



HAL
open science

**Probabilités : aspects théoriques et applications en
filtrage non linéaire, systèmes de particules et processus
sto chastiques.**

Sylvain Rubenthaler

► **To cite this version:**

Sylvain Rubenthaler. Probabilités : aspects théoriques et applications en filtrage non linéaire, systèmes de particules et processus sto chastiques.. Probabilités [math.PR]. Université Nice Sophia Antipolis, 2010. tel-01144189

HAL Id: tel-01144189

<https://hal.science/tel-01144189>

Submitted on 25 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Habilitation à diriger des recherches en
sciences
Mathématiques appliquées**

présentée par

Sylvain Rubenthaler

à l'université de Nice-Sophia Antipolis

intitulée

**Probabilités : aspects théoriques et
applications en filtrage non linéaire,
systèmes de particules et processus
stochastiques.**

rapportée par

M. Arnaud Doucet	Professeur associé à l'université de Colombie Britannique
M. Éric Moulines	Professeur à Télécom ParisTech
M. Gilles Pagès	Professeur à l'université Paris VI

et soutenue le 12 juillet 2010 devant

M. Amarjit Budhiraja	Professeur à « University of North Carolina at Chapel Hill »
M. Pierre Del Moral	Directeur de recherche à l'INRIA Bordeaux-Sud Ouest
Mme Sylvie Méléard	Professeur à l'école Polytechnique
M. Éric Moulines	Professeur à Télécom ParisTech
M. Gilles Pagès	Professeur à l'université Paris VI
M. Frédéric Patras	Directeur de recherche au CNRS, université de Nice-Sophia Antipolis
M. Denis Talay	Directeur de recherche à l'INRIA Sophia-Antipolis

Remerciements

Je voudrais remercier ici mes collègues et de manière générale tout ceux qui m'ont aidé et encouragé dans ma carrière de chercheur avec des mentions spéciales : aux rapporteurs pour leur célérité, aux membres du jury qui ont bien voulu se déplacer, à ceux qui n'ont pas pu se déplacer, à ma compagne et à ma famille, à mes collaborateurs pour leur incurable optimisme en toutes circonstances, à Philippe Maisonobe, directeur du laboratoire J. A. Dieudonné, pour son soutien constant.

Table des matières

1	Introduction	4
1.1	Liste des travaux présentés pour l'habilitation	4
1.1.1	Articles publiés	4
1.1.2	Chapitre de livre	5
1.1.3	Prépublications	5
1.2	Travaux non présentés	5
2	Filtrage, systèmes de particules	6
2.1	Définitions	6
2.1.1	Temps discret	6
2.1.2	Temps continu	7
2.2	Problématiques classiques	8
2.2.1	Stabilité du filtre par rapport à la mesure initiale	8
2.2.2	Calcul numérique du filtre optimal	9
2.2.3	Perspectives	11
2.3	Étude de la propagation du chaos.	12
2.3.1	Définitions en temps discret	12
2.3.2	Développement dans la propagation du chaos en temps discret	14
2.3.2.1	Développement	14
2.3.2.2	Combinatoire sur les arbres	15
2.3.3	Modèle génétique neutre	16
2.3.4	Propagation du chaos en temps continu	18
2.3.5	Étude de la propagation du chaos pour les systèmes de Bird et Nanbu	21
2.3.5.1	Présentation du système	21
2.3.5.2	Étude de la convergence	21
2.3.6	Perspectives	23
2.4	Application du filtrage à un problème d'optimisation de portefeuille	23
2.5	Cas d'une formule de Feynman-Kac dans laquelle le potentiel dépend de la trajectoire	26
2.5.1	Présentation du travail [11]	26
2.5.2	Perspectives	28
3	Processus stochastiques et calcul numérique	29
3.1	Approximation numérique d'une équation différentielle stochas- tique conduite par un processus de Lévy	29
3.1.1	Cas général, troncation des petits sauts	29

3.1.2	Amélioration de la vitesse de convergence dans le cas où Z est un processus subordonné	32
3.2	Optimisation de portefeuille	32
4	Articles isolés	34
4.1	Recuit simulé dans \mathbb{R}^d , application à la recherche du maximum de vraisemblance	34
4.1.1	Présentation du problème	34
4.1.2	Algorithme de recuit simulé dit « recuit simulé rapide »	34
4.1.3	Maximisation de la log-vraisemblance	36
4.2	E.D.S. pour la turbulence	37

Chapitre 1

Introduction

Depuis mon arrivée à l'université de Nice-Sophia Antipolis, j'ai travaillé en collaboration avec des mathématicien(ne)s d'horizon variés, sur des sujets variés. Le but de ce document est de souligner la cohérence de ces travaux.

1.1 Liste des travaux présentés pour l'habilitation

1.1.1 Articles publiés

- [1] Antonio CELANI, Sylvain RUBENTHALER et Dario VINCENZI : Dispersion and collapse in stochastic velocity fields on a cylinder. *138(4-5):579–597*, 2010.
- [2] Pierre DEL MORAL, Laurent MICLO, Frédéric PATRAS et Sylvain RUBENTHALER : The convergence to equilibrium of neutral genetic models. *Stochastic Anal. Appl.*, 28(1):123–143, 2010.
- [3] Sylvain RUBENTHALER, Tobias RYDÉN et Magnus WIKTORSSON : Fast simulated annealing in \mathbb{R}^d with an application to maximum likelihood estimation in state-space models. *Stochastic Process. Appl.*, 119(6):1912–1931, 2009.
- [4] Pierre DEL MORAL, Frédéric PATRAS et Sylvain RUBENTHALER : Tree based functional expansions for Feynman-Kac particle models. *Ann. Appl. Probab.*, 19(2):778–825, 2009.
- [5] Miguel MARTINEZ, Sylvain RUBENTHALER et Etienne TANRÉ : Approximations of a continuous time filter. Application to optimal allocation problems in finance. *Stoch. Anal. Appl.*, 27(2):270–296, 2009.
- [6] Nadia OUDJANE et Sylvain RUBENTHALER : Stability and uniform particle approximation of nonlinear filters in case of non ergodic signals. *Stoch. Anal. Appl.*, 23(3):421–448, 2005.
- [7] Sylvain RUBENTHALER et Magnus WIKTORSSON : Improved convergence rate for the simulation of stochastic differential equations driven by subordinated Lévy processes. *Stochastic Process. Appl.*, 108(1):1–26, 2003.

- [8] Sylvain RUBENTHALER : Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process. *Stochastic Process. Appl.*, 103(2):311–349, 2003.

1.1.2 Chapitre de livre

- [9] Pierre DEL MORAL, Frédéric PATRAS et Sylvain RUBENTHALER : A mean field theory of nonlinear filtering. In Dan CRISAN et Rozovskii BORIS, éditeurs : *Handbook on Nonlinear Filtering*. Oxford University Press.

1.1.3 Prépublications

- [10] Bruno RÉMILLARD et Sylvain RUBENTHALER : Optimal hedging in discrete and continuous time. Rapport technique G-2009-77, GERAD, HEC Montréal, November 2009.
- [11] Nicolas CHOPIN, Pierre DEL MORAL et Sylvain RUBENTHALER : Stability of Feynman-Kac formulae with path-dependent potentials. Rapport technique hal-00426415, HAL, CNRS, 2009.
- [12] Pierre DEL MORAL, Frédéric PATRAS et Sylvain RUBENTHALER : Convergence of U-statistics for interacting particle systems. Rapport technique hal-inria-00397366, HAL, INRIA, 2009.
- [13] Sylvain RUBENTHALER : Expansion of the propagation of chaos for Bird and Nanbu systems. Rapport technique hal-00355211, HAL, CNRS, 2009.

L'article [11] est accepté dans *Stochastic Processes and their Applications* sous réserve de modifications mineures. L'article [12] est accepté dans *Journal of Theoretical Probability* sous réserve de modifications mineures.

1.2 Travaux non présentés

- [14] Pierre DEL MORAL, Bruno RÉMILLARD et Sylvain RUBENTHALER : Monte carlo approximations of american options. Rapport technique 00001585-LG, Laboratoire J. A. Dieudonné, Université de Nice-Sophia Antipolis, 2006.
- [15] Pierre DEL MORAL, Bruno RÉMILLARD et Sylvain RUBENTHALER : *Une introduction aux probabilités*. Ellipses, 2006.

Chapitre 2

Filtrage, systèmes de particules

2.1 Définitions

2.1.1 Temps discret

Rappelons ici quelques notions relatives au filtrage non linéaire. Une partie de celles-ci se retrouvent dans le désormais classique ouvrage [Del04]. On se donne une collection d'espaces mesurables, des noyaux de Markov $(M_k)_{k \geq 1}$ (M_k va de E_{k-1} dans E_k) et des potentiels $(\psi_k)_{k \geq 0}$ ($\psi_k : E_k \rightarrow \mathbb{R}^+$). On dira que l'on est dans le cas homogène si $E_k = E$, $M_k = M$ indépendants de k . Soit π_0 une mesure de probabilité sur E_0 . On s'intéresse aux mesures suivantes (définies par leur action sur des fonctions tests) :

$$\gamma_k(f) = \int_{x_0, \dots, x_k} f(x_k) \pi_0(dx_0) \prod_{i=1}^k M_k(x_{i-1}, dx_i) \psi_i(x_i) \quad (2.1.1)$$

$$\eta_k(f) = \frac{\gamma_k(f)}{\gamma_k(1)}. \quad (2.1.2)$$

Les mesures $(\gamma_k)_{k \geq 0}$ sont appelées mesures de Feynman-Kac non normalisées ou filtres non normalisés. Les mesures $(\eta_k)_{k \geq 0}$ sont appelées mesures de Feynman-Kac (sous-entendu : normalisées) ou filtres (sous-entendu : normalisés). On peut donner une expression récursive de ces mesures

$$\begin{aligned} \eta_0 &= \gamma_0, \\ \eta_{k+1} &= \psi_{k+1} \cdot (\eta_k M_{k+1}), \end{aligned} \quad (2.1.3)$$

en utilisant les opérateurs sur les mesures

$$\begin{aligned} \mu M_k(dx_k) &= \int_{x \in E_{k-1}} \mu(dx_{k-1}) M_k(x_{k-1}, dx_k), \\ \psi_k \cdot \mu(dx_k) &= \frac{\psi_k(x_k) \times \mu(dx_k)}{\int_{E_k} \psi_k(z) \mu(dz)}. \end{aligned}$$

On dira que (η_k) est une suite de mesure de Feynman-Kac basée sur les noyaux (M_k) et les potentiels (ψ_k) .

Exemple 2.1.1 (Pistage radar). *Un exemple classique d'application est le pistage radar. Prenons $E_0 = E_1 = \dots = E, M_1 = M_2 = \dots = M$. Soit $(X_k)_{k \geq 0}$ une chaîne de Markov à valeurs dans E de transition M et de loi initiale π_0 . Soient $\forall k \geq 1, Y_k = h(X_k, \epsilon_k)$ avec des variables de bruit $(\epsilon_k)_{k \geq 1}$ i.i.d. (et indépendantes des autres variables). Les variables (Y_k) sont appelées observations. La suite (X_k) modélise la trajectoire d'un mobile. La suite (Y_k) modélise une série d'observations du mobile par un radar, ces observations sont bruitées.*

On suppose que, $\forall k$, la loi conditionnelle $\mathcal{L}(X_k|Y_k)$ admet une densité $x \mapsto g_k(x, Y_k)$. Posons $\forall k, \psi_k(x) = g_k(x, Y_k)$ (les observations Y_k sont fixées). On s'intéresse à la loi conditionnelle $\mathcal{L}_{\pi_0}(X_n|Y_1, \dots, Y_n)$. La formule de Bayes nous donne l'égalité :

$$\forall n \geq 1, \mathcal{L}_{\pi_0}(X_n|Y_1, \dots, Y_n) = \eta_n .$$

On parle aussi de « filtre optimal » à propos de η_n pour signifier que c'est la meilleure estimation de la position du mobile sachant les observation et connaissant le modèle d'évolution (on prend en compte toutes les observations).

Les mesures η_n n'ont en général pas d'expression analytique. L'unique exemple non trivial d'expression analytique est le filtre de Kalman : on prend, dans l'exemple ci-dessus, $E = \mathbb{R}$ (on peut aussi prendre \mathbb{R}^d), M un noyau gaussien et des observations de la forme $Y_k = aX_k + W_k$ avec des bruits i.i.d. (et indépendants des autres variables) de loi gaussienne. On montre dans ce cas que si π_0 est gaussienne alors η_n est gaussienne $\forall n$. On peut citer [CG06] pour un exemple de filtre restant dans une famille dénombrable.

2.1.2 Temps continu

Il existe une version en temps continu de cette théorie. Nous en donnons ici une version simplifiée applicable à des problèmes de pistage radar. On se donne un processus $(X_t)_{t \geq 0}$ dans un espace E , de loi initiale π_0 et un processus d'observations $(Y_t)_{t \geq 0}$ dans \mathbb{R}^d de la forme

$$Y_t = \int_0^t h(X_s) ds + W_t ,$$

avec $(W_t)_{t \geq 0}$ mouvement brownien dans \mathbb{R}^d . On définit $\forall t \geq 0$ les mesures γ_t et η_t par leur action sur des fonctions tests (la trajectoire $(Y_s)_{0 \leq s \leq t}$ est fixée) :

$$\gamma_t(f) = \mathbb{E}_{\pi_0} \left(f(U_t) \exp \left(\int_0^t h(U_s) dY_s - \int_0^t \|h(U_s)\|^2 ds \right) \right) , \quad (2.1.4)$$

$$\eta_t(f) = \frac{\gamma_t(f)}{\gamma_t(1)} , \quad (2.1.5)$$

où l'espérance est prise sur un processus $(U_s)_{s \geq 0}$ qui a même loi que $(X_s)_{s \geq 0}$ et $\|\cdot\|$ est la norme euclidienne sur \mathbb{R}^d . La mesure γ_t s'appelle la mesure de Feynman-Kac non normalisée et la mesure η_t s'appelle la mesure de Feynman-Kac normalisée (on peut remplacer « mesure de Feynman-Kac » par « filtre »). La formule suivante se démontre en utilisant le théorème de Girsanov

$$\mathcal{L}_{\pi_0}(X_t|(Y_s)_{0 \leq s \leq t}) = \eta_t .$$

2.2 Problématiques classiques

2.2.1 Stabilité du filtre par rapport à la mesure initiale

Les équations (2.1.1), (2.1.2) dépendent implicitement de la mesure initiale π_0 de la suite $(X_k)_{k \geq 0}$. Soit π'_0 une mesure de probabilité quelconque sur E_0 , définissons $\forall k, \gamma'_k, \eta'_k$ obtenu en comme γ_k, η_k en remplaçant π_0 par π'_0 dans (2.1.1), (2.1.2). On dit que le filtre est stable (par rapport à sa condition initiale) si

$$\lim_{n \rightarrow +\infty} \|\eta_n - \eta'_n\|_{VT} = 0. \quad (2.2.1)$$

On parle également d'oubli de la condition initiale.

Nous avons utilisé ici la norme en variation totale mais il existe d'autres définitions de la stabilité utilisant d'autres normes. On parle de stabilité exponentielle si

$$\limsup_{n \rightarrow +\infty} \frac{1}{n} \log(\|\eta_n - \eta'_n\|_{VT}) < 0. \quad (2.2.2)$$

Dans le cas de l'exemple 2.1.1, on suppose que le mobile entre dans le domaine de détection du radar au temps $k = 1$, la mesure π_0 représente une estimation à priori de la position du mobile. La propriété de stabilité signifie que cette estimée π_0 n'a pas une trop grande importance dans l'estimation asymptotique de la position du mobile. Nous venons ici de parler du temps discret, le même problème se pose en temps continu. On va supposer que le processus (X_k) est homogène, à valeurs dans E et de transition M , afin de simplifier les expressions.

