Contrôle stochastique par quantification et applications à la finance

Camille Illand

▶ **To cite this version:**

Camille Illand. Contrôle stochastique par quantification et applications à la finance. Probabilités [math.PR]. Université Pierre et Marie Curie - Paris VI, 2012. Français. NNT : . tel-00768570

HAL Id: tel-00768570

https://theses.hal.science/tel-00768570

Submitted on 21 Dec 2012

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.

UPMC
PARIS UNIVERSITAS

Thèse de doctorat de l'université Pierre et Marie Curie (Paris 6)

Discipline : Mathématiques

présentée par

Camille Illand

pour obtenir le grade de

Docteur en sciences de l'université Paris 6

Contrôle stochastique par quantification et applications à la finance

---

dirigée par Gilles Pagès

Soutenue le 18 décembre 2012 devant le jury composé de :

M. Vlad Bally	Université Paris-Est Marne-La-Vallée	examinateur
M. Bruno Bouchard	Université Paris-Dauphine	rapporteur
M. François Dufour	Université Bordeaux 1	rapporteur
M^{me} Nicole El Karoui	Université Pierre et Marie Curie Paris 6	examinateur
M. Gilles Pagès	Université Pierre et Marie Curie Paris 6	directeur
M. Jacques Printems	Université Paris-Est Créteil	examinateur
M. Nizar Touzi	Ecole Polytechnique	examinateur

Laboratoire de Probabilités et Modèles Aléatoires
Université Paris 6 - Pierre et Marie Curie
4, place Jussieu
75 005 Paris

École doctorale de Sciences Mathématiques de Paris Centre
Case courrier 188
4, place Jussieu
75 005 Paris

Remerciements

Je tiens tout d'abord à remercier mon directeur de thèse, Gilles Pagès. Sans lui, cette thèse n'aurait pu ni commencer ni aboutir. Il m'a été d'une grande aide quant au choix du sujet de ma thèse, il a su m'aiguiller dès le début vers le domaine de la quantification. Je le remercie également pour le temps passé à l'élaboration de la thèse, en particulier pour le lourd travail de relecture. Lors de nos échanges, ses conseils et ses corrections m'ont permis d'améliorer mon manuscrit. Enfin, je souhaite lui adresser un grand merci pour m'avoir soutenu à poursuivre ma thèse alors que ma carrière professionnelle, en temps que quant, était déjà lancée.

Merci également à mes rapporteurs de thèse, Bruno Bouchard et François Dufour, d'avoir accepté de faire partie de mon jury de thèse. Je les remercie pour tout le temps qu'ils ont accordé à la lecture de mon travail. Je les remercie vivement pour leurs remarques très précieuses. Je suis très reconnaissant envers Vlad Bally, Nicole El Karoui, Jacques Printems et Nizar Touzi qui me font l'honneur d'être parmi les membres du jury.

Je tiens à remercier les membres du LPMA, chercheurs et doctorants que j'ai rencontrés lors de ces années, tout particulièrement Abass, Sophie D., Vincent, Paul, Sophie L., Noufel, Nikos, Stavros, Sylvain, qui ont toujours su maintenir une ambiance de travail conviviale.

Cette thèse n'aurait pu s'achever sans le soutien moral de mes amis et de ma famille. Ainsi je remercie tous mes amis qui m'ont entouré ces dernières années, en particulier : Marc, Véro, Antoine, Lulu, Greg, Christophe, Thierry, mes amis lyonnais, bironnais et orléanais. Merci d'avoir partagé mes moments de détente qui, grâce à eux, sont de précieux souvenirs.

Je remercie également toute ma famille. En particulier mes parents, à qui je dédicace cette thèse. Je les remercie pour leur présence constante, toute en délicatesse. En effet, ils m'ont toujours soutenu en me laissant libre de mes choix. J'ai toujours su que je trouverais auprès d'eux la confiance et la chaleur sur lesquelles me reposer. Merci d'avoir toujours accordé beaucoup d'importance à mon éducation et d'avoir été soucieux de développer mes goûts et mes capacités. Je souhaite aussi remercier Sylvain d'être un grand frère présent et d'avoir su partager avec humour mes années d'étudiant.

Enfin, un grand merci à Camille qui me pousse à m'améliorer. Merci pour sa gentillesse, ses encouragements et sa joie de vivre, sans lesquels la rédaction de cette thèse aurait été bien plus difficile.

Table des matières

partie 1. Introduction générale 7
Résumé 8
0.1. Première partie : Résolution de problème de contrôle stochastique par quantification 9
0.2. Deuxième partie : Discrétisation en temps des EDSRs avec barrières 16
0.3. Troisième partie : Réplication robuste de dérivées sur la variance réalisée 20

partie 2. Resolution de problèmes de contrôle stochastique par quantification 25

Chapitre 1. Quantification optimale d'une variable aléatoire 26
1.1. Quantification optimale de variables aléatoires. 26
1.2. Applications aux probabilités numériques 28
1.3. Quantification de processus markovien 28
1.4. Recherche de quantifieurs optimaux 29
1.5. Exemples 30

Chapitre 2. Contrôle par quantification 35
2.1. Introduction 35
2.2. Problème à résoudre 35
2.3. Algorithmes pour l'approximation en espace 38
2.4. Preuves 45
2.5. Reconstruction de la solution 50

Chapitre 3. Exemples d'applications : 53
3.1. Introduction 53
3.2. Options bermudéennes 58
3.3. EDSR avec sauts 59
3.4. EDSR généralisée 62
3.5. EDP parabolique totalement non linéaire 64
3.6. Maximisation d'utilité exponentielle sous contrainte 66
3.7. Optimisation d'un processus de Markov contrôlé 68
3.8. Option americaine en information partielle 70
3.9. Preuves 74

Chapitre 4. Applications numériques 94
4.1. Options américaines 94
4.2. Equations différentielles stochastiques rétrogrades 99
4.3. Valorisation en volatilité incertaine 118

partie 3. A Discrete-time approximation for BSDEs with barriers 128
4.4. Introduction 129
4.5. The forward process 131
4.6. Backward Process 135
Annexe : proofs for the Euler scheme 148

partie 4. Réplication robuste de dérivés sur la variance réalisée 156
4.7. Introduction 157
4.8. Problème inverse 157
4.9. Différentes méthodes d'inversion 162
4.10. Applications pour la couverture d'options sur la variance et le spot 165
4.11. Options sur fond 171
4.12. Extension : processus de Lévy 182
4.13. Preuves 188

Bibliographie 205

Première partie

Introduction générale

Résumé

Cette thèse est constituée de trois parties pouvant être lues indépendamment.

Dans la première partie, on s'intéresse à la résolution de problème de contrôle stochastique par des méthodes de quantification. La quantification consiste à trouver la meilleure approximation d'une loi de probabilité continue par une loi de probabilité discrète avec un nombre donné N de points supportant cette loi. Nous allons expliciter un cadre de programmation dynamique "générique" qui permet de résoudre de nombreux problèmes de contrôle stochastique comme les problèmes de temps d'arrêt optimal, de maximisation d'utilité, d'équations différentielles stochastiques rétrogrades, de filtrage... Dans ce cadre, nous donnons trois schémas de discrétisation en espace associée à la quantification d'une chaîne de Markov.

Dans la deuxième partie, nous présentons un schéma numérique pour les équations différentielles stochastiques rétrogrades doublement réfléchies. Nous nous plaçons dans un cadre général qui contient des sauts et des processus progressifs dépendant de la trajectoire. On propose une approximation du type schéma d'Euler. Nous prouvons la convergence du schéma pour les équations différentielles stochastiques rétrogrades quand le nombre de pas de temps n tend vers l'infini. Nous donnons aussi la vitesse de convergence pour les *game options*.

Dans la troisième partie, on s'intéresse à la réplication des dérivés sur la variance réalisée. On propose une couverture robuste au modèle de volatilité constituée de positions dynamiques sur des options européennes. On étend ensuite cette méthodologie aux options sur fond et aux processus à saut.

0.1. Première partie : Résolution de problème de contrôle stochastique par quantification

0.1.1. Quantification optimale et intégration numérique. On se place dans un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. La quantification d'une variable aléatoire X à valeurs dans $\mathbb{R}^d$ (muni de sa norme euclidienne $|.|$) consiste en son approximation par une variable aléatoire Y prenant un nombre fini de valeurs dans $\mathbb{R}^d$. L'erreur résultant de cette discrétisation est mesurée par la norme $\mathbb{L}^p$ de $|X - Y|$. Quand $X \in \mathbb{L}^p$, on définit un N-quantifieur $\mathbb{L}^p$-optimal de X par la variable aléatoire Y^* solution de problème d'optimisation paramétré par la taille de la grille de quantification N :

$$\min \left\{ ||X - Y||_p \ , \ Y : \Omega \to \mathbb{R}^d \text{ telle que } \mathrm{Card}(Y(\Omega)) \leq N \right\}.$$

Soient Y^* une telle variable aléatoire et $\Gamma = Y^*(\Omega)$. En considérant une projection (borélienne) au plus proche voisin sur Γ $\mathrm{Proj}_\Gamma : \mathbb{R}^d \to \Gamma$ et la variable aléatoire

$$\hat{X}^\Gamma = \mathrm{Proj}_\Gamma(X),$$

on peut facilement vérifier que $\left| X - \hat{X}^\Gamma \right| \leq |X - Y^*|$: $\hat{X}^\Gamma$ est un N-quantifieur $\mathbb{L}^p$-optimal de X. À cette projection correspond un diagramme de Voronoi de $\mathbb{R}^d$ induit par Γ, c'est-à-dire une partition borélienne de l'espace $\mathbb{R}^d$ satisfaisant

$$C_i(\Gamma) \subset \left\{ \xi \in \mathbb{R}^d \ \mid \ |\xi - x_i| = \min_{1 \leq j \leq N} |\xi - x_j| \right\} = \overline{C}_i(\Gamma), \quad i = 1, ..., N$$

où $\overline{C}_i(\Gamma)$ est la fermeture de $C_i(\Gamma)$.

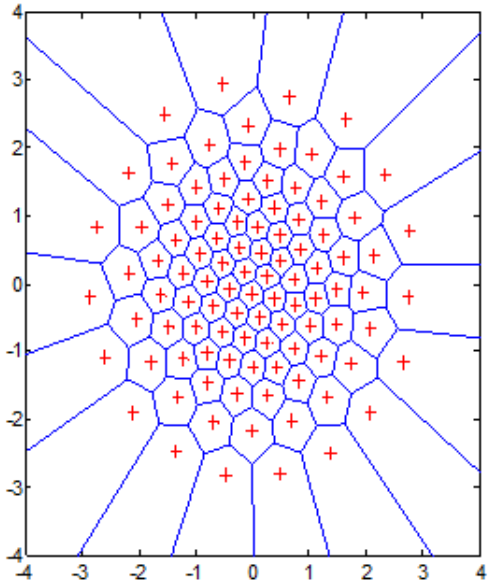


FIG. 0.1.1. Quantification optimale de la loi normale dans $\mathbb{R}^2$ avec $N = 100$ points

L'erreur de quantification $e_{N,p}(X) = \inf\left\{\left|\left|X - \hat{X}^{\Gamma}\right|\right|_p \mid \Gamma \subset \mathbb{R}^d, \mathrm{Card}(\Gamma) \leq N\right\}$ converge vers zéro quand la taille du quantifieur N tend vers l'infini. Sa vitesse de convergence asymptotique est donnée par le théorème de Zador (ref [**Zad82**], [**BW82**], [**GL00**]).

THEOREM 0.1.1. *(Théorème de Zador)*
Soit $p > 0$ et $X \in \mathbb{L}^{p+r}(\mathbb{P})$ pour $r > 0$. Alors,

$$\lim_{N\to+\infty} N^{\frac{p}{d}} \left(e_{N,p}(X)\right)^p = J_{p,d} \left|\left|\varphi\right|\right|_{\frac{d}{d+p}}$$

où $\mathbb{P}_X(d\xi) = \varphi(\xi)\lambda_d(d\xi) + \overline{\mu}(d\xi)$, $\overline{\mu} \perp \lambda_d$ (λ_d mesure de Lebesgue sur $\mathbb{R}^d$) et $J_{p,d} \in]0; +\infty[$.

Le lemme de Pierce établi dans [**LP08**] donne une vitesse de convergence non asymptotique.

LEMMA 0.1.2. *Soit $p \geq 1, \eta > 0$. Il existe une constante $C_{d,p,\eta} > 0$ et un entier $N_{d,p,\eta} \geq 1$ tels que pour toute variable aléatoire $X \in \mathbb{L}^{p+\eta}$ à valeurs dans $\mathbb{R}^d$, pour tout $N \geq N_{d,p,\eta}$,*

$$e_{N,p}(X) \leq C_{d,p,\eta} \left|\left|X\right|\right|_{p+\eta} N^{-\frac{1}{d}}.$$

Des algorithmes stochastiques d'apprentissage basés sur la simulation comme l'algorithme CLVQ et l'algorithme de Lloyd permettent de construire des quantifieurs optimaux.
En outre, les quantifieurs $\mathbb{L}^2$-optimaux vérifient la propriété projective, dite de stationnarité, à savoir que :

$$\mathbb{E}\left[X|\hat{X}^{\Gamma}\right] = \hat{X}^{\Gamma}.$$

Une application majeure de la quantification aux probabilités numériques est l'intégration numérique.

En effet, la quantification peut être vue comme une alternative à la méthode de Monte Carlo pour l'approximation numérique de l'intégrale $\mathbb{E}\left[f(X)\right]$, pour X de loi $\mathbb{P}_X$ sur $\left(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\right)$ et f une fonction borélienne de $\mathbb{R}^d$ et $f(X) \in \mathbb{L}^1$. Une formule de type "cubature" ($\hat{X}$ est une variable aléatoire discrète) donne l'estimateur :

$$\mathbb{E}\left[f(\hat{X}^{\Gamma})\right] = \sum_{x\in\Gamma} f(x)\mathbb{P}\left(\hat{X}^{\Gamma} = x\right) = \sum_{x\in\Gamma} f(x)\mathbb{P}\left(X \in C_x(\Gamma)\right),$$

où Γ est, si possible, une grille $\mathbb{L}^p$-optimale. Les pondérations $\mathbb{P}\left(X \in C_x(\Gamma)\right)$ sont des paramètres pouvant être calculés en même temps que la construction de la grille de quantification.

L'erreur d'estimation dépend de la régularité de la fonction f. Rappelons quelques résultats élémentaires.

$\diamond$ Si f est lipschitzienne de rapport $[f]_{Lip} = \sup_{x\neq y} \frac{|f(x)-f(y)|}{|x-y|}$, alors on a :

$$\begin{aligned}\left|\mathbb{E}\left[f(X)\right] - \mathbb{E}\left[f(\hat{X}^{\Gamma})\right]\right| &\leq \mathbb{E}\left[\left|f(X) - f(\hat{X}^{\Gamma})\right|\right] \\ &\leq [f]_{Lip}\left|\left|X - \hat{X}^{\Gamma}\right|\right|_r, \quad \forall r \geq 1.\end{aligned}$$

Si Γ est une N-quantification $\mathbb{L}^p$-optimale, alors le théorème de Zador donne une estimation de l'erreur :

$$\mathbb{E}\left[f(X)\right] - \mathbb{E}\left[f(\hat{X}^{\Gamma})\right] = O(N^{-\frac{1}{d}}).$$

avec les notations de Landau (o et O).

$\diamond$ Si la fonction f est 2 fois différentiable, à dérivée seconde bornée et si Γ est une N-quantification $\mathbb{L}^2$-optimale, alors la propriété de stationnarité $\mathbb{E}\left[X|\hat{X}\right] = \hat{X}$ permet d'établir une majoration d'erreur d'ordre 1 :

$$\begin{aligned}\left|\mathbb{E}\left[f(X)\right] - \mathbb{E}\left[f(\hat{X}^{\Gamma})\right]\right| &\leq [Df]_{Lip}\left|\left|X - \hat{X}^{\Gamma}\right|\right|_2^2 \\ &= O(N^{-\frac{2}{d}}).\end{aligned}$$

Dans un cadre plus général, nous allons maintenant nous intéresser à des méthodes liées à la quantification pour résoudre numériquement des problèmes de contrôle stochastique associé à une chaîne de Markov.

0.1.2. Description de problème de contrôle stochastique. Nous nous nous plaçons un cadre à temps discret. Sur un espace de probabilité filtré $(\Omega, \mathcal{F} = (\mathcal{F}_k)_{k=0,1,..,n}, \mathbb{P})$, $n \in \mathbb{N}^*$, on considère une $(\mathcal{F}_k)_{0\leq k\leq n}$-chaîne de Markov $(X_k)_{0\leq k\leq n}$ à valeurs dans $\left(\mathbb{R}^d, |.|\right)$ muni de sa norme euclidienne $|.|$.

De nombreux problèmes de contrôle optimal sont liés au principe de programmation dynamique. On pourrait citer des problèmes de temps d'arrêt optimal (options américaines ou bermudéennes en finance), des problèmes de maximisation d'utilité, d'optimisation de processus de Markov contrôlé... Pour les résoudre, la première étape est d'écrire la programmation dynamique. Commençons par illustrer cette programmation dynamique pour le problème d'enveloppe de Snell.

Exemple : La valorisation d'une option bermudéenne de payoff $h_k(X_k)$ revient à résoudre le problème de temps d'arrêt optimal :

$$Y_k = \text{esssup}_{\tau\in\mathcal{T}_{k,n}}\mathbb{E}\left[h_\tau(X_\tau)|\mathcal{F}_k\right]$$

où $\mathcal{T}_{k,n}$ représente l'ensemble des temps d'arrêts à valeurs dans $\{k,..,n\}$. L'enveloppe de Snell $\{Y_k\}_{0\leq k\leq n}$ satisfait alors $Y_k = y_k(X_k)$, $0 \leq k \leq n$, où les fonctions $y_k : \mathbb{R}^d \to \mathbb{R}$ vérifient le principe de programmation dynamique rétrograde suivant :

$$\begin{cases} y_n(x) = h_n(x), \quad x \in \mathbb{R}^d \\ y_k(x) = \max\left\{h_k(x); \mathbb{E}\left[y_{k+1}(X_{k+1})|X_k = x\right]\right\}, \quad x \in \mathbb{R}^d, 0 \leq k \leq n-1. \end{cases}$$

Ainsi, ce problème de temps d'arrêt revient à résoudre le schéma par programmation dynamique rétrograde précédent.

Dans le cadre général que nous allons adopter dans ce mémoire, la solution $Y_k = y_k(X_k)$ d'un problème de contrôle stochastique sera définie par un schéma par programmation dynamique descendante. On considère un processus de contrôle a_k $\sigma(X_k)$-mesurable à valeurs dans un compact A_k.

On suppose que $Y_k = y_k(X_k)$ vérifie le principe de progammation dynamique générique défini par $y_k = B_k(y_{k+1})$:

$$(0.1.1) \qquad \begin{cases} y_n(x) = h(x) \\ u_{k,a}(x) = \mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(x,\xi_{k+1})|X_k = x\right], \quad 0 \leq k \leq n-1, \\ v_k(x,a) = T_{k,a}\left(x, u_{k,a}(x)\right), \quad 0 \leq k \leq n-1, \\ y_k(x) = R_k(x, v_k(x,a)), \quad 0 \leq k \leq n-1, \end{cases}$$

où les fonctions $h : \mathbb{R}^d \to \mathbb{R}$, $\phi_{k,a} : \mathbb{R}^d \times \mathbb{R}^{d_\xi} \to \mathbb{R}^{d'}$, $T_{k,a} : \mathbb{R}^d \times \mathbb{R}^{d'} \to \mathbb{R}$ et $R_k : \mathbb{R}^d \times \mathbb{R}^{A_k} \to \mathbb{R}$ sont déterministes et où $(\xi_k)_{1\leq k\leq n}$ une suite de variables aléatoires indépendantes identiquement distribuées $\mathcal{F}_k$-mesurables à valeurs dans $\mathbb{R}^{d_\xi}$ (ξ_k est indépendante de $\mathcal{F}_{k-1}$).

Interprétation : Ce cadre de programmation dynamique générique permet de résoudre de nombreux problèmes en finance. Il est plus général que dans l'article [**PP04**]. Nous donnerons différents exemples d'utilisation de cette programmation dynamique générique : applications pour les options américaines, pour les équations différentielles stochastiques rétrogrades (EDSR), applications au problème de maximisation d'utilité, d'optimisation d'un processus de Markov contrôlé, de contrôle en information partielle.

La fonction h représente la fonction valeur à l'instant final n. Les fonctions $u_{k,a}$ nécessitent des calculs d'espérances conditionnelles. Les fonctions $\phi_{k,a}$ pourront être par exemple un changement de probabilité (résultat d'une transformation du type Girsanov) ou représenter des accroissements d'un mouvement brownien comme pour le schéma pour les EDSR. Les fonctions $v_k(x,a)$ représente la fonction valeur pour un contrôle $a \in A_k$ fixé. La fonction R_k représentera l'optimisation du contrôle sur $a \in A_k$: il s'agit par exemple de la réflexion dans le cas des options américaines ou bien d'une maximisation d'utilité...

Cette méthode est particulièrement intéressante quand la dimension devient grande, typiquement quand $d \geq 4$, ce qui arrive fréquemment en finance quand les options considérées portent sur plusieurs sous-jacents.

0.1.3. Schémas de quantification. Nous venons de décrire un principe de programmation dynamique générique adapté à de nombreux problèmes de contrôle stochastique. Nous allons maintenant proposer trois algorithmes liés à la quantification pour résoudre ces problèmes de contrôle stochastique : une régression sur une base quantifiée, une méthode d'arbre sur une grille de quantification et une méthode des gerbes. Nous commençons alors pour nous ramener à une structure markovienne $(X_k)_{0\leq k\leq n}$. A toute date k, nous obtenons le quantifieur $\hat{X}_k$ de X_k avec sa grille de quantification $\Gamma_k = \left\{x_i^k, 1 \leq i \leq N_k\right\}$ de cardinal N_k et avec son diagramme de Voronoi $(C_i^k)_{1\leq i\leq N_k}$. Puis, nous modifions la programmation dynamique générique (0.1.1) en remplaçant les espérances conditionnelles $\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k,\xi_{k+1})|X_k\right]$ par une estimation associée à l'un des trois algorithmes.

◇ **Cas de la régression sur une base quantifiée**

Ici, nous obtenons une discrétisation en espace par une régression linéaire sur les vecteurs orthogonaux $\left\{1_{X_k \in C_i^k}\right\}_{1 \leq i \leq N_k}$. Cette méthode peut aussi être interprétée soit comme une méthode de noyaux avec $K(x_i, X_k) = 1_{x \in C_i^k}$ ou soit comme une approximation par la sous-filtration $\sigma(\hat{X}_k)$. Les espérances conditionnelles $\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|X_k\right]$ sont alors estimées (via une méthode de Monte Carlo en général) par

$$\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right] = \sum_{i=1}^{N_k} \frac{\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})1_{X_k \in C_i^k}\right]}{\mathbb{P}\left(X_k \in C_i^k\right)} 1_{X_k \in C_i^k}.$$

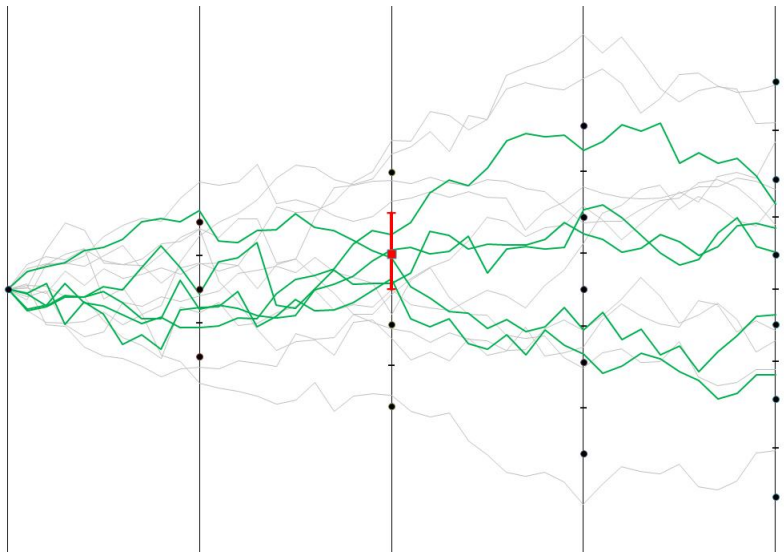


FIG. 0.1.2. Exemple d'une régression du mouvement brownien sur une base quantifiée

◇ **Algorithme par arbre de quantification**

L'application projection n'étant pas injective, le processus $\hat{X}_k$ ne conserve pas la propriété de Markov de X_k. Cependant, nous définissons des probabilités de transition à partir des $\hat{X}_k$ par :

$$\begin{aligned} p_{i,j}^k &= \mathbb{P}\left(\hat{X}_{k+1} = x_j^{k+1} | \hat{X}_k = x_i^k\right) \\ &= \mathbb{P}\left(X_{k+1} \in C_j^{k+1} | X_k \in C_i^k\right). \end{aligned}$$

Il faut supposer ici qu'il existe des fonctions $\psi_{k,a} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d'}$ telles que

$$\phi_{k,a}(X_k, \xi_{k+1}) = \psi_{k,a}(X_k, X_{k+1}).$$

Nous pouvons alors approcher les espérances conditionnelles $\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(x, \xi_{k+1})|X_k = x\right]$ pour $x \in C_i^k$ par

$$\sum_j p_{i,j}^k y_{k+1}(x_j^{k+1})\psi_{k,a}(x_i^k, x_j^{k+1}).$$

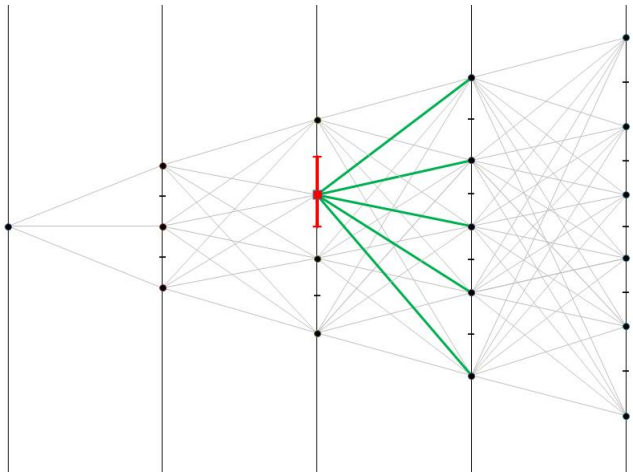


FIG. 0.1.3. Exemple d'un arbre de quantifcation approchant un mouvement brownien

◇ **Méthodes des gerbes**

Ici, nous supposons que la dynamique de la $(\mathcal{F}_k)$-chaîne de Markov $(X_k)_k$ est donnée par

$$X_{k+1} = F_k(X_k, \xi_{k+1})$$

avec F_k une fonction déterministe. On approxime le semi-groupe de transition ξ_{k+1} par la variable aléatoire discrète $\hat{\xi}_{k+1}$ indépendante de $\mathcal{F}_k$. Cette discrétisation de ξ_{k+1} peut se faire par plusieurs méthodes : par exemple par Monte Carlo, par quasi-Monte Carlo, par quantification ou par une méthode de quadrature...

Nous définissons maintenant $\mathcal{P}_{\Gamma_k}$ un prolongement dans $\mathbb{R}^d$ des fonctions de Γ_k : pour une fonction $f : \Gamma_k \to \mathbb{R}$, on a $\mathcal{P}_{\Gamma_k}(f) : \mathbb{R}^d \to \mathbb{R}$ avec $\mathcal{P}_{\Gamma_k}(f)(x) = f(x)$ $\forall x \in \Gamma_k$. Nous proposerons plusieurs types de prolongement : approximation par le plus proche voisin, par méhode de noyaux, par interpolation linéaire sur une grille cartésienne ou sur une triangulation de Delaunay.

Après avoir calculé y_{k+1} en les points de la grille $\Gamma_{k+1} = \left\{x_j^{k+1}, 1 \leq j \leq N_{k+1}\right\}$ de manière rétrograde, on obtient y_k sur la grille de quantifcation Γ_k à l'aide des estimations des espérances conditionnelles $\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|X_k = x_i^k\right]$, $x_i^k \in \Gamma_k$ définies par

$$\mathbb{E}\left[\mathcal{P}_{\Gamma_k}(y_{k+1})\left(F_k(x_i^k, \xi_{k+1})\right)\phi_{k,a}(x_i^k, \hat{\xi}_{k+1})\right], \quad 0 \leq k \leq n-1.$$

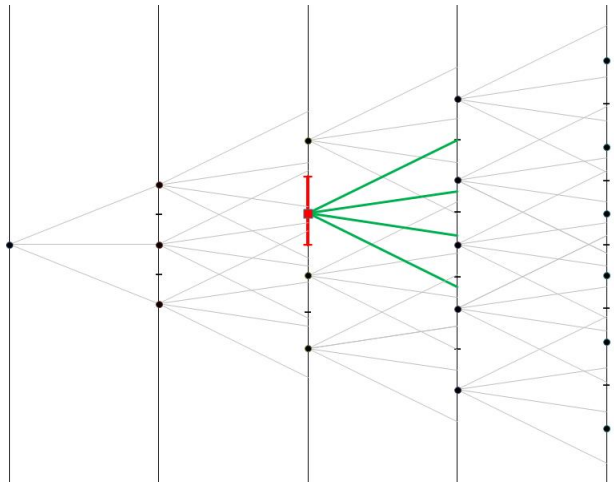


FIG. 0.1.4. Exemple d'une approximation du mouvement brownien par la méthode des gerbes

0.1.4. Convergence des schémas de quantification. Sous certaines hypothèses de régularité, nous montrons que l'erreur de l'estimation de $y_k(X_k)$ par les schémas de quantification $\hat{y}_k(X_k)$ est majorée par

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq K_{p,n} \sum_{k=1}^{n-1} \left(1 + 2 \sup_{a \in A_k} [u_{k,a}]_{\text{Lip}}\right) \left|\left|X_k - \hat{X}_k\right|\right|_p$$

Sous certaines hypothèses d'intégrabilité de $\nabla_u T_{k,a}.\phi_{k,a}$, on obtient que la constante $K_{p,n}$ peut être rendue indépendante du nombre de pas de temps n. En appliquant le lemme de Pierce, la majoration précédente de l'erreur d'approximation devient :

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq \tilde{K}_{d,p,n,\eta} \sum_{k=1}^{n-1} \left(1 + 2 \sup_{a \in A_k} [u_{k,a}]_{\text{Lip}}\right) ||X_k||_{p+\eta} N_k^{-\frac{1}{d}}$$

La convergence des schémas par quantification ne nécessite alors plus qu'une régularité lischitzienne des fonctions $u_{k,a}$.

Dans ce but, nous montrons au cas par cas, à l'aide d'arguments sur la régularité du semi-groupe, mais surtout sur le flôt de la chaîne de Markov X, que les fonctions $u_{k,a}$ vérifient la propriété lipschitzienne pour les différents problèmes de contrôle stochatique suivants : les options américaines, les équations différentielles stochastiques rétrogrades, les équations différentielles paraboliques totalement non linéaires, la maximisation d'utilité, l'optimisation d'un processus de Markov contrôlé, la maximisation d'une option bermudéenne en information partielle.

0.2. Deuxième partie : Discrétisation en temps des EDSRs avec barrières

Dans la première partie, nous avons étudié la discrétisation en espace de problèmes de contrôle stochastique, plus précisément nous avons présenté des méthodes de quantification optimale dans un cadre général de la programmation dynamique générique associée à une chaîne de Markov. Lorsque le problème de contrôle stochastique se situe en temps continu, la résolution numérique nécessite généralement de proposer une discrétisation en temps et de prouver sa convergence. Dans cette deuxième partie, nous allons ainsi nous intéresser plus particulièrement à la discrétisation en temps des équations différentielles stochatiques rétrogrades (EDSR) en mettant en évidence que c'est un cas particulier rentrant dans le cadre de la programmation dynamique générique étudiée dans la première partie.

Les équations différentielles stochastiques rétrogrades (EDSR) ont été introduites par J.-M. Bismut ([**Bis73**]) en 1973 dans le cas d'un générateur linéaire et par E. Pardoux et S. Peng ([**PP90**]) en 1990 dans le cas d'un générateur non-linéaire. De nombreux mathématiciens ont contribué et contribuent toujours à la théorie des EDSR : on pourrait par exemple citer des résultats sur des estimations à priori ([**KPQ97**]), sur les relations avec les équations aux dérivées partielles ([**BBP97**]), sur des applications à la maximisation d'utilité ([**RK00**] et dans le cadre avec sauts [**Bec05**]), pour des problèmes de couverture ([**EL05**]), pour une extension dans la cadre des diffusion avec sauts les travaux ([**TL94**]), sur une discrétisation en temps dans ([**Zha04**] et dans le cas des sauts [**BE08**]).

Différents schémas de discrétisation ont été étudiés : à l'aide d'un algorithme appelé schéma en quatre étapes ([**MPY94**]), via un théorème du type Donsker ([**BDM01**], [**PX08**]), par régression ([**Che97**], [**JLW05**]), par quantification ([**BP03a**], [**BPP05**], [**DM06**], [**DM08**]), par une méthode de Monte Carlo séquentielle ([**GL07**]), avec un schéma progressif ([**BR07**]), par calcul de Malliavin ([**BT04**]).

Les EDSRs avec une barrière sont liées aux problèmes de temps d'arrêts optimaux (comme pour les options américaines), on pourra par exemple citer ([**KKP$^+$97**]) et ([**BP03a**]), ([**BC08**]) pour les approximations numériques. Avec deux barrières, ces équations peuvent évaluer une *game option* (voir [**Kif00**]), c'est-à-dire une option où l'acheteur et le vendeur peuvent excercer le droit de acheter ou vendre l'option à des prix convenus. L'existence et l'unicité des EDSR avec doubles réflexions sont étudiées dans ([**CK96**]) dans le cadre continu et dans ([**CM08**]), ([**HH06**]) dans le cas avec sauts. Nous proposons ici une généralisation de certains schémas développés dans un cadre sans saut et sans une fonctionnelle Φ_t pour le processus progressif X dans les articles ([**Cha09**]), ([**Otm08**], ([**Xu08**]), ([**Kif06**]) : nous étendons à un processus progressif path-dépendant $X_t = (S_t, \Phi_t(S))$ où S est un processus de diffusion avec saut et Φ_t est une fonctionnelle $\mathbb{L}^\infty$-lipschitzienne.

Nous nous intéressons maintenant au problème de discrétisation des EDSR avec double réflexions avec sauts :

$$
(0.2.1) \qquad \begin{cases} Y_t = l \vee g \wedge u(X_T) + \int_t^T f(\Theta_s)ds + \left(K_T^+ - K_t^+\right) - \left(K_T^- - K_t^-\right) \\ \quad - \int_t^T Z_s dsW_s - \int_t^T \int_E U_s(e)\overline{\mu}(de, ds), \quad \forall t \in [0,T],\ \mathbb{P}-p.s. \\ l(X_t) \leq Y_t \leq u(X), \quad \forall t \in [0,T],\ \mathbb{P}-a.s. \\ \text{et } \int_0^T (Y_s - l(X_s))\, dK_s^+ = \int_0^T (Y_s - u(X_s))\, dK_s^- = 0, \quad \mathbb{P}-p.s. \end{cases}
$$

où $K^{\pm}$ sont continus et croissants et $\Theta := (X, Y, Z, \Gamma)$ avec $\Gamma := \int_E \rho(e)U(e)\lambda(de)$. Le processus progressif $X_t = \Phi_t\left((S_s)_{0\leq s\leq t}\right)$ à valeurs dans $\mathbb{R}^d$ est une fonctionnelle du processus S qui est solution de l'équation différentielle stochastique progressive :

$$S_t = S_0 + \int_0^t b(S_r)dr + \int_0^t \sigma(S_r)dW_r + \int_0^t \int_E \beta(S_{r-}, e)\overline{\mu}(de, dr).$$

Ici, W est un mouvement brownien de dimension d_W et $\overline{\mu}$ est une mesure de Poisson compensée ($\overline{\mu}(de, dr) = \mu(de, dr) - \lambda(de)ds$) indépendante de W.

Etant donnés la "grille d'approximation" $\pi = \{0 = t_0 < t_1 < .. < t_n = T\}$ (une partition de l'intervalle $[0, T]$ de module $|\pi|$) et X^π le schéma d'Euler de X, on approche la solution (0.2.1) par le schema rétrograde suivant :

$$\begin{cases} Y_T^\pi = l\left(X_T^\pi\right) \vee g\left(X_T^\pi\right) \wedge u\left(X_T^\pi\right), \\ \overline{Z}_{t_i}^\pi = \mathbb{E}\left[Y_{t_{i+1}}^\pi \frac{W_{t_{i+1}} - W_{t_i}}{t_{i+1} - t_i} | \mathcal{F}_{t_i}\right], \\ \overline{\Gamma}_{t_i}^\pi = \mathbb{E}\left[Y_{t_{i+1}}^\pi \frac{\int_E \rho(e)\overline{\mu}(de,(t_i,t_{i+1}])}{t_{i+1} - t_i} | \mathcal{F}_{t_i}\right], \\ \tilde{Y}_{t_i}^\pi = \mathbb{E}\left[Y_{t_{i+1}}^\pi | \mathcal{F}_{t_i}\right] + (t_{i+1} - t_i) f\left(X_{t_i}^\pi, \tilde{Y}_{t_i}^\pi, \overline{Z}_{t_i}^\pi, \overline{\Gamma}_{t_i}^\pi\right), \\ Y_{t_i}^\pi = l\left(X_{t_i}^\pi\right) \vee \tilde{Y}_{t_i}^\pi \wedge u\left(X_{t_i}^\pi\right). \end{cases}$$

Remarque : Si X^π est une chaîne de Markov, ce schéma discret en temps peut se réécrire dans notre programmation dynamique générique (0.1.1) avec $Y_{t_i}^\pi = y_k(X_{t_i}^\pi)$:

$$\begin{cases} y_n(x) = h(x) = l(x) \vee g(x) \wedge u(x) \\ u_k(x) = \mathbb{E}\left[y_{i+1}(X_{t_{i+1}}^\pi)\phi_k(\xi_{k+1}) | X_{t_i}^\pi = x\right] \\ y_k(x) = l(x) \vee \left(u_k^1(x) + \frac{T}{n} f(x, u_k^1(x), \frac{T}{n} u_k^2(x), \frac{T}{n} u_k^3(x))\right) \wedge g(x), \quad \forall x \in \mathbb{R}^d, \forall 0 \leq i < n. \end{cases}$$

avec

$$\begin{aligned} \phi_k(\xi_{k+1}) &= \left(1, W_{t_{k+1}} - W_{t_k}, \int_E \rho(e)\overline{\mu}(de, (t_k, t_{k+1}])\right), \\ u_k^1(x) &= \mathbb{E}\left[y_{k+1}(X_{t_{k+1}}^\pi) | X_{t_k}^\pi = x\right], \\ T_k(x, u) &= u^1 + \frac{T}{n} f(x, u^1, \frac{n}{T} u^2, \frac{n}{T} u^{3,}), \\ R_k(x, v) &= l(x) \vee v \wedge g(x), \quad \forall k \in \{0, .., n-1\}, \forall x \in \mathbb{R}^d, \forall v \in \mathbb{R}. \end{aligned}$$

Le cadre est alors adapté pour appliquer une discrétisation en espace par quantification étudiée dans la première partie.

Pour montrer la convergence du schéma, nous allons étudier l'approximation par le schéma d'Euler pour le processus progressif X, puis pour le processus rétrograde Y.

0.2.1. Convergence du processus progressif X. On note

$$\begin{aligned} \pi(t) :&= \sup\{t_i \in \pi | t_i \leq t\}, \quad \forall t \in [0, T], \\ |\pi| &:= \sup_{0\leq i<n} |t_{i+1} - t_i|, \end{aligned}$$

Ici, nous nous focalisons sur l'erreur de la discrétisation du processus progressif. Nous définissons le schéma d'Euler S^π du processus progressif S par

$$S_t^\pi = S_{\pi(t)}^\pi + b(S_{\pi(t)}^\pi)(t-\pi(t)) + \sigma(S_{\pi(t)}^\pi)(W_t - W_{\pi(t)}) + \int_E \beta(S_{\pi(t)}^\pi, e)\overline{\mu}(de, (\pi(t), t])$$

auquel nous associons le schéma d'Euler constant par morceaux $\overline{S}_t^\pi = S_{\pi(t)}^\pi$, $0 \leq t \leq T$.

On peut montrer qu'il existe une constante $C > 0$ telle que

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |S_t - S_t^\pi|^2\right] \leq C\,|\pi|\,.$$

En fait, on souhaite contrôler l'erreur du schéma d'Euler constant par morceaux $\left|S_t - S_{\pi(t)}^\pi\right|$, $\forall t \in [0,T]$. Dans le cas continu (i.e. $\beta = 0$), on a la majoration suivante

$$\mathbb{E}\left[\sup_{0\leq t\leq T} \left|S_t - S_{\pi(t)}^\pi\right|^2\right] \leq C\,|\pi|\,\ln(\frac{1}{|\pi|}).$$

Cependant, cette dernière inégalité est fausse pour une diffusion avec saut. En effet, si on considère un processus de Poisson N, le schéma d'Euler ne converge pas fortement :

$$\mathbb{E}\left[\sup_{0\leq t\leq T} \left|N_t - N_{\pi(t)}^\pi\right|^2\right] \geq \mathbb{P}\left(N_T \geq 1\right).$$

Cela pose alors problème dans le cadre path-dépendant avec saut. Par exemple, si nous considérons la fonctionnelle $\Phi_t : \mathbb{R}^{[0,T]} \to \mathbb{R}$ $\mathbb{L}^\infty$-lipschitzienne

$$\Phi_t\left((x_s)_{0\leq s\leq T}\right) = \sup_{0\leq s\leq t} x_s$$

avec $S = N$, alors le schéma d'Euler $X_t^\pi = \Phi_t\left(\left(N_{\pi(s)}^\pi\right)_{0\leq s\leq T}\right)$ ne converge pas le processus progressif $X_t = \Phi_t\left((N_s)_{0\leq s\leq T}\right)$ dans $\mathbb{L}^1$: en effet, on a

$$\begin{aligned}\mathbb{E}\left[|X_t - X_t^\pi|\right] &= \mathbb{E}\left[\left|\sup_{0\leq s\leq t} N_s - \sup_{0\leq s\leq t} N_{\pi(s)}^\pi\right|\right] = \mathbb{E}\left[\sup_{0\leq t\leq T}\left|N_t - N_{\pi(t)}^\pi\right|\right] \\ &\geq \mathbb{P}\left(N_T \geq 1\right)\end{aligned}$$

On a besoin de "tordre" un peu le temps et l'espace pour obtenir la convergence du processus càdlàg X_t^π. Ainsi, nous proposons d'utiliser la topologie de Skorokhod (au lieu d'utiliser la topologie uniforme traditionnelle qui autorise uniquement à "tordre" un peu l'espace) pour obtenir une convergence forte du schéma d'Euler $\overline{S}^\pi$.

On se place alors dans $\mathbb{D}\left([0,T], \mathbb{R}^d\right)$ l'espace des fonctions càdlàg avec la métrique de Skorokhod de $[0,T]$ dans $\mathbb{R}^d$:

$$d_{\mathbb{D}}(x,y) = \inf_{\lambda\in\Lambda}\left(\sup_{t\in[0,T]} |\lambda(t) - t| + \sup_{t\in[0,T]} |x(\lambda(t)) - y(t)|\right)$$

où Λ représente l'ensemble des bijections strictement croissantes de $[0,T]$ dans lui-même.

PROPOSITION 0.2.1. *Avec la métrique de Skorokhod $d_{\mathbb{D}}$, on obtient la même vitesse de convergence pour le schéma d'Euler que dans le cas continu avec la distance uniforme :*

$$\mathbb{E}\left[\sup_{t_i\in\pi} d_{\mathbb{D}}\left(S^{t_i},(S^{\pi})^{t_i}\right)^2\right] \leq C\,|\pi|\,,$$
$$\mathbb{E}\left[\sup_{t_i\in\pi} d_{\mathbb{D}}\left(S^{t_i},(\overline{S}^{\pi})^{t_i}\right)^2\right] \leq C\,|\pi|\ln(\frac{1}{|\pi|}).$$

où

$$x_t^{t_i} = x_{t\wedge t_i}, \quad \forall x\in\mathbb{D}([0,T],\mathbb{R}^d), \forall t,t_i\in[0,T].$$

Il suffit alors de contruire Φ régulière par rapport de la métrique $d_{\mathbb{D}}$ pour contrôler l'erreur du schéma d'Euler $X_t^{\pi} = \Phi_t\left(\left(S^{\pi}_{\pi(s)}\right)_{0\leq s\leq T}\right)$ du processus progressif path-dépendant $X_t = \Phi_t\left((S_s)_{0\leq s\leq T}\right)$: nous obtenons finalement

$$\mathbb{E}\left[\sup_{t_i\in\pi}\left|X_{t_i}-X^{\pi}_{t_i}\right|^2\right] \leq C\,|\pi|\ln(\frac{1}{|\pi|}).$$

.

0.2.2. Convergence du processus rétrograde Y. Dans la seconde section, nous nous intéressons à l'approximation du processus rétrograde Y. Nous prouvons que

$$\mathbb{E}\left[\sup_{t\in[0,T]}|Y_t-Y_t^{\pi}|^2\right] \to_{|\pi|\to 0} 0$$

dans le cas général et avec une vitesse de convergence d'au moins $|\pi|\ln\frac{1}{|\pi|}$ pour les *game options* (avec f independante de (Z,Γ)). Nous montrons que

$$\mathbb{E}\left[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\left\{\left|Z_s-\overline{Z}^{\pi}_{t_i}\right|^2+\left|\Gamma_s-\overline{\Gamma}^{\pi}_{t_i}\right|^2\right\}\right] \to_{|\pi|\to 0} 0.$$

sous des conditions de Mokobodski sur les barrières.

Notre schéma d'approximation en temps, complété par une discrétisation en espace (qui donne une estimation des différentes espérances conditionnelles, par exemple, par une méthode de quantification comme dans la première partie), permet donc d'approcher les solutions des équations différentielles stochastiques rétrogrades dans un cadre général (EDSR doublement réfléchies avec un processus progressif avec saut et dépendant de la trajectoire).

0.3. Troisième partie : Réplication robuste de dérivées sur la variance réalisée

L'étude de la volatilité est devenue centrale dans la finance quantitative. En 1973, Black et Scholes ([**BS73**]) ont introduit leur modèle lognormal où le prix d'une action possède une volatilité instantanée constante. Cependant, ce modèle est en désaccord avec les prix du marché. Ainsi, les traders qui utilisent le modèle de Black et Scholes pour la couverture doivent de façon permanente modifier les paramètres de volatilité pour obtenir les prix de marché. Or, l'introduction des modèles à volatilité stochastique permet de rester compatible avec le fait que des options de strikes et de maturités différents ont des volatilités implicites différentes.

Les marchés financiers traitent maintenant de nombreux sous-jacents autres que les actions et des indices boursiers. En particulier, ces dernières années, de nombreux contrats de swap sur la variance réalisée sont échangés sur les marchés de gré à gré et sont devenus liquides. Des dérivés dont les payoff sont des fonctions de la variance réalisée (et plus généralement aussi fonctions de l'actif risqué) sont aussi négociés de gré à gré, tel le swap de volatilité qui est la racine de la variance réalisée. Notre étude va se concentrer sur la valorisation et la couverture de ces options sur la variance réalisée. Habituellement, cela nécessite de choisir un modèle sur la volatilité (car cette dernière n'est pas directement observable) et de définir une méthode de calibration.

Pour éviter cette problématique, nous proposons dans cet article de donner une valorisation et une couverture des produits de payoff mixant actif risqué et variance réalisée par une approche robuste par rapport au modèle de volatilité, c'est-à-dire indépendante du modèle de volatilité.

Dans un espace filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$, on considère que le prix de l'actif risqué S suit l'équation de diffusion continue

$$\frac{dS_t}{S_t} = \sigma_t dW_t$$

où W est un $\mathcal{F}_t$-mouvement brownien, σ_t un processus $\mathcal{F}_t$-adapté et $\mathbb{P}$ la probabilité risque-neutre. Par souci de simplification, on suppose que les taux et les dividendes sont nuls.

Pour $0 \leq s \leq t \leq T$, on note le processus logarithme

$$X_{s,t} := \ln\left(\frac{S_t}{S_s}\right)$$

et la variance réalisée

$$\langle X\rangle_{s,t} := \int_s^t \sigma_u^2 du.$$

0.3.1. Réplication du variance swap. Un variance swap est un contrat qui paie $\langle X\rangle_{0,T}$ moins un montant convenu fixé. En appliquant la formule d'Itô, on obtient

$$\langle X\rangle_{0,T} = -2X_{0,T} + \int_0^T \frac{2}{S_u} dS_u$$

Ainsi, la stratégie auto-financée suivante replique le contrat variance swap de manière indépendante. A toute date $t \leq T$, nous portons une position statique sur le

log contrat $-2X_{0,T} = -2\ln\left(\frac{S_T}{S_0}\right)$, plus une position dynamique sur l'actif risqué S et une position en obligation :

$$\begin{aligned} 1 \quad & \text{log contrat qui paie } -2X_{0,T} = -2\ln\left(\frac{S_T}{S_0}\right) \\ \frac{2}{S_t} \quad & \text{position sur } S_T \\ \int_0^T \frac{2}{S_u} dS_u - 2 \quad & \text{en obligation} \end{aligned}$$

Sinon, le variance swap peut ainsi être répliqué indépendamment de tout modèle avec une position dynamique à la date t sur l'option européenne $P_t(X_{t,T})$ avec

$$\begin{aligned} P_t(x) &= -2x + 2\left(e^x - 1\right) - 2X_{0,t} + \int_0^t \frac{2}{S_u} dS_u \\ P_t(X_{t,T}) &= -2\ln\left(\frac{S_T}{S_t}\right) + 2\left(\frac{S_T}{S_t} - 1\right) - 2\ln\left(\frac{S_T}{S_t}\right) + \int_0^t \frac{2}{S_u} dS_u \end{aligned}$$

0.3.2. Réplication pour les options de payoff $H(X_{0,T}, \langle X\rangle_{0,T})$. Comme pour les variance swaps, on souhaite répliquer les options de payoff $H(X_{0,T}, \langle X\rangle_{0,T})$ mixant actif risqué et variance réalisée de manière robuste au modèle sur la volatilité.

Pour cela, nous commençons par définir les fonctions prix h_t et "delta" δ_t à toutes dates t pour tous niveaux de variance v sous le modèle de Black-Scholes (modèle à volatilité constante) :

$$\begin{aligned} h_t(v) &:= \mathbb{E}\left[H(X_{0,T}, \langle X\rangle_{0,T})|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right], \quad \forall v > 0, \\ \delta_t(v) &:= \frac{\partial}{\partial x}\mathbb{E}\left[H(X_{0,T}, \langle X\rangle_{0,T})|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right], \quad \forall v > 0. \end{aligned}$$

Nous proposons alors de résoudre la problème inverse : à la date t, trouver le payoff F sur $X_{t,T}$ tel que l'option européenne $F(X_{t,T})$ possède le même prix et même delta que l'option $H(X_T, \langle X\rangle_T)$ pour tous niveaux de variance déterministe :

$$\begin{aligned} \mathbb{E}\left[F(X_{0,T})|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right] &= h_t(v), \quad \forall v > 0, \\ \frac{\partial}{\partial x}\mathbb{E}\left[F(X_{0,T})|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right] &= \delta_t(v), \quad \forall v > 0. \end{aligned}$$

Nous obtenons une formulation explicite de la solution $P(h_t, \delta_t)$ de ce problème inverse :

$$\begin{aligned} P(h_t, 0)(x) &:= \sqrt{\pi} e^{\frac{x}{2}} |x| \mathcal{L}^{-1}\left(\frac{e^{\frac{1}{16\lambda}}}{\lambda^{\frac{1}{2}}} h_t\left(\frac{1}{2\lambda}\right)\right)(x^2) \\ &\quad - \frac{\sqrt{\pi}}{4} e^{\frac{x}{2}} \text{sign}(x) \mathcal{L}^{-1}\left(\frac{e^{\frac{1}{16\lambda}}}{\lambda^{\frac{3}{2}}} h_t\left(\frac{1}{2\lambda}\right)\right)(x^2), \\ P(0, \delta_t)(x) &:= \frac{\sqrt{\pi}}{2} e^{\frac{x}{2}} \text{sign}(x) \mathcal{L}^{-1}\left(\frac{e^{\frac{1}{16\lambda}}}{\lambda^{\frac{3}{2}}} \delta_t\left(\frac{1}{2\lambda}\right)\right)(x^2), \\ P(h_t, \delta_t) &:= P(h_t, 0) + P(0, \delta_t). \end{aligned}$$

où $\mathcal{L}$ représente la transformée de Laplace $\mathcal{L}f(\lambda) = \int_{\mathbb{R}^+} f(s)e^{-\lambda s}ds$.

Nous proposons de valoriser et répliquer l'option de payoff $H(X_T, \langle X \rangle_T)$ par une position dynamique sur l'option européenne de payoff $P(h_t, \delta_t)(X_{t,T})$ à chaque date t.

La réplication est exacte et indépendante du modèle sur la volatilité si l'on suppose que l'actif risqué S_t et la volatilité σ_t sont indépendants.

La réplication est en partie immunisée lorsque l'on relâche cette hypothèse d'indépendance entre S_t et σ_t. En effet, nous montrons que la couverture est robuste à l'ordre 2 en la corrélation volatilité-actif risqué.

Ensuite, nous appliquons cette couverture robuste pour différents payoffs particulièrement intéressants en finance. Pour cette réplication robuste, nous retrouvons le log contrat pour la variance swap, nous obtenons le payoff de réplication donné dans ([**CL08**]) pour les payoffs exponentielles de variance $e^{\lambda \langle X \rangle_T}$. Nous illustrons aussi la réplication robuste pour le swap de volatilité $\sqrt{\langle X \rangle_{0,T}}$ et pour le rapport de Sharpe $\frac{e^{X_T}-1}{\sqrt{\langle X \rangle_{0,T}}}$.

0.3.3. Extension aux options sur fonds. On cherche à couvrir des options sur fond de manière robuste par rapport à la dynamique du fond. On suppose que le fond est lié à un indice $S_t = e^{X_{t,T}}$ et de sa variation quadratique $\langle X \rangle_{0,T}$. Plus précisément, on considère un fond F qui est un processus markovien de loi

$$\forall t < T \quad F_T \sim G_{t,T}(F_t, X_{t,T}, \langle X \rangle_{t,T}, Z)$$

où Z est une variable aléatoire indépendante de $(X_{t,T}, \langle X \rangle_{t,T})$.

Exemple 1 : Formulation générale d'un fond mutuel benchmarké. Un fond mutuel benchmarké avec une volatilité benchmarkée suit la dynamique

$$\frac{dF_t}{F_t} = \rho\beta \frac{dS_t}{S_t} + \sqrt{(1-\rho^2)\beta^2\sigma_t^2 + \epsilon^2} dW_t^{\perp}$$

où $W^{\perp}$ est un mouvement brownien indépendant de S. Le fond F vaut à la date T

$$F_T = G_{t,T}(F_t, X_{t,T}, \langle X \rangle_{t,T}, Z)$$
$$G_{t,T}(f, x, v, z) := f e^{\rho\beta x + z\sqrt{(1-\rho^2)\beta^2 v + \epsilon^2(T-t)} - \frac{(\beta^2 - \rho\beta)v + \epsilon^2(T-t)}{2}}$$

où $Z \sim \mathcal{N}(0,1)$ indépendante de $(X_{t,T}, \langle X \rangle_{t,T})$. Son processus de volatilité satisfait $(\sigma_t^F)^2 = \beta^2(\sigma_t^S)^2 + \epsilon^2$.

Exemple 2 : Portefeuille Constant Mix. En prenant $\rho = 1$, $\epsilon = 0$, on obtient la stratégie Constant Mix qui peut être exprimée comme une puissance de l'indice corrigée par une fonction de la variation quadratique de l'indice :

$$\frac{dF_t}{F_t} = \beta \frac{dS_t}{S_t},$$
$$F_T = F_t \left(\frac{S_T}{S_t}\right)^{\beta} e^{-\frac{(\beta^2-\beta)}{2}\langle X \rangle_{t,T}}.$$

De la même manière que pour les options de payoff $H(X_{0,T}, \langle X \rangle_{0,T})$, nous obtenons une couverture robuste pour les options sur fond.

Pour la couverture robuste des options *call* sur un fond mutuel F, nous donnons une formule explicite de la position de couverture dynamique sur une option européenne sur S qui peut être approchée par la combinaison d'un *call* et d'un *put* sur l'indice sous-jacent S avec des strikes adéquats. Cette stratégie permet de définir des marchés sur la volatilité échangeable sur des fonds mutuels et permet de réduire le risque de modèle pour les options sur fond.

0.3.4. Extension aux processus à saut. Ici, nous relâchons l'hypothèse de continuité du processus de prix de l'actif risqué S en rajoutant des sauts.

Dans le cas continu, la variance quadratique $\langle X \rangle_{0,T}$ est la variable inconnue. On a cherché une réplication robuste de l'option de payoff $H(X_{0,T}, \langle X \rangle_{0,T})$ par rapport à une modélisation de la variance $\langle X \rangle_{t,T}$.

On peut aussi interpréter la variance $\langle X \rangle_{0,t}$ comme un changement de temps : nous réécrivons le processus de prix sous la forme $S_t = e^{X_{0,t}}$ avec

$$V_t X_{0,t} = L_{V_t} \quad \text{et} \quad V_t := \langle X \rangle_{0,t}$$

où $L_s = B_s - \frac{s}{2}$ est un mouvement brownien avec drift et V_t représente le changement de temps. Il s'agit maintenant d'évaluer le changement de temps qui est la variable inconnue.

En rajoutant des sauts, on étend notre processus L aux processus de Lévy. On souhaite trouver une couverture robuste d'option de payoff $H(X_{0,T}, V_T)$ par rapport à une erreur de modélisation sur le subordinateur V_t du procesus de Lévy, c'est-à-dire par rapport au changement de temps.

Le marché étant devenu incomplet, la réplication n'est plus parfaite. Nous proposons deux méthodes pour réduire le risque : une volarisation robuste par rapport au subordinateur et une couverture en moyenne variance.

Deuxième partie

Resolution de problèmes de contrôle stochastique par quantification

CHAPITRE 1

Quantification optimale d'une variable aléatoire

Issue du traitement du signal et de l'information, la quantification est une manière de discrétiser l'espace d'état d'un ensemble d'observables. Cette technique a été developpée initialement dans le but de quantifier la transmission d'information ([**GG00**]). Son utilisation s'est élargie à différents domaines comme la physique, l'informatique et, plus récemment, les probabilités numériques.

Commençons par quelques rappels sur la quantification optimale d'une variable aléatoire dans un cadre euclidien et donnons une esquisse de résultats théoriques pratiques.

1.1. Quantification optimale de variables aléatoires.

Plaçons nous dans un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et considérons une variable aléatoire X à valeurs dans $\left(\mathbb{R}^d, |.|\right)$ supposé muni de sa norme euclidenne usuelle.

Soit $N \in \mathbb{N}$. Un N-quantifieur est défini comme un ensemble

$$\Gamma := \{x_1, ..., x_N\} \subset \mathbb{R}^d$$

avec $\mathrm{Card}(\Gamma) = N$. Γ peut aussi être appelé grille. Puis, on peut discrétiser (ou quantifier) la variable X par $q(X)$ où $q : \mathbb{R}^d \to \Gamma$ est une fonction borélienne. L'inégalité suivante

$$|X(\omega) - q(X(\omega))| \geq d(X(\omega), \Gamma) = \min_{1 \leq i \leq N} |X(\omega) - x_i|, \quad \forall \omega \in \Omega$$

met en évidence que le meilleur choix pour approcher X est une projection par plus proche voisin sur Γ, notée Proj_Γ. À cette projection correspond un diagramme de Voronoi de $\mathbb{R}^d$ induit par Γ, c'est-à-dire une partition borélienne de l'espace $\mathbb{R}^d$ satisfaisant

$$C_i(\Gamma) \subset \left\{\xi \in \mathbb{R}^d \,|\, |\xi - x_i| = \min_{1 \leq j \leq N} |\xi - x_j|\right\} = \overline{C}_i(\Gamma), \quad i = 1, ..., N$$

où $\overline{C}_i(\Gamma)$ est la fermeture de $C_i(\Gamma)$. Alors la projection par plus proche voisin sur Γ vaut

$$\mathrm{Proj}_\Gamma(\xi) = \sum_{i=1}^{N} x_i 1_{C_i(\Gamma)}(\xi).$$

Puis on définit une N-quantification (de Voronoi) de X par

$$\hat{X}^\Gamma(\omega) = \sum_{i=1}^{N} x_i 1_{C_i(\Gamma)}(X(\omega)) \quad \forall \omega \in \Omega.$$

Il est clair que

$$\left|X(\omega) - \hat{X}^\Gamma(\omega)\right| = d(X(\omega), \Gamma) = \min_{x \in \Gamma} |X(\omega) - x_i|, \quad \forall \omega \in \Omega.$$

L'erreur $\mathbb{L}^p$ de quantification est définie par

$$e_p(\Gamma, X) = \left|\left|X - \hat{X}^{\Gamma}\right|\right|_p = \left\{\mathbb{E}\left[\min_{1\leq i\leq N} |X(\omega) - x_i|^p\right]\right\}^{\frac{1}{p}}.$$

Quand $X \in \mathbb{L}^p$, on définit un N-quantifieur $\mathbb{L}^p$-optimal de X par une grille Γ_N^* solution du problème d'optimisation paramétré par la taille de la grille de quantification N :

$$e_{N,p}(X) = \inf\left\{\left|\left|X - \hat{X}^{\Gamma}\right|\right|_p |\Gamma \subset \mathbb{R}^d, \mathrm{Card}(\Gamma) \leq N\right\}.$$

Des résultats dans [**Pag97**] montrent que ce problème admet une solution $\Gamma_N^* = \{x_1, ..., x_N\}$ vérifiant :

$$\mathbb{E}\left[\left|X - \hat{X}^{\Gamma}\right|^p\right] = \mathbb{E}\left[\min_{x\in\Gamma} |X(\omega) - x|^p\right].$$

L'erreur de quantification $e_{N,p}(X)$ converge vers zero quand la taille du quantifieur N tend vers l'infini. Sa vitesse de convergence asymptotique est donnée par le théorème de Zador (ref [**Zad82**], [**BW82**], [**GL00**]).

THEOREM 1.1.1. *(Théorème de Zador)*
Soit $p > 0$ et $X \in \mathbb{L}^{p+r}(\mathbb{P})$ pour $r > 0$. Alors,

$$\lim_{N\to+\infty} N^{\frac{p}{d}} \left(e_{N,p}(X)\right)^p = J_{p,d} \left|\left|\varphi\right|\right|_{\frac{d}{d+p}}$$

où $\mathbb{P}_X(d\xi) = \varphi(\xi)\lambda_d(d\xi) + \overline{\mu}(d\xi)$, $\overline{\mu} \perp \lambda_d$ (λ_d mesure de Lebesgue sur $\mathbb{R}^d$) et $J_{p,d} \in]0; +\infty[$.

REMARK 1.1.2. On montre que $J_{p,d}$ correspond à la limite pour la loi uniforme sur $[0,1]^d$:

$$J_{p,d} = \lim_{N\to+\infty} N^{\frac{p}{d}} \left(e_{N,p}(U\left([0,1]^d\right))\right)^p$$
$$J_{p,d} = \mathrm{liminf}_{N\to+\infty} N^{\frac{p}{d}} \left(e_{N,p}(U\left([0,1]^d\right))\right)^p.$$

La vraie valeur de $J_{p,d}$ est inconnue pour $d \geq 3$, mais on sait que $J_{p,1} = \frac{1}{2^r(r+1)}$ et que $J_{2,d} \sim \frac{d}{2\pi e}$ lorsque d tend vers $+\infty$.

Le lemme de Pierce établi dans [**LP08**] donne une vitesse de convergence non asymptotique.

LEMMA 1.1.3. *Soit $p \geq 1, \eta > 0$. Il existe une constante $C_{d,p,\eta} > 0$ et un entier $N_{d,p,\eta} \geq 1$ tels que pour toute variable aléatoire $X \in \mathbb{L}^{p+\eta}$ à valeurs dans $\mathbb{R}^d$, pour tout $N \geq N_{d,p,\eta}$,*

$$e_{N,p}(X) \leq C_{d,p,\eta} \left|\left|X\right|\right|_{p+\eta} N^{-\frac{1}{d}}.$$

Les quantifieurs $\mathbb{L}^2$-optimaux vérifient une propriété de stationnarité :

$$\mathbb{E}\left[X|\hat{X}^{\Gamma_N^*}\right] = \hat{X}^{\Gamma_N^*}.$$

Cette propriété permettra d'avoir des correcteurs pour des schémas d'ordre 1 pour les algorithmes de regression ou d'avoir une bonne estimation de l'incrément martingale pour les schémas par interpolation pour les équations différentielles stochastiques rétrogrades par exemple.

1.2. Applications aux probabilités numériques

La quantification peut être vue comme une alternative à la méthode de Monte Carlo pour l'approximation numérique de l'intégrale $\mathbb{E}[f(X)]$, pour X de loi $\mathbb{P}_X$ sur $\left(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\right)$ et f une fonction borélienne de $\mathbb{R}^d$ et $f(X) \in \mathbb{L}^1$. Une formule de type "cubature" ($\hat{X}$ est une variable aléatoire discrète) donne l'estimateur :

$$\mathbb{E}\left[f(\hat{X}^\Gamma)\right] = \sum_{x\in\Gamma} f(x)\mathbb{P}\left(\hat{X}^\Gamma = x\right) = \sum_{x\in\Gamma} f(x)\mathbb{P}\left(X \in C_x(\Gamma)\right),$$

où Γ est si possible une grille $\mathbb{L}^p$-optimale. Les pondérations $\mathbb{P}\left(X \in C_x(\Gamma)\right)$ sont des paramètres pouvant être calculés en même temps que la construction de la grille de quantification.

L'erreur d'estimation dépend de la régularité de la fonction f. Rappelons quelques résultats élémentaires.

$\diamond$ Si la fonction f est lipschitzienne de rapport $[f]_{Lip} = \sup_{x\neq y} \frac{|f(x)-f(y)|}{|x-y|}$, alors on a :

$$\begin{aligned}\left|\mathbb{E}[f(X)] - \mathbb{E}\left[f(\hat{X}^\Gamma)\right]\right| &\leq \mathbb{E}\left[\left|f(X) - f(\hat{X}^\Gamma)\right|\right] \\ &\leq [f]_{Lip}\left\|X - \hat{X}^\Gamma\right\|_r, \quad \forall r \geq 1.\end{aligned}$$

Si Γ est une N-quantification $\mathbb{L}^p$-optimale, alors le théorème de Zador donne une estimation de l'erreur :

$$\mathbb{E}[f(X)] - \mathbb{E}\left[f(\hat{X}^\Gamma)\right] = O(N^{-\frac{1}{d}}).$$

avec les notations de Landau (o et O).

$\diamond$ Si la fonction f est 2 fois différentiable, à dérivée seconde bornée et si Γ est une N-quantification $\mathbb{L}^2$-optimale, alors la propriété de stationnarité permet d'établir une majoration d'erreur d'ordre 1 :

$$\begin{aligned}\left|\mathbb{E}[f(X)] - \mathbb{E}\left[f(\hat{X}^\Gamma)\right]\right| &\leq [Df]_{Lip}\left\|X - \hat{X}^\Gamma\right\|_2^2 \\ &= O(N^{-\frac{2}{d}}).\end{aligned}$$

$\diamond$ Si la fonction f est convexe et $f(X) \in \mathbb{L}^1$, l'inégalité de Jensen donne la majoration suivante :

$$\mathbb{E}\left[f(\hat{X}^\Gamma)\right] \leq \mathbb{E}[f(X)].$$

L'intégration numérique par quantification est une technique déterministe de calcul d'espérance "optimisée" par rapport à un nombre de points (les grilles et les pondérations sont calculées offline). Ainsi, cette méthode peut être utilisée pour calculer rapidemment des nombreuses espérances conditionnelles.

1.3. Quantification de processus markovien

Pour les algorithmes de quantification par arbre ([**BP03a**], [**BPP05**]), il est nécessaire de quantifier un processus Markovien à temps discret $(X_k)_k$. Dans ce cas, on quantifie chaque variable X_k avec sa loi marginale et on parle de quantification marginale (notion différente de la quantification fonctionnelle de processus [**LP02**], [**PP05b**]). D'abord, à chaque pas de temps, on projette la variable aléatoire X_k

sur un N_k-quantifieur ($\mathbb{L}^p$-optimal) $\Gamma_k = \left\{x_1^{(k)}, ..., x_{N_k}^{(k)}\right\}$ suivant la règle du plus proche voisin, ce qui s'écrit :

$$\hat{X}_k = \mathrm{Proj}_{\Gamma_k}(X_k) = \sum_{i=1}^{N_k} x_i^{(k)} 1_{C_i(\Gamma_k)}(X_k).$$

Cependant, le processus quantifié $\left(\hat{X}_k\right)_k$ n'est pas une chaîne de Markov. Mais il est possible d'approximer les probabilités de transition entre deux dates à l'aide de noyaux de transition $p_{i,j}^{(k)}$ définis par

$$\begin{aligned} p_{i,j}^{(k)} &= \mathbb{P}\left(X_{k+1} \in C_j(\Gamma_{k+1}) | X_k \in C_i(\Gamma_k)\right) \\ &= \mathbb{P}\left(\hat{X}_{k+1} = x_j^{(k+1)} | \hat{X}_k = x_i^{(k)}\right). \end{aligned}$$

Ces probabilités de transition peuvent être calculées hors-ligne (par exemple par Monte Carlo). On aura ainsi une approximation de l'espérance conditionnelle

$$\begin{aligned} \mathbb{E}\left[f(X_{k+1})|X_k\right] &\sim \mathbb{E}\left[f(\hat{X}_{k+1})|\hat{X}_k\right] \\ &\sim \sum_{j=1}^{N_{k+1}} p_{i,j}^{(k)} f(x_j^{(k)}), \quad \text{si } X_k \in C_i(\Gamma_k). \end{aligned}$$

ce qui nous permettra de "forcer" un pseudo-principe de programmation dynamique sur la variable quantifiée.

1.4. Recherche de quantifieurs optimaux

La recherche d'un N-quantifieur optimal est un problème d'optimisation assez délicat, tout particulièrement dans le cas multidimensionnel. En dimension 1, il arrive que l'on dispose de formules fermées (cas de la loi uniforme) ou de formules semi-fermées (les cas des loi puissance, des lois exponentielles et de Pareto) ce qui rend le calcul plus aisé (voir [**FP02**]). Sinon on peut généralement appliquer des algorithmes d'optimisation numériques déterministes comme celle de gradient (Newton) ou de point fixe (Lloyd 1).

Cependant, la situation se complique dans le cas multidimensionnel. La solution du quantifieur optimal n'est pas unique et les calculs d'intégrales deviennent plus ardus. Une approche probabiliste est alors préférable. Présentons deux algorithmes stochastiques d'apprentissage basés sur la simulation.

1.4.1. Algorithme CLVQ. Aussi appelé algorithme de plus proche voisin ou algorithme de Kohonen à 0 voisin (ref. [**BMP90**], [**Pag97**]), l'algorithme CLVQ (Competitive Leaning Vector Quantization) est un gradient stochastique. Pour $p \geq 1$ et pour une taille de grille N fixée, on simule une suite i.i.d. $\left(\xi^k\right)_k$ dans $\mathbb{L}^2$ de loi X et on définit la suite $\left(X^k\right)_k$, partant de $X^0 \in \left(\mathbb{R}^d\right)^N$, par

$$X_i^{k+1} = X_i^k - \gamma_{k+1} 1_{\{\xi^{k+1} \in C_i(X^k)\}} \left|X_i^k - \xi^{k+1}\right|^{p-2} \left(X_i^k - \xi^{k+1}\right), \quad \forall 1 \leq i \leq N,$$

où $\gamma = (\gamma_k)_k$ est une suite à pas décroissant vers 0 et vérifiant

$$\sum_k \gamma_k = +\infty,$$
$$\sum_k \gamma_k^2 < +\infty.$$

L'algorithme se décompose en deux phases. Dans une première phase, on sélectionne le plus proche voisin et dans une seconde il s'agit d'apprentisage en modifiant le plus proche point de la valeur ξ^{k+1} et en laissant les autres points inchangés.

1.4.2. Algorithme de Lloyd 1. L'algorithme de Lloyd 1 (cf. **[Llo82]** et **[GG00]**) permet d'obtenir une quantification $\mathbb{L}^2$-optimale en se basant sur la stationarité de la solution $\mathbb{E}\left[X|\hat{X}^{\Gamma_N^*}\right] = \hat{X}^{\Gamma_N^*}$ pour $X \in \mathbb{L}^2$.

Initialisée en $X^0 \in \left(\mathbb{R}^d\right)^N$, on définit la suite $\left(X^k\right)_k$ par

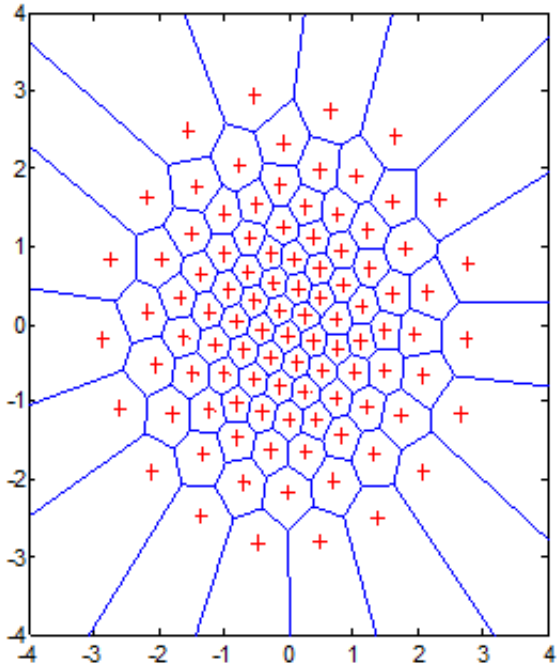
$$X_i^{k+1} = \mathbb{E}\left[X|X \in C_i(X^k)\right], \quad \forall 1 \leq i \leq N.$$

Cette espérance conditionnelle est calculée par simulation de Monte Carlo. L'algorithme converge vers l'équilibre se situant dans le bassin d'attraction de la grille initiale. La phase d'initialisation est donc primordiale.

Pour une loi X sur laquelle on a peu d'information, il peut ainsi être conseillé de commencer par initialiser avec des simulations de Monte Carlo, puis d'améliorer la grille avec l'algorithme CLVQ pendant un certain temps (pour arriver dans le bassin d'attraction). Enfin, on applique l'algorithme de Lloyd 1 (cf. **[PP03]**).

1.5. Exemples

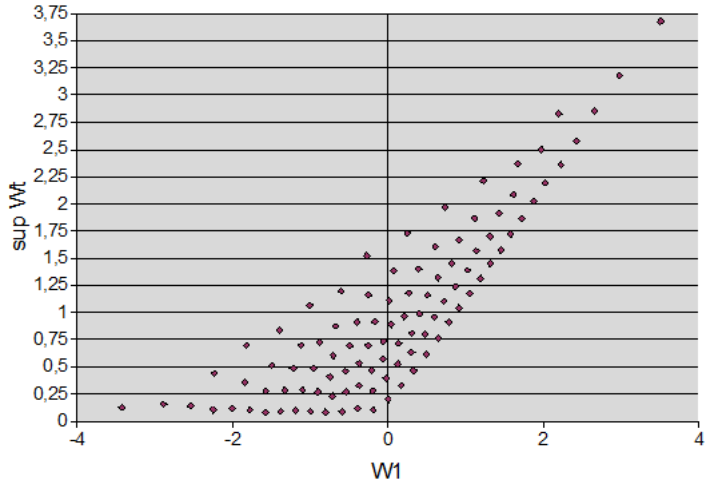
Exemple loi normale. Quantification optimale de la loi normale dans $\mathbb{R}^2$ avec $M = 100$ points :



La quantification de la loi normale permet de quantifier de nombreux processus en finance en ne faisant juste une petite transformation (passage au log, scaling, changement de probabilité par Girsanov) : dans le modèle de Black-Scholes, l'actif risqué est modélisé par une exponentielle de la loi normale : $S_t = S_0 \exp\left((r - \frac{\sigma^2}{2})t + \sigma W_t\right)$.

De plus, la quantification optimale de la loi normale permet un calcul rapide d'espérances. Ainsi, la quantification optimale pourra être utilisée pour le calcul d'espérances conditionnelles.

Exemple : Brownien et son supremum. Quantification optimale de la loi $\left(W_1, \sup_{0\leq s\leq 1} W_s\right)$ dans $\mathbb{R}^2$ avec $M = 100$ points :



Un scaling et un changement de probabilité permettent de ramener la loi de $(S_t, \sup_{0\leq s\leq t} S_s)$ sous Black-Scholes à une exponentielle de $\left(W_1, \sup_{0\leq s\leq 1} W_s\right)$.

Exemple : Option asiatique. Quantification de la loi de (S_t, A_t) où S_t suit le modèle de Black-Scholes et $A_t = \frac{1}{t}\int_0^t S_s ds$. Nous allons quantifier cette loi en trois étapes.

Tout d'abord, la grille initale est obtenue de manière aléatoire par un tirage de Monte Carlo :

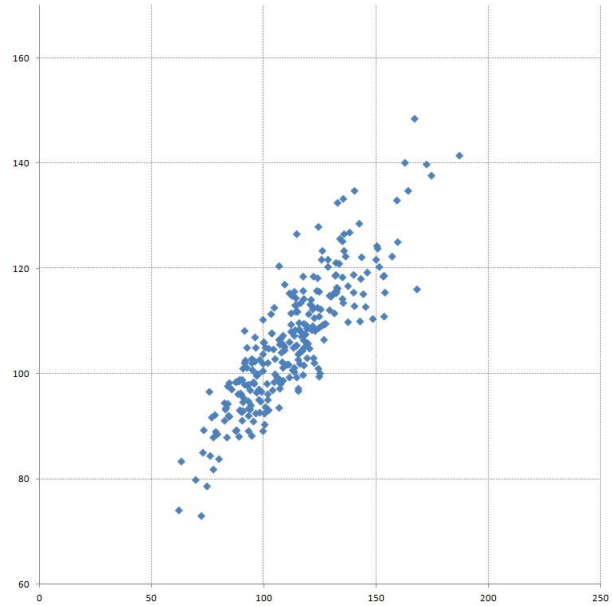


FIG. 1.5.1. Cette grille initiale pour la loi $X_t = (S_t, A_t)$ est obtenue par un tirage de Monte Carlo ($N = 270$). L'erreur de quantification vaut $e_2(\Gamma, X) = 5,70$.

