, \quad \forall \phi \in C_b(\mathbb{R}^2), \quad (\text{II.5})$$

où \mathcal{L}_u est le générateur infinitésimal de la diffusion X_t , i.e.

$$\mathcal{L}_u \phi(x) \triangleq x_2 \frac{\partial \phi(x)}{\partial x_1} + b_2^u(x) \frac{\partial \phi(x)}{\partial x_2} + \frac{\sigma^2}{2} \frac{\partial^2 \phi(x)}{\partial x_2^2}. \quad (\text{II.6})$$

En fait μ_u admet une densité $q_u(x)$ solution de l'EDP au sens fort : $\mathcal{L}_u^* q_u(x) = 0$, pour tout $x \in \mathbb{R}^2$, où \mathcal{L}_u^* désigne l'adjoint de l'opérateur \mathcal{L}_u .

II.2 APPROCHE OPTIMALE

II.2.1 L'ÉQUATION DE HAMILTON–JACOBI–BELLMAN (HJB)

On cherche une fonction $v : \mathbb{R}^2 \mapsto \mathbb{R}$ (fonction valeur) définie à une constante additive près, et une constante $\rho > 0$ telles que :

$$\min_{u \in [u_{\min}, u_{\max}]} [\mathcal{L}_u v(x) + f(u, x)] = \rho, \quad \forall x \in \mathbb{R}^2. \quad (\text{II.7})$$

Le contrôle optimal est alors défini par :

$$\hat{u}(x) = \text{Arg} \min_{u \in [u_{\min}, u_{\max}]} [\mathcal{L}_u v(x) + f(u, x)]$$

et $J(\hat{u}) = \rho$.

II.2.2 APPROXIMATION

On utilise une technique d'approximation classique : au lieu de discrétiser directement l'équation de HJB (II.7), on approche le processus X_t par un processus de Markov contrôlé en temps continu et à espace d'état discret. On résout alors exactement le problème de contrôle stochastique associé.

ÉTAPE 1 : ESPACE DISCRET

Sur la grille :

$$\mathbb{R}_h^2 \triangleq \left\{ x \in \mathbb{R}^2 ; x = (n_1 h_1 + \frac{1}{2} h_1, n_2 h_2 + \frac{1}{2} h_2), n_1, n_2 \in \mathbb{Z} \right\}$$

avec $h \triangleq (h_1, h_2)$ et $h_i > 0$, on considère les schémas aux différences finies décentrées suivants :

$$\frac{\partial \phi(x)}{\partial x_i} \simeq \begin{cases} \frac{\phi(x + e_i h_i) - \phi(x)}{h_i}, & \text{si } b_i^u(x) > 0, \\ \frac{\phi(x) - \phi(x - e_i h_i)}{h_i}, & \text{si } b_i^u(x) < 0, \end{cases} \quad (i = 1, 2) \quad (\text{II.8})$$

$$\frac{\partial^2 \phi(x)}{\partial x_2^2} \simeq \frac{\phi(x + e_2 h_2) - 2\phi(x) + \phi(x - e_2 h_2)}{h_2^2}, \quad (\text{II.9})$$

où e_i est le vecteur unité de \mathbb{R}^2 dans la $i^{\text{ème}}$ coordonnée.

Ainsi le générateur \mathcal{L}_u est approché par la matrice $\mathcal{L}_u^h \in \mathbb{R}^N \times \mathbb{R}^N$ suivante :

$$\mathcal{L}_u \phi(x) \simeq \mathcal{L}_u^h \phi(x) \triangleq \sum_{y \in \mathbb{R}_h^2} \mathcal{L}_u^h(x, y) \phi(y), \quad \forall x \in \mathbb{R}_h^2.$$

Le choix du schéma aux différences (II.8)–(II.9) implique que \mathcal{L}_u^h est le générateur infinitésimal d'un processus markovien de sauts X_t^h en temps continu et à espace d'état discret \mathbb{R}_h^2 (voir LE GLAND [12] et KUSHNER [18]). On définit $\{\xi_k^h\}_{k \in \mathbb{N}}$ la chaîne de Markov incluse associée.

On aboutit ainsi à un problème de contrôle stochastique ergodique pour un processus de Markov de générateur \mathcal{L}_u^h . La fonction coût est :

$$J_h(u) \triangleq \lim_{T \rightarrow \infty} \mathbb{E} \frac{1}{T} \int_0^T f^u(X_t^h) dt = \sum_{x \in \mathbb{R}_h^2} f^u(x) \mu_u^h(x), \quad (\text{II.10})$$

où u est un élément de la classe \mathcal{U}_h qui est simplement l'ensemble des applications de \mathbb{R}_h^2 dans $[u_{\min}, u_{\max}]$.

μ_u^h est la mesure de probabilité invariante du processus X_t^h . Cette mesure invariante est la solution du système linéaire :

$$\begin{cases} \sum_{y \in \mathbb{R}_{h,D}^2} \mathcal{L}_u^{h,D}(y, x) \mu_u^{h,D}(y) = 0, & \forall x \in \mathbb{R}_{h,D}^2, \\ \sum_{y \in \mathbb{R}_{h,D}^2} \mu_u^{h,D}(y) = 1. \end{cases} \quad (\text{II.11})$$

ÉTAPE 2 : ESPACE FINI

Soit $D \triangleq [-\bar{x}_1, \bar{x}_1] \times [-\bar{x}_2, \bar{x}_2]$ avec $\bar{x}_i > 0$, on définit la grille finie :

$$\mathbb{R}_{h,D}^2 \triangleq \mathbb{R}_h^2 \cap D \quad \text{et} \quad N \triangleq |\mathbb{R}_{h,D}^2|. \quad (\text{II.12})$$

On choisit des conditions aux bords de type Neumann ; en fait D est choisi tel que le processus n'atteigne que rarement le bord. Ce choix influe peu sur les résultats à condition de l'effectuer de telle façon que tous les états communiquent.

On obtient ainsi une approximation $\mathcal{L}_u^{h,D}$ de \mathcal{L}_u^h sous forme d'une matrice de taille $N \times N$. Il s'agit en fait du générateur infinitésimal d'un processus de Markov contrôlé $X_t^{h,D}$ en temps continu et à espace d'état fini ; $\xi_k^{h,D}$ désigne la chaîne de Markov associée.

La fonction coût est :

$$J_{h,D}(u) \triangleq \lim_{T \rightarrow \infty} \mathbb{E} \frac{1}{T} \int_0^T f^u(X_t^{h,D}) dt = \sum_{x \in \mathbb{R}_{h,D}^2} f^u(x) \mu_u^{h,D}(x),$$

où $\mu_u^{h,D}$ est la mesure de probabilité invariante du processus $X_t^{h,D}$. Cette mesure invariante est la solution du système linéaire :

$$\begin{cases} \sum_{y \in \mathbb{R}_{h,D}^2} \mathcal{L}_u^{h,D}(y, x) \mu_u^{h,D}(y) = 0, & \forall x \in \mathbb{R}_{h,D}^2, \\ \sum_{y \in \mathbb{R}_{h,D}^2} \mu_u^{h,D}(y) = 1 \end{cases}$$

où $u \in \mathcal{U}_{h,D}$, l'ensemble des applications $\mathbb{R}_{h,D}^2 \mapsto [u_{\min}, u_{\max}]$.

L'équation de Hamilton–Jacobi–Bellman associée à ce problème est :

$$\min_{u \in [u_{\min}, u_{\max}]} \left(\sum_{y \in \mathbb{R}_{h,D}^2} \mathcal{L}_u^{h,D}(x, y) v(y) + f^u(x) \right) = \rho, \quad \forall x \in \mathbb{R}_{h,D}^2, \quad (\text{II.13})$$

où ρ est une constante positive et $v : \mathbb{R}_{h,D}^2 \mapsto \mathbb{R}$ (i.e. $v \in \mathbb{R}^N$) est définie à une constante additive près. Cette équation a été étudiée dans BELLIZZI *et al* [2] à h et D fixés, elle admet une solution unique $(v, \rho) \in \mathbb{R}^N \times \mathbb{R}^+$.

II.2.3 RÉSOLUTION DU PROBLÈME APPROCHÉ

Il existe deux algorithmes itératifs pour résoudre (II.13) : le premier est implicite (itérations sur les politiques), le second explicite (itérations sur les valeurs) ; voir HOWARD [15], DOSHI [10].

Algorithme d'itération sur les politiques (ou de Howard). Partant d'une loi initiale $u^{(0)} \in \mathcal{U}_{h,D}$, l'itération $u^{(j)} \rightarrow u^{(j+1)}$ se fait en deux étapes. D'abord on résout un système linéaire : à partir de $u^{(j)}$ on calcule $(v^{(j)}, \rho^{(j)})$ la solution (v, ρ) du système linéaire :

$$\sum_{y \in \mathbb{R}_{h,D}^2} \mathcal{L}_{u^{(j)}}^{h,D}(x, y) v(y) + f^{u^{(j)}}(x) = \rho, \quad \forall x \in \mathbb{R}_{h,D}^2 \quad (\text{II.14})$$

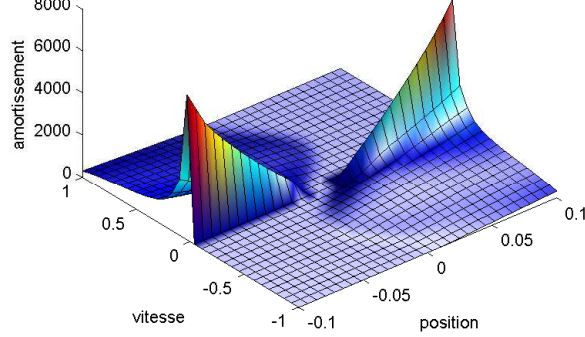


FIG. II.2 – Loi de commande optimale approchée (i.e. solution de l'équation (II.13)).

en ρ et $v(x)$, pour tout $x \in \mathbb{R}_{h,D}^2$ sauf pour un x_0 pour lequel on pose, par exemple, $v(x_0) = 0$. On résout ensuite N problèmes d'optimisation : à partir de v^j on calcule $u^{(j+1)}$ en résolvant

$$u^{(j+1)}(x) \in \text{Arg} \min_{u \in [u_{\min}, u_{\max}]} \left(\sum_{y \in \mathbb{R}_{h,D}^2} \mathcal{L}_u^{h,D}(x, y) v^{(j)}(y) + f^u(x) \right) \quad (\text{II.15})$$

pour tout $x \in \mathbb{R}_{h,D}^2$. Ces optimisations sont non linéaires et sont résolues par des méthodes itératives.

Algorithme d'itération sur les valeurs. L'équation de Hamilton–Jacobi–Bellman discrète peut être résolue à l'aide d'un schéma explicite consistant à faire apparaître artificiellement un "temps". Ce schéma s'écrit

$$\frac{v_{k+1}(x) - v_k(x)}{\delta} = \min_{u \in U} \sum_{y \in \mathbb{R}_{h,D}^2} \mathcal{L}_u(x, y) v_k(y) + f_u(x), \quad \forall x \in \mathbb{R}_{h,D}^2,$$

où $\delta > 0$ est un paramètre convenablement choisi et $v_0(x)$ est donné. On a la propriété suivante $v_k(x) \rightarrow v(x)$ et $\frac{1}{\delta}(v_{k+1}(x) - v_k(x)) \rightarrow \rho$ quand $k \rightarrow \infty$.

Le second algorithme est très simple à mettre en œuvre et demande peu de place mémoire. Mais comme ce schéma est explicite, une condition de stabilité impose de prendre δ suffisamment petit. La convergence peut être très lente.

Dans BELLIZZI *et al* [2] on a montré que l'équation de HJB (II.13) admet une unique solution et que l'algorithme de Howard (II.14)–(II.15) converge vers cette solution. L'approximation du contrôle optimal obtenue est présentée dans la Fig. II.2.

II.2.4 CONVERGENCE DE LA FONCTION COÛT

On pose $h_1 = h_2 = h$. On montre que

$$J_{h,D}(u) \rightarrow J(u) \text{ lorsque } h \downarrow 0 \text{ et } D \uparrow \mathbb{R}^2$$

pour tout $u \in \mathcal{U}$ où $h \delta_h \rightarrow 0$ et $\delta_h = \text{radius}(D_h)$. Notons que $u|_{\mathbb{R}^2_{h,D_h}} \in \mathcal{U}_{h,D_h}$. Ce résultat a été démontré dans CAMPILLO *et al* [8].