La recherche d'hypothèses assurant la stabilité a donné lieu à une vaste littérature. Nous en donnons ici une version partielle et subjective dans le but de resituer mes travaux.

- Travaux théoriques en temps continu donnent des conditions de stabilité. Voir par exemple : [Kun71, Bud03].
- Travail donnant un résultat de stabilité sous une condition d'observabilité (en temps discret) : [vH09a]. Travail donnant un résultat de stabilité exponentielle sous une condition d'observabilité (en temps discret) : [BO99]. À chaque fois, on se place dans le cas du pistage radar, l'idée est que si les observations fournissent assez d'information sur le processus $(X_k)_{k \geq 0}$, on devrait avoir la stabilité.
- En temps discret, on peut rajouter l'hypothèse de mélange fort : $\exists \lambda$ mesure de probabilité sur E , $\epsilon > 0$ tels que

$$\forall x \in E, \epsilon \lambda(dy) \leq Q(x, dy) \leq \frac{1}{\epsilon} Q(x, dy). \quad (2.2.3)$$

Sous cette hypothèse, on montre la stabilité exponentielle. Voir par exemple [LO04, AZ97, DG01]. Il est intéressant de noter que ces articles utilisent des techniques différentes, passage par la métrique de Hilbert ou décomposition astucieuse des opérateurs, mais aboutissent à des bornes équivalentes. On obtient ici un peu plus que la stabilité exponentielle : $\exists C, \alpha > 0$, $\forall \pi_0, \pi'_0$

$$\|\eta_n - \eta'_n\|_{VT} \leq C e^{-\alpha n}. \quad (2.2.4)$$

L'idée est ici que si le signal (X_k) est « suffisamment ergodique », le filtre devrait oublier sa condition initiale. La condition d'ergodicité (2.2.3) retenue est très forte. Elle n'est pas vérifiée dans le cas du filtre de Kalman, pour des paramètres pour lesquels ce filtre est exponentiellement stable.

- Des travaux plus récents ([DFMP09, KV08]) utilisent un critère d'ergodicité moins fort pour montrer la stabilité exponentielle. Si le processus repasse régulièrement dans un ensemble « petite », et si on fait diverses hypothèses dont une hypothèse de « mélange fort local » dans le « petite ensemble », alors on a stabilité exponentielle. Voir [MT09] pour la notion de « petite ensemble ».

2.2.2 Calcul numérique du filtre optimal

Puisqu'il n'existe pas d'expression analytique pour η_k , on essaie en général de le calculer numériquement par une méthode de Monte-Carlo séquentielle. Nous décrivons ici un système de particules dont la mesure empirique approche le filtre optimal. Il existe de nombreuses variantes de cet algorithme (voir par exemple [DM00]). Ces algorithmes s'appellent algorithmes génétiques, filtres à particules, filtres particuliers ...

- En $k = 0$, on tire N v.a. (que l'on appelle particules) ξ_0^1, \dots, ξ_0^N i.i.d. de loi π_0 .
- Si on a $\xi_{k-1}^1, \dots, \xi_{k-1}^N$ au temps $k-1$, on tire $\hat{\xi}_k^1, \dots, \hat{\xi}_k^N$ i.i.d. suivant $\hat{\xi}_k^i \sim M_k(\xi_{k-1}^i, \cdot)$ ($1 \leq i \leq N$). Puis on pose $\omega_i = \frac{\psi_k(\hat{\xi}_k^i)}{\sum_{1 \leq j \leq N} \psi_k(\hat{\xi}_k^j)}$, on tire $\xi_k^i \sim \sum_{1 \leq r \leq N} \omega_r \delta_{\hat{\xi}_k^r}$ (i.i.d.).

Posons $\forall n, \eta_n^N = \frac{1}{N} \sum_{i=1}^N \delta_{\xi_n^i}$, on montre que η_n^N approche η_n en bien des sens (voir [Del04] pour de nombreux résultats), par exemple :

$$\sup_{\phi, \|\phi\|_\infty \leq 1} \mathbb{E}(|\eta_n(\phi) - \eta_n^N(\phi)|) \leq \frac{C_n}{\sqrt{N}} \quad (2.2.5)$$

avec une constante C_n qui dépend de n . Le problème est de savoir, dans le cas homogène, si on peut se passer de cette dépendance en n . On dira alors que l'approximation est stable en temps.

Introduisons, $\forall k$, les opérateurs sur les mesures μ sur E_{k-1} :

$$\mu Q_k(dy) = \int_{E_{k-1}} \mu(dx) \psi_k(y) M_k(x, dy),$$

(cet opérateur Q_k est linéaire)

$$R_k(\mu)(dy) = \frac{\mu Q_k(dy)}{\mu Q_k(E_k)},$$

$$S^N(\mu) = \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_i},$$

avec des (ξ_i) i.i.d. de loi μ (cet opérateur n'est donc défini que sur les mesures de probabilité). Ce dernier opérateur S^N est un opérateur aléatoire (on tire un N -échantillon de loi μ). Notons

$$R_k^N(\mu)(dy) = S^N(\psi_k \times S^N(\mu M_k)).$$

Nous avons $\eta_n^N = R_n^N \circ \dots \circ R_1^N(\pi_0)$. Notons, $\forall i \leq j, R_{i:j} = R_j \circ R_{j-1} \circ \dots \circ R_i$, $R_{i:j}^N = R_j^N \circ R_{j-1}^N \circ \dots \circ R_i^N$. On cherche à borner $\eta_n - \eta_n^N$ en écrivant une somme télescopique

$$\eta_n - \eta_n^N = \sum_{k=1}^n R_{i+1:n} \circ R_i \circ R_{1:i-1}^N(\pi_0) - R_{i+1:n} \circ R_i^N \circ R_{1:i-1}^N(\pi_0) \quad (2.2.6)$$

On reconnaît dans les termes $R_{i+1:n} \circ R_i \circ R_{1:i-1}^N(\pi_0) - R_{i+1:n} \circ R_i^N \circ R_{1:i-1}^N(\pi_0)$:

- une erreur locale $R_i \circ R_{1:i-1}^N(\pi_0) - R_i^N \circ R_{1:i-1}^N(\pi_0)$, c'est à dire une différence entre le « vrai » opérateur R_i et son approximation par des particules R_i^N (appliqués à la même mesure $R_{1:i-1}^N(\pi_0)$)
- un transport par $R_{i+1:n}$.

L'heuristique est la suivante : on sait que l'erreur locale est d'ordre $1/\sqrt{N}$. Si le filtre est exponentiellement stable alors les opérateurs R_i sont contractants et on va pouvoir borner la somme dans (2.2.6) par C/\sqrt{N} avec une constante C ne dépendant pas de N . Pour cela, un résultat du type (2.2.1) ou (2.2.2) n'est pas suffisant. Il faut un résultat du type (2.2.4). C'est pourquoi les premiers résultats de stabilité du filtre particulière ne sont énoncés que sous l'hypothèse (2.2.3) (voir par exemple [DG01, LO04]). Dans mon travail avec Nadia Oudjane ([6]), dans le cas homogène et pour la versions « pistage » du problème, nous avons montré le résultat suivant.

Théorème 2.2.1 (Oudjane, Rubenthaler [6]). *On se place dans le cas homogène avec $E = \mathbb{R}^d$. On suppose que l'on peut écrire*

$$\begin{cases} X_{k+1} &= f(X_k) + W_k \\ Y_k &= h(X_k) + V_k, \text{ pour } k \geq 1 \end{cases}$$

avec de bruits *i.i.d.* $(V_k)_{k \geq 1}$ et des bruits $(W_k)_{k \geq 0}$ indépendants conditionnellement à $(X_i)_{i \geq 0}$ tels que $\mathcal{L}(W_k | X_k = x)$ est de densité q_x avec un encadrement par deux densités $(q_1, q_2) : \underline{m}q_1 \leq q_x \leq \overline{m}q_2$ (ceci inclut des bruits *i.i.d.*). On suppose que f est lipschitz, h est surjective avec des constantes b_0, b_1

$$h(x_1) = y_1, h(x_2) = y_2 \Rightarrow \|x_1 - x_2\| \leq b_1 + b_2 \|y_1 - y_2\|$$

(si h est inversible, ceci veut dire que h^{-1} est lipschitz).

- *Stabilité du filtre.* Si π_0 et π'_0 sont comparables (c'est à dire qu'il existe $\epsilon > 0$ tel que $\epsilon\pi_0 \leq \pi'_0 \leq \frac{1}{\epsilon}\pi_0$),

$$\forall n, \mathbb{E}(\|\eta_n - \eta'_n\|_{VT} | Y_1, \dots, Y_n) \leq t(n)$$

avec une fonction $t(n)$ explicite, ne dépendant pas des $(Y_i)_{1 \leq i \leq n}$. Dans des cas où la queue de distribution des (V_k) est « petite » devant celle des (W_k) , on montre que $t(n) \xrightarrow{n \rightarrow +\infty} 0$, à une vitesse qui peut être exponentielle, ou moins bonne.

- *Approximation uniforme par un filtre particulière.* Il existe une variation du filtre particulière η_n^N qui approche le filtre optimal avec une borne du genre

$$\sup_{\phi, \|\phi\|_\infty \leq 1} \mathbb{E}(|\eta_n(\phi) - \eta_n^N(\phi)| | Y_1, \dots, Y_n) \leq r(N),$$

avec une fonction $r(N)$ explicite, ne dépendant pas des $(Y_i)_{1 \leq i \leq n}$. Dans où la queue de distribution de (V_k) est « petite » devant celle des (W_k) , on montre que $r(N)$ est une puissance de N .

Exemple 2.2.2. – Dans le théorème ci-dessus, si les (V_k) ont des queues gaussiennes et les (W_k) ont des queues exponentielles (ce qui veut dire qu’elles décroissent respectivement en $x \mapsto e^{-C\|x\|^2}$ et $x \mapsto e^{-C\|x\|}$), alors on obtient

1. $\frac{t(n)}{n^K} \xrightarrow{n \rightarrow +\infty} 0, \forall K > 0$
2. $r(N)N^{1/2-\epsilon} \xrightarrow{N \rightarrow +\infty} 0, \forall \epsilon > 0.$

– Si $E = \mathbb{R}$ et les lois des (V_k) et (W_k) sont gaussiennes (de variances respectives s^2, σ^2), on trouve une constante explicite C telle que $\sigma/s > C$ implique le point 2 ci-dessus.

Ce résultat est le quasiment seul résultat de stabilité du filtre particulaire au-delà de l’hypothèse de mélange fort (2.2.3). Les articles [HC08, CH08] prolongent et améliorent ces techniques. La démonstration s’appuie sur une décomposition des opérateurs R_n que l’on trouve dans [DG01] pour montrer la propriété de contraction de ces opérateurs. Le résultat ne serait pas très différent si on utilisait des techniques de métrique de Hilbert pour montrer la contraction de ces opérateurs. On tronque l’espace autour des observations (si on se place dans le cas $Y_k = X_k + V_k, h = Id$, pour simplifier). On obtient alors des opérateurs R_n^Δ qui sont contractants (mais d’autant moins que Δ est grand) et qui sont proches des R_n^Δ (d’autant plus que Δ est grand). Ces opérateurs R_n^Δ peuvent donc être approchés uniformément en temps par des filtres particulières $R_n^{\Delta, N}$. Pour montrer la stabilité des R_n , on forme :

$$\begin{aligned} R_n(\pi_0) - R_n(\pi'_0) &= R_n(\pi_0) - R_n^\Delta(\pi_0) + R_n^\Delta(\pi_0) - R_n^\Delta(\pi'_0) \\ &\quad + R_n^\Delta(\pi'_0) - R_n(\pi'_0). \end{aligned}$$

Puis on choisit $\Delta(n)$ tel que les trois différences ci-dessus $\xrightarrow{n \rightarrow +\infty} 0$ dans un sens à préciser. On cherche ensuite à comparer $R_n(\pi_0)$ à un filtre particulaire : $R_n(\pi_0) - R_n^{\Delta, N}(\pi_0)$. On choisit ensuite $\Delta(n)$ tel que ces quantités $\xrightarrow{N \rightarrow +\infty} 0$ dans un sens à préciser. De tels choix de $\Delta(n)$ sont possibles sous les hypothèse que nous faisons.

2.2.3 Perspectives

Les résultats de convergence du filtre particulaire connus sont du genre :

- [6, CH08] sous des critères de comparaison des queues de distribution du bruit d’évolution et du bruit d’observation.
- [LO04, DG01] sous une hypothèse de mélange fort pour le signal.
- [vH09a, vH09b] sous des hypothèse variées sans majoration explicite en le nombre de particules N .

Il existe des résultats de stabilité du filtre optimal sous condition d’ergodicité du signal ([KV08, DFMP09]) mais ils ne sont pas écrits d’une manière qui permette d’utiliser la somme télescopique (2.2.6) pour en déduire l’uniforme convergence du filtre particulaire. Une piste de recherche évidente est la stabilité du filtre particulaire sous condition d’ergodicité du signal.

2.3 Étude de la propagation du chaos.

2.3.1 Définitions en temps discret

Dans les articles ([4, 12]), je me suis intéressé avec mes collaborateurs à l'étude de la propagation du chaos pour des systèmes génétiques. Reprenons le formalisme de 2.1, définissons les opérateurs ($\forall k \geq 1$)

$$\mu\hat{Q}_k = (\psi_{k-1} \times \mu)M_k$$

(agissant donc sur μ mesure sur E_{k-1}). On s'intéresse aux mesures définies par récurrence

$$\begin{aligned}\hat{\gamma}_1 &= \pi_0 M_1, \hat{\gamma}_{k+1} = \hat{\gamma}_k \hat{Q}_k, \\ \hat{\eta}_k &= \frac{1}{\hat{\gamma}_k(E_k)} \hat{\gamma}_k.\end{aligned}$$

Remarquons que $\hat{\eta}_1 = \hat{\gamma}_1$. Ces mesures sont encore des mesures de Feynman-Kac. Les résultats sur (η_k) s'appliquent à $(\hat{\eta}_k)$ et vice-versa. Dans cette partie, il sera plus commode de travailler avec les $(\hat{\eta}_k)$. On peut les approcher par la mesure empirique du système suivant :

- En $k = 1$, on tire Z_1^1, \dots, Z_1^N i.i.d. de loi $\hat{\gamma}_1$.
- Si on a Z_k^1, \dots, Z_k^N au temps k , on pose $\omega_j = \frac{\psi_k(Z_k^j)}{\sum_{1 \leq j \leq N} \psi_k(Z_k^j)}$ ($1 \leq j \leq N$), on tire de manière i.i.d

$$\tilde{Z}_k^r \sim \sum_{1 \leq j \leq N} \omega_j \delta_{Z_k^j}, \quad 1 \leq r \leq N$$

puis on tire de manière i.i.d. $Z_{k+1}^r \sim M_{k+1}(\tilde{Z}_{k+1}^r, \cdot)$ ($1 \leq r \leq N$).