Ensuite, la grille est ensuite améliorée par l'algorithme CLVQ :

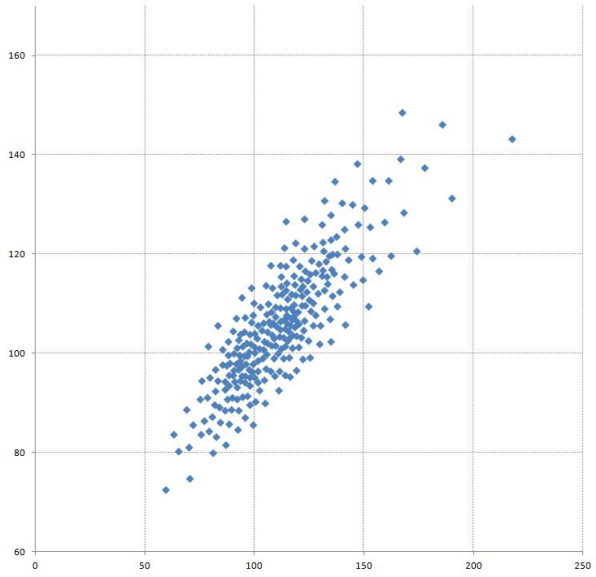


FIG. 1.5.2. En appliquant l'algorithme CLVQ à la grille initiale, on obtient une approximation de la loi $X_t = (S_t, A_t)$ est obtenue par un tirage de Monte Carlo ($N = 270$). L'erreur de quantification diminue et vaut $e_2(\Gamma, X) = 3,10$.

Finalement, on applique obtient ce qui donne une grille plus homogène en appliquant Lloyd :

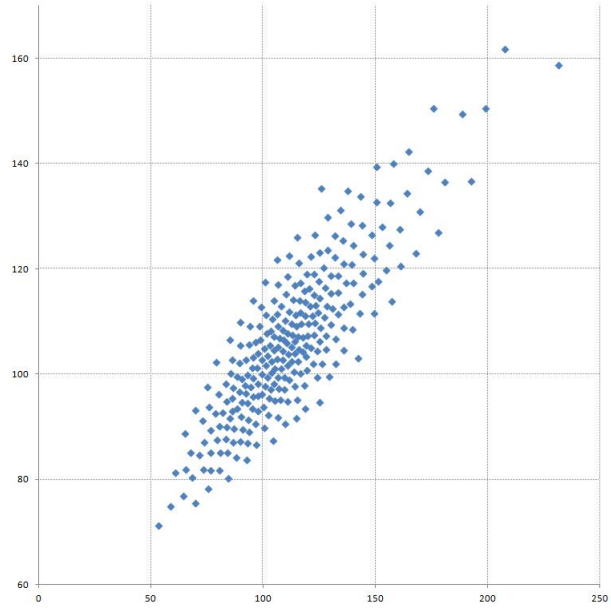


FIG. 1.5.3. Finalement, on applique l'algorithme de Lloyd pour obtnenir une grille plus régulière de la loi de $X_t = (S_t, A_t)$. L'erreur de quantification diminue et vaut $e_2(\Gamma, X) = 2,35$.

CHAPITRE 2

Contrôle par quantification

2.1. Introduction

Nous allons appliquer des méthodes de quantification pour résoudre numériquement des problèmes non linéaires associés à une chaîne de Markov en dimension d. Nous allons proposer trois méthodes :
◇ l'une par régression linéaire sur des fonctions constantes par morceaux définies par la quantification
◇ la deuxième en approximant la dynamique de la chaîne de Markov par un arbre défini sur la grille de quantification
◇ la troisième est la méthode des gerbes ; les probabilités de transition sont estimées sur chaque point de la grille.

Dans une première partie, nous allons définir un principe de programmation dynamique générique adapté à de nombreux problèmes de contrôle stochastique. Dans ce cadre, nous allons développer ces trois algorithmes liés à la quantification pour résoudre ces problèmes. Nous allons montrer la convergence et donner la vitesse de convergence des différents schémas. Dans une deuxième partie, nous allons appliquer notre algorithme de programmation dynamique génériques à quelques exemples en finance. On devra alors se ramener à une structure markovienne, quantifier le processus markovien et montrer que les schémas vérifient les hypothèses de convergence de l'algorithme.

2.2. Problème à résoudre

2.2.1. Description. Nous nous nous plaçons un cadre à temps discret. Sur un espace de probabilité filtré $(\Omega, \mathcal{F} = (\mathcal{F}_k)_{k=0,1,..,n}, \mathbb{P})$, $n \in \mathbb{N}^*$, on considère une $(\mathcal{F}_k)_{0\leq k\leq n}$-chaîne de Markov $(X_k)_{0\leq k\leq n}$ à valeurs dans $\left(\mathbb{R}^d, |.|\right)$ muni de sa norme euclidienne $|.|$. Soit $(\xi_k)_{1\leq k\leq n}$ une suite de variables aléatoires indépendantes identiquement distribuées $\mathcal{F}_k$-mesurables à valeurs dans $\mathbb{R}^{d_\xi}$ et ξ_k est indépendantes de $\mathcal{F}_{k-1}$. On considère un processus de contrôle a_k $\sigma(X_k)$-mesurable à valeurs dans un compact A_k.

De nombreux problèmes de contrôle optimal sont liés au principe de programmation dynamique. On pourrait citer des problèmes de temps d'arrêt optimal (options américaines ou bermudéennes en finance), des problèmes de maximisation d'utilité, d'optimisation de processus de Markov contrôlé... Pour les résoudre, la première étape est d'écrire la programmation dynamique.

EXAMPLE 2.2.1. Par exemple, la valorisation d'une option bermudéenne de payoff $h_k(X_k)$ revient à résoudre le problème de temps d'arrêt optimal :

$$Y_k = \text{esssup}_{\tau\in\mathcal{T}_{k,n}} \mathbb{E}\left[h_\tau(X_\tau)|\mathcal{F}_k\right]$$

où $\mathcal{T}_{k,n}$ représente l'ensemble des temps d'arrêts à valeurs dans $\{k,..,n\}$ et (Y est aussi appelé l'enveloppe de Snell de $\{h_k(X_k)\}_{0\leq k\leq n}$). L'enveloppe de Snell $\{Y_k\}_{0\leq k\leq n}$ satisfait alors $Y_k = y_k(X_k)$, $\forall 0 \leq k \leq n$, où les fonctions

$y_k : \mathbb{R}^d \to \mathbb{R}$ vérifient le principe de programmation dynamique rétrograde suivant :

$$\begin{cases} y_n(x) = h_n(x) \quad x \in \mathbb{R}^d \\ y_k(x) = \max\left\{h_k(x); \mathbb{E}\left[y_{k+1}(X_{k+1})|X_k = x\right]\right\}, \quad x \in \mathbb{R}^d, 0 \leq k \leq n-1. \end{cases}$$

Ainsi, ce problème de temps d'arrêt revient à résoudre le schéma par programmation dynamique rétrograde précédent.

Dans le cadre général que nous adoptons, on considère $Y_k = y_k(X_k)$ solution d'un problème de contrôle stochastique défini par un schéma par programmation dynamique descendante. On suppose que $Y_k = y_k(X_k)$ vérifie le principe de progammation dynamique générique défini par $y_k = B_k(y_{k+1})$:

$$\text{(2.2.1)} \qquad \begin{cases} y_n(x) = h(x) \\ u_{k,a}(x) = \mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(x,\xi_{k+1})|X_k = x\right] \\ v_k(x,a) = T_{k,a}\left(x, u_{k,a}(x)\right) \\ y_k(x) = R_k(x, v_k(x,a)), \quad 0 \leq k < n-1, \end{cases}$$

où les fonctions $h : \mathbb{R}^d \to \mathbb{R}$, $\phi_{k,a} : \mathbb{R}^d \times \mathbb{R}^{d_\xi} \to \mathbb{R}^{d'}$, $T_{k,a} : \mathbb{R}^d \times \mathbb{R}^{d'} \to \mathbb{R}$ et $R_k : \mathbb{R}^d \times \mathbb{R}^{A_k} \to \mathbb{R}$ sont déterministes.

REMARK 2.2.2. Interprétation : Ce cadre de programmation dynamique générique permet de résoudre de nombreux problèmes en finance et il est plus général que dans l'article [**PP04**]. Nous donnerons différentes exemples d'utilisation de cette programmation dynamique générique dans la partie suivante : applications pour les options américaines, pour les équations différentielles stochastiques rétrogrades (EDSR), applications au problème de maximisation d'utilité, d'optimisation d'un processus de Markov contrôlé, de contrôle en information partielle.

La fonction h représente la fonction valeur à l'instant final n. Les fonctions $u_{k,a}$ nécessitent des calculs des espérances conditionnelles. La fonctions $\phi_{k,a}$ pourront être par exemple un changement de probabilité (résultat d'une transformation du type Girsanov) ou représenter des accroissements d'un mouvement Brownien comme pour le schéma pour les EDSR. La fonction R_k représentera la réflexion dans le cas des options américaines ou bien la maximisation d'utilité.

2.2.2. Hypothèses. Pour des raisons techniques, on suppose dans la suite qu'il existe une constante $K > 0$ telle que $\forall x, \tilde{x} \in \mathbb{R}^d$, $\forall k \in \{0,..,n-1\}$, $\forall a \in A_k$

$$\begin{aligned} \left|T_{k,a}\left(x,u\right) - T_{k,a}\left(\tilde{x},\tilde{u}\right)\right| &\leq K\left(\left|x-\tilde{x}\right| + \left|u-\tilde{u}\right|\right), \\ \left|R_k\left(x,f\right) - R_k\left(\tilde{x},\tilde{f}\right)\right| &\leq K\left|x-\tilde{x}\right| + \left|f-\tilde{f}\right|_\infty, \\ \left|h(x) - h(\tilde{x})\right| &\leq K\left|x-\tilde{x}\right|. \end{aligned}$$

NOTATION 2.2.3. En notant la base orthonormée usuelle $e_i = (\delta_{i,j})_{1\leq j\leq d'}$ $1 \leq i \leq d'$ (où $\delta_{i,j}$ représente le symbole de Kronecker), nous définissons $\mathcal{S}^{\Delta T_k}$ sous-ensemble de $\mathbb{R}^{d'}$ par

$$\begin{aligned}\mathcal{S}^{\Delta T_k} :&= \Big\{\lambda = (\lambda_i)_{1\leq i\leq d'} \in \mathbb{R}^{d'} | \forall 1 \leq i \leq d', \exists a \in A_k, \exists x \in \mathbb{R}^d, \exists u \in \mathbb{R}^d, \exists \epsilon \in \mathbb{R}^*, \\ &= \lambda_i \in \{0; \frac{T_{k,a}(x, u + \epsilon e_i) - T_{k,a}(x, u)}{\epsilon}\}\Big\}.\end{aligned}$$

REMARK 2.2.4. 1/ Comme la fonction $T_{k,a}$ est lipschitzienne de rapport K, on a la majoration suivante $|\lambda|_\infty \leq K$, $\forall \lambda \in \mathcal{S}^{\Delta T_k}$.
2/ Si $T_{k,a} \in \mathcal{C}_b^1$, le support de $\nabla_u T_{k,a}$est inclus dans $\overline{\mathcal{S}^{\Delta T_k}}$.
3/ Nous avons une égalité du type du théorème des accroissements finis : soient $x \in \mathbb{R}^d$ et $u = (u_i)_{1\leq i\leq d'}, \tilde{u} = (\tilde{u}_i)_{1\leq i\leq d'} \in \mathbb{R}^{d'}$, il existe une constante $\lambda = (\lambda^i)_{1\leq i\leq d'} \in \mathcal{S}^{\Delta T_k}$ telle que

$$T_{k,a}(x, \tilde{u}) - T_{k,a}(x, u) = \lambda.(\tilde{u} - u).$$

Pour obtenir cette égalité, il suffit de considérer $\lambda = (\lambda_i)_{1\leq i\leq d'} \in \mathcal{S}^{\Delta T_k}$ défini par

$$\lambda_i = \begin{cases} \frac{T_{k,a}\left(x,\overline{u}^i\right) - T_{k,a}\left(x,\overline{u}^{i-1}\right)}{\tilde{u}_i - u_i} & \text{si } \tilde{u}_i \neq u_i \\ 0 & \text{si } \tilde{u}_i \neq u_i \end{cases}$$

avec $\overline{u}^i = (\tilde{u}_j 1_{j\geq i} + u_j 1_{j<i})_{1\leq j\leq d'}$.

De même, en appliquant la formule précédente pour des variables aléatoires X à valeurs dans $\mathbb{R}^d$ et $U, \tilde{U}$ à valeurs dans $\mathbb{R}^{d'}$, il existe une variable aléatoire Λ $\sigma(X, \tilde{U}, U)$-mesurable à valeurs dans $\mathcal{S}^{\Delta T_k}$ telle que

$$T_{k,a}\left(X, \tilde{U}\right) - T_{k,a}(X, U) = \Lambda.(\tilde{U} - U).$$

4/ Soit une fonction bornée $\epsilon_k : \mathbb{R}^d \to \mathbb{R}$ considérée comme une petite perturbation de y_{k+1}. En notant

$$u_{k,a}^\epsilon(x) = \mathbb{E}\left[(y_{k+1} + \epsilon_k)(X_{k+1})\phi_{k,a}(x, \xi_{k+1})|X_k = x\right],$$

on peut observer l'impact de la perturbation ϵ sur un pas de temps dans la programmation dynamique générique (2.2.1) :

$$\begin{aligned}T_{k,a}\left(x, u_{k,a}^\epsilon(x)\right) - T_{k,a}(x, u_{k,a}(x)) &= \lambda.\left(u_{k,a}^\epsilon(x) - u_{k,a}(x)\right) \\ &= \epsilon_k(x) \times \lambda.\phi_{k,a}(x, \xi_{k+1})\end{aligned}$$

avec $\lambda \in \mathcal{S}^{\Delta T_k}$. Il semble alors intéressant de majorer le terme $\lambda.\phi_{k,a}(x, \xi_{k+1})$ pour contrôler la stabilité de la programmation dynamique.

Suite à la remarque précédente, nous donnons des hypothèses liées à la stabilité $\mathbb{L}^p$ de schéma. Soit $p \in [1, +\infty]$ et $q \in [1, +\infty]$ son conjugué (i.e. $\frac{1}{p} + \frac{1}{q} = 1$).

ASSUMPTION 2.2.5. *Notons :*
◇ $(\mathcal{P}_p)$ *l'hypothèse suivante :* $\forall k \in \{0, .., n-1\}$, $\forall a \in A_k$, *on a*

$$|\mathbb{E}\left[|\lambda.\phi_{k,a}(x, \xi_{k+1})|^q\right]|^{\frac{1}{q}} \leq K, \quad \forall x \in \mathbb{R}^d, \forall \lambda \in \mathcal{S}^{\Delta T_k}.$$

◇ (H_p) *l'hypothèse suivante :* $\forall k \in \{0, .., n-1\}$, $\forall a \in A_k$, *on a*

$$\mathbb{E}\left[|\lambda.\phi_{k,a}(x, \xi_{k+1})|^q\right] \leq 1 + \frac{K}{n}, \quad \forall x \in \mathbb{R}^d, \forall \lambda \in \mathcal{S}^{\Delta T_k}.$$

2.3. Algorithmes pour l'approximation en espace

Nous allons développer trois méthodes d'approximation utilisant une grille de quantification $\Gamma_k = \left\{x_i^k \in \mathbb{R}^d\right\}_{1\leq i\leq N_k}$ de taille N_k. On considère les partitions boréliennes $\left\{\cup_{i=1}^{N_k} C_i^k\right\}_k$ avec $\bar{C}_i^k = \left\{x \in \mathbb{R}^d | \left|x - x_i^k\right| \leq \left|x - x_j^k\right|, \forall j \neq i\right\}$. On définit la fonction projection sur Γ_k par

$$\mathrm{proj}_{\Gamma_k}(x) = \sum_{i=1} x_i^k 1_{x\in C_i^k}.$$

On note $\hat{X}_k = \mathrm{proj}_{\Gamma_k}(X_k)$.

2.3.1. Régression sur une base quantifiée. Cette méthode peut être interprétée de plusieurs manières possibles :
◇ soit comme une approximation par la sous-filtration $\sigma(\hat{X}_k)$;
◇ soit comme une méthode de régression linéaire sur les vecteurs orthogonaux $\left\{1_{X_k\in C_i^k}\right\}_{1\leq i\leq N_k}$;
◇ soit comme une méthode de noyaux avec $K(x_i, X_k) = 1_{x\in C_i^k}$.

Les espérances conditionnelles $\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|X_k\right]$ sont alors estimées (via une méthode de Monte Carlo en général) par

$$\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right] = \sum_{i=1}^{N_k} \frac{\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})1_{X_k\in C_i^k}\right]}{\mathbb{P}\left(X_k \in C_i^k\right)} 1_{X_k\in C_i^k}.$$

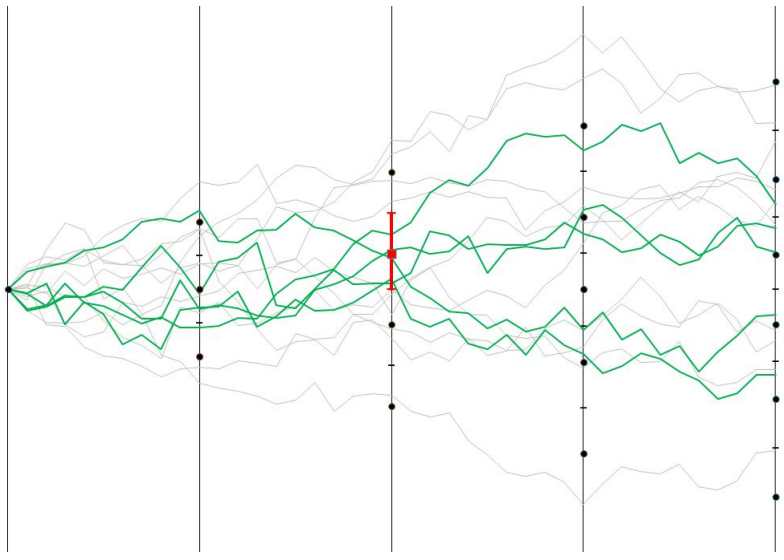


FIG. 2.3.1. Exemple d'une régression du mouvement brownien sur une base quantifiée

L'approximation en espace de la progammation dynamique générique (2.2.1) par la régression sur une base quantifié est alors déterminée par le schéma rétrograde

suivant :

$$\begin{cases} \hat{y}_n(X_n) = h(X_n) \\ \hat{v}_k(\hat{X}_k, a) = T_{k,a}\left(\hat{X}_k, \mathbb{E}\left[\hat{y}_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right]\right) \\ \hat{y}_k(\hat{X}_k) = R_k\left(\hat{X}_k, \hat{v}_k(\hat{X}_k, a)\right) \\ \hat{y}_k(X_k) = \hat{y}_k(\hat{X}_k) + r_k(\hat{X}_k, X_k), \quad \forall 0 \leq k < n, \end{cases}$$

avec $r_k(\hat{X}_k, X_k)$ un terme correcteur tentant d'approcher $y_k(X_k) - y_k(\hat{X}_k)$ que l'on peut rajouter. On suppose que

$$\left|r_k(\mathrm{proj}_{\Gamma_k}(x), x)\right| \leq C\left|\mathrm{proj}_{\Gamma_k}(x) - x\right|, \quad \forall x \in \mathbb{R}^d,$$

La proposition suivante donne la convergence de l'algorithme :

PROPOSITION 2.3.1. *Supposons $X_k \in \mathbb{L}^p$ $\forall 0 \leq k \leq n$. Il existe une constante $C \geq 0$ telle que*

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq K_p^n \sum_{k=1}^{n-1} \epsilon_k,$$

avec

$$\begin{aligned} \epsilon_k &= \left|\left|X_k - \hat{X}_k\right|\right|_p + \left|\left|\sup_{a\in A_k}\left|u_{k,a}(X_k) - \mathbb{E}\left[u_{k,a}(X_k)|\hat{X}_k\right]\right|\right|\right|_p \\ &\leq \left(1 + 2\sup_{a\in A_k}[u_{k,a}]_{Lip}\right)\left|\left|X_k - \hat{X}_k\right|\right|_p \end{aligned}$$

et avec $K_p^n \leq C^n$ sous l'hypothèse $(\mathcal{P}_p)$ et $K_p^n \leq C$ sous l'hypothèse (H_p).
Avec $\hat{X}_k$ $\mathbb{L}_p$-quantifieur optimal de X_k avec N_k points, le lemme de Pierce donne la vitesse de convergence : soit $\eta > 0$ telle que $X_k \in \mathbb{L}_{p+\eta}$, $\exists C_{d,p,\eta} > 0$, $\exists N_{d,p,\eta}$, $\forall N_k \geq N_{d,p,\eta}$

$$\left|\left|X_k - \hat{X}_k\right|\right|_p \leq C_{d,p,\eta}\,||X_k||_{p+\eta}\,N_k^{-\frac{1}{d}},$$

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq C_{d,p,\eta}K_p^n \sum_{k=1}^{n-1}\left(1 + 2\sup_{a\in A_k}[u_{k,a}]_{Lip}\right)||X_k||_{p+\eta}\,N_k^{-\frac{1}{d}}.$$

2.3.2. Algorithme par arbre de quantification. On suppose ici qu'il existe des fonctions $\psi_{k,a} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d'}$ telles que

$$\phi_{k,a}(X_k, \xi_{k+1}) = \psi_{k,a}(X_k, X_{k+1}).$$

L'approximation en espace de la progammation dynamique générique (2.2.1) par un arbre de quantification donne le schéma rétrograde suivant :

$$\begin{cases} \hat{y}_n(\hat{X}_n) = h(\hat{X}_n) \\ \hat{v}_k(\hat{X}_k, a) = T_{k,a}\left(\hat{X}_k, \mathbb{E}\left[\hat{y}_{k+1}(\hat{X}_{k+1})\psi_{k,a}(\hat{X}_k, \hat{X}_{k+1})|\hat{X}_k\right]\right) \\ \hat{y}_k(\hat{X}_k) = R_k\left(\hat{X}_k, \hat{v}_k(\hat{X}_k, a)\right), \quad \forall 0 \leq k < n. \end{cases}$$

L'application projection n'étant pas injective, le processus $\hat{X}_k$ ne conserve pas la propriété de Markov de X_k. Cependant, on définit des probabilités de transition à

partir des $\hat{X}_k$ par :

$$\begin{aligned} p_{i,j}^k &= \mathbb{P}\left(\hat{X}_{k+1} = x_j^{k+1} | \hat{X}_k = x_i^k\right) \\ &= \mathbb{P}\left(X_{k+1} \in C_j^{k+1} | X_k \in C_i^k\right). \end{aligned}$$

Alors l'algorithme devient :

$$\begin{cases} \hat{y}_n(x) = h(x), \quad \forall x \in \Gamma_n \\ \hat{u}_k(x_i^k, a) = T_{k,a}\left(x_i^k, \sum_{j=1}^{N_{k+1}} p_{i,j}^k \hat{y}_{k+1}(x_j^{k+1}) \psi_{k,a}(x_i^k, x_j^{k+1})\right) \\ \hat{y}_k(x_i^k) = P_k\left(x_i^k, \hat{u}_k(x_i^k, a)\right), \quad \forall 0 \leq k < n, \forall x_i^k \in \Gamma_k. \end{cases}$$

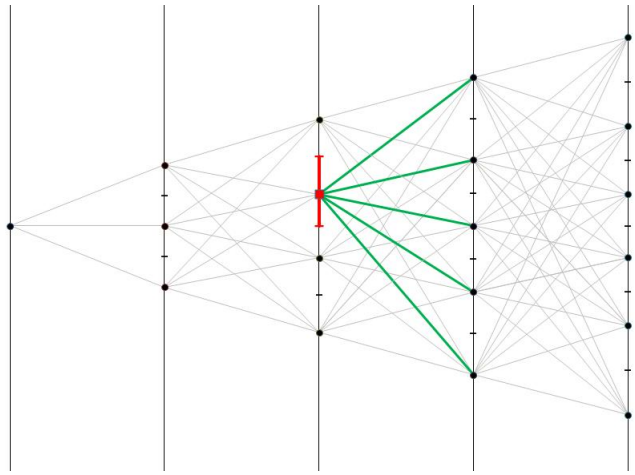


FIG. 2.3.2. Exemple d'un arbre de quantifcation approchant un mouvement brownien

La proposition suivante donne la convergence de l'algorithme :

PROPOSITION 2.3.2. *Supposons $X_k \in \mathbb{L}^p$ $\forall 0 \leq k \leq n$. Il existe une constante $C \geq 0$ telle que*

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq K_p^n \sum_{k=1}^{n-1} \epsilon_k,$$

où

$$\begin{aligned} \epsilon_k = &\left|\left|X_k - \hat{X}_k\right|\right|_p + \left|\left|\sup_{a \in A_k} \left|u_{k,a}(X_k) - \mathbb{E}\left[u_{k,a}(X_k) | \hat{X}_k\right]\right|\right|\right|_p \\ &+ \left|\left|\mathbb{E}\left[\hat{y}_{k+1}(\hat{X}_{k+1})\left(\psi_{k,a}(X_k, X_{k+1}) - \psi_{k,a}(\hat{X}_k, \hat{X}_{k+1})\right) | \hat{X}_k\right]\right|\right|_p, \end{aligned}$$

et avec $K_p^n \leq C^n$ sous l'hypothèse $(\mathcal{P}_p)$ et $K_p^n \leq C$ sous l'hypothèse (H_p).
Avec $\hat{X}_k$ $\mathbb{L}_p$-quantifieur optimal de X_k avec N_k points, le lemme de Pierce donne la vitesse de convergence : soit $\eta > 0$ tel que $X_k \in \mathbb{L}_{p+\eta}$, $\exists C_{d,p,\eta} > 0$, $\exists N_{d,p,\eta}$, $\forall N_k \geq N_{d,p,\eta}$,

$$\left|\left|X_k - \hat{X}_k\right|\right|_p \leq C_{d,p,\eta} ||X_k||_{p+\eta} N_k^{-\frac{1}{d}}$$

et

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq C_{d,p,\eta} K_p^n \sum_{k=1}^{n-1} \left(1 + 2 \sup_{a \in A_k} \left([u_{k,a}]_{Lip} + |\hat{y}|_\infty [\psi_{k,a}]_{Lip}\right)\right) ||X_k||_{p+\eta} N_k^{-\frac{1}{d}}.$$

2.3.3. Méthodes des gerbes. Ici, nous supposons que la dynamique de la $(\mathcal{F}_k)$-chaîne de Markov $(X_k)_k$ est donnée par

$$X_{k+1} = F_k(X_k, \xi_{k+1})$$

avec F_k une fonction déterministe. On approxime le semi-groupe de transition ξ_{k+1} par la variable aléatoire discrète $\hat{\xi}_{k+1}$ indépendante de $\mathcal{F}_k$. Cette discrétisation de ξ_{k+1} peut se faire par plusieurs méthodes : par exemple par Monte Carlo, par quasi-Monte Carlo, par quantification ou par une méthode de quadrature... On considère les variables aléatoires $\overline{X}_k$ et la $(\mathcal{F}_k)$-chaine de Markov $\hat{X}_k^c$ définies par

$$\begin{aligned} \hat{X}_0^c &= X_0 = \overline{X}_0 \\ \overline{X}_{k+1} &= F_k(\hat{X}_k^c, \hat{\xi}_{k+1}) \\ \hat{X}_{k+1}^c &= \mathrm{proj}_{\Gamma_{k+1}}(\overline{X}_{k+1}). \end{aligned}$$

On considère $\hat{\phi}_{k,a}$ une approximation de $\phi_{k,a}$ (cette approximation dépendra de la méthode utilisée pour $\hat{\xi}_{k+1}$).

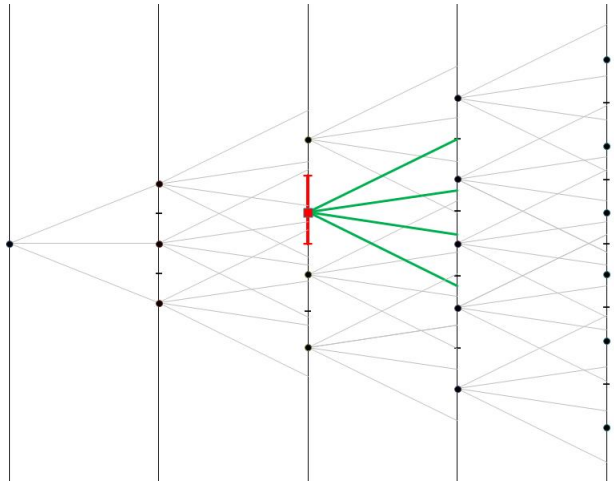


FIG. 2.3.3. Exemple d'une approximation du mouvement brownien par la méthode des gerbes

L'approximation en espace de la progammation dynamique générique (2.2.1) par la méthode des gerbres est alors déterminée par le schéma rétrograde suivant :

$$\begin{cases} \hat{y}_n(x) = h(x) \quad \forall x \in \mathbb{R}^d \\ \hat{v}_k(x_i^k, a) = T_{k,a}\left(x_i^k, \mathbb{E}\left[\hat{y}_{k+1}\left(F_k(x_i^k, \hat{\xi}_{k+1})\right)\hat{\phi}_{k,a}(x_i^k, \hat{\xi}_{k+1})\right]\right) \\ \hat{y}_k(x_i^k) = R_k\left(x_i^k, \hat{v}_k(x_i^k, a)\right) \quad \forall 0 \leq k < n, \forall x_i^k \in \Gamma_k \\ \hat{y}_k(x) = \hat{y}_k\left(\mathrm{proj}_{\Gamma_k}(x)\right) + r_k\left(\mathrm{proj}_{\Gamma_k}(x), x\right), \quad \forall 0 \leq k < n, \forall x \in \mathbb{R}^d, \end{cases}$$

avec $r_k(\text{proj}_{\Gamma_k}(x), x)$ un terme correcteur tentant d'approcher $y_k(x) - y_k(\text{proj}_{\Gamma_k}(x))$ que l'on peut rajouter. Par souci de technicité, on suppose que

$$\left|r_k(\text{proj}_{\Gamma_k}(x), x)\right| \leq C \left|\text{proj}_{\Gamma_k}(x) - x\right|.$$

Pour la méthode des gerbes, on donne des hypothèses variantes $(\mathcal{P}'_p)$ et (H'_p) au lieu de $(\mathcal{P}_p)$ et (H_p) dû à l'approximation de ξ_k par $\hat{\xi}_k$.

ASSUMPTION 2.3.3. *Pour $p \in [1, +\infty]$ et q son conjugué (i.e. $\frac{1}{p} + \frac{1}{q} = 1$), on note :*
$\diamond$ *$(\mathcal{P}'_p)$ l'hypothèse suivante : $\forall k \in \{0, .., n-1\}$, $\forall a \in A_k$,*

$$\left|\left|\lambda.\hat{\phi}_{k,a}(x, \hat{\xi}_{k+1})\right|\right|_q \leq K, \quad \forall x \in \mathbb{R}^d, \forall \lambda \in \mathcal{S}^{\Delta T_k}.$$

$\diamond$ *(H'_p) l'hypothèse suivante : $\forall k \in \{0, .., n-1\}$, $\forall a \in A_k$,*

$$\left|\left|\lambda.\hat{\phi}_{k,a}(x, \hat{\xi}_{k+1})\right|\right|_q \leq 1 + \frac{K}{n}, \quad \forall x \in \mathbb{R}^d, \forall \lambda \in \mathcal{S}^{\Delta T_k}.$$

La proposition suivante donne la convergence de l'algorithme :

PROPOSITION 2.3.4. *Supposons $X_k \in \mathbb{L}^p$ $\forall 0 \leq k \leq n$. Il existe une constante $C \geq 0$ telle que*

$$\sup_k \left|\left|y_k(\overline{X}_k) - \hat{y}_k(\overline{X}_k)\right|\right|_p \leq K_p^n \sum_{k=1}^{n-1} \epsilon_k + K_\infty^n \sum_{k=1}^{n-1} \eta_k,$$

avec

$$\begin{aligned}
\epsilon_k &= \left|\left|y_k(\overline{X}_k) - y_k(\hat{X}_k^c)\right|\right|_p + \left|\left|\overline{X}_k - \hat{X}_k^c\right|\right|_p \\
&\leq \left(1 + [y_k]_{Lip}\right) \left|\left|\overline{X}_k - \hat{X}_k^c\right|\right|_p, \\
\eta_k &= \sup_{a \in A_k, x \in \mathbb{R}^d} \left|\mathbb{E}\left[y_{k+1}\left(F_k(x, \xi_{k+1})\right)\phi_{k,a}(x, \xi_{k+1})\right] - \mathbb{E}\left[y_{k+1}\left(F_k(x, \hat{\xi}_{k+1})\right)\hat{\phi}_{k,a}(x, \hat{\xi}_{k+1})\right]\right|,
\end{aligned}$$

et avec $K_p^n \leq C^n$ sous $(\mathcal{P}'_p)$ et $K_p^n \leq C$ sous l'hypothèse (H'_p).
Avec $\hat{X}_k^c$ $\mathbb{L}_p$-quantifieur optimal de $\overline{X}_k$ avec N_k points, le lemme de Pierce donne la vitesse de convergence : soit $\eta > 0$ tel que $\overline{X}_k \in \mathbb{L}_{p+\eta}$, $\exists C_{d,p,\eta} > 0$, $\exists N_{d,p,\eta}$, $\forall N_k \geq N_{d,p,\eta}$

$$\left|\left|\overline{X}_k - \hat{X}_k^c\right|\right|_p \leq C_{d,p,\eta} \left|\left|\overline{X}_k\right|\right|_{p+\eta} N_k^{-\frac{1}{d}},$$

$$\sup_k \left|\left|y_k(\overline{X}_k) - \hat{y}_k(\overline{X}_k)\right|\right|_p \leq K_p^n C_{d,p,\eta} \sum_{k=1}^{n-1} \left(1 + [y_k]_{Lip}\right) ||X_k||_{p+\eta} N_k^{-\frac{1}{d}} + K_\infty^n \sum_{k=1}^{n-1} \eta_k.$$

Avec la grille cartésienne $\Gamma_k = \delta\mathbb{Z}^d$, on a

$$\left|\left|\overline{X}_k - \hat{X}_k^c\right|\right|_p = \left|\left|\overline{X}_k - Proj_{\Gamma_k}(\overline{X}_k)\right|\right|_p,$$

$$\sup_k \left|\left|y_k(\overline{X}_k) - \hat{y}_k(\overline{X}_k)\right|\right|_p \leq K_p^n \sum_{k=1}^{n-1} \left(1 + [y_k]_{Lip}\right) \delta + K_\infty^n \sum_{k=1}^{n-1} \eta_k.$$

Il faut maintenant avoir une majoration du terme η_k. Si $\hat{\xi}_{k+1} = \xi_{k+1}$ (et $\hat{\phi}_{k,a} = \phi_{k,a}$), alors $\eta_k = 0$. Plus généralement, sous (H_∞), K_∞^n est uniformément bornée, il faut montrer que $\eta_k = o(\frac{1}{n})$.

Nous allons proposer plusieurs méthodes pour discrétiser le semi-groupe de transition.

EXAMPLE 2.3.5. Ici, supposons que la variable aléatoire ξ_{k+1} à valeurs dans $\mathbb{R}^{d_\xi}$ est approximée par un quantificateur optimal $\hat{\xi}_{k+1}^M$ à M points. Soit $f, \phi : \mathbb{R}^{d_\xi} \to \mathbb{R}$, nous contrôlons l'approximation du semi-groupe par quantification :

$$\begin{aligned}\left|\mathbb{E}\left[f\left(\xi_{k+1}\right)\phi(\xi_{k+1})\right] - \mathbb{E}\left[f\left(\hat{\xi}_{k+1}^M\right)\phi(\hat{\xi}_{k+1}^M)\right]\right| &\leq \left|\mathbb{E}\left[\left(f(\xi_{k+1}) - f(\hat{\xi}_{k+1}^M)\right)\phi(\xi_{k+1})\right]\right| \\ &\quad + \left|\mathbb{E}\left[f(\hat{\xi}_{k+1}^M)\left(\phi(\xi_{k+1}) - \phi(\hat{\xi}_{k+1}^M)\right)\right]\right| \\ &\leq \left(|f|_{\mathrm{Lip}}\,|\phi|_\infty + |f|_\infty\,|\phi|_{\mathrm{Lip}}\right)\left|\left|\xi_{k+1} - \hat{\xi}_{k+1}^M\right|\right|_1.\end{aligned}$$

Avec $\hat{\phi}_{k,a} = \phi_{k,a} = \phi$ et $f : \xi \in \mathbb{R}^\xi \mapsto y_{k+1}\left(F_k(x,\xi)\right) \in \mathbb{R}$, on en déduit l'erreur donnée par l'approximation du semi-groupe :

$$\eta_k \leq \sup_{a\in A_k}\left\{[y_{k+1}]_{\mathrm{Lip}}\,|\nabla_\xi F_k|_\infty\,|\phi_{k,a}|_\infty + |y_{k+1}|_\infty\,|\nabla_\xi\phi_{k,a}|_\infty\right\}\left|\left|\xi_{k+1} - \hat{\xi}_{k+1}^M\right|\right|_1.$$

Le lemme de Pierce permet d'obtenir les majorations :

$$\eta_k \leq K'\,||\xi_{k+1}||_{1+\eta}\,M^{-\frac{1}{d_\xi}}$$

avec $K' = C_{d_\xi,1,\eta}\sup_{a\in A_k}\left\{[y_{k+1}]_{\mathrm{Lip}}\,|\nabla_\xi F_{k+1}|_\infty\,|\phi_{k,a}|_\infty + |y_{k+1}|_\infty\,|\nabla_\xi\phi_{k,a}|_\infty\right\}$ et

$$\sup_k\left|\left|y_k(\overline{X}_k) - \hat{y}_k(\overline{X}_k)\right|\right|_p \leq K_p^n C_{d,p,\eta}\sum_{k=1}^{n-1}\left(1 + [y_k]_{\mathrm{Lip}}\right)||X_k||_{p+\eta}\,N_k^{-\frac{1}{d}} + K_\infty^n K'\sum_{k=1}^{n-1}||\xi_{k+1}||_{1+\eta}\,M^{-\frac{1}{d_\xi}}.$$

Les espérances conditionnelles peuvent être calculées de différentes manières. Pour de nombreux problèmes, ξ_{k+1} suit une loi normale (ou bien une combinaison d'une loi normale ou d'un bruit discret (loi de Poisson...)). Par exemple, une méthode de moments (cubature), une marche aléatoire (si ξ_{k+1} est une variable gaussienne), une méthode de Monte Carlo ou la quantification de la variable ξ_{k+1} permettent d'obtenir une approximation discrète de l'espérance conditionnelle.

On souhaite que la méthode utilisée réponde aux critères suivants :

$\diamond$ Le grand nombre de calculs d'espérances conditionnelles nous incite à avoir une méthode de calcul rapide, donc avec peu de points.

$\diamond$ La méthode de discrétisation doit dépendre de la régularité de y_k ; avec une "faible" régularité, on utilisera la méthode de Monte Carlo, si la fonction y_k est lipschitzienne, on pourra approcher par quantification et si y_k est plus régulière, on pourra utiliser une méthode de quadrature.

REMARK 2.3.6. Avec $\xi \mapsto \phi_{k,a}(x,\xi)$ linéaire (par exemple pour les équations différentielles stochastiques rétrogrades) et $\hat{\xi}_{k+1}^M$ quantifieur $\mathbb{L}_2$-optimal, la formule de stationnarité permet d'obtenir l'égalité suivante :

$$\mathbb{E}\left[y_{k+1}\left(F_k(x,\hat{\xi}_{k+1})\right)\left\{\phi_{k,a}(x,\xi_{k+1}) - \phi_{k,a}(x,\hat{\xi}_{k+1})\right\}\right] = 0,$$

d'où

$$\eta_k \leq \sup_{a\in A_k}\left\{|y_{k+1}|_{\text{Lip}}\,|\nabla_\xi F_k|_\infty\,|\phi_{k,a}|_\infty\right\}\left|\left|\xi_{k+1}-\hat{\xi}^M_{k+1}\right|\right|_1.$$

Si ξ_{k+1} suit une loi normale, on peut approcher ξ_{k+1} par exemple par une loi binomiale ou trinomiale. Dans l'exemple suivant nous montrons que, sous des hypothèses de régularités des fonctions y_{k+1} et $\phi_{k,a}$ la méthode de gerbes converge en approximant ξ_{k+1} de loi normale par une loi binomiale ou trinomiale.

On note $\mathcal{C}^m_b$ l'ensemble des fonctions m-fois dérivables aux dérivées bornées.

EXAMPLE 2.3.7. 1/ Avec $\xi_{k+1}=\xi_{n,k+1}=\frac{W_{k+1}}{\sqrt{n}}$ où $\left(W^i_{k+1}\right)_{1\leq i\leq d}\sim\mathcal{N}(0,I_d)$ de loi normale, on se place dans le cas où

$$X_{k+1}=F_k\left(X_k,\xi_{n,k+1}\right).$$

Considérons l'approximation binomiale $\hat{\xi}_{n,k+1}=\frac{\hat{W}_{k+1}}{\sqrt{n}}$ définie par

$$\mathbb{P}\left(\left\{\hat{W}^i_{k+1}=\epsilon_i\right\}_{1\leq i\leq d}\right)=\frac{1}{2^d},\quad\forall\epsilon_i=\pm1,\ i=1,..,d.$$

Alors, pour $f,\phi\in\mathcal{C}^3_b$, on obtient en passant par les gradients $\nabla f,\nabla\phi$ et les hessiennes $\triangle f,\triangle\phi$ de f et de ϕ :

$$\begin{aligned}
&\mathbb{E}\left[f\left(\xi_{n,k+1}\right)\phi(\xi_{n,k+1})\right]\\
=&f(0)\phi(0)+\frac{\mathbb{E}\left[B_{k+1}\left(\triangle f(0)\phi(0)+2\nabla f(0)\nabla\phi(0)+f(0)\triangle\phi(0)\right)B^t_{k+1}\right]}{2n}+O(n^{-\frac{3}{2}})\\
=&f(0)\phi(0)+\frac{\mathbb{E}\left[\hat{B}_{k+1}\left(\triangle f(0)\phi(0)+2\nabla f(0)\nabla\phi(0)+f(0)\triangle\phi(0)\right)\hat{B}^t_{k+1}\right]}{2n}+O(n^{-\frac{3}{2}})\\
=&\mathbb{E}\left[f\left(\hat{\xi}_{n,k+1}\right)\phi(\hat{\xi}_{n,k+1})\right]+O(n^{-\frac{3}{2}}).
\end{aligned}$$

Si $y_{k+1}\in\mathcal{C}^3_b$ et $\xi\mapsto\{\phi_{k,a}(x,\xi),F_k\left(x,\xi\right)\}\in\mathcal{C}^3_b$ (avec leurs dérivées troisièmes uniformément bornées en (k,a,x)), l'erreur dûe à l'approximation du semi groupe par une loi binomiale est contrôlée et est majorée par les inégalités suivantes :

$$\eta_k\leq Cn^{-\frac{3}{2}},$$

$$\sum_{k=0}^{n-1}\eta_k\leq Cn^{-\frac{1}{2}}.$$

La vitesse de convergence convergence de l'erreur dûe à l'approximation du semi-groupe est ainsi en $O(n^{-\frac{1}{2}})$.

2/ Regardons maintenant l'approximation trinomiale.

On considère $\xi_{k+1}=\xi_{n,k+1}=\frac{W_{k+1}}{\sqrt{n}}$ avec $B_{k+1}\sim\mathcal{N}(0,I_d)$ (en dimension $d=1$). Soit $u>0$, l'approximation trinomiale $\hat{\xi}_{k+1}=\hat{\xi}_{n,k+1}=\frac{\hat{W}_{k+1}}{\sqrt{n}}$ est donnée par

$$\begin{aligned}
\mathbb{P}\left(\hat{W}_{k+1}=u\right)&=\mathbb{P}\left(\hat{W}_{k+1}=-u\right)=\frac{1}{2u^2},\\
\mathbb{P}\left(\hat{W}_{k+1}=0\right)&=1-\frac{1}{u^2}.
\end{aligned}$$

On a alors :

$$\mathbb{E}\left[(\hat{W}_{k+1})^i\right]=\mathbb{E}\left[(W_{k+1})^i\right],\quad\text{pour } i=\{0,1,2\}.$$

Avec $u=\sqrt{2}$, $\hat{W}_{k+1}$ représente une marche aléatoire à deux pas.
Avec $u=\sqrt{3}$, les cinq premiers moments sont égaux :

$$\mathbb{E}\left[(\hat{W}_{k+1})^i\right]=\mathbb{E}\left[(W_{k+1})^i\right],\quad \forall i\in\{0,1,2,3,4,5\}.$$

On pose

$$\begin{aligned}
\hat{\phi}_{n,k,a}\left(x,\frac{u}{\sqrt{n}}\right)&=\mathbb{E}\left[\frac{n\xi_{k+1}^2}{1-u^2}\phi_{k,a}(x,\xi_{k+1})\right],\\
\hat{\phi}_{n,k,a}\left(x,\frac{u}{\sqrt{n}}\right)&=\mathbb{E}\left[\left(n\xi_{k+1}^2+u\sqrt{n}\xi_{k+1}\right)\phi_{k,a}(x,\xi_{k+1})\right],\\
\hat{\phi}_{n,k,a}\left(x,-\frac{u}{\sqrt{n}}\right)&=\mathbb{E}\left[\left(n\xi_{k+1}^2-u\sqrt{n}\xi_{k+1}\right)\phi_{k,a}(x,\xi_{k+1})\right],
\end{aligned}$$

ce qui donne les égalités suivantes :

$$\mathbb{E}\left[\hat{\xi}_{n,k+1}^i\hat{\phi}_{n,k,a}\left(x,\hat{\xi}_{k+1}\right)\right]=\mathbb{E}\left[\xi_{n,k+1}^i\phi_{n,k,a}(x,\xi_{k+1})\right],\quad \text{pour } i\in\{0,1,2\}.$$

Avec $||\phi_{k,a}(x,\xi_{k+1})||_p\leq K$ et $f:\xi\mapsto y_{k+1}(F_k(x,\xi))=f(\xi)\in\mathcal{C}_b^3$, on obtient l'erreur d'approximation de $\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(x,\xi_{k+1})|X_k=x\right]$ par $\mathbb{E}\left[y_{k+1}(\overline{X}_{k+1})\hat{\phi}_{k,a}\left(x,\hat{\xi}_{k+1}\right)|X_k=x\right]$:

$$\begin{aligned}
\mathbb{E}\left[y_{k+1}(F_k(x,\xi_{k+1}))\phi_{k,a}(x,\xi_{k+1})\right]&=f(0)\mathbb{E}\left[\phi_{k,a}(x,\xi_{k+1})\right]+f'(0)\mathbb{E}\left[\xi_{k+1}\phi_{k,a}(x,\xi_{k+1})\right]\\
&\quad+\frac{f''(0)}{2}\mathbb{E}\left[\xi_{k+1}^2\phi_{k,a}(x,\xi_{k+1})\right]+O(n^{-\frac{3}{2}})\\
&=f_k(0)\mathbb{E}\left[\hat{\phi}_{k,a}\left(x,\hat{\xi}_{k+1}\right)\right]+f_k'(0)\mathbb{E}\left[\hat{\xi}_{k+1}\hat{\phi}_{k,a}\left(x,\hat{\xi}_{k+1}\right)\right]\\
&\quad+\frac{f_k''(0)}{2}\mathbb{E}\left[\hat{\xi}_{k+1}^2\hat{\phi}_{k,a}\left(x,\hat{\xi}_{k+1}\right)\right]+O(n^{-\frac{3}{2}})\\
&=\mathbb{E}\left[y_{k+1}(F_k(x,\hat{\xi}_{k+1}))\hat{\phi}_{k,a}\left(x,\hat{\xi}_{k+1}\right)\right]+O(n^{-\frac{3}{2}}).
\end{aligned}$$

On obtient finalement la majoration de l'erreur dûe à l'approximation du semi-groupe de transition par une loi trinomiale :

$$\begin{aligned}
\eta_k&\leq Cn^{-\frac{3}{2}},\\
\sum_{k=0}^{n-1}\eta_k&\leq Cn^{-\frac{1}{2}}.
\end{aligned}$$

Si $y_{k+1},F_k\in\mathcal{C}_b^3$ et$||\phi_{k,a}(x,\xi_{k+1})||_p\leq K$, la vitesse de convergence de l'erreur dûe à l'approximation du semi-groupe est ainsi en $O(n^{-\frac{1}{2}})$.

2.4. Preuves

2.4.1. Convergence du schéma par projection :

preuve de la proposition (2.3.1) : Initialement, on a $y_n(X_n)=\hat{y}_n(X_n)=h(X_n)$.

Pour $k < n$, la propriété lipschitzienne de R_k (de rapport K en x et 1 en v_k) implique :

$$\begin{aligned}\left|y_k(X_k) - \hat{y}_k(\hat{X}_k)\right| &= \left|R_k(X_k, v_k(X_k, a)) - R_k\left(\hat{X}_k, \hat{v}_k(\hat{X}_k, a)\right)\right| \\ &\leq K\left|X_k - \hat{X}_k\right| + \sup_{a\in A_k}\left|v_k(X_k, a) - \hat{v}_k(\hat{X}_k, a)\right| \\ &\leq C\left|X_k - \hat{X}_k\right| + \sup_{a\in A_k}\left|T_{k,a}\left(X_k, u_{k,a}(X_k)\right) - T_{k,a}\left(\hat{X}_k, a, \hat{u}_{k,a}(\hat{X}_k)\right)\right|,\end{aligned}$$

avec

$$\hat{u}_{k,a}(\hat{X}_k) = \mathbb{E}\left[\hat{y}_{k+1}(X_{k+1})\phi_{k,a}(x, \xi_{k+1})|\hat{X}_k\right].$$

Comme $T_{k,a}$ est lipschitzienne, on a

$$|y_k(X_k) - \hat{y}_k(X_k)| \leq C\left|X_k - \hat{X}_k\right| + C\sup_{a\in A_k}\left|u_{k,a}(X_k) - \mathbb{E}\left[u_{k,a}(X_k)|\hat{X}_k\right]\right| + \sup_{a\in A_k}|I_{k,a}|$$

où

$$I_{k,a} := T_{k,a}\left(\hat{X}_k, \mathbb{E}\left[u_{k,a}(X_k)|\hat{X}_k\right]\right) - T_{k,a}\left(\hat{X}_k, \hat{u}_{k,a}(\hat{X}_k)\right), \quad a \in A_k$$

D'après la remarque (2.2.4), il existe une variable aléatoire $\lambda_{k,a}$ $\sigma(\hat{X}_k)$-mesurable à valeurs dans $\mathcal{S}^{\Delta T_k}$ telle que

$$I_{k,a} = \lambda_{k,a}.\left(\mathbb{E}\left[u_{k,a}(X_k)|\hat{X}_k\right] - \hat{u}_{k,a}(\hat{X}_k)\right).$$

Comme $\lambda_{k,a}$ est $\sigma(\hat{X}_k)$-mesurable, on obtient

$$\begin{aligned}I_{k,a} &= \lambda_{k,a}.\left(\mathbb{E}\left[u_{k,a}(X_k)|\hat{X}_k\right] - \hat{u}_{k,a}(\hat{X}_k)\right) \\ &= \lambda_{k,a}.\left(\mathbb{E}\left[\hat{y}_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right] - \mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right]\right) \\ &= \mathbb{E}\left[\left(\hat{y}_{k+1}(X_{k+1}) - y_{k+1}(X_{k+1})\right)\lambda_{k,a}.\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right]\end{aligned}$$

L'inégalité de Hölder appliquée à $(\hat{y}_{k+1}(X_{k+1}) - y_{k+1}(X_{k+1}))$ et $\lambda_{k,a}.\phi_{k,a}(X_k, \xi_{k+1})$ donne

$$\begin{aligned}I_{k,a} &\leq \mathbb{E}\left[|y_{k+1}(X_{k+1}) - \hat{y}_{k+1}(X_{k+1})|^p\,|\hat{X}_k\right]^{\frac{1}{p}}\mathbb{E}\left[|\lambda_{k,a}.\phi_{k,a}(X_k, \xi_{k+1})|^q\,|\hat{X}_k\right]^{\frac{1}{q}} \\ &\leq K_p\mathbb{E}\left[|y_{k+1}(X_{k+1}) - \hat{y}_{k+1}(X_{k+1})|^p\,|\hat{X}_k\right]^{\frac{1}{p}},\end{aligned}$$

avec K_p borné sous $(\mathcal{P}_p)$ et $K_p \leq 1 + \frac{K}{n}$ sous l'hypothèse (H_p). L'inégalité de Jensen entraîne :

$$\begin{aligned}\left|\left|\sup_{a\in A_k} I_{k,a}\right|\right|_p &\leq \left|\left|K_p\mathbb{E}\left[|y_{k+1}(X_{k+1}) - \hat{y}_{k+1}(X_{k+1})|^p\,|\hat{X}_k\right]^{\frac{1}{p}}\right|\right|_p \\ &\leq K_p\left|\left|y_{k+1}(X_{k+1}) - \hat{y}_{k+1}(X_{k+1})\right|\right|_p\end{aligned}$$

En notant $\epsilon_k = \left|\left|X_k - \hat{X}_k\right|\right|_p + \left|\left|\sup_{a\in A_k}\left|u_{k,a}(X_k) - \mathbb{E}\left[u_{k,a}(X_k)|\hat{X}_k\right]\right|\right|\right|_p$, on a :

$$\left|\left|y_k(X_k) - \hat{y}_k(X_k)\right|\right|_p \leq \epsilon_k + K_p\left|\left|y_{k+1}(X_{k+1}) - \hat{y}_{k+1}(X_{k+1})\right|\right|_p.$$

En sommant on obtient :

$$\sup_k\left|\left|y_k(X_k) - \hat{y}_k(X_k)\right|\right|_p \leq \sum_{i=1}^{n-1} K_p^i\epsilon_i.$$

Finalement sous (H_p), on a : pour $i \leq n$, on a

$$K_p^i \leq \left(1 + \frac{K}{n}\right)^i \leq e^K$$

et par conséquent

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq e^K \sum_{i=1}^{n-1} \epsilon_i.$$

2.4.2. Convergence du schéma par arbre :
Des schémas par arbre de quantification ont été étudiés par exemple dans ([**BP03b**], [**BPP05**]) pour des problèmes de temps d'arrêts optimaux ou dans ([**PRS05**]) pour des problèmes de filtrage. Nous allons maintenant la preuve de la convergence de la méthode par arbre de quantification dans notre cadre plus général des schémas de programmation dynamique générique.

preuve de la proposition (2.3.2) : Initialement,

$$\begin{aligned} |y_n(X_n) - \hat{y}_n(X_n)| &= \left|h(X_n) - h(\hat{X}_n)\right| \\ &\leq C \left|X_n - \hat{X}_n\right|. \end{aligned}$$

Pour $k < n$, la propriété lipschitzienne de R_k implique

$$\begin{aligned} |y_k(X_k) - \hat{y}_k(X_k)| &= \left|R_k(X_k, v_k(X_k, a)) - R_k\left(\hat{X}_k, \hat{v}_k(\hat{X}_k, a)\right) - r_k(\hat{X}_k, X_k)\right| \\ &\leq K \left|X_k - \hat{X}_k\right| + \sup_{a \in A_k} \left|v_k(X_k, a) - \hat{v}_k(\hat{X}_k, a)\right|. \end{aligned}$$

En notant

$$\begin{aligned} u_{k,a}(X_k) &= \mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|X_k\right] \\ &= \mathbb{E}\left[y_{k+1}(X_{k+1})\psi_{k,a}(X_k, X_{k+1})|X_k\right], \\ \hat{u}_{k,a}(\hat{X}_k) &= \mathbb{E}\left[\hat{y}_{k+1}(\hat{X}_{k+1})\psi_{k,a}(\hat{X}_k, \hat{X}_{k+1})|\hat{X}_k\right], \end{aligned}$$

et avec $T_{k,a}$ lipschitzienne, on a

$$\begin{aligned} &\left|v_k(X_k, a) - \hat{v}_k(\hat{X}_k, a)\right| = \left|T_{k,a}\left(X_k, u_{k,a}(X_k)\right) - T_{k,a}\left(\hat{X}_k, \hat{u}_{k,a}(\hat{X}_k)\right)\right| \\ &\leq K \left|X_k - \hat{X}_k\right| + \left|T_{k,a}\left(X_k, u_{k,a}(X_k)\right) - T_{k,a}\left(X_k, \mathbb{E}\left[u_{k,a}(X_k)|\hat{X}_k\right]\right)\right| + |I_{k,a}| \\ &+ K \left|T_{k,a}\left(\hat{X}_k, \mathbb{E}\left[\hat{y}_{k+1}(\hat{X}_{k+1})\psi_{k,a}(X_k, X_{k+1})|\hat{X}_k\right]\right) - T_{k,a}\left(\hat{X}_k, \mathbb{E}\left[\hat{y}_{k+1}(\hat{X}_{k+1})\psi_{k,a}(\hat{X}_k, \hat{X}_{k+1})|\hat{X}_k\right]\right)\right| \\ &\leq K \left|X_k - \hat{X}_k\right| + K \left|\mathbb{E}\left[\hat{y}_{k+1}(\hat{X}_{k+1})\left(\psi_{k,a}(X_k, X_{k+1}) - \psi_{k,a}(\hat{X}_k, \hat{X}_{k+1})\right)|\hat{X}_k\right]\right| + |I_{k,a}| \end{aligned}$$

où

$$I_{k,a} = T_{k,a}\left(\hat{X}_k, \mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right]\right) - T_{k,a}\left(\hat{X}_k, \mathbb{E}\left[\hat{y}_{k+1}(\hat{X}_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right]\right).$$

Comme dans la preuve de la proposition *(2.3.1)*, il existe une variable aléatoire $\lambda_{k,a}$ $\sigma(\hat{X}_k)$-mesurable à valeurs dans $\mathcal{S}^{\Delta T_k}$ telle que

$$\begin{aligned} I_{k,a} &= \lambda_{,a}.\left(\mathbb{E}\left[y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right] - \mathbb{E}\left[\hat{y}_{k+1}(\hat{X}_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right]\right) \\ &= \mathbb{E}\left[\left\{y_{k+1}(X_{k+1}) - \hat{y}_{k+1}(\hat{X}_{k+1})\right\}\lambda_{k,a}.\phi_{k,a}(X_k, \xi_{k+1})|\hat{X}_k\right]. \end{aligned}$$

D'après l'inégalité de Hölder, on a

$$|I_{k,a}| \leq \mathbb{E}\left[\left|y_{k+1}(X_{k+1}) - \hat{y}_{k+1}(\hat{X}_{k+1})\right|^p |\hat{X}_k\right]^{\frac{1}{p}} \mathbb{E}\left[|\lambda_{k,a}.\phi_{k,a}(X_k, \xi_{k+1})|^q |\hat{X}_k\right]^{\frac{1}{q}}$$
$$\leq K_p \mathbb{E}\left[\left|y_{k+1}(X_{k+1}) - \hat{y}_{k+1}(\hat{X}_{k+1})\right|^p |\hat{X}_k\right]^{\frac{1}{p}}$$

avec $K_p \geq 1$ borné sous $(\mathcal{P}_p)$ et $K_p \leq 1 + \frac{K}{n}$ sous (H_p).

En notant

$$\epsilon_k = \left|\left|X_k - \hat{X}_k\right|\right|_p + \left|\left|\sup_{a\in A_k} \left|u_{k,a}(X_k) - \mathbb{E}\left[u_{k,a}(X_k)|\hat{X}_k\right]\right|\right|\right|_p$$
$$+ \left|\left|\sup_{a\in A_k} |\mathbb{E}\left[\hat{y}_{k+1}(\hat{X}_{k+1})\left(\psi_{k,a}(X_k, X_{k+1}) - \psi_{k,a}(\hat{X}_k, \hat{X}_{k+1})\right)|\hat{X}_k\right]|\right|\right|_p,$$

on obtient à l'aide de l'inégalité de Jensen :

$$||y_k(X_k) - \hat{y}_k(X_k)||_p \leq C\epsilon_k + K_p \left|\left|\mathbb{E}\left[\left|y_{k+1}(X_{k+1}) - \hat{y}_{k+1}(\hat{X}_{k+1})\right|^p |\hat{X}_k\right]^{\frac{1}{p}}\right|\right|_p$$
$$\leq C\epsilon_k + K_p \left|\left|y_{k+1}(X_{k+1}) - \hat{y}_{k+1}(\hat{X}_{k+1})\right|\right|_p$$

avec $K_p \geq 1$ borné sous $(\mathcal{P}_p)$ et $K_p \leq 1 + \frac{K}{n}$ sous (H_p). En sommant, on obtient :

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq \sum_{i=1}^n K_p^i \epsilon_i.$$

Pour $i \leq n$, on a $\left(1 + \frac{K}{n}\right)^i \leq e^K$. Finalement sous (H_p), on a :

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq e^K \sum_{i=1}^{n-1} \epsilon_i.$$

2.4.3. Convergence de la méthode des gerbes

. Des schémas par méthode de gerbes ont été étudiés. Dans ([**PP04**], [**BPP05**]), une grille de quantification est utilisée pour la discrétisation en espace (pour la variable X_k) et le semi-groupe de transition est exact (il n'est pas approché par une loi discrète). Les articles ([**DM06**], [**DM08**]) s'intéresse aux équations différentielles stochastiques progressives rétrogrades ; la discrétisation en espace (pour la variable X_k) utilise une grille cartésiennes $\epsilon\mathbb{Z}^d$, une quantification optimale de la loi normale donne l'approximation du semi-groupe de transition. Nous allons maintenant la preuve de la convergence de la méthode des gerbes dans notre cadre plus général des schémas de programmation dynamique générique.

Preuve de la proposition (2.3.4) : Initialement, on a $y_n(\overline{X}_n) = \hat{y}_n(\overline{X}_n) = h(\overline{X}_n)$.

Pour $k < n$, la propriété lipschitzienne de R_k (de rapport K en x et 1 en v_k) implique :

$$\left|y_k(\overline{X}_k) - \hat{y}_k(\overline{X}_k)\right| \leq \left|y_k(\overline{X}_k) - y_k(\hat{X}_k^c)\right| + \left|\hat{y}_k(\overline{X}_k) - \hat{y}_k(\hat{X}_k^c)\right| + \left|y_k(\hat{X}_k^c) - \hat{y}_k(\hat{X}_k^c)\right|$$
$$\leq \left|y_k(\overline{X}_k) - y_k(\hat{X}_k^c)\right| + \left|r_k\left(\hat{X}_k^c, \overline{X}_k\right)\right| + \left|R_k(\hat{X}_k^c, v_k(\hat{X}_k^c, a)) - R_k\left(\hat{X}_k, \hat{v}_k(\hat{X}_k^c, a)\right)\right|$$
$$\leq \left|y_k(\overline{X}_k) - y_k(\hat{X}_k^c)\right| + K\left|\overline{X}_k - \hat{X}_k^c\right| + \sup_{a_k\in A_k} \left|v_k(\hat{X}_k^c, a_k) - \hat{v}_k(\hat{X}_k^c, a_k)\right|.$$

Pour $a \in A_k$, la propriété de lipschitzianité de $T_{k,a}$ donne

$$\begin{aligned}
&\left|v_k(\hat{X}_k^c, a) - \hat{v}_k(\hat{X}_k^c, a)\right| \\
=&\left|T_{k,a}\left(\hat{X}_k^c, \mathbb{E}\left[y_{k+1}\left(F_k(\hat{X}_k^c, \xi_{k+1})\right)\phi_{k,a}(\hat{X}_k^c, \xi_{k+1})\right]\right) - T_{k,a}\left(\hat{X}_k^c, \mathbb{E}\left[\hat{y}_{k+1}\left(F_k(\hat{X}_k^c, \hat{\xi}_{k+1})\right)\hat{\phi}_{k,a}(\hat{X}_k^c, \hat{\xi}_{k+1})\right]\right)\right| \\
\leq&\left|T_{k,a}\left(\hat{X}_k^c, \mathbb{E}\left[y_{k+1}\left(F_k(\hat{X}_k^c, \xi_{k+1})\right)\phi_{k,a}(\hat{X}_k^c, \xi_{k+1})\right]\right) - T_{k,a}\left(\hat{X}_k^c, \mathbb{E}\left[y_{k+1}\left(F_k(\hat{X}_k^c, \hat{\xi}_{k+1})\right)\hat{\phi}_{k,a}(\hat{X}_k^c, \hat{\xi}_{k+1})\right]\right)\right| \\
&+ |I_{k,a}| \\
\leq& K\left|\mathbb{E}\left[y_{k+1}\left(F_k(x, \xi_{k+1})\right)\phi_{k,a}(x, \xi_{k+1})\right] - \mathbb{E}\left[y_{k+1}\left(F_k(x, \hat{\xi}_{k+1})\right)\hat{\phi}_{k,a}(x, \hat{\xi}_{k+1})\right]\right| + |I_{k,a}|
\end{aligned}$$

où

$$I_{k,a} = T_{k,a}\left(x, \mathbb{E}\left[y_{k+1}\left(F_k(x, \hat{\xi}_{k+1})\right)\hat{\phi}_{k,a}(x, \hat{\xi}_{k+1})\right]\right) - T_{k,a}\left(x, \mathbb{E}\left[\hat{y}_{k+1}\left(F_k(x, \hat{\xi}_{k+1})\right)\hat{\phi}_{k,a}(x, \hat{\xi}_{k+1})\right]\right).$$

Comme dans la preuve de la proposition *(2.3.1)*, il existe une constante $\lambda_{k,a}$ $\sigma(\hat{X}_k)$-mesurable à valeurs dans $\mathcal{S}^{\Delta T_k}$ telle que

$$\begin{aligned}
I_{k,a} &= \lambda_{k,a}.\left(\mathbb{E}\left[y_{k+1}\left(F_k(\hat{X}_k, \hat{\xi}_{k+1})\right)\hat{\phi}_{k,a}(\hat{X}_k, \hat{\xi}_{k+1})\right] - \mathbb{E}\left[\hat{y}_{k+1}\left(F_k(\hat{X}_k, \hat{\xi}_{k+1})\right)\hat{\phi}_{k,a}(\hat{X}_k, \hat{\xi}_{k+1})\right]\right) \\
&= \mathbb{E}\left[\left\{y_{k+1}\left(F_k(\hat{X}_k, \hat{\xi}_{k+1})\right) - \hat{y}_{k+1}\left(F_k(\hat{X}_k, \hat{\xi}_{k+1})\right)\right\}\lambda_{k,a}.\hat{\phi}_{k,a}(\hat{X}_k, \hat{\xi}_{k+1})|\hat{X}_k\right].
\end{aligned}$$

En notant

$$\eta_k = \sup_{a\in A_k, x\in\mathbb{R}^d}\left|\mathbb{E}\left[y_{k+1}\left(F_k(x, \xi_{k+1})\right)\phi_{k,a}(x, \xi_{k+1})\right] - \mathbb{E}\left[y_{k+1}\left(F_k(x, \hat{\xi}_{k+1})\right)\hat{\phi}_{k,a}(x, \hat{\xi}_{k+1})\right]\right|,$$

on a :

$$\begin{aligned}
\left|y_k(\overline{X}_k) - \hat{y}_k(\overline{X}_k)\right| \leq& \left|y_k(\overline{X}_k) - y_k(\hat{X}_k^c)\right| + K\left|\overline{X}_k - \hat{X}_k^c\right| + K\eta_k \\
&+ \sup_{a\in A_k}\left|\mathbb{E}\left[\left\{y_{k+1}(\overline{X}_{k+1}) - \hat{y}_{k+1}(\overline{X}_{k+1})\right\}\lambda_{k,a}.\hat{\phi}_{k,a}(\hat{X}_k^c, \hat{\xi}_{k+1})|\hat{X}_k^c\right]\right|.
\end{aligned}$$

En sommant (et comme $y_n = \hat{y}_n$), on obtient :

$$\left|y_k(\overline{X}_k) - \hat{y}_k(\overline{X}_k)\right| \leq \sum_{k'=k}^{n-1}\sup_{a\in A}\mathbb{E}\left[\left\{\left|y_{k'}(\overline{X}_{k'}) - y_{k'}(\hat{X}_{k'}^c)\right| + K\left|\overline{X}_{k'} - \hat{X}_{k'}^c\right| + K\eta_k\right\}\Lambda_{k,k'}(a)|\hat{X}_k^c\right]$$

avec $\Lambda_{k,k'}(a) = \Pi_{i=k}^{k'-1}\left|\lambda_{k,a_i}.\hat{\phi}_{i,a_i}(\hat{X}_i^c, \hat{\xi}_{i+1})\right|$ (et $\Lambda_{k,k} = 1$) pour $a = (a_0, .., a_{n-1}) \in A = A_0 \times .. \times A_{n-1}$.

L'inégalité de Hölder avec $\frac{1}{p} + \frac{1}{q} = 1$ donne :

$$\begin{aligned}
\left|y_k(\overline{X}_k) - \hat{y}_k(\overline{X}_k)\right| \leq& \sum_{k'=k}^{n-1}\left\{\mathbb{E}\left[\left|y_{k'}(\overline{X}_{k'}) - y_{k'}(\hat{X}_{k'}^c)\right|^p|\overline{X}_k\right]^{\frac{1}{p}}\sup_{a\in A}\mathbb{E}\left[|\Lambda_{k,k'}(a)|^q|\overline{X}_k\right]^{\frac{1}{q}}\right. \\
&\left. + K\mathbb{E}\left[\left|\overline{X}_k - \hat{X}_k^c\right|^p|\overline{X}_k\right]^{\frac{1}{p}}\mathbb{E}\left[|\Lambda_{k,k'}(a)|^q|\overline{X}_k\right]^{\frac{1}{q}} + K\eta_k\sup_{a\in A}\mathbb{E}\left[\Lambda_{k,k'}(a)|\overline{X}_k\right]\right\}.
\end{aligned} \tag{2.4.1}$$

Or,

$$\mathbb{E}\left[\left|\lambda_{i,a_i}.\hat{\phi}_i(\hat{X}_i^c, a_i, \hat{\xi}_{i+1})\right|^q|\hat{X}_i^c\right]^{\frac{1}{q}} \leq K_p$$

avec $K_p = C$ sous $(\mathcal{P}'_p)$ et $K_p = 1 + \frac{C}{n}$ sous (H'_p). Comme $\hat{X}^c_k$ est une chaîne de Markov, on a pour $k < k'$

$$\begin{aligned}\mathbb{E}\left[\Lambda_{k,k'}(a)^q|\hat{X}^c_k\right] &= \mathbb{E}\left[\Lambda_{k,k'-1}(a)^q\mathbb{E}\left[\left|\lambda_{k'-1,a_{k'-1}}.\hat{\phi}_{k'-1,a_{k'-1}}(\hat{X}^c_{k'-1},\hat{\xi}_{k'})\right|^q|\hat{X}^c_{k'-1}\right]|\hat{X}^c_k\right] \\ &\leq K_p^q\mathbb{E}\left[\Lambda_{k,k'-1}(a)^q|\hat{X}^c_k\right].\end{aligned}$$

Par une récurrence descendante, on obtient :

$$\text{esssup}_{a\in A}\mathbb{E}\left[|\Lambda_{k,k'}(a)|^q|\overline{X}_k\right]^{\frac{1}{q}} \leq K_p^n, \quad \forall 0 \leq k \leq k' \leq n$$

avec $K_p^n = C^n$ sous $(\mathcal{P}'_p)$ et $K_p^n = e^C$ sous (H'_p) et pour $p = +\infty$ ($q = 1$ pour le terme en η_k) :

$$\text{esssup}_{a\in A}\mathbb{E}\left[|\Lambda_{k,k'}(a)||\overline{X}_k\right] \leq K_{\infty,}^n \quad \forall 0 \leq k \leq k' \leq n.$$

Le passage de l'inégalité (2.4.1) à la norme $\mathbb{L}_p$ complète la preuve :

$$\begin{aligned}\left|\left|y_k(\overline{X}_k) - \hat{y}_k(\overline{X}_k)\right|\right|_p \leq K_p^n\sum_{k'=k}^{n-1}&\left\{\left|\left|y_{k'}(\overline{X}_{k'}) - y_{k'}(\hat{X}^c_{k'})\right|\right|_p + K\left|\left|\overline{X}_{k'} - \hat{X}^c_{k'}\right|\right|_p\right\} \\ &+ KK_\infty^n\sum_{k'=k}^{n-1}\eta_{k'}.\end{aligned}$$

2.5. Reconstruction de la solution

Dans cette section on précise le terme r_k. Le but est de reconstruire $\hat{y}_k$ sur $\mathbb{R}^d$ pour améliorer la convergence.

Sur la grille $\Gamma_k = \left\{x^k_1, .., x^k_{N_k}\right\}$, nous approchons $y_k(x^k_i) \simeq \mathbb{E}\left[y_k(X_k)|\hat{X}_k = x^k_i\right]$ par $\hat{y}_k(x^k_i)$. Pour $x \in \mathbb{R}^d$ avec $\text{proj}_{\Gamma_k}(x) = x^k_i$, on peut approcher $y_k(x)$ par la méthode du plus proche voisin, donc par $\hat{y}_k(x^k_i)$. Cependant, il est possible d'avoir une meilleure approximation en ajoutant un terme $r_k(x^k_i, x) \simeq y_k(x) - y_k(x^k_i)$, d'où $y_k(x) \simeq \hat{y}_k(x^k_i) + r_k(x^k_i, x)$.

Nous allons proposer plusieurs méthodes pour obtenir le terme r_k.

2.5.1. A l'aide de la grille Γ_k. Les informations données par $\left(x^k_i, \hat{y}_k(x^k_i)\right)_{1\leq i\leq N_k}$ peuvent permettre de reconstruire $\hat{y}_k$.

2.5.1.1. *Par régression linéaire.* Soit $\left\{x^k_{i_i}, .., x^k_{i_p}\right\}$ les p plus proches voisins de x^k_i dans $\Gamma_k \setminus \left\{x^k_i\right\}$ et soient $b = \{b_j\}_{1\leq j\leq m}$ un ensemble de fonctions. On fait une regression linéaire avec les plus proches voisins sur b :

$$\alpha_{k,i} = \text{argmin}_{\alpha\in\mathbb{R}^m}\sum_{j=1}^{p}\left|\hat{y}_k(x^k_i) - \hat{y}_k(x^k_{i_j}) - \alpha.b(x^k_i - x^k_{i_j})\right|^2.$$

On choisit alors r_k par :

$$r_k(x^k_i, x) = \sum_{j=1}^{p}\alpha_{k,i}.b(x - x^k_{i_j}).$$

2.5.1.2. *Par interpolation.* Considérons une grille $\{\Gamma_k\}_{0\leq k<n}$. Pour une fonction $f : \Gamma_k \to \mathbb{R}$, on définit $f^{\Gamma_k} : \mathbb{R}^d \to \mathbb{R}$ une interpolation linéaire (qui conserve le rapport de Lipschitz de f).

Deux types d'interpolation vont tout particulièrement nous intéresser. Si la grille est quelconque, on utilise l'interpolation linéaire à partir de la triangulation de Delaunay. Cependant, le nombre de simplexes de cette triangulation augmente géométriquement avec la dimension d et par conséquent le temps de calcul augmente fortement. Pour palier ce problème, il est possible dans certains cas de calculer hors-ligne les différentes espérances conditionnelles. Dans ce cas-là, il serait sans doute préférable d'obtenir une quantification duale optimale ; la grille serait alors mieux adaptée à notre schéma par interpolation (voir [**PW10**]).

Si la grille est cartésienne, i.e. $\Gamma_k = \delta\mathbb{Z}^d$, nous proposons une approximation multilineaire par morceaux. On définit les fonctions

$$\phi_z(x) = \prod_{i=1}^{d} \Phi\left(\delta^{-1}(x_i - z_i)\right), \quad \forall z \in \delta\mathbb{Z}^d, \forall x \in \mathbb{R}^d,$$

avec

$$\Phi(t) = (1 - |t|)^+ .$$

La fonction ϕ_z est à valeurs dans $[0, 1]$ et

$$\begin{aligned} \phi_z(z) &= z, \\ \phi_z(x) &= 0, \quad \text{si } |x - z|_\infty \geq \delta, \\ \sum_{z\in\delta\mathbb{Z}^d} \phi_z(x) &= 1, \\ \sum_{z\in\delta\mathbb{Z}^d} \phi_z(x) z &= x. \end{aligned}$$

L'interpolation sur une grille cartésienne sera définie par

$$f^{\delta\mathbb{Z}^d}(x) = \sum_{z\in\delta\mathbb{Z}^d} \phi_z(x) f(z).$$

Avec cette interpolation, on prendra :

$$r_k(\mathrm{Proj}_{\delta\mathbb{Z}^d}(x), x) = \sum_{z\in\delta\mathbb{Z}^d} \phi_z(x)\hat{y}_k(z) - \hat{y}_k(\mathrm{Proj}_{\delta\mathbb{Z}^d}(x)).$$

2.5.1.3. *Estimations par noyaux.* Soit $K : \mathbb{R}^d \to \mathbb{R}^+$ tel que pour $v \neq 0$, $f : t \in \mathbb{R}^+ \to K(tv)$ est décroissante et $f(t) \to_{t\to+\infty} 0$. Alors l'approximation devient

$$r_k(\hat{X}_k, X_{k+1}) = \frac{\sum_{i\in V_k(X_k)} K(X_{k+1} - x_i^k)\hat{y}_k(x_j^k)}{\sum_{i\in V_k(X_k)} K(x - x_i^k)} - \hat{y}_k(\hat{X}_k).$$

avec $V_k(X_k)$ un sous ensemble de $\{1, ..., N_k\}$ par exemple les indices des p plus proches de de X_k ou bien de $\hat{X}_k$.

2.5.2. Méthodes probabilistes. Ici, on suppose que l'on a une approximation des dérivées de y_k par une méthode probabiliste (par exemple par une represention probabiliste par du calcul de Malliavin). Par exemple, pour les EDSR, on a

$\partial_x y_k(\hat{X}_k)\sigma(\hat{X}_k) \simeq \hat{Z}_k = \mathbb{E}\left[\hat{y}_{k+1}(X_{k+1})\frac{\Delta W_{k+1}}{\Delta t_{k+1}}|\hat{X}_k\right]$. Alors on peut prendre r_k sous la forme :

$$r_k(\hat{X}_k, X_{k+1}) = \frac{\hat{Z}_k}{\sigma(\hat{X}_k)}(X_k - \hat{X}_k).$$

2.5.3. Remarques. Les méthodes par interpolation ou par noyaux donnent une reconstruction de $\hat{y}_k$ assez régulière ($\hat{y}_k$ peut conserver le même rapport de Lispchitz par la méthode par interpolation dans certain cas). Cependant, le temps de calcul est assez long pour l'interpolation linéaire sur les diagrammes de Delaunay en dimension assez grande (par exemple supérieure à 4, 5). L'estimation par noyaux nécessite de calculer les p-plus proches voisins de X_k.

Une estimation sur des fonctions locales (sur C_i^k) par une régression linéaire ou par une méthode probabiliste permet un calcul rapide de $r_k(\mathrm{Proj}_{\Gamma_k}(x), x)$.