II.3 APPROCHE SOUS-OPTIMALE

Ici on cherche la commande "optimale" parmi une classe donnée de commandes paramétrées pour le système (II.2). On considère donc le système :

$$dX_t(\theta) = b(\theta, X_t(\theta)) dt + \begin{pmatrix} 0 \\ \sigma \end{pmatrix} dW_t \quad (\text{II.16})$$

où

$$b(\theta, x) \triangleq \begin{pmatrix} b_1(\theta, x) \\ b_2(\theta, x) \end{pmatrix} = \begin{pmatrix} x_2 \\ -u(\theta, x) x_2 - \alpha x_1 - \beta \text{signe}(x_2) \end{pmatrix}.$$

Le contrôle paramétré $u : \Theta \times \mathbb{R}^2 \rightarrow U = [u_{\min}, u_{\max}]$ est tel que pour tout $x \in \mathbb{R}^2$, $\theta \rightarrow u(\theta, x)$ est C^1 , et tel que, pour tout $\theta \in \Theta$, $x \rightarrow u(\theta, x) x_2$ est C^1 sur $\mathbb{R} \times \mathbb{R}_*$ avec dérivées bornées. Ainsi u peut être discontinu en $x_2 = 0$, ce qui est le cas du contrôle optimal. Par exemple :

$$u(\theta, x) = \theta^1 + [-\theta^2 x_1 \text{signe}(x_2)]^+$$

où θ^1 et θ^2 sont positifs.

Notons que dans (II.16), le coefficient de dérive est C^1 sur $\mathbb{R} \times \mathbb{R}_*$, et discontinu en $x_2 = 0$ avec un saut d'amplitude -2β .

On considère la fonction coût :

$$J(\theta) \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \int_0^T f(\theta, X_t(\theta)) dt \quad (\text{II.17})$$

avec $f(\theta, x) \triangleq [b_2(\theta, x)]^2 = [u(\theta, x) x_2 + \alpha x_1 + \beta \text{signe}(x_2)]^2$.

Dans CAMPILLO & PARDOUX [9] on montre que (II.16) admet une solution au sens fort. Pour cela on utilise le théorème de Yamada–Watanabe (voir Karatzas & Shreve [17, Prop. V–3.20]) : on montre l'existence d'une solution au sens faible avec une technique de Girsanov, et l'unicité au sens trajectoirel.

Ce point est important : cela signifie que dans (II.17) l'espérance ne dépend pas de θ , ce qui permet de calculer la dérivée de la fonction coût comme ci-dessous.

II.3.1 ALGORITHMES D'APPROXIMATION STOCHASTIQUE

Le but est de déterminer θ^* qui minimise la fonction coût (II.17). On a déjà montré que :

$$J(\theta) = \int_{\mathbb{R}^2} f(\theta, x) \mu_\theta^X(dx) \quad (\text{II.18})$$

où μ_θ^X est l'unique mesure invariante du processus $X_t(\theta)$. Il est possible de faire appel à un algorithme de type Kiefer–Wolfowitz, ne nécessitant pas le calcul du gradient de la fonction coût, mais les algorithmes de gradient ont de meilleures propriétés de convergence. On calcule maintenant le gradient de la fonction coût. Supposons que θ est de dimension 1.

LE PROCESSUS DÉRIVÉ $\dot{X}_t(\theta) \triangleq \partial_\theta X_t(\theta)$

Le calcul de $\dot{X}_t(\theta)$ fait appel au temps local L_t de $X_t^2(\theta)$ en 0. D'après PROTTER [20] ce temps local existe et vérifie la formule de Tanaka :

$$|X_t^2| = |X_0^2| + \int_0^t \text{signe}(X_s^2) dX_s^2 + 2L_t. \quad (\text{II.19})$$

Pour tout $t > 0$, $X_t(\theta)$ est dérivable en moyenne quadratique par rapport au paramètre θ_i ($i = 1, \dots, d$) et $\dot{X}_t(\theta)$ est la solution de :

$$d\dot{X}_t(\theta) = A(\theta, X_t(\theta)) \dot{X}_t(\theta) dt + B \dot{X}_t(\theta) dL_t + C(\theta, X_t(\theta)) dt, \quad \dot{X}_0(\theta) = 0 \quad (\text{II.20})$$

avec

$$A(\theta, x) \triangleq \begin{pmatrix} 0 & 1 \\ -\alpha - \dot{u}_{x_1}(\theta, x) x_2 & -u(\theta, x) - \dot{u}_{x_2}(\theta, x) x_2 \end{pmatrix} \\ B \triangleq \begin{pmatrix} 0 & 0 \\ 0 & -\frac{2\beta}{\sigma^2} \end{pmatrix}, \quad C(\theta, x) \triangleq \begin{pmatrix} 0 \\ -\dot{u}_\theta(\theta, x) x_2 \end{pmatrix}.$$

où L_t est le temps local de X_t^2 en 0.

On peut explicitement exprimer $\dot{X}_t(\theta)$ en fonction de $X_t(\theta)$ en introduisant la résolvante $\Phi_t(\theta)$ solution de :

$$d\Phi_t(\theta) = A(\theta, X_t(\theta)) \Phi_t(\theta) dt + B \Phi_t(\theta) dL_t, \quad \Phi_0(\theta) = I, \quad (\text{II.21})$$

alors

$$\dot{X}_t(\theta) = \Phi_t(\theta) \int_0^t \Phi_s(\theta)^{-1} C(\theta, X_s(\theta)) ds. \quad (\text{II.22})$$

PROPRIÉTÉS ASYMPTOTIQUES DE $(X_t(\theta), \dot{X}_t(\theta))$

Posons :

$$\xi_t(\theta) = \begin{pmatrix} X_t(\theta) \\ \dot{X}_t(\theta) \end{pmatrix}. \quad (\text{II.23})$$

Il a déjà été montré que $X_t(\theta)$ admet une unique mesure invariante μ_θ^X . On étend le processus $X_t(\theta)$ à tout $t \in \mathbb{R}$, de telle sorte qu'il soit stationnaire de loi μ_θ^X . On peut alors résoudre (II.21) et définir $\Phi_t(\theta)$ pour tout $t \in \mathbb{R}$.

En étudiant les propriétés d'ergodicité de $\Phi_t(\theta)$ à l'aide des travaux de BOUGEROL [4, 5], on montre que s'il existe $C > 0$ tel que pour tout (θ, x)

$$\alpha + \dot{u}_{x_1}(\theta, x) x_2 \geq C > 0, \\ u(\theta, x) + \dot{u}_{x_2}(\theta, x) x_2 \geq C > 0,$$

alors

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \log \|\Phi_{-t}(\theta)^{-1}\| \leq \lambda \text{ p.s.} \quad (\text{II.24})$$

pour un certain $\lambda < 0$. Ce qui implique que l'intégrale $\int_{-\infty}^0 \Phi_t(\theta)^{-1} C(\theta, X_t(\theta)) dt$ converge p.s. On en déduit que le processus de Markov $\xi_t(\theta)$ admet une unique mesure invariante μ_θ .

LE GRADIENT DE LA FONCTION COÛT

Posons :

$$g(\theta, x) \triangleq u(\theta, x) x_2 + \alpha x_1 + \beta \text{signe}(x_2)$$

ainsi $f(\theta, x) = [g(\theta, x)]^2$. On a :

$$\begin{aligned} \dot{f}_\theta(\theta, x) &= 2g(\theta, x) \dot{u}_\theta(\theta, x) x_2, \\ \dot{f}_{x_1}(\theta, x) &= 2g(\theta, x) [\dot{u}_{x_1}(\theta, x) x_2 + \alpha], \end{aligned}$$

et, si $\delta(x_2)$ désigne la mesure de Dirac en la variable x_2 au point 0, au sens des distributions :

$$\begin{aligned} \dot{f}_{x_2}(\theta, x) &= 2g(\theta, x) [\dot{u}_{x_2}(\theta, x) x_2 + u(\theta, x) + 2\beta \delta(x_2)] \\ &= 2g(\theta, x) [\dot{u}_{x_2}(\theta, x) x_2 + u(\theta, x)] + 4\alpha\beta x_1 \delta(x_2). \end{aligned}$$

Posons :

$$\begin{aligned} \tilde{f}_{x_1}(\theta, x) &= \dot{f}_{x_1}(\theta, x) \\ \tilde{f}_{x_2}(\theta, x) &= 2g(\theta, x) [\dot{u}_{x_2}(\theta, x) x_2 + u(\theta, x)] \end{aligned}$$

et

$$F(\theta, x, y) \triangleq \dot{f}_\theta(\theta, x) + \tilde{f}_x(\theta, x) y.$$

Formellement, si l'on dérive la fonction coût (II.17) par rapport à θ , et si on intervertit la dérivation et la limite $T \rightarrow \infty$, on aboutit à :

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T F(\theta, X_t(\theta), \dot{X}_t(\theta)) dt + \int_0^T 4\alpha\beta X_t^1(\theta) \dot{X}_t^2(\theta) \delta(X_t^2(\theta)) dt \right].$$

Ceci peut être démontré de façon rigoureuse, le gradient de la fonction coût est :

$$\nabla_\theta J(\theta) = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T F(\theta, X_t(\theta), \dot{X}_t(\theta)) dt + \int_0^T 4\alpha\beta X_t^1(\theta) \dot{X}_t^2(\theta) dL_t \right]. \quad (\text{II.25})$$

On montre que le processus $(X_t(\theta), \dot{X}_t(\theta))$ admet une unique mesure invariante μ_θ ce qui montre que la limite (II.25) est correctement définie.

ALGORITHME DE GRADIENT STOCHASTIQUE

Afin de minimiser (II.18), on cherche à déterminer $\theta^* \in \Theta$ tel que :

$$\nabla_\theta J(\theta)|_{\theta=\theta^*} = 0. \quad (\text{II.26})$$

L'algorithme de gradient stochastique associé est le suivant : pour $\Delta t > 0$ fixé on pose $t_k \triangleq k \Delta t$, et on résout les équations (II.16),(II.20) avec

$$\theta = \theta_k \text{ pour } t_k \leq t < t_{k+1},$$

et θ_k donné par

$$\theta_{k+1} = \theta_k - \rho_k \left[F(\theta_k, X_{t_k}(\theta_k), \dot{X}_{t_k}(\theta_k)) \Delta t + 4\beta\alpha X_{t_k}^1(\theta_k) \dot{X}_{t_k}^2(\theta_k) \Delta L_k \right], \quad (\text{II.27})$$

où

$$\Delta L_k = L_{t_{k+1}} - L_{t_k},$$

et où le gain ρ_k est une suite de termes strictement positifs tendant convenablement vers 0.

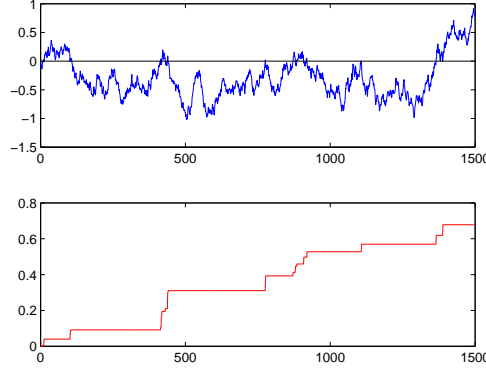


FIG. II.3 – Discrétisation du temps local L_t (courbe du bas) de la seconde composante X_t^2 du processus d'état (courbe du haut).

II.3.2 DISCRÉTISATION EN TEMPS

Pour tout $n \geq 1$ on pose $\Delta t = \frac{1}{n}$, $t_k = k \Delta t$, on approche X_{t_k} par \bar{X}_{t_k} à l'aide du schéma d'Euler :

$$\bar{X}_{t_{k+1}}^1 = \bar{X}_{t_k}^1 + \bar{X}_{t_k}^2 \Delta t, \quad (\text{II.28})$$

$$\bar{X}_{t_{k+1}}^2 = \bar{X}_{t_k}^2 - (u(\theta, \bar{X}_{t_k}) \bar{X}_{t_k}^2 + \alpha \bar{X}_{t_k}^1 + \beta \text{signe}(\bar{X}_{t_k}^2)) \Delta t + \sigma w_k \quad (\text{II.29})$$

où $w_k \triangleq W_{t_{k+1}} - W_{t_k} \sim \mathcal{N}(0, \Delta t)$.

On approche \dot{X}_t également par un schéma d'Euler :

$$\bar{\dot{X}}_{t_{k+1}}^1 = \bar{\dot{X}}_{t_k}^1 + \bar{\dot{X}}_{t_k}^2 \Delta t, \quad (\text{II.30})$$

$$\begin{aligned} \bar{\dot{X}}_{t_{k+1}}^2 = & \bar{\dot{X}}_{t_k}^2 - \left(\dot{u}_\theta(\theta, \bar{X}_{t_k}) + \dot{u}_x(\theta, \bar{X}_{t_k}) \bar{X}_{t_k} \right) \bar{X}_{t_k}^2 \Delta t \\ & - u(\theta, \bar{X}_{t_k}) \bar{X}_{t_k}^2 \Delta t - \alpha \bar{\dot{X}}_{t_k}^1 \Delta t - \frac{2\beta}{\sigma^2} \bar{\dot{X}}_{t_k}^2 \Delta \bar{L}_{t_k} \end{aligned} \quad (\text{II.31})$$

où $\Delta \bar{L}_{t_k}$ est une approximation de $L_{t_{k+1}} - L_{t_k}$.