Comme précédemment, $\hat{\eta}_k^N = \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{Z_k^i}$ est une approximation de $\hat{\eta}_k$ ($\forall k$). Définissons $\forall n$ les opérateur sur les mesures sur E_{n-1}

$$\begin{aligned}\mu\hat{Q}_n &= (\psi_{n-1} \times \mu)M_n, \\ \hat{R}_n(\mu) &= (\psi_{n-1} \cdot \mu)M_n, \\ \hat{R}_n^N(\mu) &= S^N((\psi_{n-1} \cdot \mu)M_n),\end{aligned}$$

Nous avons $\hat{\eta}_n^N = \hat{R}_n^N(\hat{\eta}_{n-1}^N)$.

Pour une mesure empirique sur N points x_1, \dots, x_N

$$m(x) = \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{x_i},$$

on remarque que le produit tensoriel peut s'écrire ($\forall q$)

$$m(x)^{\otimes q} = \frac{1}{N^q} \sum_{a \in [N]^q} \delta_{x_{a(1)}, \dots, x_{a(q)}},$$

où $\forall j \in \mathbb{N}^*$, $[j] = \{1, \dots, j\}$ et on définit

$$m(x)^{\odot q} = \frac{1}{(N)_q} \sum_{a \in \langle q, N \rangle} \delta_{x_{a(1)}, \dots, x_{a(q)}},$$

où $\langle q, N \rangle = \{a : [q] \rightarrow [N], \text{injective}\}$, $(N)_q = \text{card}(\langle q, N \rangle) = \frac{N!}{(N-q)!}$. La propriété de propagation du chaos est la convergence suivante ($\forall q$, à n fixé)

$$\mathcal{L}(Z_n^1, \dots, Z_n^N) \xrightarrow{N \rightarrow +\infty} \hat{\eta}_n^{\otimes q},$$

où « \Rightarrow » signifie la convergence étroite. Puisque les particules sont échangeables, on peut aussi écrire cette propriété sous la forme

$$\mathbb{E}((\hat{\eta}_n^N)^{\odot q}(F)) \xrightarrow{N \rightarrow +\infty} \hat{\eta}_n^{\otimes q}(F), \quad \forall F \in \mathcal{B}_b(E_n),$$

où pour tout espace mesurable E , $\mathcal{B}_b(E)$ est l'ensemble des fonctions mesurables bornées sur E . Nous noterons dans la suite $\mathbb{P}_{n,q}^N(F) = \mathbb{E}((\hat{\eta}_n^N)^{\odot q}(F))$ ($\forall F$).

Les symboles $\otimes q$ et $\odot q$ sont liés par des formules combinatoires. Définissons la combinaison linéaire d'applications :

$$L_q^N = \frac{1}{N^q} \sum_{a \in [q]^{[q]}} \frac{(N)_{|a|}}{(q)_{|a|}} a,$$

où

$$|a| = \text{card}(\text{Im}(a)).$$

Définissons l'opérateur de Markov (par son action sur les fonctions de q variables) de E^q dans E^q (son action est suffisamment générale pour que nous n'ayons pas besoin de préciser ce qu'est l'espace E)

$$D_{L_q^N}(F)(x_1, \dots, x_q) = \frac{1}{N^q} \sum_{a \in [q]^{[q]}} \frac{(N)_{|a|}}{(q)_{|a|}} F(x_{a(1)}, \dots, x_{a(q)}). \quad (2.3.1)$$

Cet opérateur peut-être vu comme une combinaison linéaire d'opérateurs de coalescence. Nous avons la formule ($\forall x_1, \dots, x_N$)

$$m(x)^{\otimes q} = m(x)^{\odot q} D_{L_q^N}. \quad (2.3.2)$$

Rappelons la formule ($\forall k \geq 1$)

$$\hat{\gamma}_k(1) = \prod_{1 \leq i \leq k-1} \hat{\eta}_i(\psi_i).$$

C'est cette formule qui induit les définitions suivantes (dans l'ordre)

$$\hat{\gamma}_k^N(1) = \prod_{1 \leq i \leq k-1} \hat{\eta}_i^N(\psi_i),$$

$$\hat{\gamma}_i^N = \hat{\gamma}_k^N(1) \times \hat{\eta}_k^N.$$

Nous définissons également ($\forall q, \forall F, \forall n$)

$$(\hat{\gamma}_n^N)^{\otimes q}(F) = (\hat{\gamma}_n^N(1))^q (\hat{\eta}_n^N)^{\otimes q}(F), \quad (2.3.3)$$

$$(\hat{\gamma}_n^N)^{\odot q}(F) = (\hat{\gamma}_n^N(1))^q (\hat{\eta}_n^N)^{\odot q}(F). \quad (2.3.4)$$

2.3.2 Développement dans la propagation du chaos en temps discret

2.3.2.1 Développement

Définition 2.3.1. Soit $(\Theta^N)_{N \geq 0}$ une suite de mesures uniformément bornées en variation totale sur E^q ($q \in \mathbb{N}^*$). La suite est dite converger faiblement vers Θ quand $N \rightarrow +\infty$ si

$$\forall F \in \mathcal{B}_b(E^q), \lim_{N \rightarrow +\infty} \Theta^N(F) = \Theta(F).$$

Si $(\Theta^N)_{N \geq 0}$ converge faiblement vers Θ quand $N \rightarrow +\infty$. La dérivée discrète de la suite $(\Theta^N)_{N \geq 0}$ est la suite de mesures $(\partial\Theta^N)_{N \geq 1}$ définie par

$$\partial\Theta^N := N[\Theta^N - \Theta].$$

Nous dirons que Θ^N est différentiable si $(\partial\Theta^N)_{N \geq 1}$ est uniformément bornée et si converge fortement vers une mesure $\partial\Theta$ appelée dérivée de Θ^N quand $N \rightarrow +\infty$.

Si $(\partial\Theta^N)$ est elle-même dérivable, on notera sa dérivée $\partial^2\Theta$ (et ainsi de suite).

Une suite (Θ^N) est dérivable jusqu'à l'ordre $k+1$ à la représentation suivante

$$\Theta^N = \sum_{0 \leq l \leq k} \frac{1}{N^l} \partial^l \Theta + \frac{1}{N^{k+1}} \partial^{k+1} \Theta^N$$

avec $\sup_{N \geq 1} \|\partial^{k+1} \Theta^N\|_{VT} < \infty$ et la convention $\partial^0 \Theta = \Theta$ pour $l = 0$.

Le théorème suivant nous dit que la suite $(\mathbb{P}_{k,q}^N)$ est dérivable à tout ordre.

Théorème 2.3.2 (Del Moral, Patras, Rubenthaler [4]). Pour tous m, q, k, F

$$\mathbb{P}_{k,q}^N(F) = \hat{\eta}_k^{\otimes q}(F) + \sum_{i=1}^m \frac{1}{N^i} \partial^i \mathbb{P}_{k,q}(F) + \frac{1}{N^{m+1}} \partial^{m+1} \mathbb{P}_{k,q}^N, \quad (2.3.5)$$

avec des mesures signées $\partial^i \mathbb{P}_{k,q}$ et $\partial^{m+1} \mathbb{P}_{k,q}^N$, et cette dernière est telle que $\sup_N \|\partial^{m+1} \mathbb{P}_{k,q}^N\|_{VT} < \infty$. Les mesures signées ci-dessus peuvent être explicitées comme des sommes de mesures sur des forêts d'arbre de coalescence (nous précisons ces notions plus loin).

Une conséquence de ce théorème est la proposition suivante. Soit

$$\begin{aligned} \mathcal{B}_0^{sym}(E^q) &= \{F \in \mathcal{B}_b(E^q), \text{ symétrique en les } q \text{ variables,} \\ &\int_E F(x_1, \dots, x_q) \eta_t(dx_q) = 0, \forall x_1, \dots, x_{q-1}\} \end{aligned}$$

Proposition 2.3.3 (Del Moral, Patras, Rubenthaler [4]). Pour toute $F \in \mathcal{B}_0^{sym}(E_n^q)$, si $k < q/2$ alors $\partial^k \mathbb{P}_{n,q}(F) = 0$ et $\partial^{q/2} \mathbb{P}_{n,q}(F) = 0$ a une forme explicite que nous n'explicitons pas ici.

Nous allons donner ici une idée du déroulement de la démonstration du théorème ci-dessus. Ce théorème et la proposition qui le suit permettent de retrouver

des résultats de convergence du filtre particulaire. Les techniques utilisées sont les mêmes que dans partie 2.3.1. Nous renvoyons donc à cette partie pour les applications de ce théorème.

À cause des relations (2.3.3), (2.3.4), on commence par écrire un développement similaire à (2.3.5) pour $\mathbb{Q}_{k,q}^N(F) := \mathbb{E}((\hat{\gamma}_k^N)^{\otimes q}(F))$ (la mesure $\mathbb{Q}_{k,q}^N$ est définie par son action sur des fonctions tests). Le calcul fondamental est le suivant

$$\begin{aligned} \mathbb{E}((\hat{\gamma}_n^N)^{\otimes q}(F)|(Z_{n-1}^i)_{1 \leq i \leq N}) &= (\hat{\gamma}_n^N(1))^q \mathbb{E}((\hat{\eta}_n^N)^{\otimes q}(F)|(Z_{n-1}^i)_{1 \leq i \leq N}) \\ &= (\hat{\gamma}_n^N(1))^q \mathbb{E}((\hat{\eta}_n^N)^{\otimes q} D_{L_q^N}(F)|(Z_{n-1}^i)_{1 \leq i \leq N}) \\ &= (\hat{\gamma}_n^N(1))^q \mathbb{E}((\hat{R}_n(\hat{\eta}_{n-1}^N))^{\otimes q} D_{L_q^N}(F)|(Z_{n-1}^i)_{1 \leq i \leq N}) \\ &= \mathbb{E}((\hat{\gamma}_{n-1}^N)^{\otimes q} \hat{Q}_{n-1}^{\otimes q} D_{L_q^N}(F)|(Z_{n-1}^i)_{1 \leq i \leq N}). \end{aligned} \quad (2.3.6)$$

Une récurrence nous donne

$$\begin{aligned} \mathbb{E}((\hat{\gamma}_n^N)^{\otimes q}(F)) &= (\hat{\gamma}_1)^{\otimes q} \hat{Q}_2^{\otimes q} D_{L_q^N} \dots \hat{Q}_n^{\otimes q} D_{L_q^N}(F) \\ (\text{par (2.3.1)}) &= \frac{1}{N^{q(n-1)}} \sum_{a \in \mathcal{A}_{n-1,q}} \frac{(N)_{|a|}}{(q)_{|a|}} \Delta_{n,q}^a(F), \end{aligned} \quad (2.3.7)$$

où $\mathcal{A}_{n-1,q} = \{(a_1, \dots, a_{n-1}), a_1, \dots, a_{n-1} \in [q]^{[q]}\}$, si $a = (a_1, \dots, a_{n-1}) \in \mathcal{A}_{n-1,q}$, $(N)_{|a|} = (N)_{|a_1|} \dots (N)_{|a_{n-1}|}$, $\Delta_{n,q}^a(F) = (\hat{\gamma}_1)^{\otimes q} \hat{Q}_2^{\otimes q} D_{a_1} \dots \hat{Q}_n^{\otimes q} D_{a_{n-1}}$. Pour $a_i \in [q]^{[q]}$, nous définissons le degré de coalescence par $\text{coal}(a_i) = q - \text{card}(\text{Im}(a_i))$. Et pour $a = (a_1, \dots, a_{n-1})$, $\text{coal}(a) = \sum_{1 \leq i \leq n-1} \text{coal}(a_i)$. Si on regroupe les termes de (2.3.7) en puissances de N , on remarque que le terme en facteur de $1/N^k$ est une combinaison linéaire des $\Delta_{n,q}^a$ avec $\text{coal}(a) \leq k$. Par ailleurs, $(\mathcal{S}_q)^n$ agit sur $\mathcal{A}_{n-1,q}$ par $(\mathcal{S}_q$ étant le groupe symétrique d'ordre q) :

$$(s_1, \dots, s_n) \cdot (a_1, \dots, a_{n-1}) = (s_1 \circ a_1 \circ s_2^{-1}, \dots, s_{n-1} \circ a_{n-1} \circ s_n^{-1}). \quad (2.3.8)$$

Les mesures $\Delta_{n,q}^a$ sont invariantes sous cette action. On regroupe donc les termes de (2.3.7) en somme sur les classes d'équivalence sous l'action de $(\mathcal{S}_q)^n$. Il faut donc calculer le cardinal de ces classes (que nous appellerons forêts dans la suite). C'est l'objet du théorème 3.8 p. 800 de [4] que nous expliquons dans la sous-section suivante. Son énoncé fait appel à des notations de combinatoires compliquées.

2.3.2.2 Combinatoire sur les arbres

Les arbres sont des graphes abstraits (avec des sommets et des arêtes) qui n'ont pas de lacet. On s'intéresse aux arbres enracinés, c'est à dire tels que l'un des sommets est différencié des autres et s'appelle « racine ». Dans toute la suite, les arbres seront enracinés. Une forêt \mathbf{f} est un produit monomial commutatif d'arbres

$$\mathbf{f} = T_1^{m_1} \dots T_k^{m_k}$$

pour k arbres T_1, \dots, T_k . Une forêt \mathbf{f} est partitionnée en niveau de manière naturelle (le niveau 0 est le niveau des racines des arbres, le niveau 1 est celui des sommets à distance 1 d'une racine, ...). La hauteur d'une forêt est le niveau le plus haut de la forêt. Pour $n \geq 0$, $q_0, \dots, q_n \in \mathbb{N}^*$, notons $\mathcal{F}_{q_0, \dots, q_n}^n$ l'ensemble

des forêts de hauteur n tels que $\forall k \in \{0, \dots, n\}$, le cardinal des sommets à la hauteur k est égal à q_k . L'opérateur B sur les arbres est l'opérateur qui enlève la racine (il s'étend aux forêts). L'opérateur B^{-1} est l'opérateur qui transforme une forêt en un arbre en ajoutant une racine commune à toutes les racines des arbres constituant la forêt. Les forêts de $\mathcal{F}_{q_0, \dots, q_n}^n$ sont en bijection avec les classes d'équivalence de $\{(a_0, \dots, a_{n-1}), a_i \in [q_i]^{[q_{i+1}]}\}$ sous l'action de $\mathcal{S}_{q_0} \times \dots \times \mathcal{S}_{q_n}$ (action similaire à celle décrite dans (2.3.8)). Notons $\#(\mathbf{f})$ le cardinal de la classe d'équivalence associée à \mathbf{f} . D'après la sous-section précédente, nous voulons ce cardinal pour \mathbf{f} avec le même nombre q d'individus à chaque niveau.

Définition 2.3.4. Soit $\mathbf{f} = T_1^{m_1} \dots T_k^{m_k}$ une forêt, le k -uplet (m_1, \dots, m_k) (non ordonné) est appelé le multi-ensemble symétrique de l'arbre $T = B^{-1}(\mathbf{f})$ et nous écrivons

$$\mathbf{s}(T) = \mathbf{s}(B^{-1}(T_1^{m_1} \dots T_k^{m_k})) = (m_1, \dots, m_k).$$

Le multi-ensemble symétrique de la forêt $\mathbf{f} = T_1^{m_1} \dots T_k^{m_k}$ est l'union disjointe des multi-ensembles symétriques de ses arbres

$$\mathbf{s}(T_1^{m_1} \dots T_k^{m_k}) = \left(\underbrace{\mathbf{s}(T_1), \dots, \mathbf{s}(T_1)}_{m_1 \text{ termes}}, \dots, \underbrace{\mathbf{s}(T_k), \dots, \mathbf{s}(T_k)}_{m_k \text{ termes}} \right).$$

Proposition 2.3.5 (Del Moral, Patras, Rubenthaler [4]). Pour tous $n \geq 0$, $q_0, \dots, q_n \in \mathbb{N}^*$, $\forall \mathbf{f} \in \mathcal{F}_{q_0, \dots, q_n}^n$,

$$\#(\mathbf{f}) = \frac{q_0! q_1! \dots q_n!}{\prod_{i=-1}^n \mathbf{s}(B^i(\mathbf{f}))!},$$

avec la notation $(m_1, \dots, m_k)! = m_1! \dots m_k!$ pour définir la factorielle d'un multi-ensemble.