Quitte à remplacer $r_k(\mathrm{Proj}_{\Gamma_k}(x), x)$ par $\mathrm{Proj}_{B\left(0, C\left|x - \mathrm{Proj}_{\Gamma_k}(x)\right|\right)}\left(r_k(\mathrm{Proj}_{\Gamma_k}(x), x)\right)$, on peut supposer pour des besoins techniques (de convergence de l'algorithme) que

$$\left|r_k(\mathrm{Proj}_{\Gamma_k}(x), x)\right| \leq C\left|x - \mathrm{Proj}_{\Gamma_k}(x)\right|.$$

CHAPITRE 3

Exemples d'applications :

3.1. Introduction

Pour les différents schémas, nous pouvons décomposer les différentes erreurs d'estimations sous la forme :

$$\varepsilon(\text{schéma}) \leq \varepsilon(\text{temps}) + \varepsilon(\text{quantif}) + \varepsilon_{\text{arbre}}(\phi) + \varepsilon_{\text{gerbes}}(\xi),$$

où :
◇ ε(temps) représente l'erreur d'approximation en temps : par exemple lorsque que l'on fait un schéma d'Euler ou bien que l'on approxime une option américaine par une option bermudéenne (voir ([**BP03a**])).
◇ ε(quantif) représente l'erreur donnée par l'approximation par quantification et est de la forme

$$\begin{aligned}\varepsilon(\text{quantif}) &= C_n \sum_{k=0}^{n-1} \left|\left| X_k - \hat{X}_k \right|\right|_p \\ &\leq C'_n \sum_{jk=0}^{n-1} N_k^{-\frac{1}{d}}.\end{aligned}$$

◇ $\varepsilon_{\text{arbre}}(\phi)$ représente l'erreur d'approximation de $\psi_{k,a}(X_k, X_{k+1})$ par $\psi_{k,a}(\hat{X}_k, \hat{X}_{k+1})$ dans l'algorithme par arbre.
◇ $\varepsilon_{\text{gerbes}}(\xi)$ représente l'erreur d'approximation de ξ_k par $\hat{\xi}_k$ dans la méthode des gerbes.

On va donner quelques applications utiles en finance. Le problème de discrétisation en temps est supposé déjà résolu.

Rappellons tout d'abord les vitesses de convergence pour les algorithmes de quantification sous les hypothèses $(\mathcal{P}_p)$:
◇ pour la méthode de régression :

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq C_{d,p,\eta} K_p^n \sum_{k=1}^{n-1} \left(1 + 2 \sup_{a \in A_k} [u_{k,a}]_{\text{Lip}}\right) ||X_k||_{p+\eta} N_k^{-\frac{1}{d}}$$

◇ pour la méthode par arbre :

$$\sup_k ||y_k(X_k) - \hat{y}_k(X_k)||_p \leq C_{d,p,\eta} K_p^n \sum_{k=1}^{n-1} \left(1 + 2 \sup_{a \in A_k} \left([u_{k,a}]_{\text{Lip}} + |\hat{y}|_\infty [\psi_{k,a}]_{\text{Lip}}\right)\right) ||X_k||_{p+\eta} N_k^{-\frac{1}{d}}$$

◇ pour la méthode des gerbes :

$$\sup_k \left|\left|y_k(\overline{X}_k) - \hat{y}_k(\overline{X}_k)\right|\right|_p \leq K_p^n C_{d,p,\eta} \sum_{k=1}^{n-1} \left(1 + [y_k]_{\text{Lip}}\right) ||X_k||_{p+\eta} N_k^{-\frac{1}{d}} + K_\infty^n \sum_{k=1}^{n-1} \eta_k$$

Pour obtenir la convergence en espace par l'un des algorithmes par quantification (par projection, par arbre et par gerbe), il suffit de prouver que l'hypothèse (H_p) ou $(\mathcal{P}_p)$ est satisfaite et que les fonctions y_k sont uniformément lipschitziennes (avec une constante indépendante de k et n).

Des arguments sur la diffusion de la chaîne de Markov X sont nécessaires pour obtenir de la régularité des fonctions y_k et $u_{k,a}$. Ainsi nous allons nous focaliser sur des propriétés lipschitziennes, soit du semi-goupe, soit du flot de X.

Les propriétés lipschitziennes du semi-groupe ont été explorées dans la littérature pour la convergence de schémas par quantification optimale (voir par exemple **[BP03a]**, **[PP04]**, **[PP05a]**).

Dans cette partie, nous allons proposer une nouvelle approche directe par régularité du flot. Cela permet d'obtenir des vitesses de convergence pour les schémas par quantification, notamment pour des options exotiques inaccessibles jusqu'alors.

NOTATION 3.1.1. Pour un chaîne de Markov X, on notera $X_{k'}^{k,x}$ la variable aléatoire $X_{k'}$ sachant $X_k = x$ (pour $k' \geq k$).

Pour $k \in \{0, .., n-1\}$, on note P_k^X la probabilité de transition de la chaîne de Markov X à l'instant k :

$$P_k^X f(x) := \mathbb{E}\left[f(X_{k+1})|X_k = x\right] = \mathbb{E}\left[f(X_{k+1}^{k,x})\right],$$

pour toute fonction f borélienne bornée.

DEFINITION 3.1.2. $\diamond$ La probabilité de transition P_k^X est lipschitzienne de rapport $\left[P_k^X\right]_{\text{Lip}}$ si pour toute fonction lipschitzienne f de $\mathbb{R}^d$ de rapport $[f]_{\text{Lip}}$, on a :

$$\left|P_k^X f(x) - P_k^X f(\tilde{x})\right| \leq \left[P_k^X\right]_{\text{Lip}} [f]_{\text{Lip}} |x - \tilde{x}|, \quad \forall x, \tilde{x} \in \mathbb{R}^d.$$

Alors on note $\left[P^X\right]_{\text{Lip}} = \max_{k\in\{0,..,n-1\}} \left[P_k^X\right]_{\text{Lip}}$.

$\diamond$ Pour $p \geq 1$, on dira que X possède un flot $\mathbb{L}_p$-lipschitzien de rapport $\left[F^X\right]_{\text{Lip}}$ si

$$\left|\left|\sup_{k\leq k'\leq n} \left|X_{k'}^{k,x} - X_{k'}^{k,\tilde{x}}\right|\right|\right|_p \leq \left[F^X\right]_p |x - \tilde{x}|, \quad \forall x, \tilde{x} \in \mathbb{R}^d, \forall 0 \leq k \leq n.$$

On note $\left[F_k^X\right]_p$ le rapport de Lipschitz pour une transition entre k et $k+1$:

$$\left|\left|X_{k+1}^{k,x} - X_{k+1}^{k,\tilde{x}}\right|\right|_p \leq \left[F_k^X\right]_p |x - \tilde{x}|, \quad \forall x, \tilde{x} \in \mathbb{R}^d, \forall 0 \leq k \leq n-1.$$

REMARK 3.1.3. 1/ D'après l'inégalité de Hölder, on a pour $p' \geq p \geq 1$

$$\left[F_k^X\right]_{p'} \geq \left[F_k^X\right]_p \quad \text{et} \quad \left[F^X\right]_{p'} \geq \left[F^X\right]_p$$

2/ Un flot lipschitzien implique une probabilité de transition lipschitizienne de rapport inférieur :

$$\left[P_k^X\right]_{\text{Lip}} \leq \left[F_k^X\right]_p.$$

En effet, considérons une fonction lipschitzienne f de $\mathbb{R}^d$, on a :

$$\begin{aligned}\left|P_k^X f(x) - P_k^X f(\tilde{x})\right| &= \left|\mathbb{E}\left[f(X_{k+1}^{k,x}) - f(X_{k+1}^{k,\tilde{x}})\right]\right| \\ &\leq [f]_{\text{Lip}} \left|\left|X_{k+1}^{k,x} - X_{k+1}^{k,\tilde{x}}\right|\right|_1 \\ &\leq [f]_{\text{Lip}} \left[F_k^X\right]_p |x - \tilde{x}|.\end{aligned}$$

Dans de nombreux cas, la chaîne de Markov $\{X_k\}_{0\leq k\leq n}$ est une discrétisation d'un processus Markovien $\{S_t\}_{0\leq t\leq T}$ sur un intervalle $[0,T]$: $X_k = S_{t_k}$ avec $t_0 = 0 < t_1 < ... < t_n = T$. Dans ce cas, on souhaite obtenir des rapports de Lipschitz du flot et des probabilités de transition de X ne dépendant pas du nombre de pas de temps n.

REMARK 3.1.4. 1/ En regardant entre deux instants k et $k+i$, le rapport de Lipschitz est contrôlé par le rapport $\left[F^X\right]_p$ si le flot de X est lipschitzien et par le rapport géométrique $\left[P^X\right]^i_{\text{Lip}}$ si sa probabilité de transition est lipschitzienne :

$$\begin{aligned}\left|\mathbb{E}\left[f(X^{k,x}_{k+i})\right] - \mathbb{E}\left[f(X^{k,\tilde{x}}_{k+i})\right]\right| &\leq \left[P^X\right]^i_{\text{Lip}} [f]_{\text{Lip}} |x-\tilde{x}| \\ &\leq \left[F^X\right]_p |x-\tilde{x}|, \quad \forall x,\tilde{x}\in\mathbb{R}^d, \forall 0\leq k\leq k+i\leq n.\end{aligned}$$

Pour contrôler indépendamment du nombre de temps n, il suffit par exemple que les probabilités de transition vérifient $\left[P^X\right]_{\text{Lip}} \leq 1+\frac{C}{n}$ (avec C indépendant de n) ; on a alors $\left[P^X\right]^n_{\text{Lip}} \leq e^C$.

2/ Si $\left[F^X_k\right]_p \leq 1+\frac{C}{n}$, alors les inégalités

$$\left[F^X\right]_p \leq \left[F^X_k\right]^n_p \leq \left(1+\frac{C}{n}\right)^n \leq e^C$$

montrent que le flot de X est lipchitzien avec un rapport de Lipschitz indépendant de n.

Lorsque les probabilités de transition ne vérifient pas une inégalité du type $\left[P^X\right]_{\text{Lip}} \leq 1+\frac{C}{n}$, on ne pourra plus montrer directement qu'il existe une constante K telle que

$$\left|\mathbb{E}\left[f(X^{0,x}_k)\right] - \mathbb{E}\left[f(X^{0,\tilde{x}}_k)\right]\right| \leq K [f]_{\text{Lip}} |x-\tilde{x}|, \quad \forall f,\forall x,\tilde{x}\in\mathbb{R}^d,\forall 0\leq k\leq n$$

à l'aide des probabilité de transition. Le rapport de Lipschitz des probabilités de transition $\left[P^X\right]^n_{\text{Lip}}$ entre X_0 et X_n n'est pas majoré uniformément en n.

Dans ce cas, il suffit de montrer que le flot de X est lipchitzien uniformément en n pour contrôler la trajectoire de X :

$$\begin{aligned}\left|\mathbb{E}\left[f(X^{0,x}_k)\right] - \mathbb{E}\left[f(X^{0,\tilde{x}}_k)\right]\right| &\leq \left|\mathbb{E}\left[\sup\left|f(X^{0,x}_k) - f(X^{0,\tilde{x}}_k)\right|\right]\right| \\ &\leq \left[F^X\right]_1 [f]_{\text{Lip}} |x-\tilde{x}|, \quad \forall f,\forall x,\tilde{x}\in\mathbb{R}^d,\forall 0\leq k\leq n.\end{aligned}$$

Ainsi, la propriété Lipschitzienne du flot va permettre de contrôler toute la trajectoire de la chaîne de Markov et de donner la régularité de solution y_k du problème de programmation dynamique générique (2.2.1). Dans l'exemple suivant, nous allons illustrer un cas où le flot est lipschitzien tandis que l'inégalité $\left[P^X_k\right]_{\text{Lip}} \leq 1+\frac{C}{n}$ n'est pas vérifiée.

EXAMPLE 3.1.5. Soit W un mouvement brownien standard et $t_k = \frac{k}{n}$, $0 \leq k \leq n$, une discrétisation en temps de $[0,1]$. On définit, pour $k\in\{0,..,n\}$, $X_k = (X^1_k, X^2_k) := (W_{t_k}, \max_{0\leq i\leq k} W_{t_i})$ une version discrète du Brownien et de son maximum. X est une chaîne de Markov car sa probabilité de transition entre X_k et X_{k+1}

dépend uniquement de X_k et de la variable aléatoire ξ_{k+1} indépendante de $\mathcal{F}_k$:

$$\begin{aligned} X_{k+1} &= \left(X_k^1 + \epsilon_{k+1}, \max\left(X_k^2, X_k^1 + \epsilon_{k+1}\right)\right), \quad \forall 0 \leq k < n, \\ &= F_k(X_k, \epsilon_{k+1}) \end{aligned}$$

où $\epsilon_{k+1} = W_{t_{k+1}} - W_{t_k} \sim \mathcal{N}(0, \frac{1}{n})$ est indépendant de X_k. En remarquant que

$$X_{k+i}^{k,x} = \left(x_1 + (W_{t_{k+i}} - W_{t_k}), \max_{0\leq j\leq i}\left(x_2, x_1 + (W_{t_{k+j}} - W_{t_k})\right)\right), \quad \forall 0 \leq k \leq k+i \leq n, \forall x_1, x_2 \in \mathbb{R}$$

et avec l'inégalité

$$\left|\max(a,b) - \max(\tilde{a},\tilde{b})\right| \leq |a-\tilde{a}| + \left|b-\tilde{b}\right|, \quad \forall a,\tilde{a},b,\tilde{b} \in \mathbb{R},$$

on obtient que le flot est lipschitzien uniformément en n :
pour $x = (x_1, x_2), \tilde{x} = (\tilde{x}_1, \tilde{x}_2) \in \mathbb{R}^2$, pour $k \in \{0, .., n\}$

$$\begin{aligned} &\left|\left|\sup_{k\leq k'\leq n}\left|X_{k'}^{k,x} - X_{k'}^{k,\tilde{x}}\right|\right|\right|_p \\ &\leq \left|\left|\sup_{k\leq k'\leq n}\left|\left(x_1 - \tilde{x}_1, \max_{0\leq j\leq i}\left(x_2, x_1 + (W_{t_{k+j}} - W_{t_k})\right) - \max_{0\leq j\leq i}\left(\tilde{x}_2, \tilde{x}_1 + (W_{t_{k+j}} - W_{t_k})\right)\right)\right|\right|\right|_2 \\ &\leq 2\left|x_1 - \tilde{x}_1\right| + \left|x_2 - \tilde{x}_2\right| \\ &\leq K\left|x - \tilde{x}\right| \end{aligned}$$

En regardant au voisinage du point $x_1 = x_2$ avec la norme euclidienne $f = |.|$ (de rapport de Lipschitz 1), on a pour $x = (x_1, x_1)$ et $\tilde{x} = (x_1 - \delta, x_1)$ avec $\delta = |x - \tilde{x}| \geq 0$ et $x_1 - \delta \geq 0$:

$$\begin{aligned} \mathbb{E}\left[\left|X_{k+1}^{k,(x_1-\delta,x_1)}\right|\right] &= \mathbb{E}\left[|(x_1 - \delta, \max(x_1, x_1 - \delta + \epsilon_{k+1}))|\right] \\ &= (x_1 - \delta)\mathbb{E}\left[|(1, 1 + \max(\delta, \epsilon_{k+1}))|\right]. \end{aligned}$$

D'où

$$\begin{aligned} &\left|\mathbb{E}\left[f\left(X_{k+1}^{k,(x_1-\delta,x_1)}\right)\right] - \mathbb{E}\left[f\left(X_{k+1}^{k,(x_1-\delta,x_1)}\right)\right]\right| \\ &= \left|x_1\mathbb{E}\left[|(1, 1 + \max(0, \epsilon_{k+1}))|\right] - (x_1 - \delta)\mathbb{E}\left[|(1, 1 + \max(\delta, \epsilon_{k+1}))|\right]\right| \\ &\geq \delta\mathbb{E}\left[|(1, 1 + \max(\delta, \epsilon_{k+1}))|\right] \\ &\geq \delta\sqrt{2} = \delta\sqrt{2}\left[f\right]_{\text{Lip}} \end{aligned}$$

On a alors que $\left[P_k^X\right]_{\text{Lip}} \geq \sqrt{2}$ et donc on ne peut pas contrôler la chaîne de Markov par le produit de probabilités de transition $\prod_{k=0}^{n-1}\left[P_k^X\right]_{\text{Lip}} \geq 2^{\frac{n}{2}}$ uniformément par rapport au nombre de pas de temps n.

D'une part, on vient de voir que travailler directement sur le flot permet de contrôler toute la trajectoire ce qui n'est pas toujours le cas avec les probabilités de transition. D'autre part, lorsque le flot est peu régulier, on montre dans la remarque suivante qu'il peut être préférable de s'intéresser à la probabilité de transition.

REMARK 3.1.6. Lorsque le flot n'est pas lipschitzien, on peut avoir l'inégalité suivante :

$$\left[P_k^X\right]_{\text{Lip}} < \left[F_k^X\right]_p = +\infty$$

Considère par exemple la chaîne de Markov

$$X_{k+1} = \text{sign}(X_k)\epsilon_{k+1}$$
$$X_0 = x_0$$

où $(\epsilon_k)_{1\leq k\leq n}$ est une suite de variables aléatoires i.i.d. de loi normale centrée $\mathcal{N}(0,1)$ et $\text{sign}(x) = 1_{s\geq 0} - 1_{x<0}$ (de norme $|\text{sign}(x)| = 1$). Alors la loi de X_{k+1} ne dépend pas X_k : $X_{k+1} \sim \mathcal{N}(0,1)$. D'où $\left[P_k^X\right]_{\text{Lip}} = 0$. Cependant, le flot n'est pas lipschitzien : on a

$$\begin{aligned}\left|\left|X_{k+1}^{k,x} - X_{k+1}^{k,-x}\right|\right|_p &= ||\text{sign}(x)\epsilon_{k+1} - \text{sign}(-x)\epsilon_{k+1}||_p, \quad \forall k \in \mathbb{N}, \\ &= 2\,||\epsilon_{k+1}||_p, \quad \forall k \in \mathbb{N}, \forall x \neq 0\end{aligned}$$

d'où

$$\left[F_k^X\right]_p \geq \frac{\left|\left|X_{k+1}^{k,x} - X_{k+1}^{k,-x}\right|\right|_p}{|2x|} \to_{x\to 0} +\infty.$$

On vient d'obtenir l'inégalité extrêmale suivante :

$$\left[P_k^X\right]_{\text{Lip}} = 0 < \left[F_k^X\right]_p = +\infty, \quad \forall k \in \mathbb{N}.$$

Dans les différentes applications que nous regarderons dans la suite, la chaîne de Markov n'aura pas un flot non régulier comme dans la remarque (3.1.6). La chaîne de Markov X dépendra souvent de la trajectoire d'une chaîne de Markov S

$$X_k = (S_k, \Phi_k(S_0, .., S_k))$$

comme dans l'exemple (3.1.5). La propriété Lipschitzienne du flot de X permettra alors de donner la régularité des fonctions y_k solutions du problème de programmation dynamique générique (2.2.1).

La proposition suivante donne la régularité des fonctions y_k et $u_{k,a}$ sous l'hypothèse du probabilité de transition lipschitzienne.

PROPOSITION 3.1.7. *Supposons la chaîne de Markov X a une probabilité de transition lipschitzienne et que $\phi_{k,a} = 1$, $\forall k \in \{0, .., n-1\}, \forall a \in A_k$, alors les fonctions y_k et $u_{k,a}$ sont lipchitziennes, $\forall k \in \{0, .., n\}$: il existe une constante $C > 0$ telle que*

$$\max_k\{[y_k]_{Lip}, [u_k]_{Lip}\} \leq C\frac{K^n\left[P^X\right]_{Lip}^n - 1}{K\left[P^X\right]_{Lip} - 1}.$$

La proposition (3.1.7) montre que le rapport de Lipschitz de y et u sont en puissance de n. Ainsi l'erreur est géométrique par rapport au nombre de pas de temps. Si le schéma nécessite une convergence en temps (c'est-à-dire que $n \to +\infty$), on souhaite réduire cette erreur. Pour cela, il faudra étudier au cas par cas (on pourra montrer la stabilité du rapport de Lipschitz de y_k par rapport à n en utilisant par exemple des propriétés du flot ou en obtenant une majoration du type $\prod_{k=1}^n \left[P_k^X\right]_{\text{Lip}} \leq C$ sur le semi-groupe de transition).

Par souci de simplification, on notera z_k pour $z_{k,a}$, si $z_{k,a}$ est indépendant de a avec $z_{k,a} = u_{k,a}$, $\phi_{k,a}$, $T_{k,a}$ ou $\psi_{k,a}$ (idem pour $v_k(x) = v_k(x,a)$). De même, on notera $\phi_{k,a}(x,\xi) = \phi_{k,a}(x), \phi_k(\xi)...$ selon la dépendance aux paramètres (a,x,ξ).

3.2. Options bermudéennes

On s'intéresse, ici, à évaluer une option américaine qui est une option donnant à son détenteur le privilège d'être exercée à tout instant durant sa durée de vie ou à sa maturité, ou à évaluer une option bermudéenne qui est une option de type américaine où les périodes d'exercice sont des dates discrètes. Afin de bénéficier du privilège de l'exercice par anticipation, l'acheteur de l'option doit exercer son option au meilleur moment possible. Cela revient à étudier un problème de temps d'arrêt optimal, qui fut développé initialement par Snell (cf [**Sne52**]) pour les processus stochastiques en temps discret. De nombreuses recherches se sont focalisées sur diverses méthodes pour évaluer une option américaine. Des approches par arbre binomiale (voir par exemple [**CRR79**]) ou par différences finies (voir [**BS77**] et [**JLL90**]) peuvent être adaptées en très faible dimension. Cependant, l'augmentation de la dimension du problème a motivé le développement de méthodes de Monte Carlo : méthode de mailles ([**BG97**]), des approches basées sur la régression ([**Che97**], [**LS01**], [**TR99**], [**TR00**]), sur le calcul de Malliavin ([**LR02**], [**BT04**]), approche martingale ([**Rog02**]), méthode par quantification ([**BP03a**], [**BP03b**], [**BPP05**]).

Etant données une chaîne de Markov $X = \{X_k\}_{0\leq k\leq n}$ et une suite de fonctions $\{g_k : \mathbb{R}^d \to \mathbb{R}\}_{0\leq k\leq n}$ lipschitziennes de rapport K, nous nous intéressons au problème de temps d'arrêt optimal :

$$Y_k = \text{esssup}_{\tau\in\mathcal{T}_{k,n}} \mathbb{E}\left[g_\tau(X_\tau)|\mathcal{F}_k\right],$$

où $\mathcal{T}_{k,n}$ représente l'ensemble des temps d'arrêts à valeurs dans $\{k,..,n\}$ et Y est aussi appelé l'enveloppe de Snell de $\{g_k(X_k)\}_{0\leq k\leq n}$.

Il est connu que l'enveloppe de Snell $\{Y_k\}_{0\leq k\leq n}$ satisfait $Y_k = y_k(X_k)$, $0 \leq k \leq n$, où les fonctions boréliennes $y_k : \mathbb{R}^d \to \mathbb{R}$ vérifient le principe de programmation dynamique suivant :

$$\begin{cases} y_n(x) = g_n(x) \quad \forall x \in \mathbb{R}^d \\ y_k(x) = \max\left\{g_k(x); \mathbb{E}\left[y_{k+1}(X_{k+1})|X_k = x\right]\right\}, \quad \forall x \in \mathbb{R}^d, \forall 0 \leq k \leq n-1. \end{cases}$$

On rentre dans notre formule de programmation dynamique générique (2.2.1) avec

$$\begin{aligned} h &= g_n, \\ \phi_{k,a} &= 1, \quad \forall 0 \leq k \leq n-1, \forall a \in A_k, \\ T_{k,a}(x,u) &= u, \quad \forall 0 \leq k \leq n-1, \forall u \in \mathbb{R}, \forall x \in \mathbb{R}^d, \forall a \in A_k, \\ R_k(x,v) &= \max\left\{g_k(x), v\right\} \quad \forall 0 \leq k \leq n-1, \forall x \in \mathbb{R}^d, \forall v \in \mathbb{R}. \end{aligned}$$

Alors l'hypothèse (H_p) est trivialement vérifiée pour tout $p \in [1,\infty]$ car $\lambda.\phi_{k,a} \in \{0;1\}$ pour $\lambda \in \mathcal{S}^{T_k} = \{0;1\}$.

La proposition suivante donne la régularité des fonctions $\{y_k\}_k$ et $\{u_k\}_k$.

PROPOSITION 3.2.1. *Supposons que la probabilité de transition de X est lipschitzienne, alors les fonctions $\{y_k\}_k$ et $\{u_k\}_k$ sont lipschitziennes de rapport majoré par $K\left[P^X\right]_{Lip}^{n-k}$.*

Si de plus, on suppose que le flot de X est lipschitzien, alors le rapport de Lipschitz de $\{y_k\}_k$ et $\{u_k\}_k$ est majoré par $K\left[F^X\right]_{Lip}$ (uniformément en k et n).

Si l'on considère que cette enveloppe de Snell discrète en temps (par exemple, il s'agit du prix d'une option bermudéenne) converge vers le problème de temps d'arrêt optimal en temps continu lorque $n \to \infty$ (en finance, il s'agit de l'approximation d'une option américaine par une option bermudéenne), alors il est préférable que le rapport de Lipschitz des fonctions $\{y_k\}_k$ soit uniformément borné indépendamment du nombre de pas de temps n. C'est la cas lorsque le flot de X est Lipschitzien ou lorsque la probabilité de transition est majorée par

$$\left[P^X\right]_{\text{Lip}} \leq 1 + \frac{C}{n}$$

avec une constante C indépendante de n (on obtient alors $\left[P^X\right]_{\text{Lip}}^n \leq e^C$). Nous avons alors l'erreur du schéma par quantification majorée par

$$\varepsilon(\text{quantif}) \leq C \sum_{jk=0}^{n-1} \left|\left|X_k - \hat{X}_k\right|\right|_p$$

avec la constante C indépendante du nombre de pas de temps n.

En supposant que $X_k \in \mathbb{L}^{p+\eta}$, on obtient finalement à l'aide du théorème de Zador

$$\varepsilon(\text{quantif}) \leq C_{d,p,\eta} \left|\left|X_k\right|\right|_{p+\eta} \sum_{jk=0}^{n-1} N_k^{-\frac{1}{d}}.$$

3.3. EDSR avec sauts

Les équations différentielles stochastiques rétrogrades (EDSR) ont été introduites par J.-M. Bismut ([**Bis73**]) en 1973 dans le cas d'un générateur linéaire et par E. Pardoux et S. Peng ([**PP90**]) en 1990 dans le cas d'un générateur non-linéaire. De nombreux mathématiciens ont contribué et contribuent toujours à la théorie des EDSR : on pourrait par exemple citer des résultats sur des estimations à priori ([**KPQ97**]), sur les relations avec les équations aux dérivées partielles ([**BBP97**]), sur des applications à la maximisation d'utilité ([**RK00**] et dans le cadre avec sauts [**Bec05**]), pour des problèmes de couverture ([**EL05**]), pour une extension dans la cadre des diffusion avec sauts les travaux ([**TL94**]), sur une discrétisation en temps dans ([**Zha04**] et dans le cas des sauts [**BE08**]).

Différents schémas de discrétisation ont été étudiés : à l'aide d'un algorithme appelé schéma en quatre étapes ([**MPY94**]), via un théorème du type Donsker ([**BDM01**], [**PX08**]), par régression ([**Che97**], [**JLW05**]), par quantification ([**BP03a**], [**DM06**], [**DM08**]), par une méthode de Monte Carlo séquentielle ([**GL07**]), avec un schéma progressif ([**BR07**]), par calcul de Malliavin ([**BT04**]).

Considérons la solution $(S,Y,Z,U) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{d'} \times \mathbb{R}^E$ de l'équation différentielle stochastique rétrograde suivante :

$$S_t = S_0 + \int_0^t b(S_r)dr + \int_0^t \sigma(S_r)dW_r + \int_0^t \int_E \beta(S_{r-},e)\overline{\mu}(de,dr),$$
$$Y_t = h(S_T) + \int_t^T f(\Theta_r)dr - \int_t^T Z_r dW_r - \int_t^T \int_E U_r(e)\overline{\mu}(de,dr),$$

où $\Theta := (S,Y,Z,\Gamma)$ avec $\Gamma = \int_E \rho(e)U(e)\lambda(de)$. W est un mouvement Brownien d_W-dimensionel, $\overline{\mu}(de,dr) = \mu(de,dr) - \lambda(de)dr$ est une mesure de Poisson compensée (avec λ une mesure finie sur $E \subset \mathbb{R}^{d_E}$) indépendante de W. On suppose les fonctions b, σ et $\beta(.,e)$, h et f sont lipschitziennes de rapport K . (Y,Z,U) peuvent être associés à la solution (unique au sens de viscosité) v de l'EDP

$$\begin{aligned} -\mathcal{L}v - f(x,v,\sigma' Dv, \mathcal{I}[v]) &= 0, \\ v(T,x) &= h(x), \end{aligned} \tag{3.3.1}$$

où

$$\begin{aligned} \mathcal{L}v(t,x) &:= \frac{\partial v}{\partial t}(t,x) + \langle b, Dv\rangle(t,x) + \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma' D^2 v\right](t,x) \\ &\quad + \int_E \left\{v(t,x+\beta(x,e)) - v(t,x) - Dv(t,x)\beta(x,e)\right\}\lambda(de), \\ \mathcal{I}[v](t,x) &:= \int_E \left\{v(t,x+\beta(x,e)) - v(t,x)\right\}\lambda(de), \end{aligned}$$

avec $Y_t = v(t,S_t)$, $Z_t = \sigma(S_{t-})' Dv(t,S_{t-})$ et $U_t(e) = v(t,S_{t-}+\beta(S_{t-},e)) - v(t,S_{t-})$.

Tout d'abord, donnons la discrétisation en temps. On se donne une grille régulière $\pi = \left\{t_k = k\frac{T}{n}, 0 \le k \le n\right\}$. On approxime le processus S par son schéma d'Euler habituel S^π :

$$S^\pi_{t_{k+1}} = S^\pi_{t_k} + b(S^\pi_{t_k})(t_{k+1}-t_k) + \sigma(S^\pi_{t_k})(W_{t_{k+1}} - W_{t_k}) + \int_E \beta(S^\pi_{t_k},e)\overline{\mu}(de,(t_k,t_{k+1}]).$$

Le schéma d'Euler $(Y^\pi, Z^\pi, \Gamma^\pi)$ est construit de manière rétrograde :

$$\begin{aligned} \overline{Z}^\pi_{t_k} &= \mathbb{E}\left[Y^\pi_{t_{k+1}} \frac{W_{t_{k+1}} - W_{t_k}}{t_{k+1}-t_k} | S^\pi_{t_k}\right] \\ \overline{\Gamma}^\pi_{t_k} &= \mathbb{E}\left[Y^\pi_{t_{k+1}} \frac{\int_E \rho(e)\overline{\mu}(de,(t_k,t_{k+1}])}{t_{k+1}-t_k} | S^\pi_{t_k}\right] \\ Y^\pi_{t_k} &= \mathbb{E}\left[Y^\pi_{t_{k+1}} | S^\pi_{t_k}\right] + \frac{T}{n} f\left(S^\pi_{t_{k+1}}, \mathbb{E}\left[Y^\pi_{t_{k+1}} | S^\pi_{t_k}\right], \overline{Z}^\pi_{t_k}, \overline{\Gamma}^\pi_{t_k}\right) \end{aligned}$$

où $Y^\pi_{t_n} = h(S^\pi_{t_n})$. L'article ([**BE08**]) donne la convergence de l'algorithme en temps (nous donnerons aussi la convergence dans le cas plus général pour des edrs avec réflexions, cf la partie suivante).

Ce schéma rétrograde est un cas particulier de notre formule de programmation dynamique (2.2.1) avec les notations suivantes : la chaîne de Markov associée vaut

$$X_k = S^\pi_{t_k}, \quad \forall 0 \le k \le n,$$

et nous avons $y_k(X_k) = Y_{t_k}^\pi$ vérifiant le principe de programmation dynamique générique (2.2.1) :

$$\phi_k(\xi_{k+1}) = \left(1, W_{t_{k+1}} - W_{t_k}, \int_E \rho(e)\overline{\mu}(de,(t_k,t_{k+1}])\right),$$
$$T_k(x,u) = u_1 + \frac{T}{n} f(x,u^1,\frac{n}{T}u^2,\frac{n}{T}u^3), \quad x\in\mathbb{R}^d, u=(u^1,u^2,u^2)\in\mathbb{R}\times\mathbb{R}^{d_W}\times\mathbb{R}^{d_E},$$
$$R_k(x,v) = v, \quad \forall x\in\mathbb{R}^d, \forall v\in\mathbb{R}.$$

Le lemme suivant permet de contrôler les diffusions avec sauts sur un pas de temps.

LEMMA 3.3.1. *Soit $\tilde{d}\in\mathbb{N}^*$. Soit $p\in[1,+\infty[$. Soient $x,\tilde{b}\in\mathbb{R}^{\tilde{d}}$, $\tilde{\sigma}\in\mathbb{R}^{\tilde{d}\times d_W}$, $\tilde{\beta}\in\mathbb{R}^{\tilde{d}\times d_E}$ vérifiant $|x|=1$ et $\left|\tilde{b}\right|, |\tilde{\sigma}|, \left|\tilde{\beta}\right|$ majoré par une constante par $M>0$. Il existe une constante $C_{p,M,T}>0$ (dépendant de p, M et T) telle que*

$$\left|\left|x+\tilde{b}\frac{T}{n}+\tilde{\sigma}(W_{t_{k+1}}-W_{t_k})+\int_E\tilde{\beta}(e)\overline{\mu}(de,(t_k,t_{k+1}])\right|\right|_p \le 1+\frac{C_{p,M,T}}{n},$$
$$\left|\left|\tilde{b}\frac{T}{n}+\tilde{\sigma}(W_{t_{k+1}}-W_{t_k})+\int_E\tilde{\beta}(e)\overline{\mu}(de,(t_k,t_{k+1}])\right|\right|_p \le \frac{C_{p,M,T}}{n^{\frac{1}{2\vee p}}}.$$

Le contrôle de la diffusion donné par le lemme précédent donne la régularité lipschitzienne du flot.

COROLLARY 3.3.2. *Le flot de X est $\mathbb{L}_p$-lipchitzien pour $p\in]1,+\infty]$ et il existe une constante $C_p>0$ (ne dependant que de p) telle que*

$$\left[F_k^X\right]_p \le 1+\frac{C_p}{n}, \quad \forall k\in\{0,..,n-1\}.$$

Le lemme (3.3.1) permet aussi d'obtenir la stabilité du schéma en temps sous $\mathbb{L}_p$ pour $p\in]1,+\infty]$.

COROLLARY 3.3.3. *L'hypothèse (H_p) est vérifiée $\forall p\in]1,+\infty]$.*

La proposition suivante nous donne la régularité de y_k, ce qui donne la vitesse de convergence grâce au théorème de Zador. La méthode des gerbes en approximant ξ_{k+1} par $\hat{\xi}_{k+1}$ par une méthode de quadrature (par exemple avec $\frac{W_{t_{k+1}}-W_{t_k}}{h}$ approximé par une loi binomiale ou trinomiale) nécessite une certaine régularité.

On note $\mathcal{C}_b^m$ l'ensemble des fonctions m-fois différentiables à dérivées bornées.

PROPOSITION 3.3.4. *Les fonctions $\{y_k\}_k$ et $\{u_k\}_k$ sont lipschitziennes de rapport bornées uniformément en k et n.*
Si on suppose, de plus, que les fonctions $b,\sigma,\beta(.,e),f,h$ appartiennent à $\mathcal{C}_b^m$ pour $m\in\mathbb{N}^*$, alors $y_k\in\mathcal{C}_b^m$ (avec une majorations des dérivées indépendantes de n et k) : il existe une constante $L\ge 0$ (indépendante de n) telle que $|D^\alpha y_k|\le L$ $\forall\alpha=(\alpha_1,..,\alpha_d)\in\mathbb{N}^d$ vérifiant $1\le\alpha_1+...+\alpha_d\le m$.

La propriété lipschitzienne de la solution y donnée par la proposition précédente permet de majorer l'erreur du schéma par quantification :

$$\varepsilon(\text{quantif}) \le C\sum_{k=0}^{n-1}\left|\left|X_k-\hat{X}_k\right|\right|_p$$

avec la constante C indépendante de n.

Le théorème de Zador permet d'obtenir l'inégalité suivante

$$\varepsilon(\text{quantif}) \leq C_{d,p,\eta} \left|\left|X_k\right|\right|_{p+\eta} \sum_{jk=0}^{n-1} N_k^{-\frac{1}{d}}.$$

REMARK 3.3.5. Ainsi dans le cas continu ($\beta = 0$) et avec tous les paramètres appartenant à $\mathcal{C}_b^3$, la convergence de l'algorithme par grille (méthode des gerbes) converge si on approxime les variables aléatoires $\xi_{k+1} = \frac{W_{t_{k+1}} - W_{t_k}}{h}$ par la loi binomiale

$$\mathbb{P}\left(\left\{\hat{\xi}_{k+1}^i = \pm 1\right\}_{1 \leq i \leq d}\right) = \frac{1}{2^d}.$$

De plus, si $X_k = W_{t_k}$, cette méthode des gerbes appliquée sur la grille de quantification optimale Γ_k (de $\overline{X}_k$) avec le nombre de points $N_k \geq (1+k)^d$ s'identifie à la méthode de l'arbre binomiale.

Le corollaire suivant donne un résultat théorique intéressant sur la régularité sur la solution de l'EDP (3.3.1). En effet, la convergence (voir [**BE08**]) du schéma discret en temps qui est uniformément régulière (3.3.4) entraîne la régularité de la solution de l'équation différentielle stochastique rétrograde.

COROLLARY 3.3.6. *Si on suppose de plus que les fonctions $b, \sigma, \beta(.,e), f, h$ appartiennent à $\mathcal{C}_b^m$ pour $m \in \mathbb{N}^*$, alors la solution $x \in \mathbb{R}^d \to y(t,x)$ de solution l'EDP (3.3.1) appartient à $\mathcal{C}_b^m$.*

3.4. EDSR généralisée

Ici, on s'interesse aux équations différentielles stochastiques rétrogrades avec double reflexions. Les EDSR avec une barrière sont liées aux problèmes de temps d'arrêts optimaux (comme pour les options américaines), on pourra citer ([**KKP$^+$97**]) et ([**BP03a**]), ([**BC08**]) pour les approximations numériques. Avec deux barrières, ces équations peuvent évaluer une *game option* (voir [**Kif00**]), c'est-à-dire une option où l'acheteur et le vendeur peuvent excercer le droit de acheter ou vendre l'option à des prix convenus. L'existence et l'unicité des EDSR avec doubles réflexions sont étudiées dans ([**CK96**]) dans le cadre continu et dans ([**CM08**]), ([**HH06**]) dans le cas avec sauts. Des schémas sont développés pour le processus progressif continu et non path-dépendant dans ([**Cha09**]), ([**Otm08**], ([**Xu08**]), ([**Kif06**]) et dans le cas général dans la partie sur la discrétisation en temps des BSDE avec doubles barrières.

Le processus progressif X est markovien, pathdependant (comme pour les options asiatiques, lookback...) dans un modèle à saut.

Considérons la diffusion S solution de l'équation

$$S_t = S_0 + \int_0^t b(S_r)dr + \int_0^t \sigma(S_r)dW_r + \int_0^t \int_E \beta(S_{r-}, e)\overline{\mu}(de, dr)$$

où W est un mouvement brownien d_W-dimensionel et où $\overline{\mu}(de,dr) = \mu(de,dr) - \lambda(de)dr$ est une mesure de Poisson compensée (avec λ une mesure finie sur $E \subset \mathbb{R}^{d_E}$) indépendante de W. On suppose les fonctions b, σ et $\beta(.,e)$ lipschitziennes et β bornées. On construit le processus de Markov $X = (S, P)$ de manière suivante : pour $t \geq s \geq 0$

$$X_t = (S_t, \Phi_t((S_s)_{0 \leq s \leq t}))$$
$$X_t = (S_t, \varphi_{s,t}(X_s, (S_u)_{s \leq u \leq t}))$$

avec Φ_t et $\varphi_{s,t}$ des fonctionnelles $\mathbb{L}^\infty$-lipschitziennes déterministes ($\Phi_t = \varphi_{0,t}$). On s'intéresse à l'équation différentielle stochastique rétrograde réfléchie suivante :

$$\begin{cases} Y_t = h(X_T) + \int_t^T f(\Theta_s)ds + \left(K_T^+ - K_t^+\right) - \left(K_T^- - K_t^-\right) \\ \quad - \int_t^T Z_s ds W_s - \int_t^T \int_E U_s(e)\overline{\mu}(de,dr) \quad \forall t \in [0,T],\ \mathbb{P}-p.s. \\ l(X_t) \leq Y_t \leq g(X) \quad \forall t \in [0,T],\ \mathbb{P}-a.s. \\ \text{et } \int_0^T (Y_s - l(X_s))\, dK_s^+ = \int_0^T (Y_s - g(X_s))\, dK_s^- = 0 \quad \mathbb{P}-a.s. \end{cases}$$

où $K^\pm$ are continuous non decreasing, $\Theta := (X,Y,Z,\Gamma)$ avec $\Gamma = \int_E \rho(e)U(e)\lambda(de)$ et $l \leq h \leq u$.

On suppose les fonctions f, h, l et g lipschitziennes.

Les articles ([**CM08**]) et ([**HH06**]) donnent l'existence et l'unicité de la solution Y sous ces hypothèses.

Notons $\pi = \{t_i\}$ avec $t_i = \frac{iT}{n}$ $\forall 0 \leq i \leq n$, $\pi(t) = \sup\{t_i \in \pi | t \geq t_i\}$ et $\left(S_{t_i}^\pi\right)_{0\leq i\leq n}$ le schéma d'Euler de S

$$S_{t_{i+1}}^\pi = S_{t_i}^\pi + b(S_{t_i}^\pi)(t_{i+1} - t_i) + \sigma(S_{t_i}^\pi)(W_{t_{i+1}} - W_{t_i}) + \int_E \beta(S_{t_i}^\pi, e)\overline{\mu}(de,(t_i,t_{i+1}]).$$

On définit alors l'approximation X^π de X par

$$X_{t_{i+1}}^\pi = \left(S_{t_{i+1}}^\pi, \varphi_{t_i,t_{i+1}}\left(X_{t_i}^\pi, (S_{\pi(s)}^\pi)_{t_i\leq s\leq t_{i+1}}\right)\right).$$

REMARK 3.4.1. $\{X_t\}_{0\leq t\leq T}$ est un processus markovien. En effet, soient $0 \leq s \leq t \leq T$ et $h : \mathbb{R}^d \to \mathbb{R}$ une fonction borélienne bornée, on a

$$\begin{aligned} \mathbb{E}\left[h(X_t)|\mathcal{F}_s\right] &= \mathbb{E}\left[h\left(S_t, \varphi_{s,t}\left(X_s, (S_u)_{s\leq u\leq t}\right)\right)|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[h\left(S_t, \varphi_{s,t}\left(X_s, (S_u)_{s\leq u\leq t}\right)\right)|X_s, S_s\right] \end{aligned}$$

car S_t est un processus markovien et $\phi_{s,t}$ est déterministe. Comme $S_t \in \sigma(X_t)$, X est Markovien :

$$\mathbb{E}\left[h(X_t)|\mathcal{F}_s\right] = \mathbb{E}\left[h(X_t)|X_s\right], \quad \forall 0 \leq s \leq t \leq T.$$

De même, son approximation en temps $X_{t_i}^\pi$ est une chaîne de Markov. En effet, comme $\left\{S_{t_i}^\pi\right\}_{0\leq i\leq n}$ est une chaîne de Markov et $S_{t_i}^\pi \in \sigma(X_{t_i}^\pi)$ pour tout $i \in \{0,...,n\}$, on a

$$\begin{aligned} \mathbb{E}\left[h(X_{t_{i+1}}^\pi)|\mathcal{F}_{t_i}\right] &= \mathbb{E}\left[h\left(S_{t_{i+1}}^\pi, \varphi_{t_i,t_{i+1}}\left(X_{t_i}^\pi, (S_{\pi(s)}^\pi)_{t_i\leq s\leq t_{i+1}}\right)\right)|\mathcal{F}_{t_i}\right] \\ &= \mathbb{E}\left[h\left(S_t, \varphi_{s,t}\left(X_s, (S_u)_{s\leq u\leq t}\right)\right)|X_{t_i}^\pi, S_{t_i}^\pi\right] \\ &= \mathbb{E}\left[h(X_{t_{i+1}}^\pi)|X_{t_i}^\pi\right] \end{aligned}$$

On s'intéresse à l'algorithme étudié dans (voir chapitre suivante) vérifiant la programmation dynamique descendante suivante :

$$\begin{cases} y_n(x) = h(x) \\ u_k(x) = \mathbb{E}\left[y_{i+1}(X_{t_{i+1}}^\pi)\phi_k(\xi_{k+1})|X_{t_i}^\pi = x\right] \\ y_k(x) = l(x) \vee \left(u_k^1(x) + \frac{T}{n}f(x, u_k^1(x), \frac{T}{n}u_k^2(x), \frac{T}{n}u_k^3(x))\right) \wedge g(x), \quad \forall x \in \mathbb{R}^d \quad \forall 0 \leq i < n. \end{cases}$$

Ce schéma peut se réécrire dans notre programmation dynamique générique (2.2.1) avec

$$\begin{aligned}
\phi_k\left(\xi_{k+1}\right) &= \left(1, W_{t_{k+1}} - W_{t_k}, \int_E \rho(e)\overline{\mu}(de, (t_k, t_{k+1}])\right), \\
u_k^1(x) &= \mathbb{E}\left[y_{k+1}(X^\pi_{t_{k+1}})|X^\pi_{t_k} = x\right], \\
T_k\left(x, u\right) &= u^1 + \frac{T}{n} f(x, u^1, \frac{n}{T}u^2, \frac{n}{T}u^{3,}), \\
R_k(x, v) &= l\left(x\right) \vee v \wedge g\left(x\right), \quad \forall k \in \{0, .., n-1\}, \forall x \in \mathbb{R}^d, \forall v \in \mathbb{R}.
\end{aligned}$$

Le chapitre suivant donnera la convergence de cette discrétisation en temps. Elle montrera aussi que les fonctions sont y_k et u_k sont uniformément lipschitziennes.

En outre, l'hypothèse (H_p) est vérifiée $\forall p \in]1, +\infty]$ (la preuve est identique que celle du corollaire (3.3.3)).

Ainsi, la régularité de y_k et u_k et la vérification de l'hypothèse (H_p) entraînent la stabilité du schéma en temps sous $\mathbb{L}_p$ pour $p \in]1, +\infty]$ et la convergence des schémas par quantification :

$$\varepsilon(\text{quantif}) \leq C \sum_{k=0}^{n-1} \left|\left|X_k - \hat{X}_k\right|\right|_p$$

avec la constante C indépendante de n. On obtient finalement à l'aide du théorème de Zador

$$\varepsilon(\text{quantif}) \leq C_{d,p,\eta} \left|\left|X_k\right|\right|_{p+\eta} \sum_{jk=0}^{n-1} N_k^{-\frac{1}{d}}.$$

3.5. EDP parabolique totalement non linéaire

On dit que (Y, Z, Γ, A) est une solution d'une EDSR du second ordre (voir **[CSTV06]**) si

$$\begin{aligned}
dX_t &= \sigma(t, X_t)dW_t, \\
dY_t &= -f(t, X_t, Y_t, Z_t, \Gamma_t)dt + Z_t dX_t, \\
dZ_t &= A_t dt + \Gamma_t dX_t, \\
Y_T &= h(X_T).
\end{aligned} \tag{3.5.1}$$

Considérons l'équation aux dérivées partielles parabolique non-linéaire suivante :

$$\begin{aligned}
\mathcal{L}v + f(., v, Dv, D^2v) = 0 \quad &\text{sur } [0, T) \times \mathbb{R}^d, \\
v(1, .) = h \quad &\text{sur } \mathbb{R}^d,
\end{aligned} \tag{3.5.2}$$

où $\mathcal{L}$ est l'opérateur linéaire du second ordre défini par :

$$\mathcal{L}v := \frac{\partial v}{\partial t} + \langle b, Dv\rangle + \frac{1}{2}\text{Tr}\left[\sigma\sigma' D^2 v\right]$$

avec b, σ de $\mathbb{R}_+ \times \mathbb{R}^d$ dans $\mathbb{R}^d$ et $\mathbb{M}^d$ et la partie non linéaire est donnée par la fonction $f : (t, x, r, p, \gamma) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d \to f\left(t, x, r, p, \gamma\right) \in \mathbb{R}$.

Alors sous de bonnes conditions de régularités de v solutions de l'EDP (3.5.2), l'article ([**CSTV06**]) montre que

$$\begin{cases} Y_t = v(t, X_t) \\ Z_t = Dv(t, X_t) \\ \Gamma_t = D^2 v(t, X_t) \\ A_t = (\frac{\partial}{\partial t} + \mathcal{L})Dv(t, X_t) \end{cases}$$

est solution de l'EDSR du second ordre (3.5.1), où $\mathcal{L}$ est le générateur infinitésimal du processus X.

Ainsi, la résolution numérique de la 2-EDSR (3.5.1) permet de déterminer la solution de l'EDP (3.5.2) d'une manière plus adaptée que la méthode des différences finies lorsque la dimension est importante. Ces équations sont appliquées en finance par exemple pour des problèmes de sur-réplication en modèle de volatilité incertaine (voir [**ALP95**], [**GV02**]), sous contraintes gamma (voir [**ST00**], [**CST05**]), de couverture avec coûts de transaction (voir [**Lel85**]).

Soient $n \in \mathbb{N}^*$ et la discrétisation en temps $\{t_0 < t_1 < .. \leq t_n = T\}$ avec $t_k = \frac{kT}{n}$ pour $k = 0, ..., n$, on définit la discrétisation par le schéma d'Euler de la diffusion correspondant à l'opérateur $\mathcal{L}$:

$$X_{k+1} = \overline{X}_{t_{k+1}} = \overline{X}_{t_k} + b(t_k, \overline{X}_{t_k})\frac{T}{n} + \sigma(t_k, \overline{X}_{t_k})\left(W_{t_{k+1}} - W_{t_k}\right).$$

On estime $v(t_k, x)$ par $y_k(x)$ définie de manière rétrograde par programmation dynamique :

$$y_n(x) = h(x), \quad \forall x \in \mathbb{R}^d,$$

$$y_k(x) = \mathbb{E}\left[y_{k+1}(X_{k+1}^{k,x})\right] + \frac{T}{n} f\left(t_k, x, \mathbb{E}\left[y_{k+1}(X_{k+1}^{k,x})\phi_k(x, W_{t_{k+1}} - W_{t_{ik}})\right]\right), \quad \forall 0 \leq k \leq n-1, \forall x \in \mathbb{R}^d.$$

On rentre dans notre formule de programmation dynamique générique (2.2.1) en prenant

$$\phi_k(x, w) = \left(1, (\sigma')^{-1}\frac{w}{\frac{T}{n}}, (\sigma')^{-1}\frac{w^2 - \frac{T}{n}I_d}{(\frac{T}{n})^2}\sigma^{-1}\right), \quad \forall x \in \mathbb{R}^d, \forall w \in \mathbb{R}^d$$

$$T_k(x, u) = u_1 + \frac{T}{n} f(t_k, x, u), \quad \forall k \in \{0, .., n-1\}, \quad \forall x \in \mathbb{R}^d, \forall u = (u_1, u_2, u_3) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d \times d},$$

$$R_k = 1.$$

On suppose de plus que la fonction f est bornée, lipschitzienne en (r, p, γ) et elliptique et strictement dominée par la diffusion de l'opérateur $\mathcal{L}$, i.e. il existe $\epsilon > 0$ tel que

$$0 \leq \nabla_\gamma f \leq (1 - \epsilon)\sigma\sigma'.$$

Le lemme suivant donne la stabilité $\mathbb{L}_\infty$ de l'algorithme.

LEMMA 3.5.1. *Les hypothèses (H_∞) et $(\mathcal{P}_p)$ pour $p \in [1; +\infty[$ sont vérifiées (d'où une stabilité $\mathbb{L}_\infty$ de l'algorithme).*

Dans l'article ([**FTW11**]), la convergence en temps est montrée et la propriété lipschtzienne de y_k (uniformément en k) est prouvé. En la combinant avec la condition (H_∞) (qui est vérifiée), on a la convergence par les trois méthodes de

quantification proposées avec une vitesse géométrique en le nombre de pas de temps n :

$$\varepsilon(\text{quantif}) \leq C^n \sum_{k=0}^{n-1} \left|\left| X_k - \hat{X}_k \right|\right|_p ,$$

et à l'aide du théorème de Zador :

$$\varepsilon(\text{quantif}) \leq C_{d,p,\eta} \, ||X_k||_{p+\eta} \sum_{jk=0}^{n-1} N_k^{-\frac{1}{d}}.$$

3.6. Maximisation d'utilité exponentielle sous contrainte

L'utilité permet de modéliser les préférences d'un agent financier (voir [**Mer71**] un article pionner sur l'allocation de portefeuille), de donner un prix d'une option par indifférence (notion introduite par [**HN89**]) dans un marché incomplet.

Soit X une chaîne de Markov à valeurs dans $\mathbb{R}^d$. On considère un portefeuille de richesse finale $h(X_n) - \sum_{i=0}^{n-1} \pi_i.(X_{i+1} - X_i)$ avec les contraintes $\pi \in A = \{A_0, ..A_{n-1}\}$. On souhaite maximiser l'espérance de l'utilité de ce portefeuille sur l'ensemble des stratégies admissibles (i.e. $\pi_k \in A_k$, $\forall k \in \{0, .., n-1\}$), c'est-à-dire résoudre la maximisation suivante

$$v_0 = \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[u\left(h(X_n) - \sum_{i=k}^{n-1} \pi_i.(X_{i+1} - X_i) \right) \right]$$

où $u : x \in \mathbb{R} \to -e^{-\gamma x}$ est la fonction d'utilité exponentielle avec une aversion au risque γ.

Comme la fonction exponentielle est additive, cela revient à optimiser un portefeuille sous les contraintes $\pi_k \in A_k$:

$$y_k(X_k) = \inf_{\pi \in \mathcal{A}} \mathbb{E}\left[\exp\left(-\gamma \left(h(X_n) - \sum_{i=k}^{n-1} \pi_i.(X_{i+1} - X_i) \right) \right) | X_k \right]$$

avec $y_k(X_0) = v_0$. La fonction valeur y vérifie le principe de programmation dynamique rétrograde suivant :

$$\begin{aligned} y_n &= \exp -\gamma h, \\ y_k(x) &= \inf_{\pi_k \in A_k} \mathbb{E}\left[e^{\gamma \pi_k . (X_{k+1}^{k,x} - x)} y_{k+1}(X_{k+1}^{k,x}) \right], \quad \forall k \in \{0, ..., n-1\}, \forall x \in \mathbb{R}^d, \end{aligned}$$

qui peut se réécrire sous la formulation de la programmation dynamique générique (2.2.1) :

$$\begin{aligned} y_n(x) &= \exp -\gamma h(x), \\ \phi_{k,\pi_k}(x, \xi_{k+1}) &= e^{\gamma \pi_k (X_{k+1}^{k,x} - x)}, \\ u_{k,a}(x) &= \mathbb{E}\left[e^{a.(X_{k+1}^{k,x} - x)} y_{k+1}(X_{k+1}^{k,x}) \right], \quad \forall a \in A_k, \\ T_{k,\pi}(x, u) &= u, \\ R_k(x, u) &= \inf_{a \in A_k} u(x, a), \quad \forall k \in \{0, ..., n-1\}, \forall x \in \mathbb{R}^d, \forall u \in \mathbb{R}^{d \times A_k}. \end{aligned}$$

On considère la grille $\pi = \left\{t_i = i\frac{T}{n} | 0 \leq i \leq n\right\}$. Considérons le schéma d'Euler $S \in \mathbb{R}^{d_S}$ d'une diffusion avec saut :

$$S_{k+1} = S_k + b(S_k)(t_{k+1} - t_k) + \sigma(S_k)\left(W_{t_{k+1}} - W_{t_k}\right) + \int_E \beta(S_k, e)\overline{\mu}(de, (t_k, t_{k+1}]),$$

où W est un mouvement Brownien d_W-dimensionel et $\overline{\mu}(de, dr = \mu(de, dr) - \lambda(de)dr$ est une mesure de Poisson compensée (avec λ une mesure finie sur $E \subset \mathbb{R}^{d_E}$) indépendante de W.

On construit le processus de Markov $X_k = (S, P) \in \mathbb{R}^d$ de manière suivante : pour $k' \geq k$

$$X_{k'} = (S_{k'}, P_{k'}) = \left(S_k, \Phi_{k,k'}\left(X_k, (S_i)_{k \leq i \leq k'}\right)\right),$$

avec Φ $\mathbb{L}^\infty$-lipschitzien : il existe une constante $C \geq 0$ telle que $\forall X, X' \in \mathbb{R}^d$, $\forall \{S_i\}_{0 \leq i \leq n}, \{S'_i\}_{0 \leq i \leq n} \in \mathbb{R}^{d_S \times (n+1)}$

$$|\Phi_{k,k'}(X, (S_i)_{k \leq i \leq k'}) - \Phi_{k,k'}(X', (S'_i)_{k \leq i \leq k'})| \leq C|X - X'| + C \sup_{k \leq i \leq k'} |S_i - S'_i|, \quad \forall 0 \leq k \leq k' \leq n.$$

On note $\Phi_k(X_k, S_k, S_{k+1}) = \Phi_{k,k+1}(X_k, S_k, S_{k+1})$.

On suppose que les fonctions b, σ et β sont bornées et Lispchitziennes, alors le corollaire (3.3.2) implique que X a un flot $\mathbb{L}_2$-lipschitzien :

$$\mathbb{E}\left[\sup_{k \leq i \leq n} \left|X_i^{k,x} - X_i^{k,\tilde{x}}\right|^2\right] \leq C_p |x - \tilde{x}|^2, \quad \forall x, \tilde{x} \in \mathbb{R}^d.$$

On suppose maintenant que $A_k = \left(A_k^{(S)}, 0\right)$ avec $A_k^{(S)} \subset \mathbb{R}^{d_S}$ compact (en finance, cela revient à dire que l'on peut acheter ou vendre dynamiquement que de l'actif S et non tout X).

Le lemme suivant donne la stabilité de l'algorithme sous $\mathbb{L}_p$.

LEMMA 3.6.1. *Pour $p \in]1, +\infty]$, l'hypothèse (H_p) est vérifiée.*

La proposition suivante donne la régularité de l'approximation y_k et permet d'obtenir la vitesse de convergence des algorithmes par quantification optimale :

PROPOSITION 3.6.2. *Les fonctions y_k et $u_{k,a}$ sont bornées et lipschitziennes uniformément en k, n et a.*

La propriété lipschitzienne de la solution y donnée par la proposition précédente permet de majorer l'erreur du schéma par quantification :

$$\varepsilon(\text{quantif}) \leq C \sum_{k=0}^{n-1} \left|\left|X_k - \hat{X}_k\right|\right|_p$$

avec la constante C indépendante de n. On obtient finalement à l'aide du théorème de Zador

$$\varepsilon(\text{quantif}) \leq C_{d,p,\eta} ||X_k||_{p+\eta} \sum_{jk=0}^{n-1} N_k^{-\frac{1}{d}}.$$

3.7. Optimisation d'un processus de Markov contrôlé

3.7.1. Introduction. On considère un processus adapté de contrôle $a : \Omega \times \{0, .., n-1\} \to A$. On note $\mathcal{A}$ un ensemble de contrôle admissible. Soit un contrôle donné $a = (a_0, a_1, ..., a_{n-1}) \in \mathcal{A}$, on considère un processus de Markov contrôlé X à valeurs dans $\mathbb{R}^d$ et défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P}^a)$ avec $X_0 = x_0$ et un noyau de transition de type suivant :

$$X_{k+1} = F_k(X_k, \xi_{k+1}, a_k),$$

i.e.

$$\mathbb{P}^a\left(X_{k+1} \in dy | X_k = x\right) = \mathbb{P}_k^{a_k}(x, dy).$$

Ainsi, on suppose que la distribution de X_{k+1} conditionnée à $\mathcal{F}_k$ est gouvernée par un noyau de transition $\mathbb{P}^{a_k}(X_r, dy)$ (sur un pas de temps) qui est à son tour contrôlé par a_k. Dans ce cadre, nous considérons le problème de contrôle optimal suivant :

$$y_k(X_k) = \sup_{a \in \mathcal{A}} \mathbb{E}^a \left[\sum_{k \le i < n} \frac{1}{n} f_i(X_i, a) + h(X_n) | X_k \right]$$

On suppose maintenant qu'il existe une mesure de référence $\mathbb{P}^*$ équivalente à $\mathbb{P}^a$ telle que

$$\mathbb{E}^a\left[G(X_{k+1})|X_k\right] = \mathbb{E}^*\left[G(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|X_k\right], \quad a \in A,$$

avec $\mathbb{P}_k^*(x, dy) = \mathbb{P}^*(X_{k+1} \in dy | X_k = x)$ et la fonction $\phi_{k,a}(x, y)$ satisfait $\phi_k \ge 0$ et $\mathbb{E}^*\left[\phi_{k,a}(X_k, \xi_{k+1})|X_k\right] = 1$.

Alors le principe de programmation dynamique générique (2.2.1) devient :

$$\begin{aligned}
y_n(x) &= h(x), \\
y_k(x) &= \sup_{a \in \mathcal{A}} \mathbb{E}^a \left[\frac{f_k(X_k, a)}{n} + y_{k+1}(X_{k+1}) | X_k = x \right] \\
&= \sup_{a \in A} \mathbb{E}^* \left[\frac{f_k(X_k, a)}{n} + y_{k+1}(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1}) | X_k = x \right], \quad \forall 0 \le k \le n-1, \forall x \in \mathbb{R}^d.
\end{aligned}$$

ce qui donne dans la formulation du schéma (2.2.1) : $T_{k,a}(x, u) = \frac{f_k(x,a)}{n} + u$ et $R_k(x, v) = \sup_{a \in A} v(a)$.

Comme $\mathbb{E}^*\left[|\phi_{k,a}(X_k, \xi_{k+1})|\,|X_k\right] = 1$, l'hypothèse (H_∞) est vérifiée si on suppose la fonction f_k lipschitzienne.

3.7.2. Exemple. On cherche à approcher la solution de l'inégalité variationnelle du type

$$\begin{aligned}
\max\left\{\sup_{a \in A}(\mathcal{L}^a y + f(t, x, y, a)), y - g\right\} = 0, \\
y(T, x) = h(x), \quad \forall x \in \mathbb{R}^d,
\end{aligned}$$

avec $h \ge g$ et où $\mathcal{L}^a$ est le générateur infinitésimal d'une diffusion avec saut :

$$\mathcal{L}^a u = \partial_t u + \mu_a(x)\partial_x u + \frac{1}{2}\mathrm{Tr}(\partial_{xx} u \sigma_a(x)\sigma_a^T(x)) + \sum_{e \in E}\left(u(t, x + \beta(x, e)) - u(t, x)\right)\nu_a(e)$$

pour $u \in \mathcal{C}_b^2$ et avec E un ensemble fini. On suppose que $\sigma_a \sigma_a^T \ge \epsilon I_d \ \forall a \in A$ et que les fonctions f et g sont lipschitziennes.

En prenant X_k défini par le schéma d'Euler associé au générateur infinitésimal $\mathcal{L}^a$ (avec un pas de temps $\frac{T}{n} \leq \inf_{a,e} \frac{1}{2\nu_a(e)}$) :

$$X_{k+1} = X_k + \mu_a(X_k)\frac{T}{n} + \sigma_a(X_k)\Delta W^a_{k+1} + \sum_{e\in E} \beta(X_k, e) b^{(e)}_{k+1,a}$$

avec $\Delta W^a_{k+1} \sim \mathcal{N}(0, \frac{T}{n} I_d)$ sous $\mathbb{P}^a$ et $b^{(e)}_{k+1,a} \in \mathcal{F}_{k+1}$ indépendantes de $\mathcal{F}_k$, de loi de Bernouilli $\mathbb{P}^a(b^{(e)}_{k+1,a} = 1) = \nu_a(e)\frac{T}{n} = 1 - \mathbb{P}^a(b^{(e)}_{k+1,a} = 0)$.
Définissons $\nu_* : E \to \mathbb{R}$ par $\nu_*(e) = \sup_{a\in A} \nu_a(e)$ et prenons σ_* vérifiant $\sigma_*(x)\sigma^T_*(x) \geq \sigma_a(x)\sigma^T_a(x)$, $\forall x \in \mathbb{R}^d, \forall a \in A$. Soit $b^{(e)}_{k+1} \in \mathcal{F}_{k+1}$ indépendantes de $\mathcal{F}_k$, de loi de Bernouilli $\mathbb{P}^*(b^{(e)}_{k+1,a} = 1) = \nu_a(e)\frac{T}{n} = 1 - \mathbb{P}^*(b^{(e)}_{k+1,a} = 0)$.

Sous $\mathbb{P}^*$, X suit la dynamique

$$X_{k+1} = X_k + \mu_*(X_k)\frac{T}{n} + \sigma_*(X_k)\Delta W_{k+1} + \sum_{e\in E} \beta(X_k, e) b^{(e)}_{k+1}$$

avec $\Delta W_{k+1} \sim \mathcal{N}(0, \frac{T}{n} I_d)$ sous $\mathbb{P}^*$. Les propriétés de σ_* et ν_* ont été choisies pour que X soit plus diffus sous $\mathbb{P}^*$ que sous $\mathbb{P}^a$, ce qui va rendre l'algorithme plus stable (en norme $\mathbb{L}^2$).

Supposons qu'il existe une fonction $\phi \geq 0$ et ξ_{k+1} v.a. telle que

$$\begin{aligned} \mathbb{E}^a\left[G(X_{k+1})|X_k\right] &= \mathbb{E}^*\left[G(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|X_k\right], \\ \mathbb{E}^*\left[\phi_{k,a}(X_k, \xi_{k+1})|X_k\right] &= 1. \end{aligned}$$

Nous pouvons alors définir notre approximation de la solution de l'inégalité variationnelle par programmation dynamique rétrograde générique (2.2.1) :

$$\begin{aligned} y_n(x) &= h(x), \\ v_k(x,a) &= T_{k,a}\left(x, \mathbb{E}^*\left[G(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|X_k = x\right]\right), \\ y_k(x) &= \max\left\{\sup_{a\in A} v_k(x,a); g(x)\right\}, \end{aligned}$$

avec $T_{k,a}(x,u) = u + \frac{T}{n} f\left(\frac{kT}{n}, x, u, a\right)$.

La convergence du problème de contrôle en temps discret au cas continu pourrait être prouvé par des arguments probabilistes (voir [**KD85**]) ou par une approche avec des solutions de viscosité (voir [**BS91**]). On suppose maintenant que la discrétisation en temps converge, on ne va s'intéresser qu'à la discrétisation en espace par quantification.

Le lemme suivant précise la forme de $\phi_{k,a}$ pour le processus de Markov contrôlé.

LEMMA 3.7.1. *On a*

$$\phi_{k,a}(X_k, \xi_{k+1}) = Q_{k,a}(X_k, W_{\frac{k+1}{n}T} - W_{\frac{k}{n}T}) \prod_{e\in E} \phi^{(e)}_a(b^{(e)}_{k+1})$$

avec

$$\xi_{k+1} = \left(W_{\frac{k+1}{n}T} - W_{\frac{k}{n}T}, \left(b^{(e)}_{k+1}\right)_{e\in E}\right)$$

et

$$\phi_a^{(e)}(b) = b\frac{1-\nu_a(e)\frac{T}{n}}{1-\nu_*(e)\frac{T}{n}} + (1-b)\frac{\nu_a(e)}{\nu_*(e)},$$

$$Q_{k,a}(x,w) = \frac{\left|\det(\sigma_*\sigma_*^T(x))\right|^{\frac{1}{2}}}{|\det(\sigma_a\sigma_a^T(x))|^{\frac{1}{2}}} e^{\frac{|w|^2 - \left|\sigma_a^{-1}(x)\sigma_*(x)w + \frac{T}{n}(\mu_a-\mu_*)(x)\right|^2}{2\frac{T}{n}}}.$$

La convergence du schéma par quantification nécessite les hypothèses $(\mathcal{P}_p)$ ou (H_p) qui sont vérifiées par le lemme suivant.

LEMMA 3.7.2. *Les hypothèses $(\mathcal{P}_p)$ et (H_∞) sont validées pour $p \in]1, +\infty]$.*
S'il n'y a pas de contrôle sur σ_a (i.e. $\sigma_* = \sigma_a$), l'hypothèse (H_p) est vérifiée pour $p \in]1, \infty]$.

Pour obtenir la vitesse de convergence des algorithmes, on a besoin de connaître la régularité de y_k donnée par la proposition suivante :

PROPOSITION 3.7.3. *Les fonctions y_k et $u_{k,a}$ sont lipschitziennes uniformément par rapport à k et n.*

La propriété lipschitzienne de la solution y donnée par la proposition précédente permet de majorer l'erreur du schéma par quantification :

$$\varepsilon(\text{quantif}) \leq C \sum_{k=0}^{n-1} \left|\left| X_k - \hat{X}_k \right|\right|_p$$

avec la constante C indépendante de n. On obtient finalement à l'aide du théorème de Zador

$$\varepsilon(\text{quantif}) \leq C_{d,p,\eta} \left|\left|X_k\right|\right|_{p+\eta} \sum_{jk=0}^{n-1} N_k^{-\frac{1}{d}}.$$

REMARK 3.7.4. Cette méthode de résolution de problèmes d'optimisation d'un processus de Markov contrôlé a de nombreuses applications : problèmes d'investissement irréversible, modèles de versement de dividendes d'une firme dans [**CTZ03**], valorisation d'options avec des paramètres incertains (surréplication en volatilité incertaine σ_a dans [**ALP95**], maximisation du prix par rapport à un ensemble de probabilités martingales en marché incomplet).

3.8. Option americaine en information partielle

3.8.1. Introduction. Nous allons maintenant nous intéresser à un domaine de contrôle en finance : les problèmes avec observation partielle (voir par exemple [**KB61**], [**Ben92**]). Plus précisément, nous souhaitons évaluer une option américaine en information partielle. Des schémas numériques ont été proposé par quantification (voir [**PP05a**], [**PRS05**]) et par une méthode de filtre à particules et regression (voir [**Lud09**]). Nous nous placerons dans un cadre plus général avec une démonstration différente de la convergence de notre schéma (reposant sur le flot du processus observable).

On considère un temps discret $\{t_0 = 0 < t_1 < .. < t_n = T\}$ (avec $t_{i+1} - t_i \leq \frac{C}{n}$), un processus partiellement observé (V, S) où V représente le processus d'état ou de signal qui peut ne pas être observé, tandis que S représente l'observation. Le processus signal $\{V_k\}_{0\leq k\leq n}$ est une chaîne de Markov à valeurs dans

$E = \{v_1, .., v_m\}$ un ensemble discret ($\sharp(E) = m$). L'observation S_k à valeurs dans $\mathbb{R}^{d_S}$ est telle que la paire (V, S) est une chaîne de Markov. La loi de S_k conditionnellement à (V_{k-1}, S_{k-1}, V_k) notée $q_k(V_{k-1}, S_{k-1}, V_k, ds)$ admet une densité $g_k(V_{k-1}, S_{k-1}, V_k, s)$.

Notons $(\mathcal{F}_k^S)$ la filtration engendrée par l'observation S et Π_k la loi de V_k conditionnellement au filtre $\mathcal{F}_k^S$

$$\Pi_k^i = \mathbb{P}\left(V_k = v_i | \mathcal{F}_k^S\right), \quad \forall 1 \leq i \leq m.$$

Π_k peut être identifiée à un vecteur aléatoire à valeurs dans le simplexe K_m de $\mathbb{R}^m$ de dimension $m-1$:

$$K_m = \left\{ \pi = (\pi^i) \in \mathbb{R}^m | \pi^i \geq 0, 1 \leq i \leq m; |\pi|_1 = \sum_{i=1}^m \pi^i = 1 \right\}.$$

Le couple $(\Pi_k, S_k)_{0 \leq k \leq n}$ est une $\left(\mathbb{P}, \left(\mathcal{F}_k^S\right)_k\right)$-chaîne de Markov.

Soit μ la loi discrète initiale. La formule de Bayes donne la loi de Π à valeurs dans $\mathbb{R}^m$:

$$\begin{aligned}
\Pi_0^i &= \mathbb{P}(V_0 = v_i) \\
&= \mu^i, \quad \forall 1 \leq i \leq m. \\
\Pi_k &= \left\{\mathbb{P}\left(V_k = v_i | \Pi_{k-1}, S_{k-1}, S_k\right)\right\}_{1 \leq i \leq m} \\
&= \left\{ \frac{\mathbb{P}\left(\{V_k = v_i\} \bigcap \Pi_{k-1} | S_{k-1}, S_k\right)}{\mathbb{P}\left(\Pi_{k-1} | S_{k-1}, S_k\right)} \right\}_{1 \leq i \leq m} \\
&= \frac{GP_k(S_{k-1}, S_k)^T \Pi_{k-1}}{|GP_k(S_{k-1}, S_k)^T \Pi_{k-1}|_1}.
\end{aligned}$$

où $GP_k(S_{k-1}, S_k)$ est une matrice aléatoire de taille $m \times m$ donnée par :

$$GP_k(S_{k-1}, S_k)_{i,j} = g_k(v_i, S_{k-1}, v_j, S_k) \mathbb{P}\left(V_k = v_j | V_{k-1} = v_i\right)$$

On considère P un processus dépendant de la trajectoire de S à valeurs dans $\mathbb{R}^{d_P}$ sous la forme :

$$\begin{aligned}
P_{k'} &= \Phi_{k,k'}\left(P_k, (S_i)_{k \leq i \leq k'}\right) \\
&= \Phi_{k'}\left(P_{k'-1}, S_{k'-1}, S_{k'}\right).
\end{aligned}$$

On note $Z_k = (S_k, P_k) \in \mathbb{R}^d$ avec $d = d_S + d_P$. Par exemple, P pourra représenter une option asiatique ou lookback en finance.

Alors (V_k, S_k, P_k) est une $(\mathbb{P}, (\mathcal{F}_k)_k)$-chaîne de Markov à valeurs dans $E \times \mathbb{R}^{d_S} \times \mathbb{R}^{d_P}$ et $X_k = (\Pi_k, S_k, P_k)$ est une $\left(\mathbb{P}, \left(\mathcal{F}_k^S\right)_k\right)$-chaîne de Markov à valeurs dans $K_m \times \mathbb{R}^{d_S} \times \mathbb{R}^{d_P}$.

Simulation de la $\mathcal{F}_k^S$-chaîne de Markov $(X_k)_k$. Pour $k = 0$, X_0 est un vecteur déterministe valant $X_0 = (\mu, S_0, P_0)$. Pour $k \geq 1$, partant de $X_{k-1} = (\Pi_{k-1}, S_{k-1}, P_{k-1})$,

$\diamond$ on simule V_{k-1} selon la loi de probabilité Π_{k-1}, puis on simule V_k sachant V_{k-1} ;

$\diamond$ on simule S_k selon la probabilité la probabilité de transition $q_k(V_{k-1}, S_{k-1}, V_k, ds)$, puis on calcule $P_k = \Phi_{k-1}(P_{k-1}, S_{k-1}, S_k)$;

$\diamond$ on calcule Π_k selon la formule

$$\Pi_k = \frac{GP_k(S_{k-1}, S_k)^T \Pi_{k-1}}{|GP_k(S_{k-1}, S_k)^T \Pi_{k-1}|_1}.$$

3.8.2. Valorisation d'une option américaine en information partielle. Ici, on peut considérer par exemple que S représente les actifs risqués et V le processus de rendement et/ou la volatilité de l'actif.

On cherche maintenant à résoudre le problème de temps d'arrêt optimal sous observation partielle :

$$u_0 = \sup_{\tau \in \mathcal{T}_{0,n}^S} \mathbb{E}\left[\sum_{0\leq k<\tau} \frac{f_k(V_k, Z_k)}{n} + h_\tau(V_\tau, Z_\tau)\right].$$

où $\mathcal{T}_{k,n}^S$ est l'ensemble des temps d'arrêt adapté à la filtration observable $\mathcal{F}_u^S = \sigma\left(S_i, 1 \leq i \leq u\right) = \sigma\left(Z_i, 1 \leq i \leq u\right)$ à valeurs dans $\{k, ..., n\}\}$.