Pour calculer cette approximation on note que, d'après (II.19)

$$L_t = \frac{1}{2}|X_t^2| - \frac{1}{2}|X_0^2| - \frac{1}{2} \int_0^t \text{signe}(X_s^2) dX_s^2. \quad (\text{II.32})$$

On approche L_t de la façon suivante : dans (II.32) on remplace X_t^2 par l'interpolation polygonale du processus en temps discret $\bar{X}_{t_k}^2$ donné par (II.29)

$$\bar{X}_t^2 \triangleq \sum_{k \geq 0} \frac{(t_{k+1} - t) \bar{X}_{t_k}^2 + (t - t_k) \bar{X}_{t_{k+1}}^2}{t_{k+1} - t_k} \mathbf{1}_{[t_k, t_{k+1}[}(t).$$

On obtient l'approximation :

$$\bar{L}_t = \frac{1}{2}|\bar{X}_t^2| - \frac{1}{2}|\bar{X}_0^2| - \frac{1}{2} \int_0^t \text{signe}(\bar{X}_{t(s)}^2) d\bar{X}_s^2, \quad (\text{II.33})$$

où $t(s) = t_k$ if $s \in [t_k, t_{k+1}[$. Un petit calcul conduit à la formule (voir CAMPILLO & PARDOUX [9]) :

$$\Delta \bar{L}_{t_k} = \bar{L}_{t_{k+1}} - \bar{L}_{t_k} \triangleq \begin{cases} |\bar{X}_{t_{k+1}}^2| & \text{si } \bar{X}_{t_{k+1}}^2 \bar{X}_{t_k}^2 < 0, \\ 0 & \text{sinon.} \end{cases} \quad (\text{II.34})$$

la Figure (II.3) donne un exemple de réalisation de cette approximation.

On a montré que ces schémas convergent en moyenne quadratique : pour tout $t \geq 0$ et $\theta \in \Theta$,

$$\bar{X}_t(\theta) \xrightarrow[n \rightarrow \infty]{L^2} X_t(\theta) \quad \bar{\dot{X}}_t(\theta) \xrightarrow[n \rightarrow \infty]{L^2} \dot{X}_t(\theta) \quad \bar{L}_t(\theta) \xrightarrow[n \rightarrow \infty]{L^2} L_t(\theta).$$

II.4 D'AUTRES MODÈLES

Le modèle précédent admet de nombreuses extensions. On présente maintenant des modèles à deux degrés de liberté (deux systèmes de suspension en parallèle ou en série). En plus du critère de confort présenté ici, nous avons également proposé une modélisation d'un critère de tenue de route. Enfin il est possible de faire appel à des modélisations plus "fines" de l'accélération du profil de la route ou bien du terme de frottement sec. Dans tous les cas cela conduit à une augmentation de la dimension de l'espace d'état.

2-DDL EN PARALLÈLE

Les équations du mouvement du modèle à deux degrés de liberté représenté à la Figure II.4 sont :

$$\begin{aligned} m \ddot{z} &= -c_1 (\dot{z} - l_1 \dot{\theta} - \dot{e}_1) - K_1 (z - l_1 \theta - e_1) - F_1 \text{signe}(\dot{z} - l_1 \dot{\theta} - \dot{e}_1) \\ &\quad - c_2 (\dot{z} + l_2 \dot{\theta} - \dot{e}_2) - K_2 (z + l_2 \theta - e_2) - F_2 \text{signe}(\dot{z} + l_2 \dot{\theta} - \dot{e}_2) \end{aligned}$$

et

$$\begin{aligned} I \ddot{\theta} &= -l_1 [-c_1 (\dot{z} - l_1 \dot{\theta} - \dot{e}_1) - K_1 (z - l_1 \theta - e_1) - F_1 \text{signe}(\dot{z} - l_1 \dot{\theta} - \dot{e}_1)] \\ &\quad + l_2 [-c_2 (\dot{z} + l_2 \dot{\theta} - \dot{e}_2) - K_2 (z + l_2 \theta - e_2) - F_2 \text{signe}(\dot{z} + l_2 \dot{\theta} - \dot{e}_2)] \end{aligned}$$

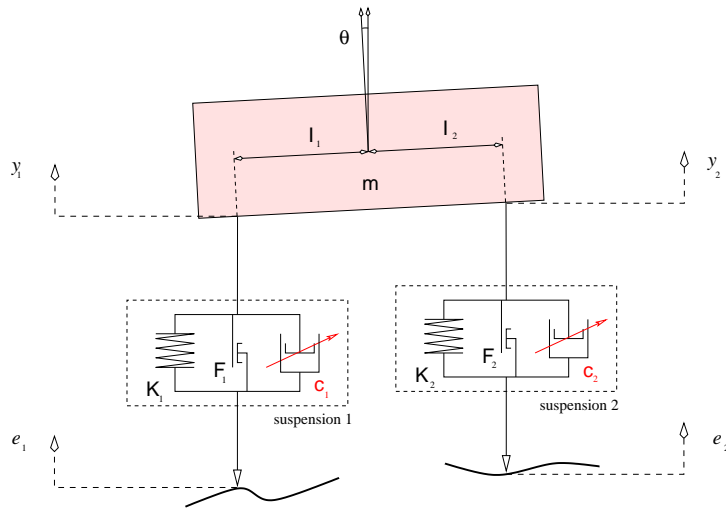
Pour modéliser le confort de roulement on prend en compte l'accélération verticale et l'accélération angulaire :

$$J = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \int_0^T [\rho_1 |\ddot{z}|^2 + \rho_2 |\ddot{\theta}|^2] dt.$$

où ρ_1 et ρ_2 sont des facteurs de pondération.

Soit x_i le déplacement relatif de la suspension i : $x_1 = z - l_1 \theta - e_1$ et $x_2 = z + l_2 \theta - e_2$. On suppose que \ddot{e}_1 et \ddot{e}_2 sont des bruits blancs gaussiens indépendants $\ddot{e}_i = -\sigma dW_i/dt$.

Le vecteur d'état est $X = [x_1, \dot{x}_1, x_2, \dot{x}_2]^*$.



z	déplacement absolu vertical de la masse m
θ	déplacement angulaire de la masse
m	masse
I	tenseur inertiel par rapport au centre de gravité
suspension i :	
K_i	raideur
F_i	coefficient de frottement sec
c_i	coefficient d'amortissement (variable)
l_i	distance au centre de gravité
e_i	entrée aléatoire

FIG. II.4 – Modèle à 2 DDL en parallèle

MODÈLE À 2DDL EN SÉRIE

Dans les modèles précédents la fonction coût ne tient compte que d'un critère de confort. Dans ce dernier exemple on tient compte à la fois du confort de roulement et de la tenue de route. Le modèle présenté à la Figure II.5 représente un quart de véhicule.

L'équation du mouvement donne :

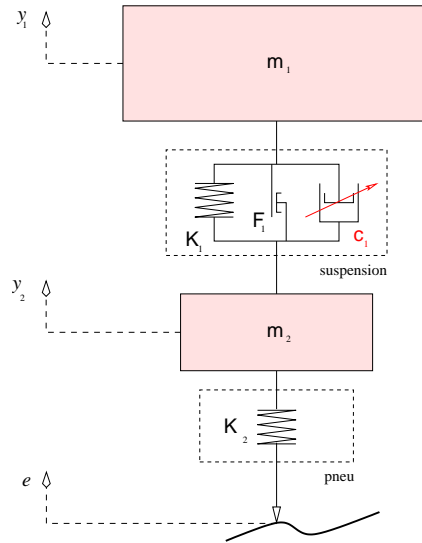
$$\begin{aligned} m_1 \ddot{y}_1 &= -c_1 (\dot{y}_1 - \dot{y}_2) - K_1 (y_1 - y_2) - F_1 \text{signe}(\dot{y}_1 - \dot{y}_2), \\ m_2 \ddot{y}_2 &= +c_1 (\dot{y}_1 - \dot{y}_2) + K_1 (y_1 - y_2) + F_1 \text{signe}(\dot{y}_1 - \dot{y}_2) - K_2 (y_2 - e). \end{aligned}$$

Posons $x_1 = y_1 - y_2$ et $x_2 = y_2 - e$, on obtient :

$$\begin{aligned} \ddot{x}_1 &= -\left(\frac{1}{m_1} + \frac{1}{m_2}\right) [c_1 \dot{x}_1 + K_1 x_1 + F_1 \text{signe}(\dot{x}_1)] + \frac{K_2}{m_2} x_2, \\ \ddot{x}_2 &= +\frac{1}{m_2} [c_1 \dot{x}_1 + K_1 x_1 + F_1 \text{signe}(\dot{x}_1)] - \frac{K_2}{m_2} x_2 - \ddot{e}. \end{aligned}$$

La fonction coût tient compte des deux termes :

- pour le confort : accélération du compartiment passager \ddot{y}_1 ,
- pour la tenue de route : le débattement du pneu $y_2 - e$,



m_1	quart de la masse suspendue
m_2	masse de la roue avec un demi-essieu
K_1	raideur
F_1	coefficient de frottement sec
c_1	coefficient d'amortissement (variable)
K_2	raideur du pneu

FIG. II.5 – Modèle à 2DDL en série

ainsi que d'une contrainte technique :

– l'intervalle de débattement de la suspension $y_1 - y_2$.

On choisit :

$$J = \lim_{T \rightarrow \infty} \frac{1}{T} E \int_0^T [|\dot{y}_1|^2 + \rho_1 \psi_1(y_2 - e) + \rho_2 \psi_2(y_1 - y_2)] dt$$

avec par exemple : $\psi_1(y) = y^2$ et $\psi_2(y) = 0$ pour $|y| \leq R$, et $\psi_2(y) = |y - R|$ for $|y| > R$, où R est borné par la course maximale de la suspension. ρ_1 and ρ_2 sont des constantes de pondération gérant le compromis confort/tenue de route.

On suppose que \ddot{e} est un bruit blanc gaussien standard $\ddot{e} = -\sigma dW_t/dt$. La variable d'état est $X = [x_1, \dot{x}_1, x_2, \dot{x}_2]^*$.

BIBLIOGRAPHIE

- [1] J. ALANOLY AND S. SANKAR. Semi-active force generators for shock isolation. *Journal of Sound and Vibration*, 126(1) :145–156, 1988.
- [2] S. BELLIZZI, F. CAMPILLO, R. BOUC, AND É. PARDOUX. Contrôle optimal semi-actif de suspension de véhicule. In A. Bensoussan and J. L. Lions, editors, *Analysis and Optimization of Systems, Antibes 1988*, volume 111 of *Lecture Notes in Control and Information Sciences*, pages 689–699. Springer-Verlag, 1988.
- [3] R. BOUC. Modèle mathématique d'hystérésis. *Acustica*, 23(3) :16–25, 1971.

- [4] P. BOUGEROL. Comparaison des exposants de Lyapounov des processus markoviens multiplicatifs. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 24(4) :439–489, 1988.
- [5] P. BOUGEROL. Théorèmes limite pour les systèmes linéaires à coefficients markoviens. *Probability Theory and Related Fields*, 78 :193–221, 1988.
- [6] F. CAMPILLO. Optimal ergodic control for a class of nonlinear stochastic systems. In *Proceedings of the 28th IEEE Conference on Decision and Control, Tampa*, pages 1190–1195. December 1989.
- [7] F. CAMPILLO, F. CÉROU, AND É. PARDOUX. Numerical method for ergodic stochastic control problems : Application to vehicle suspension systems. In *Proceedings of the 31st IEEE Conference on Decision and Control*, pages 2056–2061. Tucson, December 1992.
- [8] F. CAMPILLO, F. LE GLAND, AND É. PARDOUX. Approximation of a stochastic ergodic control problem. In J. Descusse, M. Fliess, A. Isidori, and D. Leborgne, editors, *New Trends in Nonlinear Control Theory, Nantes 1988*, volume 122 of *Lecture Notes in Control and Information Sciences*, pages 379–395. Springer-Verlag, 1989.
- [9] F. CAMPILLO AND É. PARDOUX. Numerical methods in ergodic optimal stochastic control and application. In I. Karatzas and D. Ocone, editors, *Applied Stochastic Analysis, Rutgers University 1991*, volume 177 of *Lecture Notes in Control and Information Sciences*, pages 59–73. Springer-Verlag, 1992.
- [10] B.T. DOSHI. Continuous time control of Markov processes on an arbitrary state space : average return criterion. *Stochastic Processes and their Applications*, 4 :55–77, 1976.
- [11] S. N. ETHIER AND T. G. KURTZ. *Markov Processes – Characterization and Convergence*. John Wiley & Sons, New York, 1986.
- [12] F. LE GLAND. *Estimation de Paramètres dans les Processus Stochastiques, en Observation Incomplète – Applications à un Problème de Radio-Astronomie*. Thèse de Docteur-Ingénieur, Université de Paris IX – Dauphine, 1981.
- [13] R. M. GOODDALL AND W. KORTUM. Active controls in ground transportation — A review of the state-of-the-art and future potential. *Vehicle Dystems Dynamics*, 12 :225–257, 1983.
- [14] A. C. HEINRICHER AND F. MARTINS. Optimal control for the local time of a diffusion : numerical approximation. In *Stochastic Dynamics and Reliability of Nonlinear Ocean Systems DE Vol. 77*, pages 7–14. ASME 1994
- [15] R. A. HOWARD. *Dynamic Programming and Markov Processes*. John Wiley & Sons, New York, 1960.
- [16] D. HROVAT. Survey of advanced suspension developments and related optimal control applications. *Automatica*, 33(10) :1781–1817, October 1997.
- [17] I. KARATZAS AND S. E. SHREVE. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1988.
- [18] H. J. KUSHNER. *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, volume 129 of *Mathematics in Science and Engineering*. Academic Press, New York, 1977.
- [19] M. S. MENDIONDO AND R. H. STOCKBRIDGE. Long term average control of a local time process. In Z. Hou and J. Filar and A. Chen, editors, *Markov Processes and Controlled Markov Chains*, pages 425–441. Kluwer, 2002.
- [20] P. PROTTER. *Stochastic integration and differential equations : a new approach*, volume 21 of *Applications of Mathematics*. Springer-Verlag, 1990.