Ce calcul suffit à classer les termes de (2.3.7) par puissances de N , et donc à obtenir un développement pour $\mathbb{Q}_{k,q}^N$. Pour obtenir un développement pour $\mathbb{P}_{k,q}^N$, nous avons développé dans [4] une combinatoire plus compliquée à base d'arbres ayant des feuilles colorées.

La publication [9] est un article de synthèse sur la propagation du chaos pour les filtres particulières.

2.3.3 Modèle génétique neutre

Dans le cas homogène, le système de particules décrit dans la partie 2.3.1 modélise l'évolution d'une population de taille fixée soumise à un noyau de mutation $M = M_k$ ($\forall k$) et à un potentiel $\psi = \psi_k$ ($\forall k$). Les individus de la population au temps $k-1$ qui ont le plus grand potentiel ψ_{k-1} sont ceux qui ont la plus grande probabilité d'avoir des descendants (on dit que l'individu r descend de l'individu j si $\tilde{Z}_k^r = Z_{k-1}^j$). Dans le cas où $\psi = 1$, ce système de particules s'appelle le modèle de Wright-Fisher (voir par exemple [Dur02, Ewe04]). Dans [2], nous décrivons la mesure invariante d'un tel système (à N fixé) en utilisant les outils développés dans [4].

Soit $Z_n^{\mathbf{i}_n}$ avec $\mathbf{i}_n = (i_1, \dots, i_n) \in [N]^n$, $n \geq 1$, une collection de v.a. à valeurs dans E définies par récurrence

- $Z_1^{\mathbf{i}_1}$ avec $\mathbf{i}_1 \in [N]$ est une collection de N variables i.i.d. de loi η

- Si les variables $Z_{n-1}^{\mathbf{i}_{n-1}}$ sont données, les variables $Z_n^{\mathbf{i}_n}$ avec $\mathbf{i}_n = (\mathbf{i}_{n-1}, i_n)$ sont une collection de N variables i.i.d. de loi $\delta_{Z_{n-1}^{\mathbf{i}_{n-1}}} M$.

Posons, pour $(b_2, \dots, b_n) \in \mathcal{A}_{n-1, N}$ telles que $\forall i, \exists c_i, b_{i-1} = c_i \circ b_i$ (on dira que la suite (b_2, \dots, b_n) est faiblement décroissante)

$$\mathcal{X}_n^{(b_2, \dots, b_n)} = (Z_n^{b_n(i), \dots, b_2(i)})_{1 \leq i \leq N}.$$

Avec cette notation, nous avons $\forall F$

$$\mathbb{E}(\hat{\eta}_n^N(F)) = \mathbb{E}_{\hat{\gamma}_1}(F(\mathcal{X}_n^{A_{2,n}, A_{3,n}, \dots, A_n})),$$

où les variables $(A_p)_{2 \leq p \leq n}$ sont i.i.d. de loi uniforme dans $\mathcal{A} := [N]^{[N]}$, $A_{p,n} = A_p \circ A_{p+1} \circ \dots \circ A_n$. L'indice $\hat{\gamma}_1$ dans la formule ci-dessus signifie que les $Z_1^{\mathbf{i}_1}$ forment une famille i.i.d. de loi $\hat{\gamma}_1$. Ce résultat est une conséquence du calcul (2.3.6). Soit $(B_p)_{p \geq 2}$ la chaîne de Markov dans \mathcal{A} faiblement décroissante définie par : $B_2 = A_2, B_{p+1} = A_{p+1}B_p$. Soit

$$\partial\mathcal{A} = \{a \in \mathcal{A} : |a| = 1\}.$$

Cet ensemble est absorbant pour la chaîne (B_p) . Soit T le temps d'entrée de la chaîne dans $\partial\mathcal{A}$. Les théorèmes suivant nous donnent l'existence d'une mesure stationnaire pour la chaîne $((Z_n^i)_{1 \leq i \leq N})_{n \geq 2}$ (à valeurs dans E^N), une description de cette mesure stationnaire, ainsi que la vitesse de convergence vers cette mesure stationnaire.

Théorème 2.3.6 (Del Moral, Miclo, Patras, Rubenthaler [2]). *La chaîne (B_p) est absorbée en temps fini par $\partial\mathcal{A}$ et B_T est uniformément distribué dans $\partial\mathcal{A}$. Nous avons de plus, $\forall n \geq 2$*

$$\mathbb{P}(T > n) \leq K \left(\frac{n-2}{N} \vee 1 \right) \exp \left(- \left(\frac{n-2}{N} - 1 \right)_+ \right)$$

avec

$$K = e \prod_{l \in \mathbb{N}^*} \left(1 - \frac{2}{(l+1)(l+2)} \right)^{-1}.$$

Notons $\forall F$,

$$\Gamma_{\hat{\gamma}_1}(F) = \mathbb{E}_{\hat{\gamma}_1}(F(\mathcal{X}_T^{B_T, \dots, B_2})).$$

Théorème 2.3.7 (Del Moral, Miclo, Patras, Rubenthaler [2]). *La mesure $\Gamma_{\hat{\gamma}_1}$ est une mesure invariante pour la chaîne de Markov $((Z_n^i)_{1 \leq i \leq N})_{n \geq 2}$ (à valeurs dans E^N). Si nous faisons de plus l'hypothèse que $\exists \lambda > 0, \delta < \infty$ tels que $\forall n \geq 0$,*

$$\beta(M^n) := \sup_{(x,y) \in E^2} \|M^n(x, \cdot) - M^n(y, \cdot)\|_{VT} \leq \delta e^{-\lambda n},$$

alors $\exists K' > 0$ (constante universelle) telle que $\forall n \geq N + \frac{1}{\lambda} + 2$,

$$\|\hat{\eta}_n^N - \Gamma_{\hat{\gamma}_1}\|_{VT} \leq \delta K' \frac{n}{N} \exp \left(- \frac{n-2}{N + \frac{1}{\lambda}} \right).$$

Si η est une mesure sur E , et (b_2, \dots, b_n) est une suite faiblement décroissante de $[N]^{[N]}$, on note $(\forall F)$

$$\eta_{(b_2, \dots, b_n)}(F) = \mathbb{E}_\eta(F(\mathcal{X}_n^{b_2, \dots, b_n})).$$

On note

$$\begin{aligned} \rho(b_2, \dots, b_n) &= (|b_2|, \dots, |b_n|), \\ |\rho(b_2, \dots, b_n)| &= |b_2| \times \dots \times |b_n|, \end{aligned}$$

$$\begin{aligned} (N)_{\rho(b_2, \dots, b_n)} &= (N)_{|b_2|} \times \dots \times (N)_{|b_n|}, \\ N^{|\rho(b_2, \dots, b_n)|} &= N^{|b_2|} \times \dots \times N^{|b_n|}. \end{aligned}$$

Ces définitions s'étendent naturellement aux classes d'équivalences de $\{(b_2, \dots, b_n), b_i \in [N]^{[N]}\}$ sous l'action de $(\mathcal{S}_N)^{n-1}$ (action décrite dans (2.3.8)). On identifie ces classes d'équivalence à des forêts. Pour \mathbf{t} un arbre, on appelle $ht(\mathbf{t})$ sa hauteur. Pour $1 \leq q_1 \leq \dots \leq q_n$, on note $\mathcal{T}_n(q_1, \dots, q_n)$ l'ensemble des arbres de hauteur n avec q_i sommets au niveau i . Notons

$$\mathcal{T} = \cup_{n \geq 0} \cup_{1 \leq q_1 \leq \dots \leq q_n} \mathcal{T}_n(q_1, \dots, q_n).$$

La proposition suivante décompose la mesure stationnaire $\Gamma_{\hat{\gamma}_1}$ en combinaison linéaires sur les généalogies possibles jusqu'au premier ancêtre commun des N particules.

Proposition 2.3.8 (Del Moral, Miclo, Patras, Rubenthaler [2]).

$$\Gamma_{\hat{\gamma}_1} = \sum_{\mathbf{t} \in \mathcal{T}} \frac{(N)_{\rho(\mathbf{t})}}{N^{|\rho(\mathbf{t})|} \prod_{1 \leq i \leq ht(\mathbf{t})} \mathbf{s}(B^i(\mathbf{t}))!} (\hat{\gamma}_1)_{\mathbf{t}}(F).$$

2.3.4 Propagation du chaos en temps continu

Soit $(X_t)_{t \geq 0}$ un processus de Markov dans $E = \mathbb{R}^d$ de générateur (possiblement inhomogène) L_t . On suppose que (X_t) a des trajectoires continues. On se donne une famille de fonctions potentielles bornées $(V_t)_{t \geq 0}$, $0 \leq V_t \leq V_\infty$. Notons $\gamma_0 = \eta_0$ la loi de X_0 . Nous nous intéressons aux flots $(\gamma_t)_{t \geq 0}$, $(\eta_t)_{t \geq 0}$ solutions de (pour des fonctions tests f suffisamment régulières)

$$\begin{aligned} \frac{d}{dt} \gamma_t(f) &= \gamma_t(L_t(f)) - \gamma_t(f), \\ \frac{d}{dt} \eta_t(f) &= \eta_t(L_t(f)) + \eta_t(f(\eta_t(V_t) - V_t)). \end{aligned}$$

Nous avons

$$\gamma_t(f) = \mathbb{E} \left(f(X_t) \exp \left(- \int_0^t V_s(X_s) ds \right) \right), \quad \eta_t(f) = \frac{\gamma_t(f)}{\gamma_t(1)}.$$

Il existe un système de particules qui approche η_t . En $t = 0$, on prend un vecteur aléatoire $\Xi_t = (\xi_t^1, \dots, \xi_t^N)$ une famille i.i.d. de loi η_0 . Chaque entrée ξ_t^i diffuse suivant le générateur L_t indépendamment des autres entrées. Chaque entrée a une horloge exponentielle de paramètre V_∞ (indépendante des variables définies jusqu'à maintenant). Quand l'horloge de ξ_t^i sonne, par exemple en τ , alors ξ_t^i peut

- sauter aléatoirement (et uniformément) sur une particule de la famille (y compris elle-même), avec une probabilité $\frac{V_\tau(\xi_\tau^i)}{V_\infty}$
- rester au même endroit avec probabilité $1 - \frac{V_\tau(\xi_\tau^i)}{V_\infty}$.

Définissons

$$\begin{aligned}\eta_t^N &= m(\Xi_t) = m(\xi_t^1, \dots, \xi_t^m), \\ \gamma_t^N &= \gamma_t^N(1) \times \eta_t^N, \text{ avec } \gamma_t^N(1) = \exp\left(-\int_0^t \eta_s^N(V_s) ds\right), \\ (\gamma_t^N)^{\odot q} &= (\gamma_t^N(1))^q \times (\eta_t^N)^{\odot q}.\end{aligned}$$

Introduisons un système de particules auxiliaires $(\hat{\xi}_t^1, \dots, \hat{\xi}_t^q)$. Les particules $(\hat{\xi}_0^1, \dots, \hat{\xi}_0^q)$ sont i.i.d. de loi η_0 . Les particules diffusent indépendamment les unes des autres suivant L_t et subissent des sauts. Pour chaque $(i, j) \in [q]^2$, $i \neq j$, il existe une horloge exponentielle de paramètre V_∞/N et un processus ponctuel de Poisson $T_{(i,j)}^1, T_{(i,j)}^2, \dots$ associé de paramètre V_∞/N . Les $T_{(i,j)}^1, T_{(i,j)}^2, \dots$ sont appelés les temps de sonnerie. À un temps de sonnerie $t \in \{T_{(i,j)}^1, T_{(i,j)}^2, \dots\}$,

$$\hat{\xi}_t^i \begin{cases} \leftarrow \hat{\xi}_{t-}^j & \text{avec proba. } \frac{V_t(\hat{\xi}_t^i)}{V_\infty} \\ \leftarrow \hat{\xi}_{t-}^i & \text{avec proba. } 1 - \frac{V_t(\hat{\xi}_t^i)}{V_\infty}. \end{cases}$$

La notation $\hat{\xi}_t^i \leftarrow \hat{\xi}_t^j$ signifie que $\hat{\xi}_t^i$ saute sur $\hat{\xi}_t^j$ (ou est substitué par $\hat{\xi}_t^j$). Si $q = N$, ce système a même loi que Ξ_t .

Fixons $\forall F \in \mathcal{B}_b(E^q)$,

$$F^e((\hat{\xi}_s^1, \dots, \hat{\xi}_s^q)_{0 \leq s \leq t}) = F(\hat{\xi}_t^1, \dots, \hat{\xi}_t^q) \exp\left(-\int_0^t [V_s(\hat{\xi}_s^1) + \dots + V_s(\hat{\xi}_s^q)] ds\right)$$

et

$$E_{t,k}(F) = \mathbb{E}\left(F^e((\hat{\xi}_s^1, \dots, \hat{\xi}_s^q)_{0 \leq s \leq t}) \mid k \text{ sonneries sur } [0; t]\right).$$

Remarquons que les $E_{t,k}(F)$ ne dépendent pas de N . Posons $\forall q \in \mathbb{N}^*$, $t \geq 0$, $F \in \mathcal{B}_b(E^q)$,

$$\mathbb{P}_{t,q}^N(F) = \mathbb{E}((\eta_t^N)^{\odot q}(F)), \quad \mathbb{Q}_{t,q}^N(F) = \mathbb{E}((\gamma_t^N)^{\odot q}(F)).$$

(Dans la suite, le temps t est fixé.)

Théorème 2.3.9 (Del Moral, Patras, Rubenthaler [12]). *Les suites $(\mathbb{P}_{t,q}^N)_{N \geq 1}$ et $(\mathbb{Q}_{t,q}^N)_{N \geq 1}$ sont différentiables à tout ordre (les coefficients sont explicites).*

Proposition 2.3.10 (Del Moral, Patras, Rubenthaler [12]). *Si $F \in \mathcal{B}_0^{\text{sym}}(E^q)$, pour $r < q/2$, $\partial^r \mathbb{Q}_{t,q}^N(F) = \partial^r \mathbb{P}_{t,q}^N(F) = 0$, et si, de plus, q est pair*

$$\partial^{q/2} \mathbb{Q}_{t,q}(F) = \frac{(q(q-1)V_\infty t)^{q/2}}{(q/2)!} E_{t,q/2}(F), \quad (2.3.9)$$

$$\partial^{q/2} \mathbb{P}_{t,q}(F) = \frac{\partial^{q/2} \mathbb{Q}_{t,q}(F)}{\gamma_t(1)^q}. \quad (2.3.10)$$

On sait que $\forall f$ bornée,

$$\eta_t^N(f) \xrightarrow[N \rightarrow +\infty]{\text{P.S.}} \eta_t(f). \quad (2.3.11)$$

La formule (2.3.10) ci-dessus se simplifie si $F = (f_1 \otimes \dots \otimes f_q)_{sym}$ avec $f_1, \dots, f_q \in \mathcal{B}_0^{sym}(E)$ (l'indexation sym dans $(\dots)_{sym}$ signifie $\forall \phi \in \mathcal{B}_b(E^q)$, $\phi_{sym}(x_1, \dots, x_q) = \frac{1}{q!} \sum_{\sigma \in \mathcal{S}_q} \phi(x_{\sigma(1)}, \dots, x_{\sigma(q)})$). Nous avons alors

$$N^{q/2} \mathbb{E}((\eta_t^N)^{\odot q}(F)) \xrightarrow[N \rightarrow +\infty]{} \sum_{I_q \in \mathcal{I}_q} \prod_{\{i,j\} \in I_q} W_t(f_i \otimes f_j), \quad (2.3.12)$$

où \mathcal{I}_q est l'ensemble des partitions de $[q]$ en paires. La mesure W_t est explicite mais nous ne l'explicitons pas ici, sa définition demande l'introduction d'un autre système auxiliaire de particules. On peut alors démontrer un théorème limite vectoriel. Ce théorème est basé uniquement sur (2.3.11) et (2.3.12). On peut le réécrire pour les systèmes exposés dans la partie 2.3.2 (à condition de pousser les résultats de la proposition 2.3.3 jusqu'à une formule de type (2.3.12)) et dans la partie 2.3.5.