Introduisons les fonctions :

$$\begin{aligned}
\hat{f}_k(\pi, z) &= \sum_{i=1}^m f_k(v_i, z)\pi^i \\
&= F_k(z).\pi, \\
\hat{h}_k(\pi, z) &= \sum_{i=1}^m h_k(v_i, z)\pi^i \\
&= H_k(z).\pi,
\end{aligned}$$

où $F_k(z) = \left(f_k(v_i, z)\right)_{1\leq i\leq m}$ et $H_k(z) = \left(h_k(v_i, z)\right)_{1\leq i\leq m}$. Or, pour $\tau \in \mathcal{T}_{0,n}^Y$, on a

$$\begin{aligned}
\mathbb{E}\left[\sum_{k=0}^{\tau} \frac{f_k(V_k, Z_k)}{n} + h_\tau(V_\tau, Z_\tau)\right] &= \mathbb{E}\left[\sum_{j=0}^n 1_{\tau=j}\left\{\sum_{0\leq k<j} \frac{f_k(V_k, Z_k)}{n} + h_\tau(V_\tau, Z_\tau)\right\}\right] \\
&= \mathbb{E}\left[\sum_{j=0}^n 1_{\tau=j}\left\{\sum_{0\leq k<j} \mathbb{E}\left[\frac{f_k(V_k, Z_k)}{n}|\mathcal{F}_k^S\right] + \mathbb{E}\left[h_j(V_j, Z_j)|\mathcal{F}_j^S\right]\right\}\right] \\
&= \mathbb{E}\left[\sum_{j=0}^n 1_{\tau=j}\left\{\sum_{0\leq k<j} \frac{\hat{f}_k(\Pi_k, Z_k)}{n} + \hat{h}_j(\Pi_j, Z_j)\right\}\right] \\
&= \mathbb{E}\left[\sum_{0\leq k<\tau}^{\tau} \frac{\hat{f}_k(X_k)}{n} + \hat{h}_\tau(X_\tau)\right],
\end{aligned}$$

et par conséquent

$$u_0 = \sup_{\tau \in \mathcal{T}_{0,n}^S} \mathbb{E}\left[\sum_{0\leq k<\tau}^{\tau} \frac{\hat{f}_k(X_k)}{n} + \hat{h}_\tau(X_\tau)\right].$$

Comme X est une $\mathcal{F}^S$-chaîne de Markov, il s'agit maintenant de résoudre un problème d'arrêt optimal en information complète. Pour $0 \leq k \leq n$, on considère les

fonctions $y_k : K_m \times \mathbb{R}^d \to \mathbb{R}$ définies par :

$$y_k(x) = \sup_{\tau \in \mathcal{T}_{k,n}^S} \mathbb{E}\left[\sum_{0\le k<\tau} \frac{f_k(V_k, Z_k)}{n} + h_\tau(V_\tau, Z_\tau)|X_k = x\right]$$
$$= \sup_{\tau \in \mathcal{T}_{k,n}^X} \mathbb{E}\left[\sum_{k\le i<\tau} \frac{\hat{f}_i(X_i^{k,x})}{n} + \hat{h}_\tau(X_\tau^{k,x})\right].$$

Par le principe de programmation dynamique, on a $u_0 = y_0(X_0)$ (avec les mêmes notations que sur la partie sur la convergence des algorithmes par quantification) avec la fonction $y_k : K_m \times \mathbb{R}^d \to \mathbb{R}$ définie par la formule rétrograde :

$$y_n(x) = \hat{h}_n(x), \quad \forall x \in K_m \times \mathbb{R}^d,$$
$$y_k(x) = \sup\left\{\hat{h}_k(x); \frac{\hat{f}_k(x)}{n} + \mathbb{E}\left[y_{k+1}(X_{k+1})|X_k = x\right]\right\}, \quad \forall x \in K_m \times \mathbb{R}^d,$$
$$= \sup_{\tau \in \mathcal{T}_{k,n}^S} \mathbb{E}\left[\sum_{k\le i<\tau}^{\tau} \frac{\hat{f}_k(X_i^{k,x})}{n} + \hat{h}_\tau(X_\tau^{k,x})\right].$$

On a réussi à formuler l'algorithme dans le cadre de la programmation dynamique rétrograde (2.2.1) avec $h = h_n$, $T_{k,a} = I_d$ et

$$R_k(x, v) = \sup\left\{\hat{h}_k(x); \frac{\hat{f}_k(x)}{n} + v\right\}, \quad \forall x \in \mathbb{E}^d, \forall v \in \mathbb{R}.$$

LEMMA 3.8.1. *Les fonctions $\hat{f}_k$ et $\hat{h}_k$ sont bornées par K et lipschitziennes de rapport $2K$.*

Les fonctions $\hat{f}_k$ et $\hat{g}_k$ étant uniformément lipschitziennes, l'hypothèse (H_1) est trivialement vérifiée. Pour obtenir la convergence des algorithmes par quantification, il suffit de montrer que y_k est uniformément lipschitzienne. Pour cela, nous avons besoin de régularité de la chaîne de Markov X.

ASSUMPTION 3.8.2. *On suppose que les fonctions f_k et h_k sont bornées lipschitziennes, que S a un flot lipschitzien et la fonction Φ est lipschitzienne. Plus précisément, on suppose qu'il existe une constante $K \ge 0$ telle que $\forall k \in \{0, .., n-1\}$, $\forall v \in E$, $\forall \{s_i\}_{0\le i\le n}, \{\overline{s}_i\}_{0\le i\le n} \in \mathbb{R}^{d_S}$, $\forall p, \overline{p} \in \mathbb{R}^{d_P}$,*

$$\sup_k \left\{|f_k|_\infty, |h_k|, [f_k]_{Lip}, [g_k]_{Lip}\right\} \le K,$$
$$|\Phi_{k,k'}(p, (s_i)_{k\le i\le k'}) - \Phi_{k,k'}(\overline{p}, (\overline{s}_i)_{k\le i\le k'})| \le K\left\{|p - \overline{p}| + \sup_{k\le i\le k'} |s_i - \overline{s}_i|\right\}$$
$$\mathbb{E}\left[\sup_{k\le i\le n}\left|S_i^{k,(v,s_k)} - S_i^{k,(v,\overline{s}_k)}\right|\right] \le K|s_k - \overline{s_k}|.$$

LEMMA 3.8.3. *Sous cette hypothèse, le flot de Z est lipschitzien, c'est-à-dire il existe une constante $C > 0$ telle que $\forall 0 \le k \le n$, $\forall v \in E$, $\forall z, z' \in \mathbb{R}^d$*

$$\mathbb{E}\left[\sup_{k\le k'\le n}\left|Z_{k'}^{k,(v,z)} - Z_{k'}^{k,(v,\tilde{z})}\right|\right] \le C|z - \tilde{z}|, \quad \forall z, \tilde{z} \in \mathbb{R}^d.$$

REMARK 3.8.4. Cette dernière inégalité est vérifiée si, par exemple, S est le schéma d'Euler d'une diffusion avec saut :

$$S_{k+1} = S_k + b(S_k, V_k, V_{k+1})\frac{T}{n} + \sigma(S_k, V_k)(W_{(k+1)\frac{T}{n}} - W_{\frac{kT}{n}}) + \int_A \beta(S_k, a, V_k)\mu_S\left(da, (k\frac{T}{n}, (k+1)\frac{T}{n}]\right),$$

$$V_{k+1} = T_k\left(V_k, (W_s)_{\frac{kT}{n}\le s\le (k+1)\frac{T}{n}}, (\mu_S(da, ds))_{a\in A, \frac{kT}{n}<s\le (k+1)\frac{T}{n}}, \epsilon_{k+1}\right),$$

où les fonctions b, σ et β sont bornées et lipschitziennes par rapport à la variable S_k.

PROPOSITION 3.8.5. *y_k est bornée et lipschitzienne avec un rapport de Lipschitz indépendant de k et n.*

La propriété lipschitzienne de la solution y donnée par la proposition précédente permet de majorer l'erreur du schéma par quantification :

$$\varepsilon(\text{quantif}) \le C\sum_{k=0}^{n-1}\left|\left|X_k - \hat{X}_k\right|\right|_p$$

avec la constante C indépendante de n.

REMARK 3.8.6. L'article [**PRS05**] donne une vitesse de convergence en exponentielle du nombre de pas de temps n. Ici, en montrant que y_k est lipschitzienne uniformément par rapport à k et n, on obtient une vitesse de convergence linéaire par rapport à n. On a utilisé la propriété de Lipschiz du flot de Z sans avoir d'hypothèses contraignantes sur la densité g_k de la loi de S_k conditionnelle à (V_{k-1}, S_{k-1}, V_k) ; nous avons juste besoin de l'existence pour pouvoir simuler le couple (Π_k, S_k). A l'aide d'hypothèses assez fortes sur la densité g_k, l'article [**PRS05**] montre que la fonction valeur y_k est lipschitzienne avec un rapport de Lipschitz en exponentielle de n.

3.9. Preuves

3.9.1. Options bermudéennes. La proposition (3.1.7) donne le rapport de Lipschitz de y_k et $u_{k,a}$ pour le schéma avec $\phi = 1$ quand la probabilité de transition de X est lipschitzienne .

DÉMONSTRATION. Supposons que la probabilité de la chaîne de Markov X est lipschitzienne.

Initialement, on a $[y_n]_{\text{Lip}} = [h]_{\text{Lip}} \le K$.

Pour $k < n$, comme on considére $\phi_{k,a} = 1$ ($u_{k,a}$ est indépendant de a), on a $u_k = P_k^X y_{k+1}$ avec le rapport de Lipschitz

$$[u_k]_{\text{Lip}} \le \left[P^X\right]_{\text{Lip}}[y_{k+1}]_{\text{Lip}}\,.$$

Comme les fonctions T_k et R_k sont lipschitziennes de rapport K et $y_k(x) = R_k\left(x, T_k(x, u_k(x))\right)$ $\forall x \in \mathbb{R}^d$, on a

$$\begin{aligned}[y_k]_{\text{Lip}} &\le [R_k]_{\text{Lip}} + [R_k]_{\text{Lip}}[T_k]_{\text{Lip}} + [R_k]_{\text{Lip}}[u_k]_{\text{Lip}} \\ &\le (K + K^2) + K\left[P^X\right]_{\text{Lip}}[y_{k+1}]_{\text{Lip}}\,.\end{aligned}$$

En sommant, on obtient finalement

$$[y_k]_{\text{Lip}} \le (K + K^2)\sum_{i=k}^{n-1} K^{i-k}\left[P^X\right]_{\text{Lip}}^{i-k}$$

et donc

$$\max_k\{[y_k]_{\mathrm{Lip}}\,,[u_k]_{\mathrm{Lip}}\}\leq C\frac{K^n\left[P^X\right]_{\mathrm{Lip}}^n-1}{K\left[P^X\right]_{\mathrm{Lip}}-1}.$$

□

La proposition (3.2.1) donne le rapport de Lipschitz de y_k et $u_{k,a}$ pour le problème d'enveloppe de Snell.

Démonstration. Supposons la probabilité de transition de X lipschitzienne. Pour $k=n$, $y_n=g_n$ est lipschitzien de rapport K. Pour $k<n$, comme $u_k=P_k^Xy_{k+1}$, on a :

$$[u_k]_{\mathrm{Lip}}=\left[P_k^X\right]_{\mathrm{Lip}}[y_{k+1}]_{\mathrm{Lip}}\,.$$

L'inégalité $\left|\max(a,b)-\max(\tilde{a},\tilde{b})\right|\leq\max(|a-\tilde{a}|\,,\left|b-\tilde{b}\right|)$ et la lipschtziannité de g_k de rapport K impliquent

$$\begin{aligned}|y_k(x)-y_k(\tilde{x})|&=|\max\left(g_k(x),u_k(x)\right)-\max\left(g_k(\tilde{x}),u_k(\tilde{x})\right)|\,,\quad\forall x,\tilde{x}\in\mathbb{R}^d,\\&\leq\max\left(|g_k(x)-g_k(\tilde{x})|\,,|u_k(x)-u_k(\tilde{x})|\right),\quad\forall x,\tilde{x}\in\mathbb{R}^d,\\&\leq\max\left([g_k]_{\mathrm{Lip}}\,|x-\tilde{x}|\,,\left[P_k^X\right]_{\mathrm{Lip}}[y_{k+1}]_{\mathrm{Lip}}\,|x-\tilde{x}|\right),\quad\forall x,\tilde{x}\in\mathbb{R}^d.\end{aligned}$$

D'où $[y_k]_{\mathrm{Lip}}\leq\max\left(K,\left[P_k^X\right]_{\mathrm{Lip}}[y_{k+1}]_{\mathrm{Lip}}\right)$ et on obtient finalement de manière rétrograde

$$\begin{aligned}[y_k]_{\mathrm{Lip}}&\leq\max\left(K,\prod_{i=k}^{n-1}\left[P_i^X\right]_{\mathrm{Lip}}[y_n]_{\mathrm{Lip}}\right)\\&\leq K\left[P^X\right]_{\mathrm{Lip}}^{n-k-1}.\end{aligned}$$

Supposons de plus que la chaîne de Markov X a un flot $\mathbb{L}_1$-lipschitzien. Alors $\{y_k\}_k$ et $\{u_k\}_k$ sont lipschitziennes de rapport $K\left[F^X\right]_{\mathrm{Lip}}$. En effet, on a

$$\begin{aligned}|y_k(x)-y_k(\tilde{x})|&=\left|\sup_{\tau\in\mathcal{T}_{k,n}}\mathbb{E}\left[g_\tau(X_\tau^{k,x})\right]-\sup_{\tau\in\mathcal{T}_{k,n}}\mathbb{E}\left[g_\tau(X_\tau^{k,\tilde{x}})\right]\right|\\&\leq\left|\sup_{\tau\in\mathcal{T}_{k,n}}\mathbb{E}\left[g_\tau(X_\tau^{k,x})-g_\tau(X_\tau^{k,\tilde{x}})\right]\right|\\&\leq\sup_{k\leq i\leq n}[g_i]_{\mathrm{Lip}}\,\mathbb{E}\left[\sup_{k\leq i\leq n}\left|X_i^{k,x}-X_i^{k,\tilde{x}}\right|\right]\\&\leq K\left[F^X\right]_{\mathrm{Lip}}|x-\tilde{x}|\,,\quad\forall x,\tilde{x}\in\mathbb{R}^d,\\|u_k(x)-u_k(\tilde{x})|&=\left|\sup_{\tau\in\mathcal{T}_{k+1,n}}\mathbb{E}\left[g_\tau(X_\tau^{k,x})\right]-\sup_{\tau\in\mathcal{T}_{k+1,n}}\mathbb{E}\left[g_\tau(X_\tau^{k,\tilde{x}})\right]\right|\\&\leq K\left[F^X\right]_{\mathrm{Lip}}|x-\tilde{x}|\,,\quad\forall x,\tilde{x}\in\mathbb{R}^d.\end{aligned}$$

□

3.9.2. EDSR. Le lemme (3.3.1) permet de contrôler les diffusions avec sauts sur un pas de temps.

DÉMONSTRATION. Soit $k \in \{0, .., n-1\}$. Soit $p > 0$, soit $x \in \mathbb{R}^{\tilde{d}}$ tel que $|x| = 1$ et soient $\tilde{b} \in \mathbb{R}^{\tilde{d}}$, $\tilde{\sigma} \in \mathbb{R}^{\tilde{d} \times d_W}$ et $\tilde{\beta} \in \mathbb{R}^{\tilde{d} \times d_E}$ bornés par M. En notant

$$A_k = \tilde{b}\frac{T}{n} + \tilde{\sigma}(W_{t_{k+1}} - W_{t_k}) + \int_E \tilde{\beta}(e)\overline{\mu}(de, (t_k, t_{k+1}]),$$

il faut montrer que $\exists C_{p,M} > 0$ telle que

$$||A_k||_p \leq \frac{C_{p,M}}{n^{\frac{1}{2\vee p}}} \quad \text{et} \quad ||x + A_k||_p \leq 1 + \frac{C_{p,M}}{n}$$

Soit $m \geq 2$, on a

$$\begin{aligned}\mathbb{E}\left[\left|\int_E \tilde{\beta}(e)\overline{\mu}(de, (t_k, t_{k+1}])\right|^m\right] &\leq C_m(t_{k+1} - t_k)\mathbb{E}\left[\int_E \left|\tilde{\beta}(e)\right|^m \lambda(de)\right] \\ &\leq C_m K^m \lambda(E)(t_{k+1} - t_k).\end{aligned}$$

A l'aide de l'inégalité de Jensen, on obtient l'existence d'une constante $C'_m > 0$ telle que

$$\begin{aligned}\left|\left|\int_E \tilde{\beta}(e)\overline{\mu}(de, (t_k, t_{k+1}])\right|\right|_p &\leq \left|\left|\int_E \tilde{\beta}(e)\overline{\mu}(de, (t_k, t_{k+1}])\right|\right|_{2\vee p} \\ &\leq \frac{C'_m K}{n^{\frac{1}{2\vee p}}}.\end{aligned}$$

Comme $\left|\left|W_{t_{k+1}} - W_{t_k}\right|\right|_p \leq \frac{C_p}{\sqrt{n}}$, on obtient finalement

$$\begin{aligned}||A_k||_p &\leq C\left|\left|\tilde{b}\frac{T}{n}\right|\right|_p + C\left|\tilde{\sigma}\right|\left|\left|W_{t_{k+1}} - W_{t_k}\right|\right|_p + C\left|\left|\int_E \tilde{\beta}(e)\overline{\mu}(de, (t_k, t_{k+1}])\right|\right|_p \\ &\leq \frac{C_{p,M}}{n^{\frac{1}{2\vee p}}}.\end{aligned}$$

Notons $\overline{p} = \inf\{i \in 2\mathbb{N} | i > p\}$ ($\overline{p}$ est paire et $\overline{p} \geq 2$). En appliquant la formule du binôme de Newton, on obtient

$$\begin{aligned}\mathbb{E}\left[|x + A_k|^{\overline{p}}\right] &= \mathbb{E}\left[\left(1 + x.A_k + |A_k|^2\right)^{\frac{\overline{p}}{2}}\right] \\ &= 1 + \frac{\overline{p}}{2}\mathbb{E}\left[x.A_k + |A_k|^2\right] + \sum_{j=2}^{\frac{\overline{p}}{2}} \frac{\overline{p}!}{(\overline{p} - j)!j!}\mathbb{E}\left[\left|x.A_k + |A_k|^2\right|^j\right] \\ &\leq 1 + \frac{\overline{p}}{2}\left|\mathbb{E}\left[x.A_k\right]\right| + C_{\overline{p}}\sum_{j=2}^{\overline{p}} \mathbb{E}\left[|A_k|^j\right].\end{aligned}$$

Comme

$$\max\left\{\left|\mathbb{E}\left[x.A_k\right]\right|; \mathbb{E}\left[|A_k|^j\right]\right\} \leq \frac{C_{\overline{p}}}{n}, \quad \forall 2 \leq j \leq \overline{p},$$

on a $\mathbb{E}\left[|x+A_k|^{\overline{p}}\right] \leq 1+\frac{C_{\overline{p}}}{n}$. Finalement, en appliquant l'inégalité de Jensen, on obtient le résultat souhaité :

$$\begin{aligned}||x+A_k||_p &\leq ||x+A_k||_{\overline{p}} \\ &\leq 1+\frac{C_{\overline{p}}}{n}.\end{aligned}$$

□

Le corollaire (3.3.2) montre que le flot de la chaîne de Markov X pour les EDSR avec saut est lipschitzien.

Démonstration. Soient $k \in \{0,..,n-1\}$ et $x, \tilde{x} \in \mathbb{R}^d$ avec $x \neq \tilde{x}$, on a

$$\frac{X_{k+1}^{k,x}-X_{k+1}^{k,\tilde{x}}}{|x-\tilde{x}|} = \frac{x-\tilde{x}}{|x-\tilde{x}|}+\frac{b(x)-b(\tilde{x})}{|x-\tilde{x}|}\frac{T}{n}+\frac{\sigma(x)-\sigma(\tilde{x})}{|x-\tilde{x}|}(W_{t_{k+1}}-W_{t_k})+\int_E \frac{\beta(x,e)-\beta(\tilde{x},e)}{|x-\tilde{x}|}\overline{\mu}(de,(t_k,t_{k+1}])$$

Comme les fonctions b, σ et β sont lipschitziennes, on peut appliquer le lemme (3.3.1) ce qui donne

$$\left|\left|\frac{X_{k+1}^{k,x}-X_{k+1}^{k,\tilde{x}}}{|x-\tilde{x}|}\right|\right|_p \leq 1+\frac{C_p}{n}.$$

Comme $\left[F_k^X\right]_p \leq 1+\frac{C_p}{n}$, le flot de X est lipschitzien car $\left[F^X\right]_p \leq \left[F_k^X\right]_p^n \leq e^{C_p}$. □

Le corollaire (3.3.3) montre que l'hypothèse (H_p) est vérifiée dans le cas des équations différentielles strochastiques rétrogrades avec sauts.

Démonstration. Soient $p \in]1,+\infty]$ et q son conjugué. Soit $\lambda = (\lambda_y, \lambda_z, \lambda_E) \in \mathbb{R} \times \mathbb{R}^{d_W} \times \mathbb{R}^{d_E}$ appartenant à $\mathcal{S}^{\Delta T_k}$. Comme

$$\phi_k(\xi_{k+1}) = \left(1, W_{t_{k+1}}-W_{t_k}, \int_E \rho(e)\overline{\mu}(de,(t_k,t_{k+1}])\right),$$

$$T_k(x,u) = u_1 + \frac{T}{n}f(x,u^1,\frac{n}{T}u^2,\frac{n}{T}u^3), \quad \forall x \in \mathbb{R}^d, \forall u=(u^1,u^2,u^2) \in \mathbb{R}\times\mathbb{R}^{d_W}\times\mathbb{R}^{d_E},$$

et comme f est lipschitzienne, on a

$$|\lambda_y| \leq 1+\frac{T}{n}[f]_{\text{Lip}} \quad \text{et} \quad \max(|\lambda_z|_\infty, |\lambda_E|_\infty) \leq [f]_{\text{Lip}}$$

En appliquant le lemme (3.3.1) à

$$\lambda.\phi_k(\xi_{k+1}) = \lambda_y + \lambda_z.\left(W_{t_{k+1}}-W_{t_k}\right) + \lambda_E.\int_E \rho(e)\overline{\mu}(de,(t_k,t_{k+1}])$$

on obtient le résultat souhaité :

$$||\lambda.\phi_k(\xi_{k+1})||_q \leq 1+\frac{K}{n}$$

. □

La proposition (3.3.4) donne la régularité des fonctions $\{y_k\}_k$ et $\{u_k\}_k$ dans le cas de la discrétisation des équations différentielles stochastiques rétrogrades.

DÉMONSTRATION. 1/ Montrons que les fonctions $\{y_k\}_k$ sont lipschitziennes. Pour $k = n$, on a $y_n = h$ lipschitzienne. Soient $k < n$ et $x, \tilde{x} \in \mathbb{R}^d$, la propriété lipschitzienne de T_k implique

$$\begin{aligned}|y_k(x) - y_k(\tilde{x})| &\leq \left|T_k(x, \mathbb{E}\left[y_{k+1}(X_{k+1}^{k,x})\phi_k(\xi_{k+1})\right]) - T_k(\tilde{x}, \mathbb{E}\left[y_{k+1}(X_{k+1}^{k,\tilde{x}})\phi_k(\xi_{k+1})\right])\right| \\ &\leq [T_k]_{\text{Lip}}\,|x - \tilde{x}| + |I_k|\end{aligned}$$

où

$$I_k = T_k\left(\tilde{x}, \mathbb{E}\left[y_{k+1}(X_{k+1}^{k,x})\phi_k(\xi_{k+1})\right]\right) - T_k(\tilde{x}, \mathbb{E}\left[y_{k+1}(X_{k+1}^{k,\tilde{x}})\phi_k(\xi_{k+1})\right].$$

Il existe alors $\lambda \in \mathcal{S}^{\Delta T_k}$ telle que

$$\begin{aligned}I_k &= \lambda_k.\left\{\mathbb{E}\left[y_{k+1}(X_{k+1}^{k,x})\phi_k(\xi_{k+1})\right] - \mathbb{E}\left[y_{k+1}(X_{k+1}^{k,\tilde{x}})\phi_k(\xi_{k+1})\right]\right\} \\ &= \mathbb{E}\left[\left\{y_{k+1}(X_{k+1}^{k,x}) - y_{k+1}(X_{k+1}^{k,\tilde{x}})\right\}\lambda_k.\phi_k(\xi_{k+1})\right].\end{aligned}$$

L'inégalité de Cauchy-Schwarz et la vérification de l'hypothèse (H_2) par la proposition (3.3.3) impliquent

$$\begin{aligned}|I_k| &\leq \left|\left|y_{k+1}(X_{k+1}^{k,x}) - y_{k+1}(X_{k+1}^{k,\tilde{x}})\right|\right|_2 ||\lambda_k.\phi_k(\xi_{k+1})||_2 \\ &\leq [y_{k+1}]_{\text{Lip}}\,(1 + \frac{C}{n})\left|\left|X_{k+1}^{k,x} - X_{k+1}^{k,\tilde{x}}\right|\right|_2.\end{aligned}$$

Comme le flot de $\left[F_k^X\right]_2 \leq 1 + \frac{C}{n}$ d'après le corollaire (3.3.2), on obtient

$$\left|\left|X_{k+1}^{k,x} - X_{k+1}^{k,\tilde{x}}\right|\right|_2 \leq \left(1 + \frac{C}{n}\right)|x - \tilde{x}|,$$

et en regroupant, on vérifie l'inégalité

$$[y_k]_{\text{Lip}} \leq \frac{C}{n} + [y_{k+1}]_{\text{Lip}}\,(1 + \frac{C}{n}).$$

En appliquant le lemme de Gronwall, on obtient que les fonctions y_k sont uniformément lipschitziennes. Il en est de même pour u_k d'après l'inégalité suivante : $\forall 0 \leq k < n$, $\forall x, \tilde{x} \in \mathbb{R}^d$,

$$\begin{aligned}|u_k(x) - u_k(\tilde{x})| &\leq \left|\mathbb{E}\left[\{y_{k+1}(X_{k+1}^{k,x}) - y_{k+1}(X_{k+1}^{k,\tilde{x}})\}\left(1, W_{t_{k+1}} - W_{t_k}, \int_E \rho(e)\overline{\mu}(de, (t_k, t_{k+1}])\right)\right]\right| \\ &\leq \left|\left|y_{k+1}(X_{k+1}^{k,x}) - y_{k+1}(X_{k+1}^{k,\tilde{x}})\right|\right|_2 \left|\left|\left(1, W_{t_{k+1}} - W_{t_k}, \int_E \rho(e)\overline{\mu}(de, (t_k, t_{k+1}])\right)\right|\right|_2 \\ &\leq C\,[y_{k+1}]_{\text{Lip}}\left|\left|X_{k+1}^{k,x} - X_{k+1}^{k,\tilde{x}}\right|\right|_2 \\ &\leq C\,[y_{k+1}]_{\text{Lip}}\,|x - \tilde{x}|.\end{aligned}$$

2/ Supposons $b, \sigma, \beta(., e)$ et g appartiennent à $\mathcal{C}_b^m$ pour $m \in \mathbb{N}^*$. Notons

$$L_k^m = \sup_{1 \leq \alpha \leq m}\left|\frac{\partial^\alpha y_k}{\partial x^\alpha}\right|, \quad \forall 0 \leq k \leq n.$$

Par souci de simplification des notations, on se place en dimension $d = 1$.

Montrons par récurrence sur $m \geq 1$ que les fonctions $\{L_k^m\}_{0 \leq k \leq n}$ sont bornées uniformément.

Pour $m = 1$, on vient de montrer en (1) que la fonction $\{y_k\}_k$ est uniformément lipschitzienne, donc L_k^m est bornée $\forall 0 \leq k \leq n$.

Supposons $m \geq 2$ et que l'hypothèse de récurrence est vérifiée $\forall i \in \{1,..,m-1\}$, c'est-à-dire $L_k^i < +\infty \ \forall i \in \{1,..,m-1\}$, $\forall 0 \leq k \leq n$.

Pour $k = n$, on a $y_n = g \in \mathcal{C}_b^m$, donc L_n^m bornée.

Soit $k < n$. Avec la fonction $u_k = (u_{k,1}, u_{k,2}, u_{k,3})$ définie par

$$u_{k,i}(x) = \mathbb{E}\left[y_{k+1}(X_{k+1}^{k,x}) r_{k+1,i}\right]$$

et $r_{k+1} = (r_{k+1,1}, r_{k+1,2}, r_{k+1,3}) = \left(1, \frac{W_{t_{k+1}} - W_{t_k}}{t_{k+1}-t_k}, \frac{\int_E \rho(e)\overline{\mu}(de,(t_k,t_{k+1}])}{t_{k+1}-t_k}\right)$, on peut réécrire la fonction y_k sous la forme

$$y_k(x) = u_{k,1}(x) + \frac{T}{n} f\left(x, u_k(x)\right), \quad \forall x \in \mathbb{R}^d.$$

On note θ_k la fonction $\theta(x) = (x, u_{k,1}(x), u_{k,2}(x), u_{k,3}(x))$.

Pour majorer la dérivée m-ième de y_k, on va d'abord chercher contrôler des dérivées de u_k :

$$\begin{aligned}\frac{\partial^\alpha}{\partial x^\alpha} u_k(x) &= \frac{\partial^\alpha}{\partial x^\alpha} \mathbb{E}\left[y_{k+1}(X_{k+1}^{k,x}) r_{k+1}\right] \\ &= \mathbb{E}\left[\frac{\partial^\alpha}{\partial x^\alpha} y_{k+1}(X_{k+1}^{k,x}) r_{k+1}\right], \quad \forall x \in \mathbb{R}, \forall 1 \leq \alpha \leq m.\end{aligned}$$

Pour $\alpha \geq 1$, on utilise la formule de Faa di Bruno pour la dérivée α-ième de la fonction composée $x \to y_{k+1}(X_{k+1}^{k,x})$ exprimée en fonction des polynômes de Bell $B_{\alpha,i}(x_1, ..., x_{\alpha-i+1})$:

$$\begin{aligned}\frac{\partial^\alpha}{\partial x^\alpha}\left(y_{k+1}(X_{k+1}^{k,x})\right) &= \sum_{i=1}^{\alpha} \frac{\partial^i y_{k+1}}{\partial x^i}(X_{k+1}^{k,x}) B_{\alpha,i}\left(\frac{\partial X_{k+1}^{k,x}}{\partial x}, \frac{\partial^2 X_{k+1}^{k,x}}{\partial x^2}, .., \frac{\partial^{\alpha-i+1} X_{k+1}^{k,x}}{\partial x^{\alpha-i+1}}\right) \\ &= \frac{\partial^\alpha y_{k+1}}{\partial x^\alpha}\left(\frac{\partial X_{k+1}^{k,x}}{\partial x}\right)^\alpha + \sum_{i=1}^{\alpha-1} \frac{\partial^i y_{k+1}}{\partial x^i}(X_{k+1}^{k,x}) B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1 \leq u \leq \alpha-i+1}\right)\end{aligned}$$

où $B_{\alpha,i}(x_1, ..., x_{\alpha-i+1})$ représente les polynomes de Bell donnés par :

$$B_{\alpha,i}(x_1, .., x_{\alpha-i-1}) = \sum \frac{\alpha!}{j_1! j_2! ... j_{\alpha-i+1}!}\left(\frac{x_1}{1!}\right)^{j_1}\left(\frac{x_2}{2!}\right)^{j_2} ... \left(\frac{x_{\alpha-i+1}}{(\alpha-i+1)!}\right)^{\alpha-i+1}$$

où la somme porte sur toutes les suites $j_1, j_2, ..., j_{\alpha-i+1}$ d'entiers tels que $\sum_{u=1}^{\alpha-i+1} j_u = i$ et $\sum_{u=1}^{\alpha-i+1} u j_u = \alpha$.

On remarque que l'on a les inégalités suivantes : soit $\tilde{p} \in \mathbb{N}^*$, soient $\left(\tilde{\epsilon}_i, \tilde{b}_i, \tilde{\sigma}_i, \tilde{\beta}_i\right)_{1 \leq i \leq \tilde{p}} \in \left(\{0;1\} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^E\right)^{\tilde{p}}$ bornés par K :

$$\sup_{1 \leq i \leq \tilde{p}} \left\{\left|\tilde{b}_i\right|, \left|\tilde{\sigma}_i\right|, \left|\tilde{\beta}_i\right|_\infty\right\} \leq K$$

alors en posant

$$P_i = \tilde{\epsilon}_i + \tilde{b}_i h + \tilde{\sigma}_i(W_{t+h} - W_t) + \int_E \tilde{\beta}_i \overline{\mu}(de, (t, t+h]), \quad \forall 1 \leq i \leq \tilde{p},$$

avec $h \leq T$, il existe une constante $C_{K,\tilde{p},T}$ (dépendante de K, de $\tilde{p}$ et de T) telle que

$$\begin{aligned}
\left|\mathbb{E}\left[\prod_{1\leq i\leq \tilde{p}} P_i\right]\right| &\leq \begin{cases} C_{K,\tilde{p},T}h & \text{si } \prod_{1\leq i\leq \tilde{p}} \tilde{\epsilon}_i = 0 \\ 1 + C_{K,\tilde{p},T}h & \text{si } \prod_{1\leq i\leq \tilde{p}} \tilde{\epsilon}_i = 1 \end{cases} \\
\left\|\prod_{1\leq i\leq \tilde{p}} P_i\right\|_2 &\leq \begin{cases} C_{K,\tilde{p},T}\sqrt{h} & \text{si } \prod_{1\leq i\leq \tilde{p}} \tilde{\epsilon}_i = 0 \\ 1 + C_{K,\tilde{p},T}h & \text{si } \prod_{1\leq i\leq \tilde{p}} \tilde{\epsilon}_i = 1 \end{cases}
\end{aligned} \tag{3.9.1}$$

Or,

$$\begin{aligned}
\frac{\partial X_{k+1}^{k,x}}{\partial x} &= 1 + b'(x)(t_{k+1} - t_k) + \sigma'(x)(W_{t_{k+1}} - W_{t_k}) + \int_E \frac{\partial \beta(x,e)}{\partial x}\overline{\mu}(de,(t_k,t_{k+1}]), \quad \forall x \in \mathbb{R}, \\
\frac{\partial^j X_{k+1}^{k,x}}{\partial x^j} &= b^{(j)}(x)(t_{k+1} - t_k) + \sigma^{(j)}(x)(W_{t_{k+1}} - W_{t_k}) + \int_E \frac{\partial^j \beta(x,e)}{\partial x^j}\overline{\mu}(de,(t_k,t_{k+1}]), \quad \forall x \in \mathbb{R}, \forall 2 \leq j \leq m.
\end{aligned}$$

et comme $b^{(j)}$, $\sigma^{(j)}$ et $\frac{\partial^j \beta(x,e)}{\partial x^j}$ sont bornées, les inégalités (3.9.1) permettent d'obtenir les majorations suivantes : pour $1 \leq i < m$, $u \in \{2,..,m\}$ et pour $j_1,...,j_m \in \mathbb{N}^*$ tel que $j_u \geq 1$:

$$\begin{aligned}
\left|\mathbb{E}\left[\left(\frac{\partial X_{k+1}^{k,x}}{\partial x}\right)^{j_1} ... \left(\frac{\partial^m X_{k+1}^{k,x}}{\partial^m x}\right)^{j_m}\right]\right| &\leq C\alpha_n, \\
\left\|\left(\frac{\partial X_{k+1}^{k,x}}{\partial x}\right)^{j_1} ... \left(\frac{\partial^m X_{k+1}^{k,x}}{\partial^m x}\right)^{j_m}\right\|_2 &\leq C\sqrt{\alpha_n}, \\
\left|\mathbb{E}\left[\left(\frac{\partial X_{k+1}^{k,x}}{\partial x}\right)^{j_1} ... \left(\frac{\partial^m X_{k+1}^{k,x}}{\partial^m x}\right)^{j_m} r_{k+1}\right]\right| &\leq C, \\
\left\|\left(\frac{\partial X_{k+1}^{k,x}}{\partial x}\right)^{j_1} ... \left(\frac{\partial^m X_{k+1}^{k,x}}{\partial^m x}\right)^{j_m} r_{k+1}\right\|_2 &\leq \frac{C}{\sqrt{n}}
\end{aligned}$$

avec

$$\alpha_n = \begin{cases} \frac{1}{n} & \text{si } \exists u \in \{2,..,m\} \text{ tel que } j_u \geq 1 \\ 1 & \text{sinon} \end{cases}$$

Cela implique que pour $i < \alpha$

$$\begin{aligned}
\left|\mathbb{E}\left[B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1\le u\le \alpha-i+1}\right)\right]\right| &\le C\alpha_n,\\
\left\|B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1\le u\le \alpha-i+1}\right)\right\|_2 &\le C\sqrt{\alpha_n},\\
\left|\mathbb{E}\left[B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1\le u\le \alpha-i+1}\right) r_{k+1}\right]\right| &\le C,\\
\left\|B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1\le u\le \alpha-i+1}\right) r_{k+1}\right\|_2 &\le \frac{C}{\sqrt{n}}
\end{aligned}$$

avec

$$\alpha_n = \begin{cases} \frac{1}{n} & \text{si } i < \alpha \\ 1 & \text{si } i = \alpha. \end{cases}$$

En combinant les inégalités précédentes avec l'inégalité de Cauchy-Schwarz, on obtient les inégalités suivantes pour une fonction $f \in \mathcal{C}_b^1$ et pour $i \le \alpha$

$$\begin{aligned}
&\left|\mathbb{E}\left[f(X_{k+1}^{k,x})B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right) r_{k+1}\right]\right|\\
\le &\left|\mathbb{E}\left[\left(f(X_{k+1}^{k,x}) - f(x)\right) B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right) r_{k+1}\right]\right| + \left|\mathbb{E}\left[f(x)B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right) r_{k+1}\right]\right|\\
\le &\left|\mathbb{E}\left[[f]_{Lip}\left|X_{k+1}^{k,x} - x\right| \left|B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right) r_{k+1}\right|\right]\right| + |f|_\infty \left|\mathbb{E}\left[B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right) r_{k+1}\right]\right|\\
\le &[f]_{Lip}\left\|X_{k+1}^{k,x} - x\right\|_2 \left\|B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1\le u\le \alpha-i+1}\right) r_{k+1}\right\|_2 + |f|_\infty C\\
&\qquad (3.9.2)\\
\le &C\left(|f|_\infty + [f]_{Lip}\right).
\end{aligned}$$

et

$$
\begin{aligned}
&\left|\mathbb{E}\left[f(X_{k+1}^{k,x})B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right)\right]\right| \\
\leq &\left|\mathbb{E}\left[\left(f(X_{k+1}^{k,x})-f(x)\right)B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right)\right]\right| + \left|\mathbb{E}\left[f(x)B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right)\right]\right| \\
\leq &\left|\mathbb{E}\left[[f]_{Lip}\left|X_{k+1}^{k,x}-x\right|\left|B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right)\right|\right]\right| + |f|_\infty\left|\mathbb{E}\left[B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right)\right]\right| \\
\leq &[f]_{Lip}\left|\left|X_{k+1}^{k,x}-x\right|\right|_2\left|\left|B_{\alpha,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1\leq u\leq \alpha-i+1}\right)\right|\right|_2 + |f|_\infty\frac{C}{n} \\
&(3.9.3) \\
\leq &\frac{C}{n}\left(|f|_\infty + [f]_{Lip}\right)
\end{aligned}
$$

Pour $\alpha < m$, l'inégalité (3.9.2) implique la majoration des dérivées α-ième de u_k :

$$
\begin{aligned}
\left|\frac{\partial^\alpha}{\partial x^\alpha}u_k(x)\right| &= \left|\mathbb{E}\left[\sum_{i=1}^{\alpha}\frac{\partial^i y_{k+1}}{\partial x^i}(X_{k+1}^{k,x})B_{\alpha,i}\left(\frac{\partial X_{k+1}^{k,x}}{\partial x},\frac{\partial^2 X_{k+1}^{k,x}}{\partial x^2},..,\frac{\partial^{\alpha-i+1}X_{k+1}^{k,x}}{\partial x^{\alpha-i+1}}\right)\right]\right| \\
&\leq \sum_{i=1}^{\alpha}\left|\mathbb{E}\left[\frac{\partial^i y_{k+1}}{\partial x^i}(X_{k+1}^{k,x})B_{\alpha,i}\left(\frac{\partial X_{k+1}^{k,x}}{\partial x},\frac{\partial^2 X_{k+1}^{k,x}}{\partial x^2},..,\frac{\partial^{\alpha-i+1}X_{k+1}^{k,x}}{\partial x^{\alpha-i+1}}\right)\right]\right| \\
&\leq \sum_{i=1}^{\alpha}C\left(\left|\frac{\partial^i y_{k+1}}{\partial x^i}\right|_\infty + \left[\frac{\partial^i y_{k+1}}{\partial x^i}\right]_{Lip}\right) \\
&\leq CL_{k+1}^{\alpha+1}.
\end{aligned}
$$

En dérivant m fois la fonction $y_k(x) = u_{k,1}(x) + \frac{T}{n}f\left(x, u_k(x)\right)$, on obtient

$$
\begin{aligned}
\frac{\partial^m}{\partial x^m}y_k(x) = \frac{\partial^m}{\partial x^m}u_{k,1}(x) &+ \frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}\frac{\partial f\left(\theta(x)\right)}{\partial u_\epsilon}\frac{\partial^m}{\partial x^m}u_{k,\epsilon}(x) \\
&+ \frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}w_{k,\epsilon}(f,u,x)\frac{\partial^{m-1}}{\partial x^{m-1}}u_{k,\epsilon}(x) + \frac{T}{n}R_{k,m}(x)
\end{aligned}
$$

où $R_{k,m}(x)$ est une fonction des dérivées de f et des dérivées α-ième de u_k avec $0 \leq \alpha \leq m-2$ et est bornée avec un rapport dépendant de L_{k+1}^{m-1} (car $f \in \mathcal{C}_b^m$ et $u_k \in \mathcal{C}_b^m$) et où la fonction

$$
w_{k,\epsilon}(f,u,x) = \sum_{1\leq i\leq 4}\frac{\partial^2 f\left(x,u_k(x)\right)}{\partial u_\epsilon \partial\theta_i}\frac{\partial\theta_i}{\partial x}(x), \quad \forall\epsilon\in\{1;2;3\}
$$

est bornée (par une contante dépendante de L_{k+1}^{m-1}). D'où

$$
\left|\frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}w_{k,\epsilon}(f,u,x)\frac{\partial^{m-1}}{\partial x^{m-1}}u_{k,\epsilon}(x) + \frac{T}{n}R_{k,m}(x)\right| \leq C\frac{1+L_{k+1}^m}{n}
$$

Majorons maintenant le terme principal,

$$\begin{aligned}
&\frac{\partial^m}{\partial x^m}u_{k,1}(x)+\frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}\frac{\partial f\left(x,u_k(x)\right)}{\partial u_\epsilon}\frac{\partial^m}{\partial x^m}u_{k,\epsilon}(x)\\
=&\mathbb{E}\left[\frac{\partial^m}{\partial x^m}(y_{k+1}(X_{k+1}^{k,x}))\left\{1+\frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}\frac{\partial f\left(x,u_k(x)\right)}{\partial u_\epsilon}r_{k+1,\epsilon}\right\}\right]\\
=&\mathbb{E}\left[\frac{\partial^m y_{k+1}}{\partial x^m}\left(\frac{\partial X_{k+1}^{k,x}}{\partial x}\right)^m\left\{1+\frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}\frac{\partial f\left(x,u_k(x)\right)}{\partial u_\epsilon}r_{k+1,\epsilon}\right\}\right]\\
+&\sum_{i=1}^{m-1}\mathbb{E}\left[\frac{\partial^i y_{k+1}}{\partial x^i}(X_{k+1}^{k,x})B_{m,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_u\right)\left\{1+\frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}\frac{\partial f\left(x,u_k(x)\right)}{\partial u_\epsilon}r_{k+1,\epsilon}\right\}\right].
\end{aligned}$$

Nous avons, pour $i<m$,

$$\begin{aligned}
&\left|\mathbb{E}\left[\frac{\partial^i y_{k+1}}{\partial x^i}(X_{k+1}^{k,x})B_{m,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1\le u\le m-i+1}\right)\left\{1+\frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}\frac{\partial f\left(x,u_k(x)\right)}{\partial u_\epsilon}r_{k+1,\epsilon}\right\}\right]\right|\\
\le&\left|\mathbb{E}\left[\frac{\partial^i y_{k+1}}{\partial x^i}(X_{k+1}^{k,x})B_{m,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1\le u\le m-i+1}\right)\right]\right|\\
+&C\frac{T}{n}\left|\mathbb{E}\left[\frac{\partial^i y_{k+1}}{\partial x^i}(X_{k+1}^{k,x})B_{m,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1\le u\le m-i+1}\right)r_{k+1}\right]\right|\\
\le&\frac{C}{n}\left(\left|\frac{\partial^i y_{k+1}}{\partial x^i}\right|_\infty+\left[\frac{\partial^i y_{k+1}}{\partial x^i}\right]_{Lip}\right)\\
\le&\frac{C}{n}L_{k+1}^m
\end{aligned}$$

L'inégalité de Hölder permet de majorer le dernier terme

$$\begin{aligned}
&\left|\mathbb{E}\left[\frac{\partial^m y_{k+1}}{\partial x^m}\left(\frac{\partial X_{k+1}^{k,x}}{\partial x}\right)^m\left\{1+\frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}\frac{\partial f\left(x,u_k(x)\right)}{\partial u_\epsilon}r_{k+1,\epsilon}\right\}\right]\right|\\
\le&\left|\frac{\partial^m y_{k+1}}{\partial x^m}\right|_\infty\left\|\left(\frac{\partial X_{k+1}^{k,x}}{\partial x}\right)^m\right\|_2\left\|1+\frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}\frac{\partial f\left(x,u_k(x)\right)}{\partial u_\epsilon}r_{k+1,\epsilon}\right\|_2\\
\le&L_{k+1}^m\left(1+\frac{C}{n}\right)
\end{aligned}$$

En combinant les inégalités précédentes, on obtient :

$$\begin{aligned}\left|\frac{\partial^m}{\partial x^m}y_k(x)\right|_\infty &\leq \frac{T}{n}\left|R_{k,m}(x)\right| + \left|\mathbb{E}\left[\frac{\partial^m y_{k+1}}{\partial x^m}\left(\frac{\partial X_{k+1}^{k,x}}{\partial x}\right)^m\left\{1+\frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}\frac{\partial f\left(x,u_k(x)\right)}{\partial u_\epsilon}r_{k+1,\epsilon}\right\}\right]\right| \\ &+ \left|\mathbb{E}\left[\frac{\partial^i y_{k+1}}{\partial x^i}(X_{k+1}^{k,x})B_{m,i}\left(\left(\frac{\partial^u X_{k+1}^{k,x}}{\partial x^u}\right)_{1\leq u\leq m-i+1}\right)\left\{1+\frac{T}{n}\sum_{\epsilon\in\{1;2;3\}}\frac{\partial f\left(x,u_k(x)\right)}{\partial u_\epsilon}r_{k+1,\epsilon}\right\}\right]\right| \\ &\leq L_{k+1}^m\left(1+\frac{C}{n}\right)+\frac{C}{n},\quad \forall x\in\mathbb{R}.\end{aligned}$$

D'où

$$L_k^m \leq L_{k+1}^m\left(1+\frac{C}{n}\right)+\frac{C}{n}.$$

L'inégalité de Gronwall permet de conclure en donnant la majoration pour tout $k\in\{0,...,n\}$

$$\sup_{0\leq k\leq n} L_k^m \leq C$$

□

Le corollaire (3.3.6) donne la régularité de la solution de l'EDSR.

Démonstration. Suppose de plus que les fonctions $b,\sigma,\beta(.,e),f,h$ appartiennent à $\mathcal{C}_b^m$ pour $m\in\mathbb{N}^*$, alors le schéma discret avec n pas de temps $\pi=\{t_k\}=k\frac{T}{n}|0\leq k\leq n\}$ appartient à $\mathcal{C}_b^m$ avec une majoration des dérivées indépendante de n. La convergence du schéma discret quand $n\to+\infty$ (voir ([**BE08**])) permet de conclure. □

3.9.3. EDP parabolique totalement non linéaire. Le lemme (3.5.1) donne la stabilité $\mathbb{L}_\infty$ de l'algorithme pour les EDP paraboliques totalement non linéaires.

Démonstration. Soient $k\in\{0,...,n-1\}$ et $x\in\mathbb{R}^d$. Soit $\lambda=(\lambda_y,\lambda_z,\lambda_\gamma)\in\mathcal{S}^{\Delta T_k}$. En notant $B_{\frac{T}{n}}=W_{t_{k+1}}-W_{t_k}\sim\mathcal{N}(0;\frac{T}{n})$, on obtient

$$\lambda.\phi_k(x,B_{\frac{T}{n}})=\lambda_y+\lambda_z.\left(\sigma'\right)^{-1}B_{\frac{T}{n}}+\lambda_\gamma.\left(\sigma'\right)^{-1}\left(\frac{B_{\frac{T}{n}}^2}{\frac{T}{n}}-I_d\right)\sigma^{-1}.$$

Comme

$$T_k\left(x,u\right)=u_y+\frac{T}{n}f\left(t_k,x,u\right),\quad \forall x\in\mathbb{R}^d,\forall u=(u_y,u_z,u_\gamma)\in\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^{d\times d}$$

avec f lipschitzienne, on a

$$|\lambda_y|\leq 1+\frac{T}{n}\left[f\right]_{\text{Lip}}\quad\text{et}\quad\max\left\{\left|\lambda_z\right|_\infty,\left|\lambda_\gamma\right|_\infty\right\}\leq\frac{T}{n}\left[f\right]_{\text{Lip}}.$$

L'inégalité

$$\begin{aligned}\left|\left|\lambda.\phi_k(x,B_{\frac{T}{n}})\right|\right|_p &\leq \left|\left|\lambda_y\right|\right|_p+\left|\left|\lambda_z.\left(\sigma'\right)^{-1}B_{\frac{T}{n}}\right|\right|_p+\left|\left|\lambda_\gamma.\left(\sigma'\right)^{-1}\left(\frac{B_{\frac{T}{n}}^2}{\frac{T}{n}}-I_d\right)\sigma^{-1}\right|\right|_p \\ &\leq C,\quad\forall p\in[1,+\infty[,\end{aligned}$$

montre que l'hypothèse $(\mathcal{P}_p)$ est vérifiée pour tout $p\in[1,+\infty[$.

Pour l'hypothèse (H_1), on a :

$$\begin{aligned}
\mathbb{E}\left[\left|\lambda.\phi_k(x, B_{\frac{T}{n}})\right|\right] &= \mathbb{E}\left[\left|\lambda_x + \lambda_z.\left(\sigma'\right)^{-1} B_{\frac{T}{n}} + \lambda_\gamma.\left(\sigma'\right)^{-1}\left(\frac{B^2_{\frac{T}{n}}}{\frac{T}{n}} - I_d\right)\sigma^{-1}\right|\right] \\
&\leq \mathbb{E}\left[\left|1 - \lambda_\gamma.\left(\sigma'\right)^{-1}\sigma^{-1} + \lambda_z.\left(\sigma'\right)^{-1} B_{\frac{T}{n}}\right| + \lambda_\gamma.\left(\sigma'\right)^{-1}\frac{B^2_{\frac{T}{n}}}{\frac{T}{n}}\sigma^{-1}\right] + \frac{C}{n} \\
&\leq \mathbb{E}\left[\left|1 - P_k + \nabla_p f.\left(\sigma'\right)^{-1} B_{\frac{T}{n}}\right|\right] + P_k + \frac{T}{n}h,
\end{aligned}$$

avec $P_k = \nabla_\gamma f.\left(\sigma'\right)^{-1}\sigma^{-1} \geq 0$ et $1 - P_k \geq \epsilon$. Or, pour $a > 0$, l'inégalité de Cauchy-Schwarz implique

$$\begin{aligned}
\mathbb{E}\left[\left|a + bB_{\frac{T}{n}}\right|\right] &= a\mathbb{E}\left[\left|1 + \frac{b}{a}B_{\frac{T}{n}}\right|\right] \\
&\leq a\mathbb{E}\left[\left|1 + \frac{b}{a}B_{\frac{T}{n}}\right|^2\right]^{\frac{1}{2}} \\
&\leq a(1 + \frac{b^2}{a^2}\frac{T}{n}).
\end{aligned}$$

D'où,

$$\begin{aligned}
\mathbb{E}\left[\left|\lambda.\phi_k(x, B_{\frac{T}{n}})\right|\right] &\leq (1 - P_k)\,(1 + \frac{C}{n}) + P_k + \frac{C}{n} \\
&\leq 1 + \frac{C}{n}.
\end{aligned}$$

□

3.9.4. Maximisation d'utilité. Le lemme (3.6.1) donne la stabilité de l'algorithme pour le problème de maximisation d'utilité.

Démonstration. Soient $k \in \{0, ..., n-1\}$ et $\lambda \in \mathcal{S}^{\Delta T_k}$. Comme $T_{k,\pi} = I_d$, $\lambda \in \mathcal{S}^{\Delta T_k} = \{0; 1\}$. Soient $x = (s, m) \in \mathbb{E}^d$, $p \in]1, +\infty]$ et $q \in [1, +\infty[$ le conjugué de p, on a

$$\begin{aligned}
\mathbb{E}\left[|\lambda.\phi_{k,\pi_k}(x, \xi_{k+1})|^q\right] &\leq \sup_{\pi_k \in A_k} \mathbb{E}\left[e^{q\gamma\pi_k.(X^{k,x}_{k+1} - x)}\right] \leq \sup_{\pi_k \in A_k} \mathbb{E}\left[e^{q\gamma\pi_k.(S^{k,s}_{k+1} - s)}\right] \\
&\leq \sup_{\pi_k \in A_k} \mathbb{E}\left[e^{q\gamma\pi_k.(S^{k,s}_{k+1} - s)}\right],
\end{aligned}$$

car $A_k = \left(A_k^{(S)}, 0\right)$. Comme $A_k^{(S)} \subset \mathbb{R}^{d_S}$ est compact, on obtient

$$\begin{aligned}
\mathbb{E}\left[|\lambda.\phi_{k,\pi_k}(x, \xi_{k+1})|^q\right] &\leq \sup_{\pi_k \in A_k} \mathbb{E}\left[e^{q\gamma\pi_k.(X^{k,x}_{k+1} - x)}\right] \\
&\leq \sup_{\pi_k \in A_k} \mathbb{E}\left[e^{q\gamma\pi_k.(S^{k,s}_{k+1} - s)}\right] \\
&\leq 1 + \frac{C}{n}.
\end{aligned}$$

□

La proposition (3.6.2) donne la régularité lipschitzienne de y et u.

DÉMONSTRATION. Pour $k=n$, $x\in\mathbb{R}^d\to y_n(x)=\exp(-\gamma h(x))$ est bornée et lipschitziennne car la fonction h est bornée et lipchtizienne.

Soit $k<n$.

Comme $\sup_{\pi_k\in A_k}\mathbb{E}\left[e^{p\gamma\pi_k.(X_{k+1}^{k,x}-x)}\right]\leq 1+\frac{C_p}{n}$, les fonctions y_k et $u_{k,a}$ sont bornées : en effet, pour $x\in\mathbb{R}^d$, on a :

$$\begin{aligned}\sup\{|y_k(x)|,|u_{k,a}(x)|\}&\leq\sup_{\pi\in\mathcal{A}}\left|\mathbb{E}\left[e^{-\gamma\left(h(X_n^{k,x})-\sum_{i=k}^{n-1}\pi_i.(X_{i+1}^{k,x}-X_i^{k,x})\right)}\right]\right|\\&\leq\left|e^{-\gamma h}\right|_\infty\sup_{\pi\in A}\mathbb{E}\left[e^{\gamma\sum_{i=k}^{n-1}\pi_i.(X_{i+1}^{k,x}-X_i^{k,x})}\right]\\&\leq\left|e^{-\gamma h}\right|_\infty\left(1+\frac{C_p}{n}\right)^{n-k}\\&\leq e^{C_p}\left|e^{-\gamma h}\right|_\infty,\quad\forall 0\leq k\leq n,\forall x\in\mathbb{R}^d.\end{aligned}$$

Montrons la propriété lipschitzienne sur y_k de manière rétrograde. Soit $x=(s,p),\ \tilde{x}=(\tilde{s},\tilde{p})\in\mathbb{R}^{d_S\times d_P}$, on a, $\forall 0\leq k\leq n,\forall x\in\mathbb{R}^d$,

$$\begin{aligned}|y_k(x)-y_k(\tilde{x})|&\leq\left|\inf_{\pi_k\in A_k}\mathbb{E}\left[e^{\gamma\pi_k.(X_{k+1}^{k,x}-x)}y_{k+1}(X_{k+1}^{k,x})\right]-\inf_{\pi_k\in A_k}\mathbb{E}\left[e^{\gamma\pi_k.(X_{k+1}^{k,\tilde{x}}-\tilde{x})}y_{k+1}(X_{k+1}^{k,\tilde{x}})\right]\right|\\&\leq\sup_{\pi_k\in A_k}\left|\mathbb{E}\left[e^{\gamma\pi_k.(X_{k+1}^{k,x}-x)}y_{k+1}(X_{k+1}^{k,x})-e^{\gamma\pi_k.(X_{k+1}^{k,\tilde{x}}-\tilde{x})}y_{k+1}(X_{k+1}^{k,\tilde{x}})\right]\right|\\&\leq\sup_{\pi_k\in A_k^{(S)}}\mathbb{E}\left[e^{\gamma\pi_k.(S_{k+1}^{k,x}-s)}\left|y_{k+1}(X_{k+1}^{k,x})-y_{k+1}(X_{k+1}^{k,\tilde{x}})\right|\right]\\&\quad+\sup_{\pi_k\in A_k^{(S)}}\left|\mathbb{E}\left[\left(e^{\gamma\pi_k.(S_{k+1}^{k,x}-s)}-e^{\gamma\pi_k.(S_{k+1}^{k,\tilde{s}}-\tilde{s})}\right)y_{k+1}(X_{k+1}^{k,\tilde{x}})\right]\right|.\end{aligned}$$

L'inégalité de Jensen implique :

$$\begin{aligned}\mathbb{E}\left[e^{\gamma\pi_k.(S_{k+1}^{k,x}-s)}\left|y_{k+1}(X_{k+1}^{k,x})-y_{k+1}(X_{k+1}^{k,\tilde{x}})\right|\right]&\leq\left|\left|e^{\gamma\pi_k.(S_{k+1}^{k,x}-s)}\right|\right|_2\left|\left|y_{k+1}(X_{k+1}^{k,x})-y_{k+1}(X_{k+1}^{k,\tilde{x}})\right|\right|_2\\&\leq\left(1+\frac{C_2}{n}\right)\left|\left|y_{k+1}(X_{k+1}^{k,x})-y_{k+1}(X_{k+1}^{k,\tilde{x}})\right|\right|_2.\end{aligned}$$

En notant $\overline{X}_{k+1}^{k,x}=(s,\Phi_k(x,s,s))\in\mathcal{F}_k$ pour $x=(s,m)\in\mathbb{R}^d$ (et $X_{k+1}^{k,x}=\left(s,\Phi_k(x,s,S_{k+1}^{k,s})\right)$), on a l'inégalité suivante :

$$\left|X_{k+1}^{k,x}-\overline{X}_{k+1}^{k,x}\right|\leq C\left|S_{k+1}^{k,s}-s\right|.$$

Majorons maintenant le second terme

$$\begin{aligned}&\left|\mathbb{E}\left[\left(e^{\gamma\pi_k.(S_{k+1}^{k,s}-s)}-e^{\gamma\pi_k.(S_{k+1}^{k,\tilde{s}}-\tilde{s})}\right)y_{k+1}(X_{k+1}^{k,\tilde{x}})\right]\right|\\\leq&\left|\mathbb{E}\left[\left(e^{\gamma\pi_k.(S_{k+1}^{k,s}-s)}-e^{\gamma\pi_k.(S_{k+1}^{k,\tilde{s}}-\tilde{s})}\right)y_{k+1}(\overline{X}_{k+1}^{k,\tilde{x}})\right]\right|\\&+\left|\mathbb{E}\left[\left(e^{\gamma\pi_k.(S_{k+1}^{k,s}-s)}-e^{\gamma\pi_k.(S_{k+1}^{k,\tilde{s}}-\tilde{s})}\right)\left(y_{k+1}(X_{k+1}^{k,\tilde{x}})-y_{k+1}(\overline{X}_{k+1}^{k,\tilde{x}})\right)\right]\right|.\end{aligned}$$

$y_{k+1}(\overline{X}_{k+1}^{k,\tilde{x}})$ étant déterministe, on a

$$\begin{aligned}\left|\mathbb{E}\left[\left(e^{\gamma\pi_k.(S_{k+1}^{k,s}-s)}-e^{\gamma\pi_k.(S_{k+1}^{k,\tilde{s}}-\tilde{s})}\right)y_{k+1}(\overline{X}_{k+1}^{k,\tilde{x}})\right]\right|&\leq|y_{k+1}|_\infty\left|\mathbb{E}\left[e^{\gamma\pi_k.(S_{k+1}^{k,s}-s)}-e^{\gamma\pi_k.(S_{k+1}^{k,\tilde{s}}-\tilde{s})}\right]\right|\\&\leq|y_{k+1}|_\infty C\frac{|x-\tilde{x}|}{n}.\end{aligned}$$

L'inégalité de Hölder implique :

$$\begin{aligned}&\left|\mathbb{E}\left[\left(e^{\gamma\pi_k.(S_{k+1}^{k,s}-s)}-e^{\gamma\pi_k.(S_{k+1}^{k,\tilde{s}}-\tilde{s})}\right)\left(y_{k+1}(X_{k+1}^{k,\tilde{x}})-y_{k+1}(\overline{X}_{k+1}^{k,\tilde{x}})\right)\right]\right|\\&\leq\left|\left|e^{\gamma\pi_k.(S_{k+1}^{k,s}-s)}-e^{\gamma\pi_k.(S_{k+1}^{k,\tilde{s}}-\tilde{s})}\right|\right|_2\left|\left|y_{k+1}(X_{k+1}^{k,\tilde{x}})-y_{k+1}(\overline{X}_{k+1}^{k,\tilde{x}})\right|\right|_2.\end{aligned}$$

Or,

$$\begin{aligned}\left|\left|e^{\gamma\pi_k.(S_{k+1}^{k,s}-s)}-e^{\gamma\pi_k.(S_{k+1}^{k,\tilde{s}}-\tilde{s})}\right|\right|_2&\leq C\frac{|s-\tilde{s}|}{\sqrt{n}}\\&\leq C\frac{|x-\tilde{x}|}{\sqrt{n}},\\\left|\left|y_{k+1}(X_{k+1}^{k,\tilde{x}})-y_{k+1}(\overline{X}_{k+1}^{k,\tilde{x}})\right|\right|_2&\leq[y_{k+1}]_{\mathrm{Lip}}\left|\left|X_{k+1}^{k,\tilde{x}}-\overline{X}_{k+1}^{k,\tilde{x}}\right|\right|_2\\&\leq C\,[y_{k+1}]_{\mathrm{Lip}}\left|\left|S_{k+1}^{k,\tilde{s}}-\tilde{s}\right|\right|_2\\&\leq C\frac{[y_{k+1}]_{\mathrm{Lip}}}{\sqrt{n}}.\end{aligned}$$

En regroupant les termes, on obtient :

$$\begin{aligned}|y_k(x)-y_k(\tilde{x})|\leq&\left(1+\frac{C_2}{n}\right)\left|\left|y_{k+1}(X_{k+1}^{k,x})-y_{k+1}(X_{k+1}^{k,\tilde{x}})\right|\right|_2\\&+C\left\{|y_{k+1}|_\infty+[y_{k+1}]_{\mathrm{Lip}}\right\}\frac{|x-\tilde{x}|}{n}.\end{aligned}$$

puis de même pour $k'\geq k$:

$$\begin{aligned}\left|\left|y_{k'}(X_{k'}^{k,x})-y_{k'}(X_{k'}^{k,\tilde{x}})\right|\right|_2\leq&\left(1+\frac{C_2}{n}\right)\left|\left|y_{k'+1}(X_{k'+1}^{k,x})-y_{k'+1}(X_{k'+1}^{k,\tilde{x}})\right|\right|_2\\&+C\left\{|y_{k'+1}|_\infty+[y_{k'+1}]_{\mathrm{Lip}}\right\}\frac{\left|\left|X_{k'+1}^{k,x}-X_{k'+1}^{k,\tilde{x}}\right|\right|_2}{n}.\end{aligned}$$

En sommant, on obtient

$$\begin{aligned}|y_k(x)-y_k(\tilde{x})|\leq&\sum_{k'=k}^{n-1}C\left(1+\frac{C_2}{n}\right)^{k'-k}\left\{|y_{k'+1}|_\infty+[y_{k'+1}]_{\mathrm{Lip}}\right\}\frac{\left|\left|X_{k'+1}^{k,x}-X_{k'+1}^{k,\tilde{x}}\right|\right|_2}{n}\\&+\left(1+\frac{C_2}{n}\right)^{n-k}\left|\left|y_n(X_n^{k,x})-y_n(X_n^{k,\tilde{x}})\right|\right|_2.\end{aligned}$$

Comme $\left(1+\frac{C_2}{n}\right)^k\leq e^{C_2}$, $\forall 0\leq k\leq n$, la fonction y_n est lipschitzienne et comme X a un flot $\mathbb{L}_2$-lipschitzien, on a, $\forall 0\leq k\leq n$,

$$|y_k(x)-y_k(\tilde{x})|\leq|x-x'|\,C\left\{\sum_{k'=k}^{n-1}\frac{\left\{|y_{k'+1}|_\infty+[y_{k'+1}]_{\mathrm{Lip}}\right\}}{n}+[y_n]_{\mathrm{Lip}}\right\},\quad\forall x,\tilde{x}\in\mathbb{R}^d,$$

$$[y_k]_{\mathrm{Lip}}\leq C\left\{\sum_{k'=k}^{n-1}\frac{\left\{|y_{k'+1}|_\infty+[y_{k'+1}]_{\mathrm{Lip}}\right\}}{n}+[y_n]_{\mathrm{Lip}}\right\}.$$

En posant

$$a_i = C\left\{\sum_{k=i}^{n-1}\frac{\left\{|y_{k+1}|_\infty + [y_{k+1}]_{\text{Lip}}\right\}}{n} + [y_n]_{\text{Lip}}\right\},$$

on a

$$\begin{aligned} a_{i-1} &= a_i + C\frac{[y_i]_{\text{Lip}} + |y_i|_\infty}{n} \\ &\leq a_i\left(1+\frac{C^2}{n}\right) + \epsilon_i \end{aligned}$$

avec $\epsilon_i = C\frac{\sup_i |y_i|_\infty}{n}$. Le lemme de Gronwall (discret) permet finalement d'obtenir le rapport de Lipschitz pour y :

$$\begin{aligned} \sup_{0\leq i\leq n}[y_i]_{\text{Lip}} &\leq \sup_{1\leq i\leq n} a_i \\ &\leq e^{C^2}\left\{a_n + \sum_{i=1}^n \epsilon_i\right\} \\ &\leq C'\left\{[y_n]_{\text{Lip}} + \sup_i |y_i|_\infty\right\} \\ &\leq C''. \end{aligned}$$

De même, on obtient le rapport de Lipschitz de u :

$$\begin{aligned} |u_{k,a}(x) - u_{k,a}(\tilde{x})| &= \left|\mathbb{E}\left[e^{\gamma a.(S^{k,s}_{k+1}-s)}y_{k+1}(X^{k,x}_{k+1}) - e^{\gamma a.(X^{k,\tilde{s}}_{k+1}-\tilde{s})}y_{k+1}(X^{k,\tilde{x}}_{k+1})\right]\right| \\ &\leq \left|\mathbb{E}\left[e^{\gamma a.(S^{k,s}_{k+1}-s)}\left\{y_{k+1}(X^{k,x}_{k+1}) - y_{k+1}(X^{k,\tilde{x}}_{k+1})\right\}\right]\right| \\ &\quad + \left|\mathbb{E}\left[e^{\gamma a.(S^{k,s}_{k+1}-s)} - e^{\gamma a.(X^{k,\tilde{s}}_{k+1}-\tilde{s})}y_{k+1}(X^{k,\tilde{x}}_{k+1})\right]\right| \\ &\leq C\left|x-\tilde{x}\right|, \quad \forall 0\leq k\leq n-1, \forall x\in\mathbb{R}^d. \end{aligned}$$

□

3.9.5. Optimisation d'un processus de Markov contrôlé. Le lemme (3.7.1) décrit le changement de probabilité $\phi_{k,a}$ entre $\mathbb{P}^*$ et $\mathbb{P}^a$.

Démonstration. Décomposons X_{k+1} avec sa partie continue et sa partie à saut. On a

$$\mathbb{E}^a\left[G(b^{(e)}_{k+1,a})\right] = \mathbb{E}^a\left[G(b^{(e)}_{k+1,a})\phi^{(e)}_a(b^{(e)}_{k+1})\right]$$

où, pour $b=\{0;1\}$,

$$\begin{aligned} \phi^{(e)}_a(b) &= b\frac{\mathbb{P}^a(b^{(e)}_{k+1,a}=1)}{\mathbb{P}^*(b^{(e)}_{k+1}=1)} + (1-b)\frac{\mathbb{P}^a(b^{(e)}_{k+1,a}=0)}{\mathbb{P}^*(b^{(e)}_{k+1}=0)} \\ &= b\frac{1-\nu_a(e)\frac{T}{n}}{1-\nu_*(e)\frac{T}{n}} + (1-b)\frac{\nu_a(e)}{\nu_*(e)}. \end{aligned}$$

Pour le terme en ΔW^a_{k+1}, on a

$$\mathbb{E}^a\left[G\left(\mu_a(x)\frac{T}{n} + \sigma_a(x)\Delta W^a_{k+1}\right)\right] = \int_{\mathbb{R}^d} G\left(\mu_a(x)\frac{T}{n} + \sqrt{\frac{T}{n}}\sigma_a(x)u\right)e^{-\frac{|u|^2}{2}}\frac{du}{(2\pi)^{\frac{d}{2}}}.$$

Le changement de variable $v = \sigma_*^{-1}(x)\left(\sigma_a(x)u + (\mu_a(x) - \mu_*(x))\sqrt{\frac{T}{n}}\right)$ donne

$$\begin{aligned}\mathbb{E}^a\left[G\left(\mu_a(x)\frac{T}{n} + \sigma_a(x)\Delta W_{k+1}^a\right)\right] &= \int_{\mathbb{R}^d} G\left(\mu_*(x)\frac{T}{n} + \sqrt{\frac{T}{n}}\sigma_*(x)v\right) Q_{k,a}(x, \sqrt{\frac{T}{n}}v)e^{-\frac{|v|^2}{2}}\frac{dv}{(2\pi)^{\frac{d}{2}}} \\ &= \mathbb{E}^*\left[G\left(\mu_*(X_k)\frac{T}{n} + \sigma_*(X_k)\Delta W_{k+1}\right) Q_{k,a}(x, \Delta W_{k+1})\right],\end{aligned}$$

où

$$Q_{k,a}(x,w) = \frac{\left|\det(\sigma_*\sigma_*^T(x))\right|^{\frac{1}{2}}}{|\det(\sigma_a\sigma_a^T(x))|^{\frac{1}{2}}}\exp\frac{|w|^2 - \left|\sigma_a^{-1}(x)\sigma_*(x)w + \frac{T}{n}(\mu_a - \mu_*)(x)\right|^2}{2\frac{T}{n}}.$$

Par indépendance de $\left(\Delta W_{k+1}^a, \left(b_{k+1,a}^{(e)}\right)_{e\in E}\right)$, on a

$$\mathbb{E}^a\left[G(X_{k+1})|X_k\right] = \mathbb{E}^*\left[G(X_{k+1})\phi_{k,a}(X_k, \xi_{k+1})|X_k\right]$$

où $\xi_{k+1} = \left(W_{\frac{k+1}{n}T} - W_{\frac{k}{n}T}, \left(b_{k+1}^{(e)}\right)_{e\in E}\right)$ et $\phi_k \geq 0$ vaut

$$\phi_{k,a}(X_k, \xi_{k+1}) = Q_{k,a}(X_k, W_{\frac{k+1}{n}T} - W_{\frac{k}{n}T})\prod_{e\in E}\phi_a^{(e)}(b_{k+1}^{(e)}).$$

□

Le lemme (3.7.2) montre que les hypothèses $(\mathcal{P}_p)$ et (H_p) sont vérifiées pour les processus de Markov contrôlé.

Démonstration. Comme $T_{k,a}(x,u) = \frac{f_k(x,a)}{n} + u$, on a $\mathcal{S}^{\Delta T_k} = \{0;1\}$. Il suffit de considérer $\lambda = 1$.

Soit $p \in]1, +\infty[$ et $q \in]1, +\infty[$ son conjugué (i.e. $\frac{1}{p} + \frac{1}{q} = 1$). Soit $x \in \mathbb{R}^d$, il faut majorer à majorer le terme

$$\begin{aligned}\mathbb{E}^*\left[\phi_{k,a}(x, \xi_{k+1})^q\right] &= \mathbb{E}^*\left[Q_{k,a}(x, W_{\frac{k+1}{n}T} - W_{\frac{k}{n}T})^q\left|\prod_{e\in E}\phi_a^{(e)}(b_{k+1}^{(e)})\right|^q\right] \\ &= \mathbb{E}^*\left[Q_{k,a}(x, W_{\frac{k+1}{n}T} - W_{\frac{k}{n}T})^q\right]\prod_{e\in E}\mathbb{E}^*\left[\left(\phi_a^{(e)}(b_{k+1}^{(e)})\right)^q\right].\end{aligned}$$

Or,

$$\begin{aligned}\mathbb{E}^*\left[\left(\phi_a^{(e)}(b_{k+1}^{(e)})\right)^q\right] &= \phi_a^{(e)}(0)^q\mathbb{P}^*\left(b_{k+1}^{(e)} = 0\right) + \phi_a^{(e)}(1)^q\mathbb{P}^*\left(b_{k+1}^{(e)} = 1\right) \\ &= \left(1 - \nu_*(e)\frac{T}{n}\right)\left(\frac{\nu_a(e)}{\nu_*(e)}\right)^q + \nu_*(e)\frac{T}{n}\left(\frac{1 - \nu_a(e)\frac{T}{n}}{1 - \nu_*(e)\frac{T}{n}}\right)^q \\ &\leq 1 + C_q\frac{T}{n},\end{aligned}$$

et quelques calculs permettent d'obtenir

$$\begin{aligned}
&\mathbb{E}^*\left[Q_{k,a}(x, W_{\frac{k+1}{n}T} - W_{\frac{k}{n}T})^q\right] \\
&= \frac{\left|\det(\sigma_*\sigma_*^T(x))\right|^{\frac{q}{2}}}{\left|\det(\sigma_a\sigma_a^T(x))\right|^{\frac{q}{2}}} \int_{\mathbb{R}^d} e^{\frac{(q-1)|v|^2}{2} - \frac{q}{2}\left|\sigma_a^{-1}(x)\sigma_*(x)v + \sqrt{\frac{T}{n}}(\mu_a - \mu_*)(x)\right|^2} \frac{dv}{(2\pi)^{\frac{d}{2}}} \\
&= \frac{\left|\det(\Gamma_a(x))\right|^{\frac{q}{2}}}{\left|\det(q\Gamma_a(x) + (1-q)I_d)\right|^{\frac{1}{2}}} e^{\frac{T}{n}\left|R_a^{(q)}(\mu_a - \mu_*)\right|^2},
\end{aligned}$$

avec $\Gamma_a = \sigma_*^T(\sigma_a^T)^{-1}\sigma_a^{-1}\sigma_*$ et $(R_a^{(q)})^T R_a^{(q)} = \frac{1}{2}\sigma_*^T(\sigma_a^T)^{-1}\left(q\Gamma_a + (1-q)I_d\right)\sigma_a^{-1}\sigma_* - \frac{q}{2}I_d$.

Comme $\sigma_*\sigma_*^T \geq \sigma_a\sigma_a^T$, on a

$$\begin{aligned}
\Gamma_a &\geq I_d, \\
q\Gamma_a(X_k) + (1-q)I_d &\geq I_d.
\end{aligned}$$

D'où

$$\frac{\left|\det(\Gamma_a(x))\right|^{\frac{q}{2}}}{\left|\det(q\Gamma_a(x) + (1-q)I_d)\right|^{\frac{1}{2}}} < +\infty.$$

Comme les fonctions $\mu_a - \mu_*$ et $\sigma_a^{-1}\sigma_*$ sont bornées, on obtient $\mathbb{E}^*\left[Q_{k,a}(x, W_{\frac{k+1}{n}T} - W_{\frac{k}{n}T})^q\right]$ bornée et donc la condition $(\mathcal{P}_p)$ est vérifiée.

Si $\sigma_* = \sigma_a$, alors $\Gamma_a = I_d$, d'où

$$\begin{aligned}
\mathbb{E}^*\left[\phi_{k,a}(x, a, \xi_{k+1})^q\right] &\leq \mathbb{E}^*\left[Q_{k,a}(x, W_{\frac{k+1}{n}T} - W_{\frac{k}{n}T})^q\right] \prod_{e\in E} \mathbb{E}^*\left[\left(\phi_a^{(e)}(b_{k+1}^{(e)})\right)^q\right] \\
&\leq e^{\frac{T}{n}\left|R_a^{(q)}(\mu_a - \mu_*)\right|^2} \prod_{e\in E}\left(1 + C_q\frac{T}{n}\right) \\
&\leq 1 + \frac{C}{n},
\end{aligned}$$

par conséquent la condition (H_p) est vérifiée. □

La proposition (3.7.3) donne les propriétés de Lipschitz pour la solution y_k du problème d'optimisation d'un processus de Markov contrôlé.

DÉMONSTRATION. Montrons la propriété lipchitzienne pour y_k (ce qui impliquera que $u_{k,a}$ est lipschitzienne). Pour $k = n$, on a $y_n = h$ lipschitzienne.

Pour $k < n$, on définit la fonction v_k par

$$v_k(x) = \sup_{a\in\mathcal{A}} \mathbb{E}^a\left[\frac{f_k(x,a)}{n} + y_{k+1}(X_{k+1}^{k,x})\right], \quad \forall x \in \mathbb{R}^d.$$

On a alors $y_k = v_k \vee g$ et

$$\begin{aligned}
|v_k(x) - v_k(x')| &= \left| \sup_{a\in\mathcal{A}} \mathbb{E}^a \left[\frac{f_k(x,a)}{n} + y_{k+1}(X_{k+1}^{k,x}) \right] - \sup_{a\in\mathcal{A}} \mathbb{E}^a \left[\frac{f_k(x',a)}{n} + y_{k+1}(X_{k+1}^{k,x'}) \right] \right| \\
&\leq \sup_{a\in\mathcal{A}} \left| \mathbb{E}^a \left[\frac{f_k(x,a) - f_k(x',a)}{n} + y_{k+1}(X_{k+1}^{k,x}) - y_{k+1}(X_{k+1}^{k,x'}) \right] \right| \\
&\leq \frac{[f_k]_{\mathrm{Lip}}}{n} |x - x'| + [y_{k+1}]_{\mathrm{Lip}} \sup_{a\in\mathcal{A}} \mathbb{E}^a \left[\left| X_{k+1}^{k,x} - X_{k+1}^{k,x'} \right| \right] \\
&\leq |x - x'| \left\{ \frac{[f_k]_{\mathrm{Lip}}}{n} + [y_{k+1}]_{\mathrm{Lip}} \left(1 + \frac{C}{n} \right) \right\}, \quad \forall x, x' \in \mathbb{R}^d.
\end{aligned}$$

Avec l'inégalité

$$\begin{aligned}
|y_k(x) - y_k(x')| &= |v_k(x) \vee g(x) - v_k(x') \vee g(x')| \\
&\leq \sup \left(|v_k(x) - v_k(x')|, |g(x) - g(x')| \right), \quad \forall x, x' \in \mathbb{R}^d,
\end{aligned}$$

on obtient

$$\begin{aligned}
[y_k]_{\mathrm{Lip}} &\leq [v_k]_{\mathrm{Lip}} \vee [g]_{\mathrm{Lip}} \\
&\leq \left(\frac{[f_k]_{\mathrm{Lip}}}{n} + [y_{k+1}]_{\mathrm{Lip}} \left(1 + \frac{C}{n} \right) \right) \vee [g]_{\mathrm{Lip}}.
\end{aligned}$$

Le lemme de Gronwall (discret) appliqué à $\alpha_k = [y_k]_{\mathrm{Lip}} \vee [g]_{\mathrm{Lip}}$ permet d'obtenir la majoration suivante :

$$\sup_{0\leq k\leq n} [y_k]_{\mathrm{Lip}} \leq e^C \left\{ [h]_{\mathrm{Lip}} + \sup_{0\leq k\leq n} [f_k]_{\mathrm{Lip}} \right\} \vee [g]_{\mathrm{Lip}}.$$

□

Trivialement, on obtient alors que $u_{k,a}$ est aussi uniformément lipschitzienne avec

$$u_{k,a}(x) = \mathbb{E}^* \left[y_{k+1}(X_{k+1}^{k,x}) \phi_{k,a}(x, \xi_{k+1}) \right] = \mathbb{E}^a \left[y_{k+1}(X_{k+1}^{k,x}) \right].$$

3.9.6. Option américaine en information partielle. Le lemme (3.8.1) montre que les fonctions $\hat{f}$ et $\hat{h}$ sont lipschitziennes ce qui permet d'avoir la régularité nécéssaire pour T_k et R_k dans la formulation du schéma par programmation dynamique générique (2.2.1).