- [21] S. TAKAHASHI, T. KANEKO, AND K. TAKAHASHI. A damping force control method which reduce energy to the vehicle body. *JSAE Review*, 8(3) :95–98, 1987.

CHAPITRE III

HOMOGENÉISATION

Deux raisons ont été à l'origine de cette étude : un contrat industriel avec Benoît Nœtinger de l'IFP et une volonté de collaboration sur le thème de l'homogénéisation entre l'Inria Sophia–Antipolis et l'équipe de probabilités du LATP (Centre de Mathématiques et Informatique, Château–Gombert, Université de Provence). Au LATP, en plus d'Étienne Pardoux, Sergeï Kozlov était également très intéressé par ce projet. Le but initial était de développer des méthodes de Monte Carlo pour la calcul de coefficients effectifs de milieux hétérogènes. Dans un premier temps à l'aide de chaînes de Markov sur des grilles, ensuite à l'aide de méthodes sans grille. La disparition prématurée de Sergeï Kozlov a profondément modifié ce projet. Étienne Pardoux s'est intéressé à l'approche probabiliste du thème. Durant cette période j'ai eu le plaisir de travailler avec des étudiants comme Elisabeth Remy ou Antoine Lejay, ainsi que de collaborer avec Andrey Piatnitski et Marina Kleptsina.

III.1 UN PROBLÈME D'HOMOGENÉISATION D'OPÉRATEUR PARABOLIQUE ALÉATOIRE

On présente essentiellement les résultats obtenus dans CAMPILLO, KLEPTSINA & PIATNITSKI [6] avec une extension présentée dans CAMPILLO, KLEPTSINA & PIATNITSKI [8]. Les résultats proposés dans CAMPILLO & PIATNITSKI [12] sont très techniques et ne sont pas présentés ici.

III.1.1 INTRODUCTION

Considérons un opérateur à coefficients fortement hétérogènes qui se comporte à grande échelle comme un opérateur de même nature mais à coefficients constants. Un tel opérateur présente donc des propriétés d'homogénéisation. On s'intéresse au problème d'homogénéisation pour des opérateurs aléatoires paraboliques avec terme elliptique symétrique dans le cas d'un grand potentiel, les coefficients sont des fonctions rapidement oscillantes aussi bien en espace qu'en temps.

Le problème d'homogénéisation pour différentes structures aléatoires a été très étudié dans la littérature physique comme mathématique, voir par exemple JIKOV *et al* [17] et sa bibliographie.

Dans le cas multidimensionnel les premiers résultats rigoureux pour des opérateurs aléatoires elliptiques sous forme divergence à coefficients aléatoires, ont notamment été obtenus par KOZLOV [19]. Une autre approche a été développée par PAPANICOLAOU & VARADHAN [31]. Plus tard, beaucoup d'autres structures aléatoires ont été étudiées, parmi lesquelles les milieux poreux aléatoires (voir par exemple HORNUNG [16]), les équations avec des termes d'ordre inférieur comme dans AVELLANEDA & MAJDA [2, 3], FANNJIANG & PAPANICOLAOU [13, 14], systèmes d'équations, modèles non linéaires... Une méthode efficace d'homogénéisation de structures aléatoires a été proposée dans BOURGEAT *et al* [5] où la technique développée plus tôt pour des micro-structures périodiques (ALLAIRE [1]) a été généralisée au cas aléatoire.

Actuellement, il y a plusieurs approches mathématiques qui permettent d'examiner des problèmes d'homogénéisation en milieux aléatoires, mais dans tous les modèles étudiés, l'aspect aléatoire des variables spatiales et la présence d'un groupe de transformation préservant une certaine mesure de probabilité, sont supposés.

Dans notre modèle nous supposons que le milieu possède une structure microscopique périodique tandis que la dynamique du système est aléatoire et, de plus, diffusive. Les équations sans potentiel ont été précédemment considérées dans KLEPTSZYNA & PIATNITSKI [18].

On considère l'opérateur parabolique :

$$u(t, x) \rightarrow \frac{\partial}{\partial t} u(t, x) - \operatorname{div} \left[a\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}\right) \nabla u(t, x) \right] - \frac{1}{\varepsilon^\beta} c\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}\right) u(t, x)$$

où ε est un petit paramètre positif, $\alpha, \beta > 0$, les coefficients $a^{ij}(z, y)$ et $c(z, y)$ sont périodiques en z et ξ_t est un processus de diffusion stationnaire.

Le paramètre α représente le rapport entre les longueurs d'échelle microscopique en espace et en temps ; dans le cas "auto-similaire" $\alpha = 2$, un couplage entre les moyennisations en espace et en temps apparaît. Si $\beta < \frac{\alpha}{2} \wedge 1$ alors la présence du potentiel $c(z, y)$ est négligeable dans un sens approprié. Si, au contraire, $\beta > \frac{\alpha}{2} \wedge 1$ alors, en général, la famille des solutions n'est pas compacte. La seule exception est le cas où $\alpha > 2$ et où la valeur moyenne de $c(z, y)$ en z est égale à zéro pour tout y . Dans ce cas, le choix approprié est $\beta = \frac{\alpha}{2}$ au lieu de $\beta = 1$. On traite donc le cas $\beta = \frac{\alpha}{2} \wedge 1$.

On s'intéresse au comportement asymptotique ($\varepsilon \downarrow 0$) de la solution du problème de Cauchy suivant :

$$\frac{\partial}{\partial t} u^\varepsilon(t, x) = \operatorname{div} \left[a\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}\right) \nabla u^\varepsilon(t, x) \right] + \frac{1}{\varepsilon^{1 \wedge \frac{\alpha}{2}}} c\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}\right) u^\varepsilon(t, x), \quad (\text{III.1})$$

avec $u^\varepsilon(0, x) = u_0(x)$, $x \in \mathbb{R}^n$, $t \in [0, T]$ et $\alpha > 0$. Les coefficients $a(z, y)$ et $c(z, y)$ sont périodiques en z i.e. z appartient au tore unité $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$.

Dans le cas particulier $\alpha = 1$, PARDOUX & PIATNITSKI [33] ont étendu le résultat présenté ci-dessous au cas d'un terme d'ordre 0 non linéaire dans le second membre de (III.1).

Le processus ξ_t est un processus de diffusion stationnaire, à valeurs dans \mathbb{R}^d , solution de :

$$d\xi_t = B(\xi_t) dt + \sigma(\xi_t) dW_t \quad (\text{III.2})$$

où W_t est un processus de Wiener standard. Supposons que ce processus possède de bonnes propriétés de mélange et de localisation. Notre approche exige une décroissance suffisamment rapide de la densité de la mesure invariante de ξ_t à l'infini ; par exemple, une condition de type Has'minskii $B(y) \cdot \frac{y}{|y|} < -c$, $c > 0$ est suffisante. En fait, en utilisant PARDOUX & VERETENNIKOV [34], on fera une hypothèse plus faible (cf. Hypothèse III.1.3).

On définit le générateur infinitésimal de la diffusion ξ_t :

$$\mathcal{L}g(y) \triangleq \sum_{k,l=1}^d q^{kl}(y) \frac{\partial^2 g(y)}{\partial y_k \partial y_l} + \sum_{k=1}^d B^k(y) \frac{\partial g(y)}{\partial y_k}, \quad \text{avec } q \triangleq \frac{1}{2} \sigma \sigma^*, \quad (\text{III.3})$$

et l'opérateur :

$$\mathcal{A}^\varepsilon h(x) \triangleq \operatorname{div} \left(a\left(\frac{x}{\varepsilon}, y\right) \nabla h(x) \right). \quad (\text{III.4})$$

\mathcal{A} désigne \mathcal{A}^ε pour $\varepsilon = 1$. On note que, appliqué à une fonction $f(z, y)$, \mathcal{L} agit sur la fonction $y \rightarrow f(z, y)$ à z fixé, et \mathcal{A}^ε agit sur la fonction $z \rightarrow f(z, y)$ à y fixé.

Hypothèse III.1.1 *Les coefficients $a(z, y)$, $c(z, y)$, et $q(y)$ ainsi que leurs dérivées d'ordre 1 en z et y sont uniformément bornés en $(z, y) \in \mathbb{T}^n \times \mathbb{R}^d$. La fonction $B(y)$ ainsi que ses dérivées d'ordre 1 sont à croissance au plus polynomiale.*

Hypothèse III.1.2 *Les opérateurs \mathcal{L} et \mathcal{A} sont uniformément elliptiques.*

Hypothèse III.1.3 *Il existe des constantes $\mu > -1$, $R_1 > 0$ et $C_3 > 0$ telles que*

$$B(y) \cdot \frac{y}{|y|} \leq -C_3 |y|^\mu \quad \text{pour tout } y \text{ tel que } |y| > R_1.$$

Sous les Hypothèses III.1.1, III.1.2 et III.1.3, le processus ξ_t admet une unique mesure invariante de densité régulière $\rho(\cdot)$ donnée par :

$$\mathcal{L}^* \rho = 0 \text{ sur } \mathbb{R}^d \text{ avec } \int_{\mathbb{R}^d} \rho(y) dy = 1. \quad (\text{III.5})$$

De plus, $\rho(\cdot)$ décroît plus rapidement que n'importe quelle puissance négative de $|y|$ lorsque $|y| \rightarrow \infty$; en fait, $\rho(y) \leq c_1 \exp(-c|y|^{1+\mu})$ avec $c > 0$.

Notations. ∇_z, ∇_y sont les gradients en espace selon z et y respectivement. Lorsque il n'y a pas d'ambiguïté sur l'argument de la fonction on utilise la notation ∇ , par exemple $\nabla u^\varepsilon(t, x)$ est $\nabla_x u^\varepsilon(t, x)$. (\cdot, \cdot) désigne le produit scalaire sur $L^2(\mathbb{R}^n)$ et $\|\cdot\|$ la norme associée. Pour une fonction $(z, y) \rightarrow f(z, y)$, on pose :

$$\begin{aligned} \overline{\langle f(\cdot, \cdot) \rangle} &= \int_{\mathbb{T}^n} \int_{\mathbb{R}^d} f(z, y) \rho(y) dy dz, \\ \langle f(\cdot, y) \rangle &= \int_{\mathbb{T}^n} f(z, y) dz, \\ \overline{f(z, \cdot)} &= \int_{\mathbb{R}^d} f(z, y) \rho(y) dy; \end{aligned}$$

où $\rho(y)$ est la densité de la mesure invariante du processus ξ_t . Pour un processus $(t, x) \rightarrow w(t, x)$, $w(t)$ désigne une application $x \rightarrow w(t, x)$. Ainsi $\|w(t)\| = (\int_{\mathbb{R}^n} |u(t, x)|^2 dx)^{1/2}$. Pour une fonction $(z, y) \rightarrow g(z, y)$ on pose $g^\varepsilon = g(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha})$.

On suppose finalement que :

$$\overline{\langle c(\cdot, \cdot) \rangle} = 0.$$

On peut toujours se ramener à cette situation en considérant la nouvelle fonction inconnue $\tilde{u}^\varepsilon(t, x) = e^{-t\overline{\langle c \rangle}/\varepsilon} u^\varepsilon(t, x)$.

III.1.2 RÉSULTATS DE CONVERGENCE

La structure du problème limite diffère singulièrement selon que $\alpha = 2$, $\alpha < 2$ ou $\alpha > 2$.

- Si $\alpha \leq 2$ alors la suite des lois de u^ε converge faiblement vers l'unique solution du problème de martingale limite comprenant le mouvement brownien unidimensionnel. La formule pour les coefficients du problème limite est différente selon que $\alpha = 2$ ou $\alpha < 2$.
- Si $\alpha > 2$ alors la suite u^ε converge en probabilité dans un espace fonctionnel approprié vers la solution du problème de Cauchy pour l'opérateur parabolique avec des coefficients constants déterministes. Ce résultat ressemble au résultat classique d'homogénéisation ; la seule différence est que nous obtenons la convergence dans la probabilité. La convergence presque sûre est une question en suspens.