Théorème 2.3.11 (Del Moral, Patras, Rubenthaler [12]). *Pour toutes $f_1, \dots, f_q \in \mathcal{B}_0(E)$,*

$$\sqrt{N}(\eta_t^N(f_1), \dots, \eta_t^N(f_q)) \xrightarrow[N \rightarrow +\infty]{\text{loi}} \mathcal{N}(0, K),$$

avec $K(i, j) = \eta_t(f_i, f_j) + \mathbb{E}(W_t(f_i \otimes f_j))$.

La démonstration de ce théorème est élémentaire. Elle repose sur le calcul d'une transformée de Fourier.

Dans [12], nous nous intéressons aussi au comportement asymptotique des U -statistiques du système $(\xi_t^i)_{1 \leq i \leq N}$ (à la suite de [dlPG99, DM83, Lee90, RW97]). Une U -statistique (d'ordre q , $q \in \mathbb{N}^*$) est une statistique des particules de la forme $(\eta_t^N)^{\odot q}(F)$ avec $F \in \mathcal{B}_b(E^q)$. Nous obtenons la proposition suivante en utilisant la décomposition de Hoeffding d'une fonction de q plusieurs variables (voir [dlPG99, Lee90]).

Proposition 2.3.12 (Del Moral, Patras, Rubenthaler [12]). *Pour toute fonction $F \in \mathcal{B}_b(E^q)$,*

$$\sqrt{N}((\eta_t^N)^{\odot q}(F) - (\eta_t)^{\otimes q}) \xrightarrow[N \rightarrow +\infty]{} \mathcal{N}(0, q^2 \eta_t((F^{(1)})^2)),$$

où $F^{(1)}(x_1) = \int_{E^{q-1}} F_{sym}(x_1, x_2, \dots, x_q) (\eta_t)^{\otimes q-1}(dx_2, \dots, dx_q)$.

Si on prend $F \in \mathcal{B}_0^{sym}(E^q)$ de la forme $F = (f_1 \otimes \dots \otimes f_q)_{sym}$ avec les $f_i \in \mathcal{B}_0(E)$ alors nous avons le théorème suivant.

Théorème 2.3.13 ([Del Moral, Patras, Rubenthaler [12])). *Nous avons la convergence en loi (quand $N \rightarrow +\infty$)*

$$N^{\frac{q}{2}} (\eta_t^N)^{\odot q}(F) \rightarrow \sum_{k=0}^{\lfloor \frac{q}{2} \rfloor} (-1)^k \sum_{1 \leq i_1 < \dots < i_{q-2k} \leq q} I(f_{i_1}) \dots I(f_{i_{q-2k}}) \sum_J \eta_t(f_{J_1}) \dots \eta_t(f_{J_k}).$$

Ici, J parcourt les partitions de $[q] - \{i_1, \dots, i_{q-2k}\}$ en paires ordonnées $J_1 = \{j_1, j'_1\}, \dots, J_k = \{j_k, j'_k\}$ (où $j_i < j'_i$ et $j_1 < \dots < j_k$). De plus, $\lceil x \rceil$ est la partie entière supérieure du nombre réel x (le plus petit entier supérieur ou égal à x). Les $I(f_i), i = 1, \dots, q$ sont des intégrales de Wiener des fonctions f_i . Elles forment une famille gaussienne avec $\mathbb{E}[I(f_i)] = 0$ et $\mathbb{E}[I(f_i)I(f_j)] = \eta_t(f_i f_j) + \mathbb{E}(W_t(f_i \otimes f_j))$.

Dans le cas particulier $F = f^{\otimes q}$ avec $\eta_t(f^2) = 1$, nous obtenons la convergence en loi :

$$N^{\frac{q}{2}}(\eta_t^N)^{\odot q}(F) \rightarrow H_k(I(f))$$

où H_k est le k -ème polynôme de Hermite.

2.3.5 Étude de la propagation du chaos pour les systèmes de Bird et Nanbu

On s'intéresse dans cette partie à des systèmes de particules en interaction dont la mesure empirique approche la solution d'une équation intégral-différentielle appelée « équation de Boltzmann mollifiée ». Ce système est étudié dans [GM94, GM97, GM99, Mé198]. Nous ne rentrerons pas dans les détails physiques de la modélisation (voir [GM99] pour ces détails), ils ne sont pas nécessaires aux résultats mathématiques.

2.3.5.1 Présentation du système

On a N particules $Z^{N,1}, \dots, Z^{N,N}$ dans $E = \mathbb{R}^d$, qui évoluent dans le temps. On se donne un noyau de Markov μ de $E \times E$ dans $E \times E$ tel que $\mu(x, y, \dots) = \mu(y, x, \dots), \forall x, y \in E$ et tel que $\sup_{x, y \in E} \mu(x, y, E, E) \leq \Lambda < +\infty$.

- À $t = 0$, ces particules sont tirées de manière i.i.d. suivant une loi P_0 .
- Ces particules bougent dans E de deux manières différentes.
 - Elles évoluent indépendamment les unes des autres suivant un générateur de Markov L (homogène).
 - Chaque ensemble de particules de cardinal 2 ($\{i, j\}$ avec $i \neq j$) possède une horloge exponentielle de paramètre $\Lambda/(N-1)$. Quand cette horloge sonne, on dit qu'il y a interaction. S'il y a interaction pour les particules no. i, j au temps t , on tire $U \sim \mathcal{U}([0; 1])$ (indépendante de toutes les autres variables). Si $U \leq \frac{\mu(Z_{t-}^{N,i}, Z_{t-}^{N,j}, E, E)}{\Lambda}$, on dit qu'il y a un saut. On tire alors $(H, K) \sim \frac{\mu(Z_{t-}^{N,i}, Z_{t-}^{N,j}, \dots)}{\mu(Z_{t-}^{N,i}, Z_{t-}^{N,j}, E, E)}$ et on pose

$$\begin{cases} Z_t^{N,i} = Z_{t-}^{N,i} + H \\ Z_t^{N,j} = Z_{t-}^{N,j} + K. \end{cases}$$

2.3.5.2 Étude de la convergence

L'article [GM97] nous dit que $\forall t$

$$P_t^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{Z_t^{N,i}} \xrightarrow{N \rightarrow +\infty} P_t, \quad (2.3.13)$$

Entre 0 et t_1 , on ajoute toutes les particules qui ont interagité avec la particule 3 (par exemple la particule 5 en t_2). On recommence la même construction pour la particule 5, ... Sur notre figure : 1 a interagité avec 3 sur $[0; t]$, 3 a interagité avec 5 sur $[0; t_1]$, 5 a interagité avec 2 sur $[0; t_2]$, 2 a interagité avec 4 sur $[0; t]$ (et réciproquement). Pour une distance s à t , on compte le nombre de particules K_s dans le graphe au-dessus de 1, 2 et L_s le nombre d'interactions entre particules de cette population sur $[t - s; t]$. Sur notre figure : $K_{s_1} = 4$, $L_{s_1} = 0$, $K_{s_2} = 5$, $L_{s_2} = 1$,

La démonstration du théorème ci-dessus est basée sur la décomposition

$$\mathbb{E}(F(Z_t^{N,1}, \dots, Z_t^{N,q})) = \sum_{l \geq 0} \mathbb{E}(F(Z_t^{N,1}, \dots, Z_t^{N,q}) | L_t = l) \mathbb{P}(L_t = l).$$

On montre que $\mathbb{P}(L_t = l)$ est d'ordre l par des techniques de couplage et en utilisant des résultats classiques de croissance de population ([AN72]).

J'en déduis une proposition similaire à la proposition 2.3.3.

Proposition 2.3.15 (Rubenthaler, [13]). *Pour tout $q \in \mathbb{N}^*$, $\forall t \geq 0$, $\forall F \in \mathcal{B}_0^{sym}(E^q)$,*

$$N^{q/2} \mathbb{E}(F(Z_t^{N,1}, \dots, Z_t^{N,q})) \xrightarrow{N \rightarrow +\infty} \text{limite explicite}.$$

On a alors l'équivalent du théorème 2.3.13.

Théorème 2.3.16 (Rubenthaler [13]). *Pour toutes $f_1, \dots, f_q \in \mathcal{B}_0(E)$,*

$$\sqrt{N} (P_t^N(f_1), \dots, P_t^N(f_q)) \xrightarrow{N \rightarrow +\infty} \text{loi } \mathcal{N}(0, K),$$

avec K explicite.

Ce théorème s'étend de manière immédiate à \tilde{P}_t .

2.3.6 Perspectives

Ces méthodes devraient permettre d'étudier le système suivant. On regarde un processus de fragmentation de l'intervalle $[0; 1]$ (tel que décrit dans [Ber06]). On arrête la fragmentation d'un morceau dès que sa taille est plus petite qu'un $\epsilon > 0$. Les fragments ainsi arrêtés ont une loi asymptotique (leur taille, divisée par ϵ converge en loi quand $\epsilon \rightarrow 0$). À des fins d'estimation des paramètres de la fragmentation, on aimerait avoir un TCL sur la mesure empirique des fragments arrêtés (voir [HK08]).

2.4 Application du filtrage à un problème d'optimisation de portefeuille

On s'intéresse au prix d'un actif risqué S_t satisfaisant l'E.D.S suivante

$$\frac{dS_t}{S_t} = \left(\mu(t) + \frac{\sigma^2}{2} \right) dt + \sigma dB_t,$$

où (B_t) est un mouvement brownien standard et $(\mu(t))$ est un processus à valeurs dans un ensemble $\{\mu_1, \mu_2\}$ de générateur infinitésimal matriciel

$$G = \begin{bmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{bmatrix},$$

$\lambda_1, \lambda_2 > 0$ (ce qui signifie que l'on reste un temps exponentiel de paramètre λ_i dans l'état μ_i avant de sauter dans l'autre état). On dispose également d'un actif non risqué (S_t^0) évoluant suivant

$$\frac{dS_t^0}{S_t^0} = r dt.$$

Le processus $X_t = \log(S_t)$ vérifie

$$dX_t = \mu(t)dt + \sigma dB_t.$$

Le processus (S_t) modélise un actif dont la dérive est soumise à des ruptures. On investit une proportion π_t de notre richesse dans S_t et le reste dans S_t^0 (cette proportion varie avec le temps), Posons $\mathcal{F}_t^X = \sigma((X_s)_{0 \leq s \leq t})$, $F_t = \mathbb{P}(\mu(t) = \mu_1 | \mathcal{F}_t^X)$, $\mu^{opt}(t) = \mu_1 F_t + \mu_2(1 - F_t)$. On suppose que l'on démarre avec une richesse $W_0^{x,\pi} = x$. Notons $W_T^{x,\pi}$ la richesse à un temps $T > 0$ (obtenue avec la stratégie π). Nous voulons maximiser $\mathbb{E}(\log(W_T^{x,\pi}))$. La stratégie d'allocation optimale est la suivante (sous la contrainte $\pi \in [0; 1]$) :

$$\pi_t^* = \text{proj}_{[0;1]} \left\{ \frac{\mu^{opt}(t) - r}{\sigma^2} \right\}.$$

On sait que (F_t) vérifie une équation de Kushner-Stratonovich (voir [KO88], p. 90) :

$$F_t = p_0 + \int_0^t (-\lambda_1 F_s + \lambda_2(1 - F_s)) ds + \int_0^t \frac{\mu_1 - \mu_2}{\sigma} F_s(1 - F_s) d\bar{B}_s,$$

où

$$\bar{B}_t = \frac{1}{\sigma} \left(X_t - \int_0^t \mu^{opt}(s) ds \right)$$

est un mouvement brownien standard et $p_0 = \mathbb{P}(\mu(0) = \mu_1)$. Supposons que l'on dispose des prix (S_t) sur une grille temporelle arbitrairement petite, on pourrait approcher (F_t) par un schéma d'Euler. Dans [5], nous proposons une méthode, inspiré du filtrage, qui converge plus vite que le schéma d'Euler.

On suppose que l'on observe les prix (S_t) sur un pas de temps δ . Soit $(\tilde{\mu}_k)_{k \geq 0}$ la chaîne de Markov à valeurs dans $\{\mu_1, \mu_2\}$ de matrice transition

$$\begin{bmatrix} e^{-\lambda_1 \delta} & 1 - e^{-\lambda_1 \delta} \\ 1 - e^{-\lambda_2 \delta} & e^{-\lambda_2 \delta} \end{bmatrix}.$$

Soit $(\tilde{X}_k)_{k \geq 0}$ la suite aléatoire définie par

$$\begin{aligned} \tilde{X}_0 &= \log(S_0) \\ \tilde{X}_{k+1} &= \tilde{X}_k + \tilde{\mu}_k \delta + \sigma \sqrt{\delta} U_k, \end{aligned}$$

avec des (U_k) i.i.d. $\sim \mathcal{N}(0, 1)$. Nous définissons (à X_0, \dots, X_k fixés)

$$\bar{F}_k = \mathbb{P}(\tilde{\mu}_k = \mu_1 | \tilde{X}_0 = X_0, \dots, \tilde{X}_k = X_k).$$

Cette quantité se calcule récursivement de manière aisée.

Puisque $(\tilde{\mu}_k, \tilde{X}_k)_{k \geq 0}$ évolue presque comme $(\mu(k\delta), X_{k\delta})$, \bar{F}_k doit approcher $F_{k\delta}$. Ce résultat est l'objet du théorème suivant.

Théorème 2.4.1 (Martinez, Rubenthaler, Tanré [5]). *Pour tout $N \in \mathbb{N}$, $\exists C_{N\delta}$ (dépendant continument de $N\delta$ et des paramètres du problème) tel que :*

$$\mathbb{E} \left(\sup_{0 \leq k \leq N} (F_{k\delta} - \bar{F}_k)^2 \right) \leq C_{N\delta} \delta^2. \quad (2.4.1)$$

Il existe une extension $(\bar{\bar{F}}_t)_{t \geq 0}$ de $(\bar{F}_k)_{k \geq 0}$ en temps continu (telle que $\forall k, \bar{\bar{F}}_{k\delta} = \bar{F}_k$) telle que $\forall t_0$:

$$\mathbb{E} \left(\sup_{0 \leq t \leq t_0} (\bar{\bar{F}}_t - F_t)^2 \right) \leq C_{t_0} \delta^2,$$

où C_{t_0} dépend seulement de t_0 et des paramètres du problème.