Démonstration. Il suffit de constater que

$$\begin{aligned}
\left| \hat{f}_k \right|_\infty &= \sup_{(\pi,z)\in K_m \times \mathbb{R}^d} \left| \hat{f}_k(\pi, z) \right| \\
&\leq \sup_{\pi\in K_m} \sum_{i=1}^m |f_k|_\infty \pi^i \\
&\leq K, \\
\left| \hat{h}_k \right|_\infty &\leq K,
\end{aligned}$$

et que pour $(\pi,z),(\overline{\pi},\overline{z})\in K_m\times\mathbb{R}^d$, on a :

$$\begin{aligned}\left|\hat{f}_k(\pi,z)-\hat{f}_k(\overline{\pi},\overline{z})\right| &\leq \left|\hat{f}_k(\pi,z)-\hat{f}_k(\pi,\overline{z})\right|+\left|\hat{f}_k(\overline{\pi},\overline{z})-\hat{f}_k(\overline{\pi},\overline{z})\right| \\ &\leq \left|\sum_{i=1}^m \left(f_k(v_i,z)-f_k(v_i,\overline{z})\right)\pi^i\right|+\left|\sum_{i=1}^m f_k(v_i,\overline{z})\left(\overline{\pi}^i-\pi^i\right)\right| \\ &\leq [f_k]_{\mathrm{Lip}}\,|z-\overline{z}|+|f_k|_\infty\,|\overline{\pi}-\pi|_1\,.\end{aligned}$$

□

Le lemme (3.8.3) donne la régularité de processus de Markov Z en information partielle.

Démonstration. Comme la fonction Φ est lipschitzienne et que S a un flot lipchitzien, alors Z a un flot lipschitzien. En effet pour $k\in\{0,..,n\}$, $v\in E$ et pour $z=(s,p),\tilde{z}=(\tilde{s},\tilde{p})\in\mathbb{R}^d$, on a

$$\begin{aligned}&\mathbb{E}\left[\sup_{k\leq k'\leq n}\left|Z_{k'}^{k,(v,z)}-Z_{k'}^{k,(v,\tilde{z})}\right|\right] \\ &\leq \mathbb{E}\left[\sup_{k\leq k'\leq n}\left|\left(s,\Phi_{k,k'}\left(p,(S_i^{k,(v,s)})_{k\leq i\leq k'}\right)\right)-\left(\tilde{s},\Phi_{k,k'}\left(p,(S_i^{k,(v,\tilde{s})})_{k\leq i\leq k'}\right)\right)\right|\right] \\ &\leq C\,|p-\tilde{p}|+C\mathbb{E}\left[\sup_{k\leq i\leq n}\left|S_i^{k,(v,s)}-S_i^{k,(v,\tilde{s})}\right|\right] \\ &\leq C\,|z-\tilde{z}|\,.\end{aligned}$$

□

La proposition (3.8.5) indique que la solution $\{y_k\}_{0\leq k\leq n}$ de l'option bermudéenne en information partielle est uniformément lipschitzienne.

Démonstration. Pour $k=n$, la fonction $y_n=\hat{h}_n$ est bornée et Lipschitzienne.

Soit $k<n$, l'inégalité

$$\begin{aligned}|y_k(x)| &= \left|\sup_{\tau\in\mathcal{T}_n^X}\mathbb{E}\left[\sum_{k\leq i<\tau}\frac{\hat{f}_i(X_i^{k,x})}{n}+\hat{h}_\tau(X_\tau^{k,x})\right]\right| \\ &\leq \sup_{0\leq i\leq n}\left|\hat{f}_i\right|_\infty+\sup_{0\leq i\leq n}\left|\hat{h}_i\right|_\infty\end{aligned}$$

montre que y_k est bornée (avec une majoration indépendante de k et n).

Soit $(\pi,z),(\overline{\pi},\overline{z})\in K_m\times\mathbb{R}^d$. Pour $k\leq k'<n$, en notant $X_{k'}^{k,(\pi,z)}=\left(\Pi_{k'}^{k,\pi},Z_{k'}^{k,(\pi,z)}\right)$ la variable $X_{k'}$ conditionnellement à $X_k=(\pi,z)$, on a :

$$\begin{aligned}|y_k((\pi,z))-y_k((\overline{\pi},\overline{z}))| &= |y_k((\pi,z))-y_k(\overline{\pi},\overline{z})| \\ &\leq |y_k((\pi,z))-y_k(\pi,\overline{z})|+|y_k((\pi,\overline{z}))-y_k(\overline{\pi},\overline{z})|\,.\end{aligned}$$

La propriété lipschitzienne de $\hat{f}_i$, $\hat{h}_i$ (d'après le lemme (3.8.1)) et du flot de Z (d'après le lemme (3.8.3)) permet de contrôler le premier terme :

$$
\begin{aligned}
&|y_k((\pi,z)) - y_k(\pi,\overline{z})| \\
=&\left| \sup_{\tau\in\mathcal{T}_n^X} \mathbb{E}\left[\sum_{k\le i<\tau} \frac{\hat{f}_i(X_i^{k,(\pi,z)})}{n} + \hat{h}_\tau(X_\tau^{k,(\pi,z)})\right] - \sup_{\tau\in\mathcal{T}_n^X} \mathbb{E}\left[\sum_{k\le i<\tau} \frac{\hat{f}_i(X_i^{k,(\pi,\overline{z})})}{n} + \hat{h}_\tau(X_\tau^{k,(\pi,\overline{z})})\right]\right| \\
\le& \sup_{\tau\in\mathcal{T}_n^X} \left|\mathbb{E}\left[\sum_{k\le i<\tau} \frac{\hat{f}_i(\Pi_i^{k,\pi}, Z_i^{k,(\pi,z)}) - \hat{f}_i(\Pi_i^{k,\pi}, Z_i^{k,(\pi,\overline{z})})}{n} + \left(\hat{h}_\tau(\Pi_\tau^{k,\pi}, Z_\tau^{k,(\pi,z)}) - \hat{h}_\tau(\Pi_i^{k,\pi}, Z_\tau^{k,(\pi,\overline{z})})\right)\right]\right| \\
\le& \sup_{\tau\in\mathcal{T}_n^X} \mathbb{E}\left[\sum_{k\le i<\tau} \frac{\left[\hat{f}_i\right]_{\mathrm{Lip}} \left|Z_i^{k,(\pi,z)} - Z_i^{k,(\pi,\overline{z})}\right|}{n} + \left[\hat{h}_\tau\right]_{\mathrm{Lip}} \left|Z_\tau^{k,(\pi,z)} - Z_\tau^{k,(\pi,\overline{z})}\right|\right] \\
\le& 2K'\mathbb{E}\left[\sup_{k\le i\le n} \left|Z_i^{k,(\pi,z)} - Z_i^{k,(\pi,\overline{z})}\right|\right] \\
\le& 2K'C\left|z-\overline{z}\right|.
\end{aligned}
$$

Pour le second terme, on va utiliser le fait que

$$
\begin{aligned}
&|y_k((\pi,\overline{z})) - y_k(\overline{\pi},\overline{z})| \\
\le&|\sup_{\tau\in\mathcal{T}_n^X} \mathbb{E}[\sum_{k\le i<\tau} \frac{1}{n}\hat{f}_i(X_i^{k,(\pi,\overline{z})}) + \hat{h}_\tau(X_\tau^{k,(\pi,\overline{z})})] - \sup_{\tau\in\mathcal{T}_n^X} \mathbb{E}[\sum_{k\le i<\tau} \frac{1}{n}\hat{f}_i(X_i^{k,(\overline{\pi},\overline{z})}) + \hat{h}_\tau(X_\tau^{k,(\overline{\pi},\overline{z})})]| \\
\le& \sup_{\tau\in\mathcal{T}_n^X} |\mathbb{E}[\sum_{k\le i<\tau} \frac{1}{n}\hat{f}_i(X_i^{k,(\pi,\overline{z})}) + \hat{h}_\tau(X_\tau^{k,(\pi,\overline{z})})] - \mathbb{E}[\sum_{k\le i<\tau} \frac{1}{n}\hat{f}_i(X_i^{k,(\overline{\pi},\overline{z})}) + \hat{h}_\tau(X_\tau^{k,(\overline{\pi},\overline{z})})]| \\
\le& \sup_{\tau\in\mathcal{T}_n^X} |\mathbb{E}[\sum_{k\le i<\tau} \frac{1}{n}\hat{f}_i(V_i,Z_i) + \hat{h}_\tau(V_\tau,Z_\tau)|(\Pi_k,Z_k) = (\pi,\overline{z})] \\
&- \mathbb{E}[\sum_{k\le i<\tau} \frac{1}{n}\hat{f}_i(V_i,Z_i) + \hat{h}_\tau(V_\tau,Z_\tau)|(\Pi_k,Z_k) = (\overline{\pi},\overline{z})]|.
\end{aligned}
$$

Comme $\mathcal{T}_n^X \subset \mathcal{T}_n$ où $\mathcal{T}_n$ est l'ensemble des temps d'arrêt $(\mathcal{F}_i)_{0\le i\le n}$ mesurable et comme (V_i,Z_i) est une $(\mathcal{F}_i)_{0\le i\le n}$ chaîne de Markov, on a

$$
\begin{aligned}
|y_k((\pi,\overline{z})) - y_k(\overline{\pi},\overline{z})| &\le \sup_{\tau\in\mathcal{T}_n} |\mathbb{E}[\sum_{k\le i<\tau} \frac{1}{n}\hat{f}_i(V_i,Z_i) + \hat{h}_\tau(V_\tau,Z_\tau)|(V_k,Z_k) = (v_j,\overline{z})](\pi^j - \overline{\pi}^j)| \\
&\le (\sup_{0\le i\le n} |f_i|_\infty + \sup_{0\le i\le n} |h_i|_\infty) \sum_{j=1}^m |\pi_j - \overline{\pi}_j|.
\end{aligned}
$$

En regroupant les deux inégalités, on obtient que y_k est lipschitzienne :

$$
|y_k((\pi,z)) - y_k(\overline{\pi},\overline{z})| \le 2K'C\left|z-z'\right| + 2K'\left|\pi-\overline{\pi}\right|_1.
$$

□

CHAPITRE 4

Applications numériques

4.1. Options américaines

Soit $\{X_{t_k}\}_{0\leq k\leq n}$ une chaîne de Markov. On cherche à calculer

$$\text{Esssup}_{\tau\in\mathcal{T}_{t_k,t_n}}\mathbb{E}_{t_k}\left[h(t_\tau,X_\tau)\right]$$

où $\mathcal{T}_{t_k,t_n}$ est l'ensemble des temps d'arrêts à valeurs dans $\{t_i\geq t_k\}$. Nous allons considérer deux types d'options exotiques courantes en finance :

$\diamond$ le cas des options lookback où la chaîne de Markov est définie par un actif risqué S_{t_k} et son maximum $\sup_{0\leq u\leq t_k}S_u$;

$\diamond$ le cas des options asiatiques où la chaîne de Markov est définie par un actif risqué S_{t_k} et son moyenne $\int_0^{t_k}S_u du$.

La convergence des schémas par quantification est prouvée dans la partie précédente et l'erreur de quantification est donnée par

$$\varepsilon(\text{quantif})\leq C\sum_{k=0}^{n}\left|\left|X_{t_k}-\hat{X}_{t_k}\right|\right|_p$$

avec $\hat{X}_{t_k}$ une quantification de la variable aléatoire X_{t_k}.

4.1.1. Option Lookback. On suppose que S suit le modèle de Black-Scholes, c'est-à-dire

$$\frac{dS_t}{S_t}=(r-\mu)dt+\sigma dW_t$$

où W est un mouvement Brownien standart.

On souhaite valoriser l'option américaine suivante :

$$\begin{aligned}P_0&=\text{Esssup}_{\tau\in\mathcal{T}_T}\mathbb{E}^{\mathbb{P}}\left[h\left(\tau,S_\tau,\sup_{0\leq t\leq\tau}S_t\right)\right]\\&=\text{Esssup}_{\tau\in\mathcal{T}_T}\mathbb{E}^{\mathbb{P}}\left[h\left(\tau,S_0e^{(r-\frac{\sigma^2}{2})\tau+\sigma W_\tau},\sup_{0\leq t\leq\tau}S_0e^{(r-\frac{\sigma^2}{2})t+\sigma W_t}\right)\right].\end{aligned}$$

4.1.1.1. *Transformation du problème.* Nous allons transformer le problème pour se ramener au processus Markovien $\left(B_t,\sup_{0\leq s\leq\tau}B_t\right)$ (où B est un mouvement Brownien standart) que l'on pourra quantifier hors-ligne (indépendamment des paramètres de drift r, μ et de volatilité σ).

Considérons la probabilité $\mathbb{Q}$ définie par

$$\frac{\partial\mathbb{Q}}{\partial\mathbb{P}}_{|t}=\exp\left(-\theta W_t-\frac{\theta^2}{2}t\right),$$

$$\theta=\frac{r-\mu-\frac{\sigma^2}{2}}{\sigma}.$$

et B un mouvement Brownien sous $\mathbb{Q}$ défini par

$$B_t = W_t + \theta t.$$

Le théorème de Girsanov nous donne le changement de probabilité suivant :

$$P_0 = \text{Esssup}_{\tau \in \mathcal{T}_T} \mathbb{E}^{\mathbb{Q}} \left[e^{\theta B_\tau - \frac{\theta^2 \tau}{2}} h\left(\tau, S_0 e^{\sigma B_\tau}, \sup_{0 \leq t \leq \tau} S_0 e^{\sigma B_t}\right)\right].$$

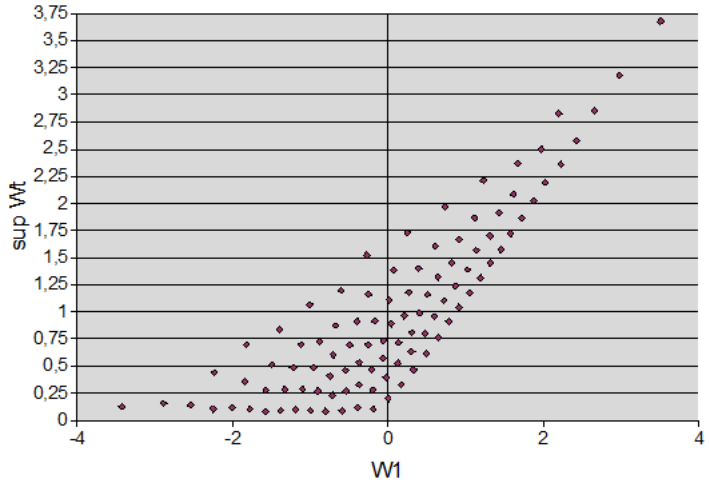
A l'aide d'un scaling, le processus $B'_s = \frac{B_{Ts}}{\sqrt{T}}$ $(s \in [0,1])$ est un mouvement brownien standart, ce qui implique que

$$P_0 = \text{Esssup}_{\tau \in \mathcal{T}_{[0,1]}} \mathbb{E}\left[\tilde{h}\left(\tau, B'_\tau, \sup_{0 \leq s \leq \tau} B'_s\right)\right]$$

avec

$$\tilde{h}(t,x,m) = e^{\theta\sqrt{T}x - \frac{\theta^2 T}{2} t} h\left(t\sqrt{T}, S_0 e^{\sigma\sqrt{T}x}, S_0 e^{\sigma\sqrt{T}m}\right).$$

Comme $\left(B'_{t_k}, \sup_{0 \leq s \leq t_k} B'_s\right)$ et $\sqrt{t_k}\left(W_1, \sup_{0 \leq s \leq 1} W_s\right)$ ont même loi, il suffit alors de quantifier une fois pour toute la loi de $\left(W_1, \sup_{0 \leq s \leq 1} W_s\right)$.
Quantification optimale de la loi $\left(W_1, \sup_{0 \leq s \leq 1} W_s\right)$ dans $\mathbb{R}^2$ avec $M = 100$ points :



4.1.1.2. *Résultats numériques.* On applique la méthode de quantification par arbre avec une restriction sur le nombre de cellules pour le calcul des probabilités de transition.
Le payoff est $h(t, S_t, \sup_{0 \leq u \leq t} S_u)$ avec $h(t,x,m) = e^{-rt} \sup(m - x, 0)$.

Prenons les paramètres suivants : $r = 0.1$, $T = 1$ $S_0 = 100$. Le nombre de trajectoires pour le calcul des probabilités de transition est 10^6. Le nombre de pas de temps vaut $n = 5$.

Résultats numériques avec une volatilité $\sigma = 20\%$:

N1	Nn	Prix
10	18	12,291
30	52	12,730
50	86	12,958
100	171	13,180
150	257	13,328
200	342	13,405
250	428	13,467
300	500	13,397
350	599	13,405

FIG. 4.1.1. Tableau des résultats numériques pour une option lookback américaine avec une volatilité $\sigma = 20\%$: comparaison selon la taille de la grille de quantification.

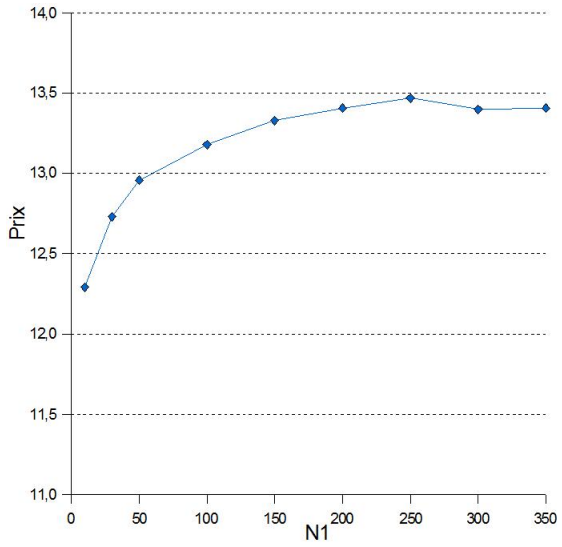


FIG. 4.1.2. Graphique des résultats numériques pour une option lookback américaine avec une volatilité $\sigma = 20\%$: comparaison selon la taille de la grille de quantification.

Résultats numériques avec une volatilité $\sigma = 40\%$:

N1	Nn	Prix
10	18	28,626
30	52	30,708
50	86	31,520
100	171	31,926
150	257	32,160
200	342	32,156
250	428	32,284
300	500	32,143
350	599	32,139

FIG. 4.1.3. Tableau des résultats numériques pour une option lookback américaine avec une volatilité $\sigma = 40\%$: comparaison selon la taille de la grille de quantification.

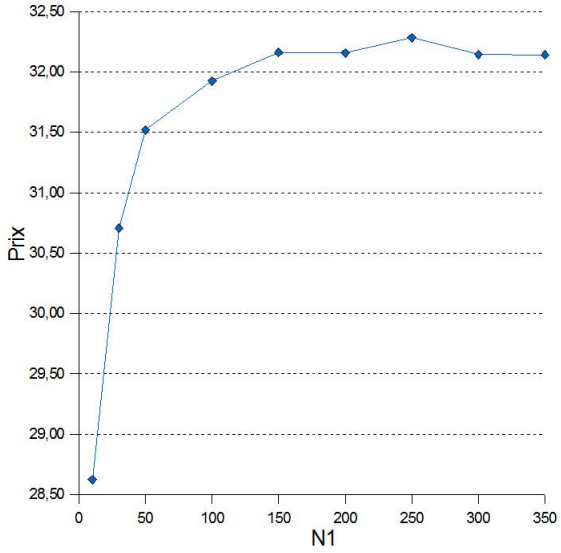


FIG. 4.1.4. Graphique des résultats numériques pour une option lookback américaine avec une volatilité $\sigma = 40\%$: comparaison selon la taille de la grille de quantification.

4.1.2. Option asiatique. Ici, on considère le problème de la valorisation d'une option américaine asiatique : on pose $X_t = (S_t, A_t)$ avec $A_t = \frac{1}{t}\int_0^t S_s ds$. Le payoff est

$$h(t, S_t, A_t) = e^{-rt} \sup\{S_t - A_t, 0\}.$$

4.1.2.1. *Réduction de l'erreur avec des variables de contrôle.* Nous allons améliorer la convergence de l'algorithme de quantification à l'aide de variables de contrôle.

On peut récrire le problème d'enveloppe de Snell sous la forme

$$Y_0 = \text{Esssup}_{\tau \in \mathcal{T}^\pi} \mathbb{E}_{t_k}\left[h(t_\tau, \hat{X}_\tau)\right] = \text{Esssup}_{\tau \in \mathcal{T}^\pi} \hat{\mathbb{E}}\left[\hat{h}\left(\tau, \hat{X}_\tau\right)\right] + P_0(X_0)$$

où $\hat{X}_k$ est une quantification de la variable aléatoire X_k et où les fonctions P et $\hat{h}$ sont définies selon la variable de contrôle.
◇ Dans le cas sans variable de contrôle, on a

$$\hat{h}\left(t_k, X_{t_k}\right) = h\left(t_k, \hat{X}_k\right),$$
$$P_0 = 0.$$

◇ Dans le cas "control 1" : la variable de contrôle est l'option européenne de maturité T :

$$\hat{h}\left(t_k, X_{t_k}\right) = h\left(t_k, \hat{X}_k\right) - P_k(\hat{X}_k),$$
$$P_k(\hat{X}_k) = \mathbb{E}\left[h\left(T, X_T\right)|\hat{X}_k\right].$$

◇ Dans le cas "control 2" : la variable de contrôle est l'option européenne de maturité T et la barrière est conditionnée par rapport à $\sigma(\hat{X}_k)$:

$$\hat{h}\left(t_k, X_{t_k}\right) = \mathbb{E}\left[h\left(t_k, X_{t_k}\right)|\hat{X}_k\right] - P_k(\hat{X}_k),$$
$$P_k(\hat{X}_k) = \mathbb{E}\left[h\left(t_k, X_T\right)|\hat{X}_k\right].$$

4.1.2.2. *Résultats numériques.* Les paramètres de l'option sont : $S_0 = 100$, $r = 0.1$, $T = 1$. Le nombre de pas de temps vaut $n = 20$.

Pour la construction de la grille de quantification, on commence par appliquer l'algorithme CLVQ avec $M_{init} = 100000$ trajectoires, puis on régularise cette grille en utilisant l'algorithme de Lloyd (50 fois avec 20000 trajectoires).

On utilise $M = 10^6$ trajectoires pour le calcul des probabilités de transition.

N1	Normal	Control1	Control2
10	8,058	8,150	8,151
50	8,377	8,373	8,388
100	8,387	8,411	8,433
150	8,393	8,417	8,412
200	8,392	8,419	8,448
250	8,425	8,441	8,446
300	8,419	8,457	8,478
350	8,436	8,472	8,481
400	8,427	8,469	8,490
450	8,437	8,469	8,477
500	8,441	8,474	8,483
550	8,439	8,482	8,492
600	8,428	8,467	8,487
650	8,444	8,478	8,499
700	8,451	8,500	8,515
750	8,443	8,488	8,501

FIG. 4.1.5. Tableau des résultats numériques pour une option américaine asiatique en fontion de la taille de la grille N_1. On compare les trois cas (sans variable de contrôle, le cas control 1 et control 2).

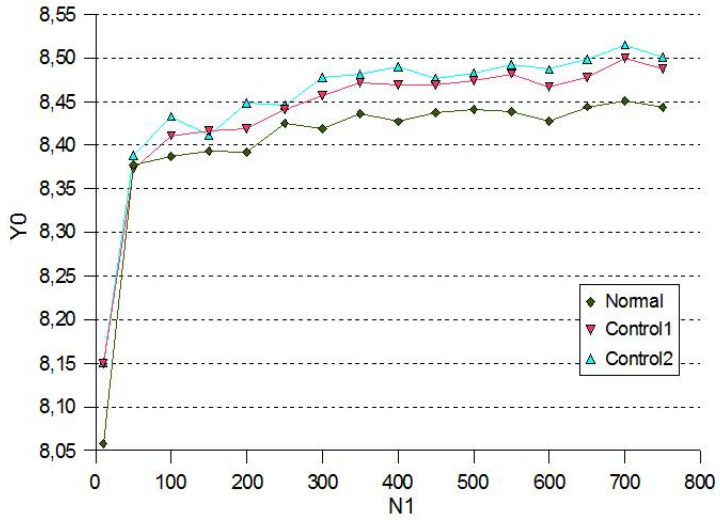


FIG. 4.1.6. Comparaison des trois méthodes (sans variable de contrôle, le cas control 1 et control 2) par rapport à la taille de la grille N_1 pour une option américaine asiatique.

4.2. Equations différentielles stochastiques rétrogrades

Nous appliquons maintenant les algorithmes par quantification pour la résolution d'équations différentielles stochastiques rétrogrades (EDSR).

Soit X_t un processus Markovien.

L'EDSR est définie par

$$Y_t = h(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_s^T Z_s dW_s$$
$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s$$

On considère $\pi = \{t_0 = 0 < t_1 < ... < t_n = T\}$ une partition en temps et une grille $\Gamma_k = \{x_1^k, ..., x_{N_k}^k\}$ une grille de quantification avec sa partition associée $\overline{C_i^{(k)}} = \{x \in \mathbb{R}^d | \left|x - x_i^k\right| \leq \left|x - x_j^k\right| \quad \forall 1 \leq j \leq n\}$. On note $\text{Proj}_{\Gamma_k}(x) = \sum_{i=1}^{N_k} x_i^k 1_{x \in C_i^{(k)}}$.

La convergence des schémas par quantification est prouvée dans la partie précédente et l'erreur de quantification est donnée par

$$\varepsilon(\text{quantif}) \leq C \sum_{k=0}^{n} \left|\left| X_{t_k} - \hat{X}_{t_k} \right|\right|_p$$

avec $\hat{X}_{t_k}$ une quantification de la variable aléatoire X_{t_k}.

Pour la reconstruction d'une fonction $g : \mathbb{R}^d \to \mathbb{R}$ à partir de Γ_k, on notera le reste de l'approximation par le plus proche voisin selon la méthode

$\diamond\, r_{\text{interp}}(g, \Gamma_k, x)$ pour l'interpolation linéaire sur la grille Γ_k (interpolation linéaire classique sur une grille cartésienne, sinon il s'agit d'une interpolation à l'aide de la

triangulation de Delaunay) :

$$g(x) \simeq g(\mathrm{Proj}_{\Gamma_k}(x)) + r_{\mathrm{interp}}(g, \Gamma_k, x)$$

$\diamond$ $r_0 = 0$, $r_1(g, \mathrm{Proj}_{\Gamma_k}(x), x)$ et $r_2(g, \mathrm{Proj}_{\Gamma_k}(x), x)$ pour les regressions par des polynomes locaux de degré 1 et 2 à l'aide des p-plus proches voisins.

$$g(x) \simeq g(\mathrm{Proj}_{\Gamma_k}(x)) + r_\epsilon(g, \mathrm{Proj}_{\Gamma_k}(x), x) \quad \epsilon = \{0; 1; 2\}$$

avec

$$r_\epsilon(g, \mathrm{Proj}_{\Gamma_k}(x), x) = \sum_{j=1}^{p} \alpha_k(\mathrm{Proj}_{\Gamma_k}(x)).b_\epsilon(x - \mathrm{Proj}_{\Gamma_k}(x))$$

où pour $x_i^k \in \Gamma_k$et $\left\{x_{i_i}^k, .., x_{i_p}^k\right\}$ les p plus proches voisins de x_i^k dans $\Gamma_k \setminus \left\{x_i^k\right\}$ et pour $b = \{b_j\}$ l'ensemble des polynômes de degré inférieur à ϵ (et supérieur à 1), α_k est défini par régression linéaire

$$\alpha_k(x_i^k) = \mathrm{argmin}_{\alpha \in \mathbb{R}^d} \sum_{j=1}^{p} \left| \hat{y}_k(x_{i_j}^k) - \hat{y}_k(x_i^k) - \alpha.b(x_{i_j}^k - x_i^k) \right|^2$$

On tire M trajectoires $\left\{W_{t_k}^{(m)}, X_{t_k}^{\pi,(m)}\right\}_{0 \leq k \leq n}$ où $X_{t_k}^\pi$ est le schéma d'Euler de X_{t_k} :

$$X_{t_{k+1}}^\pi = X_{t_k}^\pi + (t_{k+1} - t_k) b(X_{t_k}^\pi) + \sigma(X_{t_k}^\pi)(W_{t_{k+1}} - W_{t_k})$$

4.2.1. EDSR linéaire. Nous allons maintenant nous intéresser ici l'équation différentielle stochastique rétrograde linéaire suivante :

$$\begin{aligned} X_t &= W_t \in \mathbb{R}^d, \\ f(t, x, y, z) &= by + c\frac{z.1}{\sqrt{d}} + r, \quad \forall t \in\in [0, T], x \in \mathbb{R}^d, y \in \mathbb{R}, z \in \mathbb{R}^d, \\ h(x) &= \frac{|x|}{\sqrt{d}}, \quad \forall x \in \mathbb{R}^d. \end{aligned}$$

La solution est indépendante de la dimension d du processus Markovien X. Nous pourrons ainsi tester la robustesse de nos algorithmes par rapport à la dimension de l'espace d'état.

On peut résoudre l'EDSR et la solution Y_0 à la date initiale est

$$Y_0 = \exp\left((b - \frac{c^2}{2})T\right) \mathbb{E}\left[\xi \exp(\frac{c}{\sqrt{d}} W_T)\right] + \frac{r}{b}\left(\exp(bT) - 1\right).$$

On pourra ainsi comparer nos résultats numériques issus de nos algorithmes avec ceux obtenus par la méthode de Monte Carlo.

Donnons un exemple : on prend les paramètres

$$\begin{aligned} b &= c = 1, \\ r &= 0, \\ T &= 1. \end{aligned}$$

Le résultat par Monte Carlo donne $Y_0 = 3.171$ et $Z_0 = 1.856$. Or $\mathbb{E}[Y_T] = \mathbb{E}\left[|W_1|\right] \simeq 0.7$ et $\mathbb{E}\left[Z_T\right] = \mathbb{E}\left[\mathrm{sign}(W_T)\right] = 0$, ce qui montre que le resultat va dépendre essentiellement de la robustesse du calcul de Y_{t_k} et de Z_{t_k} (qui sont peu expliqués par leur valeur finale).

4.2.1.1. *Algorithme par régression linéaire.* Le schéma par régression linéaire d'ordre $\epsilon \in \{0;1;2\}$ est donné par programmation dynamique rétrograde :

Initialement,

$$Y_T^{(m)} = h(X_T^{(m)}), \quad \forall 1 \leq m \leq M.$$

Puis $0 \leq k \leq n-1$, on calcule

$$\overline{y}_k^M(x_i^k) = \frac{\sum_{m=1}^M 1_{X_{t_k}^{\pi,(m)} \in C_i^{(k)}} Y_{t_{k+1}}^{(m)}}{\sum_{m=1}^M 1_{X_{t_k}^{\pi,(m)} \in C_i^{(k)}}}, \quad 1 \leq i \leq N_k, 0 \leq k \leq n-1,$$

$$z_k^M(x_i^k) = \frac{\sum_{r=1}^m 1_{X_{t_k}^{\pi,r)} \in C_i^{(k)}} Y_{t_{k+1}}^{(m)} \frac{W_{t_{k+1}}^{(m)} - W_{t_k}^{(m)}}{t_{k+1}-t_k}}{\sum_{r=1}^m Y_{t_{k+1}}^{(m)} 1_{X_{t_k}^{\pi,(m)} \in C_i^{(k)}}}, \quad 1 \leq i \leq N_k, 0 \leq k \leq n-1,$$

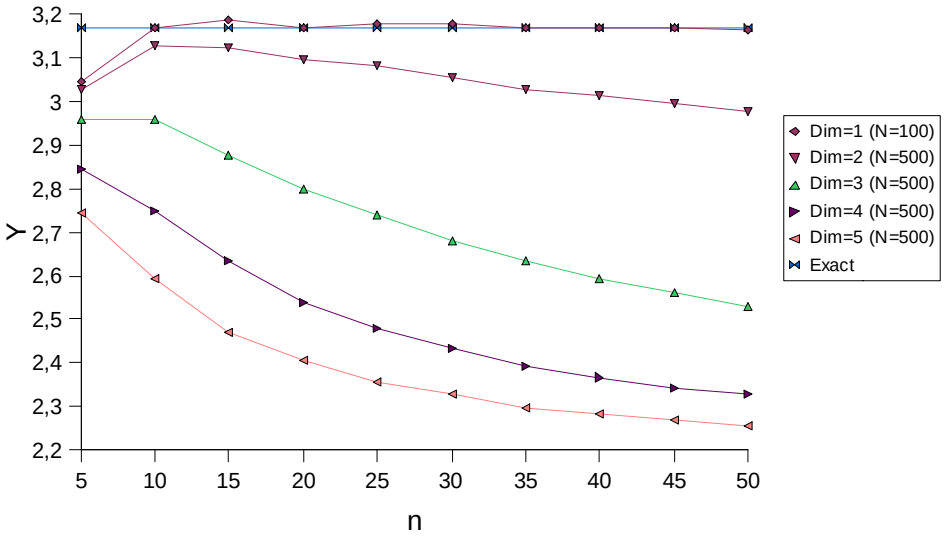
$$y_k^M(x_i^k) = \overline{y}_k^M(x_i^k) + (t_{k+1} - t_k) f(t_k, x_i^k, \overline{y}_k^M(x_i^k), z_k^M(x_i^k)), \quad 1 \leq i \leq N_k, 0 \leq k \leq n-1,$$

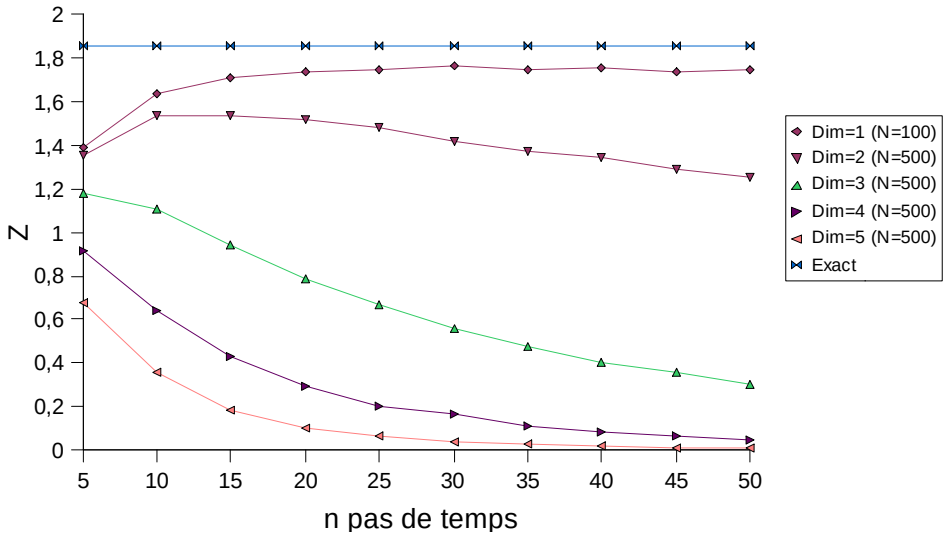
puis

$$Y_{t_k}^{(m)} = y_k^M(\mathrm{Proj}_{\Gamma_k}(X_{t_k}^{\pi,(m)})) + r_\epsilon(y_k^M, \mathrm{Proj}_{\Gamma_k}(X_{t_k}^{\pi,(m)}), X_{t_k}^{\pi,(m)}). \tag{4.2.1}$$

Le calcul de $y_k^M(x_i^k)$ et $z_k^M(x_i^k)$ nécessite une recherche d'un grand nombre de plus proches voisins. Pour accélérer ce calcul, on utilisera un algorithme rapide (PAT) de recherche de plus proche voisin basé sur un arbre d'axes principaux (voir [**McN01**]).

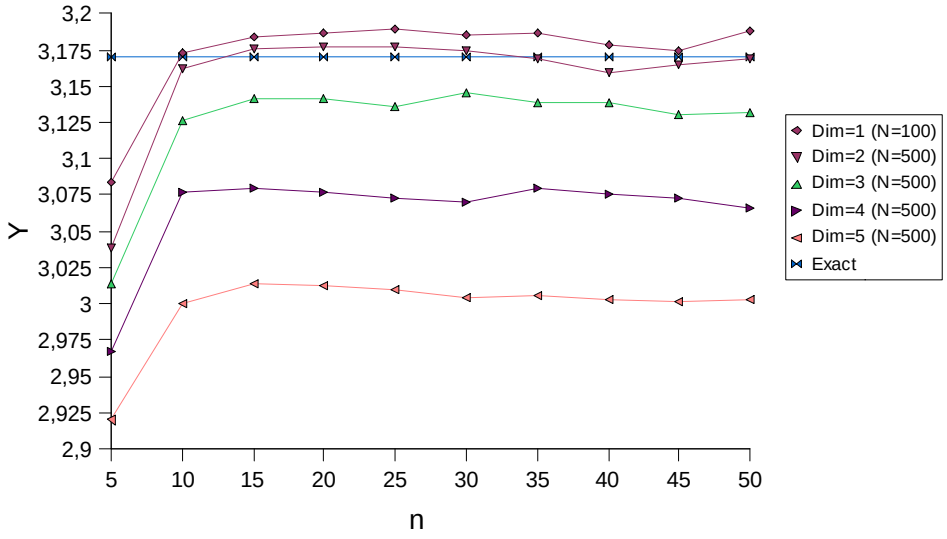
Avec $r_0 = 0$, on obtient pour la fonction valeur Y et son gradient :

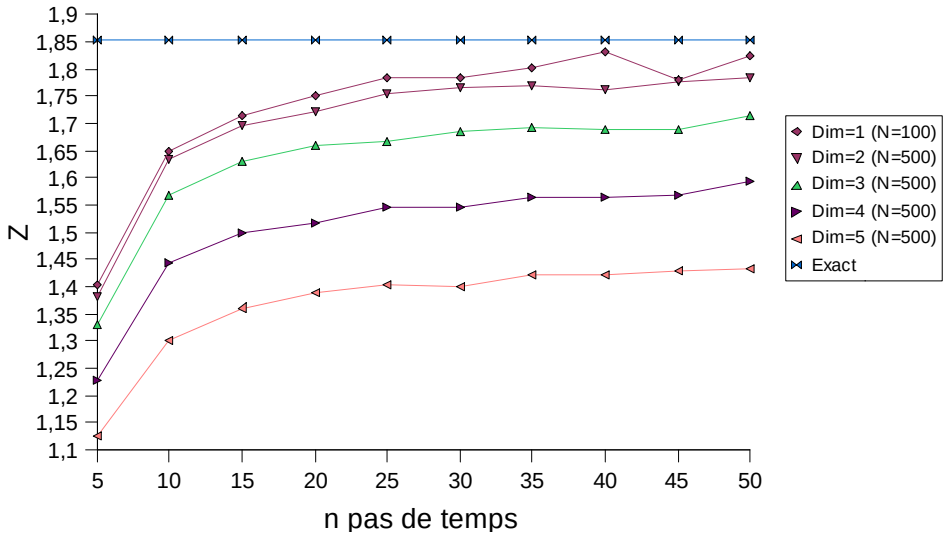




La convergence est améliorée avec le schéma à l'ordre 2 ($\epsilon = 2$).

Avec r_2 polynôme d'ordre 2, on obtient pour la fonction valeur Y et son gradient :





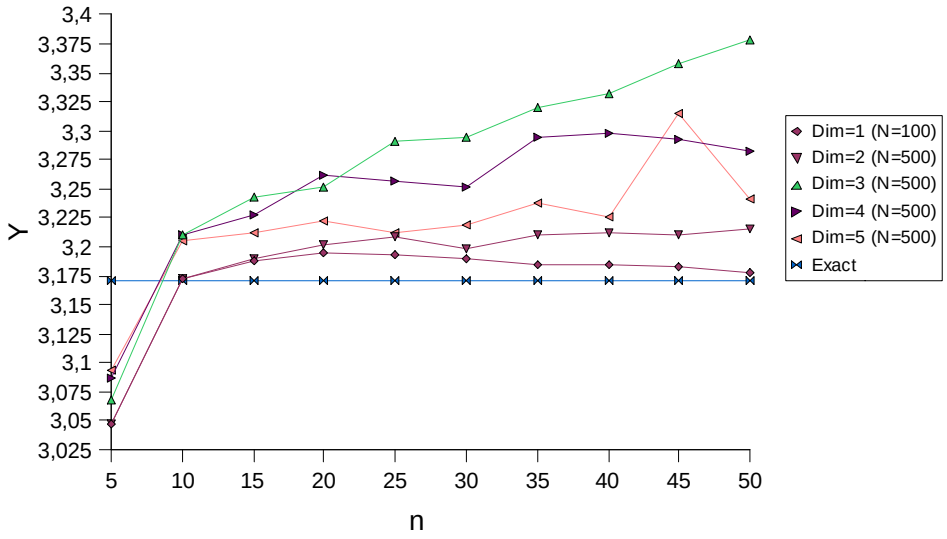
Il est possible d'améliorer cette algorithme à l'aide d'un schéma progressif donné par l'article de Bender et Denk ([**BR07**]) en remplaçant (4.2.1) par

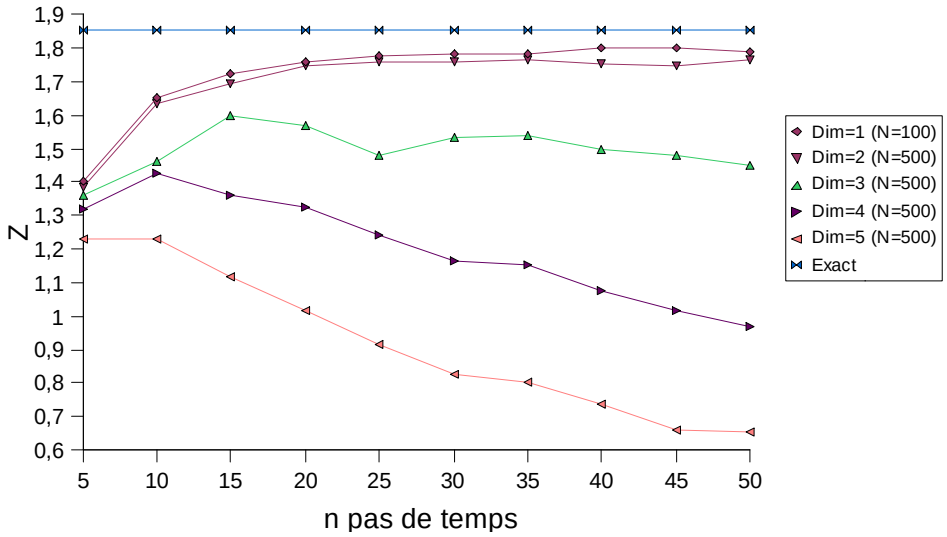
$$Y_{t_k}^{(m)} = Y_{t_{k+1}}^{(m)} + r_\epsilon \left(\tilde{f}_k, \mathrm{Proj}_{\Gamma_k}(X_{t_k}^{\pi,(m)}), X_{t_k}^{\pi,(m)} \right)$$

avec

$$\tilde{f}_k(x) = (t_{k+1} - t_k) f\left(t_k, x, \overline{y}_k^M(x), z_k^M(x)\right), \quad \forall x \in \Gamma_k.$$

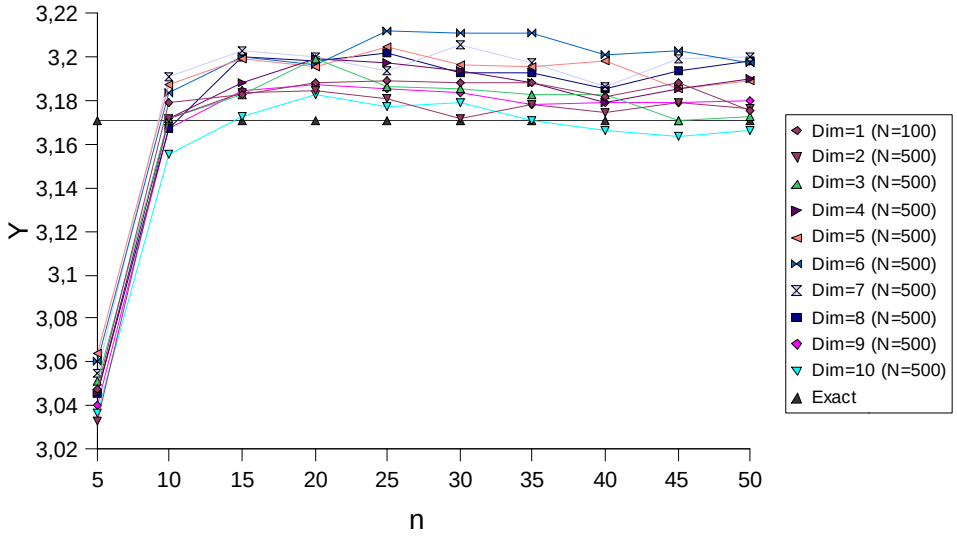
Avec $r_0 = 0$, on obtient pour la fonction valeur Y et son gradient :

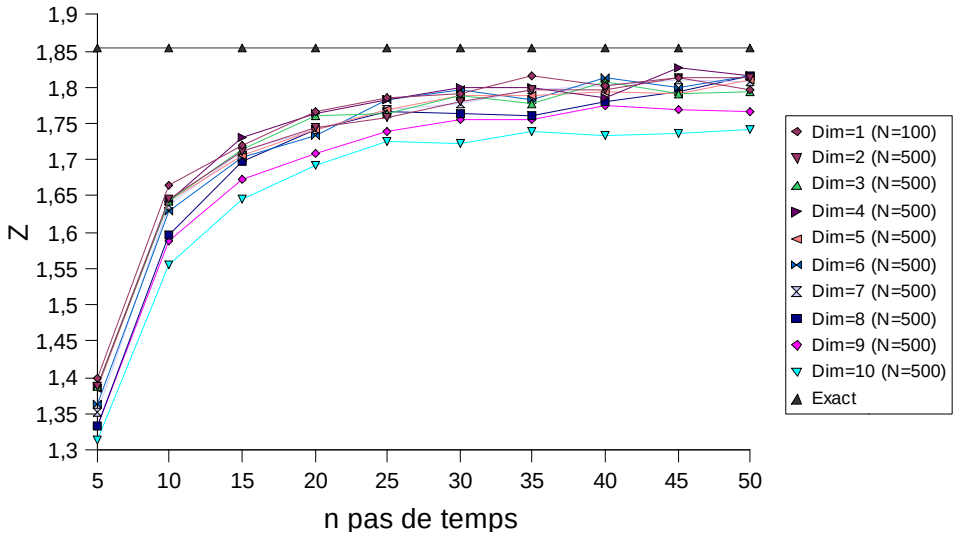




La convergence est améliorée avec le schéma à l'ordre 2 ($\epsilon = 2$).

Avec r_2 polynôme d'ordre 2, on obtient pour la fonction valeur Y et son gradient :





Ce dernier schéma fonctionne bien en grande dimension.

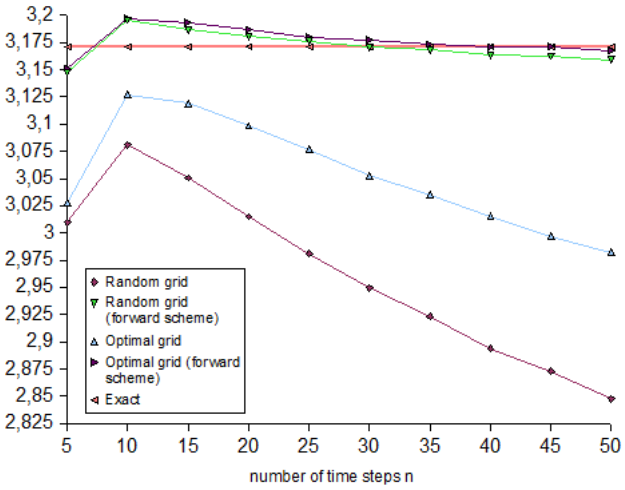
A la date t_k, il s'agit d'une regression linéaire sur N_k vecteurs orthogonaux $1_{x \in C_i^{(k)}}$. En reconstruisant à l'aide de polynômes locaux, on a une approximation de $Y_{t_k} = y_k(X_{t_k})$ sur p vecteurs indépendants avec $p = N_k(1+d)$ si $\epsilon = 1$ et $p = N_k(1 + d + \frac{d(d+1)}{2})$ si $\epsilon = 2$ ($X_{t_k} \in \mathbb{R}^d$). Ainsi, cette reconstruction locale permet d'être robuste avec la dimension d avec un temps de calcul équivalent.

On regarde maintenant l'impact de la grille de quantification optimale.

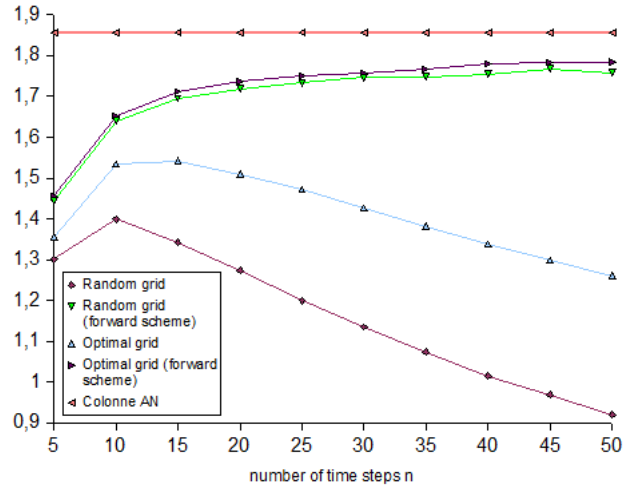
Impact de la grille de quantification optimale. On regarde l'impact du grille aléatoire par rapport à une grille optimale sur le schéma de [**JLW05**] et le schéma progressif de [**BR07**].

Les deux graphes donnent les approximations par régression en dimension 2.

Prix Y_0 :



Incrément martingale Z_0 :



4.2.1.2. *Algorithme par arbre.* On suppose que X est une diffusion elliptique et on note

$$\psi_k(x,\tilde{x}) = \sigma^{-1}(x)\left\{\frac{\tilde{x}-x}{t_{k+1}-t_k} - b(x)\right\}$$

impliquant $\frac{W^{(m)}_{t_{k+1}} - W^{(m)}_{t_k}}{t_{k+1}-t_k} \simeq \psi_k(X^\pi_{t_k}, X^\pi_{t_{k+1}})$.

L'algorithme par arbre est alors défini de manière rétrograde

$$\hat{y}_n(x) = h(x) \quad x \in \Gamma_n$$

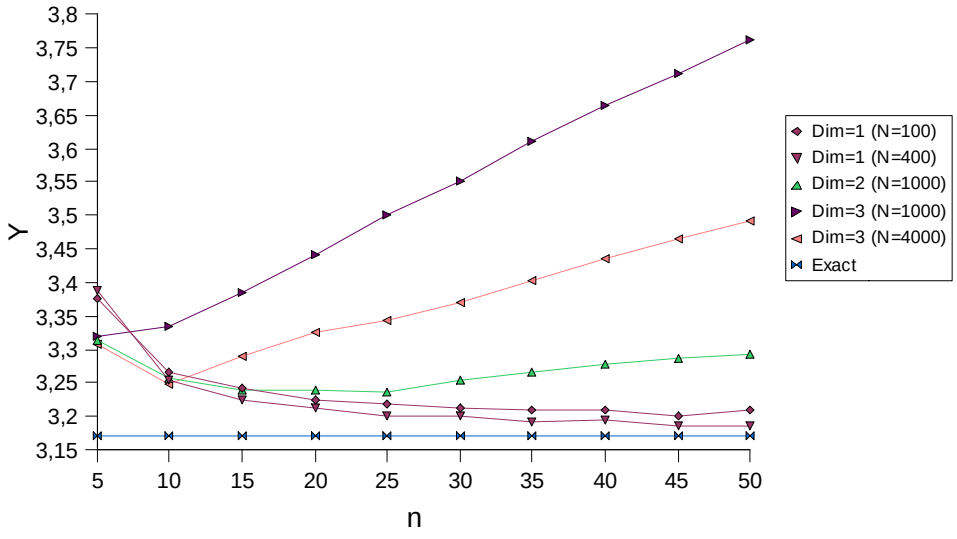
et pour $0 \leq k \leq n-1$ et $x_i^k \in \Gamma_k$

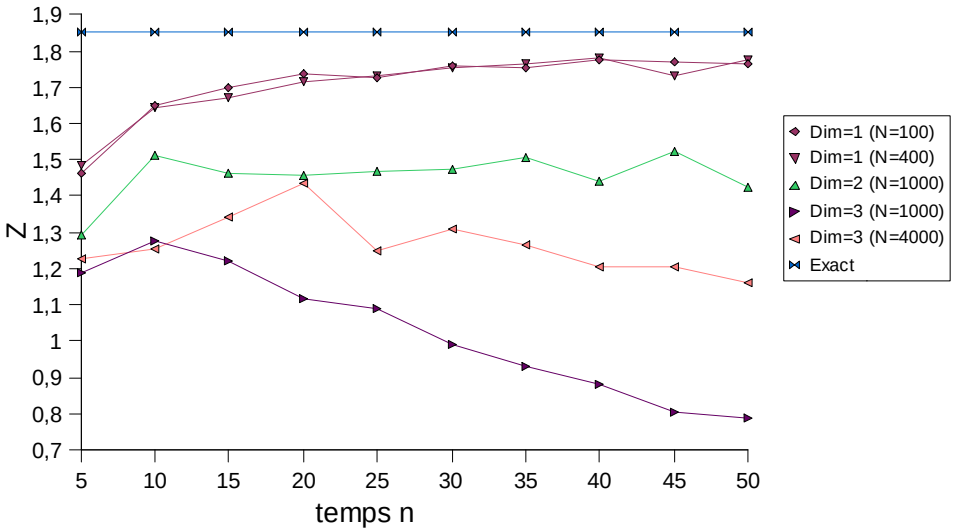
$$\begin{aligned}
\overline{y}_k(x_i^k) &= \sum_{j=1}^{N_{k+1}} p_{i,j}^{k,M} \hat{y}_{k+1}(x_j^{k+1}) \\
\hat{z}_k(x_i^k) &= \sum_{j=1}^{N_{k+1}} p_{i,j}^{k,M} \hat{y}_{k+1}(x_j^{k+1}) \psi_k(x_i^k, x_j^{k+1}) \\
\hat{y}_k(x_i^k) &= \overline{y}_k(x_i^k) + (t_{k+1} - t_k) f(t_k, x_i^k, \overline{y}_k(x_i^k), \hat{z}_k(x_i^k))
\end{aligned}$$

où les probabilités de transition $p_{i,j}^{k,M}$ peuvent être pré-calculées hors ligne :

$$\begin{aligned}
p_{i,j}^{k,M} &= \frac{\sum_{m=1}^{M} 1_{X_{t_k}^{\pi,(m)} \in C_i^{(k)}} 1_{X_{t_{k+1}}^{\pi,(m)} \in C_j^{(k)}}}{\sum_{m=1}^{M} 1_{X_{t_k}^{\pi,(m)} \in C_i^{(k)}}} \\
&\simeq \mathbb{P}\left(X_{k+1}^{\pi} \in C_j^{k+1} | X_k^{\pi} \in C_i^k\right)
\end{aligned}$$

Ici, on a approché $\psi_k(X_{t_k}^{\pi}, X_{t_{k+1}}^{\pi})$ par $\psi_k\left(\mathrm{pr}_{\Gamma_k}(X_{t_k}^{\pi}), \mathrm{pr}_{\Gamma_{k+1}}(X_{t_{k+1}}^{\pi})\right)$ ce qui rajoute une erreur non négligeable quand la dimension augmente.





Amélioration par extrapolation de Romberg : L'algorithme par arbre avec un grille de quantification optimale de taille $N_k = N$ donne une estimation de l'erreur pour $\hat{y}^N$:

$$\left|\left|\hat{y}_0^\infty(X_0) - \hat{y}_0^N(X_0)\right|\right|_2 \leq \frac{C_n}{N^{\frac{1}{d}}}$$

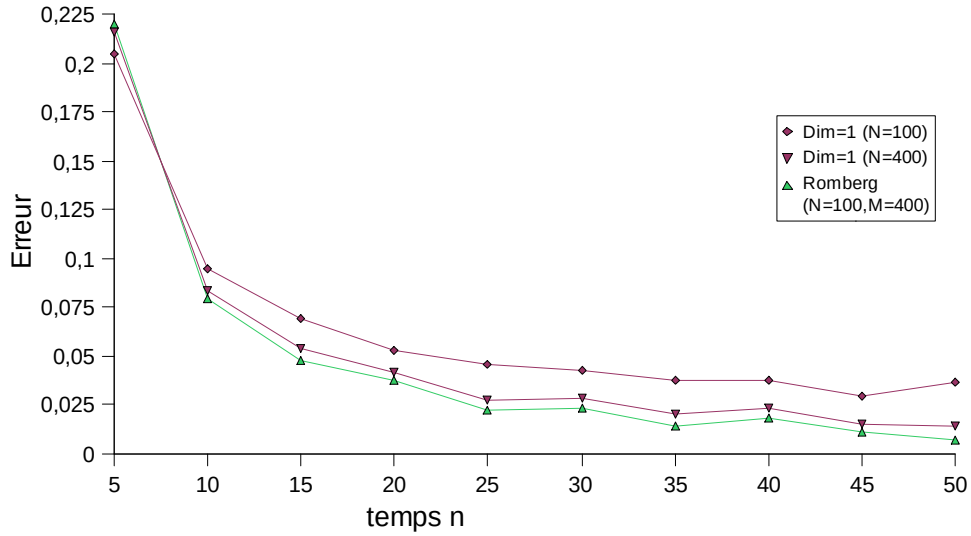
où C_n est une constante indépendante de la taille de grille N (mais dépendant du nombre de pas de temps n). On fait la conjecture que

$$\hat{y}_0^N(X_0) = \hat{y}_0^\infty(X_0) + \frac{C_n}{N^{\frac{1}{d}}} + o(\frac{1}{N^{\frac{1}{d}}}).$$

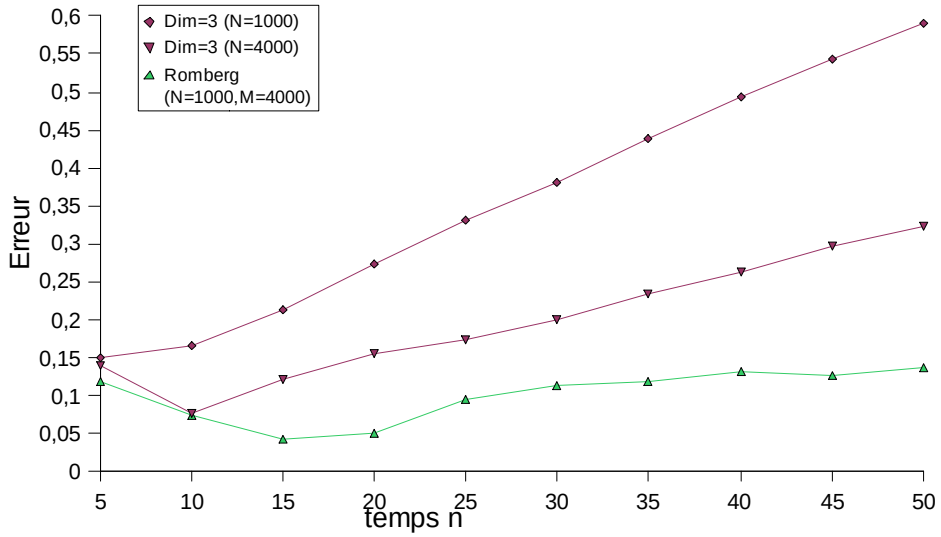
On regarde alors l'approximation de Romberg : pour $N \ll N'$ (on prendra dans les exemple $N' = 4N$)

$$\begin{aligned}\hat{y}_0^{N,N'}(X_0) &= \frac{N'^{\frac{1}{d}}\hat{y}_0^{N'}(X_0) - N^{\frac{1}{d}}\hat{y}_0^N(X_0)}{N'^{\frac{1}{d}} - N^{\frac{1}{d}}} \\ &= \hat{y}_0^\infty(X_0) + o(\frac{1}{N^{\frac{1}{d}}}).\end{aligned}$$

En dimension 1 :



En dimension 3 :



4.2.1.3. *Méthode par gerbes.* On approche la loi normale $\xi_{k+1} = \frac{W^{(m)}_{t_{k+1}} - W^{(m)}_{t_k}}{\sqrt{t_{k+1}-t_k}} \sim \mathcal{N}(0; I_{d_W})$ de manière discrète $\left(\hat{\xi}\right)$ avec M points. On note

$$F_k(x, B) = x + (t_{k+1} - t_k)b(x) + \sqrt{t_{k+1} - t_k}\sigma(x)B$$

i.e. $X^\pi_{k+1} = F_k(X^\pi_k, \frac{W^{(m)}_{t_{k+1}} - W^{(m)}_{t_k}}{\sqrt{t_{k+1}-t_k}})$. On considère $\gamma = d/\mathbb{E}\left[\hat{\xi}^T\hat{\xi}\right]$, ce qui implique $\mathbb{E}\left[\xi^T_{k+1}\xi_{k+1}\right] = \gamma\mathbb{E}\left[\hat{\xi}^T\hat{\xi}\right]$.

L'algorithme par la méthode des gerbes est alors défini de manière rétrograde

$$\hat{y}_n(x) = h(x), \quad x \in \mathbb{R}^d,$$

et pour $0 \leq k \leq n-1$

$$\begin{aligned}
\overline{y}_k(x_i^k) &= \mathbb{E}\left[\hat{y}_{k+1}\left(F_k(x_i^k, \hat{\xi})\right)\right], \quad x_i^k \in \Gamma_k, \\
\hat{z}_k(x_i^k) &= \mathbb{E}\left[\hat{y}_{k+1}\left(F_k(x_i^k, \hat{\xi})\right)\frac{\gamma\hat{\xi}}{\sqrt{t_{k+1}-t_k}}\right], \quad x_i^k \in \Gamma_k, \\
\hat{y}_k(x_i^k) &= \overline{y}_k(x_i^k) + (t_{k+1}-t_k) f(t_k, x_i^k, \overline{y}_k(x_i^k), \hat{z}_k(x_i^k)), \quad x_i^k \in \Gamma_k,
\end{aligned}$$

puis on prolonge sur $\mathbb{R}^d$

$$\hat{y}_k(x) = \hat{y}_k(x_i^k) + r(\mathrm{Proj}_{\Gamma_k}(x), x)$$

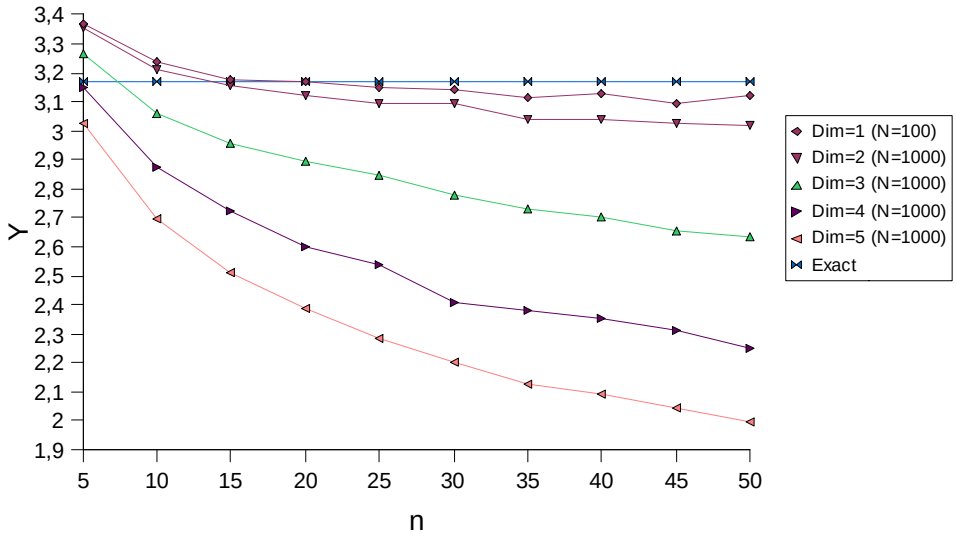
avec $r(\mathrm{Proj}_{\Gamma_k}(x), x) = r_{\mathrm{interp}}(\hat{y}_k, \Gamma_k, x)$ pour le prolongement par interpolation linéaire et $r(\mathrm{Proj}_{\Gamma_k}(x), x) = r_\epsilon(\hat{y}_k, \mathrm{Proj}_{\Gamma_k}(x), x)$ pour le prolongement par des polynômes d'ordre inférieure ϵ.

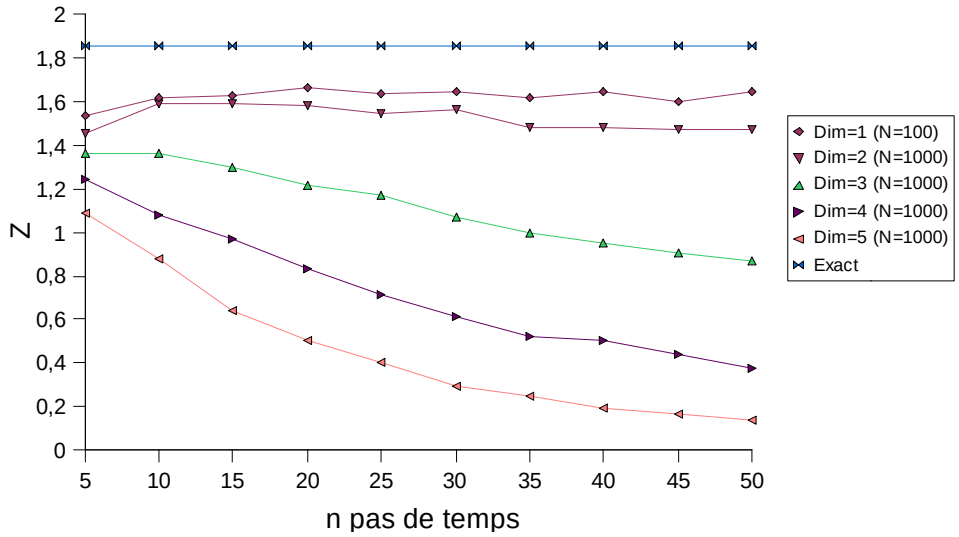
On s'interessera à deux discrétisations du semi-groupe ξ :

$\diamond$ Soit par quantification $\mathbb{L}^2$-optimale avec M points. On a alors la propriété de stationnarité $\mathbb{E}\left[\xi|\hat{\xi}\right] = \hat{\xi}$ et une estimation de l'erreur $\left|\left|\xi - \hat{\xi}\right|\right|_p \leq C_p M^{\frac{1}{d_W}}$. Par symétrie, $\mathbb{E}\left[\hat{\xi}^{2p+1}\right] = 0$. Le paramètre γ permet d'avoir une meilleure approximation du gradient z_k. On a $\gamma = \gamma_M \to_{M\to+\infty} 1$. Pour une fonction G régulière et par symétrie de $\hat{\xi}$ (et ξ), on a

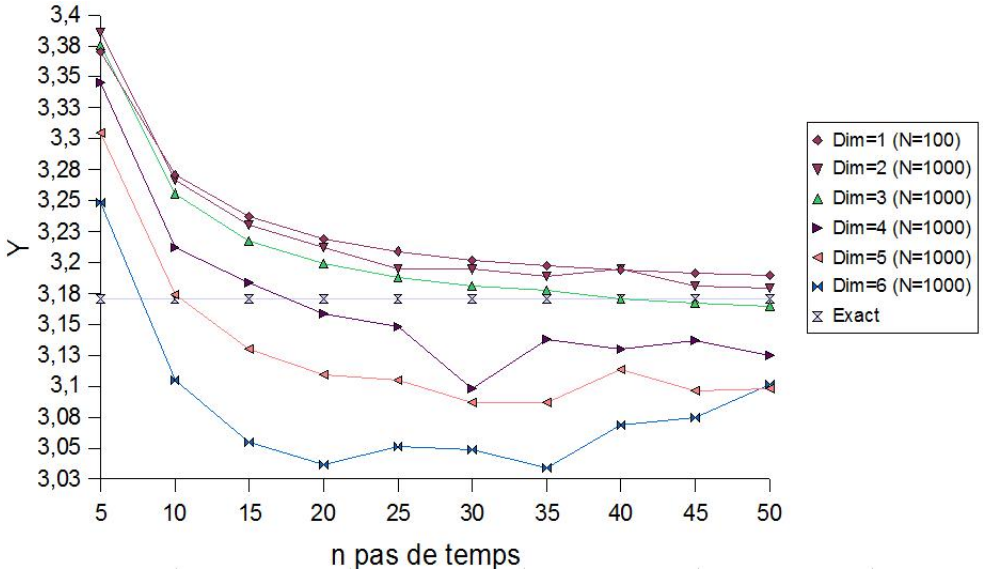
$$\begin{aligned}
\mathbb{E}\left[G\left(x + \sqrt{t_{k+1}-t_k}\hat{\xi}\right)\frac{\gamma\hat{\xi}}{\sqrt{t_{k+1}-t_k}}\right] &= \nabla G(x)\gamma\mathbb{E}\left[\hat{\xi}^T\hat{\xi}\right] + O(t_{k+1}-t_k) \\
&= \mathbb{E}\left[G\left(x + W_{t_{k+1}} - W_{t_k}\right)\frac{W_{t_{k+1}} - W_{t_k}}{t_{k+1}-t_k}\right] + O(t_{k+1}-t_k)
\end{aligned}$$

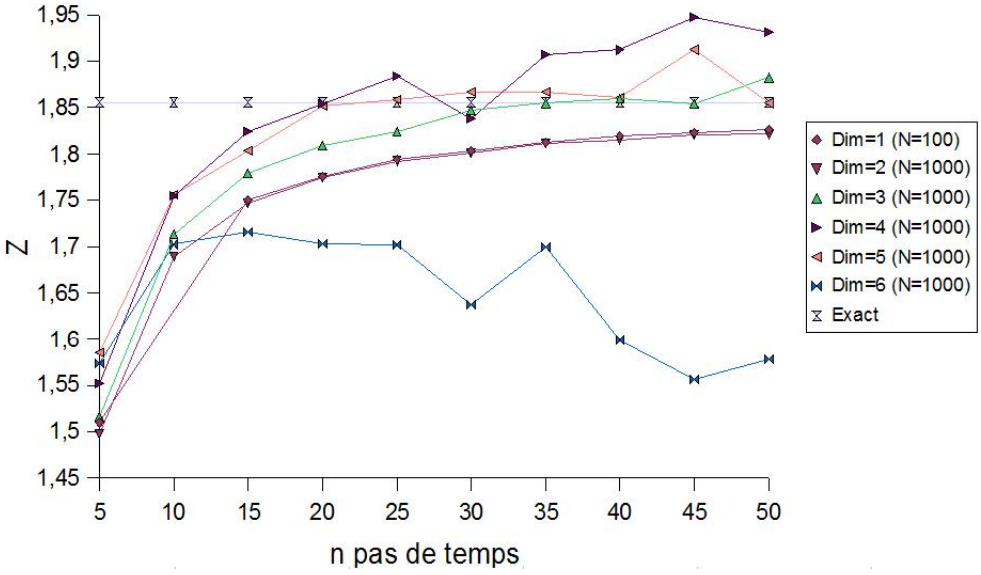
A l'ordre 0 :





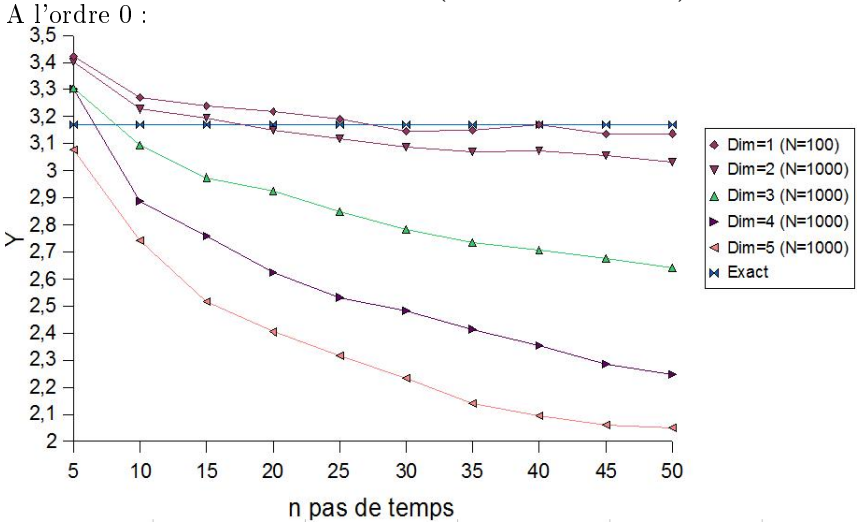
A l'ordre 2, l'erreur d'approximation diminue :

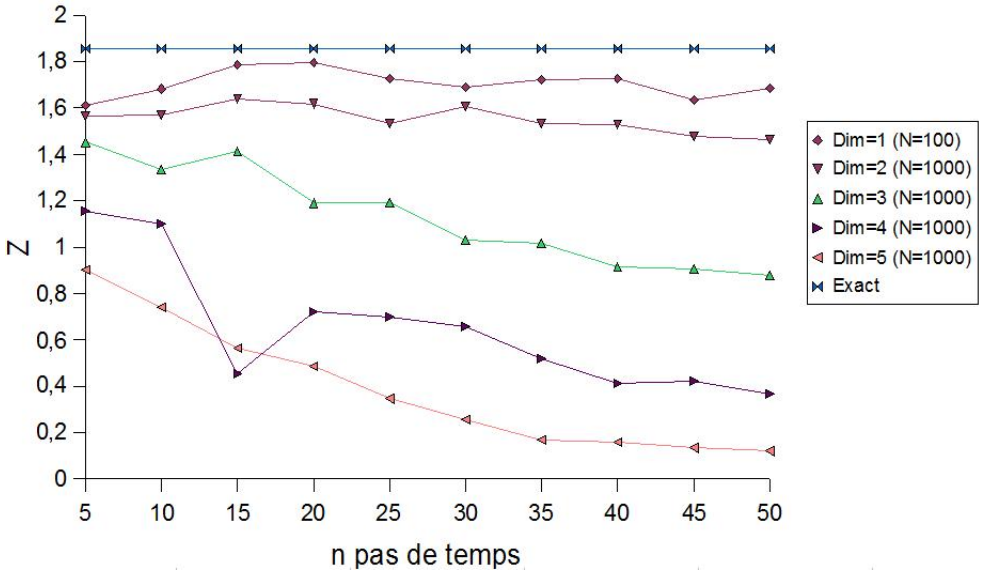




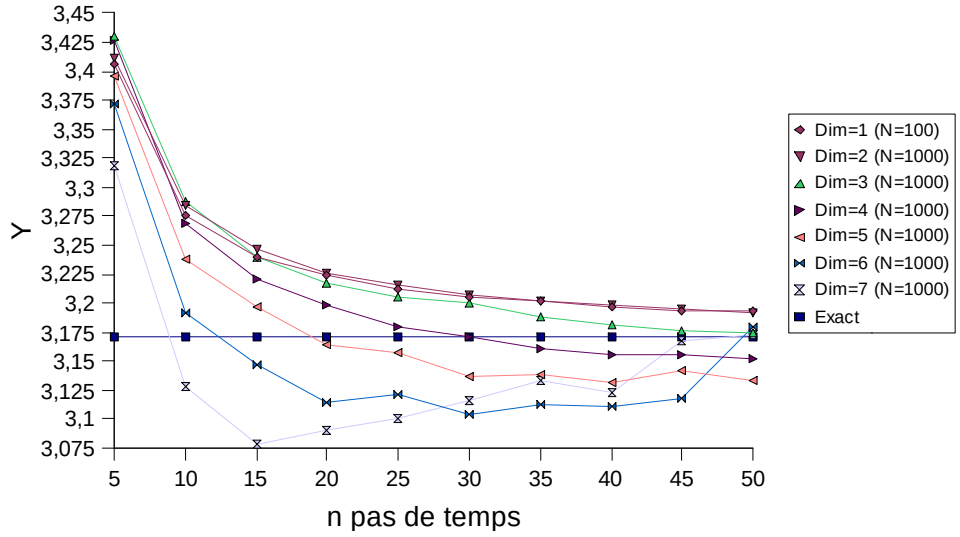
$\diamond$ Soit par une marche aléatoire à p pas. A un pas, on prend $\hat{\xi}$ à valeurs dans les $2d_W$ points $\left\{\pm\sqrt{d_W}e_i\right\}_{1\leq i\leq d_W}$ avec $\mathbb{P}(\hat{\xi} = \pm\sqrt{d_W}e_i) = \frac{1}{2d_W}$. Pour nos tests numériques, on regardera une marche aléatoire à deux pas, c'est-à-dire $\hat{\xi}$ à valeurs dans les $(1 + d_W + \frac{d_W(d_W+1)}{2})$ points $\left\{\pm\sqrt{\frac{d_W}{2}}e_i \pm \sqrt{\frac{d_W}{2}}e_i\right\}_{1\leq i,\leq d_W}$ avec $\mathbb{P}(\hat{\xi} = 0) = \frac{1}{2d_W}$, $\mathbb{P}(\hat{\xi} = \pm\sqrt{2d_W}e_i) = \frac{1}{4d_W^2}$ et $\mathbb{P}\left(\hat{\xi} = \sqrt{\frac{d_W}{2}}(\pm e_i \pm e_j)\right) = \frac{1}{2d_W}$ $(i \neq j)$.