Ces résultats prouvent que la présence de termes aléatoires dans la dynamique peut profondément influencer le comportement limite des solutions. Il est intéressant de noter que dans le cas particulier $c(z, y) \equiv 0$ l'équation limite est toujours déterministe KLEPTSYNA & PIATNITSKI [18].

Notons $L_w^2(\mathbb{R}^n)$ l'espace $L^2(\mathbb{R}^n)$ muni de la topologie de la convergence faible. Posons :

$$\Omega_T \triangleq L_w^2([0, T[; H^1(\mathbb{R}^n)) \cap \mathcal{C}([0, T]; L_w^2(\mathbb{R}^n)) \quad (\text{III.6})$$

qui sera munit du supremum de la topologie de la convergence uniforme sur $\mathcal{C}([0, T]; L_w^2(\mathbb{R}^n))$ et de la topologie faible sur $L^2([0, T[; H^1(\mathbb{R}^n))$. Ω_T est muni de sa tribu borélienne \mathcal{F} . Sur (Ω_T, \mathcal{F}) on définit la loi de probabilité suivante :

$$Q^\varepsilon \triangleq \text{loi}(u^\varepsilon(t); 0 \leq t \leq T).$$

le comportement asymptotique de Q^ε dépend du fait que $\alpha < 2$, $\alpha = 2$, ou $\alpha > 2$.

- (i) Si $\alpha < 2$ alors $Q^\varepsilon \Rightarrow \hat{Q}$ dans Ω_T lorsque $\varepsilon \downarrow 0$ où \hat{Q} est la loi de la solution \hat{u} de l'EDPS suivante :

$$d\hat{u}(t, x) = \left[\text{div}(\hat{a} \nabla \hat{u}(t, x)) + \hat{c} \hat{u}(t, x) \right] dt + \lambda \hat{u}(t, x) d\hat{W}_t \quad (\text{III.7})$$

avec $\hat{u}(0, x) = u_0(x)$, où $(t, x) \in [0, T] \times \mathbb{R}^n$, \hat{W}_t est un mouvement brownien standard à valeurs dans \mathbb{R} et :

$$\begin{cases} \hat{a} &= \overline{\langle a(I + \nabla_z \Psi) \rangle}, \\ \hat{c} &= \overline{\langle G c \rangle} \text{ qui est égal à } \overline{q \nabla G \cdot \nabla G}, \\ \lambda^2 &= \overline{2q \nabla G \cdot \nabla G}, \end{cases} \quad (\text{III.8})$$

et les fonctions $G \in \bar{H}_\rho^1(\mathbb{R}^d)$ et $\Psi^i \in \bar{H}_\rho^1(\mathbb{T}^n \times \mathbb{R}^d)$ sont solutions de :

$$\mathcal{L}G(y) = -\langle c(\cdot, y) \rangle, \quad (\text{III.9})$$

$$\mathcal{A}\Psi^i(z, y) = -\sum_{j=1}^n a_{z_j}^{ij}(z, y), \quad (\text{III.10})$$

pour $(z, y) \in \mathbb{T}^n \times \mathbb{R}^d$ et $1 \leq i \leq n$.

(ii) Si $\alpha = 2$ alors $Q^\varepsilon \Rightarrow \hat{Q}$ dans Ω_T lorsque $\varepsilon \downarrow 0$ où \hat{Q} est la loi de la solution \hat{u} de l'EDPS suivante :

$$d\hat{u}(t, x) = \left[\operatorname{div}(\hat{a} \nabla \hat{u}(t, x)) - \hat{b} \cdot \nabla \hat{u}(t, x) + \hat{c} \hat{u}(t, x) \right] dt + \lambda \hat{u}(t, x) d\hat{W}_t, \quad (\text{III.11})$$

avec $\hat{u}(0, x) = u_0(x)$, où $(t, x) \in [0, T] \times \mathbb{R}^n$, \hat{W}_t est un mouvement brownien standard à valeurs dans \mathbb{R} , et :

$$\begin{cases} \hat{a} &= \overline{\langle a(I + \nabla_z \Psi) \rangle}, \\ \hat{b} &= \overline{\langle \Psi c + a \nabla_z G \rangle}, \\ \hat{c} &= \overline{\langle G \bar{c} \rangle}, \\ \lambda^2 &= 2q \overline{\langle \nabla_y G \rangle} \cdot \overline{\langle \nabla_y G \rangle}, \end{cases} \quad (\text{III.12})$$

et les fonctions $G, \Psi^j \in \bar{H}_p^1(\mathbb{T}^n \times \mathbb{R}^d)$ sont les solutions de :

$$(\mathcal{A} + \mathcal{L})G(z, y) = -c(z, y), \quad (\text{III.13})$$

$$(\mathcal{A} + \mathcal{L})\Psi^j(z, y) = -\sum_{i=1}^n a_{z_i}^{ij}(z, y) \quad (\text{III.14})$$

pour $(z, y) \in \mathbb{T}^n \times \mathbb{R}^d$ et $1 \leq j \leq n$.

(iii) Si $\alpha > 2$ alors u^ε converge en probabilité vers \hat{u} la solution du problème de Cauchy limite suivant :

$$\hat{u}_t(t, x) = \operatorname{div}(\hat{a} \nabla \hat{u}(t, x)) + \hat{c} \hat{u}(t, x), \quad \hat{u}(0, x) = u_0(x) \quad (\text{III.15})$$

avec $(t, x) \in [0, T] \times \mathbb{R}^n$ et :

$$\hat{a} = \langle \bar{a}(I + \nabla_z \Psi) \rangle, \quad \hat{c} = \langle G \bar{c} \rangle, \quad (\text{III.16})$$

où les fonctions $G, \Psi^i \in \bar{H}^1(\mathbb{T}^n)$ sont solutions de :

$$\bar{\mathcal{A}}G(z) = -\overline{c(z, \cdot)}, \quad (\text{III.17})$$

$$\bar{\mathcal{A}}\Psi^i(z) = -\overline{\sum_{j=1}^n a_{z_j}^{ij}(z, \cdot)} \quad (\text{III.18})$$

pour $z \in \mathbb{T}^n$ et $1 \leq i \leq n$, et l'opérateur $\bar{\mathcal{A}}$ est défini par :

$$\bar{\mathcal{A}}f(z) = \operatorname{div}(\overline{a(z, \cdot)}) \nabla f(z). \quad (\text{III.19})$$

On en déduit que :

$$P\text{-}\lim_{\varepsilon \downarrow 0} \|u^\varepsilon - \hat{u}\|_{L^2([0, T] \times \mathbb{R}^n)} = 0 \quad (\text{III.20})$$

Comparaison avec les systèmes sans potentiel ou sans entrée bruitée : Il est intéressant de comparer les problèmes limites (III.7), (III.11), (III.15) avec les problèmes limites pour l'équation sans potentiel :

$$\frac{\partial}{\partial t} u(t, x) = \operatorname{div}(a(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}) \nabla u(t, x)) \quad (\text{III.21})$$

et l'équation sans entrée bruitée :

$$\frac{\partial}{\partial t} u(t, x) = \operatorname{div}(a(\frac{x}{\varepsilon}) \nabla u(t, x)) + \frac{1}{\varepsilon} c(\frac{x}{\varepsilon}) u(t, x). \quad (\text{III.22})$$

D'après [18], l'absence de potentiel dans (III.21) conduit toujours à un système homogénéisé déterministe :

$$\frac{\partial}{\partial t} \bar{u}(t, x) = \sum_{1 \leq i, j \leq n} \bar{a}^{ij} \frac{\partial^2}{\partial x_i \partial x_j} \bar{u}(t, x),$$

cet opérateur ne présente ni terme aléatoire ni terme d'ordre inférieur. Le problème limite pour (III.22) prend la forme suivante :

$$\frac{\partial}{\partial t} \bar{u}(t, x) = \sum_{1 \leq i, j \leq n} \bar{a}^{ij} \frac{\partial^2}{\partial x_i \partial x_j} \bar{u}(t, x) + \bar{c} \bar{u}(t, x)$$

où \bar{c} est appelé "terme étrange".

Commentaires sur l'équation limite (III.11) : L'aspect du terme du premier ordre $\hat{b} \cdot \nabla$ dans la partie dérive du problème limite (III.11) est intéressant. Dans [6] on donne un exemple en dimension 1 pour lequel ce terme est non nul.

III.1.3 DÉMONSTRATIONS

En général, l'espérance de la norme de u^ε n'admet pas de borne supérieure uniforme en ε . Nous ne pouvons donc pas appliquer la technique standard afin d'obtenir des résultats de compacité faible. Pour contourner cette difficulté on décompose u^ε de la façon suivante :

$$u^\varepsilon(t, x) = \exp\left(\frac{1}{\varepsilon^{1 \wedge \frac{\alpha}{2}}} \int_0^t \langle c(\cdot, \xi_{s/\varepsilon^\alpha}) \rangle ds\right) \times v^\varepsilon(t, x). \quad (\text{III.23})$$

PROBLÈME D'HOMOGENÉISATION AUXILIAIRE

Un calcul direct permet d'établir que la fonction auxiliaire v^ε introduite en (III.23) satisfait :

$$\frac{\partial}{\partial t} v^\varepsilon(t, x) = \operatorname{div} \left[a\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}\right) \nabla v^\varepsilon(t, x) \right] + \frac{1}{\varepsilon^{1 \wedge \frac{\alpha}{2}}} \tilde{c}\left(\frac{x}{\varepsilon}, \xi_{t/\varepsilon^\alpha}\right) v^\varepsilon(t, x), \quad (\text{III.24})$$

avec $v^\varepsilon(0, x) = u_0(x)$, $x \in \mathbb{R}^n$, $t \in [0, T]$, et $\tilde{c}(z, y) = c(z, y) - \langle c(\cdot, y) \rangle$.

Nous avons les estimées *a priori* suivantes : il existe c tel que uniformément en $\varepsilon > 0$,

$$\begin{aligned} (\text{E1}) \quad & \sup_{t \leq T} \|v^\varepsilon(t)\|^2 \leq c, \\ (\text{E2}) \quad & \int_0^T \|\nabla v^\varepsilon(s)\|^2 ds \leq c, \\ (\text{E3}) \quad & |(\phi, v^\varepsilon(t) - v^\varepsilon(s))| \leq c|t - s|^{1/2} \end{aligned}$$

pour toute fonction test $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$. Soit \tilde{Q}^ε la mesure de probabilité de v^ε sur (Ω_T, \mathcal{F}) . D'après VIOT [36] et BOUC & PARDOUX [4, Th. 2.5], la famille $\{\tilde{Q}^\varepsilon\}_{\varepsilon > 0}$ est tendue.

PROBLÈMES D'HOMOGENÉISATION AUXILIAIRES LIMITES

On montre la convergence de v^ε en probabilité et que la limite est solution d'une équation parabolique à coefficients constants. On traite les cas $\alpha < 2$, $\alpha = 2$ et $\alpha > 2$ séparément.

(i) Cas $\alpha < 2$: soit $\hat{v}(t, x)$ la solution du problème de Cauchy suivant :

$$\frac{\partial}{\partial t} \hat{v} = \operatorname{div}(\hat{a} \nabla \hat{v}), \quad (\text{III.25})$$

avec $\hat{v}(x, 0) = u_0(x)$ et où \hat{a} est définie en (III.8).

Pour toute fonction test $\varphi(x, t) \in C^\infty([0, T]; C_0^\infty(\mathbb{R}^n))$, on a :

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} \mathbb{E} \sup_{t \leq T} \left| (\phi(t), v^\varepsilon(t)) - (\phi(0), u_0) - \int_0^t (\phi_s(s), v^\varepsilon(s)) ds \right. \\ \left. - \int_0^t \left(\sum_{ij} \hat{a}^{ij} \frac{\partial^2 \phi(s)}{\partial x_i \partial x_j}, v^\varepsilon(s) \right) ds \right| = 0 \end{aligned} \quad (\text{III.26})$$

On en déduit que v^ε converge en probabilité vers la solution \hat{v} de (III.25).

En effet, on introduit la fonctionnelle continue et bornée $\Phi_\phi(u)$ sur Ω_T :

$$\begin{aligned} \Phi_\phi(u) \triangleq 1 \wedge \sup_{t \leq T} \left| (\phi(t), u(t)) - (\phi(0), u_0) - \int_0^t ((\phi_s(s), u(s)) ds \right. \\ \left. - \int_0^t \left(\sum_{ij} \hat{a}^{ij} \phi_{x_i x_j}(s), u(s) \right) ds \right|. \end{aligned}$$

d'après (III.26), on a $\lim_{\varepsilon \downarrow 0} \mathbb{E} \Phi_\phi(v^\varepsilon) = 0$, donc pour tout point d'adhérence Q de la famille \tilde{Q}^ε , on a $\mathbb{E}_Q \Phi_\phi = 0$ et donc toute mesure limite Q est concentrée sur la solution faible de l'équation déterministe (III.25). L'unicité de la solution de cette dernière équation implique la convergence en probabilité cherchée.