Remarquons que la vitesse d'un schéma d'Euler pour la quantité dans (2.4.1) est en δ . Si les paramètres du problème sont mal spécifiés, c'est à dire si on fait les calculs avec $\bar{\lambda}_i, \bar{\mu}_i$ ($i = 1, 2$) à la place de λ_i, μ_i , on obtient un filtre « mal spécifié » $(\hat{F}_k)_{k \geq 0}$ à la place de $(\bar{F}_k)_{k \geq 0}$. En utilisant des techniques classiques de stabilité du filtre, nous montrons le théorème suivant.

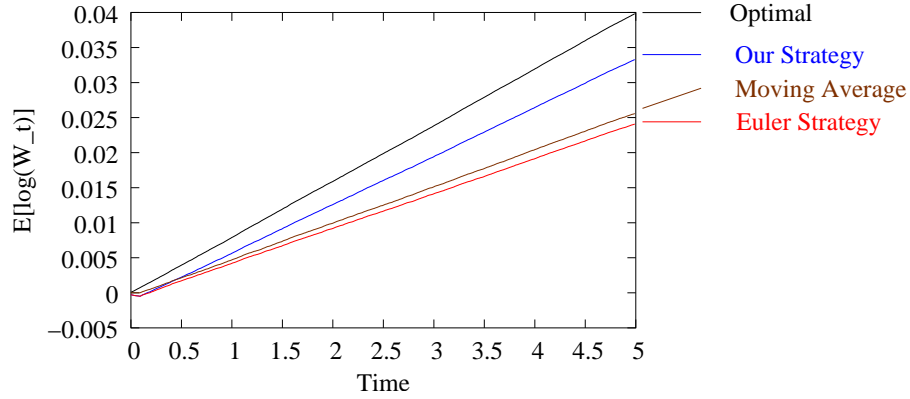
Théorème 2.4.2. *Nous avons les bornes uniformes*

$$\sup_{t \geq 0} \mathbb{E}(|F_{\lfloor t/\delta \rfloor \delta} - \hat{F}_{\lfloor t/\delta \rfloor}|) \leq C(\delta + \sup_{i=1,2} (|\lambda_i - \bar{\lambda}_i| + |\mu_i - \bar{\mu}_i|),$$

où C dépend continument des paramètres.

Nous injectons ensuite ce calcul de F_t dans la stratégie et comparons la richesse obtenue avec la richesse obtenue par d'autres méthodes (voir graphe ci-dessous où nous traçons $\mathbb{E}(\log(W_t))$ en fonction de t , W_t est la richesse au temps t).

- « Optimal » est la courbe obtenue avec la stratégie optimale.
- « Our strategy » est la courbe obtenue avec notre algorithme.
- « Euler strategy » est la courbe obtenue en calculant F_t par un schéma d'Euler.
- « Moving average » est une stratégie où l'on estime la dérive $\mu(t)$ par un $\bar{\mu}$ en faisant une moyenne sur une fenêtre mobile. On investit ensuite une proportion π_t qui correspond à ce qu'il faudrait investir si $\mu(t)$ était constante égale à μ .



2.5 Cas d'une formule de Feynman-Kac dans laquelle le potentiel dépend de la trajectoire

2.5.1 Présentation du travail [11]

On s'intéresse à un modèle avec 2 processus : (pour un processus $(U_n)_{n \geq 0}$ et $q \leq p$, on note $U_{q:p} = U_q, \dots, U_p$)

- $(\Lambda_n)_{n \geq 0}$ une chaîne de Markov dans un espace E , de transition M , de loi initiale π_0 .
- $(Y_n)_{n \geq 1}$ tel que Y_n admet la probabilité conditionnelle $\Psi_n(y_n | y_{1:n-1}, \lambda_{0:n})$ sachant $\Lambda_{0:n} = \lambda_{0:n}$, $Y_{1:n-1} = y_{1:n-1}$.

On veut calculer $\forall n$, la loi $\pi_n = \mathcal{L}(\Lambda_n | Y_{1:n-1} = y_{1:n-1})$. On suppose que la suite $(Y_n)_{n \geq 0}$ est fixée égale à $(y_n)_{n \geq 0}$. On pose $\Psi_n(\lambda_{0:n}) = \Psi_n(y_n | y_{1:n-1}, \lambda_{0:n})$. Notons $\forall n$, $\eta_n = \mathcal{L}(\Lambda_{0:n} | Y_{1:n} = y_{1:n})$ (c'est une loi sur E^{n+1}). Introduisons, $\forall q \in \mathbb{N}^*$, l'opérateur

$$\begin{aligned} \bar{M}_{q+1} : \mathcal{M}(E^q) &\rightarrow \mathcal{M}(E^{q+1}) \\ \mu &\mapsto \bar{M}_{q+1}(dx_1, \dots, dx_{q+1}) \\ &= \mu(dx_1, \dots, dx_q) M(x_q, dx_{q+1}), \end{aligned}$$

où, pour tout espace F , $\mathcal{M}(F)$ représente les mesures sur F . Nous avons la relation de récurrence

$$\eta_n = \Psi_n \cdot (\eta_{n-1} \bar{M}_n).$$

C'est pourquoi on parle encore de mesure de Feynman-Kac. La mesure π_n est une marginale de la mesure η_n ($\forall n$). Nous voulons montrer que la mesure π_n peut-être approchée uniformément en temps par n système de particules. Le modèle que nous venons de décrire englobe les modèles décrits dans [DGA00, CL00].

Pour nous ramener à des mesures de Feynman-Kac plus classiques (où le potentiel ne dépend pas du chemin), nous tronquons les potentiels. Plus précisément, nous introduisons, pour un $p \in \mathbb{N}^*$, la suite de mesures $(\eta_n^p)_{n \geq 0}$ définie par récurrence par :

- $\eta_0^p = \eta_0$, $\eta_1^p = \eta_1$, ..., $\eta_{p-1}^p = \eta_{p-1}$
- Pour $k \geq p$, η_k^p est une mesure sur E^p . Définissons M_p l'opérateur de Markov de E^p dans E^p par

$$\mu M_p(dx_1, \dots, dx_p) = \int_{x_0 \in E} \mu(dx_0, dx_1, \dots, dx_{p-1}) M_p(dx_{p-1}, dx_p).$$

Posons $\forall n$,

$$\Psi_n^p(x_1, \dots, x_p) \begin{cases} = \Psi_n(0, \dots, 0, x_1, \dots, x_p) & \text{si } n-1 \geq p \\ = \Psi_n(x_{p-(n-1)}, \dots, x_p) & \text{si } n-1 < p. \end{cases}$$

Les $0, \dots, 0$ signifient que nous mettons là un élément arbitraire de E .
Alors

$$\eta_k^p = \Psi_k^p(\eta_{k-1}^p M_p).$$

On reconnaît une équation de même type que (2.1.3). On peut donc approcher les mesures (η_k^p) par un filtre particulière $\eta_k^{p,N}$ (à N particules, construit sur le modèle de la partie 2.2.2). Ce filtre particulière approche en particulier la mesure η_k . Pour que ce filtre particulière soit stable en temps et approche $(H^p \eta_k)$ (quand $n, p \rightarrow +\infty$) ($H^p \eta_k$ est la marginale de η_k en les p dernières coordonnées), il faut :

- Un noyau M_p fortement mélangeant (comme (2.2.3)). Même si on suppose que M est fortement mélangeant, ce n'est pas le cas. Mais $(M_p)^p$ (itéré p fois) est fortement mélangeant. Nous détaillons les hypothèse nécessaires plus bas. On dit que M_p est d'autant moins mélangeant que p est grand (il faut itérer de plus en plus pour obtenir quelque chose de mélangeant). On subodore que les résultats de stabilité du filtre particulière vont être meilleures pour p petit.
- Des potentiels Ψ_n^p « proches » des potentiels Ψ_n . Ceci est d'autant plus vrai que p est grand.

Ces remarques amènent à des exigences contradictoires sur p . Nous faisons les hypothèse suivantes. (C'est une simplification par rapport à [11], où la chaîne (Λ_n) n'est pas homogène.)

Hypothèse 1. *Le noyau M est fortement mélangeant, i.e. $\exists \varepsilon \in (0, 1)$ such that*

$$\varepsilon \xi(A) \leq M(\lambda_{n-1}, A) \leq \frac{1}{\varepsilon} \xi(A)$$

for some $\xi \in \mathcal{M}_+(E)$, and for any Borel set $A \subset E$, any $\lambda_{n-1} \in E$.

Hypothèse 2. *Pour p assez grand,*

$$|\Psi_n(\lambda_{0:n}) - \Psi_n^p(\lambda_{n-p+1:n})| \leq \phi_n \tau^p \{ \Psi_n(\lambda_{0:n}) \wedge \Psi_n^p(\lambda_{n-p+1:n}) \}$$

fpour des constantes ϕ_n et τ , $\phi_n > 0$, $0 < \tau < 1$, et tout $\lambda_{0:n} \in E^{n+1}$. Par abus de notation, $\tilde{\Psi}_n^p = \Psi_n^p$ pour $p > n$.

Hypothèse 3. *Il existe des constantes a_n, b_n , $n \geq 0$, $a_n \geq 1$, $b_n \geq 1$, telles que*

$$\frac{1}{a_n} \leq \Psi_n(\lambda_{0:n}) \leq b_n, \quad \frac{1}{a_n} \leq \tilde{\Psi}_n^p(\lambda_{(n-p+1):n}) \leq b_n$$

pour tout $\lambda_{0:n} \in E^{n+1}$.

Sous les hypothèses (1,3), le noyau $(M_p)^p$ est mélangeant. Posons $\forall k \geq 0$, μ, μ' probabilités sur $E^{(k+1) \wedge p}$

$$\tilde{\varepsilon}_{k,p}^2 = \frac{\varepsilon_k^2}{(a_k \dots a_{k+p-2})(b_k \dots b_{k+p-2})}, \quad \tilde{\rho}_{k,p} = \frac{1 - \tilde{\varepsilon}_{k,p}^2}{1 + \tilde{\varepsilon}_{k,p}^2}.$$

Soit η_n^N un filtre particulaire approchant η_n (à N particules, construit sur le modèle de la partie 2.2.2) Posons (pour une fonction test f fixée)

$$\mathcal{E}_{n,N}^p(y_{1:n}) = \sup_{f: \|f\|_\infty=1} \mathbb{E}_N (|\langle H^p \eta_n - H^p \eta_n^N, f \rangle| \mid Y_{0:n} = y_{0:n}) .$$

C'est une estimation de l'erreur du filtre particulaire.

Théorème 2.5.1 (Chopin, Del Moral, Rubenthaler [11]). *Pour tout fonction test f ,*

$$\mathcal{E}_{n,N}^p(y_{1:n}) \leq \frac{12}{\log 3} \sum_{i=1}^n \frac{\delta_i}{\tilde{\varepsilon}_{i+1,p}^2 \tilde{\varepsilon}_{i+p+1,p}^2} \prod_{j=2}^{\lfloor (n-i)/p \rfloor - 1} \tilde{\rho}_{i+jp+1} \quad (2.5.1)$$

où

$$\delta_i = \tau^p \phi_i + \frac{2a_i b_i}{\sqrt{N}}. \quad (2.5.2)$$

Nous comparons $H^p \eta_n$ et $H^p \eta_n^N$ en passant par les différences suivantes : $H^p \eta_n - \eta_n^p$, $\eta_n^p - \eta_n^{p,N}$, $\eta_n^{p,N} - H^p \eta_n^N$. Le filtre particulaire η_n^N est celui utilisé dans [DGA00, CL00]. L'équation (2.5.2) montre quel choix de $p, N \rightarrow +\infty$ on doit faire pour avoir $\mathcal{E}_{n,N}^p \rightarrow 0$ quand $p, N \rightarrow +\infty$. Nous donnons des exemples de tels calculs dans [11].

2.5.2 Perspectives

Je travaille actuellement avec Dan Crisan (Imperial College, London) sur le filtrage en temps continu (voir partie 2.1.2). On regarde les observations à des instants discrets (par exemple entiers) Y_0, Y_1, \dots . Le potentiel dépend de la trajectoire (voir formule (2.1.4)). Fixons $\forall k, \bar{X}_k = (X_k, X_{k+1})$. Les lois $\mathcal{L}(\bar{X}_k | (Y_s)_{0 \leq s \leq k+1})$ forment une suite de Feynman-Kac basée sur le noyau de transition de (\bar{X}_k) et les potentiels

$$\psi_{k+1}(x_1, x_1) = \mathbb{E} \left(\exp \left(\int_k^{k+1} h(X_s) dY_s - \frac{1}{2} \int_k^{k+1} h(X_s)^2 ds \right) \right) .$$

Notons \bar{M} le noyau de (\bar{X}_k) , si le noyau de (X_k) est fortement mélangeant (ce qui va arriver si on tronque correctement l'espace d'état) alors \bar{M}^2 est mélangeant. On voit que l'on retrouve une problématique similaire à [11].

Chapter 3

Processus stochastiques et calcul numérique

3.1 Approximation numérique d'une équation différentielle stochastique conduite par un processus de Lévy

3.1.1 Cas général, troncation des petits sauts

On s'intéresse à l'équation suivante

$$X_t = X_0 + \int_0^t f(X_{s-}) dZ_s, \quad (3.1.1)$$

où Z est un processus de Lévy dans \mathbb{R} et $f : \mathbb{R} \rightarrow \mathbb{R}$ est C^2 avec f, f', f'' bornées. Le processus Z a les caractéristiques (b, c, F) , c'est à dire qu'il a la fonction caractéristique suivante ($a > 0$)

$$\mathbb{E}(e^{iuZ_t}) = \exp \left[t \left(iub - \frac{cu^2}{2} + \int F(dx)(e^{iux} - 1 - iux\mathbf{1}_{|x| \leq a}) \right) \right].$$

Pour simuler des trajectoires de (X_t) , on a recours à un schéma d'Euler. Pour cela, il faut savoir simuler les incréments de (Z_t) , ce qui n'est possible que dans des cas particuliers. On introduit une suite d'ensembles $D_n =]-\infty; -1/v_n] \cap [1/u_n; +\infty[$ ($u_n, v_n > 0, \xrightarrow[n \rightarrow +\infty]{} +\infty$). Posons

$$M_t^n := \sum_{s \leq t} \Delta Z_s \mathbf{1}_{D_n \cap [-a, a]}(\Delta Z_s) - t \int_{D_n \cap [-a, a]} xF(dx),$$

$$N_t = \sum_{s \leq t} \Delta Z_s \mathbf{1}_{|\Delta Z_s| > a},$$

$$Z_t^n := bt + \sqrt{c}W_t + M_t^n + N_t = b_n t + \sqrt{c}W_t + N_t^n$$

où

$$N_t^n := \sum_{s \leq t} \Delta Z_s \mathbf{1}_{D_n}(\Delta Z_s)$$

et

$$b_n := b - \int_{D_n \cap [-a, a]} x F(dx).$$

Le processus Z^n est le processus Z dans lequel on a enlevé les «petits» sauts, et que l'on a compensé correctement. On appelle T_n^1, T_n^2, \dots les temps de saut de Z^n . On définit une discrétisation (t_i^n) du temps par

$$t_0^n = 0, t_{i+1}^n = \inf\{t_i^n + 1/k_n, T_j^n : T_j^n > t_i^n\}.$$

Posons

$$\theta_n := F(D_n), \bar{\theta}_n := \sup(\theta_n, k_n).$$

Nous proposons le schéma d'approximation suivant.