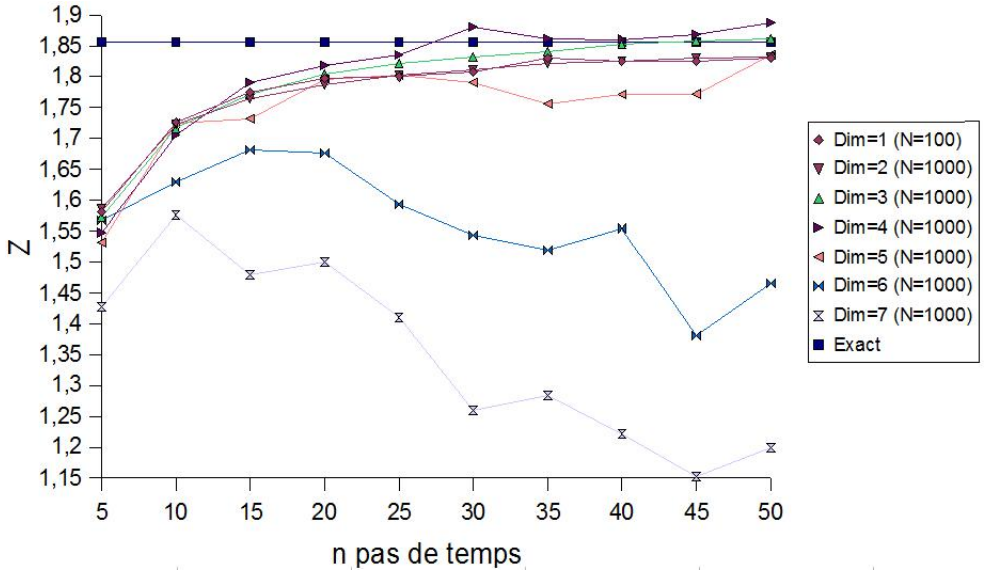
A l'ordre 0 :





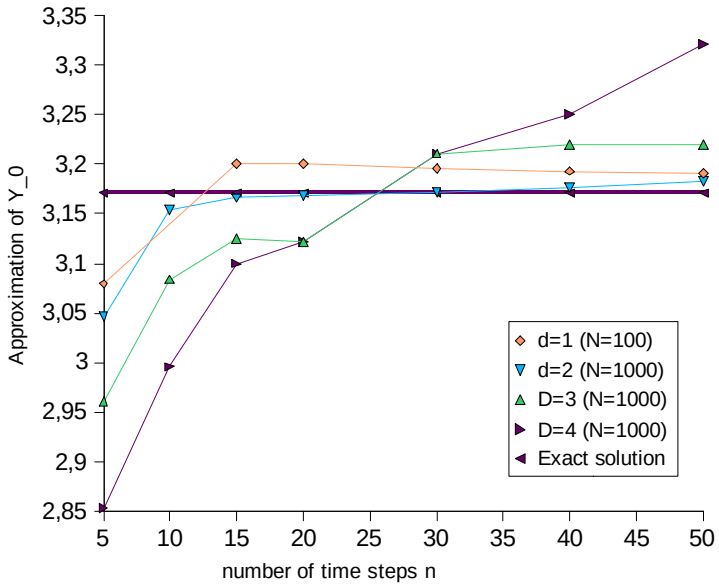
A l'ordre 2, l'erreur numérique diminue :





4.2.1.4. *par interpolation sur un grille quantifiée.* On résout par l'algorithme par interpolation avec $M = 100$ points pour le quantificateur de la loi de transition.

N	100	1000	1000	1000
Nbre pas n	d=1 (N=100)	d=2 (N=1000)	D=3 (N=1000)	D=4 (N=1000)
3	3,19	3,16	3,09	2,98
5	3,25	3,21	3,12	3
10	3,24	3,2	3,13	3,04
15	3,22	3,19	3,15	3,08
20	3,21	3,18	3,17	3,14
25	3,2	3,18	3,2	3,18
30	3,2	3,18	3,23	3,22
35	3,19	3,18	3,27	3,25
40	3,19	3,18	3,3	3,28
45	3,19	3,18	3,33	3,3
50	3,19	3,19	3,36	3,32



4.2.2. EDSR avec contrainte. On considère maintenant une EDSR avec contrainte sur Y et Z. On considère l'EDSR suivante

$$Y_t = h(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds + (A_T - A_t) - \int_t^T Z_s dW_s \tag{4.2.2}$$

telle que

$$d_C(Y_s, Z_s) = 0 \tag{4.2.3}$$

où Γ est un ensemble fermé de $\mathbb{R} \times \mathbb{R}^d$ et d_C est la distance à C. Il est clair que la fonction distance d_C est Lispchtizienne.

Peng ([**Pen99**]) définit la plus petite f-supersolution sur $[0, T]$ de l'EDSR (4.2.2) avec la contrainte (4.2.3).

On définit l'EDSR suivante par pénalisation :

$$Y_t = h(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds + p\int_t^T d_C(Y_s, Z_s)ds - \int_t^T Z_s dW_s$$

La monotonie induit la convergence par pénalisation.

La convergence du schéma en espace est donnée par les résultats précédents pour les ESDR en remarquant que les paramètres du générateurs sont maintenant lipschitziens de rapport p:

$$Y_t^{(p)} \to_{p\to\infty} Y_t$$
$$\left|\left|Y_{t_k}^{(p)} - \hat{Y}_{t_k}^{\pi,(p)}\right|\right| \leq p\epsilon(\text{global})$$

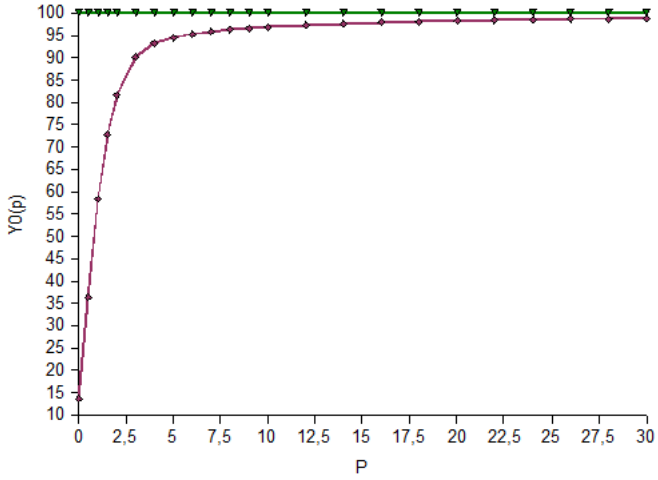
On va chercher maintenant à valoriser une option (put) américaine avec l'interdiction de vente à découvert comme contrainte sur le portefeuille. L'EDSR à résoudre devient alors

$$Y_t^{(p)} = (K - X_\tau)^+ - \int_t^T r Y_s^{(p)} ds + p \int_t^T \left(\frac{Z_s^{(p)}}{\sigma} \right)^- ds - \int_t^T Z_s^{(p)} dW_s$$
$$\tau = \inf \left\{ s \geq t | Y_s^{(p)} \leq (K - X_s)^+ \right\}$$

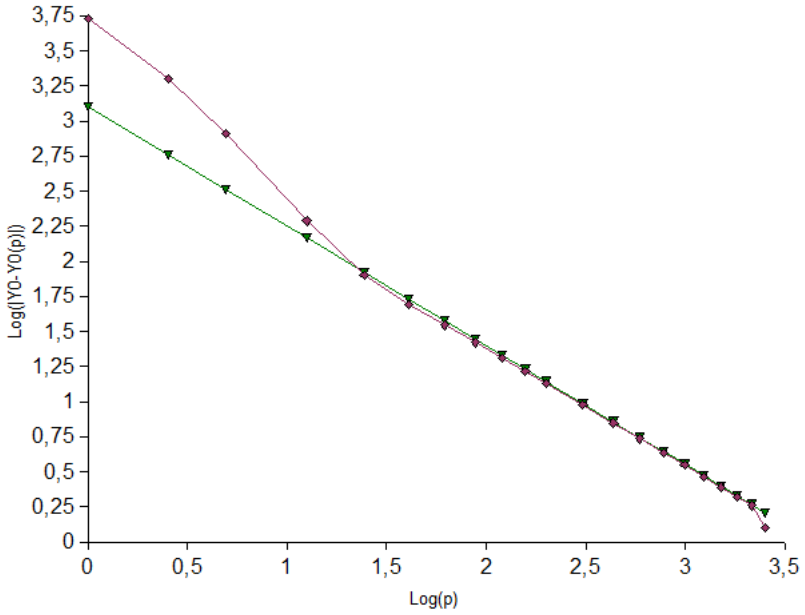
On se place dans le modèle de Black-Scholes avec $\sigma = 40\%$, $r = 0.05$ et $T = 1$. La solution est:

$$Y_t = \begin{cases} K, & t \in [0, T[\\ (K - X_T)^+, & t = T \end{cases}$$
$$\pi_t = 0$$

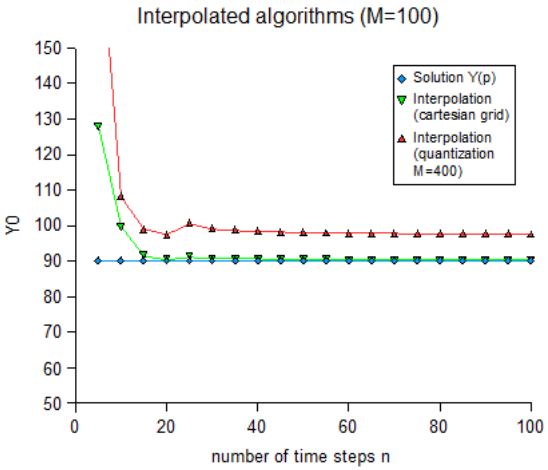
Le premier schéma donne la convergence du schéma pénalisé en fonction de p. On a utilisé le schéma par interpolation avec une (très grande) grille homogène et beaucoup de pas de temps.



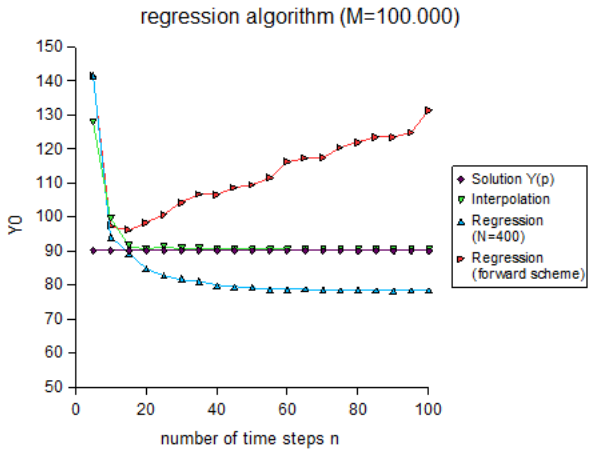
La seconde courbe compare le logarithme de l'erreur. Il semble que $Y_0 - Y_0^{(p)} = O(\frac{1}{p})$.



Le graphe suivant illustre les schémas par interpolation avec un pénalisation $p = 3$.



Le graphe suivant illustre les schémas par régression avec un pénalisation $p = 3$.



4.3. Valorisation en volatilité incertaine

Le modèle à volatilité incertaine développé par [**ALP95**] suppose que l'actif risqué suit l'équation de diffusion contrôlé suivante sous la mesure risque-neutre :

$$\frac{dS_t}{S_s} = \sigma_t dW_t$$

où le processus de volatilité σ_t est incertain, adapté à la filtration Brownienne et appartient à l'intervalle $[\sigma_{min}, \sigma_{max}]$.

On cherche à sur-répliquer une option $h(S_T)$ de maturité T. Le modèle à volatilité incertaine selectionne la volatilité à toute date qui maximise le prix de l'option ; cela revient à regarder le pire des scénarios.

En notant Ξ l'ensemble des volatilités σ, adaptées à la filtration Brownienne, à valeurs dans l'intervalle $[\sigma_{min}, \sigma_{max}]$, le prix de sur-réplication de l'option est à la date t

$$Y_t = \sup_{\sigma \in \Xi} \mathbb{E}\left[h(S_T)|\mathcal{F}_t\right]. \tag{4.3.1}$$

Cette valorisation de l'option est aussi la solution (au sens de viscosité) de l'équation d'Hamilton-Jacobi-Bellman (HJB) avec un contrôle sur la coefficient de diffusion (voir [**Pha09**]) :

$$\begin{aligned} \frac{\partial}{\partial t} y(t,x) + \sup_{\sigma \in [\sigma_{min}, \sigma_{max}]} \left\{ \frac{x^2\sigma^2}{2} \frac{\partial^2}{\partial x^2} y(t,x) \right\} &= 0, \quad \forall t \in [0,T[, \forall x \in \mathbb{R}, \\ y(T,x) &= h(x), \quad \forall x \in \mathbb{R}, \end{aligned}$$

ce qui peut se réécrire sous la forme

$$\begin{aligned} \frac{\partial}{\partial t} y(t,x) + \frac{1}{2} x^2 \Sigma\left(\frac{\partial^2}{\partial x^2} y(t,x)\right)^2 \frac{\partial^2}{\partial x^2} y(t,x) = 0, \quad \forall t \in [0,T[, \forall x \in \mathbb{R}, \\ y(T,x) = h(x), \quad \forall x \in \mathbb{R}, \end{aligned} \tag{4.3.2}$$

avec $\Sigma(v) = \sigma_{min} 1_{v<0} + \sigma_{max} 1_{v \geq 0}$.

Nous allons proposer deux méthodes numériques : on considére soit qu'il s'agit de résoudre une equation aux dérivées partielles de second ordre totalement non linéraire à l'aide d'EDSR du second ordre (voir [**CSTV06**] pour la 2-EDSR et

[**FTW11**] pour sa discrétisation en temps), soit qu'il s'agit d'une optimisation d'un processus de Markov contrôlé.

4.3.1. Résolution par contrôle stochastique. On considère $\pi = \{t_0 = 0 < t_1 < ... < t_n = T\}$ et Ξ_π l'ensemble des volatilités σ adaptées, constantes sur chaque intervalle $[t_k, t_{k+1}[$. Le problème de maximisation (4.3.1) devient

$$Y^\pi_{t_k} = \sup_{\sigma \in \Xi_\pi} \mathbb{E}\left[h(S_T)|\mathcal{F}_{t_k}\right], \quad \forall t_k \in \pi.$$

Alors le processus discret Y^π vérifie le principe de programmation dynamique et nous avons $Y^\pi_{t_k} = y^\pi_k(S_{t_k})$ où la fonction y^π_k est définie de manière rétrograde :

$$\begin{aligned} y^\pi_n &= h \\ y^\pi_k(x) &= \sup_{\sigma \in [\sigma_{min}, \sigma_{max}]} \mathbb{E}\left[y^\pi_{k+1}(S_{t_{k+1}})|S_{t_k} = x\right], \quad \forall 0 \leq k < n, \forall x \in \mathbb{R}. \end{aligned}$$

Pour calculer ces espérances conditionnelles, on peut par exemple utiliser la méthode des gerbes après avoir discrétiser l'espace d'état (en prenant un grille $\delta\mathbb{Z}$ ou bien en quantifiant la variable S_{t_k} en fixant sa volatilité à σ_{max} pour construire la grille). Il suffit alors de calculer

$$y^\pi_k(x) = \sup_{\sigma \in [\sigma_{min}, \sigma_{max}]} \mathbb{E}\left[y^\pi_{k+1}\left(xexp(\sigma\Delta W^\pi_{k+1} - \frac{\sigma^2 \Delta t^\pi_{k+1}}{2})\right)\right], \quad \forall 0 \leq k < n, \forall x \in \mathbb{R}$$

avec $\Delta W^\pi_{k+1} = W_{t_{k+1}} - W_{t_k}$ et $\Delta t^\pi_{k+1} = t_{k+1} - t_k$.

On peut aussi se ramener à une optimisation d'un processus de Markov contrôlé (comme dans l'application dans la partie précédente) en considérant le processus de Markov X vérifiant l'EDS suivant

$$\frac{dX_t}{X_t} = \sigma_{max} dW_t$$

Nous avons alors

$$y^\pi_k(x) = \sup_{\sigma \in [\sigma_{min}, \sigma_{max}]} \mathbb{E}\left[\phi_{k,a}(\Delta W^\pi_{k+1}) y^\pi_{k+1}(X_{t_{k+1}})|X_{t_k} = x\right], \quad \forall 0 \leq k < n, \forall x \in \mathbb{R}$$

avec

$$\phi_{k,\sigma}(w) = \frac{\sigma_{max}}{\sigma} \exp\left(-\frac{(\frac{\sigma_{max}}{\sigma}\frac{\Delta W^\pi_{k+1}}{\Delta t^\pi_{k+1}} + \frac{\sqrt{\Delta t^\pi_{k+1}}}{2\sigma}(\sigma^2 - \sigma^2_{max}))^2 + (\frac{\Delta W^\pi_{k+1}}{\Delta t^\pi_{k+1}})^2}{2}\right)$$

ce qui rentre dans le cadre de nos algorithmes par quantification (2.2.1).

4.3.2. EDSR du second ordre. On dit que (Y, Z, Γ, A) est une solution d'une EDSR du second ordre (voir [**CSTV06**]) si

$$\begin{aligned} dX_t &= \sigma(t, X_t) dW_t, \\ dY_t &= -f(t, X_t, Y_t, Z_t, \Gamma_t)dt + Z_t dX_t, \\ dZ_t &= A_t dt + \Gamma_t dX_t, \\ Y_T &= h(X_T). \end{aligned} \tag{4.3.3}$$

On considère v la solution de l'EDP parabolique totalement non-linéaire suivante

$$\begin{aligned} \frac{\partial}{\partial t}v(t,x) + \frac{1}{2}\mathrm{Tr}(D^2v(t,x)\sigma\sigma'(t,X_t)) + f\left(t,x,v(t,x),Dv(t,x),D^2v(t,x)\right) &= 0 \\ v(T,x) &= h(x) \end{aligned} \tag{4.3.4}$$

Alors sous de bonnes conditions de régularités de v solutions de l'EDP (4.3.4), l'article ([**CSTV06**]) montre que

$$\begin{cases} Y_t = v(t,X_t) \\ Z_t = Dv(t,X_t) \\ \Gamma_t = D^2v(t,X_t) \\ A_t = (\frac{\partial}{\partial t} + \mathcal{L})Dv(t,X_t) \end{cases}$$

est solution de la 2-BSDE (4.4.1), où $\mathcal{L}$ est le générateur infinitésimal du processus X.

En notant $\overline{\sigma} = \sqrt{\frac{\sigma_{min}^2+\sigma_{max}^2}{2}}$ et $\delta\sigma^2 = \frac{\sigma_{max}^2-\sigma_{min}^2}{2}$, l'EDP (4.3.2) pour la valorisation en volatilité incertaine devient

$$\frac{\partial}{\partial t}y(t,x) + \frac{1}{2}\overline{\sigma}^2x^2\frac{\partial^2}{\partial x^2}y(t,x) + \delta\sigma^2x^2\left|\frac{\partial^2}{\partial x^2}y(t,x)\right| = 0, \quad \forall t \in [0,T[, \forall x \in \mathbb{R}.$$

Pour que la 2-BSDE (4.3.3) permet de résoudre la valorisation en modèle de volatilité incertaine, il suffit d'identifier les EDP (4.3.4) et (4.3.2). Il suffit donc par exemple de prendre

$$\begin{aligned} \sigma(t,x) &= x\overline{\sigma}, \quad \forall t \in [0,T[, \forall x \in \mathbb{R}, \\ f(t,x,y,z,\gamma) &= \delta\sigma^2x^2\left|\gamma\right|, \quad \forall t \in [0,T[, \forall x,y,z,\gamma \in \mathbb{R}. \end{aligned}$$

La discrétisation en temps est déterminée par le schéma dans [**CSTV06**] :

◇ condition terminale :

$$Y_T^\pi = h(X_T)$$

◇ processus rétrograde

$$\begin{aligned} Y_{t_k}^\pi &= \mathbb{E}\left[Y_{t_{k+1}}^\pi | X_{t_k}\right] + \Delta t_{k+1}^\pi \times f\left(t_k, X_{t_k}, Y_{t_k}^\pi, Z_{t_k}^\pi, \Gamma_{t_k}^\pi\right), \\ Z_{t_k}^\pi &= \sigma(X_{t_k})^{-1}\mathbb{E}\left[Y_{t_{k+1}}^\pi \frac{\Delta W_{k+1}^\pi}{\Delta t_{k+1}^\pi} | X_{t_k}\right], \\ \Gamma_{t_k}^\pi &= \sigma(X_{t_k})^{-1}\mathbb{E}\left[Z_{t_{k+1}}^\pi \frac{\Delta W_{k+1}^\pi}{\Delta t_{k+1}^\pi} | X_{t_k}\right], \quad \forall 0 \leq k < n. \end{aligned}$$

Nous allons aussi considérer le schéma donné dans [[**FTW11**]] où Γ_{t_k} est calculé en tant que dérivée seconde du processus Y, c'est-à-dire on remplace le terme Γ dans l'algorithme précédent par :

$$\Gamma_{t_k} = \sigma(X_{t_k})^{-2}\mathbb{E}\left[Y_{t_{k+1}}^\pi \frac{\left|\Delta W_{k+1}^\pi\right|^2 - \Delta t_{k+1}^\pi}{\Delta t_{k+1}^\pi} | X_{t_k}\right], \quad \forall 0 \leq k < n.$$

Nous allons appliquer le méthode des gerbes pour la discrétisation spatiale.On approxime ΔW_{k+1}^π par $\Delta\hat{W}_{k+1}^\pi = \sqrt{\Delta t_{k+1}^\pi}\gamma\hat{B}_{k+1}$ où $\hat{B}$ est une quantification $\mathbb{L}_2$-optimale de la loi normale et γ corrige le moment d'ordre 2 de la quantification, i.e.

$$\mathbb{E}\left[\hat{B}\hat{B}^*\right] = \left(\gamma\gamma^*\right)^{-1}.$$

En notant

$$\begin{aligned}\rho &= \frac{\mathbb{E}\left[\left(\frac{\Delta W_{t_{k+1}}}{\sqrt{\Delta t^\pi_{k+1}}}\right)^4 - \left(\frac{\Delta W_{t_{k+1}}}{\sqrt{\Delta t^\pi_{k+1}}}\right)^2\right]}{\mathbb{E}\left[\hat{B}^4 - \hat{B}^2\right]} \\ &= \frac{2}{\mathbb{E}\left[\left(\gamma\hat{B}\right)^4 - 1\right]}\end{aligned}$$

et $X_{t_{k+1}} = F(X_{t_k}, W_{t_{k+1}} - W_{t_k})$, on approxime $X_{t_{k+1}}$ par $F(X_{t_k}, \Delta\hat{W}^\pi_{k+1})$, alors notre schéma 1 (déduit de la discrétisation dans [**CSTV06**]) devient

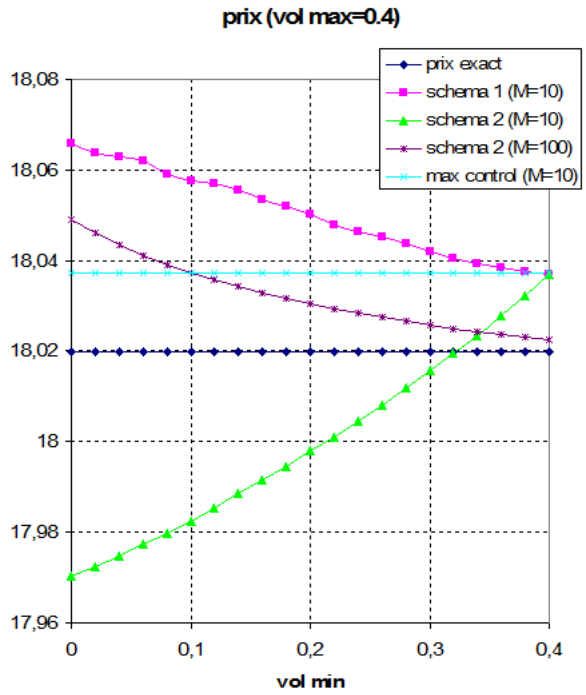
$$\begin{aligned}\hat{y}^\pi_k(x) &= \mathbb{E}\left[\hat{y}^\pi_{k+1}\left(F(x, \Delta\hat{W}^\pi_{k+1})\right)\right] + \Delta t^\pi_{k+1} \times f\left(t_k, x, \hat{y}^\pi_k(x), \hat{z}^\pi_k(x), \hat{\Gamma}^\pi_k(x)\right), \\ \hat{z}^\pi_k(t_k, x) &= \sigma(x)^{-1}\mathbb{E}\left[\hat{y}^\pi_{k+1}\left(F(x, \Delta\hat{W}^\pi_{k+1})\right)\frac{\Delta\hat{W}^\pi_{k+1}}{\Delta t^\pi_{k+1}}\right], \\ \hat{\Gamma}^\pi_k(x) &= \sigma(x)^{-1}\mathbb{E}\left[\hat{z}^\pi_{k+1}\left(F(x, \Delta\hat{W}^\pi_{k+1})\right)\frac{\Delta\hat{W}^\pi_{k+1}}{\Delta t^\pi_{k+1}}\right].\end{aligned}$$

On peut aussi considérer le schéma 2 (déduit dans [**FTW11**]) en remplaçant le terme $\hat{\Gamma}^\pi_k$ par

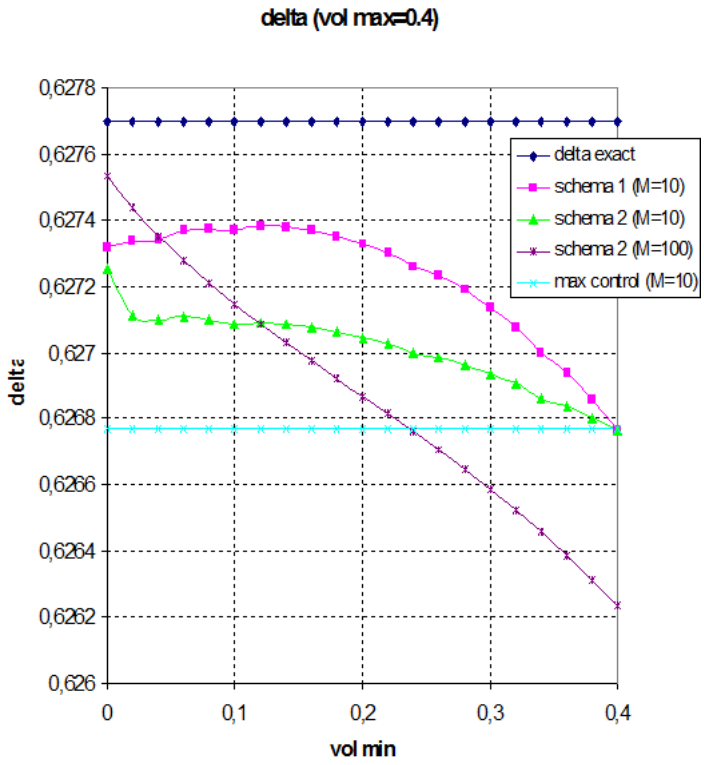
$$\hat{\Gamma}^\pi_k(x) = \sigma(x)^{-2}\mathbb{E}\left[\hat{y}^\pi_k\left(F(x, \Delta\hat{W}^\pi_{k+1})\right)\frac{\left|\Delta\hat{W}^\pi_{k+1}\right|^2 - \Delta t_k}{(\Delta t^\pi_{k+1})^2}\right]\rho.$$

4.3.3. Application numérique. On prend maintenant l'exemple du call : $h(x) = (x-K)^+$ avec les paramètres suivants $r = 0.05$, $\sigma_{max} = 0.4$, $T = 1$, $S_0 = 100$ et $K = 100$.

Prix du call Y_0 :



Delta Z_0 du call :



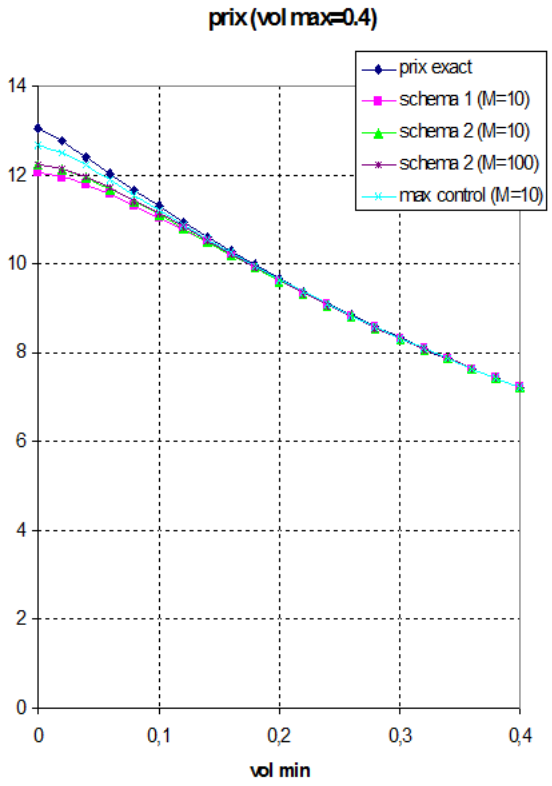
L'algorithme est très efficace. On a agrandi les graphiques pour les mettre à une échelle adéquate pour comparer les différentes méthodes.

Par convexité de payoff, $\sigma^* = \sigma_{max}$.

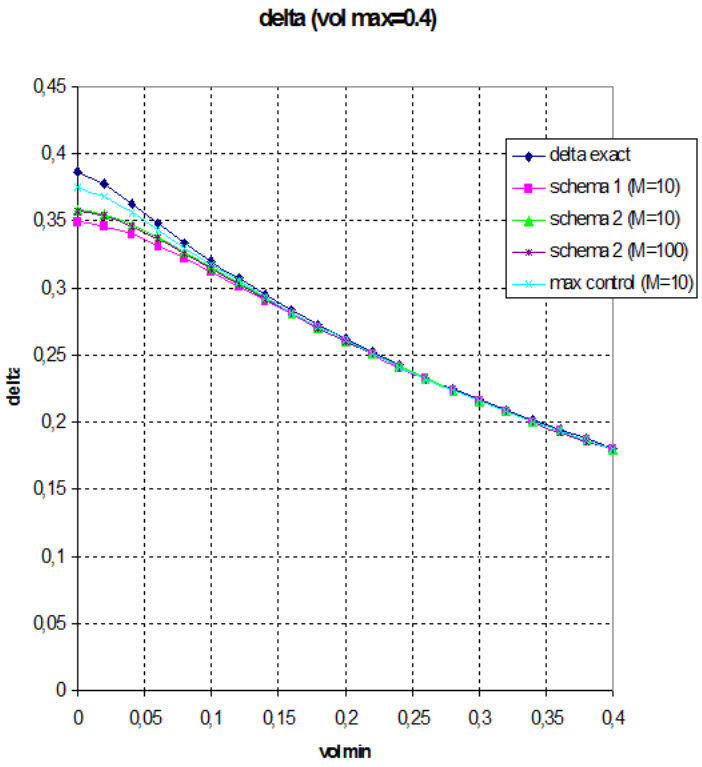
Exemple du bull spread : $g(x) = (x - K_{min})^+ - (x - K_{max})^+$. $r = 0.05$, $\sigma_{max} = 40\%$, $T = 1$, $S_0 = 100$, $K_{min} = 100$ et $K_{max} = 120$.

Comparaison des différents schémas :

Prix du call spread :

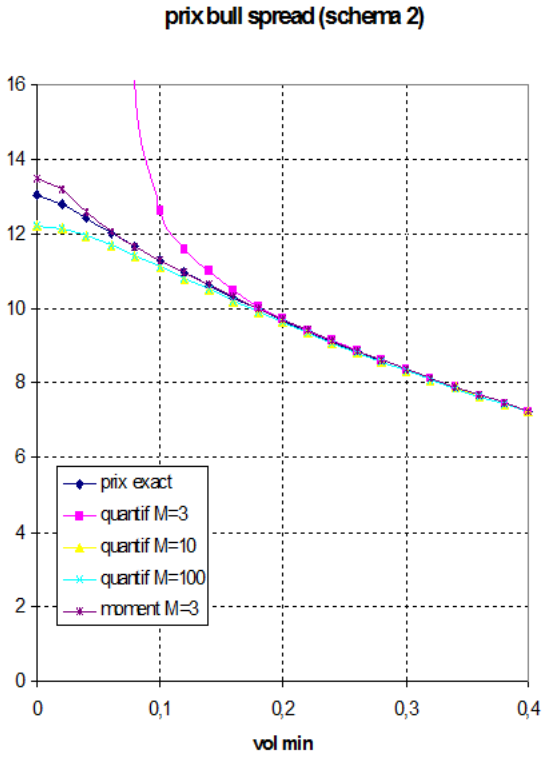


Delta du call spread :

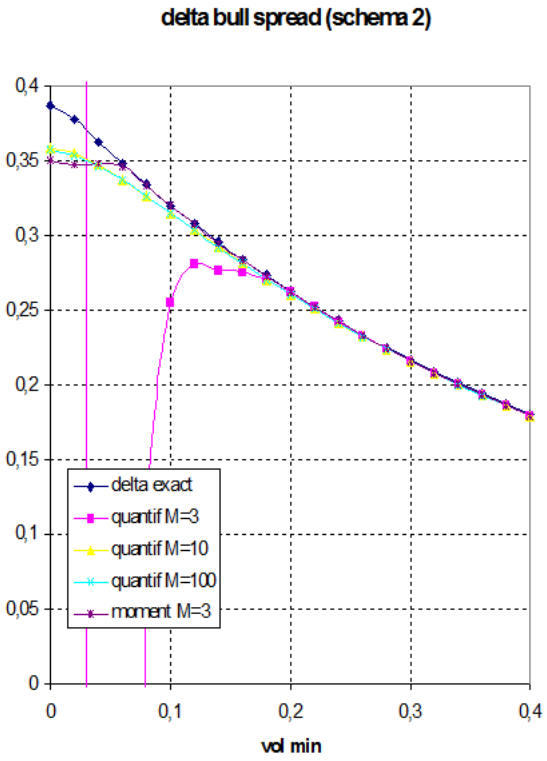


Comparaison selon la taille de la grille de quantification M (de l'incrément brownien ΔW) avec la méthode 2

Prix du call spread :



Delta du call spread :


delta bull spread (schema 2)
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
0
0,1
0,2
0,3
0,4
vol min
delta exact
quantif M=3
quantif M=10
quantif M=100
moment M=3

Troisième partie

A Discrete-time approximation for BSDEs with barriers

We study a numerical scheme for doubly reflected backward stochastic differential equations (RBSDE) with possible path-dependant values and with jumps. We propose an Euler scheme's type approximation. We prove the convergence of the scheme for BSDE when the number of time steps n goes to infinity. We give the rate of convergence for game options.

4.4. Introduction

In this paper, we are interested in the problem of discretization of a backward stochastic differential equations (BSDE) with two reflected barriers with path-dependent values and jumps :

$$\left\{\begin{array}{l} Y_t = l \vee g \wedge u(X_T) + \int_t^T f(\Theta_s)ds + \left(K_T^+ - K_t^+\right) - \left(K_T^- - K_t^-\right) \\ \quad - \int_t^T Z_s dsW_s - \int_t^T \int_E U_s(e)\overline{\mu}(de,ds), \quad \forall t \in [0,T],\ \mathbb{P}-a.s. \\ l(X_t) \leq Y_t \leq u(X) \quad \forall t \in [0,T],\ \mathbb{P}-a.s. \\ \text{and } \int_0^T (Y_s - l(X_s))\, dK_s^+ = \int_0^T (Y_s - u(X_s))\, dK_s^- = 0, \quad \mathbb{P}-a.s. \end{array}\right. \tag{4.4.1}$$

where $K^{\pm}$ are continuous non decreasing, $\Theta := (X, Y, Z, \Gamma)$ with $\Gamma := \int_E \rho(e)U(e)\lambda(de)$. The forward process $X_t = \Phi_t\left((S_s)_{0\leq s\leq t}\right)$ is a $\mathbb{R}^d$-valued pathdependent functionnal of a process S which is solution of a forward SDE :

$$S_t = S_0 + \int_0^t b(S_r)dr + \int_0^t \sigma(S_r)dW_r + \int_0^t \int_E \beta(S_{r-}, e)\overline{\mu}(de, dr).$$

Here, W is a d_W-dimensional Brownian motion and $\overline{\mu}$ an independent compensated Poisson measure $\overline{\mu}(de,dr) = \mu(de,dr) - \lambda(de)ds$.

Given $\pi = \{0 = t_0 < t_1 < .. < t_n = T\}$ a partition of the time interval $[0,T]$ with modulus $|\pi|$ and X^π the Euler scheme of X, we approximate the solution (4.4.1) by the backward scheme :

$$\left\{\begin{array}{l} Y_T^\pi = l\left(X_T^\pi\right) \vee g\left(X_T^\pi\right) \wedge u\left(X_T^\pi\right), \\ \overline{Z}_{t_i}^\pi = \mathbb{E}\left[Y_{t_{i+1}}^\pi \frac{W_{t_{i+1}} - W_{t_i}}{t_{i+1}-t_i} | \mathcal{F}_{t_i}\right], \\ \overline{\Gamma}_{t_i}^\pi = \mathbb{E}\left[Y_{t_{i+1}}^\pi \frac{\int_E \rho(e)\overline{\mu}(de,(t_i,t_{i+1}])}{t_{i+1}-t_i} | \mathcal{F}_{t_i}\right], \\ \tilde{Y}_{t_i}^\pi = \mathbb{E}\left[Y_{t_{i+1}}^\pi | \mathcal{F}_{t_i}\right] + (t_{i+1}-t_i) f\left(X_{t_i}^\pi, \tilde{Y}_{t_i}^\pi, \overline{Z}_{t_i}^\pi, \overline{\Gamma}_{t_i}^\pi\right), \\ Y_{t_i}^\pi = l\left(X_{t_i}^\pi\right) \vee \tilde{Y}_{t_i}^\pi \wedge u\left(X_{t_i}^\pi\right). \end{array}\right.$$

For BSDE without reflexion (in (4.4.1) one should take $l = -\infty$ and $u = +\infty$, thereby $K^{\pm} = 0$), such approximations have been studied by [**BT04**] and [**Zha04**], see also [**BE08**] and [**JLW05**] for BSDE with jump. The simply reflected BSDEs (in (4.4.1) one should take $u = +\infty$, thereby $K^- = 0$) are related to optimal stopping problem (American option in finance) see [**KKP$^+$97**] and see [**BP03a**], [**BC08**] for numerical approximation.

Existence and unicity for reflected and doubly reflected BSDEs are studied in [**CK96**] for the continuous case and in [**CM08**], [**HH06**] for the case with jumps. Some numerical schemes are developped in [**Cha09**], [**Otm08**] for a continuous non-pathdependent forward process $X = S$.

In our numerical scheme for R2BSDE, we are able to extend to a forward pathdependent process $X_t = (S_t, \Phi_t(S))$ where S is a jump diffusion process and Φ_t is a $\mathbb{L}^\infty$-Lipschitz functionnal (see section 2 for precise definition).

In the first section, we study the forward process X and its approximation X^π by the Euler Scheme X^π . We show new results for the strong convergence of this scheme in the jump framework for both the uniform norm and the Skorokhod metric. We prove some properties of the forward process : some regularity on time and on the flow.

In the second section, we are interested in the approximation of the backward process Y. We prove that

$$\mathbb{E}\left[\sup_{t\in[0,T]}|Y_t - Y_t^\pi|^2\right] \to_{|\pi|\to 0} 0$$

in a general case and with a rate of convergence of at least $|\pi|\ln\frac{1}{|\pi|}$ for the game option (with f independent of (Z,Γ), see [**Kif00**]). It allows us to show that

$$\mathbb{E}\left[\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\left\{\left|Z_s - \overline{Z}_{t_i}^\pi\right|^2 + \left|\Gamma_s - \overline{\Gamma}_{t_i}^\pi\right|^2\right\}\right] \to_{|\pi|\to 0} 0.$$

under Mokobodski conditions.

We work with finite time horizon $T > 0$, a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$. Let $W = (W_t)_{0\le t\le T}$ a d_W-dimensional standart Brownian motion and μ a Poisson measure with intensity $\nu(de,dt) = \lambda(de)dt$, for some finite measure λ in $E \subset \mathbb{R}^m$ for some $m \ge 1$. Suppose W and μ independent under $\mathbb{P}$.

NOTATION. We denote by :
$\diamond$ $\mathbb{E}_t$ the conditionnal expectation $\mathbb{E}[.|\mathcal{F}_t]$ $\forall t \in [0,T]$;
$\diamond$ $\overline{\mu} := \mu - \lambda$ is the compensated measure associated to μ ;
$\diamond$ $\mathbb{M}^{n,d}$ the set of all matrices $n \times d$ (and simply $\mathbb{R}^d = \mathbb{M}^{d,1}$ and $\mathbb{M}^d = \mathbb{M}^{d,d}$), $|x| := \left(\sum_{i,j}|x_{i,j}|^2\right)^{\frac{1}{2}}$ the Euclidian norm on $\mathbb{M}^{n,d}$.
$\diamond$ $\mathbb{D}\left([0,T],\mathbb{R}^d\right)$ (or $\mathbb{D}([0,T])$ to simplify) is the space of all càdlàg (right continuous with left limits) functions : $[0,T] \to \mathbb{R}^d$ with the topology J_1 metric (it is the Skorokhod space with the Skorokhod topology see in [**Bil99**])

$$d_{\mathbb{D}}(x,y) = \inf_{\lambda\in\Lambda}\left(\sup_{t\in[0,T]}|\lambda(t)-t| + \sup_{t\in[0,T]}|x(\lambda(t)) - y(t)|\right)$$

where Λ denotes the set of all strictly increasing continuous bijections from $[0,T]$ to itself.

We consider $(H,|.|)$ a Hilbert space (usually $H = \mathbb{R}^d$).
$\diamond$ $\mathbb{L}^2(H)$ is the space of square integrable H-valued ($\mathcal{F}_T$ -measurable) random variable V such that $||V||_{\mathbb{L}^2} = \mathbb{E}\left[|V|^2\right]^{\frac{1}{2}} < +\infty$.
$\diamond$ $\mathbb{H}^2(H)$ is the space of progressively measurable $\mathbb{R}^d$-valued process V such that $||V||_{\mathbb{H}^2} = \mathbb{E}\left[\int_0^T |V_t|^2\,dt\right]^{\frac{1}{2}} < +\infty$.
$\diamond$ $\mathbb{H}^2_\lambda(H)$ is the set of mappings $V : \Omega\times[0,T]\times E \to H$ which are $\mathcal{F}_t$-measurable and $||V||_{\mathbb{H}^2_\lambda} = \mathbb{E}\left[\int_0\int_E |V_t(e)|^2\,\lambda(e)dt\right]^{\frac{1}{2}} < +\infty$.
$\diamond$ $\mathcal{S}^2(H)$ is the set of H-valued adapted càdlàg process V such that $||V||_{\mathcal{S}^2} =$

$\mathbb{E}\left[\sup_{0\leq t\leq T}|V_t|^2\right]^{\frac{1}{2}} < +\infty$.

$\diamond$ For a partition $\pi = \{t_0 = 0 < t_1 < ... < t_n = T\}$, $\mathcal{S}^2_\pi(H)$ is the set of H-valued $\{\mathcal{F}_{t_i}\}_{t_i\in\pi}$-adapted process V such that $||V||_{\mathcal{S}^2_\pi} = \mathbb{E}\left[\sup_{t_i\in\pi}|V_{t_i}|^2\right]^{\frac{1}{2}} < +\infty$.

We denote $\mathbb{L}^2, \mathbb{H}^2, \mathcal{S}^2, \mathcal{S}^2_\pi$ for simplicity, when the context is clear.

Let $\pi = \{0 = t_0 < t_1 < .. < t_n = T\}$ be a partition of the time $[0,T]$. Throughout this paper, we shall use the notation

$$\begin{aligned}
|\pi| &:= \sup_{0\leq i<n}|t_{i+1} - t_i|, \\
\pi(t) &:= \sup\{t_i \in \pi | t_i \leq t\}, \quad \forall t \in [0,T], \\
\overline{\pi}(t) &:= \inf\{t_i \in \pi | t_i \geq t\}, \quad \forall t \in [0,T], \\
\Delta^\pi_{i+1} &:= t_{i+1} - t_i, \quad 0 \leq i < n, \\
\Delta^\pi_{i+1}W &:= W_{t_{i+1}} - W_{t_i}, \quad 0 \leq i < n.
\end{aligned}$$

4.5. The forward process

Consider $K > 0$, two K-Lipschitz continuous functions $b : \mathbb{R}^{d_S} \to \mathbb{R}^{d_S}$, $\sigma : \mathbb{R}^{d_S} \to \mathbb{M}^{d_S,d_W}$ and a measurable map $\beta : \mathbb{R}^{d_S} \times E \to \mathbb{R}^{d_S}$ such that :

$$\begin{aligned}
\sup_{e\in E}|\beta(0,e)| &\leq K, \\
\sup_{e\in E}|\beta(x,e) - \beta(\tilde{x},e)| &\leq K|x - \tilde{x}|, \quad \forall x, \tilde{x} \in \mathbb{R}^{d_S}.
\end{aligned}$$

Then, it is known that, for any initial condition $S_0 \in \mathbb{R}^{d_S}$, the forward stochastic differential equation

$$S_t = S_0 + \int_0^t b(S_r)dr + \int_0^t \sigma(S_r)dW_r + \int_0^t\int_E \beta(S_{r-},e)\overline{\mu}(de,dr)$$

exists and has a unique solution $\{S_t\}_{0\leq t\leq T}$ in $\mathcal{S}^2_{[0,T]}$, see **[FK85]**.

Given $\pi = \{0 = t_0 < t_1 < .. < t_n = T\}$ be a partition of the time $[0,T]$, we approximate S by its Euler scheme S^π defined by

$$S^\pi_0 = S_0,$$

$$S^\pi_{t_{i+1}} = S^\pi_{t_i} + b(S^\pi_{t_i})\Delta^\pi_{i+1} + \sigma(S^\pi_{t_i})\Delta^\pi_{i+1}W + \int_E \beta(S^\pi_{t_i},e)\overline{\mu}(de,(t_i,t_{i+1}]), \quad i = 0,..,n-1,$$

and we set

$$S^\pi_t = S^\pi_{\pi(t)} + b(S^\pi_{\pi(t)})(t - \pi(t)) + \sigma(S^\pi_{\pi(t)})(W_t - W_{\pi(t)}) + \int_E \beta(S^\pi_{\pi(t)},e)\overline{\mu}(de,(\pi(t),t]),$$

$$\overline{S}^\pi_t = S^\pi_{\pi(t)}.$$

In the following, $C > 0$ will denote a generic constant independent of i and n that may take different values from line to line. It is known that (see **[BE08]** for example) :

$$\begin{aligned}
\mathbb{E}\left[\sup_{s\leq u,v\leq(s+h)}|S_u - S_v|^2|\mathcal{F}_s\right] &\leq C\left(1 + |S_s|^2\right)h, \quad \forall s \in [0,T], \forall h \geq 0, \\
\sup_{0\leq i\leq n-1}\mathbb{E}\left[\sup_{t_i\leq t<t_{i+1}}\left|S^\pi_t - S^\pi_{t_i}\right|^2\right] &\leq C|\pi|.
\end{aligned}$$

LEMMA 4.5.1. *The strong error due to the Euler scheme is known : there exists a constant $C > 0$ such that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|S_t - S_t^\pi|^2\right] \leq C\,|\pi|\,.$$

In the continuous case (i.e. $\beta = 0$), we have

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|S_t - S_{\pi(t)}^\pi\right|^2\right] \leq C\,|\pi|\ln(\frac{1}{|\pi|}).$$

see [**Zha04**].

The last inequality is false for a diffusion with jumps. Indeed, if we consider a Poisson process N_t for example, we have

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|N_t - N_{\pi(t)}^\pi\right|^2\right] \geq \mathbb{P}\left(N_T \geq 1\right).$$

So, $\overline{N}^\pi \nrightarrow N$ in $\mathcal{S}_T^2$. We need to wiggle space and time a bit to obtain the convergence to the càdlàg processus. Therefore, we use the Skorokhod topology (whereas the traditional topology of uniform convergence only allows us to "wiggle space a bit") for the strong convergence of the Euler scheme $\overline{S}^\pi$.

PROPOSITION 4.5.2. *With the J_1 metric on $\mathbb{D}$, we obtain the same convergence rate as in the continuous case with the uniform distance :*

$$\mathbb{E}\left[\sup_{t_i\in\pi} d_{\mathbb{D}}\left(S^{t_i}, (S^\pi)^{t_i}\right)^2\right] \leq C\,|\pi|\,,$$
$$\mathbb{E}\left[\sup_{t_i\in\pi} d_{\mathbb{D}}\left(S^{t_i}, (\overline{S}^\pi)^{t_i}\right)^2\right] \leq C\,|\pi|\ln(\frac{1}{|\pi|}).,$$

where we denote

$$x_t^{t_i} = x_{t\wedge t_i}, \quad \forall x \in \mathbb{D}([0,T]), \forall t, t_i \in [0,T].$$

DEFINITION 4.5.3. A functionnal $\Phi : \mathbb{D}([0,T]) \to \mathbb{R}$ is called $\mathbb{L}^\infty$-Lipschitz, if there exists a constant K such that

$$|\Phi(x) - \Phi(x')| \leq K \sup_{0\leq t\leq T}|x(t) - x'(t)|\,, \quad \forall x, x' \in \mathbb{D}([0,T]).$$

Now we define the forward process X.
Let $K > 0$. Suppose that for each $0 \leq s \leq t \leq T$, there exits a fonction $\Phi_t : \mathbb{D}\left([0,T], \mathbb{R}^{d_S}\right) \to \mathbb{R}^{d_P}$ and $\varphi_{s,t} : \mathbb{R}^{d_P} \times \mathbb{D}\left([s,T], \mathbb{R}^{d_S}\right) \to \mathbb{R}^{d_P}$ such that for $0 \leq s \leq t \leq T$, for all $x \in \mathbb{D}\left([0,T], \mathbb{R}^d\right)$

$$\Phi_t(x) = \varphi_{s,t}\left(\Phi_s(x), \{x(u)\}_{s\leq u\leq t}\right).$$

We suppose that all $\varphi_{s,t}$ (and Φ_t also) is $\mathbb{L}^\infty$-Lipschitz with a uniformly constant K.

With the path-dependant process $P_t = \Phi_t(S_t)$, we define

$$X_t = (S_t, P_t) = (S_t, \Phi_t(S_t)) \in \mathbb{R}^d$$

with $d = d_S + d_P$. Since S is Markovian, one can easily see that X is Markovian.

Using the Euler Scheme of S, we approximate $X_t = (S_t, \Phi_t(S))$ by

$$X_t^\pi = \left(S_{\pi(t)}^\pi, P_t^\pi\right) = \left(S_{\pi(t)}^\pi, \Phi_t(S_\pi^\pi)\right).$$

.

Moreover, we suppose that :

$$|\Phi_t(x) - \Phi_t(\tilde{x})| \leq K d_{\mathbb{D}}(x^t, \tilde{x}^t) + \frac{K}{t}\int_0^t |x_s - \tilde{x}_s|\, ds, \quad \forall 0 \leq t \leq T. \tag{4.5.1}$$

REMARK 4.5.4. The functionnals $\Phi_t(x) = \sup_{0\leq s\leq t} x(s)$ (for the lookback option), $\Phi_t(x) = \int_0^t x(s)ds$ and $\Phi_t(x) = \frac{1}{t}\int_0^t x(s)ds$ (for asian option) are $\mathbb{L}^\infty$ Lipschitz uniformly in t and is satisfied (4.5.1).

LEMMA 4.5.5. *The strong error of the scheme is given by*

$$||X - X^\pi||_{\mathcal{S}_\pi^2} = \mathbb{E}\left[\sup_{t_i\in\pi} \left|X_{t_i} - X_{t_i}^\pi\right|^2\right] \leq C\,|\pi|\ln(\frac{1}{|\pi|}). \tag{4.5.2}$$

REMARK 4.5.6. It is possible to approximate X_{t_i} by $\tilde{X}_{t_i}^\pi = \left(S_{t_i}^\pi, \Phi_{t_i}(S^\pi)\right)$. This estimation is less convenient to use than $\Phi\left(\overline{S}^\pi\right) = \Phi\left(\left(S_{\pi(t)}^\pi\right)_{0\leq t\leq T}\right)$ but the error is smaller :

$$\mathbb{E}\left[\sup_{t_i\in\pi}\left|X_{t_i} - \tilde{X}_{t_i}^\pi\right|^2\right] \leq C\mathbb{E}\left[\sup_{t\in[0,T]}|S_{t_i} - S_t^\pi|^2\right] \leq C\pi$$

with only $\mathbb{L}^\infty$-Lipschitz property required by Φ. For example, we can simulate $\Phi(S^\pi)$ for asian and lookback option in continuous case :
$\diamond$ For lookback option $\tilde{P}_{t_k}^\pi = \Phi_{t_k}(S^\pi) = \sup_{0\leq t\leq t_k} S_t^\pi$, we have

$$\tilde{P}_{t_{k+1}}^\pi = \sup\left(\tilde{P}_{t_k}^\pi, S_{t_k}^\pi + \sigma(S_{t_k}^\pi)\sqrt{t_{k+1} - t_k}\epsilon_{k+1}\right)$$

where $\epsilon_{k+1} = \sup_{0\leq s\leq 1}\left(B_s + s\sqrt{t_{k+1}-t_k}\sigma^{-1}(S_{t_k}^\pi)b(S_{t_k}^\pi)\right)$ (where $B_s = \frac{W_{t_k+s(t_{k+1}-t_k)} - W_{t_k}}{\sqrt{t_{k+1}-t_k}}$ is a standard Brownian motion by scaling) can be simulate.

$\diamond$ For asian option $\tilde{P}_{t_k}^\pi = \Phi_{t_k}(X^\pi) = \frac{1}{t_k}\int_0^{t_k} S_t^\pi dt$, we have

$$\begin{aligned}\tilde{P}_{t_k}^\pi &= \frac{1}{t_{k+1}}\int_0^{t_k} S_t^\pi dt + \frac{1}{t_{k+1}}\int_{t_k}^{t_{k+1}} \left\{S_{t_k}^\pi + b(S_{t_k}^\pi)(t - t_k) + \sigma(S_{t_k}^\pi)(W_t - W_{t_k})\right\} dt \\ &= \frac{t_k}{t_{k+1}}P_{t_k}^\pi + \frac{t_{k+1} - t_k}{t_{k+1}}\left\{\frac{S_{t_k}^\pi + S_{t_{k+1}}^\pi}{2} + \epsilon_{k+1}\sqrt{t_{k+1} - t_k}\right\}\end{aligned}$$

where $\epsilon_{k+1} = \frac{1}{(t_{k+1}-t_k)^{\frac{3}{2}}}\int_{t_k}^{t_{k+1}}\left(W_t - \frac{W_{t_k}+W_{t_{k+1}}}{2}\right)ds \sim \mathcal{N}(0; \frac{1}{3})$ is independent of $\left(S_{t_k}^\pi, S_{t_{k+1}}^\pi\right)$. This scheme is used in **[LT01]** to simulate asian option $\int_0^T S_s ds$.

For $0 \leq s \leq t \leq T$ and for a Markov process M, we denote $M_t^{(s,m)}$ for M_t conditioning to $M_s = m$.
The following lemma gives the Lipschitz property of the flow of X.

LEMMA 4.5.7. *There exists a constant $L \geq 0$ that $\forall t \in [0,T]$, $\forall 0 \leq k \leq n$, $\forall x = (s,p), x' = (s',p') \in \mathbb{R}^{d_S} \times \mathbb{R}^{d_P}$*

$$\mathbb{E}\left[\sup_{t\leq u\leq T}\left|S_u^{(t,s)} - S_u^{(t,s')}\right|^2\right] + \mathbb{E}\left[\sup_{t_k\leq s\leq T}\left|S_s^{\pi,(t_k,s)} - S_s^{\pi,(t_k,s')}\right|^2\right] \leq L\left|s-s'\right|^2,$$

$$\mathbb{E}\left[\sup_{t\leq u\leq T}\left|X_u^{(t,x)} - X_u^{(t,x')}\right|^2\right] + \mathbb{E}\left[\sup_{t_k\leq u\leq T}\left|X_u^{\pi,(t_k,x)} - X_u^{\pi,(t_k,x')}\right|^2\right] \leq L\left|x-x'\right|^2.$$

We want the process X to be càdlàg and therefore we need some time regularity for Φ_t. Then we suppose that $t \to \Phi_t(S)$ is right continuous with left hand limits.

We also suppose that $R(\epsilon) \to 0$ as $\epsilon \to 0$ with

$$R(\epsilon) = \mathbb{E}\left[\sup_{0\leq t\leq u\leq (t+\epsilon)\wedge T}\left|\Phi_u(S^t) - \Phi_t(S^t)\right|^2\right].$$

Then, this implies $\mathbb{E}\left[\sup_{0\leq t\leq T}\mathbb{E}\left[\sup_{t\leq u\leq (t+\epsilon)\wedge T}\left|X_u - X_t\right|^2\right]\right] \to 0$ as $\epsilon \to 0$. Indeed

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\leq t\leq T}\mathbb{E}_t\left[\sup_{t\leq u\leq (t+\epsilon)\wedge T}\left|X_u - X_t\right|^2\right]\right] &\leq C\mathbb{E}\left[\sup_{0\leq t\leq T}\mathbb{E}_t\left[\sup_{t\leq u\leq (t+\epsilon)\wedge T}\left|S_u - S_t\right|^2\right]\right] \\
&\quad + C\mathbb{E}\left[\sup_{0\leq t\leq u\leq (t+\epsilon)\wedge T}\mathbb{E}_t\left[\sup_{t\leq u\leq (t+\epsilon)\wedge T}\left|\Phi_u(S^t) - \Phi_t(S^t)\right|^2\right]\right] \\
&\leq C\mathbb{E}\left[\sup_{0\leq t\leq T}\mathbb{E}_t\left[\sup_{t\leq u\leq (t+\epsilon)\wedge T}\left|S_u - S_t\right|^2\right]\right] + CR(\epsilon) \\
&\leq C\left\{\epsilon + R(\epsilon)\right\}.
\end{aligned}$$

ASSUMPTION 4.5.8. *We denotes (At) the assumption : there exists a constant $C \geq 0$ such that $\forall \epsilon \geq 0$, $R(\epsilon) \leq C\epsilon \ln \frac{1}{\epsilon}$. Then the assumption (At) gives a time control of X :*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\mathbb{E}_t\left[\sup_{t\leq u\leq (t+\epsilon)\wedge T}\left|X_u - X_t\right|^2\right]\right] \leq C\epsilon \ln \frac{1}{\epsilon}. \tag{4.5.3}$$

REMARK 4.5.9. $\diamond$ If there exists C such that

$$\left|\Phi_{t+h}(x^t) - \Phi_t(x^t)\right| \leq C(1 + \sup_{0\leq u\leq t}|x(u)|)h \quad \forall h \geq 0, \forall t, \forall x \in \mathbb{D}([0,T], \mathbb{R}^{d_S})$$

then (Ht) is verified.

$\diamond$ For the lookback option, we have $\Phi_{t+h}(x^t) - \Phi_t(x^t) = \sup_{0\leq u\leq t+\epsilon} x(u\wedge t) - \sup_{0\leq u\leq t} x(u) = 0$ $\forall x \in \mathbb{D}([0,T], \mathbb{R}^{d_S})$ and (At) is verified.

$\diamond$ For the asian option $\Phi_t(x)) = \int_0^t x(u)du$, we have $\Phi_{t+h}(x^t) - \Phi_t(x^t) = hx(t)$ $\forall x \in \mathbb{D}([0,T], \mathbb{R}^{d_S})$ and (At) is satisfied.

$\diamond$ For the asian option $\Phi_t(x)) = \frac{1}{t}\int_0^t x(u)du$, assumption (At) is satisfied.

Indeed, we have for $h \geq 0$

$$\begin{aligned}
\Phi_{t+h}(S^t) - \Phi_t(S^t) &= \frac{1}{t+h}\int_0^{t+h} S_{s\wedge t}ds - \frac{1}{t}\int_0^t S_s ds \\
&= \frac{1}{t}\int_0^t (S_{\lambda_{t,h}(s)} - S_s)ds.
\end{aligned}$$

where $\lambda_{t,h}(s) = \left(\frac{t+h}{t}s\right) \wedge t \ \forall 0 \leq s,t \leq T$ (with $|\lambda_{t,h}(s) - s \wedge t| \leq h$). With $C_\epsilon = \epsilon \ln \frac{1}{\epsilon}$, we get

$$\begin{aligned}
R(\epsilon) &= \mathbb{E}\left[\sup_{0\leq t\leq t+h\leq (t+\epsilon)\wedge T}\left|\frac{1}{t}\int_0^t (S_{\lambda_{t,h}(s)} - S_s)ds\right|^2\right] \\
&\leq \mathbb{E}\left[\sup_{0\leq t\leq t+h\leq (t+\epsilon)\wedge T}\frac{1}{t}\int_0^t \left|S_{\lambda_{t,h}(s)} - S_s\right|^2 ds\right] \\
&\leq \mathbb{E}\left[\sup_{0\leq t\leq t+h\leq (t+\epsilon)\wedge T}\frac{1}{t}\int_0^{t\wedge C_\epsilon} \left|S_{\lambda_{t,h}(s)} - S_s\right|^2 ds\right] + \mathbb{E}\left[\sup_{C_\epsilon\leq t\leq t+h\leq (t+\epsilon)\wedge T}\frac{1}{t}\int_{C_\epsilon}^t \left|S_{\lambda_{t,h}(s)} - S_s\right|^2 ds\right] \\
&\leq C\mathbb{E}\left[\sup_{0\leq s\leq C_\epsilon}|S_s - S_0|^2\right] + \int_{C_\epsilon}^T \mathbb{E}\left[\sup_{t\in[0,T],h\in[0,\epsilon]}\left|S_{\lambda_{t,h}(s)} - S_s\right|^2\right]\frac{ds}{s} \\
&\leq CC_\epsilon + C\epsilon\left\{\ln T - \ln C_\epsilon\right\} \\
&\leq C\left|\pi\right|\ln\frac{1}{|\pi|}.
\end{aligned}$$

4.6. Backward Process

Now, we shall use the following standing assumptions : the determinist functions $f : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{dz} \times \mathbb{R}^m \to \mathbb{R}$, $g : \mathbb{R}^d \to \mathbb{R}$, $l : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ and $u : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ are Lipschitz continuous with Lipschitz constant K.

REMARK 4.6.1. This condition permits us to treat the situations $l = -\infty$ or $u = +\infty$, corresponding RBSDE with a single lower barrier $l(X_t)$ and upper barrier $u(X_t)$ or a classical BSDE (without barrier).

We consider now the R2BSDE :

$$\left\{\begin{array}{l}
Y \in \mathcal{S}^2,\ Z \in \mathbb{H}^2,\ U \in \mathbb{H}^2,\ K^{\pm} \in \mathcal{S}^2, \\
Y_t = \tilde{g}(X_T) + \int_t^T f(\Theta_s)ds + \left(K_T^+ - K_t^+\right) - \left(K_T^- - K_t^-\right), \\
\quad - \int_t^T Z_s dsW_s - \int_t^T \int_E U_s(e)\overline{\mu}(de,dr), \quad \forall t \in [0,T],\ \mathbb{P}-a.s., \\
l(X_t) \leq Y_t \leq u(X_t) \quad \forall t \in [0,T],\ \mathbb{P}-a.s. \\
\text{and } \int_0^T (Y_s - l(X_s))\, dK_s^+ = \int_0^T (Y_s - u(X_s))\, dK_s^- = 0, \quad \mathbb{P}-a.s.,
\end{array}\right.$$

where $K^{\pm}$ are continuous non decreasing, $\Theta := (X, Y, Z, \Gamma)$ with $\Gamma := \int_E \rho(e)U(e)\lambda(de)$ and $\tilde{g} = l \vee g \wedge u$.

With these hypothesis, we have unicity of the solution of the R2BSDE (4.4.1). In order to ensure existence and unicity of the solution of this equation (4.4.1) (see **[CM08]** and **[HH06]**), we assume that Mokobodski condition is fullfilled : there exist two supermartingales $J^+, J^- \in \mathcal{S}^2_{[0,T]}$ such that $l(X_t) \leq J_t^+ - J_t^- \leq u(X_t)$ over $t \in [0,T]$.

ASSUMPTION 4.6.2. *We denote (Am) the assumption : $\exists L > 0, \forall \pi = \{0 = t_0 < ... < t_n = T\}$ such that there exist two supermartingales $J^{\pi,+}, J^{\pi,-} \in \mathcal{S}^2_{[0,T]}$ such that $l(X_t^\pi) \leq J_t^{\pi,+} - J_t^{\pi,-} \leq u(X_t^\pi)$ over $t \in [0,T]$ and $\mathbb{E}\left[|J^{\pi,\pm}|^2\right] \leq L$.*

We then approximate (Y, Z, Γ) by $\left(Y^\pi, \overline{Z}^\pi, \overline{\Gamma}^\pi\right)$ defined by the backward implicit scheme

$$\begin{cases} Y_T^\pi = \tilde{g}\left(X_T^\pi\right), \\ \overline{Z}_{t_i}^\pi = \mathbb{E}\left[Y_{t_{i+1}}^\pi \frac{\Delta^\pi W_{i+1}}{\Delta_{i+1}^\pi} | \mathcal{F}_{t_i}\right], \\ \overline{\Gamma}_{t_i}^\pi = \mathbb{E}\left[Y_{t_{i+1}}^\pi \frac{\int_E \rho(e)\overline{\mu}(de,(t_i,t_{i+1}])}{\Delta_{i+1}^\pi} | \mathcal{F}_{t_i}\right], \\ \tilde{Y}_{t_i}^\pi = \mathbb{E}\left[Y_{t_{i+1}}^\pi | \mathcal{F}_{t_i}\right] + \Delta_{i+1}^\pi f\left(\Theta_{t_i}^\pi\right), \\ Y_{t_i}^\pi = l\left(X_{t_i}^\pi\right) \vee \tilde{Y}_{t_i}^\pi \wedge u\left(X_{t_i}^\pi\right), \quad \forall 0 \leq i < n, \end{cases}$$

where $\Theta_t^\pi = \left(X_{\pi(t)}^\pi, \tilde{Y}_{\pi(t)}^\pi, \overline{Z}_{\pi(t)}^\pi, \overline{\Gamma}_{\pi(t)}^\pi\right)$ $\forall t \in [0, T]$.

Since $\left(X_{t_i}^\pi\right)_{0 \leq i \leq n}$ is a Markov chain, there exist determinist functions $y_{t_i}^\pi : \mathbb{R}^d \to \mathbb{R}$, $\tilde{y}_{t_i}^\pi; \mathbb{R}^d \to \mathbb{R}$, $z_{t_i}^\pi : \mathbb{R}^d \to \mathbb{R}$ and $\gamma_{t_i}^\pi : \mathbb{R}^d \to \mathbb{R}$ $\forall 0 \leq i \leq n$ such that $\left(y_{t_i}^\pi(X_{t_i}^\pi), \tilde{y}_{t_i}^\pi(X_{t_i}^\pi), z_{t_i}^\pi(X_{t_i}^\pi), \gamma_{t_i}^\pi(X_{t_i}^\pi)\right) = \left(Y_{t_i}^\pi, \tilde{Y}_{t_i}^\pi, \overline{Z}_{t_i}^\pi, \overline{\Gamma}_{t_i}^\pi\right)$, then $\left(y_{t_i}^\pi, \tilde{y}_{t_i}^\pi, z_{t_i}^\pi, \gamma_{t_i}^\pi\right)$ is defined by the backward implicit scheme

$$\begin{cases} y_T^\pi(x) = \tilde{g}\left(x\right), \\ z_{t_i}^\pi(x) = \mathbb{E}\left[y_{t_{i+1}}^\pi(X_{t_{i+1}}^\pi) \frac{\Delta^\pi W_{i+1}}{\Delta_{i+1}^\pi} | X_{t_i}^\pi = x\right], \\ \gamma_{t_i}^\pi(x) = \mathbb{E}\left[y_{t_{i+1}}^\pi(X_{t_{i+1}}^\pi) \frac{\int_E \rho(e)\overline{\mu}(de,(t_i,t_{i+1}])}{\Delta_{i+1}^\pi} | X_{t_i}^\pi = x\right] \\ \tilde{y}_{t_i}^\pi(x) = \mathbb{E}\left[y_{t_{i+1}}^\pi(X_{t_{i+1}}^\pi) | X_{t_i}^\pi = x\right] + \Delta_{i+1}^\pi f\left(\theta_{t_i}^\pi(x)\right), \\ y_{t_i}^\pi(x) = l\left(x\right) \vee \tilde{y}_{t_i}^\pi(x) \wedge u\left(x\right) \quad \forall x \in \mathbb{R}^d, \quad \forall 0 \leq i < n, \end{cases}$$

where $\theta_{t_i}^\pi = \left(I_d, \tilde{y}_{t_i}^\pi, z_{t_i}^\pi, \gamma_{t_i}^\pi\right)$ $\forall t_i \in \pi$.

Let us introduce the continuous time scheme $\left(Y_t^\pi, \tilde{Y}_t, Z_t^\pi, \Gamma_t\right)_{0 \leq t \leq T}$ associated to $\left(Y_{t_i}^\pi, \overline{Z}_{t_i}^\pi, \overline{\Gamma}_{t_i}^\pi\right)_{t_i \in \pi}$. By martingale representation theorem, there exists $(Z^\pi, U^\pi) \in \mathbb{H}^2 \times \mathbb{L}_\lambda^2$ such that

$$Y_{t_{i+1}}^\pi = \mathbb{E}_{t_i}\left[Y_{t_{i+1}}^\pi\right] + \int_{t_i}^{t_{i+1}} Z_t^\pi dW_t + \int_{t_i}^{t_{i+1}} \int_E U_t^\pi(e)\overline{\mu}(de, dt),$$

and we define $\Gamma_t^\pi := \int_E \rho(e) U_s^\pi(e)\lambda(de)$. We observe that

$$\overline{Z}_{t_i}^\pi = \mathbb{E}_{t_i}\left[\frac{1}{\Delta_{i+1}^\pi}\int_{t_i}^{t_{i+1}} Z_t^\pi dt\right],$$
$$\overline{\Gamma}_{t_i}^\pi = \mathbb{E}_{t_i}\left[\frac{1}{\Delta_{i+1}^\pi}\int_{t_i}^{t_{i+1}} \Gamma_t^\pi dt\right].$$

We define also $\left(\overline{Z}_{t_i}, \overline{\Gamma}_{t_i}\right)$ by

$$\overline{Z}_{t_i} = \mathbb{E}_{t_i}\left[\frac{1}{\Delta_{i+1}^\pi}\int_{t_i}^{t_{i+1}} Z_t dt\right],$$
$$\overline{\Gamma}_{t_i} = \mathbb{E}_{t_i}\left[\frac{1}{\Delta_{i+1}^\pi}\int_{t_i}^{t_{i+1}} \Gamma_t dt\right].$$

Then $\overline{Z}_{t_i}$ ($\overline{\Gamma}_{t_i}$, $\overline{Z}_{t_i}^\pi$ and $\overline{\Gamma}_{t_i}^\pi$ respectively) is the best approximation of Z (Γ, Z^π and Γ^π respectively) in $\mathbb{H}_{[t_i,t_{i+1}]}^2$ by $\mathcal{F}_{t_i}$-measurable random variable (which is constant

on $[t_i, t_{i+1}))$:

$$\mathbb{E}_{t_i}\left[\int_{t_i}^{t_{i+1}} \left|V_t - \overline{V}_{t_i}\right|^2 dt\right] = \inf_{V_i \in \mathbb{L}^2(\mathcal{F}_{t_i})} \mathbb{E}_{t_i}\left[\int_{t_i}^{t_{i+1}} |V_t - V_i|^2 dt\right]$$

for $\left(V_t, \overline{V}_{t_i}\right) = (Z_t, Z_{t_i})$, $\left(\Gamma_t, \overline{\Gamma}_{t_i}\right)$, $\left(Z_t^\pi, \overline{Z}_{t_i}^\pi\right)$ or $\left(\Gamma_t, \overline{\Gamma}_{t_i}^\pi\right)$. We prolonge $\overline{V}$ on $[0,T)$ by $\overline{V}_t = \overline{V}_{\pi(t)}$.

We can define the càdlàg process $\tilde{Y}^\pi$ on $[t_i, t_{i+1})$ and the làdcàg process Y^π on $(t_i, t_{i+1}]$ by

$$\tilde{Y}_t^\pi = Y_t^\pi = Y_{t_{i+1}}^\pi + (t_{i+1} - t) f(\Theta_{t_i}^\pi) - \int_s^{t_{i+1}} Z_t^\pi dW_s - \int_t^{t_{i+1}} U_s(e)\overline{\mu}(de, ds) \quad \forall t \in]t_i, t_{i+1}[$$

and the increasing process $K^{\pi,\pm}$ by

$$K_t^{\pi,\pm} = \sum_{t_i \le t, t_i \in \pi} \Delta K_{t_{i+1}}^{\pi,\pm} \text{ and } K^\pi = K^{\pi,+} - K^{\pi,-},$$

$$\Delta K_{t_i}^{\pi,+} = \left(l(X_{t_i}^\pi) - \tilde{Y}_{t_i}^\pi\right)^+ \text{ and } \Delta K_{t_i}^{\pi,-} = \left(\tilde{Y}_{t_i}^\pi - u(X_{t_i}^\pi)\right)^+.$$

Then, we have the following representation

$$\tilde{Y}_t^\pi = Y_T^\pi + \int_t^T f(\Theta_s^\pi) ds + K_T^\pi - K_t^\pi - \int_t^T Z_s^\pi dW_s - \int_t^T U_s(e)\overline{\mu}(de, ds), \quad \forall t \in [0,T].$$

Our main result is the following theorem.

THEOREM 4.6.3. *We have the convergence of the approximation :*

$$||Y - Y^\pi||_{\mathcal{S}^2} \to 0$$

as $|\pi| \to 0$.
If assumption (Am) *is satisfied, then*

$$||Y - Y^\pi||_{\mathcal{S}^2} + \left|\left|Z - \overline{Z}^\pi\right|\right|_{\mathbb{H}^2} + \left|\left|\Gamma - \overline{\Gamma}^\pi\right|\right|_{\mathbb{H}^2} \to_{|\pi| \to 0} 0.$$

Moreover, if assumption (At) *is satisfied and if* $f(x, y, z, \gamma)$ *is independent of* (z, γ)*, then the rate of convergence is given by :*

$$||Y - Y^\pi||_{\mathcal{S}^2} \le C\sqrt{|\pi| \ln \frac{1}{|\pi|}}.$$

.

To prove this theorem, we need some properties of Y^π given by the following lemma.

LEMMA 4.6.4. *We suppose* $K|\pi| < 1$. $\left(y_{t_i}^\pi, \tilde{y}_{t_i}^\pi, z_{t_i}^\pi, \gamma_{t_i}^\pi\right)$ *are Lipschitz continuous. There exists a constant* $C \ge 0$ *independent of* π *such that :*

$$\sup\left\{\left|y_{t_i}^\pi(x)\right|, \left|\tilde{y}_{t_i}^\pi(x)\right|, \left|z_{t_i}^\pi(x)\right|, \left|\gamma_{t_i}^\pi(x)\right|\right\} \le C\left(1 + |x|\right), \quad \forall i, \forall x \in \mathbb{R}^d,$$

$$\mathbb{E}\left[\sup_{s \in [0,T], e \in E} \left(|Y_s^\pi|^2 + |Z_s^\pi|^2 + |\Gamma_s^\pi|^2 + |U_s^\pi(e)|^2\right)\right] \le C, \quad \forall t \in [0,T].$$

Assume (Am), then $||K^{\pi,\pm}||_{\mathbb{L}^2} \le C$.

DÉMONSTRATION. 1/ First, we show that $\left|\tilde{y}_{t_i}^{\pi}(x)\right| \leq C\left(1+|x|\right)$. For $t_i = T$, the Lipschitz property of (g,l,u) implies that $\tilde{g} = l \vee g \wedge u$ is Lipschitz continuous with the Lipschitz constant K and gives

$$\begin{aligned}\left|\tilde{y}_{t_i}^{\pi}(x)\right| &\leq \left|\tilde{g}\left(x\right)\right| \\ &\leq C\left(1+|x|\right).\end{aligned}$$

Let $t_i \leq t_k < T$. From the Lipschitz property of f, there exists an $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{d_Z} \times \mathbb{R}^m$-valued $\sigma(X_{t_k}^{\pi,(t_i,x)})$-measurable random variable $(\lambda_{k,x}, \lambda_{k,y}, \lambda_{k,z}, \lambda_{k,\gamma})$, essentially bounded by K such that :

$$\begin{aligned}f\left(\theta_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)})\right) = f(0) &+ \lambda_{k,x}.X_{t_k}^{\pi,(t_i,x)} + \lambda_{k,y}\tilde{y}_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)}) \\ &+ \lambda_{k,z}.z_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)}) + \lambda_{k,\gamma}.\gamma_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)}).\end{aligned}$$

The definition of $\tilde{y}_{t_k}^{\pi}, z_{t_k}^{\pi}$ gives

$$\begin{aligned}\tilde{y}_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)}) = \mathbb{E}_{t_k}&\left[y_{t_{k+1}}^{\pi}(X_{t_{k+1}}^{\pi,(t_i,x)})\left\{1+\lambda_{k,z}.\Delta^{\pi}W_{k+1} + \lambda_{k,\gamma}.\int_E \rho(e)\overline{\mu}(de,(t_k,t_{k+1}])\right\}\right] \\ &+ \Delta_{k+1}^{\pi}\left\{f(0) + \lambda_{k,x}.X_{t_k}^{\pi,(t_i,x)}\right\} + \lambda_{k,y}\Delta_{k+1}^{\pi}\tilde{y}_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)}).\end{aligned}$$

From Cauchy-Schwarz inequality and the inequality $|x| \leq 1 + |x|^2$, we have

$$\begin{aligned}&\left|\tilde{y}_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)})\right|\left(1-K\left|\pi\right|\right) \leq \Delta_{k+1}^{\pi}\left\{f(0) + K\left|X_{t_k}^{\pi,(t_i,x)}\right|\right\} \\ &+ \mathbb{E}_{t_k}\left[\left|y_{t_{k+1}}^{\pi}(X_{t_{k+1}}^{\pi,(t_i,x)})\right|^2\right]^{\frac{1}{2}}\mathbb{E}_{t_k}\left[\left|1+\lambda_{k,z}.\Delta^{\pi}W_{k+1} + \lambda_{k,\gamma}.\int_E \rho(e)\overline{\mu}(de,(t_k,t_{k+1}])\right|^2\right]^{\frac{1}{2}} \\ &\leq \mathbb{E}_{t_k}\left[\left|y_{t_{k+1}}^{\pi}(X_{t_{k+1}}^{\pi,(t_i,x)})\right|^2\right]^{\frac{1}{2}}\left(1+2K\left|\pi\right|\right) + \Delta_{k+1}^{\pi}\left\{f(0) + K\left|X_{t_k}^{\pi,(t_i,x)}\right|\right\}.\end{aligned}$$

Since $\lambda_{k,z}$ and $\lambda_{k,\gamma}$ are $\sigma(X_{t_k}^{\pi,(t_i,x)})$-measurable and bounded by K,

$$\mathbb{E}\left[\left|1+\lambda_{k,z}.\Delta^{\pi}W_{k+1} + \lambda_{k,\gamma}.\int_E \rho(e)\overline{\mu}(de,(t_k,t_{k+1}])\right|^2 |X_{t_k}^{\pi,(t_i,x)}\right] \leq 1 + 2K\Delta_{k+1}^{\pi},$$

and then

$$\left|\tilde{y}_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)})\right| \leq \mathbb{E}\left[\left|y_{t_{k+1}}^{\pi}(X_{t_{k+1}}^{\pi,(t_i,x)})\right|^2 |X_{t_k}^{\pi,(t_i,x)}\right]^{\frac{1}{2}}\left(1+C\Delta_{k+1}^{\pi}\right) + \Delta_{i+1}^{\pi}C\left\{1+\left|X_{t_k}^{\pi,(t_i,x)}\right|\right\}.$$

Consider the stopping time

$$\tau = \inf_{t_k \in \pi}\left\{t_k \geq t_i | \tilde{y}_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)}) \leq l(X_{t_k}^{\pi,(t_i,x)} \text{ or } \tilde{y}_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)}) \geq u(X_{t_k}^{\pi,(t_i,x)})\right\} \wedge T.$$

Then, we have $\tilde{y}_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)}) = y_{t_k}^{\pi}(X_{t_k}^{\pi,(t_i,x)})$ for $t_i \leq t_k < \tau$ and $\tilde{y}_{\tau}^{\pi}(X_{\tau}^{\pi,(t_i,x)}) = l(X_{\tau}^{\pi,(t_i,x)}), u(X_{\tau}^{\pi,(t_i,x)})$ or $g(X_{\tau}^{\pi,(t_i,x)})$, therefore

$$\left|\tilde{y}_{\tau}^{\pi}(X_{\tau}^{\pi,(t_i,x)})\right| \leq C\left(1+\left|X_{\tau}^{\pi,(t_i,x)}\right|\right).$$

Summing up from t_i to τ with $\tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) = y^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k})$, this provides using Jensen inequality :

$$\begin{aligned}\left|\tilde{y}^\pi_{t_i}(x)\right| &\le \mathbb{E}\left[\left|y^\pi_\tau(X^{\pi,(t_i,x)}_\tau)\right|^2\right]^{\frac{1}{2}} \prod_{k=i}^{n-1}\left(1+C\Delta^\pi_{k+1}\right)\\ &+ \prod_{k=i}^{n-1}\left(1+C\Delta^\pi_{k+1}\right)\mathbb{E}\left[1+\sup_{i\le k\le n}\left|X^{\pi,(t_i,x)}_{t_k})\right|^2\right]\\ &\le C\left\{1+\mathbb{E}\left[1+\sup_{i\le k\le n}\left|X^{\pi,(t_i,x)}_{t_k})\right|^2\right]\right\}\\ &\le C\left\{1+|x|\right\},\end{aligned}$$

and

$$\begin{aligned}\left|y^\pi_{t_i}(x)\right| &= \left|l(x)\vee\tilde{y}^\pi_{t_i}(x)\wedge u(x)\right|\\ &\le C\left\{1+|x|\right\}.\end{aligned}$$

2/ We show now the Lipschitz property of $y^\pi_{t_i}$ and $\tilde{y}^\pi_{t_i}$.