(ii) Cas $\alpha = 2$: On considère la solution $\hat{v}(x, t)$ du problème de Cauchy suivant,

$$\frac{\partial}{\partial t} \hat{v} = \operatorname{div}(\hat{a} \nabla \hat{v}) - \hat{b} \nabla \hat{v} + \hat{c} \hat{v} \quad (\text{III.27})$$

avec $\hat{v}(x, 0) = u_0(x)$, où \hat{a} et \hat{b} sont définis en (III.8); $\hat{c} = \overline{\langle \tilde{G} \tilde{c} \rangle}$. On utilise la même technique que la cas précédent.

(iii) Cas $\alpha > 2$: De la même façon on introduit la fonction $\hat{v}(x, t)$ solution du problème de Cauchy :

$$\frac{\partial}{\partial t} \hat{v} = \operatorname{div}(\hat{a} \nabla \hat{v}) + \hat{c} \hat{v} \quad (\text{III.28})$$

avec $\hat{v}(x, 0) = u_0(x)$, \hat{a} et \hat{c} sont définis en (III.16). On montre de la même façon que v^ε converge vers \hat{v} .

CONVERGENCE EN LOI DU PROBLÈME INITIAL

L'analyse utilisée ici est essentiellement basée sur les propriétés de la densité de la mesure invariante de ξ_t et d'une solution d'une équation de Poisson, établies sur l'espace produit de \mathbb{R}^d et du tore \mathbb{T}^n . Ces propriétés proviennent de PARDOUX & VERETENNIKOV [34] et ont nécessité quelques extensions techniques (voir [6, Appendix]).

On s'intéresse ici à la loi limite de l'argument de l'exponentielle de la décomposition (III.23). Le passage à la limite dans le produit (III.23) s'appuie sur la nature déterministe de la limite de v^ε .

(i) Cas $\alpha \leq 2$: D'après [34, Theorem 3] :

$$J_t^{\varepsilon, \alpha} = \frac{1}{\varepsilon^{\alpha/2}} \int_0^t \langle c(\cdot, \xi_{s/\varepsilon^\alpha}) \rangle ds \xrightarrow[\varepsilon \downarrow 0]{} \lambda \hat{W}_t \quad \text{dans } \mathcal{C}[0, T]$$

où \hat{W}_t est un mouvement brownien standard, scalaire et

$$\lambda^2 \triangleq 2 \overline{\nabla G \cdot \nabla G} \quad \text{avec} \quad \mathcal{L}G = -\langle c(\cdot, y) \rangle.$$

Cette définition de λ coïncide avec celles données dans les cas $\alpha < 2$ et $\alpha = 2$. Pour $\alpha \leq 2$ la famille $\exp(J_t^{\varepsilon, \alpha})$ converge en loi vers $\exp(\lambda \hat{W}_t)$.

Afin de passer à la limite dans le produit (III.23), c'est à dire dans $\exp(J_t^{\varepsilon, \alpha}) v^\varepsilon$, on note que :

$$\exp(J_t^{\varepsilon, \alpha}) v^\varepsilon = \exp(J_t^{\varepsilon, \alpha}) \hat{v} + \exp(J_t^{\varepsilon, \alpha}) (v^\varepsilon - \hat{v}) \quad (\text{III.29})$$

L'application $\nu(\cdot) \rightarrow \nu(\cdot) \hat{v}$ est continue de $\mathcal{C}[0, T]$ dans Ω_T . Le premier terme du membre de droite de (III.29) converge en loi vers $\exp(\lambda \hat{W}_t) \hat{v}$ dans Ω_T as $\varepsilon \rightarrow 0$. Le second terme converge en probabilité vers 0 (voir [6] pour les détails).

Pour montrer que le processus limite est solution des EDPS (III.7) ou (III.11) on peut appliquer la formule de Itô au produit $\exp(\lambda \hat{W}_t) \hat{v}(t, x)$ et utiliser les équations homogénéisées auxiliaires (III.25) ou (III.27) respectivement.

(ii) Cas $\alpha > 2$: $\frac{1}{\varepsilon} \int_0^t \langle c(\cdot, \xi_{s/\varepsilon^\alpha}) \rangle ds$ converge en probabilité vers 0 (i.e. $\lambda = 0$) et :

$$\mathbb{P}\text{-}\lim_{\varepsilon \rightarrow 0} \exp\left(\frac{1}{\varepsilon} \int_0^t \langle c(\cdot, \xi_{s/\varepsilon^\alpha}) \rangle ds\right) = 1. \quad (\text{III.30})$$

On passe alors à la limite dans le produit (III.23) et on utilise le fait que v^ε converge en probabilité dans Ω_T vers la solution \hat{v} de (III.28).

III.2 MÉTHODES DE MONTE CARLO POUR UN MILIEU HÉTÉROGÈNE

III.2.1 INTRODUCTION

Le sous-sol est constitué de matériaux ayant la capacité d'emmagasiner, de laisser s'écouler et de restituer un liquide. Il peut ainsi jouer le rôle de *réservoir* comme dans le cas de gisements pétroliers. On peut distinguer deux types de roches réservoirs : les *pores* (vides de petites dimensions entre les grains de forme et de grosseur variable) et les *fissures* (fentes allongées).

La *perméabilité* est une caractéristique physique qui représente la facilité qu'a un matériau à permettre le transfert de fluide au travers d'un réseau connecté. La grande hétérogénéité du sous-sol conduit à modéliser la perméabilité comme un champ aléatoire $\{k(x); x \in D\}$ sur le domaine réservoir D .

La *loi de Darcy* permet de relier un débit à un gradient de pression via la perméabilité $k(x)$. Pour un écoulement monophasique, elle s'écrit

$$u(x) = -\frac{k(x)}{\mu} \nabla p(x)$$

où $u(x)$ est le flux, μ la viscosité du liquide et $p(x)$ la pression. À cette équation on adjoint des conditions limites comme des flux nuls aux bords ($u(x) \cdot \vec{n} = 0$) et des pressions données aux points de forage.

Le terme $K(x) = \frac{k(x)}{\mu}$, appelé *coefficient de perméabilité* ou conductivité hydraulique, est une caractéristique du réservoir i.e. de l'écoulement du fluide donné dans un matériau donné.

Dans le cas d'un fluide incompressible, comme le pétrole, $\text{div}u(x) = 0$. L'équation précédente devient

$$-\text{div}\left(\frac{k(x)}{\mu} \nabla p(x)\right) = 0.$$

HOMOGENÉISATION

Un gisement pétrolier est un milieu poreux/fissuré formé d'un ou plusieurs réservoirs rocheux souterrains. Pour optimiser l'exploitation d'un gisement il est préalablement nécessaire d'en estimer les caractéristiques et de pouvoir simuler différentes stratégies de production. Il est également impératif de suivre l'évolution de ces caractéristiques en cours de l'exploitation.

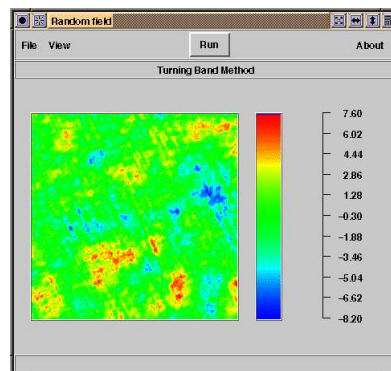


FIG. III.1 – Un exemple de champ de perméabilité 2D.

À partir de connaissances géologiques et de différents types de mesures (observations aux puits de forage, mesures sismiques, mesures satellitaires,...), les géo[logues | physiciens | statisticiens] proposent des modèles probabilistes du gisement : ils peuvent proposer différentes réalisations (2D ou 3D) de champ de perméabilité (cf. Fig. III.1) sur une grille.

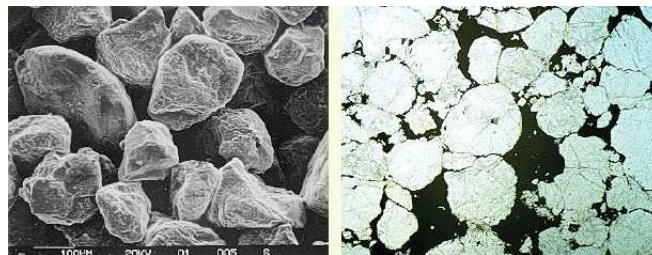


FIG. III.2 – Roches sédimentaires poreuses.

Ce champ de perméabilité détermine les conditions d'écoulement dans le gisement. Se pose alors un problème d'échelle : la perméabilité est déterminée à l'échelle des pores (cf. Fig. III.2), en dessous du mm , le gisement est à l'échelle de plusieurs kilomètres voire centaines de kilomètres. Les codes de simulation traitent typiquement des grilles de taille 10^6 à 10^7 ; une grille pertinente sur le plan physique serait de taille 10^{10} à 10^{18} .

L'homogénéisation est alors utilisée comme technique de changement d'échelle (upscaling) : un champ de perméabilité homogénéisé est généré sur une grille plus grossière sur laquelle les codes de simulation peuvent tourner.

À petite échelle, l'écoulement du fluide est décrite par la loi de Darcy

$$-\operatorname{div}\left(K\left(\frac{x}{\varepsilon}\right) \nabla p_{\varepsilon}(x)\right) = 0$$

pour $\varepsilon > 0$ petit. L'homogénéisation consiste à étudier le comportement asymptotique de l'équation précédente et à prouver la convergence vers

$$-\operatorname{div}\left(K_0 \nabla p_0(x)\right) = 0$$

où K_0 est un coefficient déterministe appelé coefficient homogénéisé ou *coefficient de perméabilité effective*.

III.2.2 MARCHES ALÉATOIRES POUR LE CALCUL DE COEFFICIENTS EFFECTIFS

L'idée sous-jacente aux méthodes probabilistes est très simple. Il s'agit de simuler une particule (ou plusieurs particules en parallèle) dont le déplacement est conditionné par le champ de perméabilité, la facilité avec laquelle cette particule s'éloigne de son point de départ donnera des indications précises sur la perméabilité effective du milieu (cf. Fig. III.3).

Ces idées ont été initialement proposées par MC CARTHY [22, 23, 25, 24] et NEUTINGER [28] comme alternatives aux méthodes numériques non probabilistes (différences finies, volumes finis, etc.).

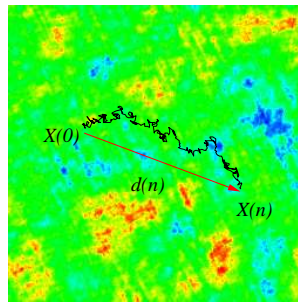


FIG. III.3 – Calcul de perméabilité effective par marche aléatoire.

La première idée est de construire une marche aléatoire X_n sur la grille dont la probabilité de transition dépend des valeurs locales des perméabilités, on peut par exemple choisir la moyenne harmonique des perméabilités (cf. Fig. III.4)

$$\mathbb{P}(X_{n+1} = y | X_n = x) = \begin{cases} \frac{1}{4} \frac{2k(x)k(y)}{k(x)+k(y)} & \text{si } y \text{ est l'un des 4 voisins de } x, \\ 1 - \sum_{y' \text{ voisin de } x} \frac{1}{4} \frac{2k(x)k(y')}{k(x)+k(y')} & \text{si } y = x, \\ 0 & \text{sinon.} \end{cases}$$

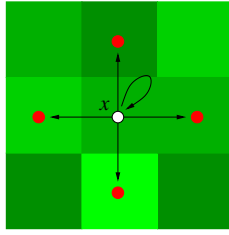


FIG. III.4 – Définir une marche aléatoire à partir du champ de perméabilité.

On peut alors estimer le coefficient de diffusion (cf. Fig. III.3) :

$$D \simeq \frac{1}{N} |X_0 - X_N|^2$$

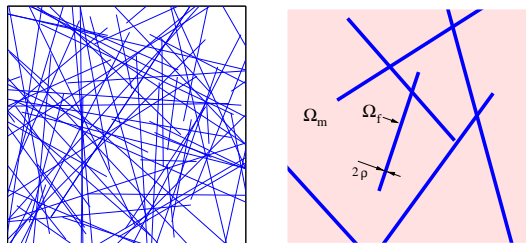
il existe enfin une relation entre le coefficient de diffusion et le coefficient de conductivité (relation de Nernst–Einstein) qui est le coefficient de perméabilité.

Il existe des extensions en temps continu de cet algorithme permettant d'échapper plus aisément aux zones très imperméables. En pratique plusieurs marches aléatoires indépendantes sont générées indépendamment et en parallèles, il est donc nécessaire de pouvoir générer des nombres pseudo-aléatoires en parallèle. Ces travaux ont été traités dans la première partie de la thèse d'Élisabeth Remy [35].

III.2.3 SIMULATION "EXACTE" POUR DES MILIEUX FISSURÉS

On présente un algorithme de simulation d'une diffusion dans un milieu poreux fissuré : il s'agit de simuler le couple instant/position de première atteinte de l'interface matrice/fissures, puis de suivre le comportement de la particule dans la fissure (voir CAMPILLO & LEJAY [9, 10, 11]). Le but est de calculer le coefficient d'échange dans le modèle à double porosité.