1. On simule \tilde{X}_0^n de même loi que X_0 ,
2. On simule le prochain temps de saut de Z^n (les temps entre les sauts de Z^n sont de loi exponentielle de paramètre θ_n).
3. Si nous n'avons pas $b_n = c = 0$, entre les temps de saut de Z^n , T_i^n et T_{i+1}^n , \tilde{X}^n est défini récursivement par : $\tilde{X}_{t_{j+1}^n}^n = \tilde{X}_{t_j^n}^n + f(\tilde{X}_{t_j^n}^n)b_n(t_{j+1}^n - t_j^n) + f(\tilde{X}_{t_j^n}^n)\sqrt{c}(W_{t_{j+1}^n} - W_{t_j^n})$ si t_{j+1}^n n'est pas un temps de saut de Z^n (nous devons simuler les incréments $W_{t_{j+1}^n} - W_{t_j^n}$ si $c \neq 0$)
si $b_n = 0$ et $c = 0$ alors \tilde{X}^n est constant entre T_i^n et T_{i+1}^n ,
4. Si $t_{j+1}^n = T_i^n$, nous introduisons :

$$\tilde{X}_{T_i^n-}^n = \tilde{X}^n(t_j^n) + f(\tilde{X}^n(t_j^n))b_n(t_{j+1}^n - t_j^n) + f(\tilde{X}^n(t_j^n))\sqrt{c}(W_{t_{j+1}^n} - W_{t_j^n})$$

et posons

$$\tilde{X}_{T_i^n}^n = \tilde{X}_{T_i^n-}^n + f(\tilde{X}_{T_i^n-}^n)\Delta Z_{T_j^n}^n$$

puis on retourne au point (2.).

Théorème 3.1.1 (Rubenthaler [8]). *Si Z n'est pas un processus de Poisson composé et que :*

- i) $\exists w_n \xrightarrow{n \rightarrow +\infty} +\infty$ tel que $w_n(Z - Z^n) \xrightarrow{n \rightarrow +\infty} M$
- ii) la suite (k_n) est telle que $\rho_n := \inf\left(w_n, \frac{\bar{\theta}_n}{b_n^2}, \frac{\sqrt{\bar{\theta}_n}}{c}\right) \xrightarrow{n \rightarrow +\infty} +\infty$

alors :

$$\rho_n(X - \tilde{X}^n) \text{ est tendue}$$

Si nous avons de plus :

iii) $\frac{k_n}{\theta_n} \xrightarrow{n \rightarrow +\infty} A \in [0, +\infty]$, $\frac{\rho_n}{w_n} \xrightarrow{n \rightarrow +\infty} \alpha_1 \in [0, 1]$, $\frac{\rho_n}{(\theta_n/b_n^2)} \xrightarrow{n \rightarrow +\infty} \alpha_2 \in [0, 1]$,
 $\frac{\rho_n}{(\sqrt{\theta_n}/c)} \xrightarrow{n \rightarrow +\infty} \alpha_3 \in [0, 1]$
 alors :

$$\rho_n(X - \tilde{X}^n) \xRightarrow{n \rightarrow +\infty} \alpha_1 U + V \left(\alpha_2 c(A), \alpha_3 \sqrt{c(A)} \right), \quad (3.1.2)$$

où

$$c(A) = \begin{cases} \sup(1, A) \left(1 - \frac{e^{-1/A}}{A(1-e^{-1/A})} \right) & \text{if } A \in (0, +\infty) \\ 1 & \text{if } A = 0 \\ 1/2 & \text{if } A = +\infty, \end{cases} \quad (3.1.3)$$

et U est la solution de :

$$U_t := \int_0^t f'(X_{s-}) U_{s-} dZ_s + \int_0^t f(X_{s-}) dM_s \quad (3.1.4)$$

et pour tous $a_1, a_2 \in \mathbb{R}$, $V(a_1, a_2)$ est la solution de :

$$V_t(a_1, a_2) = \int_0^t f'(X_{s-}) V_{s-}(a_1, a_2) dZ_s + \int_0^t f'(X_{s-}) f(X_{s-}) d(a_1 s + a_2 B_s). \quad (3.1.5)$$

où B est un mouvement brownien et B, M, Z sont indépendants.

Ce théorème appelle les remarques suivantes.

- Les convergences en loi sont entendues pour la topologie de Skorohod dans $\mathbb{D}(\mathbb{R}^+, \mathbb{R})$.
- Nous sommes libres de choisir les suites (u_n) , (v_n) , (k_n) . Il reste donc une optimisation à faire pour avoir la suite (ρ_n) la plus « grande » possible (à nombre d'opérations fixées) (ce que je fais dans [8]).
- Sous la condition $u^2 \int_{|x| < 1/u} x^2 F(dx) \xrightarrow{u \rightarrow +\infty} +\infty$, $\exists (w_n)$ tel que le point (i) ci-dessus a lieu avec M un mouvement brownien.

Les démonstration sont basées sur des résultats de stabilité des E.D.S que l'on trouve dans [KP91a, KP91c, KP91b, MS91, JMP89]. On écrit

$$\tilde{X}_t = X_0 + \int_0^t f(\tilde{X}_{s-}^n) dZ_s^n.$$

Puis, les processus Z et Z^n étant définis sur le même espace,

$$X_t - \tilde{X}_t^n = \int_0^t f(X_{s-}) - f(\tilde{X}_{s-}^n) dZ_s + \int_0^t f(\tilde{X}_{s-}^n) d(Z_s - Z_s^n).$$

Supposons pour simplifier que la différence $f(X_{s-}) - f(\tilde{X}_{s-}^n)$ soit une fonction de $X_{s-} - \tilde{X}_{s-}^n$. Puisque l'on a suppose que $w_n(Z - Z^n)$ converge en loi pour une certaine suite w_n , les résultats de stabilité des E.D.S nous disent que $w_n(X - \tilde{X}^n)$ converge vers la solution de (3.1.4). On peut coupler $w_n(Z - Z^n)$ à sa limite pour améliorer la vitesse de convergence du schéma, c'est ce qui est fait dans [Fou].

3.1.2 Amélioration de la vitesse de convergence dans le cas où Z est un processus subordonné

On reprend l'équation (3.1.1) dans laquelle on remplace Z par $Y_t = \sigma W_t + bt + Z_t(V_t)$ où V est un subordonateur, W est un brownien standard, $\sigma, b \geq 0$. On suppose que Z est un processus stable (donc on sait simuler ses incréments). En utilisant les idées de [Bon82], nous écrivons

$$V_t = D_V t + \sum_{k \geq 1} g(T_k) \mathbf{1}_{U_k \leq t},$$

avec D_V la dérive de V , $(T_k)_{k \geq 1}$ un processus ponctuel de Poisson homogène, $(U_k)_{k \geq 0}$ une suite i.i.d. $\sim \mathcal{U}([0; 1])$, g l'inverse continu à droite de $F_V(\cdot; +\infty[)$ (F_V est la mesure de Lévy de V)

$$g(u) = \inf \{t > 0 : F_V(\cdot; +\infty[) < u\}.$$

Nous posons $\forall n$

$$V_t^n = D_V t + \sum_{k: T_k \leq a_n} g(T_k) \mathbf{1}_{U_k \leq t},$$

(a_n) est une suite de troncation à choisir,

$$\tilde{V}_t^n = \sum_{k: T_k > a_n} g(T_k) \mathbf{1}_{U_k \leq t}.$$

Nous proposons l'approximation suivante des incréments de Y sur un pas de temps de longueur $1/n$:

$$\begin{aligned} \delta \bar{Y}_k^n &= Z \left(V_{\frac{k-1}{n}} + V_{\frac{k}{n}}^n - V_{\frac{k-1}{n}}^n + \left(\mathbb{E} \tilde{V}_{\frac{1}{n}}^n \right) \wedge \left(\tilde{V}_{\frac{k}{n}}^n - \tilde{V}_{\frac{k-1}{n}}^n \right) \right) - Z \left(V_{\frac{k-1}{n}} \right) \\ &\quad + Z^k \left(\left[\mathbb{E} \tilde{V}_{\frac{1}{n}}^n - \left(\tilde{V}_{\frac{k}{n}}^n - \tilde{V}_{\frac{k-1}{n}}^n \right) \right]^+ \right) + \sigma \left(W_{\frac{k}{n}} - W_{\frac{k-1}{n}} \right) + \frac{1}{n} b \end{aligned} \quad (3.1.6)$$

Dans [7], nous obtenons un théorème (th. 2.4, Rubenthaler & Wiktorsson, 2003) du même genre que le théorème 3.1.1, avec une meilleure vitesse de convergence.

Le type de convergence obtenue (convergence en loi pour les trajectoires) est utile quand on veut estimer une fonctionnelle continue de la trajectoire par une méthode de Monte-Carlo.

3.2 Optimisation de portefeuille

On s'intéresse aux prix (en temps discret) de d actifs. Ces prix forment une suite aléatoire $(S_k)_{k \geq 0}$ (à valeurs dans \mathbb{R}^d). Une stratégie d'investissement $(\phi_k)_{k \geq 0}$ est une suite aléatoire prévisible (dans \mathcal{F}_S). Si on investit la richesse V_0 (sur les d actifs) et que l'on applique la stratégie ϕ , on obtient la richesse au temps k :

$$V_k = V_0 + \sum_{j=1}^k \phi_j^T \Delta_j$$

avec $\Delta_k = \beta_k S_k - \beta_{k-1} S_{k-1}$ (la suite prévisible β_k représente le taux d'intérêt). On cherche à construire un portefeuille qui couvre une option C (une variable

\mathcal{F}_n -mesurable, $n \in \mathbb{N}^*$). Il n'existe pas forcément de couverture exacte et nous cherchons à minimiser le risque quadratique $\mathbb{E}((G(V_0, \phi))^2)$, où $G = G(V_0, \phi) = \beta_n C - V_n$. Dans [10], nous calculons la stratégie optimale dans divers cas. Nous calculons aussi la stratégie optimale dans un modèle à temps continu (analogue à celui en temps discret que nous venons de décrire), dans le cas où les actifs sont l'exponentielle d'un processus de Lévy à changement de régime. Un processus de Lévy à changement de régime est un processus de Markov qui est Lévy par morceaux (avec des caractéristiques fixes) et qui change de caractéristiques à des instants fixés.

Dans le cas où les S_k ci-dessus sont les valeurs à des instants k/n d'un processus de Lévy à changements de régime, nous montrons dans [10], que la stratégie optimale en temps discret converge vers la stratégie optimale en temps continu quand $n \rightarrow +\infty$. Cette démonstration utilise les résultats de stabilité des E.D.S de [KP91a, KP91c, KP91b, MS91, JMP89].

Chapitre 4

Articles isolés

4.1 Recuit simulé dans \mathbb{R}^d , application à la recherche du maximum de vraisemblance

4.1.1 Présentation du problème

On suppose que l'on dispose d'une suite d'observations $(Y_k)_{0 \leq k \leq n}$ d'une chaîne de Markov $((S_k, Y_k))_{k \geq 0}$ (S_k, Y_k dans \mathbb{R}^d). La chaîne (S_k) a un noyau de transition Q et $\mathcal{L}(Y_k | S_{0:k}) = r(\cdot | S_k)$. Les données Q et r dépendent d'un paramètre caché θ ($\in \Theta \subset \mathbb{R}^d$). Nous voulons estimer ce paramètre en utilisant le maximum de vraisemblance. C'est à dire que nous cherchons le maximum de la fonction de θ

$$l_n(\theta) = \sum_{k=1}^n \log \int r_\theta(y_k | s) \pi_{k|k-1}^\theta(ds),$$

où $p_\theta(y_k | y_{1:k-1})$ est la distribution conditionnelle de Y_k sachant $Y_{1:k-1}$ et $\pi_{k|k-1}^\theta$ est le filtre prédictif $\mathbb{P}_\theta(S_k \in \cdot | y_{1:k-1})$ (tout ceci calculé avec le paramètre θ). La mesure $\pi_{k|k-1}^\theta$ est approchée par un filtre particulière $\pi_{k|k-1}^{\theta, N}$ (du genre de celui décrit dans la partie 2.2.2). On obtient alors une approximation

$$l_n^N(\theta) = \sum_{k=1}^n \log \int r_\theta(y_k | s) \pi_{k|k-1}^{\theta, N}(ds).$$

Nous cherchons à maximiser $l_n^N(\theta)$ dans le but de maximiser $l_n(\theta)$.

4.1.2 Algorithme de recuit simulé dit « recuit simulé rapide »

Supposons que l'on cherche à maximiser une fonction $\psi : \Theta \rightarrow \mathbb{R}$. On se donne un noyau de Markov K (de Θ dans Θ) et un schéma de température $(\beta_n)_{n \geq 0}$ (c'est à dire une suite tendant vers $+\infty$). On fixe $\theta_0 \in \Theta$, puis :

- Si on a construit θ_n , on tire $Z \sim K(\theta_n, \cdot)$.
- On prend $\theta_{n+1} = Z$ avec probabilité

$$f(\beta_{n+1}(\psi(\theta_n) - \psi(Z))_+),$$

et sinon $\theta_{n+1} = \theta_n$.

Remarquons que si $f = \exp$, ce schéma est le recuit simulé. La fonction d'acceptation f doit vérifier l'hypothèse suivante.

Hypothèse 4. 1. $f(0) = 0$ et f est décroissante, convexe

2. $f(t) \xrightarrow{t \rightarrow +\infty} 0$

3. $\frac{1}{t} \int_0^t f(z) dz \xrightarrow{t \rightarrow +\infty} 0$

4. Pour toute suite décroissante $(a_n)_{n \geq 1}$ de $]0; 1]$, \exists suite croissante $(\beta_n)_{n \geq 1}$ telle que $\forall c > 0$,

$$\sup_{n \geq 1} \frac{1}{a_n \beta_n} \int_0^{c\beta_n} f(t) dt < \infty.$$

Nous faisons l'hypothèse suivante sur la fonction ψ .

Hypothèse 5. 1. La fonction ψ a un seul maximum global θ_{\max} situé dans l'intérieur de Θ . Nous notons $\psi_{\max} = \psi(\theta_{\max})$.

2. La fonction ψ est localement C^3 autour de θ_{\max} et la forme quadratique $\psi''(\theta_{\max})$ est définie négative.

Nous faisons l'hypothèse suivante sur le noyau K . Posons $\forall \epsilon > 0$, $U^\epsilon = \{x \in \Theta : \psi(x) > \psi_{\max} - \epsilon\}$.

Hypothèse 6. 1. Pour tout $\epsilon > 0$, $\bar{K} = \sup_{x \in (U^\epsilon)^c} K(x, (U^\epsilon)^c) < 1$.