The Lipschitz property of $\tilde{g}$ gives the result for $t_i = T$.

Let $t_i \le t_k < T$ and $x, x' \in \mathbb{R}^d$. From the Lipschitz property of f, there exists $(\lambda_x, \lambda_y, \lambda_z, \lambda_\gamma) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{d_Z} \times \mathbb{R}^m$ $\mathcal{F}_{t_k}$-measurable, essentially bounded by K such that :

$$\begin{aligned}&f\left(\theta^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k})\right) - f\left(\theta^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k})\right) = \lambda_y.\{\tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}))\}\\ &+\lambda_z.\{z^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - z^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k})\} + \lambda_\gamma.\{\gamma^\pi_{t_k}(x) - \gamma^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k})\} + \lambda_x.\{X^{\pi,(t_i,x)}_{t_k} - X^{\pi,(t_i,x')}_{t_k}\}.\end{aligned}$$

The definition of $\tilde{y}^\pi_{t_k}, z^\pi_{t_k}$, gives

$$\begin{aligned}\tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) &= \Delta^\pi_{k+1}\lambda_x.(X^{\pi,(t_i,x)}_{t_k} - X^{\pi,(t_i,x')}_{t_k})\\ &+ \mathbb{E}_{t_k}\left[\left(y^\pi_{t_{k+1}}(X^{\pi,(t_k,x)}_{t_{k+1}}) - y^\pi_{t_{k+1}}(X^{\pi,(t_k,x')}_{t_{k+1}})\right)\xi_{k+1}\right]\end{aligned}$$

where

$$\xi_{k+1} = 1 + \Delta^\pi_{k+1}\lambda_y + \lambda_z.\Delta^\pi W_{k+1} + \lambda_\gamma.\int_E \rho(e)\overline{\mu}(de,(t_k,t_{k+1}]).$$

The Cauchy-Schwarz inequality implies

$$\begin{aligned}&\left|\tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k})\right| \le \Delta^\pi_{k+1}C\left|X^{\pi,(t_i,x)}_{t_k} - X^{\pi,(t_i,x')}_{t_k}\right|\\ &+ \mathbb{E}_{t_k}\left[\left|y^\pi_{t_{k+1}}(X^{\pi,(t_k,x)}_{t_{k+1}}) - y^\pi_{t_{k+1}}(X^{\pi,(t_k,x')}_{t_{k+1}})\right|^2\right]\mathbb{E}_{t_k}\left[|\xi_{k+1}|^2\right].\end{aligned}$$

Since $\mathbb{E}_{t_k}\left[|\xi_{k+1}|^2\right] \le \left(1+2K\Delta^\pi_{k+1}\right)$, we have

$$\begin{aligned}&\left|\tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k})\right| \le \Delta^\pi_{k+1}C\left|X^{\pi,(t_i,x)}_{t_k} - X^{\pi,(t_i,x')}_{t_k}\right|\\ &+ \mathbb{E}_{t_k}\left[\left|y^\pi_{t_{k+1}}(X^{\pi,(t_k,x)}_{t_{k+1}}) - y^\pi_{t_{k+1}}(X^{\pi,(t_k,x')}_{t_{k+1}})\right|^2\right]\left(1+C\Delta^\pi_{k+1}\right).\end{aligned} \tag{4.6.1}$$

Let $t_i < T$. We consider the stopping time

$$\tau = \inf_{t_k \in \pi} \left\{ t_k \geq t_i \middle| \left| y^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - y^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) \right| \leq K \left| X^{\pi,(t_i,x)}_{t_k} - X^{\pi,(t_i,x')}_{t_k} \right| \right\}.$$

Since $y^\pi_T = \tilde{g}$ are K-Lipschitz , we have

$$\left| y^\pi_T(X^{\pi,(t_i,x)}_T) - y^\pi_T(X^{\pi,(t_i,x')}_T) \right| \leq K \left| X^{\pi,(t_i,x)}_{t_k} - X^{\pi,(t_i,x')}_{t_k} \right|$$

and consequently $\tau \leq T$.
For $t_i \leq t_k < \tau$, let us show that :

$$\left| y^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - y^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) \right| \leq \left| \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) \right|. \tag{4.6.2}$$

Indeed, supposing $\left| y^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - y^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) \right| \neq \left| \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) \right|$, then $y^\pi_{t_k}(X^{\pi,(t_i,\tilde{x})}_{t_k}) \neq \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,\tilde{x})}_{t_k})$ for $\tilde{x} = x$ or x'. By symmetry (on (x,x') and (l,u)), we can restrict in the case of $y^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) = l(X^{\pi,(t_i,x)}_{t_k})$. Then

$$\begin{aligned} y^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) - y^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) &\geq l(X^{\pi,(t_i,x')}_{t_k}) - l(X^{\pi,(t_i,x)}_{t_k}) \\ &\geq -K \left| X^{\pi,(t_i,x')}_{t_k} - X^{\pi,(t_i,x)}_{t_k} \right|. \end{aligned}$$

Since $\left| y^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - y^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) \right| > K \left| X^{\pi,(t_i,x)}_{t_k} - X^{\pi,(t_i,x')}_{t_k} \right|$ by definition of τ $(t_k < \tau)$, we have

$$y^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) - y^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) > K \left| X^{\pi,(t_i,x')}_{t_k} - X^{\pi,(t_i,x)}_{t_k} \right|$$

and

$$y^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) > l(X^{\pi,(t_i,x)}_{t_k}) + K \left| X^{\pi,(t_i,x')}_{t_k} - X^{\pi,(t_i,x)}_{t_k} \right| \geq l(X^{\pi,(t_i,x')}_{t_k}).$$

Therefore $y^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) = \tilde{y}^\pi_{t_k} \wedge u(X^{\pi,(t_i,x)}_{t_k}) \leq \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k})$, we obtain (4.6.2) :

$$\begin{aligned} \left| y^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) - y^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) \right| &= y^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) - y^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) \\ &\leq \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) - \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}). \end{aligned}$$

Combining the inequalities (4.6.1) and (4.6.2), we have for $t_k < \tau$:

$$\begin{aligned} \left| \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x)}_{t_k}) - \tilde{y}^\pi_{t_k}(X^{\pi,(t_i,x')}_{t_k}) \right| &\leq \Delta^\pi_{k+1} C \left| X^{\pi,(t_i,x)}_{t_k} - X^{\pi,(t_i,x')}_{t_k} \right| \\ &\quad + \mathbb{E}_{t_k} \left[\left| \tilde{y}^\pi_{t_{k+1}}(X^{\pi,(t_i,x)}_{t_{k+1}}) - \tilde{y}^\pi_{t_{k+1}}(X^{\pi,(t_i,x')}_{t_{k+1}}) \right|^2 \right]^{\frac{1}{2}} \left(1 + C\Delta^\pi_{k+1}\right). \end{aligned}$$

Summing up this last inequality from t_i to τ and we deduce

$$\begin{aligned} \left| \tilde{y}^\pi_{t_i}(x) - \tilde{y}^\pi_{t_i}(x') \right| &\leq \sum_{k=i}^{n-1} \Delta^\pi_{k+1} \mathbb{E} \left[\left| X^{\pi,(t_i,x)}_{t_k} - X^{\pi,(t_i,x')}_{t_k} \right|^2 \right]^{\frac{1}{2}} \prod_{j=i}^{k} \left(1 + C\Delta^\pi_{j+1}\right) \\ &\quad + \prod_{j=i}^{k} \left(1 + C\Delta^\pi_{j+1}\right) \mathbb{E} \left[\left| y^\pi_\tau(X^{\pi,(t_i,x')}_\tau) - y^\pi_\tau(X^{\pi,(t_i,x)}_\tau) \right|^2 \right]^{\frac{1}{2}} \\ &\leq C\mathbb{E} \left[\sup \left| X^{\pi,(t_i,x)}_{t_k} - X^{\pi,(t_i,x')}_{t_k} \right|^2 \right]^{\frac{1}{2}}. \end{aligned}$$

The definition of τ and the inequality $\prod_{j=i}^{k}\left(1+C\Delta_{j+1}^{\pi}\right)\leq e^{CT}$ permit us deduce

$$\left|\tilde{y}_{t_i}^{\pi}(x)-\tilde{y}_{t_i}^{\pi}(x')\right|\leq C\mathbb{E}\left[\sup\left|X_{t_k}^{\pi,(t_i,x)}-X_{t_k}^{\pi,(t_i,x')}\right|^2\right]^{\frac{1}{2}}.$$

The Lipschitz property of the flow of X^{π} given by lemma (4.5.7) shows :

$$\left|\tilde{y}_{t_i}^{\pi}(x)-\tilde{y}_{t_i}^{\pi}(x')\right|\leq C\left|x-x'\right|,$$

and $y_{t_i}^{\pi}=l\vee\tilde{y}_{t_i}^{\pi}\wedge u$ are also Lipschitz-continuous (since $l,u,\tilde{y}_{t_i}^{\pi}$ are Lipschitz-continous).

3/ By definition of $X_{t_{i+1}}^{\pi,(t_i,x)}$ and because the functionnal Φ is $\mathbb{L}^{\infty}$-Lipschitz, we have :

$$\mathbb{E}\left[\left|X_{t_{i+1}}^{\pi,(t_i,x)}-\mathbb{E}\left[X_{t_{i+1}}^{\pi,(t_i,x)}\right]\right|^2\right]\leq C\left(1+|x|^2\right)\Delta_{i+1}^{\pi}.$$

It follows that

$$\begin{aligned}\left|z_{t_i}^{\pi}(x)\right|&=\left|\mathbb{E}\left[\tilde{y}_{t_{i+1}}^{\pi}(X_{t_{i+1}}^{\pi,(t_i,x)})\frac{\Delta^{\pi}W_{i+1}}{\Delta_{i+1}^{\pi}}\right]\right|\\&\leq\left|\mathbb{E}\left[\tilde{y}_{t_{i+1}}^{\pi}\left(\mathbb{E}\left[X_{t_{i+1}}^{\pi,(t_i,x)}\right]\right)\frac{\Delta^{\pi}W_{i+1}}{\Delta_{i+1}^{\pi}}\right]\right|\\&\quad+\left|\mathbb{E}\left[\left(\tilde{y}_{t_{i+1}}^{\pi}\left(X_{t_{i+1}}^{\pi,(t_i,x)}\right)-\tilde{y}_{t_{i+1}}^{\pi}\left(\mathbb{E}\left[X_{t_{i+1}}^{\pi,(t_i,x)}\right]\right)\right)\frac{\Delta^{\pi}W_{i+1}}{\Delta_{i+1}^{\pi}}\right]\right|.\end{aligned}$$

$\tilde{y}_{t_{i+1}}^{\pi}\left(\mathbb{E}\left[X_{t_{i+1}}^{\pi,(t_i,x)}\right]\right)$ is determinist, then $\mathbb{E}\left[\tilde{y}_{t_{i+1}}^{\pi}\left(\mathbb{E}\left[X_{t_{i+1}}^{\pi,(t_i,x)}\right]\right)\frac{\Delta^{\pi}W_{i+1}}{\Delta_{i+1}^{\pi}}\right]=0$. The Cauchy-Schwarz inequality and the Lipschitz property of $\tilde{y}_{t_{i+1}}^{\pi}$ imply :

$$\begin{aligned}\left|z_{t_i}^{\pi}(x)\right|&\leq\mathbb{E}\left[\left|\tilde{y}_{t_{i+1}}^{\pi}\left(X_{t_{i+1}}^{\pi,(t_i,x)}\right)-\tilde{y}_{t_{i+1}}^{\pi}\left(\mathbb{E}\left[X_{t_{i+1}}^{\pi,(t_i,x)}\right]\right)\right|^2\right]^{\frac{1}{2}}\left[\left|\frac{\Delta^{\pi}W_{i+1}}{\Delta_{i+1}^{\pi}}\right|^2\right]^{\frac{1}{2}}\\&\leq C\mathbb{E}\left[\left|X_{t_{i+1}}^{\pi,(t_i,x)}-\mathbb{E}\left[X_{t_{i+1}}^{\pi,(t_i,x)}\right]\right|^2\right]^{\frac{1}{2}}\left|\Delta_{i+1}^{\pi}\right|^{-\frac{1}{2}}\\&\leq C\left(1+|x|\right).\end{aligned}$$

Applying the same reasoning to γ_{t_i} gives that $\left|\gamma_{t_i}^{\pi}(x)\right|\leq C\left(1+|x|\right)$ uniformly in π.

4/ Representation gives

$$\begin{aligned}Z_t^{\pi}&=\mathbb{E}_t\left[y_{\overline{\pi}(t)}^{\pi}\left(X_{\overline{\pi}(t)}^{\pi}\right)\frac{W_{\overline{\pi}(t)}-W_t}{\overline{\pi}(t)-t}\right],\\\Gamma_t^{\pi}&=\int_E\rho(e)\mathbb{E}_t\left[y_{\overline{\pi}(t)}^{\pi}\left(\left(S_{\overline{\pi}(t)}+\beta(S_{\pi(t)}^{\pi},e),\Phi_t(S_{\pi}^{\pi}+1_t\beta(S_{\pi(t)}^{\pi},e))\right)\right)-\tilde{y}_{\overline{\pi}(t)}^{\pi}\left(X_{\overline{\pi}(t)}^{\pi}\right)\right]\lambda(de),\\U_s^{\pi}(e)&=\mathbb{E}_t\left[y_{\overline{\pi}(t)}^{\pi}\left(\left(S_{\overline{\pi}(t)}+\beta(S_{\pi(t)}^{\pi},e),\Phi_t(S_{\pi}^{\pi}+1_t\beta(S_{\pi(t)}^{\pi},e))\right)\right)-\tilde{y}_{\overline{\pi}(t)}^{\pi}\left(X_{\overline{\pi}(t)}^{\pi}\right)\right],\end{aligned}$$

and from the Lipschitz property of y^{π} we deduce

$$\mathbb{E}\left[\sup_{0\leq t\leq T,e\in E}\left(|Z_t^{\pi}|^2+|\Gamma_t^{\pi}|^2+U_s^{\pi}(e)\right)\right]\leq C$$

with M is independent of π.

5/ Suppose (Am) satisfied. Let us show that $\mathbb{E}\left[\left|K_T^{\pi,\pm}\right|^2\right]$ is bounded independently of π. For $t_i \in \pi$, we define

$$V_{t_i}^\pi = \mathbb{E}_{t_i}\left[\tilde{g}(X_T^\pi) + \sum_{t_k\in\pi, t_i\le t_k<T} \Delta_k^\pi f(\Theta_{t_k}^\pi)\right],$$

$$\tilde{L}_{t_i}^\pi = \begin{cases} l(X_{t_i}^\pi) - V_{t_i} & \text{if } t_i < T \\ 0 & \text{if } t_i = T \end{cases}$$

$$\tilde{U}_{t_i}^\pi = \begin{cases} u(X_{t_i}^\pi) - V_{t_i} & \text{if } t_i < T \\ 0 & \text{if } t_i = T. \end{cases}$$

The dynamic programming principle gives a representation of Y^π

$$Y_{t_i}^\pi = \text{Esssup}_{\tau\in\mathcal{T}_{t_i}^\pi}\text{Essinf}_{\theta\in\mathcal{T}_{t_i}^\pi}\mathbb{E}_{t_i}\left[\sum_{t_i\le t_k<\tau\wedge\theta} \Delta_k^\pi f(\Theta_{t_k}^\pi) + +l(X_\tau^\pi)1_{\tau<\theta\wedge T} + u(X_\theta^\pi)1_{\theta\le\tau<T} + \tilde{g}(X_T^\pi)\right]$$
$$= V_{t_i}^\pi + D_{t_i}^\pi,$$

where $\mathcal{T}_t^\pi$ denotes the set of stopping times $\tau \in \pi$ such that $\tau \ge t$ and

$$D_{t_i}^\pi := \text{Esssup}_{\tau\in\mathcal{T}_{t_i}^\pi}\text{Essinf}_{\theta\in\mathcal{T}_{t_i}^\pi}\mathbb{E}\left[\tilde{L}_\tau^\pi 1_{\tau<\theta} + \tilde{U}_\theta^\pi 1_{\tau\ge\theta}\right].$$

For a process $P \in \mathcal{S}_\pi^2$, we denote $\{S(P)_{t_k}\}_{t_k\in\pi}$ its (discret) Snell envelope which is defined by

$$S(P)_{t_i} = \text{esssup}_{\tau\in\mathcal{T}_{t_i}^\pi}\mathbb{E}_{t_i}[P_\tau], \quad \forall t_i \in \pi.$$

$S(P)$ is the smallest supermartingale such that a.s. $S(P) \ge P$.

Let us consider the supermartingale $\tilde{J}_{t_i}^{\pi,+}, \tilde{J}_{t_i}^{\pi,-} \in \mathcal{S}_\pi^2$ defined by :

$$\tilde{J}_T^{\pi,+} = \tilde{J}_T^{\pi,-} = 0,$$

$$\tilde{J}_{t_i}^{\pi,+} = J_{t_i}^{\pi,+} - \mathbb{E}_{t_i}\left[\tilde{g}(X_T^\pi)^+\right] - \mathbb{E}_{t_i}\left[\sum_{t_k\in\pi, t_i\le t_k<T} \Delta_k^\pi f(\Theta_{t_k}^\pi)^+\right],$$

$$\tilde{J}_{t_i}^{\pi,-} = J_{t_i}^{\pi,-} - \mathbb{E}_{t_i}\left[\tilde{g}(X_T^\pi)^-\right] - \mathbb{E}_{t_i}\left[\sum_{t_k\in\pi, t_i\le t_k<T} \Delta_k^\pi f(\Theta_{t_k}^\pi)^-\right], \quad \forall 0 \le i < n,$$

where $J^{\pi,\pm} \in \mathcal{S}_\pi^2$ are the supermartingale of the assumption (Am). Then

$$\tilde{L}_{t_i}^\pi \le \tilde{J}_{t_i}^{\pi,+} - \tilde{J}_{t_i}^{\pi,-} \le \tilde{U}_{t_i}^\pi, \quad \forall t_i \in \pi, \tag{4.6.3}$$

and

$$\mathbb{E}\left[\sup_{t_i\in\pi}\left|\tilde{J}_{t_i}^{\pi,\pm}\right|^2\right] \le C\mathbb{E}\left[\sup_{t_i\in\pi}\left(\left|J_{t_i}^{\pi,\pm}\right|^2 + \left|f(\Theta_{t_i}^\pi)\right|^2\right)\right]$$
$$\le C.$$

Let us consider the sequences $\left\{C_{t_k}^{n,\pm}\right\}_{t_k\in\pi}$ $\mathcal{F}_{t_k}$-adapted defined by : $C_{t_k}^{0,\pm} = 0$ and $\forall n \ge 0$

$$C^{n+1,+} = S(\tilde{L}^\pi + C^{n,-}) \text{ and } C^- = S(-\tilde{U}^\pi + C^{n,+}).$$

The inequality (4.6.3) implies by induction :

$$0 \le C^{n,+} \le C^{n+1,+} \le \tilde{J}_{t_i}^{\pi,+} \text{ and } 0 \le C^{n,-} \le C^{n+1,-} \le \tilde{J}_{t_i}^{\pi,-}.$$

Then, the sequences $C^{n,\pm}$ converge pointwisely to a supermartingale $C^{\pm}$ which satisfies

$$0 \leq C^{+} \leq \tilde{J}^{\pi,+}_{t_i} \text{ and } 0 \leq C^{-} \leq \tilde{J}^{\pi,-}_{t_i}$$

and

$$C^{+} = S(\tilde{L}^{\pi} + C^{-}) \text{ and } C^{-} = S(-\tilde{U}^{\pi} + C^{+}).$$

By Doob-Meyer decomposition, there exist martingales $M^{\pi,\pm}$ and non-decreasing processes $A^{\pi,\pm}_{t_i} = \sum_{k=0}^{i-1} \Delta A^{\pi,\pm}_{t_k}$ ($A^{\pi,\pm}_0 = 0$) such that $C^{\pm} = M^{\pi,\pm} + A^{\pi,\pm}$ (where $\Delta A^{\pi,\pm}_{t_k} = \mathbb{E}_{t_k}\left[C^{\pm}_{t_{k+1}} - C^{\pm}_{t_k}\right]$). We can easily verify that :

$$\begin{aligned} C^{+}_{t_i} - C^{-}_{t_i} &= \tilde{L}^{\pi}_{t_i} \vee \mathbb{E}_{t_i}\left[C^{+}_{t_{i+1}} - C^{-}_{t_{i+1}}\right] \wedge \tilde{U}^{\pi}_{t_i} \\ &= D^{\pi}_{t_i}, \end{aligned}$$

and then $\forall 0 \leq i < n$

$$\begin{aligned} Y^{\pi}_{t_i} &= C^{+}_{t_i} - C^{-}_{t_i} + V_{t_i} \\ &= \mathbb{E}_{t_i}\left[C^{+}_{t_{i+1}} - C^{-}_{t_{i+1}} + V_{t_{i+1}}\right] + \Delta^{\pi}_i f(\Theta^{\pi}_{t_i}) + \Delta A^{\pi,+}_{t_i} - \Delta A^{\pi,-}_{t_i} \\ &= \tilde{Y}^{\pi}_{t_i} + \Delta K^{\pi,+}_{t_i} - \Delta K^{\pi,-}_{t_i}. \end{aligned}$$

By uniqueness, we have

$$\begin{aligned} \Delta K^{\pi,+}_{t_i} &= \left(l(X^{\pi}_{t_i}) - \tilde{Y}^{\pi}_{t_i}\right)^{+} = \Delta A^{\pi,+}_{t_i}, \\ \Delta K^{\pi,-}_{t_i} &= \left(\tilde{Y}^{\pi}_{t_i} - u(X^{\pi}_{t_i})\right)^{+} = \Delta A^{\pi,-}_{t_i}. \end{aligned}$$

Finally, $\left|\left|K^{\pi,\pm}_T\right|\right|_2$ is bounded by a constant independent of π :

$$\begin{aligned} \mathbb{E}\left[\left|K^{\pi,\pm}_T\right|^2\right] &= \mathbb{E}\left[\left|A^{\pi,\pm}_T\right|^2\right] \\ &\leq C\left|\left|\tilde{J}^{\pi,\pm}_{t_i}\right|\right|^2_{\mathcal{S}^2_\pi} \leq C. \end{aligned}$$

□

We can now prove the Theorem (4.6.3) that shows the convergence of the discrete-time scheme.

Démonstration. (Proof of Theorem (4.6.3))

1/ For $V = Y, Z, \overline{Z}, \Gamma, \overline{\Gamma}, U$ or $\overline{U}$, we set $\delta V = V - V^{\pi}$.

Let $t \in [0,T]$. We consider the times τ, τ^c and τ^d defined by

$$\tau^c = \inf\left\{s \geq t | \int_t^s 1_{\delta Y_s > 0} dK^{+}_s + \int_t^s 1_{\delta Y_s < 0} dK^{-}_s > 0\right\} \wedge T,$$

$$\tau^d = \inf\left\{t_j \in \pi \cap [t,T] | 1_{\delta Y_{t_j} < 0}\left(l(X^{\pi}_{t_j}) - \tilde{Y}^{\pi}_{t_j}\right)^{+} + 1_{\delta Y_{t_j} > 0}\left(\tilde{Y}^{\pi}_{t_j} - u(X^{\pi}_{t_j})\right)^{+} > 0\right\} \wedge T,$$

$$\tau = \tau^c \wedge \tau^d.$$

τ^c is the entry time of a continous $\mathcal{F}_t$-adapted process in a open subset of $\mathbb{R}$ and τ^d is the entry time of a $\mathcal{F}_{t_i}$-adapted, discrete-time process in a open subset of $\mathbb{R}$, therefore τ, τ^c and τ^d are $\mathcal{F}_t$-stopping times.

We use Ito's formula between t and τ :

$$|\delta Y_t|^2 = |\delta Y_\tau|^2 + \sum_{t_k\in\pi, t\le t_k<\tau}\left(|\delta Y_{t_k}|^2 - \left|Y_{t_k}-\tilde{Y}^\pi_{t_k}\right|^2\right) + 2\int_t^\tau \delta Y_s\left(dK^+_s - dK^-_s\right)$$

(4.6.4)

$$+\int_t^\tau \delta Y_s\eta_s ds - \int_t^\tau |\delta Z_s|^2 ds - \int_t^\tau\int_E |\delta U_s(e)|^2\lambda(de)ds + M_\tau - M_t.$$

with $\eta_s = 2\left(f(\Theta_s) - f(\Theta^\pi_s)\right) + 2\int_E \delta U_s(e)\lambda(de)ds$ and where $M_t = 2\int_0^t \delta Y_s\delta Z_s dW_s + \int_0^t\int_E |\delta U_s(e)|^2\,\overline{\mu}(de,ds)$ is a martingale. Here we try to estimate each term of this equality.

If $t_k < \tau \le \tau^d$ for $t_k\in\pi$, we have

$$|\delta Y_{t_k}| - \left|Y_{t_k}-\tilde{Y}^\pi_{t_k}\right| \le 0. \tag{4.6.5}$$

Indeed, there are three cases by definition on τ^d.
◇ If $\tilde{Y}^\pi_{t_k}\in[l(X^\pi_{t_k}), u(X^\pi_{t_k})]$, then $\tilde{Y}^\pi_{t_k} = Y^\pi_{t_k}$.
◇ If $\tilde{Y}^\pi_{t_k} < l(X^\pi_{t_k})$ and $\delta Y_{t_k}\ge 0$, then $\tilde{Y}^\pi_{t_k} - Y_{t_k} < l(X^\pi_{t_j}) - Y_{t_k} = -\delta Y_{t_k}\le 0$ and $\left|Y_{t_k}-\tilde{Y}^\pi_{t_k}\right| > |\delta Y_{t_k}|$.
◇ If $\tilde{Y}^\pi_{t_k} > u(X^\pi_{t_k})$ and $\delta Y_{t_k}\le 0$, then $\tilde{Y}^\pi_{t_k} - Y_{t_k} > u(X^\pi_{t_k}) - Y_{t_k} = -\delta Y_{t_k}\ge 0$ and $\left|Y_{t_k}-\tilde{Y}^\pi_{t_k}\right| > |\delta Y_{t_k}|$.

For $s\le t\le\tau\le\tau^c$, the definition of τ^c implies

$$\int_s^t \delta Y_t\left(dK^+_t - dK^-_t\right)\le 0. \tag{4.6.6}$$

For $|\delta Y_\tau|^2$, we consider all the different cases :
◇ If $\tau = T$, we have $|\delta Y_\tau|\le K|X_T - X^\pi_T|$.
◇ If $\tau = \tau^c\in]t_k, t_{k+1}[$ $(t_k\in\pi)$ and if $1_{Y_\tau = l(X_\tau)}\delta Y_\tau > 0$, then

$$\begin{aligned}|\delta Y_\tau| &= l(X_\tau) - Y^\pi_\tau\\ &= l(X_\tau) - \mathbb{E}_\tau\left[Y^\pi_{\overline{\pi}(\tau)}\right] - (\overline{\pi}(\tau)-\tau)f(\Theta^\pi_{t_k})\\ &\le \mathbb{E}_\tau\left[l(X_\tau) - l(X^\pi_{t_{k+1}})\right] + |\pi|\left|f(\Theta^\pi_{t_k})\right|.\end{aligned}$$

◇ If $\tau = \tau^c\in]t_k, t_{k+1}[$ and if $1_{Y_\tau = u(X_\tau)}\delta Y_\tau < 0$, then

$$\begin{aligned}|\delta Y_\tau| &= \mathbb{E}_\tau\left[Y^\pi_{t_{k+1}}\right] - (t_{k+1}-\tau)f\left(\Theta^\pi_{t_k}\right) - u(X_\tau)\\ &\le \mathbb{E}_\tau\left[u(X^\pi_{t_{k+1}}) - u(X_\tau)\right] + |\pi|\left|f\left(\Theta^\pi_{t_k}\right)\right|.\end{aligned}$$

◇ If $\tau = \tau^d\in\pi$ and if $1_{\delta Y_\tau<0}\left(l(X^\pi_\tau) - \tilde{Y}^\pi_\tau\right) > 0$, then

$$\begin{aligned}|\delta Y_\tau| &= l(X^\pi_\tau) - Y_\tau\\ &\le l(X^\pi_\tau) - l(X_\tau).\end{aligned}$$

◇ If $\tau = \tau^d\in\pi$ and if $1_{\delta Y_\tau>0}\left(\tilde{Y}^\pi_\tau - u(X^\pi_\tau)\right) > 0$, then

$$\begin{aligned}|\delta Y_\tau| &= Y_\tau - u(X^\pi_\tau)\\ &\le u(X_\tau) - u(X^\pi_\tau).\end{aligned}$$

By the inequality $\left|\Theta^\pi_{t_k}\right| \leq C\left(1+\left|X^\pi_{t_k}\right|\right)\ \forall t_k \in \pi$ given by lemma (4.6.4) and by the Lipschitz property of (f,g,u), we get :

$$|\delta Y_\tau| \leq K\mathbb{E}_\tau\left[\left|X_\tau - X_{\overline{\pi}(\tau)}\right| + \left|X_{\overline{\pi}(\tau)} - X^\pi_{\overline{\pi}(\tau)}\right|\right] + |\pi|\, C\left(1+\left|X^\pi_{\pi(\tau)}\right|\right),$$

$$\mathbb{E}_t\left[|\delta Y_\tau|^2\right] \leq C\mathbb{E}_t\left[\sup_{0\leq s\leq T}\mathbb{E}_s\left[\left|X_s - X_{\overline{\pi}(s)}\right|^2\right]\right] + C\mathbb{E}_t\left[\sup_{t_k\in\pi}\left|X_{t_k} - X^\pi_{t_k}\right|^2\right]$$

$$+ |\pi|\, C\mathbb{E}_t\left(1+\sup_{t\leq s\leq T}\left|X^\pi_{\pi(s)}\right|^2\right). \tag{4.6.7}$$

We note $0 \leq K_f \leq K$ a constant such that $\forall x \in \mathbb{R}^d$, $\forall y \in \mathbb{R}$, $\forall z, z' \in \mathbb{R}^{d_Z}$, $\forall \gamma, \gamma' \in \mathbb{R}^m$

$$|f(x,y,z,\gamma) - f(x,y,z',\gamma')| \leq K_f\left(|z-z'| + |\gamma-\gamma'|\right),$$

and $K_f = 0$ if $f(x,y,z,\gamma)$ is independent of (z,γ). From the Lipschitz property of f and the inequality $ab \leq \frac{\alpha a^2}{2} + \frac{b^2}{2\alpha}$ (for $\alpha > 0$) , we get

$$\begin{aligned}
\delta Y_s(f(\Theta_s) - f(\Theta^\pi_s)) &\leq |\delta Y_s|\left\{K\left|X_s - X^\pi_{\pi(s)}\right| + K\left|Y_s - \tilde{Y}^\pi_{\pi(s)}\right| + K_f\left|Z_s - \overline{Z}^\pi_s\right| + K_f\left|\Gamma_s - \overline{\Gamma}^\pi_s\right|\right\} \\
&\leq \frac{K}{\alpha}\left\{\left|X_s - X^\pi_{\pi(s)}\right|^2 + \left|Y^\pi_s - \tilde{Y}^\pi_{\pi(s)}\right|^2 + \left|\delta\overline{Z}_s\right| + \left|\delta\overline{\Gamma}_s\right|\right\} \\
&\quad + \frac{K_f}{2\alpha}\left\{\left|Z_s - \overline{Z}^\pi_s\right|^2 + \left|\Gamma_s - \overline{\Gamma}^\pi_s\right|^2\right\} + |\delta Y_s|^2\left\{K + \alpha(2K+K_f)\right\},
\end{aligned} \tag{4.6.8}$$

and

$$\begin{aligned}
\delta Y_s\int_E \delta U_s(e)\lambda(de) &\leq \frac{\alpha}{2}|\delta Y_s|^2 + \frac{1}{2\alpha}\left|\int_E \delta U_s(e)\lambda(de)\right|^2 \\
&\leq \frac{\alpha}{2}|\delta Y_s|^2 + \frac{\lambda(E)}{2\alpha}\int_E |\delta U_s(e)|^2\,\lambda(de).
\end{aligned} \tag{4.6.9}$$

Combining (4.6.5), (4.6.6), (4.6.7), (4.6.8) et (4.6.9) with (4.6.4), we obtain

$$\begin{aligned}
|\delta Y_t|^2 &\leq C\mathbb{E}_t\left[\sup_{0\leq s\leq T}\mathbb{E}_s\left[\left|X_s - X_{\overline{\pi}(s)}\right|^2\right]\right] + C\mathbb{E}_t\left[\sup_{t_k\in\pi}\left|X_{t_k} - X^\pi_{t_k}\right|^2\right] \\
&+ \mathbb{E}_t\left[|\pi|\, C\left(1+\sup_{t_k\in\pi}\left|X^\pi_{t_k}\right|^2\right)\right] + \frac{K}{\alpha}\mathbb{E}_t\left[\int_t^T\left|\tilde{Y}^\pi_s - \tilde{Y}^\pi_{\pi(s)}\right|^2 ds\right] \\
&+ \left\{2K + \alpha(1+2K+2K_f)\right\}\mathbb{E}_t\left[\int_t^T |\delta Y_s|^2\, ds\right] + \mathbb{E}_t\left[\int_t^\tau P_s(\alpha)ds\right]
\end{aligned} \tag{4.6.10}$$

where

$$P_s(\alpha) = \frac{K_f}{\alpha}\left|Z_s - \overline{Z}^\pi_s\right|^2 - |\delta Z_s|^2 + \frac{K_f}{\alpha}\left|\Gamma_s - \overline{\Gamma}^\pi_s\right|^2 + \left(\frac{\lambda(E)}{2\alpha} - 1\right)\int_E |\delta U_s(e)|^2\,\lambda(de).$$

For $t_k \in \pi \backslash \{T\}$, we have

$$\begin{aligned}\mathbb{E}_t\left[\int_{t_k}^{t_{k+1}} P_s(\alpha)ds\right] \leq \frac{2K_f}{\alpha}\mathbb{E}_t\left[\int_{t_k}^{t_{k+1}} \left|Z_s - \overline{Z}_s\right|^2 + \left|\Gamma_s - \overline{\Gamma}_s\right|^2 ds\right] \\ + \mathbb{E}_t\left[(t_{k+1}-t_k)\frac{2K_f}{\alpha}\left|\delta\overline{Z}_{t_k}\right|^2 - \int_{t_k}^{t_{k+1}} |\delta Z_s|^2 ds\right] \\ + \mathbb{E}_t\left[(t_{k+1}-t_k)\frac{2K_f}{\alpha}\left|\delta\overline{\Gamma}_{t_k}\right|^2 - \int_{t_k}^{t_{k+1}} \left(\frac{\lambda(E)}{2\alpha} - 1\right)\int_E |\delta U_s(e)|^2 \lambda(de)ds\right].\end{aligned}$$

Applying Jensen inequality, we obtain

$$\begin{aligned}\mathbb{E}_t\left[\left|\delta\overline{Z}_{t_k}\right|^2\right] &\leq \mathbb{E}_t\left[\frac{1}{t_{k+1}-t_k}\int_{t_k}^{t_{k+1}} |\delta Z_s|^2 ds\right] \\ \mathbb{E}_t\left[\left|\delta\overline{\Gamma}_{t_k}\right|^2\right] &\leq \mathbb{E}_t\left[\frac{1}{t_{k+1}-t_k}\int_{t_k}^{t_{k+1}} |\delta\Gamma_s|^2 ds\right] \\ &\leq K\lambda(E)\mathbb{E}_t\left[\frac{1}{t_{k+1}-t_k}\int_{t_k}^{t_{k+1}}\int_E |\delta U_s(e)|^2 \lambda(de)ds\right].\end{aligned}$$

With $\alpha^* > \sup\left\{2K_f, (2KK_f + \frac{1}{2})\lambda(E)\right\}$, we get

$$\mathbb{E}_t\left[\int_{t_k}^{t_{k+1}} P_s(\alpha^*)ds\right] \leq \frac{2K_f}{\alpha^*}\mathbb{E}_t\left[\int_{t_k}^{t_{k+1}} \left\{\left|Z_s - \overline{Z}_s\right|^2 + \left|\Gamma_s - \overline{\Gamma}_s\right|^2\right\} ds\right].$$

Moreover,

$$P_s(\alpha^*) \leq \frac{2K_f}{\alpha^*}\left\{\left|Z_s^\pi - \overline{Z}_s^\pi\right|^2 + \left|\Gamma_s^\pi - \overline{\Gamma}_s^\pi\right|^2\right\},$$

therefore

$$\begin{aligned}\mathbb{E}_t\left[\int_t^\tau P_s(\alpha)ds\right] &= \mathbb{E}_t\left[\int_t^{\overline{\pi}(t)} P_s(\alpha^*)ds + \int_{\pi(\tau)}^\tau P_s(\alpha^*)ds\right] + \mathbb{E}_t\left[\int_{\overline{\pi}(t)}^{\pi(\tau)} P_s(\alpha^*)ds\right] \\ &\leq CK_f\,|\pi|\,\mathbb{E}_t\left[\sup_{0\leq s\leq T}|Z_s^\pi|^2 + \sup_{0\leq s\leq T}|\Gamma_s^\pi|^2\right] \\ &\quad + CK_f\mathbb{E}_t\left[\int_0^T \left\{\left|Z_s - \overline{Z}_s\right|^2 + \left|\Gamma_s - \overline{\Gamma}_s\right|^2\right\} ds\right].\end{aligned}$$

The inequality (4.6.10) becomes

$$|\delta Y_t|^2 \leq C\mathbb{E}_t\left[R^\pi\right] + C\mathbb{E}_t\left[\int_t^T |\delta Y_s|^2 ds\right].$$

where $R^\pi = R_0^\pi + K_f R_1^\pi$ with

$$\begin{aligned}R_0^\pi &= \sup_{0\leq s\leq T}\mathbb{E}_s\left[\left|X_s - X_{\overline{\pi}(s)}\right|^2\right] + \sup_{t_k\in\pi}\left|X_{t_k} - X_{t_k}^\pi\right|^2 + C\int_0^T \left|\tilde{Y}_s^\pi - \tilde{Y}_{\pi(s)}^\pi\right|^2 ds \\ &\quad + C\,|\pi|\left\{1 + \sup_{t_k\in\pi}\left|X_{t_k}^\pi\right|^2 + \sup_{0\leq s\leq T}|Z_s^\pi|^2 + \sup_{0\leq s\leq T}|\Gamma_s^\pi|^2\right\}, \\ R_1^\pi &= C\int_0^T \left\{\left|Z_s - \overline{Z}_s\right|^2 + \left|\Gamma_s - \overline{\Gamma}_s\right|^2\right\} ds.\end{aligned}$$

Appling the Gronwall Lemma, we obtain :

$$|\delta Y_t|^2 \leq C\mathbb{E}_t\left[R^\pi\right],$$
$$\mathbb{E}\left[\sup_{0\leq s\leq T}|\delta Y_s|^2\right] \leq C\mathbb{E}\left[R^\pi\right].$$

From lemma (4.6.4), we get

$$\mathbb{E}\left[\sup_{t_k\in\pi}\left|X^\pi_{t_k}\right|^2 + \sup_{0\leq s\leq T}|Z^\pi_s|^2 + \sup_{0\leq s\leq T}|\Gamma^\pi_s|^2\right] \leq C$$
$$\mathbb{E}\left[\left|\tilde{Y}^\pi_s - \tilde{Y}^\pi_{\pi(s)}\right|^2\right] \leq C\mathbb{E}\left[\int_{\pi(s)}^s (|Z^\pi_u|^2 + \int_E |U^\pi_u(e)|^2\,\lambda(de) + |f(\Theta^\pi_u)|^2)du\right]$$
$$\mathbb{E}\left[\int_0^T \left|Y^\pi_s - Y^\pi_{\pi(s)}\right|^2 ds\right] \leq C\,|\pi|\,.$$

Since $||X - X^\pi||^2_{\mathcal{S}^2_\pi} \leq C\,|\pi|\ln(\frac{1}{|\pi|})$ by Lemma (4.5.2) and
$\mathbb{E}\left[\sup_{0\leq s\leq T}\mathbb{E}_s\left[\left|X_s - X_{\overline{\pi}(s)}\right|^2\right]\right] \to 0$ as $|\pi| \to 0$, we have $\mathbb{E}\left[R^\pi_0\right] \to 0$ as $|\pi| \to 0$.
Otherwise, Jensen's inequality implies

$$\begin{aligned}\left|\left|Z - \overline{Z}\right|\right|^2_{\mathbb{H}^2} &\leq \left|\left|Z - \overline{Z}\right|\right|^2_{\mathbb{H}^2} + \left|\left|\overline{Z} - \overline{Z}^\pi\right|\right|^2_{\mathbb{H}^2}\\ &\leq \left|\left|Z - \overline{Z}\right|\right|_{\mathbb{H}^2} + ||Z - Z^\pi||^2_{\mathbb{H}^2}\,.\end{aligned}$$

Z is $\mathcal{F}_t$-adapted, there exists a sequence of adapted processes $(Z^n)_n$ constant on $[t_i, t_{i+1})$ converging to Z in $\mathbb{H}^2$. The inequality $\left|\left|Z - \overline{Z}\right|\right|_{\mathbb{H}^2_{[t_i,t_{i+1}]}} \leq \inf_{z_{i\in\mathcal{F}_{t_i}}}\left|\left|Z - \overline{Z}\right|\right|_{\mathbb{H}^2_{[t_i,t_{i+1}]}}$ implies $\left|\left|Z - \overline{Z}\right|\right|_{\mathbb{H}^2} \leq ||Z - Z^n||_{\mathbb{H}^2}$ and using the same reasoning to Γ, we obtain the convergence of $\left(\overline{Z}, \overline{\Gamma}\right)$ to (Z, Γ) in $\mathbb{H}^2$ and consequently the convergence of Y^π to Y in $\mathcal{S}^2$ since $\mathbb{E}\left[R^\pi_1\right] = C\left|\left|Z - \overline{Z}\right|\right|^2_{\mathbb{H}^2} + C\left|\left|\Gamma - \overline{\Gamma}\right|\right|^2_{\mathbb{H}^2} \to 0$ as $|\pi| \to 0$.

2/ If the assumption (At) is satisfied and $f(x, y, z, \gamma)$ is independent of (z, γ), then the rate of the convergence of the algorithm is known :

$$\begin{aligned}||\delta Y||^2_{\mathcal{S}^2} &\leq C\mathbb{E}\left[R_0\right]\\ &\leq C\,|\pi|\ln(\frac{1}{|\pi|}) + \mathbb{E}\left[\sup_{0\leq s\leq T}\mathbb{E}_s\left[\left|X_s - X_{\overline{\pi}(s)}\right|^2\right]\right]\\ &\leq C\,|\pi|\ln(\frac{1}{|\pi|}).\end{aligned}$$

3/ Assume (Am). Jensen's inequality implies

$$\begin{aligned}&\left|\left|Z - \overline{Z}^\pi\right|\right|^2_{\mathbb{H}^2}\left|\left|\Gamma - \overline{\Gamma}^\pi\right|\right|^2_{\mathbb{H}^2}\\ \leq\ &\left|\left|Z - \overline{Z}\right|\right|^2_{\mathbb{H}^2} + \left|\left|\overline{Z} - \overline{Z}^\pi\right|\right|^2_{\mathbb{H}^2} + \left|\left|\Gamma - \overline{\Gamma}\right|\right|^2_{\mathbb{H}^2} + \left|\left|\overline{\Gamma} - \overline{\Gamma}^\pi\right|\right|^2_{\mathbb{H}^2}\\ \leq\ &\left|\left|Z - \overline{Z}\right|\right|_{\mathbb{H}^2} + ||Z - Z^\pi||^2_{\mathbb{H}^2} + \left|\left|\Gamma - \overline{\Gamma}\right|\right|^2_{\mathbb{H}^2} + ||\Gamma - \Gamma^\pi||^2_{\mathbb{H}^2}\,.\end{aligned}$$

Since $\left(\overline{Z},\overline{\Gamma}\right) \rightarrow_{|\pi|\rightarrow 0} (Z,\Gamma)$ in $\mathbb{H}^2$, we need to prove the convergence of $||Z-Z^\pi||_{\mathbb{H}^2}+||\Gamma-\Gamma^\pi||_{\mathbb{H}^2}$ to 0 to conclude the theorem. Using Ito's formula between 0 and T, we deduce

$$\begin{aligned}
&\mathbb{E}\left[|\delta Y_0|^2\right]+\mathbb{E}\left[\int_0^T|\delta Z_s|^2\,ds+\int_0^T\int_E|\delta U_s(e)|^2\,\lambda(de)ds\right]\\
\leq&\mathbb{E}\left[|\delta Y_T|^2\right]+\mathbb{E}\left[\int_0^T\delta Y_s\eta_s ds\right]\\
+&\mathbb{E}\left[\sum_{t_k\in\pi}\left(|\delta Y_{t_k}|^2-\left|Y_{t_k}-\tilde{Y}^\pi_{t_k}\right|^2\right)\right]+\mathbb{E}\left[2\int_0^T|\delta Y_s|\left(dK^+_s+dK^-_s\right)\right] \qquad (4.6.11)
\end{aligned}$$

with

$$||\eta||_{\mathbb{H}^2}=2\left|\left|\int_E\delta U(e)\lambda(de)+f(\Theta)-f(\Theta^\pi)\right|\right|_{\mathbb{H}^2}\leq C.$$

Combining (4.6.11) with the inequality

$$|\delta Y_{t_k}|^2-\left|Y_{t_k}-\tilde{Y}^\pi_{t_k}\right|^2\leq\left|Y_{t_k}-\tilde{Y}^\pi_{t_k}\right|\left(\Delta K^{\pi,+}_{t_k}+\Delta K^{\pi,-}_{t_k}\right)+\left|\Delta K^{\pi,+}_{t_k}\right|^2+\left|\Delta K^{\pi,-}_{t_k}\right|^2$$

and using Cauchy-Schwarz inequality, it follows that

$$\begin{aligned}
&||Z-Z^\pi||^2_{\mathbb{H}^2}+||\Gamma-\Gamma^\pi||^2_{\mathbb{H}^2}\\
\leq&\mathbb{E}\left[\int_0^T|\delta Z_s|^2\,ds+\lambda(E)\,|\rho|^2_\infty\int_0^T\int_E|\delta U_s(e)|^2\,\lambda(de)ds\right]\\
\leq&C\,||\delta Y||_{\mathcal{S}^2}\left\{||\delta Y||_{\mathcal{S}^2}+||V^\pi||_{\mathbb{H}^2}+\left|\left|K^+_T\right|\right|_{\mathbb{L}^2}+\left|\left|K^-_T\right|\right|_{\mathbb{L}^2}\right\}\\
+&C\left\{\left|\left|K^{\pi,+}_T\right|\right|_{\mathbb{L}^2}+\left|\left|K^{\pi,-}_T\right|\right|_{\mathbb{L}^2}\right\}\left\{||\delta Y||_{\mathcal{S}^2}+\left|\left|\Delta K^{\pi,+}_{t_k}\right|\right|_{\mathcal{S}^2_\pi}+\left|\left|\Delta K^{\pi,-}_{t_k}\right|\right|_{\mathcal{S}^2_\pi}\right\}.
\end{aligned}$$

We have proved that $||\delta Y||_{\mathcal{S}^2}\rightarrow 0$ as $|\pi|\rightarrow 0$ and we have

$$\begin{aligned}
\left|\left|\Delta K^{\pi,+}_{t_k}\right|\right|_{\mathcal{S}^2_\pi}&=\left|\left|\sup_{t_k\in\pi}\left(l(X^\pi_{t_k})-\tilde{Y}^\pi_{t_k}\right)^+\right|\right|_{\mathbb{L}^2}\\
&\leq\left|\left|\sup_{t_k\in\pi}\left(l(X^\pi_{t_k})-\mathbb{E}_{t_k}\left[l(X^\pi_{t_{k+1}})+\Delta^\pi_k f(X^\pi_{t_k},\tilde{Y}^\pi_{t_k},\overline{Z}^\pi_{t_k},\overline{\Gamma}^\pi_{t_k})\right]\right)^+\right|\right|_{\mathbb{L}^2}\\
&\leq C\,|\pi|+C\mathbb{E}\left[\sup_{t_k\in\pi}\left|\mathbb{E}_{t_k}\left[X^\pi_{t_k}-X^\pi_{t_{k+1}}\right]\right|^2\right]^{\frac{1}{2}}\rightarrow_{|\pi|\rightarrow 0}0.
\end{aligned}$$

We suppose that assumption (Am) is verified. By lemma (4.6.4), $\left|\left|K^{\pi,\pm}_T\right|\right|_{\mathbb{L}^2}\leq C$. Therefore,$||Z-Z^\pi||^2_{\mathbb{H}^2}+||\Gamma-\Gamma^\pi||^2_{\mathbb{H}^2}\rightarrow 0$ as $|\pi|\rightarrow 0$. □

Annexe : proofs for the Euler scheme

We now prove the lemma (4.5.1) : $\exists C>0$ such that $||S-S^\pi||_{\mathcal{S}^2}\leq C\sqrt{|\pi|}$.

Démonstration. (proof of lemma (4.5.1))

We decompose the semi-martingale $S_t-S^\pi_t=M_t+A_t$ where M is a martingale and A_t a predictable process with finite variation :

$$M_t = \int_0^t \left(\sigma(S_t) - \sigma(S^\pi_{\pi(r)})\right) dW_r + \int_0^t \int_E \left(\beta(S_{r-}, e) - \beta(S^\pi_{\pi(r)-}, e)\right) \overline{\mu}(de, dr),$$
$$A_t = \int_0^t \left(b(S_r) - b(S^\pi_{\pi(r)})\right) dr.$$

Consider $\alpha(t) = \mathbb{E}\left[\sup_{0\le u\le t} |S_u - S^\pi_u|^2\right]$, we have

$$\begin{aligned}\alpha(t) &= \mathbb{E}\left[\sup_{0\le u\le t} |A_u + M_u|^2\right] \\ &\le 2\mathbb{E}\left[\sup_{0\le u\le t} |A_u|^2\right] + 2\mathbb{E}\left[\sup_{0\le u\le t} |M_u|^2\right].\end{aligned}$$

We deduce from Burkholder-Davis-Gundy inequality that

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le u\le t} |M_u|^2\right] &\le D_2 \mathbb{E}\left[[M]_t\right] \\ &\le D_2 \int_0^t \mathbb{E}\left[\left|\sigma(S_r) - \sigma(S^\pi_{\pi(r)})\right|^2 + \int_E \left|\beta(S_{r-}, e) - \beta(S^\pi_{\pi(r)-}, e)\right|^2 \lambda(de)\right] dr, \\ \mathbb{E}\left[\sup_{0\le u\le t} |A_u|^2\right] &\le C\mathbb{E}\left[\int_0^t \left|b(S_r) - b(S^\pi_{\pi(r)})\right|^2 dr\right].\end{aligned}$$

Combining these inequalities with the Lipschitz property of b, σ and $\beta(., e)$, it follows that

$$\begin{aligned}\alpha(t) &\le C\mathbb{E}\left[\int_0^t \left|S_r - S^\pi_{\pi(r)}\right|^2 dr\right] \\ &\le C\int_0^t \mathbb{E}\left[\left|S_{\pi(r)} - S^\pi_{\pi(r)}\right|^2 dr\right] + C\int_0^t \mathbb{E}\left[\left|S_r - S_{\pi(r)}\right|^2 dr\right] \\ &\le C\int_0^t \alpha(r)dr + C\left|\pi\right|.\end{aligned}$$

Gronwall lemma applied to $\alpha(t)$ leads to the required result. □

The proposition (4.5.2) gives the error of the Euler scheme with the Skorokhod topology :

$$\mathbb{E}\left[\sup_{t_i\in\pi} d_{\mathbb{D}}\left(S^{t_i}, (S^\pi)^{t_i}\right)^2\right] \le C\left|\pi\right|,$$
$$\mathbb{E}\left[\sup_{t_i\in\pi} d_{\mathbb{D}}\left(S^{t_i}, (\overline{S}^\pi)^{t_i}\right)^2\right] \le C\left|\pi\right| \ln(\frac{1}{|\pi|}).$$

Démonstration. (proof of proposition (4.5.2))
For $x, y \in \mathbb{D}$, we have $d_{\mathbb{D}}(x, y) \le |x - y|_\infty$, therefore

$$\mathbb{E}\left[\sup_{t_i\in\pi} d_{\mathbb{D}}\left(S^{t_i}, (S^\pi)^{t_i}\right)^2\right] \le \mathbb{E}\left[\sup_{0\le t\le T} |S_t - S^\pi_t|^2\right] \le C\left|\pi\right|.$$

$d_{\mathbb{D}}$ is a metric, so

$$d_{\mathbb{D}}\left(S^{t_i}, (S^\pi)^{t_i}\right) \le d_{\mathbb{D}}\left(S^{t_i}, (S^\pi)^{t_i}\right) + d_{\mathbb{D}}\left((S^\pi)^{t_i}, (\overline{S}^\pi)^{t_i}\right).$$

We have

$$d_{\mathbb{D}}\left((S^\pi)^{t_i},(\overline{S}^\pi)^{t_i}\right)=\inf_{\lambda\in\Lambda}\left(\sup_{t\in[0,T]}|\lambda(t)-t|+\sup_{t\in[0,T]}\left|S^\pi_{t\wedge t_i}-\overline{S}^\pi_{\pi(\lambda(t))\wedge t_i}\right|\right)$$

where Λ denotes the set of all strictly increasing continuous bijections from $[0,T]$ to itself.

We define $C^\pi(i)=\alpha K\left(1+\left|S^\pi_{t_i}\right|\right)\sqrt{|\pi|\ln\left(\frac{1}{\pi}\right)}$, $\forall 0\le i\le n-1$, with

$$\sup_{e\in E}\left\{|b(x)|,|\sigma|,|\beta(x,e)|\left(1\vee\lambda(E)\right)\right\}\le K\left(1+|x|\right),\quad\forall x\in\mathbb{R}^d,$$

and with $\alpha\ge 1$ determinated later. Let us define $(t^\pi_i,\theta^\pi_i)_{0\le i\le n}$ by

$$\begin{cases}t^\pi_0=t_0=0\\ \theta^\pi_i=\inf\left\{s\ge t_i\left|\left|S^\pi_s-S^\pi_{t_i}\right|\ge C_{|\pi|}\right.\right\}\wedge t_{i+1} & \forall 0\le i<n-1\\ t^\pi_i=\frac{t_i+\theta^\pi_i}{2} & \forall 1\le i<n\\ \theta^\pi_{n-1}=t^\pi_n=T\end{cases}$$

We now consider $\lambda^\pi\in\Lambda$ a piecewise linear function defined by : for $t\in[t_i,t_{i+1}]$

$$\lambda^\pi(t)=\begin{cases}t^\pi_i & \text{if } t=t_i\\ t_{i+1} & \text{if } t=\theta^\pi_i\\ t^\pi_{i+1} & \text{if } t=t_{i+1}\\ t^\pi_i+\frac{t-t_i}{\theta^\pi_i-t_i}(t_{i+1}-t^\pi_i) & \text{if } t\in[t_i,\theta^\pi_i]\\ t_{i+1}+\frac{t-\theta^\pi_i}{t_{i+1}-\theta^\pi_i}(t^\pi_{i+1}-t_{i+1}) & \text{if } t\in[\theta^\pi_i,t_{i+1}]\end{cases}$$

Then $\forall t\in[0,T]$, we have $0\le\lambda^\pi(t)-t\le 2|\pi|$ and for $t\in[t_i,t_{i+1}]$

$$\pi(\lambda^\pi(t))=t_i1_{t<\theta^\pi_i}+t_{i+1}1_{t\ge\theta^\pi_i}.$$

As $\lambda^\pi\in\Lambda$, we obtain :

$$d_{\mathbb{D}}\left((S^\pi)^{t_i},(\overline{S}^\pi)^{t_i}\right)\le\sup_{t\in[0,T]}|\lambda^\pi(t)-t|+\sup_{t\in[0,T]}\left|S^\pi_{t\wedge t_i}-\overline{S}^\pi_{\pi(\lambda^\pi(t))\wedge t_i}\right|.$$

and then

$$\sup_{0\le i\le n}d_{\mathbb{D}}\left((S^\pi)^{t_i},(\overline{S}^\pi)^{t_i}\right)\le 2|\pi|+\sup_{0\le i<n-1}\sup_{t_i\le t<t_{i+1}}\left|S^\pi_t-S^\pi_{\pi(\lambda^\pi(t))}\right|+\sup_{t\in[t_{n-1},T]}\left|S^\pi_t-S^\pi_{t_{n-1}}\right|$$

The definition of C_π and the construction of $\pi\circ\lambda^\pi$ give us $\forall 0\le i<n-1$

$$\begin{aligned}\sup_{t_i\le t<t_{i+1}}\left|S^\pi_t-S^\pi_{\pi(\lambda^\pi(t))}\right|&\le\sup_{t^\pi_i\le t<\theta^\pi_i<T}\left|S^\pi_t-S^\pi_{t_i}\right|+\sup_{\theta^\pi_i\le t<t_{i+1}}\left|S^\pi_t-S^\pi_{t_{i+1}}\right|1_{\theta^\pi_i<t_{i+1}}\\&\le C_\pi(i)+\sup_{\theta^\pi_i\le t<t_{i+1}}\left|S^\pi_t-S^\pi_{t_{i+1}}\right|1_{\theta^\pi_i<t_{i+1}}.\end{aligned}$$

From the inequalities

$$\mathbb{E}\left[\sup_{t\in[t_{n-1},T]}\left|S^\pi_t-S^\pi_{t_{n-1}}\right|^2\right]\le C|\pi|,$$

$$\mathbb{E}\left[\sup_{0\le i<n-1}C^2_\pi(i)\right]\le C|\pi|\ln\left(\frac{1}{|\pi|}\right),$$

we obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le i\le n} d_{\mathbb{D}}\left((S^\pi)^{t_i},(\overline{S}^\pi)^{t_i}\right)\right] &\le \mathbb{E}\left[\left(2\left|\pi\right| + \sup_{0\le i<n-1}\sup_{t_i\le t<t_{i+1}}\left|S^\pi_t - S^\pi_{\pi(\lambda^\pi(t))}\right| + \sup_{t\in[t_{n-1},T]}\left|S^\pi_t - S^\pi_{t_{n-1}}\right|\right)^2\right]\\ &\le C\left|\pi\right|^2 + C\mathbb{E}\left[\sup_{t\in[t_{n-1},T]}\left|S_t - S_{t_{n-1}}\right|^2\right] + C\mathbb{E}\left[\sup_{0\le i<n-1} C^2_\pi(i)\right]\\ &\quad + C\mathbb{E}\left[\sup_{0\le i<n-1}\sup_{\theta^\pi_i\le t<t_{i+1}}\left|S^\pi_t - S^\pi_{t_{i+1}}\right|^2 1_{\theta^\pi_i<t_{i+1}}\right]\\ &\le C\left|\pi\right|\ln\left(\frac{1}{|\pi|}\right) + C\sum_{i=0}^{n-1}\mathbb{E}\left[1_{\theta^\pi_i<t_{i+1}}\mathbb{E}\left[\sup_{\theta^\pi_i\le t<t_{i+1}}\left|S^\pi_t - S^\pi_{t_{i+1}}\right|^2|\mathcal{F}_{\theta^\pi_i}\right]\right].\end{aligned}$$

We deduce from the Burkholder-Davis-Gundy inequality that

$$\begin{aligned}\mathbb{E}\left[\sup_{\pi_{C_\pi}(i)\le t<t_{i+1}}\left|S^\pi_t - S^\pi_{t_{i+1}}\right|^2|\mathcal{F}_{\theta^\pi_i}\right] &\le C\mathbb{E}\left[\langle S^\pi\rangle_{[\theta^\pi_i,t_{i+1}]}|\mathcal{F}_{\theta^\pi_i}\right]\\ &\le C\left(1+\left|S^\pi_{t_i}\right|^2\right)(t_{i+1}-t_i).\end{aligned}$$

It remains to estimate

$$\begin{aligned}&\mathbb{E}\left[1_{\theta^\pi_i<t_{i+1}}|\mathcal{F}_{t_i}\right]\\ \le&\mathbb{P}\left(\sup_{0\le u<t_{i+1}-t_i}\left|b(S^\pi_{t_i})u + \sigma(S^\pi_{t_i})(W_{t_i+u}-W_{t_i}) + \int_E \beta(S^\pi_{\pi(t)},e)\overline{\mu}(de,(t_i,t_i+u])\right| \ge C^\pi(i)|\mathcal{F}_{t_i}\right)\\ \le&\mathbb{P}\left(\sup_{t_i\le u<t_{i+1}}|W_{t_i+u}-W_{t_i}|\left|\sigma(S^\pi_{t_i})\right| \ge \frac{C^\pi(i)}{2}|\mathcal{F}_{t_i}\right)\\ +&\mathbb{P}\left(\sup_{t_i\le u<t_{i+1}}\left|\int_E\beta(S^\pi_{\pi(t)},e)\mu(de,(t_i,t_{i+u}])\right| + \left|b(S^\pi_{t_i})u + \int_E\beta(S^\pi_{\pi(t)},e)\lambda(de)u\right| \ge \frac{C^\pi(i)}{2}|\mathcal{F}_{t_i}\right)\\ \le&\mathbb{P}\left(\sup_{t_i\le u<t_{i+1}}|W_{t_i+u}-W_{t_i}| \ge \frac{\sqrt{|\pi|\ln(\frac{1}{|\pi|})}}{2}\right) + \mathbb{P}\left(\left|\int_E\mu(de,(t_i,t_{i+1}])\right|>0\right),\end{aligned}$$

since $\left|b(S^\pi_{t_i})u + \int_E\beta(S^\pi_{\pi(t)},e)\lambda(de)u\right| \le \frac{C^\pi(i)}{2} = \frac{\alpha K\left(1+\left|X^\pi_{t_i}\right|\right)\sqrt{|\pi|\ln\left(\frac{1}{|\pi|}\right)}}{2}$ (with $|\pi| \le \sqrt{|\pi|\ln\left(\frac{1}{|\pi|}\right)}$, i.e. $|\pi|$ enough small) and $\left|\sigma(S^\pi_{t_i})\right| \le K\left(1+\left|X^\pi_{t_i}\right|\right)$. We have

$$\mathbb{P}\left(\left|\int_E\mu(de,(t_i,t_{i+1}])\right|>0\right) \le C(t_{i+1}-t_i).$$

Since $B_s = \frac{W_{t_i+s|\pi|}-W_{t_i}}{\sqrt{\pi}}$ is a standard Brownian motion by scaling and since $\mathbb{E}\left[e^{\beta \sup_{0\leq u<1}|B_s|^2}\right] < +\infty$ for $0 < \beta < \frac{1}{2}$, we deduce :

$$\begin{aligned}
\mathbb{P}\left(\sup_{t_i\leq u<t_{i+1}} |W_{t_i+u} - W_{t_i}| \geq \frac{\sqrt{|\pi|\ln(\frac{1}{\pi})}}{2}\right) &\leq \mathbb{P}\left(\sup_{0\leq u<1} |B_s| \geq \frac{\alpha\sqrt{\ln(\frac{1}{\pi})}}{2}\right) \\
&\leq \mathbb{E}\left[e^{\beta\left\{\sup_{0\leq u<1}|B_s|^2 - \left(\frac{\alpha\sqrt{\ln(\frac{1}{\pi})}}{2}\right)^2\right\}}\right] \\
&\leq \mathbb{E}\left[e^{\beta \sup_{0\leq u<1}|B_s|^2}\right] |\pi|^{\frac{\beta\alpha^2}{4}} \\
&\leq C\,|\pi|\,,
\end{aligned}$$

with $\alpha > \frac{2}{\sqrt{\beta}}$. Then $\mathbb{E}\left[1_{\theta_i^\pi < t_{i+1}}|\mathcal{F}_{t_i}\right] \leq C\,|\pi|$.

Combining the inequalities leads to

$$\begin{aligned}
\mathbb{E}\left[d_{\mathbb{D}}(S_\pi, S)^2\right] &\leq C\,|\pi|\ln\left(\frac{1}{\pi}\right) + C\sum_{i=0}^{n-1}(t_{i+1}-t_i)\mathbb{E}\left[\mathbb{E}\left[1_{\pi_{C_\pi}(i)<t_{i+1}}|\mathcal{F}_{t_i}\right]\left(1+\left|S_{t_i}^\pi\right|^2\right)\right] \\
&\leq C\,|\pi|\ln\left(\frac{1}{\pi}\right) + C\sum_{i=0}^{n-1}(t_{i+1}-t_i)\sup_{0\leq k\leq n}\mathbb{E}\left[\left(1+\left|S_{t_i}^\pi\right|^2\right)\right]|\pi| \\
&\leq C\,|\pi|\ln\left(\frac{1}{\pi}\right).
\end{aligned}$$

□

The lemma (4.5.5) determines the strong error for the approximation of X by X^π : $||X - X^\pi||^2_{\mathcal{S}^2_\pi} \leq C\,|\pi|\ln(\frac{1}{|\pi|})$.

Démonstration. (Proof of Lemma (4.5.5))

Inequality (4.5.2) and Proposition (4.5.2) imply

$$\begin{aligned}
\mathbb{E}\left[\sup_{t_i\in\pi}\left|X_{t_i} - X_{t_i}^\pi\right|^2\right] &\leq C\mathbb{E}\left[\sup_{t_i\in\pi}\left|S_{t_i} - S_{t_i}^\pi\right|^2\right] + C\mathbb{E}\left[\sup_{t_i\in\pi}|\Phi_{t_i}(S) - \Phi_{t_i}(S_\pi^\pi)|^2\right] \\
&\leq C\,|\pi| + C\mathbb{E}\left[\sup_{t_i\in\pi} d_{\mathbb{D}}\left(S^{t_i}, (S_\pi^\pi)^{t_i}\right)^2\right] + C\mathbb{E}\left[\sup_{t_i\in\pi}\left(\frac{1}{t_i}\int_0^{t_i}\left|S_s - S_{\pi(s)}^\pi\right| ds\right)^2\right] \\
&\leq C\,|\pi|\ln(\frac{1}{|\pi|}) + C\mathbb{E}\left[\sup_{t_i\in\pi}\frac{1}{t_i}\int_0^{t_i}\left|S_s - S_{\pi(s)}^\pi\right|^2 ds\right].
\end{aligned}$$

Then, with $C_{|\pi|} = C\,|\pi|\ln\frac{1}{|\pi|}$, we have

$$\begin{aligned}
\mathbb{E}\left[\sup_{t_i\in\pi}\frac{1}{t_i}\int_0^{t_i}\left|S_s - S_{\pi(s)}^\pi\right|^2 ds\right] &\leq C\mathbb{E}\left[\sup_{0\leq s\leq C_{|\pi|}}\left|S_s - S_{\pi(s)}^\pi\right|^2\right] + C\mathbb{E}\left[\sup_{t_i\in\pi, t_i\geq C_{|\pi|}}\frac{1}{t_i}\int_0^{t_i}\left|S_s - S_{\pi(s)}^\pi\right|^2 ds\right] \\
&\leq CC_{|\pi|} + C\sum_{t_i\in\pi, t_i\geq C_{|\pi|}}\frac{1}{t_{i+1}}\int_{t_i}^{t_{i+1}}\mathbb{E}\left[\left|S_s - S_{t_i}^\pi\right|^2\right] ds.
\end{aligned}$$

By $\mathbb{E}\left[\left|S_s - S^\pi_{t_i}\right|^2\right] \leq C\,|\pi|$ for $s \in [t_i, t_{i+1}]$, we obtain the inequality

$$\sum_{t_i\in\pi, t_i\geq C_{|\pi|}} \frac{1}{t_{i+1}} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[\left|S_s - S^\pi_{t_i}\right|^2\right] ds \leq C\,|\pi| \int_{C_{|\pi|}}^{T} \frac{ds}{s}$$
$$\leq C\,|\pi| \ln\left(\frac{1}{|\pi|}\right).$$

concluding the lemma. □

The lemma (4.5.7) shows that the flows of $(S_t, X_t)_{t\in[0,T]}$ and $\left(S^\pi_{t_i}, X^\pi_{t_i}\right)_{t_i\in\pi}$ are Lipchitz continuous.

Démonstration. We give the proof of the Lipschitz property of the flow of S and X (the reasonning is the same for S^π and X^π). We decompose the semi-martingale $S^{(t,s)}_{t+u} - S^{(t,s')}_{t+u} = M_u + A_u$ where M is a martingale and A a predictable finite variation process :

$$M_u = \int_0^u \left(\sigma(S^{(t,s)}_{t+r}) - \sigma(S^{(t,s')}_{t+r})\right) dW_{t+r} + \int_0^u \int_E \left(\beta(S^{(t,s)}_{t+r-}, e) - \beta(S^{(t,s')}_{t+r-}, e)\right) \overline{\mu}(de, dr),$$
$$A_u = \int_0^u \left(b(S^{(t,s)}_{t+r}) - b(S^{(t,s')}_{t+r})\right) dr.$$

Consider $\alpha(u) = \mathbb{E}\left[\sup_{0\leq r\leq u}\left|S^{(t,s)}_{t+r} - S^{(t,s')}_{t+r}\right|^2\right]$, we have

$$\alpha(u) = \mathbb{E}\left[\sup_{0\leq r\leq u} |A_u + M_u|^2\right]$$
$$\leq 2\mathbb{E}\left[\sup_{0\leq u\leq t} |A_u|^2\right] + 2\mathbb{E}\left[\sup_{0\leq u\leq t} |M_u|^2\right].$$

We deduce from Burkholder-Davis-Gundy inequality and Lipschitz property of b, σ and $\beta(., e)$ that

$$\mathbb{E}\left[\sup_{0\leq r\leq u} |M_r|^2\right] \leq D_2\mathbb{E}\left[[M]_u\right]$$
$$\leq D_2\mathbb{E}\left[\int_0^u \left|\sigma(S^{(t,s)}_{t+r}) - \sigma(S^{(t,s')}_{t+r})\right|^2 dr\right]$$
$$+ D_2\mathbb{E}\left[\int_0^u \int_E \left|\beta(S^{(t,s')}_{t+r-}, e) - \beta(S^{(t,s')}_{t+r-}, e)\right|^2 \lambda(de)dr\right]$$
$$\leq C\int_0^u \alpha(r)dr,$$
$$\mathbb{E}\left[\sup_{0\leq u\leq t} |A_u|^2\right] \leq C\mathbb{E}\left[\int_0^u \left|b(S^{(t,s)}_{t+r}) - b(S^{(t,s)}_{t+r})\right|^2 dr\right]$$
$$\leq C\int_0^u \alpha(r)dr.$$

Gronwall lemma applied to $\alpha(u) \leq C \int_0^u \alpha(r)dr$ gives :

$$\begin{aligned}\mathbb{E}\left[\sup_{t\leq u\leq T}\left|S_u^{(t,s)} - S_u^{(t,s')}\right|^2\right] &= \alpha(T-t) \\ &\leq C\alpha(0) = C\left|s-s'\right|^2.\end{aligned}$$

Since φ is uniformly $\mathbb{L}^\infty$-Lipschitz, we have $X_u^{(t,s)} = \left(S_u^{(t,s)}, \varphi_{t,u}(p, S^{(t,s)}\right)$ and

$$\begin{aligned}\mathbb{E}\left[\sup_{t\leq u\leq T}\left|X_u^{(t,(s,p))} - X_u^{(t,(s',p'))}\right|^2\right] &\leq \mathbb{E}\left[\sup_{t\leq u\leq T}\left|\left(S_u^{(t,s)}, \varphi_{t,u}(p, S^{(t,s)}\right) - \left(S_u^{(t,s)}, \varphi_{t,u}(p', S^{(t,s')}\right)\right|^2\right] \\ &\leq C\left|p-p'\right|^2 + C\mathbb{E}\left[\sup_{t\leq u\leq T}\left|S_u^{(t,s)} - S_u^{(t,s')}\right|^2\right] \\ &\leq C\left|p-p'\right|^2 + C\left|s-s'\right|^2.\end{aligned}$$

□

Quatrième partie

Réplication robuste de dérivés sur la variance réalisée

4.7. Introduction

L'étude de la volatilité est devenue centrale dans la finance quantitative. En 1973, Black et Scholes ([**BS73**]) ont introduit leur modèle lognormal où le prix d'une action possède une volatilité instantanée constante. Cependant, ce modèle est en désaccord avec les prix du marché. Ainsi, les traders qui utilisent le modèle de Black et Scholes pour la couverture doivent de façon permanente modifier les paramètres de volatilité pour obtenir les prix de marché. Or, l'introduction des modèles à volatilité stochastique permet de rester compatible avec le fait que des options de strikes et de maturités différents ont des volatilités implicites différentes.

Les marchés financiers traitent maintenant de nombreux sous-jacents autres que les actions et des indices boursiers. En particulier, ces dernières années, de nombreux contrats de swap sur la variance réalisée sont échangés de gré-à-gré et sont devenus liquides. Des dérivés dont les payoff sont des fonctions de la variance réalisée (et plus généralement aussi fonctions de l'actif risqué) sont aussi négociés de gré-à-gré, tel le swap de volatilité qui est la racine de la variance réalisée. Notre étude va se concentrer sur la valorisation et la couverture de ces options sur la variance réalisée. Habituellement, cela nécessite de choisir un modèle sur la volatilité (car cette dernière n'est pas directement observable) et de définir une méthode de calibration.

Pour éviter cette problématique, nous proposons dans cet article de donner une valorisation et une couverture des produits de payoff mixant actif risqué et variance réalisée qui est robuste au modèle de volatilité, c'est-à-dire, si possible, indépendante du modèle de volatilité. Dans une première partie, nous donnons la méthode de réplication des dérivés de la variance et du spot, par une couverture en traitant dynamiquement des options européennes. Il s'agit d'un problème inverse : il faut trouver une option européenne possédant les mêmes prix et les mêmes deltas donnés pour tout niveau de volatilité dans le modèle de Black-Scholes. On obtient alors une robustesse par rapport à la corrélation entre l'actif risqué et sa volatilité (avec une erreur du deuxième ordre en la corrélation). Il s'agit d'une extension de la méthode de réplication donnée dans l'article ([**CL08**]). On obtient de plus une formule fermée pour l'inversion. Ensuite, nous proposons différentes méthodes pour résoudre cette inversion. Nous illustrons ensuite cette couverture avec le swap de volatilité et le ratio de Sharpe.

Puis, nous étendons cette méthologie cette couverture pour les options sur fonds en y ajoutant un effet d'asymétrie sur les diffusions et en y intégrant des bruits orthogonaux. On peut alors couvrir de manière robuste des options sur fonds (fond benchmarké, fond à volatilité benchmarké), sur des stratégies tel le constant mix, le CPPI. En particulier, nous obtenons une couverture simple (en traitant un call et un put) pour la réplication d'un call sur fond. Ensuite, nous rajoutons des sauts dans notre modèle : on travaille alors avec un processus de Lévy avec un changement de temps et la robustesse de la couverture concerne alors le subordinateur au lieu de la variance réalisée (dans le modèle continu). Finalement, nous donnons les différentes preuves dans la dernière partie.

4.8. Problème inverse

Considérons une maturité $T > 0$ et un espace filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$. On se place dans le cas où les taux sont nuls par souci de simplication. On suppose que

l'actif risqué S suit l'équation de diffusion :

$$\frac{dS_t}{S_t} = \sigma_t dW_t$$

où W est un $\mathcal{F}_t$-mouvement Brownien, σ_t un processus $\mathcal{F}_t$-adapté et $\mathbb{P}$ est la probabilité risque-neutre. On note $\mathbb{E}_t$ l'espérance conditionnelle par rapport à $\mathcal{F}_t$.

Pour $0 \leq s \leq t \leq T$, on note le processus logarithme

$$X_{s,t} = \ln\left(\frac{S_t}{S_s}\right)$$

et la variance réalisée

$$\langle X \rangle_{s,t} = \int_s^t \sigma_u^2 du.$$

Par souci de simplification, on note $X_t := X_{0,t}$ et $\langle X \rangle_t = \langle X \rangle_{0,t}$. On suppose que σ_t est indépendante de W.

On cherche à couvrir une option de payoff

$$\xi_T = H(X_T, \langle X \rangle_T) \in \mathbb{L}^2$$

de maturité T.

REMARK 4.8.1. Souvent, l'option à répliquer est $H\left(X_T, V_T^n\right)$ où V_T^n est un proxi de la variance réalisée :

$$V_T^n := \sum_{0<t_n\leq T} \left(\log(\frac{S_{t_n}}{S_{t_{n-1}}})\right)^2 = \sum_{0<t_n\leq T} \left(\log(\frac{S_{t_n}}{S_{t_{n-1}}})\right)^2$$

C'est le cas par exemple dans le contrat de variance swap. En fixant T, V_T^n converge vers $\langle X \rangle_T$ quand n tend vers l'infini (avec un pas temps $\sup_i(t_{i+1} - t_i)$ qui tend vers 0). On approximera alors V_T^n par $\langle X \rangle_T$.