Lorsque l'on approche une fonctionnelle d'un processus de diffusion arrêté à l'aide d'une méthode de Monte Carlo on est souvent tenté de simuler ce processus de diffusion. Ce n'est pas toujours une bonne stratégie. Dans le cas traité ici, il est possible de calculer directement la loi de la diffusion arrêtée et d'en déduire une bonne approximation, ce qui explique l'appellation "exacte" de ce genre de simulation.

FIG. III.5 – Un domaine fissuré Ω comprenant un réseau de fissures Ω_f (fissures de largeur 2ρ) et la matrice Ω_m .

MODÈLE À DOUBLE POROSITÉ

On souhaite déterminer le coefficient d'échange dans un modèle à double porosité. Soit $\Omega = \Omega_f \cup \Omega_m \subset \mathbb{R}^2$ un domaine borné avec $\Omega_f \cap \Omega_m = \emptyset$. On suppose que le milieu est périodique, Ω est donc identifié à tout l'espace. Ω_m (resp. Ω_f) est la *matrice*, c'est à dire le milieu poreux (resp. le réseau des fissures); voir Fig. III.5.

La pression $p(t, x, y)$ d'un fluide dans un tel milieu à l'instant t au point (x, y) est une solution de :

$$\frac{\partial}{\partial t} p(t, x, y) = Ap(t, x, y), \quad A = \operatorname{div}(a(x, y) \nabla). \quad (\text{III.31})$$

Le coefficient a est la diffusivité du milieu, on suppose qu'elle est de la forme :

$$a(x, y) = [a_m \mathbf{1}_{\Omega_m}(x, y) + a_f \mathbf{1}_{\Omega_f}(x, y)] \times \operatorname{Id},$$

avec $a_f \gg a_m > 0$.

L'équation (III.31) est écrite à l'échelle des pores, alors qu'un réservoir pétrolier peut faire plusieurs kilomètres de long. Une des façons d'étudier la pression consiste à remplacer (III.31) par le système :

$$\begin{cases} \Phi_m \frac{\partial}{\partial t} P_m &= a_m \Delta P_m - \alpha(P_m - P_f), \\ \Phi_f \frac{\partial}{\partial t} P_f &= a_f \Delta P_f + \alpha(P_m - P_f), \end{cases} \quad (\text{III.32})$$

où $\Phi_{\Omega_\bullet} = \operatorname{Mesure}(\Omega_\bullet) / \operatorname{Mesure}(\Omega)$ avec $\bullet = m$ ou f ; P_m et P_f sont les pressions moyennes dans la matrice et dans les fissures sur un volume V donné.

Le coefficient α est appelé *coefficient d'échange*. Le modèle (III.32) est le *modèle à double porosité*, présenté ici en régime permanent (voir WARREN & ROOT [37]). Pour l'utiliser il est nécessaire d'estimer au préalable le coefficient d'échange α .

On a proposé une méthode de Monte Carlo s'appuyant sur le fait que α et toute autre propriété effective de ce milieu est lié à la loi d'une particule brownienne de vitesse $2a_m$ dans la matrice Ω_m et de vitesse $2a_f$ dans le réseau de fissures Ω_f . C'est à dire au processus de diffusion (X, Y) associé au générateur infinitésimal A défini par (III.31); voir LEJAY [20] et NĚTINGER *et al* [29].

DE LA MATRICE VERS LE RÉSEAU DE FISSURES

Les fissures sont supposées de largeur nulle. L'algorithme se fonde sur la simulation du couple instant/position de sortie d'un domaine de forme simple, à savoir le carré : la loi jointe de $\tilde{\tau} = \inf\{t \geq 0 \mid W_t \notin \square\}$ et $W_{\tilde{\tau}}$ est connue explicitement comme une série dont les termes d'ordre supérieur peuvent être négligés, voir MILSTEIN & TRETYAKOV [26].

L'algorithme qui en découle est simple, voir Fig. III.6.

SIMULATION D'UNE PARTICULE DANS LES FISSURES

La difficulté réside dans le traitement de la particule à l'interface matrice/fissures. Dans une fissure (unique) on peut considérer les composantes longitudinale X_t et transversale Y_t du mouvement de la particule (cf. Fig. III.7).

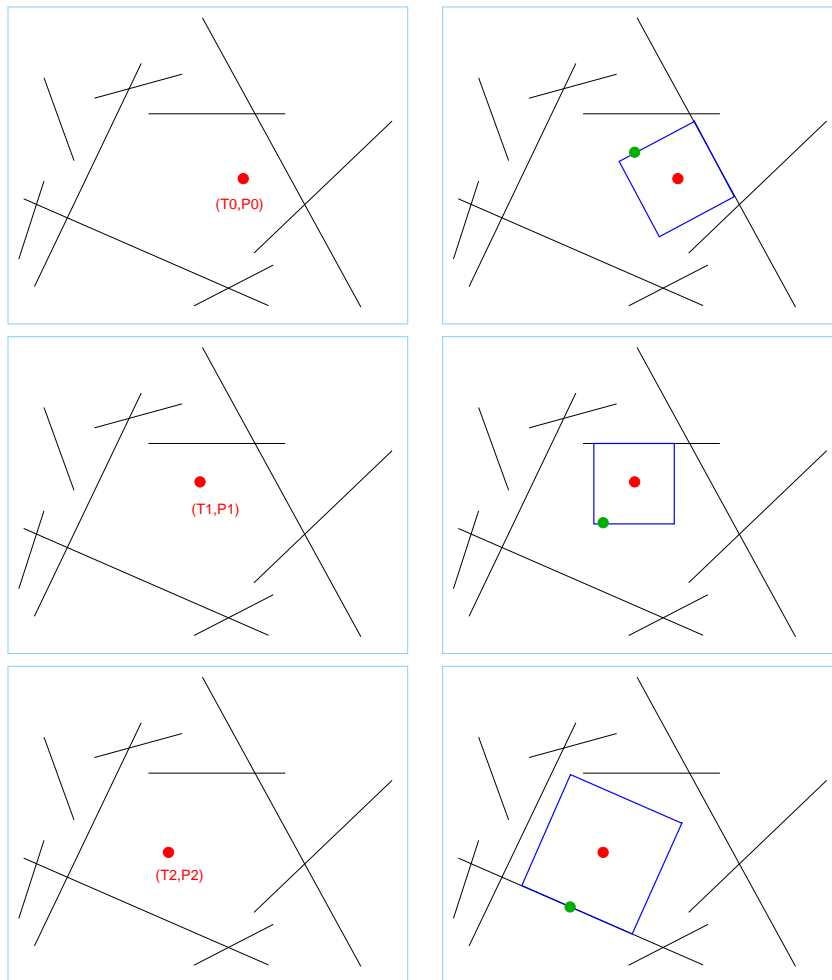


FIG. III.6 – Simulation matrice→fissures : Initialement dans la matrice (t_0, P_0) , on connaît la loi du couple temps/position d'atteinte du bord d'un carré centré en P_0 . Il s'agit alors de déterminer le plus grand carré inclus dans la matrice, centré en P_0 et n'intersectant pas une fissure. À partir d'un certain ordre ce carré s'appuiera sur une fissure et la particule atteindra cette fissure avec une probabilité $\frac{1}{4}$ (si tout le côté est inclus dans la fissure).

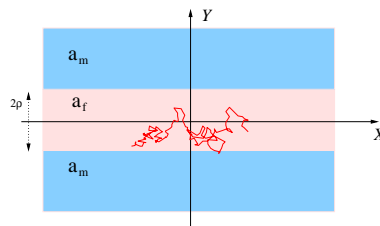


FIG. III.7 – Composantes longitudinale X_t et transversale Y_t du mouvement de la particule le long de la fissure.

L'équation satisfaite par X_t est relativement simple. Celle correspondante à Y_t l'est moins, elle fait intervenir les temps locaux du processus Y_t aux points $-\rho$ et ρ (aux interfaces). Néanmoins, Y_t admet une représentation faisant intervenir des mouvements browniens asymétriques ("skew Brownian motions"). La première idée consiste à s'inspirer de la construction d'une trajectoire de mouvement brownien asymétrique à partir de la trajectoire d'un mouvement brownien, cette construction est présentée à la Fig. III.8.

Toutefois il n'est pas possible de "suivre" le mouvement brownien dans toutes ses excursions. La seconde idée, empruntée à Erika Hausenblas, consiste donc à négliger les excursions inférieures à une durée η fixée à l'avance (cf. Fig. III.9).

Une fois regroupées, ces idées permettent de proposer un algorithme en deux étapes : on simule d'abord le temps de sortie t_{sortie} et la composante transversale y_{sortie} correspondante (cf. Fig. III.10), puis de simuler la composante longitudinale x_{sortie} associée (cf. Fig. III.11).

Nous ne disposons pas pour le moment d'analyse numérique de ce genre d'approximation. Ce travail est toujours en cours.

BIBLIOGRAPHIE

- [1] G. ALLAIRE. Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23 :1482–1518, 1992.
- [2] M. AVELLANEDA AND A. J. MAJDA. An integral representation and bounds on the effective diffusivity in passive advection by laminar and turbulent flows. *Communications in Mathematical Physics*, 138(2) :339–391, 1991.
- [3] M. AVELLANEDA AND A. J. MAJDA. Simple examples with features of renormalization for turbulent transport. *Philosophical Transactions of the Royal Society of London, Series A*, 346 :205–233, 1994.
- [4] R. BOUC AND É. PARDOUX. Asymptotic analysis of PDEs with wide-band noise disturbances, and expansion of the moments. *Stochastic Analysis and Application*, 2(4) :369–422, 1984.
- [5] A. BOURGEAT, S. LUCKHAUS, AND A. MIKELIĆ. A rigorous result for a double porosity model of immiscible two-phase flow. *Comptes Rendus à l'Académie des Sciences*, 320, Série I, pages 1289–1294, 1994.
- [6] F. CAMPILLO, M. KLEPTSINA, AND A. PIATNITSKI. Homogenization of random parabolic operator with large potential. *Stochastic Processes and Their Applications*, 93 :57–85, 2001.

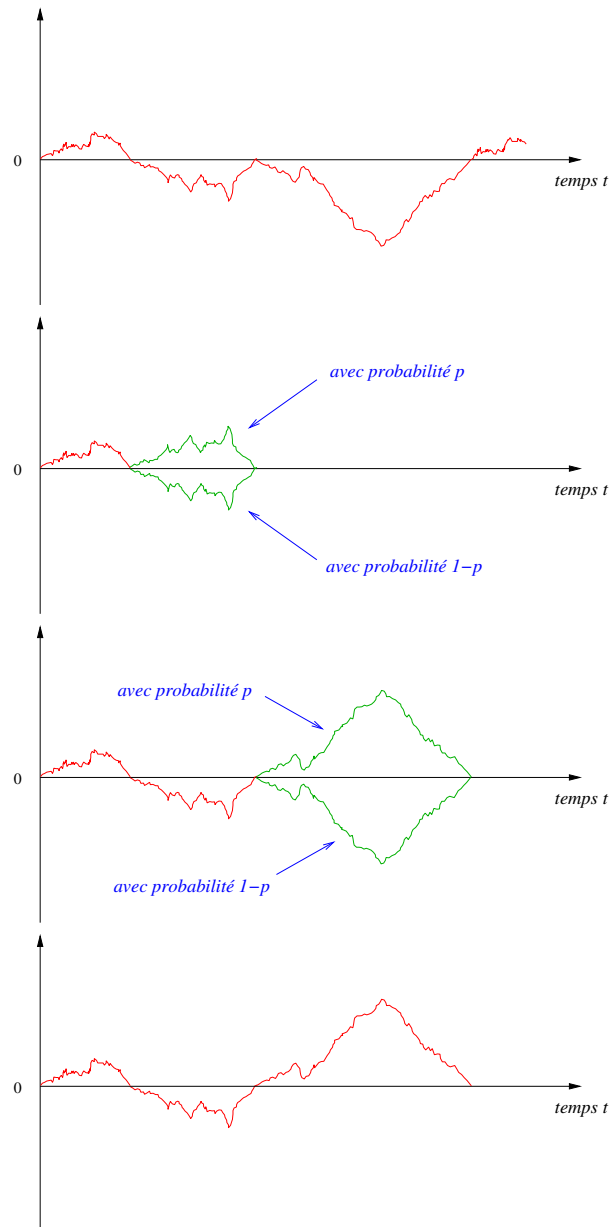


FIG. III.8 – “Construire” une trajectoire d’un mouvement brownien asymétrique de paramètre γ à partir d’une trajectoire de mouvement brownien : on redistribue les excursions à l’aide d’un tirage i.i.d. de Bernoulli de probabilité $p = \frac{1}{2}(1 + \gamma)$ indépendant du processus. Dans notre cas le coefficient γ dépend des perméabilités : l’excursion a une probabilité plus importante de faire une excursion dans la fissure que de la faire dans la matrice.