2. Nous avons $\forall x, y$, $K(x, dy) \leq \lambda(y) dy$ avec une fonction bornée λ .

Fixons une suite (a_n) de $]0; 1]$ convergeant vers 0 quand $n \rightarrow +\infty$. On prend une suite (β_n) telle que $\forall c > 0$, il existe $b_c > 0$ telle que

$$\frac{1}{\beta_n} \int_0^{c\beta_n} f(t) dt \leq b_c a_n.$$

Théorème 4.1.1 (Rubenthaler, Rydén, Wiktorsson [3]). Pour tout $n \geq 1$ et pour tout $0 < \delta < 1$,

$$\mathbb{P}(\psi(\theta_n) > \psi_{\max} - \epsilon) \geq 1 - \frac{\bar{K}}{1 - \bar{K}} \left\{ (1 + C'' a_1) \bar{K}^{n(1-\delta)} + C'' a_{\lfloor n\delta \rfloor + 1} \right\},$$

où $C'' = \max(C'/\bar{K}, 1)$ avec

$$C' = \left(C_{\epsilon, \epsilon''} + \sup_{x \in U^\epsilon} \frac{K(x, (U^{\epsilon''})^c)}{(\epsilon'' - \epsilon)} \right) b_{\epsilon'' - \epsilon},$$

$$C_{\epsilon, \epsilon''} = \|\lambda\|_\infty \times \sup_{\|z\| < \epsilon''} |\det(\text{Jac}_\xi(z))| \times \sup_{v \in [\epsilon, \epsilon'']} \frac{S_{\sqrt{v}, d}}{\sqrt{v}}$$

et $S_{t,d}$ est la surface de la sphère de rayon t dans \mathbb{R}^d et $\text{Jac}_\xi(z)$ est la matrice jacobienne de ξ en z . Si nous supposons de plus que $1 - a_k \leq a_k/a_{k-1}$ pour tout k , alors

$$\mathbb{P}(\psi(\theta_n) > \psi_{\max} - \epsilon) \geq 1 - \frac{\bar{K}}{1 - \bar{K}} a_n (1 + C'').$$

La démonstration de ce théorème repose sur des estimations fines de queues de mesures de Gibbs.

Corollaire 4.1.2 (Rubenthaler, Rydén, Wiktorsson [3]). *Si nous prenons $f(t) = 1/(1+t)$ et $\beta_n = (n+1)^\gamma \log((n+1)^\gamma)$, $a_n = (n+1)^{-\gamma}$ avec $0 < \gamma \leq 1$, la deuxième borne du théorème 4.1.1 est $1 - C_\epsilon(n+1)^{-\gamma}$.*

Notons que la vitesse de convergence est algébrique (alors qu'elle est logarithmique pour le recuit classique).

4.1.3 Maximisation de la log-vraisemblance

On suppose que l'on cherche à maximiser une fonction ψ dont on ne connaît que des approximations ψ^N .

Hypothèse 7. *Pour tout $N \geq 1$, nous pouvons calculer une approximation déterministe ou stochastique ψ^N de ψ telle que*

$$\mathbb{E}|\psi^N(x) - \psi(x)| \leq \frac{b_1}{\sqrt{N}}, \text{ pour tout } x \in \Theta.$$

C'est le cas de notre fonction l_n . Nous fixons ici $f(t) = 1/(1+t)$, $\beta_n = (1/a_n) \log(1/a_n)$, $N = N_n = \lceil \beta_n^2 \rceil$. Nous introduisons la suite aléatoire $(\bar{\theta}_n)_{n \geq 0}$ définie par : $\bar{\theta}_0$ quelconque et

1. si on a $\bar{\theta}_n$, on fait une proposition $Z \sim K(\bar{\theta}_n, \cdot)$
2. on pose $\bar{\theta}_{n+1} = Z$ avec probabilité

$$f(\beta_{n+1}(\psi^{N_{n+1}}(\bar{\theta}_n) - \psi^{N_{n+1}}(Z))_+),$$

et $\bar{\theta}_{n+1} = \bar{\theta}_n$ sinon.

Si N_n était constant, nous serions en train de maximiser ψ^N . Ici, la fonction ψ est approchée de plus en plus finement, et maximisée en même temps. L'intérêt d'utiliser un algorithme de type recuit pour la maximisation est qu'il n'est pas gêné par l'imprécision sur ψ .

Théorème 4.1.3 (Rubenthaler, Rydén, Wiktorsson (2009) [3]). *Pour tout $n \geq 1$, $0 < \delta < 1$,*

$$\mathbb{P}(\psi(\bar{\theta}_n) > \psi_{\max} - \epsilon) \geq 1 - \frac{\bar{C}}{1 - \bar{C}} \left((1 + C'' a_1) \bar{C}^{m(\delta-1)} + C'' a_{\lfloor n\delta \rfloor + 1} \right),$$

où $\bar{C} = (2b_1 + \bar{K})/(2b_1 + 1)$, C' est définie dans le théorème 4.1.1 et $C'' = \max(C'(2b_1 + 1)^2/(2b_1 + \bar{K}), 1)$. *Si nous supposons de plus que $1 - a_k \leq a_k/a_{k-1}$, $\forall k$ alors*

$$\mathbb{P}(\psi(\bar{\theta}_n) > \psi_{\max} - \epsilon) \geq 1 - \frac{\bar{C}}{1 - \bar{C}} a_n (1 + C'').$$

Corollaire 4.1.4 (Rubenthaler, Rydén, Wiktorsson (2009) [3]). *Avec $a_n = (n+1)^{-\gamma}$ ($\gamma > 0$), nous pouvons exprimer le taux de convergence en fonction du temps de calcul W*

$$\mathbb{P}(\psi(\bar{\theta}_n) > \psi_{\max} - \epsilon) \leq 1 - \frac{C''' (\log W)^{2\gamma/(2\gamma+1)}}{W^{\gamma/(2\gamma+1)}}.$$

Nous comparons notre algorithme à un algorithme tirant $(\bar{\theta}_n)_{n \geq 0}$ i.i.d. dans Θ et gardant le meilleur (pour ψ^N avec un certain N). Notre algorithme est toujours meilleur en terme de temps de calcul comparé à la précision.

4.2 E.D.S. pour la turbulence

Soit une E.D.S. dans $\Omega = \mathbb{R} \times [-\pi L; \pi L[$ ($L > 0$) de la forme

$$dR_t = \sqrt{2}\sigma(R_t)dB_t, \quad R_0 = r \in \Omega, \quad (4.2.1)$$

où B est un mouvement brownien sur Ω et σ est une matrice 2×2 . On peut voir cette équation comme la projection de l'équation dans \mathbb{R}^2 :

$$d\tilde{R}_t = \sqrt{2}\tilde{\sigma}(\tilde{R}_t)d\tilde{B}_t$$

où $\tilde{\sigma}$ est la fonction $2\pi L$ périodique en la deuxième variable définie de manière adéquate. L'équation (4.2.1) modélise la séparation (le vecteur) entre deux particules de colorant plongées dans un fluide turbulent (théorique puisque tout se passe sur la surface d'un cylindre). Précisons la matrice σ :

$$\sigma = \begin{bmatrix} \sqrt{d_{11}^\kappa(r)} & 0 \\ \frac{d_{12}^\kappa(r)}{\sqrt{d_{12}^\kappa(r)}} & \sqrt{d_{22}^\kappa(r) - \frac{(d_{12}^\kappa(r))^2}{d_{11}^\kappa(r)}} \end{bmatrix},$$

$$\kappa > 0, \quad \xi \in [0; 1], \quad \xi' = \xi + 1,$$

$$d_{11}(r) \underset{|r_1| \rightarrow +\infty}{\sim} \mathcal{D}_1 |r_1|^{1+\xi} + \chi_1,$$

$$d_{22}(r) \underset{|r_1| \rightarrow +\infty}{\sim} \mathcal{D}_2 |r_1|^{1+\xi} + \chi_2,$$

$$\lim_{|r_1| \rightarrow +\infty} d_{12}(r) = \lim_{|r_1| \rightarrow +\infty} d_{12}(r) = 0$$

avec des constantes $\mathcal{D}_1, \mathcal{D}_2, \chi_1, \chi_2 > 0$. Dans [1], nous montrons par des techniques de fonction de Lyapunov que R_t a une mesure invariante. Pour des raisons physiques que nous passons sous silence, ceci montre que l'énergie cinétique du système se diffuse à la fois vers les petites échelles et vers les grandes échelles, à cause du phénomène de turbulence. Le cylindre que nous avons utilisé représente une dimension intermédiaire entre la dimension 3 (où l'énergie se transfère des grandes vers les petites échelles) et la dimension 2 (où c'est l'inverse).

Bibliographie

- [AN72] Krishna B. ATHREYA et Peter E. NEY : *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.
- [AZ97] Rami ATAR et Ofer ZEITOUNI : Exponential stability for nonlinear filtering. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(6):697–725, 1997.
- [Ber06] Jean BERTOIN : *Random fragmentation and coagulation processes*, volume 102 de *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.
- [BO99] A. BUDHIRAJA et D. OCONE : Exponential stability in discrete-time filtering for non-ergodic signals. *Stochastic Process. Appl.*, 82(2): 245–257, 1999.
- [Bon82] Lennart BONDESSON : On simulation from infinitely divisible distributions. *Adv. in Appl. Probab.*, 14(4):855–869, 1982.
- [Bud03] A. BUDHIRAJA : Asymptotic stability, ergodicity and other asymptotic properties of the nonlinear filter. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(6):919–941, 2003.
- [CG06] Mireille CHALEYAT-MAUREL et Valentine GENON-CATALOT : Computable infinite-dimensional filters with applications to discretized diffusion processes. *Stochastic Process. Appl.*, 116(10):1447–1467, 2006.
- [CH08] D. CRISAN et K. HEINE : Stability of the discrete time filter in terms of the tails of noise distributions. *J. Lond. Math. Soc. (2)*, 78(2):441–458, 2008.
- [CL00] Rong CHEN et Jun S. LIU : Mixture Kalman filters. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 62(3):493–508, 2000.
- [Del04] Pierre DEL MORAL : *Feynman-Kac formulae*. Probability and its Applications (New York). Springer-Verlag, New York, 2004. Genealogical and interacting particle systems with applications.
- [DFMP09] R. DOUC, G. FORT, E. MOULINES et P. PRIOURET : Forgetting the initial distribution for hidden Markov models. *Stochastic Process. Appl.*, 119(4):1235–1256, 2009.
- [DG01] Pierre DEL MORAL et Alice GUIONNET : On the stability of interacting processes with applications to filtering and genetic algorithms. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(2):155–194, 2001.

- [DGA00] A. DOUCET, S. GODSILL et C. ANDRIEU : On sequential monte-carlo methods for bayesian filtering. *Stat. Comput.*, 10(3):197–208, 2000.
- [dlPG99] Víctor H. de la PEÑA et Evarist GINÉ : *Decoupling*. Probability and its Applications (New York). Springer-Verlag, New York, 1999. From dependence to independence, Randomly stopped processes. U -statistics and processes. Martingales and beyond.
- [DM83] E. B. DYNKIN et A. MANDELBAUM : Symmetric statistics, Poisson point processes, and multiple Wiener integrals. *Ann. Statist.*, 11(3): 739–745, 1983.
- [DM00] Pierre DEL MORAL et Laurent MICLO : Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering. In *Séminaire de Probabilités, XXXIV*, volume 1729 de *Lecture Notes in Math.*, pages 1–145. Springer, Berlin, 2000.
- [Dur02] Rick DURRETT : *Probability models for DNA sequence evolution*. Probability and its Applications (New York). Springer-Verlag, New York, 2002.
- [Ewe04] Warren J. EWENS : *Mathematical population genetics. I*, volume 27 de *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, second édition, 2004. Theoretical introduction.
- [Fou] Nicolas FOURNIER : Simulation and approximation of lévy-driven stochastic differential equations. Rapport technique, à paraître dans ESAIM PS.
- [GM94] Carl GRAHAM et Sylvie MÉLÉARD : Chaos hypothesis for a system interacting through shared resources. *Probab. Theory Related Fields*, 100(2):157–173, 1994.
- [GM97] Carl GRAHAM et Sylvie MÉLÉARD : Stochastic particle approximations for generalized Boltzmann models and convergence estimates. *Ann. Probab.*, 25(1):115–132, 1997.
- [GM99] Carl GRAHAM et Sylvie MÉLÉARD : Probabilistic tools and Monte-Carlo approximations for some Boltzmann equations. In *CEMRACS 1999 (Orsay)*, volume 10 de *ESAIM Proc.*, pages 77–126 (electronic). Soc. Math. Appl. Indust., Paris, 1999.
- [HC08] Kari HEINE et Dan CRISAN : Uniform approximations of discrete-time filters. *Adv. in Appl. Probab.*, 40(4):979–1001, 2008.
- [HK08] Marc HOFFMANN et Nathalie KRELL : Statistical analysis of self-similar conservative fragmentation chains. Rapport technique hal-00261291, HAL-CNRS, mars 2008.
- [JMP89] A. JAKUBOWSKI, J. MÉMIN et G. PAGÈS : Convergence en loi des suites d'intégrales stochastiques sur l'espace \mathbf{D}^1 de Skorokhod. *Probab. Theory Related Fields*, 81(1):111–137, 1989.
- [KO88] T. G. KURTZ et D. L. OCONE : Unique characterization of conditional distributions in nonlinear filtering. *Ann. Probab.*, 16(1):80–107, 1988.
- [KP91a] Thomas G. KURTZ et Philip PROTTER : Characterizing the weak convergence of stochastic integrals. In *Stochastic analysis (Durham,*

- 1990), volume 167 de *London Math. Soc. Lecture Note Ser.*, pages 255–259. Cambridge Univ. Press, Cambridge, 1991.
- [KP91b] Thomas G. KURTZ et Philip PROTTER : Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19(3):1035–1070, 1991.
- [KP91c] Thomas G. KURTZ et Philip PROTTER : Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. In *Stochastic analysis*, pages 331–346. Academic Press, Boston, MA, 1991.
- [Kun71] Hiroshi KUNITA : Asymptotic behavior of the nonlinear filtering errors of Markov processes. *J. Multivariate Anal.*, 1:365–393, 1971.
- [KV08] M. L. KLEPTSZYNA et A. Yu. VERETENNIKOV : On discrete time ergodic filters with wrong initial data. *Probab. Theory Related Fields*, 141(3-4):411–444, 2008.
- [Lee90] A. J. LEE : *U-statistics*, volume 110 de *Statistics : Textbooks and Monographs*. Marcel Dekker Inc., New York, 1990. Theory and practice.
- [LO03] François LEGLAND et Nadia OUDJANE : A robustification approach to stability and to uniform particle approximation of nonlinear filters : the example of pseudo-mixing signals. *Stochastic Process. Appl.*, 106(2):279–316, 2003.
- [LO04] François LE GLAND et Nadia OUDJANE : Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filters. *Ann. Appl. Probab.*, 14(1):144–187, 2004.
- [Mél98] Sylvie MÉLÉARD : Convergence of the fluctuations for interacting diffusions with jumps associated with Boltzmann equations. *Stochastics Stochastics Rep.*, 63(3-4):195–225, 1998.
- [MS91] Jean MÉMIN et Leszek SŁOMIŃSKI : Condition UT et stabilité en loi des solutions d'équations différentielles stochastiques. In *Séminaire de Probabilités, XXV*, volume 1485 de *Lecture Notes in Math.*, pages 162–177. Springer, Berlin, 1991.
- [MT09] Sean MEYN et Richard L. TWEEDIE : *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second édition, 2009. With a prologue by Peter W. Glynn.
- [RW97] Gian-Carlo ROTA et Timothy C. WALLSTROM : Stochastic integrals : a combinatorial approach. *Ann. Probab.*, 25(3):1257–1283, 1997.
- [Tav04] Simon TAVARÉ : Ancestral inference in population genetics. In *Lectures on probability theory and statistics*, volume 1837 de *Lecture Notes in Math.*, pages 1–188. Springer, Berlin, 2004.
- [vH08] Ramon van HANDEL : Discrete time nonlinear filters with informative observations are stable. *Electron. Commun. Probab.*, 13:562–575, 2008.
- [vH09a] Ramon van HANDEL : Observability and nonlinear filtering. *Probab. Theory Related Fields*, 145(1-2):35–74, 2009.
- [vH09b] Ramon van HANDEL : Uniform time average consistency of Monte Carlo particle filters. *Stochastic Process. Appl.*, 119(11):3835–3861, 2009.