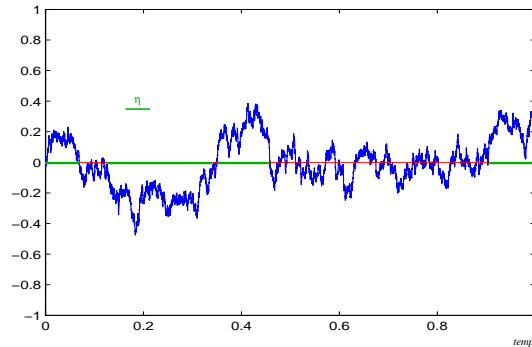


FIG. III.9 – Une idée utilisée ici et empruntée à Erika Hausembles, consiste à négliger les petites excursions (i.e. celles dont la durée est inférieure à η). Pour ces petites excursions on suppose que la particule ne s'est pas suffisamment éloignée de l'interface pour décider si elle est dans la matrice ou dans les fissures.

- [7] F. CAMPILLO, M. KLEPTSZYNA, AND A. PIATNITSKI. Homogenization of random parabolic operator with large potential. In B. Grigelionis et al., editor, *Probability Theory and Mathematical Statistics, Proceedings of the 7th Vilnius Conference*. TEV. Vilnius – VSP. Utrecht, 1999.
- [8] F. CAMPILLO, M. KLEPTSZYNA, AND A. PIATNITSKI. Homogenization of random parabolic operator with large potential in locally periodic media. In *Xth International Symposium on Applied Stochastic Models and Data Analysis (ASMDA)*, pages 277–282, Compiègne, Juin 12–15 2001.
- [9] F. CAMPILLO AND A. LEJAY. A Monte Carlo method without grid to compute the exchange coefficient in the double porosity model. Part I : From the matrix to fissures. Rapport de Recherche RR-4048, INRIA, November 2000. <http://www.inria.fr/rrrt/rr-4048.html>.
- [10] F. CAMPILLO AND A. LEJAY. A Monte Carlo method to compute the exchange coefficient in the double porosity model. *Monte Carlo Methods and Applications*, 7(1-2) :65–72, 2001.
- [11] F. CAMPILLO AND A. LEJAY. A Monte Carlo method without grid for a fractured porous domain model. *Monte Carlo Methods and Applications*, 8(2) :129–148, 2002.
- [12] F. CAMPILLO AND A. L. PIATNITSKI. Effective diffusion in vanishing viscosity. In J.-L. Lions D. Cioranescu, editor, *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. XIV (Paris, 1997/1998)*, volume 31 of *Studies in Mathematics and its Applications*, pages 133–145. North-Holland, Amsterdam, 2002.
- [13] A. FANNJIANG AND G. PAPANICOLAOU. Diffusion in turbulence. *Probability Theory and Related Fields*, 105 :279–334, 1996.
- [14] A. FANNJIANG AND G. PAPANICOLAOU. Convection enhanced diffusion for random flows. *Journal of Statistical Physics*, 88 :1033–1076, 1997.
- [15] E. HAUSEMBLAS. A numerical scheme using Itô excursions for simulating local time resp. stochastic differential equations with reflection. *Osaka Journal of Mathematics*, 36(1) :105–137, 1999.
- [16] U. HORNUNG, editor. *Homogenization and porous media*, volume 6 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 1997.

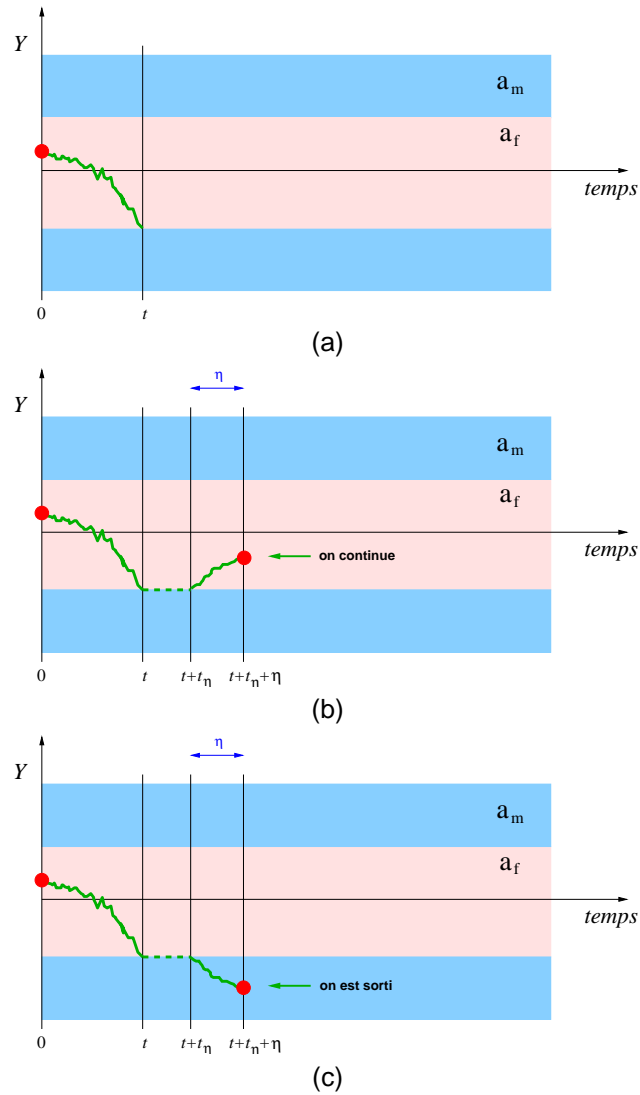


FIG. III.10 – Fissures \rightarrow matrice, simulation de la composante transversale et du temps de sortie : (a) à l'instant 0 on est en $y \in]-\rho, +\rho[$, on connaît la loi du couple temps/position transversale d'atteinte de l'interface que l'on peut simuler [temps= t]. Après un temps exponentiel τ avant la première excursion plus longue que η une excursion peut nous ramener dans la fissure (b) ou ne faire sortir dans la matrice (c). Dans le premier cas on poursuit cette simulation, dans le second on conserve le temps de sortie t_{sortie} et la composante y_{sortie} correspondante.

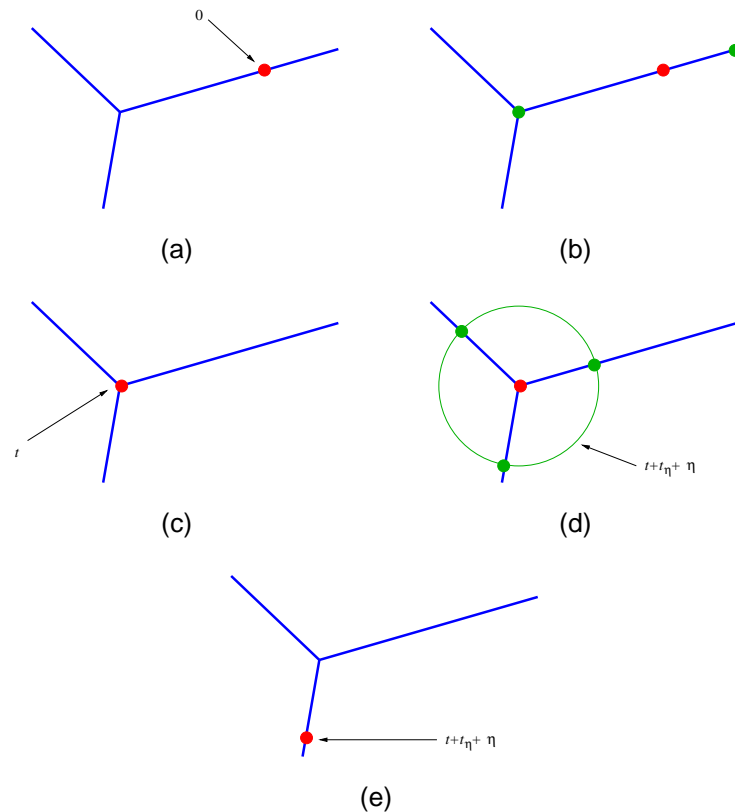


FIG. III.11 – Fissures \rightarrow matrice, simulation de la composante longitudinale : (a) à l'instant initial la particule est à l'intérieur d'un segment [temps=0]. (b) et (c) on simule l'instant/le sommet atteint [temps=t]. (d) on simule le temps d'attente avant la première excursion $> \eta$ et la distance atteinte à partir de là en un instant η [temps= $t+t_\eta+\eta$]. (e) on choisit un des points au hasard [temps= $t+t_\eta+\eta$]. À chaque instant il faut tester si l'on n'a pas dépassé le temps t_{sortie} précédemment calculer (cf. Fig.III.10), dès que cet instant est dépassé on simule la position longitudinale x_{sortie} correspondante (conditionnée par le fait que l'on connaît la position de la particule à un instant précédent t_{sortie} et sa position après t_{sortie}).

- [17] V. V. JIKOV, S. M. KOZLOV, AND O. A. OLEINIK. *Homogenization of Differential Operators and Integral Functionals*. Springer–Verlag, 1994.
- [18] M. L. KLEPTSYNA AND A. L. PIATNITSKI. Homogenization of random parabolic operators. *GAKUTO International Series, Mathematical Sciences and Applications*, 9 :241–255, 1997.
- [19] S. M. KOZLOV. Averaging of random operators. *Math. USSR Sbornik*, 37(2) : 167–180, 1980.
- [20] A. LEJAY. *Méthodes Probabilistes pour l’Homogénéisation des Opérateurs Sous Forme Divergence : Cas Linéaires et Semi-Linéaires*. Thèse de doctorat, Université de Provence, 2000.
- [21] A. LEJAY. Simulating a diffusion on a graph. Application to reservoir engineering. *Monte Carlo Methods and Applications*, 9(3) :241–256, 2003.
- [22] J. F. MC CARTHY. Effective permeability of sandstone–shale reservoirs by random walk method. *Journal of Physics A : Mathematical and General*, 23 :L445–L451, 1990.
- [23] J. F. MC CARTHY. Analytical models of the effective permeability of sand–shale reservoirs. *Geophysical Journal International*, 105 :513–527, 1991.
- [24] J. F. MC CARTHY. Reservoir characterization : Efficient random–walk methods for upscaling and image selection. In *Proceedings of the SPE Asia Pacific Oil & Gas Conference & Exhibition, Singapore, 8–10 February*, pages 159–171, 1993.
- [25] J. F. MC CARTHY. Continuous–time random walks on random media. *Journal of Physics A : Mathematical and General*, 26 :2495–2503, 1993.
- [26] G. N. MILSTEIN AND M. V. TRETYAKOV. Simulation of a space-time bounded diffusion. *The Annals of Applied Probability*, 9(3) :732–779, 1999.
- [27] G. N. MILSTEIN AND M. V. TRETYAKOV. *Stochastic numerics for mathematical physics*. Scientific Computation. Springer–Verlag, Berlin, 2004.
- [28] B. NÆTINGER. The effective permeability of heterogeneous porous media. *Transport in Porous Media*, 15 :99–127, 1994.
- [29] B. NÆTINGER, T. ESTEBENET, AND M. QUINTARD. Up-scaling flow in fractured media : equivalence between the large scale averaging theory and the continuous time random walk method. *Transport in Porous Media*, 43(3) :581–596, 2001.
- [30] Y. OUKNINE. “Skew–Brownian motion” and derived processes. *Theory of Probability and its Applications*, 35(1) :163–169, 1990.
- [31] G. C. PAPANICOLAOU AND S. R. S. VARADHAN. Diffusions with random coefficients. In G. Kallianpur, P. R. Krishnaiah, and J. K. Ghosh, editors, *Statistics and Probability : An Essay in Honor of C. R. Rao*, pages 547–552. North–Holland, 1982.
- [32] É. PARDOUX. Quelques méthodes probabilistes pour les équations aux dérivées partielles. In *Actes du 30ème Congrès d’Analyse Numérique : CANum ’98 (Arles, 1998)*, volume 6 of *ESAIM Proc.*, pages 91–109. SMAI, Paris, 1998.
- [33] É. PARDOUX AND A. L. PIATNITSKI. Homogenization of a nonlinear random parabolic partial differential equation. *Stochastic Processes and their Applications*, 104(1) :1–27, 2003.
- [34] É. PARDOUX AND A. YU. VERETENNIKOV. On the Poisson equation and diffusion approximation. I. *The Annals of Probability*, 29(3) :1061–1085, 2001.
- [35] É. REMY. *Homogénéisation d’opérateurs aux différences : Approches analytique et probabiliste*. Thèse, Université de Provence, 1999.

-
- [36] M. VIOT. *Solutions Faibles d'Équations aux Dérivées Partielles Stochastiques Non Linéaires*. Thèse de doctorat d'état, Université Paris VI, 1976.
- [37] J. E. WARREN AND P. J. ROOT. The behavior of naturally fractured reservoirs. *The Society of Petroleum Engineers Journal*, 3(3) :245–255, 1963.