La formule de Clark-Ocone donne la réplication de l'option ξ_T :

$$\begin{aligned} H\left(X_T, \langle X \rangle_T\right) &= \mathbb{E}\left[H\left(X_T, \langle X \rangle_T\right) | \mathcal{F}_t, \langle X \rangle_{t,T}\right] + \int_t^T \mathbb{E}_s\left[D_s H\left(X_T, \langle X \rangle_T\right)\right] dW_s \\ &= h_t(\langle X \rangle_{t,T}) + \int_t^T \delta_s\left(\langle X \rangle_{s,T}\right) \sigma_s dW_s, \end{aligned}$$

où D_s désigne la dérivée de Malliavin (voir ([**Nua95**])) et avec $\forall v \geq 0$

$$\begin{aligned} h_t(v) &= \mathbb{E}\left[H\left(X_T, \langle X \rangle_T\right) | \mathcal{F}_t, \langle X \rangle_{t,T} = v\right], \\ \delta_t(v) &= \mathbb{E}\left[\frac{\partial}{\partial x} H\left(X_T, \langle X \rangle_T\right) | \mathcal{F}_t, \langle X \rangle_{t,T} = v\right]. \end{aligned}$$

On parlera de couverture robuste lorsque la réplication fonctionne pour toute variance réalisée $\langle X \rangle_{t,T}$ indépendante de W.
La couverture robuste dynamique à la date t par une option européenne de payoff $P_t(X_{t,T})$ de maturité T doit vérifier

$$\begin{aligned} \mathbb{E}\left[P_t\left(X_{t,T}\right) | \mathcal{F}_t, \langle X \rangle_{t,T} = v\right] &= h_t(v), \\ \frac{\partial}{\partial x}\mathbb{E}\left[P_t\left(X_{t,T}\right) | \mathcal{F}_t, \langle X \rangle_{t,T} = v\right] &= \delta_t(v). \end{aligned} \tag{4.8.1}$$

On obtient alors la réplication sous l'hypothèse d'indépendance entre

$$H\left(X_T, \langle X\rangle_T\right) = P_t\left(X_{t,T}\right) + \int_t^T \frac{\partial P_s\left(X_{s,T}\right)}{\partial x}\sigma_s dW_s.$$

EXAMPLE 4.8.2. Lorsque la volatilité devient déterministe, la réplication de l'option $H\left(X_T, \langle X\rangle_T\right)$ ne nécéssite que l'actif risqué comme instrument de couverture :

$$P_t(X_{t,T}) = h_t + \delta_t\left(\frac{S_T}{S_t} - 1\right).$$

Par exemple, dans le cadre de Black-Scholes (avec une volatilité $\sigma_s = \sigma$ constante), en appliquant le lemme d'Itô à $u(t, X_t) = \mathbb{E}\left[H\left(X_T, \langle X\rangle_T\right)|\mathcal{F}_t\right]$

$$\begin{aligned} H\left(X_T, \langle X\rangle_T\right) &= u(T, X_T) \\ &= u(t, X_t) + \int_t^T \frac{\partial}{\partial x}u(s, X_s)dX_s + \int_t^T (\frac{\partial u}{\partial s}ds + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}d\langle X\rangle_s) \\ &= u(t, X_t) + \int_t^T \frac{\partial}{\partial x}u(s, X_s)\sigma dW_s \\ &= u(t, X_t) + \int_t^T \frac{\partial}{\partial x}u(s, X_s)\frac{dS_s}{S_s} \end{aligned}$$

on obtient la réplication de l'option : à la date t, le prix de l'option est $u(t, X_t) = h_t(\sigma^2(T-t)) = \mathbb{E}\left[H\left(X_T, \langle X\rangle_T\right)|\mathcal{F}_t\right]$ et il suffit d'annuler le terme en delta $\frac{\partial}{\partial x}u(s, X_s) = \delta_t(\sigma^2(T-t))$ avec position dynamique en quantité $\frac{\partial}{\partial x}u(t, X_t)/S_t$ sur l'actif risqué S_t.

REMARK 4.8.3. Cette réplication robuste suppose que le marché soit sans friction (par exemple sans coût de transaction) et que toutes les options européennes de maturité T existent et sont liquides. Par exemple, il suffit de supposer que dans le marché, pour une maturité T, tous les calls et puts sont négociables, alors l'option européenne $P_t(X_{t,T})$ est valorisable et réplicable à l'aide de la formule de Carr-Madan (**[CM02]**) par une combinaison linéaire de calls et de puts :

$$\begin{aligned} P_t(X_{t,T}) &= P_t(\ln(\frac{S_T}{S_t})) \\ &= P_t(0) + \frac{P_t'(0)}{S_t}(S_T - S_t) + \int_0^{S_t}(K - S_T)^+\frac{P"(\ln(\frac{K}{S_t})) - P'(\ln(\frac{K}{S_t}))}{K^2}dK \\ &+ \int_{S_t}^{+\infty}(S_T - K)^+\frac{P"(\ln(\frac{K}{S_t})) - P'(\ln(\frac{K}{S_t}))}{K^2}dK. \end{aligned}$$

4.8.1. Solution. Pour obtenir une solution du problème (4.8.1), nous définissons une fonction P de (h_t, δ_t) à l'aide de la transformée de Laplace.

DEFINITION 4.8.4. Définissons (au sens des distributions) les fonctions $P(h_t,0)$, $P(0,\delta_t)$ et $P(h_t,\delta_t)$ par

$$\begin{aligned}
P(h_t,0)(x) &= \sqrt{\pi}e^{\frac{x}{2}}\,|x|\,\mathcal{L}^{-1}\left(\frac{e^{\frac{1}{16\lambda}}}{\lambda^{\frac{1}{2}}}h_t\left(\frac{1}{2\lambda}\right)\right)(x^2)\\
&\quad -\frac{\sqrt{\pi}}{4}e^{\frac{x}{2}}\mathrm{sign}(x)\mathcal{L}^{-1}\left(\frac{e^{\frac{1}{16\lambda}}}{\lambda^{\frac{3}{2}}}h_t\left(\frac{1}{2\lambda}\right)\right)(x^2),\\
P(0,\delta_t)(x) &= \frac{\sqrt{\pi}}{2}e^{\frac{x}{2}}\mathrm{sign}(x)\mathcal{L}^{-1}\left(\frac{e^{\frac{1}{16\lambda}}}{\lambda^{\frac{3}{2}}}\delta_t\left(\frac{1}{2\lambda}\right)\right)(x^2),\\
P(h_t,\delta_t) &= P(h_t,0)+P(0,\delta_t).
\end{aligned}$$

où $\mathcal{L}$ est la transformée de Laplace :

$$\mathcal{L}f(\lambda)=\int_{\mathbb{R}^+}f(s)e^{-\lambda s}ds.$$

Nous devons maintenant nous placer dans un espace adapté aux fonctions $P(h_t,0)$, $P(0,\delta_t)$ et $P(h_t,\delta_t)$.

DEFINITION 4.8.5. On note $\mathbb{H}_t(\mathbb{R})$ l'ensemble des fonctions $g_t:\mathbb{R}\to\mathbb{R}$ $\mathcal{F}_t$-mesurables telles que

$$||g_t||_{\mathbb{H}_t}=\mathbb{E}\left[|g_t(X_{t,T})|+|g_t(X_{t,T})X_{t,T}|\,|\mathcal{F}_t,\langle X\rangle_{t,T}=v\right]<+\infty\quad\forall v>0.$$

La proposition suivante donne la résolution du problème d'inversion (4.8.1).

PROPOSITION 4.8.6. *A la date t, l'option européenne P_t vérifiant (4.8.1) est unique dans $\mathbb{H}_t(\mathbb{R})$ et vaut $P_t=P(h_t,\delta_t)$. Si $P_t\in\mathbb{H}_t(\mathbb{R})$ vérifiant (4.8.1), alors $P_t=P(h_t,\delta_t)$ avec $P(h_t,0),P(0,\delta_t)\in\mathbb{H}_t(\mathbb{R})$. Réciproquement si $P(h_t,0),P(0,\delta_t)\in\mathbb{H}_t(\mathbb{R})$, alors $P(h_t,\delta_t)$ vérifie (4.8.1).*

Nous avons décomposé le payoff $P_t=P(h_t,\delta_t)=P(h_t,0)+P(0,\delta_t)$ en deux options.

La partie $P(h_t,0)(X_{t,T})$ donne le même prix et est delta-neutre :

$$\mathbb{E}\left[P(h_t,0)(X_{t,T})|\mathcal{F}_t,\langle X\rangle_{t,T}=v\right]=h_t(v)\quad\text{et}\quad\mathbb{E}\left[P(0,\delta_t)(X_{t,T})|\mathcal{F}_t,\langle X\rangle_{t,T}=v\right]=0$$

et $P(0,\delta_t)(X_{t,T})$ a un prix nul et le même delta que $H\left(X_T,\langle X\rangle_T\right)$

$$\mathbb{E}\left[P'(h_t,0)(X_{t,T})|\mathcal{F}_t,\langle X\rangle_{t,T}=v\right]=0\quad\text{et}\quad\mathbb{E}\left[P'(0,\delta_t)(X_{t,T})|\mathcal{F}_t,\langle X\rangle_{t,T}=v\right]=\delta_t(v).$$

4.8.2. Impact de la corrélation. Nous supprimons ici maintenant l'hypothèse d'indépendance entre l'actif risqué et la volatilité. Nous regardons l'impact de la corrélation sur notre couverture robuste. On considère

$$dS_t^{(\rho)}=\sqrt{1-\rho^2}\sigma_tS_t^{(\rho)}dW_t^1+\rho\sigma_tS_t^{(\rho)}dW_t^2$$

où $|\rho|\le 1$, W^1 et W^2 sont des mouvements Brownien $\mathcal{F}_t$-adaptés et σ et W^2 sont adaptés à une filtration $\mathcal{H}_t\subseteq\mathcal{F}_t$ avec $\mathcal{H}_t$ et $\mathcal{F}_t^{W^1}$ indépendants. On a

$$\begin{aligned}
h_t(v) &= \mathbb{E}\left[H\left(\ln(S_t^{(0)})+X_{t,T},\langle X\rangle_T\right)|\mathcal{F}_t,\langle X\rangle_{t,T}=v\right],\\
\delta_t(v) &= \mathbb{E}\left[\frac{\partial}{\partial x}H\left(\ln(S_t^{(0)})+X_T,\langle X\rangle_T\right)|\mathcal{F}_t,\langle X\rangle_{t,T}=v\right].
\end{aligned}$$

La proposition suivante nous donne une estimation de l'erreur induite par la corrélation.

PROPOSITION 4.8.7. *Cette couverture est robuste à l'ordre 2 en la corrélation volatilité-actif risqué :*

$$\mathbb{E}_t\left[H\left(\ln S_T^{(\rho)}, \langle X\rangle_T\right)\right] = \mathbb{E}_t\left[P_t\left(\ln(\frac{S_T^{(\rho)}}{S_t^{(\rho)}})\right)\right] + O(\rho^2).$$

Dans la suite, on parlera de couverture ρ-neutre pour la couverture dynamique avec des options européennes vérifiant (4.8.1).

4.9. Différentes méthodes d'inversion

Trouver la fonction P_t revient à résoudre une inversion de Laplace. Dans cette partie, nous allons proposer plusieurs méthodes pour obtenir le payoff P_t.

En remarquant que l'inversion est linéaire en h_t et δ_t, l'idée pricinpale est de décomposer h_t et δ_t à l'aide de fonction de bases inversibles pour obtenir P_t.

4.9.1. A l'aide de table de transformations de Laplace. On utilise la table de transformée inverse de Laplace et des propriétés de la transformations de Laplace comme la convolution $\mathcal{L}(f*g) = \mathcal{L}f\mathcal{L}g$ ou la translation $\mathcal{L}(f(t)^{-at}) = \mathcal{L}f(.+a)$. Si la fonction est assez simple (par exemple pour le swap de volatilité, le ratio de Sharpe), cette méthode est très pratique et permet d'obtenir une formule fermée. Le lemme suivant nous incite à utiliser $v \in \mathbb{R}_+^* \to v^\alpha e^{\beta v - \frac{\theta}{v}}$ comme fonctions de base pour obtenir l'inversion.

LEMMA 4.9.1. *Soit $\alpha > -\frac{1}{2}$ et $\alpha' > -\frac{3}{2}$. On a les inversions suivantes :*

$$P(v^\alpha e^{\beta v - \frac{\theta}{v}}, 0)(x) = \sqrt{\frac{\pi}{2}} e^{\frac{x}{2}} |x| \left(\frac{x^2 - 2\theta}{2\beta + \frac{1}{4}}\right)^{\frac{\alpha - \frac{1}{2}}{2}} I_{\alpha - \frac{1}{2}}\left(\sqrt{(2\beta + \frac{1}{4})(x^2 - 2\theta)}\right) 1_{x^2 \geq 2\theta}$$
$$- \frac{\sqrt{\pi}}{2\sqrt{2}} e^{\frac{x}{2}} sign(x) \left(\frac{x^2 - 2\theta}{2\beta + \frac{1}{4}}\right)^{\frac{\alpha + \frac{1}{2}}{2}} I_{\alpha + \frac{1}{2}}\left(\sqrt{(2\beta + \frac{1}{4})(x^2 - 2\theta)}\right) 1_{x^2 \geq 2\theta},$$

$$P(0, v^{\alpha'} e^{\beta v - \frac{\theta}{v}})(x) = \sqrt{\frac{\pi}{2}} e^{\frac{x}{2}} sign(x) \left(\frac{x^2 - 2\theta}{2\beta + \frac{1}{4}}\right)^{\frac{\alpha' + \frac{1}{2}}{2}} I_{\alpha' + \frac{1}{2}}\left(\sqrt{(2\beta + \frac{1}{4})(x^2 - 2\theta)}\right) 1_{x^2 \geq 2\theta},$$

où I_ν est la fonction de Bessel modifiée de première espèce. Avec $\beta = -\frac{1}{8}$ et $\theta = 0$, on a

$$P(v^\alpha e^{-\frac{v}{8}}, 0)(x) = \frac{\sqrt{\pi}}{2^\alpha} e^{\frac{x}{2}} |x|^{2\alpha} \left(\frac{1}{\Gamma(\alpha - \frac{1}{2})} - \frac{x}{4\Gamma(\alpha + \frac{1}{2})}\right),$$
$$P(0, v^{\alpha'} e^{-\frac{v}{8}})(x) = \frac{\sqrt{\pi}}{2^{\alpha+1}\Gamma(\alpha + \frac{1}{2})} e^{\frac{x}{2}} x |x|^{2\alpha'},$$

où Γ désigne la fonction gamma.

4.9.2. Inversion de la transformée de Laplace. Soit $F(p) = \mathcal{L}f(p)$. Alors

$$f(x) = \mathcal{L}^{-1}F(x) = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{pt} F(p) dp$$

où γ est choisi pour que l'intégrale soit convergente, ce qui implique que γ soit supérieur à la partie réelle de toute singularité de $F(p)$.

4.9.3. Méthode numérique : algorithme de Stehfest. La méthode précédente nécéssite de connaître h_t et δ_t dans le plan complexe, ce qui n'est généralement pas pratique.

Pour inverser la transformation de Laplace, on peut utiliser l'algorithme de Stehfest (voir ([**Ste70**])). Soit $F(p) = \mathcal{L}f(p)$. Pour $x \geq 0$, on approxime $f(x)$ par

$$f_N(x) = \frac{\ln(2)}{x} \sum_{j=1}^{N} Q_j F\left(j \frac{\ln(2)}{x}\right)$$

où les coefficients Q_j sont donnés par

$$Q_j = (-1)^{\frac{N}{2}+j} \sum_{k=\frac{j+1}{2}}^{\min(j,\frac{N}{2})} \frac{k^{\frac{N}{2}}\,(2k)!}{(\frac{N}{2}-k)!k!(k-1)!(j-k)!(2k-1)!}.$$

On doit prendre N pair (il est conseillé de le prendre entre 10 et 30). Ici, l'algorithme a l'avantage d'être simple, très rapide et ne demande de connaître la fonction F que sur $\mathbb{R}^+$, ce qui semble plus naturel sachant que l'on travaille avec la variance $\langle X\rangle_{t,T} \in \mathbb{R}^+$.

4.9.4. Méthode de Carr et Lee. Carr et Lee ([**CL08**]) travaillent des fonctions exponentielles en utilisant l'inégalité suivante :

$$\mathbb{E}_t\left[e^{\lambda\left(\langle X\rangle_T-\langle X\rangle_t\right)+\mu(X_T-X_t)}\right] = \mathbb{E}_t\left[e^{\frac{p(\mu,\lambda)^2}{2}\langle X\rangle_{t,T}}\right] = \mathbb{E}_t\left[e^{\left(\frac{1}{2}\pm p(\mu,\lambda)\right)(X_T-X_t)}\right]$$

avec $p(\mu,\lambda) = \sqrt{\frac{1}{4}+2\lambda+\mu^2}-\mu$. En appliquant le lemme (4.9.1), on retrouve leur méthode de couverture ρ-neutre.

PROPOSITION 4.9.2. *Pour la couverture de l'option* $H(X_T,\langle X\rangle_T) = e^{\lambda\langle X\rangle_T+\mu X_T}$, *on a* $h_t(v) = \delta_t(v)/\mu = e^{\lambda\langle X\rangle_t+\mu X_t}e^{\frac{p(\mu,\lambda)^2-p(\mu,\lambda)}{2}v}$ *et*

$$P(h_t,0)(x) = e^{\lambda\langle X\rangle_t+\mu X_t}\left\{e^{\frac{x}{2}}\cosh\left(xp(\mu,\lambda)\right) - \frac{e^{\frac{x}{2}}\sinh\left(xp(\mu,\lambda)\right)}{2p(\mu,\lambda)}\right\},$$

$$P(0,\delta_t)(x) = e^{\lambda\langle X\rangle_t+\mu X_t}\mu e^{\frac{x}{2}}\frac{\sinh\left(xp(\mu,\lambda)\right)}{p(\mu,\lambda)}$$

$$\begin{aligned} P_t(x) &= e^{\lambda\langle X\rangle_t+\mu X_t}e^{\frac{x}{2}}\left\{\cosh\left(xp(\mu,\lambda)\right)+\frac{\mu-\frac{1}{2}}{p(\mu,\lambda)}\sinh\left(xp(\mu,\lambda)\right)\right\} \\ &= e^{\lambda\langle X\rangle_t+\mu X_t}\left\{\left(\frac{1}{2}+\theta(\mu,\lambda)\right)e^{\left(\frac{1}{2}+p(\mu,\lambda)\right)x} + (\frac{1}{2}-\theta(\mu,\lambda))e^{\left(\frac{1}{2}-p(\mu,\lambda)\right)x}\right\} \end{aligned}$$

avec $\theta(\mu,\lambda) = \frac{\mu-\frac{1}{2}}{2p(\mu,\lambda)}$.

Soit une option de payoff (sur la variance et le spot)

$$f(\langle X\rangle_T, X_T) = \int F(\mu,\lambda)e^{\mu X_T+\lambda\langle X\rangle_T}m(d\mu,d\lambda)$$

où m est une mesure qui peut être obtenue par exemple à l'aide de la transformée de Fourier. Sous de bonnes conditions d'intégrabilité sur la mesure m (pour pouvoir appliquer Fubini), cette option peut être couverte avec une position dynamique à la date t par

$$\begin{aligned} P_t(X_{t,T}) = \int F(\mu,\lambda)e^{\mu X_t+\lambda\langle X\rangle_t}\Bigg(&\left(\frac{1}{2}+\theta(\mu,\lambda)\right)e^{\left(\frac{1}{2}+p(\mu,\lambda)\right)X_{t,T}} \\ &+(\frac{1}{2}-\theta(\mu,\lambda))e^{\left(\frac{1}{2}-p(\mu,\lambda)\right)X_{t,T}}\Bigg)m(d\mu,d\lambda). \end{aligned}$$

4.9.5. Approximation du payoff P_t. P_t peut être mal définie comme fonction (par exemple s'il s'agit d'une distribution, P_t peut contenir une masse de Dirac...) ou bien l'inversion peut être compliquée. On propose dans cette section de donner une estimation de P_t en inversant une approximation de (h_t, δ_t) avec des éléments de la forme $v \to v^j e^{-\frac{v}{8}}$. On suppose qu'il existe une constante $M > 0$ telle que $\langle X \rangle_{t,T} \leq M$ et que (h_t, δ_t) est continue sur $[0, M]$.

4.9.5.1. *Approximation uniforme.* La proposition suivante nous donne l'estimation du payoff de couverture dynamique par une approximation uniforme à l'aide des polynômes de Berstein.

PROPOSITION 4.9.3. *Soit*

$$b_{n,k}^{h_t} = \frac{1}{M^k} \sum_{j=0}^{k} e^{\frac{jM}{8k}} h_t(\frac{jM}{n}) \binom{n}{k} \binom{k}{j} (-1)^{k-j},$$

$$b_{n,k}^{\delta_t} = \frac{1}{M^k} \sum_{j=0}^{k} e^{\frac{jM}{8k}} (\frac{jM}{kn}) \binom{n}{k} \binom{k}{j} (-1)^{k-j}.$$

En considérant

$$P_{n,t}(x) = \sqrt{\pi} e^{\frac{x}{2}} \sum_{k=0}^{n} \left(b_{n,k}^{h_t} \frac{x^{2k}}{2^k} \{ \frac{1}{\Gamma(k - \frac{1}{2})} - \frac{x}{4\Gamma(k + \frac{1}{2})} \} + b_{n,k}^{\delta_t} \frac{x^{2k+1}}{2^{k+1}\Gamma(k + \frac{1}{2})} \right)$$

on a

$$\lim_{n \to +\infty} \mathbb{E} \left[P_{n,t}(X_{t,T}) | \mathcal{F}_t, \langle X \rangle_{t,T} \right] = h_t(\langle X \rangle_{t,T}),$$

$$\lim_{n \to +\infty} \frac{\partial}{\partial x} \mathbb{E} \left[P_{n,t}(X_{t,T}) | \mathcal{F}_t, \langle X \rangle_{t,T} \right] = \delta_t(\langle X \rangle_{t,T}).$$

4.9.5.2. *Approximation par régression linéaire.* La proposition suivante nous donne l'estimation du payoff de couverture dynamique par une approximation par une régression linéaire.

PROPOSITION 4.9.4. *Considérons μ une mesure sur $[0, M]$ et Q la distribution de $\langle X \rangle_{t,T}$. On suppose que Q est absolument continue par rapport à μ et que $\frac{dQ}{d\mu} \in \mathbb{L}^2(\mu)$. Considérons $\sum_{k=0}^{n} a_{n,k}^{h_t} v^k$ et $\sum_{k=0}^{n} a_{n,k}^{\delta_t} v^k$ les $\mathbb{L}^2(\mu)$- projections de $v \to e^{\frac{v}{8}} h_t(v)$ et $v \to e^{\frac{v}{8}} \delta_t(v)$ sur l'espace $\{1, v, ..., v^n\}$. En considérant*

$$P_{n,t}(x) = \sqrt{\pi} e^{\frac{x}{2}} \sum_{k=0}^{n} \left(a_{n,k}^{h_t} \frac{x^{2k}}{2^k} \{ \frac{1}{\Gamma(k - \frac{1}{2})} - \frac{x}{4\Gamma(k + \frac{1}{2})} \} + a_{n,k}^{\delta_t} \frac{x^{2k+1}}{2^{k+1}\Gamma(k + \frac{1}{2})} \right)$$

on a

$$\lim_{n \to +\infty} \mathbb{E}_t \left[P_{n,t}(X_{t,T}) \right] = \mathbb{E}_t \left[h_t(\langle X \rangle_{t,T}) \right],$$

$$\lim_{n \to +\infty} \frac{\partial}{\partial x} \mathbb{E}_t \left[P_{n,t}(X_{t,T}) \right] = \mathbb{E}_t \left[\delta_t(\langle X \rangle_{t,T}) \right].$$

4.10. Applications pour la couverture d'options sur la variance et le spot

Dans cette section, on donne quelques exemples de couverture robuste sur des produits dérivés sur la variance et l'actif sous-jacent.

4.10.1. Cas puissance $\langle X\rangle_T^\alpha$. On cherche à couvrir une option de payoff $H\left(X_T, \langle X\rangle_T\right) = \langle X\rangle_T^\alpha$ avec $\alpha > -\frac{1}{2}$. Alors

$$\begin{aligned} h_0(v) &= v^\alpha, \\ \delta_0(v) &= 0. \end{aligned}$$

La formule obtenue par le lemme (4.9.1) pour $\langle X\rangle_{t,T}^\alpha e^{\beta\langle X\rangle_{t,T} - \frac{\theta}{\langle X\rangle_{t,T}}}$ avec $\beta = \theta = 0$ donne

$$P_0(x) = e^{\frac{x}{2}}\sqrt{\pi}2^{\alpha-1}\left|x\right|^{\alpha+\frac{1}{2}}\left\{I_{\alpha-\frac{1}{2}}\left(\frac{|x|}{2}\right) - \operatorname{sign}(x) I_{\alpha+\frac{1}{2}}\left(\frac{|x|}{2}\right)\right\}$$

où I_ν est la fonction de Bessel modifiée de première espèce.

4.10.1.1. *Variance swap.* Pour le swap de variance $\langle X\rangle_{t,T}$ ($\alpha = 1$), on a

$$\begin{aligned} P_t(x) &= \sqrt{\pi}e^{\frac{x}{2}}\left|x\right|^{\frac{3}{2}}\left\{I_{\frac{1}{2}}\left(\frac{|x|}{2}\right) - \operatorname{sign}(x) I_{\frac{3}{2}}\left(\frac{|x|}{2}\right)\right\} \\ &= -2x + 2e^x - 2. \end{aligned}$$

Le payoff de couverture robuste pour le variance swap est

$$P_t(X_{t,T}) = -2X_{t,T} + 2e^{X_{t,T}} - 2 = -2\log(\frac{S_T}{S_t}) + 2\frac{S_T}{S_t} - 2.$$

On retrouve ainsi la réplication par le log contrat, proposée et établie dans ([**Neu94**]), ([**KDZ99**]) et ([**CM02**]) qui ne nécessite cependant pas l'hypothèse d'indépendance de (W_t) et de (σ_t).

4.10.1.2. *Volatility swap.* La volatilité réalisée $\sqrt{\langle X\rangle_T}$ est la racine carrée de la variance réalisée $\langle X\rangle_T$. L'inégalité de Jensen montre que l'espérance de la racine carrée d'une variable est plus faible que la racine carrée de l'espérance. La différence entre $\sqrt{\langle X\rangle_T}$ et $\langle X\rangle_T$ peut être vue comme un ajustement de convexité. Le swap de volatilité admet une surréplication modèle indépendant. En effet, le swap de volatilité peut être surcouvert avec de la variance réalisée qui elle-même admet une couverture modèle indépendant.

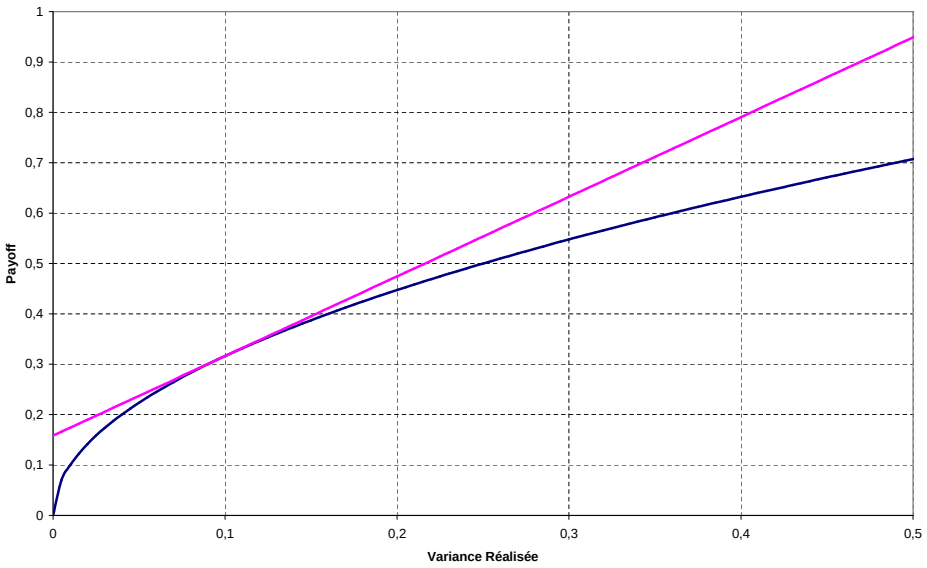


FIG. 4.10.1. Le volatility swap admet une surréplication modèle indépendant

Cependant, la valorisation du swap de volatilité est généralement considérée comme très sensible au modèle sur la volatilité instantanée. **[CL08]** et **[FG05]** donnent la valorisation du swap de volatilité à l'aide d'un problème inverse.
Nous allons résoudre ce problème en donnant une valorisation et une couverture modèle indépendant du swap de volatilité avec notre méthodologie (qui annule l'erreur du premier ordre sur la corrélation).

PROPOSITION 4.10.1. *Pour le volatility swap* $\sqrt{\langle X \rangle_T}$*, le payoff de réplication robuste vaut à la date* $t = 0$ *(*$\alpha = \frac{1}{2}$*)*

$$P_0(x) = \sqrt{\frac{\pi}{2}} e^{\frac{x}{2}} |x| \left(I_0\left(\frac{x}{2}\right) - I_1\left(\frac{x}{2}\right)\right)$$

et à la date t

$$P_t(x) = \sqrt{\frac{\pi}{2}} |x| e^{\frac{x}{2}} \left(I_0\left(\frac{x}{2}\right) - I_1\left(\frac{x}{2}\right)\right) \left\{ 1 + \frac{\sqrt{\langle X \rangle_t}}{\sqrt{\pi}|x|} e^{-\frac{x^2}{\langle X \rangle_t}} - erfc\left(\frac{|x|}{\sqrt{\langle X \rangle_t}} \right) \right\}$$
$$- \sqrt{\langle X \rangle_t} \frac{e^{\frac{x}{2}}}{\sqrt{2}} \int_0^1 e^{\frac{x^2}{\langle X \rangle_t}(u^2-1)} \left\{ \left(I_0\left(\frac{x}{2}\right) - I_0\left(\frac{ux}{2}\right)\right) - \left(I_1\left(\frac{x}{2}\right) - I_1\left(\frac{ux}{2}\right) u\right) \right\} \frac{udu}{(1-u^2)^{\frac{3}{2}}}$$

où $erfc$ *désigne la fonction d'erreur complémentaire* $erfc(a) = \frac{2}{\sqrt{\pi}} \int_a^\infty e^{-\frac{y^2}{2}} dy$.

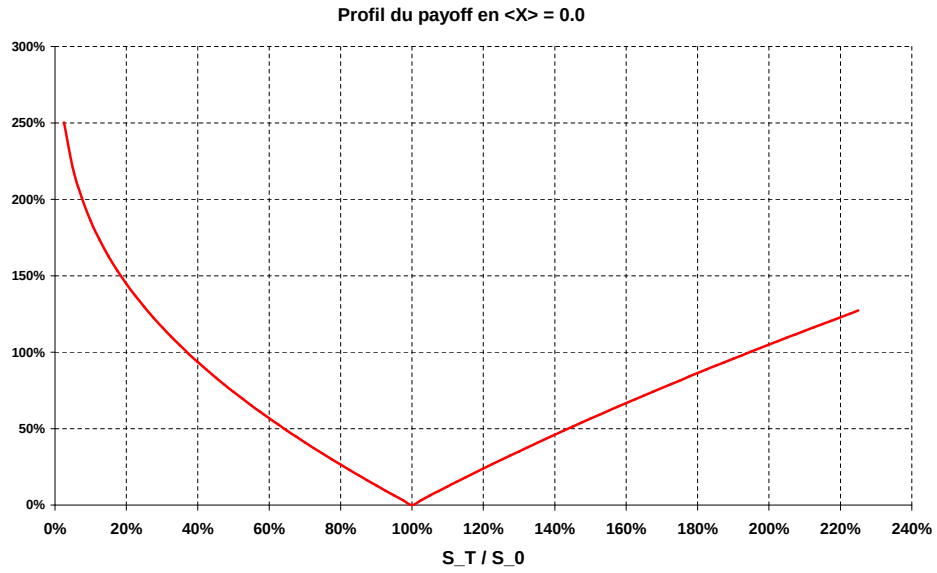


FIG. 4.10.2. Profil du payoff qui couvre le swap de volatilité en t=0

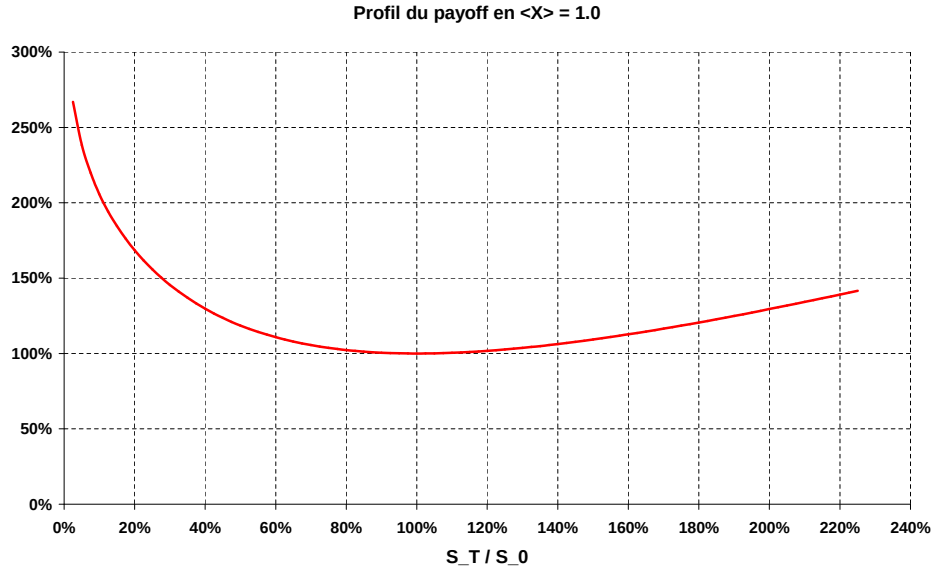


FIG. 4.10.3. Profil du payoff qui couvre le swap de volatilité en t

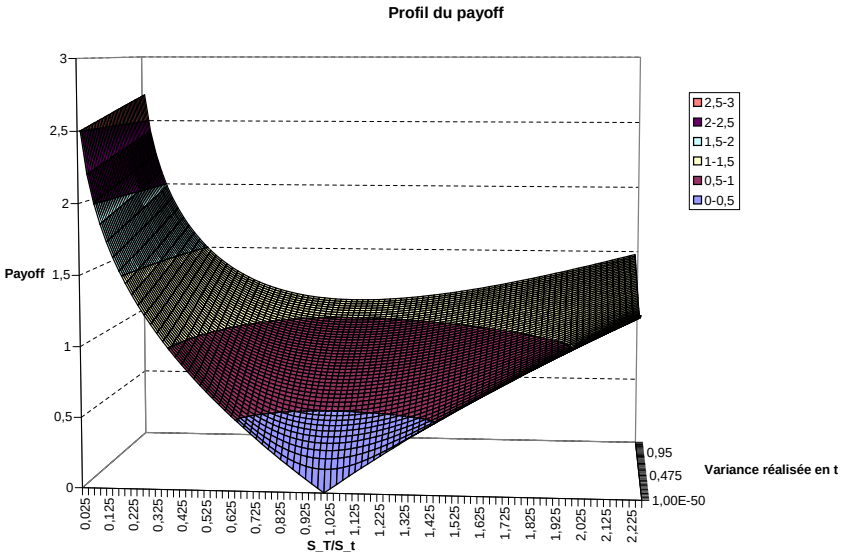


FIG. 4.10.4. Evolution de la couverture

4.10.2. Sharpe Ratio $\frac{e^{X_T}-1}{\sqrt{\langle X\rangle_T}}$. Le ratio de Sharpe vise à mettre en rapport la performance et le risque d'un portefeuille. Il se définit pour un portefeuille, comme la sur-performance obtenue par rapport à un placement sans risque, divisée par la volatilité du portefeuille. Lorsqu'il est positif, plus le ratio de Sharpe est élevé, plus la prise de risque est remunérée. Un ratio de Sharpe négatif signifie simplement que le portefeuille n'a pas sur-performé un placement sans risque.

Nous montrons maintenant comment le répliquer de façon modèle indépendant.

PROPOSITION 4.10.2. *Pour le Sharpe Ratio* $\frac{e^{X_T}-1}{\sqrt{\langle X\rangle_T}}$, *le payoff de couverture robuste* $P_t = P(h_t, 0) + P(0, \delta_t)$ *vaut en* $t = 0$

$$P(h_0, 0) = 0$$
$$P(0, \delta_0)(x) = \sqrt{\frac{\pi}{2}} e^{\frac{x}{2}} sign(x) I_0\left(\frac{x}{2}\right)$$

et en t

$$P(h_t, 0)(x) = \frac{\left(e^{X_t} - 1\right)}{\sqrt{2}\left\langle X\right\rangle_t^2} x^2 e^{\frac{x}{2}} \int_0^1 e^{\frac{(u^2-1)\langle X\rangle_t}{x^2}} \left\{ I_1\left(\frac{u\,|x|}{2}\right) - sign(x) I_0\left(\frac{u\,|x|}{2}\right) u \right\} (1-u^2) du$$
$$P(0, \delta_t)(x) = \sqrt{2} e^{X_t} \frac{x^2 e^{\frac{x}{2}} sign(x)}{\langle X\rangle_t^2} \int_0^1 e^{\frac{(u^2-1)\langle X\rangle_t}{x^2}} I_0\left(\frac{u\,|x|}{2}\right) u(1-u^2) du.$$

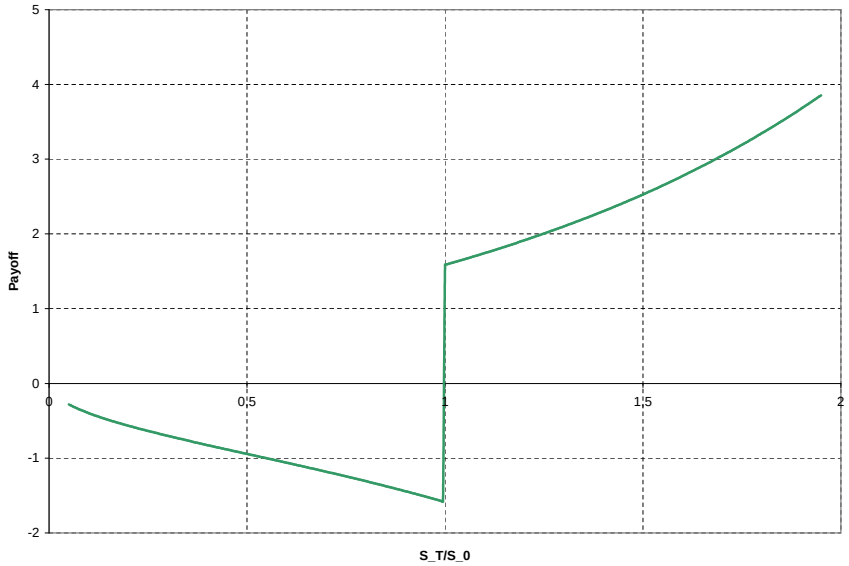


FIG. 4.10.5. Payoff de couverture du Ratio de Sharpe en t=0

Remarque : Dans le cas de non-corrélation volatilité-actif risqué, nous avons $\mathbb{E}_0[\frac{e^{X_T}-1}{\sqrt{\langle X\rangle_T}}] = 0$.

Cependant, le prix devient non nul avec l'apparition d'une corrélation. On peut vérifier la robustesse de cette méthode rho-neutre par rapport à la corrélation en simulant les prix en fonction de la corrélation dans le modèle d'Heston (voir ([**Hes93**])) :

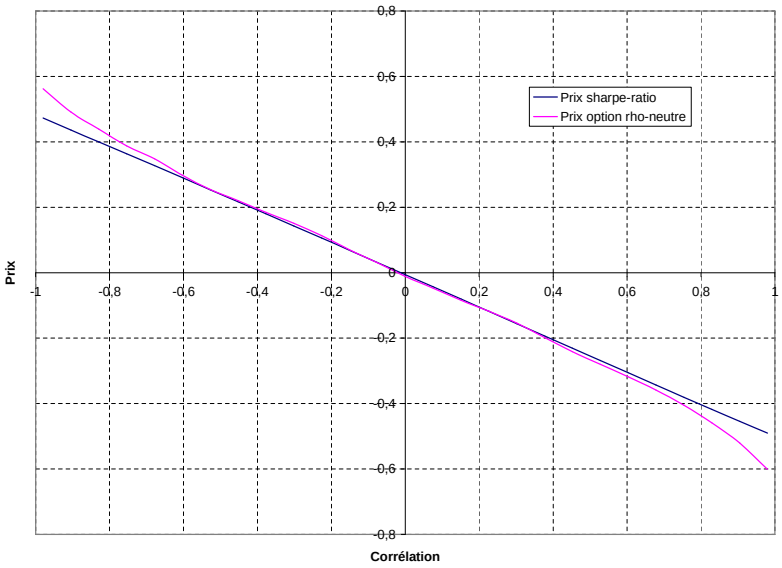


FIG. 4.10.6. Prix du ratio de Sharpe et de l'option rho-neutre $P_0(X_T)$ par rapport à la corrélation

4.11. Options sur fond

4.11.1. Processus asymétriques. On suppose maintenant que l'actif risqué S suit la dynamique :

$$dS_t = (a_S + b_S S_t)\sigma_t dW_t^S$$

avec (σ_t) et (W_t^S) indépendants. Les termes (a_S, b_S) sont supposés connus (par exemple obtenus par calibration) et permettent de mieux prendre en compte le smile de volatilité que dans le cas usuel ($(a_S, b_S) = (0; 1)$). Ca revient à rajouter une correlation positive ou négative entre W^S et la volatilité. Alors pour $t > s$, la résolution donne :

$$S_t = \begin{cases} S_s + a_S X_{s,t} & \text{si } b_S = 0 \\ \left(S_s + \frac{a_S}{b_S}\right) e^{b_S X_{s,t}} - \frac{a_S}{b_S} & \text{si } b_S \neq 0 \end{cases}$$

où

$$\begin{aligned} X_{s,t} &= \int_s^t \sigma_u dW_u^S - \frac{b_S}{2} V_{s,t} \\ &= \begin{cases} \frac{1}{b_S} \ln\left(\frac{a_S + b_S S_t}{a_S + b_S S_s}\right) & \text{si } b_S \neq 0 \\ a_S(S_t - S_s) & \text{si } b_S = 0 \end{cases} \\ V_{s,t} &= \int_s^t \sigma_u^2 du. \end{aligned}$$

On considère des fonds $F_t = \left(F_t^1, ..., F_t^d\right)$ qui suivent la dynamique

$$dF_t^i = (a_F^i + b_F^i F_t^i)(\beta_i \sigma_t dW_t^S + dL_{V_t}^{1,i} + dL_t^{2,i}), \quad \forall 1 \leq i \leq d,$$

où $\{L^{\epsilon,i}\}_{\epsilon \in \{1;2\}, 1 \leq i \leq d}$ sont des processus de Lévy (c'est-à-dire des processus à accroissements indépendants et stationnaires, voir ([**Ber98**])) indépendants de (σ, W^S) et où $(a_F^i, b_F^i)_{1 \leq i \leq d}$ sont supposés connus. On note

$$\begin{aligned} \psi_{\epsilon,i}(p) &= \ln(\mathbb{E}[e^{pL_t^{\epsilon,i}}]) \quad \forall p \in \mathbb{C}, \forall \epsilon \in \{1; 2\}, \forall 1 \leq i \leq d, \\ L_{s,t}^{\epsilon,i} &= L_t^{\epsilon,i} - L_s^{\epsilon,i}. \end{aligned}$$

On suppose que $\{L^{\epsilon,i}\}_{\epsilon \in \{1;2\}, 1 \leq i \leq d}$ admettent des moments exponentiels, c'est-à-dire

$$|\psi_{\epsilon,i}(p)| < +\infty, \quad \forall p \in \mathbb{C}, \forall \epsilon \in \{1; 2\}, \forall 1 \leq i \leq d.$$

Alors la résolution donne pour $s \leq t$

$$F_t^i = \begin{cases} F_s^i + a_F^i\{\beta_i(X_{s,t} + \frac{b_S}{2} V_{s,t}) + L_{V_t}^{1,i} - L_{V_s}^{1,i} + L_t^{2,i} - L_s^{2,i}\} & \text{si } b_F^i = 0 \\ (\frac{a_F^i}{b_F^i} + F_s^i) e^{b_F^i(\beta_i X_{s,t} + L_{V_s,V_t}^{1,i} + L_{s,t}^{2,i}) + \{\frac{\beta_i^2 b_S - \beta_i^2 (b_F^i)^2}{2} - \psi_{1,i}(b_F^i)\} V_{s,t} - \psi_{1,i}(b_F^i)(t-s)} - \frac{a_F^i}{b_F^i} & \text{si } b_F^i \neq 0 \end{cases}$$

On cherche à couvrir l'option de payoff $H(F_T)$. On peut étendre cette méthode de réplication dynamique à la couverture en variance moyenne des dérivés sous la forme ξ_T telle que

$$\begin{aligned} H(F_T) &= h_t(V_{t,T}) + \int_t^T \delta_s(V_{s,T})\, \sigma_s dW_s^S + M_T - M_t \\ &= h_t(V_{t,T}) + \int_t^T \delta_s(V_{s,T})\, \frac{dS_s}{a_S + b_S S_s} + M_T - M_t \end{aligned}$$

où M est une martingale vérifiant $[dM_s, dW_s^S] = 0$ et

$$h_t(v) = \mathbb{E}\left[H(F_T)|\mathcal{F}_t, V_{t,T} = v\right],$$
$$\delta_t(v) = \mathbb{E}\left[\frac{D_t^W H(F_T)}{\sigma_t}|\mathcal{F}_t, V_{t,T} = v\right],$$

où D_t^W est la dérivée de Malliavin par rapport W (voir ([**Nua95**])). On peut réécrire δ_t sous la forme

$$\begin{aligned}\delta_t(v) &= \frac{\partial}{\partial x}\mathbb{E}\left[H(F_T)|\mathcal{F}_t, V_{t,T} = v\right] \\ &= \mathbb{E}\left[\nabla H(F_T).\frac{\partial F_T}{\partial x}|\mathcal{F}_t, V_{t,T} = v\right] \\ &= \mathbb{E}\left[H(F_T)(X_{s,t} + \frac{b_S V_{t,T}}{2})|\mathcal{F}_t, V_{t,T} = v\right]\end{aligned}$$

avec

$$\frac{\partial F_T^i}{\partial x} = \begin{cases} a_F^i \beta_i & \text{si } b_F^i = 0 \\ b_F^i \beta_i (\frac{a_F^i}{b_F^i} + F_s^i) & \text{si } b_F^i \neq 0 \end{cases}$$

On a alors l'extension de la proposition (4.8.6) qui nous donne la couverture dynamique en moyenne variance par des options européennes sur S_T.

PROPOSITION 4.11.1. *Avec*

$$P(h_t, 0)(\lambda) = \sqrt{\pi}\frac{e^{\frac{b_S^2}{16}\frac{1}{\lambda}}}{\sqrt{\lambda}} h_t\left(\frac{1}{2\lambda}\right),$$
$$P(0, \delta_t)(\lambda) = \frac{\sqrt{\pi}}{2}\lambda^{-\frac{3}{2}} e^{\frac{b_S^2}{16}\frac{1}{\lambda}} \delta_t\left(\frac{1}{2\lambda}\right) - \frac{\sqrt{\pi} b_S}{4}\lambda^{-\frac{3}{2}} e^{\frac{b_S^2}{16}\frac{1}{\lambda}} h_t\left(\frac{1}{2\lambda}\right),$$

la couverture dynamique avec l'option européenne $P_t(X_{t,T}) = P(h_t, 0)(X_{t,T}) + P(0, \delta_t)(X_{t,T})$ est robuste par rapport à la variance V :

$$\mathbb{E}\left[P_t(X_{t,T})|\mathcal{F}_t, V_{t,T} = v\right] = h_t(v),$$
$$\frac{\partial}{\partial x}\mathbb{E}\left[P_t(X_{t,T})|\mathcal{F}_t, V_{t,T} = v\right] = \delta_t(v). \tag{4.11.1}$$

REMARK 4.11.2. Il s'agit bien d'une extension de la première section sur la couverture rho-neutre des options des options sur variance réalisée et sur l'actif S. En effet, en prenant $a_S = a_F^1 = b_F^2 = \beta_2 = L^{1,1} = L^{2,2} = L^{1,2} = 0$, $b_S = \beta_1 = a_F^2 = 1$ et $L_t^{2,1} = t$, on obtient une option sur l'actif risqué et de la variance réalisée : $(F_t^1, F_t^2) = (S_t, \langle X\rangle_t)$.

4.11.2. Cas d'un fond. On cherche à couvrir l'option sur un fond $H(F_T)$ ($d = 1$, $F_T = F_T^1$) de manière robuste par rapport à V.

On suppose que le fond F suit la dynamique :

$$dF_t = (a_F + b_F F_t)\left(\beta\rho\sigma_t dW_t^S + \beta\sigma_t\sqrt{1-\rho^2} dW_t^\perp + dL_t\right) \tag{4.11.2}$$

avec W^S et $W^\perp$ indépendants et L_t un processus de Levy indépendant de $(W^S, W^\perp, \sigma)$. On a alors :

si $b_F \neq 0$

$$\begin{aligned} F_T &= \left(F_t + \frac{a_F}{b_F}\right) e^{b_F\beta\rho\int_t^T \sigma_u dW_u^S + b_F\beta\sqrt{1-\rho^2}\int_t^T \sigma_t dW_u^\perp - \frac{b_F^2\beta^2}{2}\int_t^T \sigma_u^2 du + b_F(L_T - L_t) - (T-t)\psi(b_F)} - \frac{a_F}{b_F} \\ &= \left(F_t + \frac{a_F}{b_F}\right) e^{b_F\beta\rho X_{t,T} + b_F\beta\sqrt{1-\rho^2}\sqrt{V_{t,T}}N + \frac{\rho\beta b_F b_S - b_F^2\beta^2}{2}V_{t,T} + \xi_{t,T}} - \frac{a_F}{b_F} \end{aligned}$$

et si $b_F = 0$

$$F_T = F_t + a_F(\beta\rho X_{t,T} + \beta\sqrt{1-\rho^2}\sqrt{V_{t,T}}N + \frac{\rho\beta b_S}{2}V_{t,T} + \xi_{t,T})$$

où $N \sim \mathcal{N}(0,1)$ indépendant de $\mathcal{F}_t$, de $X_{t,T}$ et de $V_{t,T} = \int_t^T \sigma_s^2 ds$ et où $\xi_{t,T} = b_F(L_T - L_t) - (T-t)\psi(b_F)$ avec $\psi(p) = \ln \mathbb{E}\left[e^{pL_1}\right]$. Notons

$$H_{t,\xi}(x) = \begin{cases} \mathbb{E}_t\left[H\left(\left(F_t + \frac{a_F}{b_F}\right)e^{b_F\beta x + \xi_{t,T}} - \frac{a_F}{b_F}\right)\right] & \text{si } b_F \neq 0 \\ \mathbb{E}_t\left[H\left(F_t + a_F(\beta x + \xi_{t,T})\right)\right] & \text{si } b_F = 0 \end{cases}$$

et $Q^0_{t,\xi}$ et $Q^1_{t,\xi}$ les parties paires de $y \to H_{t,\xi}(y)\, e^{-\frac{b_F\beta}{2}y}$ et de $y \to H'_{t,\xi}(y)\, e^{-\frac{b_F\beta}{2}y}$, c'est-à-dire

$$\begin{aligned} Q^0_{t,\xi}(x) &= \frac{H_{t,\xi}(y)\, e^{-\frac{b_F\beta}{2}y} + H_{t,\xi}(-y)\, e^{\frac{b_F\beta}{2}y}}{2}, \\ Q^1_{t,\xi}(x) &= \frac{H'_{t,\xi}(y)\, e^{-\frac{b_F\beta}{2}y} + H'_{t,\xi}(-y)\, e^{\frac{b_F\beta}{2}y}}{2}. \end{aligned}$$

PROPOSITION 4.11.3. *Le payoff $P_t(X_{t,T})$ de couverture robuste (qui vérifie (4.11.1)) vaut :*

$$\begin{aligned} P_t(x) = e^{\frac{b_S}{2}x}Q^0_{t,\xi}(x) - e^{\frac{b_S}{2}x}|x|\int_{[0,|x|]} Q^0_{t,\xi}(y)\sqrt{x^2-y^2}J_1\left(\mu_\beta\sqrt{x^2-y^2}\right)\frac{4dy}{\mu_\beta} \\ + e^{\frac{b_S}{2}x}sign(x)\int_{[0,|x|]}\left(\rho Q^1_{t,\xi}(y) - \frac{b_S Q^1_{t,\xi}(y)}{2}\right)J_0\left(\mu_\beta\sqrt{x^2-y^2}\right)dy \end{aligned}$$

avec $\mu_\beta = \frac{\sqrt{b_F^2\beta^2 - b_S^2}}{2}$ et où J_ν désigne la fonction de Bessel de première espèce.

COROLLARY 4.11.4. *Supposons $b_S, b_F \neq 0$. Dans le cas du call $H(F_T) = (F_T - K)^+$, on a l'approximation du payoff ρ-neutre par un call et un put :*

$$P_t(X_{t,T}) \simeq (F_t - K)^+ + \rho m_t^0(S_T - S_t) + (1-\rho)\, m_t^c\,(S_T - k_t^c)^+ + (1+\rho)\, m_t^p\,(k_t^p - S_T)^+$$

où $l_S(x) = a_S + b_S x$, $l_F(x) = a_F + b_F x$ *et*

$$\begin{aligned}
k_t^c &= \left(\frac{a_S}{b_S} + S_t\right) e^{\frac{b_S}{b_F\beta}|\ln(\frac{l_F(F_t)}{l_F(K)})|} - \frac{a_S}{b_S}, \\
k_t^p &= \left(\frac{a_S}{b_S} + S_t\right) e^{-\frac{b_S}{b_F\beta}|\ln(\frac{l_F(F_t)}{l_F(K)})|} - \frac{a_S}{b_S}, \\
m_t^c &= \frac{\beta}{2}\frac{\sqrt{l_F(F_t)l_F(K)}}{l_S(S_t)} e^{-\frac{b_S}{2b_F\beta}|\ln(\frac{l_F(F_t)}{l_F(K)})|}, \\
m_t^p &= \frac{\beta}{2}\frac{\sqrt{l_F(F_t)l_F(K)}}{l_S(S_t)} e^{\frac{b_S}{2b_F\beta}|\ln(\frac{l_F(F_t)}{l_F(K)})|}, \\
m_t^0 &= \beta\frac{l_F(F_t)}{l_S(S_t)}.
\end{aligned}$$

4.11.3. Applications. On suppose ici que les dynamiques de S et du fond F sont respectivement données par

$$\frac{dS_t}{S_t} = \sigma_t dW_t^S \quad \text{et} \quad \frac{dF_t}{F_t} = \sqrt{v_t^F} dW_t^F$$

(ie $a_S = a_F = 0$ et $b_S = b_F = 1$) avec

$$v_t^F = \beta^2\sigma_t + \varepsilon^2$$

On peut réécrire F sous la forme (4.11.2) :

$$\begin{aligned}
\frac{dF_t}{F_t} &= \beta\rho\sigma_t dW_t^S + \beta\sqrt{1-\rho^2}\sigma_t dW_t^\perp + \epsilon dW_t^\epsilon \\
&= \beta\rho\frac{dS_t}{S_t} + \beta\sqrt{1-\rho^2}\sigma_t dW_t^\perp + \epsilon dW_t^\epsilon
\end{aligned}$$

où W^S, $W^\perp$ et W^ϵ sont des mouvements Brownien indépendants (et aussi indépendants de σ). Alors

$$F_T = F_t e^{\beta\rho X_{t,T} + \beta\sqrt{1-\rho^2}\int_t^T \sigma_s dW_s^\perp + \epsilon(W_T^\epsilon - W_t^\epsilon) - \frac{\beta^2\langle X\rangle_{t,T} + \epsilon^2(T-t)}{2}}$$

Donnons quelques exemples de tels fonds :

◇ Applications aux stratégies dynamiques Constant-Mix et CPPI

On s'intéresse aux stratégies dynamiques pour l'allocation d'actif, plus précisément aux stratégies Constant Mix et Constant Proportion Portfolio Insurance (CPPI) (voir [**PS88**]).

Constant Mix : En prenant les taux d'intérêt (on utilise les valeurs forwards de la stratégies et des prix), la gestion Constant-Mix M est un stragégie avec effet de levier (avec un multiplicateur m) : il s'agit une option mixte telle que

$$\frac{dM_t}{M_t} = m\frac{dS_t}{S_t} \Rightarrow M_T = M_t e^{mX_{t,T} + \frac{m(1-m)}{2}\langle X\rangle_{t,T}},$$

c'est-à-dire $\beta = m$, $\rho = 1$ et $\epsilon = 0$ avec les notations précédentes pour les options sur fond.

Le Constant Mix est un portefeuille optimal pour une utilité CRRA dans le modèle de Black-Scholes.

CPPI : Le CPPI est un fond à coussin (avec une valeur plancher P) avec un effet de levier La richesse au dessus du plancher P_t est investi avec un effet de levier

(avec un multiplicateur m). Avec des taux d'intérêt nuls, le CPPI C peut être vu comme la combinaison d'un Constant-Mix et d'une valeur plancher P

$$dC_t = m(C_t - P)\frac{dS_t}{S_t} \Rightarrow\ C_T = P + (C_t - P)e^{mX_{t,T} + \frac{m(1-m)}{2}\langle X\rangle_{t,T}}.$$

$\diamond$ Un fond à volatilité benchmarkée dont la dynamique vaut :

$$\begin{cases} \frac{dF_t}{F_t} = \sqrt{v_t^F}\, dW_t^{\perp} \\ \frac{dS_t}{S_t} = \sqrt{\sigma_t} dW_t^S \end{cases}$$

avec $v_t^F = \beta^2\sigma_t^2 + \varepsilon^2$ $(\rho = 0)$
$\diamond$ Un fond benchmarké dont la dynamique vaut :

$$\frac{dF_t}{F_t} = \beta\frac{dS_t}{S_t} + \varepsilon dW_t^{\perp} \Rightarrow F_T = F_t e^{\beta X_{t,T} + \epsilon\sqrt{T-t}Z - \frac{(\beta^2-\beta)\langle X\rangle_{t,T} + \epsilon^2(T-t)}{2}}$$

où $Z \sim \mathcal{N}(0;1)$ indépendante de $X_{t,T}$.
$\diamond$ Un fond benchmarké avec une volatilité benchmarkée dont la dynamique vaut :

$$\frac{dF_t}{F_t} = \rho\beta\frac{dS_t}{S_t} + \sqrt{(1-\rho^2)\beta^2 v_t^S + \varepsilon^2}\, dW_t^{\perp}$$

avec $v_t^F = \beta^2\sigma_t + \varepsilon^2$,

On cherche à couvrir $H(F_T)$. On note P_t le payoff de couverture robuste donné par la proposition (4.11.3) et

$$\Delta_t^F = \frac{\partial}{\partial F_t}\mathbb{E}_t\left[H(F_T)\right],$$
$$\Delta_t^S = \mathbb{E}_t\left[P_t'(X_{t,T})\right].$$

Pour répliquer $H(F_T)$, la stratégie robuste est donnée à la date t par une position dynamique avec :
$\diamond$ Δ_t^F futures sur F_T
$\diamond$ une option $P_t(\ln(\frac{S_T}{S_t}))$
$\diamond$ $-\Delta_t^S$ futures sur S_T
Remarque : La position Δ_t^F futures sur F_T permet de se couvrir en delta (en F). La position $P_t(\ln(\frac{S_T}{S_t})) - \Delta_t^S S_T$ est delta nulle et permet de se couvrir en véga.

Exemple du call. On cherche à couvrir le call $(F_T - K)^+$. Avec

$$H_{t,\xi}(x) = \mathbb{E}_t\left[\left(F_t e^{\beta x + \epsilon\sqrt{T-t}Z - \frac{\epsilon^2(T-t)}{2}} - K\right)^+\right], \quad \text{avec } Z \sim N(0,1)$$

et avec $Q_{t,\xi}^0(x)$ et $Q_{t,\xi}^1(x)$ les parties paires de $H_{t,\xi}(x)\, e^{-\frac{\beta x}{2}}$ et de $H'_{t,\xi}(x)\, e^{-\frac{\beta x}{2}}$), on obtient à partir de la proposition (4.11.3)

$$P_t(x) = e^{\frac{x}{2}} Q_{t,\xi}^0(x) - e^{\frac{x}{2}}|x| \int_{[0,|x|]} Q_{t,\xi}^0(x)(y)\sqrt{x^2-y^2}\, J_1\left(\frac{\sqrt{\beta^2-1}}{2}\sqrt{x^2-y^2}\right)\frac{8dy}{\sqrt{\beta^2-1}}$$
$$+ e^{\frac{x}{2}}\mathrm{sign}(x)\int_{[0,|x|]}\left(\rho Q_{t,\xi}^1(y)\, e^{-\frac{\beta y}{2}} + \frac{Q_{t,\xi}^0(y)}{2}\right) J_0\left(\frac{\sqrt{\beta^2-1}}{2}\sqrt{x^2-y^2}\right) dy$$

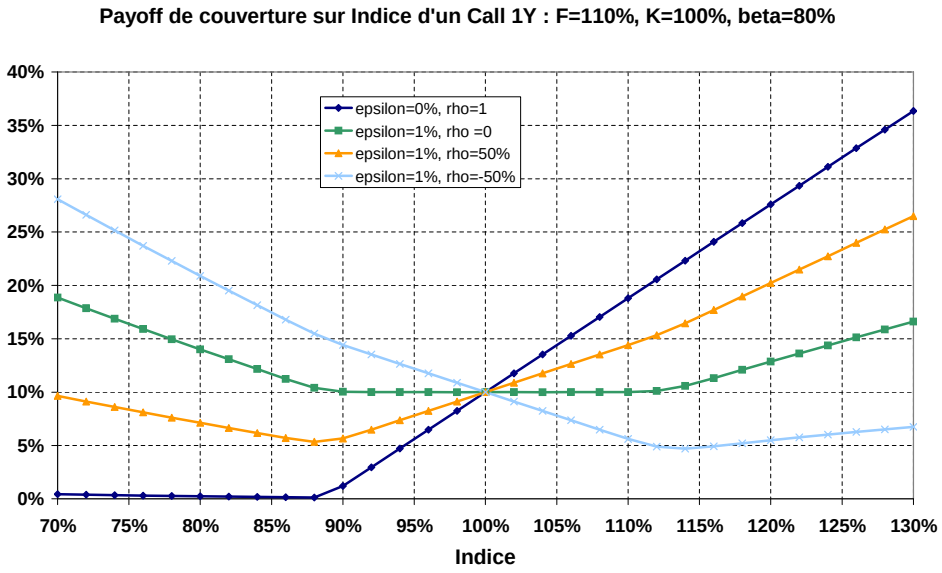


FIG. 4.11.1. Profil sur l'indice pour couvrir le véga du call sur le fond, $f = 110\%$, $K = 100\%$ et $\beta = 80\%$

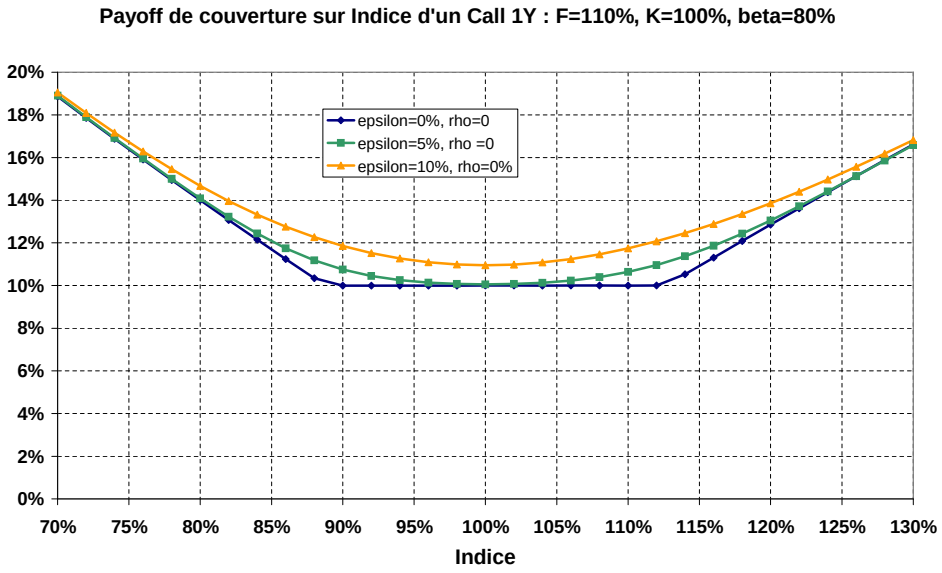


FIG. 4.11.2. Profil sur l'indice en fonction de ϵ

Si ε^2 est petit, le profil du payoff sur S peut être approximé par un call et un put (sur S_T) qui sont donnés par le corollaire (4.11.4) : $P_t(X_{t,T}) \simeq P_t^{app}(S_T)$

$$\begin{aligned} P_t^{app}(S_T) =& (F_t - K)^+ + \rho\beta F_t \left(\frac{S_T}{S_t} - 1\right) + (1-\rho)\frac{\beta}{2}\sqrt{F_t K}\alpha_t^{-\frac{1}{2}}\left(\frac{S_T}{S_t} - \alpha_t\right)^+ \\ &+ (1+\rho)\frac{\beta}{2}\sqrt{F_t K}\alpha_t^{\frac{1}{2}}\left(\alpha_t^{-1} - \frac{S_T}{S_t}\right)^+ \end{aligned}$$

avec $\alpha_t = \exp(\frac{1}{\beta}\left|\ln\left(\frac{F_t}{K}\right)\right|)$.

Ce call et ce put permettent de se couvrir en véga (et au premier ordre en correlation). Il reste à se couvrir en delta. Les deltas sur les options sur le sous-jacent S sont supposés connus. Il reste à calculer le delta du call sur le fond F de manière modèle indépendant. Nous avons :

$$\begin{aligned} \Delta_t^F &= \frac{\partial \mathbb{E}_t\left[(F_T - K)^+\right]}{\partial F_t}, \\ \Delta_t^S &= \frac{\partial \mathbb{E}_t\left[P_t^{app}(S_T)\right]}{\partial S_t}, \end{aligned}$$

La stratégie de couverture à la date t a une position dynamique avec :
- Δ_t^F contrats futures sur F_T
- une option $P_t^{app}(S_T)$
- $-\Delta_t^S$ contrats futures sur S_T

REMARK 4.11.5. Δ_t^F peut être calculé de manière robuste en la corrélation (avec une erreur à l'ordre 2). En effet, comme

$$\Delta_t^F = \mathbb{E}_t\left[1_{F_T \geq K}\frac{F_T}{F_t}\right]$$

il suffit de trouver le payoff ρ-neutre sur S_T en appliquant le résultat de la proposition (4.11.3) à $1_{F_T \geq K}\frac{F_T}{F_t}$.

Simulation du marché. Pour la simulation, on suppose que S_t suit le modèle de Heston (cf ([**Hes93**])) :

$$\begin{aligned} \frac{dS_t}{S_t} &= \sqrt{v_t^S}dW_t^S, \\ dv_t^S &= k(v_\infty^S - v_t^S)dt + \xi\sqrt{v_t^S}dW_t^v, \end{aligned}$$

et pour le fond

$$\frac{dF_t}{F_t} = \sqrt{v_t^F}dW_t^F \quad \text{avec } v_t^F = \beta^2 v_t^S + \varepsilon^2$$

On suppose : $\left\langle dW_t^S, dW_t^v\right\rangle = \rho^S$, $\left\langle dW_t^F, dW_t^S\right\rangle = 0$ et $\left\langle dW_t^F, dW_t^v\right\rangle = \rho^F$ (avec $(\rho^S)^2 + (\rho^F)^2 < 1$). L'hypothèse d'indépendance entre σ et W^S n'est pas respectée (si $\rho^S \neq 0$), on va tester la robustesse par rapport à cette corrélation.

Pour les simulations, les valeurs standards des paramètres sont

$$\begin{aligned} F_0 &= 1 \\ S_0 &= 1 \\ T &= 1 \\ K &= 1 \\ \beta &= 0.8 \\ \epsilon &= 0.05 \\ v_\infty^S &= 0.033 \\ v_0^S &= 0.022 \\ k &= 0.62 \\ \xi &= 0.55 \end{aligned}$$

et les corrélations (ρ^S et $\rho^\perp$) sont les paramètres étudiés.

Fond à volatilité benchmarkée. Dans le cas d'un fond à volatilité benchmarkée, c'est-à-dire $\rho = 0$ (ie $\langle dW_t^S, dW_t^F \rangle = 0$), la couverture ρ-neutre est efficace.

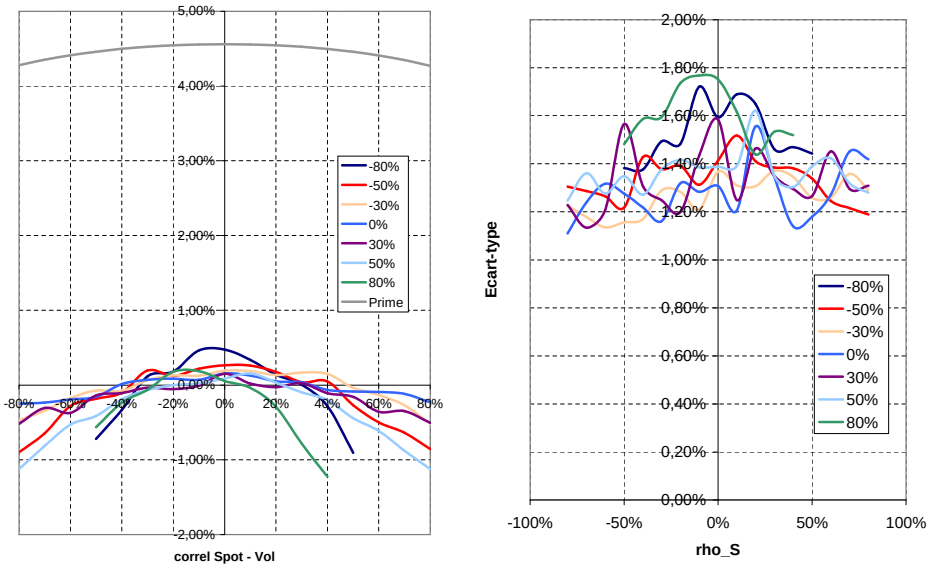


FIG. 4.11.3. Etude du P&L (moyenne et écart-type) pour une couverture hebdomadaire (1 courbe = 1 niveau de corrélation Fonds - Vol)

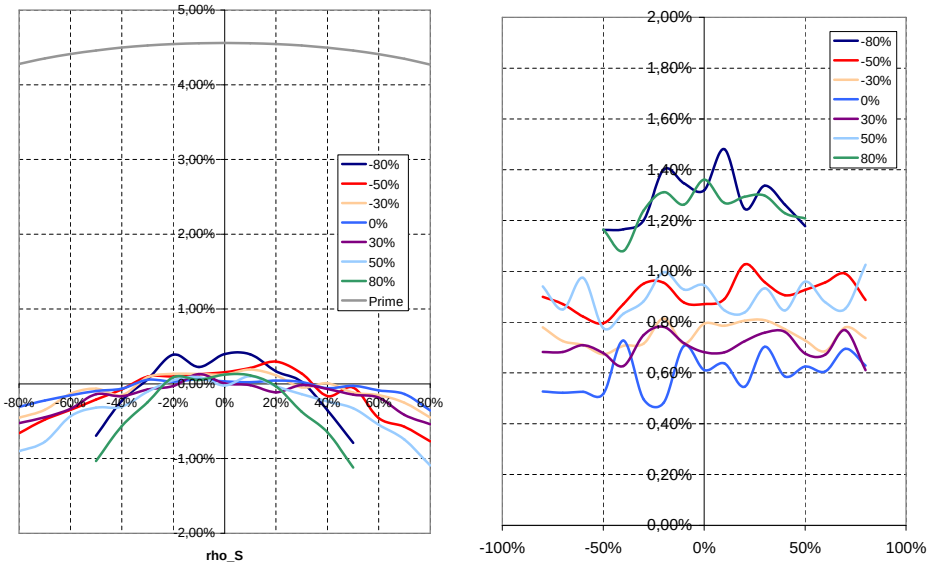


FIG. 4.11.4. Etude du P&L (moyenne et écart-type) pour une couverture journalière (1 courbe = 1 niveau de corrélation Fonds - Vol)

Fond benchmarké. On suppose que le fond satisfait la relation suivante

$$\frac{dF_t}{F_t} = \beta \frac{dS_t}{S_t} + \varepsilon dW_t^{\perp}$$

où $\left\langle \frac{dS}{S}, dW^{\perp} \right\rangle = 0$ et $\left\langle dv^S, dW^{\perp} \right\rangle = 0$.

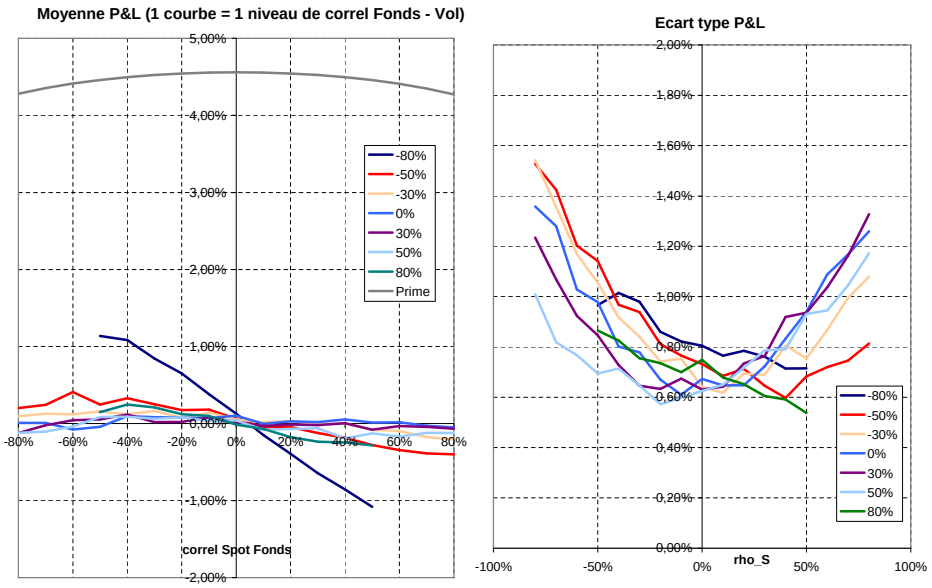


FIG. 4.11.5. Etude du P&L (moyenne et écart-type) pour une couverture hebdomadaire (1 courbe = 1 niveau de corrélation Fonds - Vol)

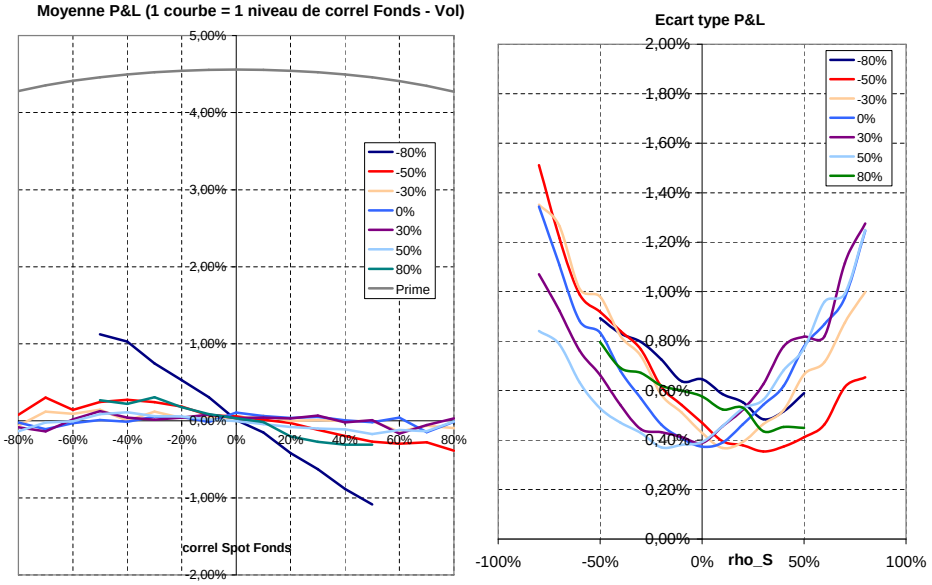


FIG. 4.11.6. Etude du P&L (moyenne et écart-type) pour une couverture journalière (1 courbe = 1 niveau de corrélation Fonds - Vol)

Ces réplications (journalières et hebdomadaires), indépendantes du modèle, sont acceptables. On constate que la couverture du call sur fonds benchmarké est légèrement plus efficace que celle d'un fonds à volatilité benchmarkée.

Plusieurs raisons expliquent l'imperfection de la couverture :
◇ Nous avons approximé les payoffs par des calls et des puts (cette approximation devient moins bonne quand $\varepsilon^2 (T-t)$ devient grand).
◇ Les prix et les deltas (Δ_t^S et Δ_t^F) lors de la couverture ont été calculés de manière rho-neutre, c'est-à-dire qu'ils sont supposés exactes à corrélation nulle et qu'ils annulent le premier ordre du à la corrélation non nulle. La forme des courbes (paraboliques et non affines) confirme l'annulation de cette erreur de premier ordre.
◇ La couverture se fait en temps discret (et non en temps continu). La couverture modèle indépendant journalière est plus fiable que celle hebdomadaire (surtout avec une faible corrélation).

4.11.4. Application aux paniers rebalancés. On considère le panier rebalancé constitué de deux actifs risqués de prix S_t^1 et S_t^2. La dynamique de son prix est la suivante

$$\frac{dB_t}{B_t} = \alpha \frac{dS_t^1}{S_t^1} + (1-\alpha)\frac{dS_t^2}{S_t^2}.$$

On réécrit la dynamique du panier dans le numéraire S_t^2

$$\frac{d\left(\frac{B_t}{S_t^2}\right)}{\left(\frac{B_t}{S_t^2}\right)} = \alpha \frac{d\left(\frac{S_t^1}{S_t^2}\right)}{\left(\frac{S_t^1}{S_t^2}\right)}.$$

On considère un call sur la valeur terminale du panier $(B_T - K)_+$. On peut donc appliquer les résultats précédents dans le numéraire S_t^2 et réécrire le résultat dans le numéraire cash, ce qui donne le profil dynamique répliquant suivant

$$\tilde{P}_t(S_T^1, S_T^2) \simeq (B_t - K)^+ + \alpha B_t \left(\frac{S_T^1}{S_t^1} - \frac{S_T^2}{S_t^2}\right) + \alpha\sqrt{B_t K} e^{\frac{|\ln(\frac{B_t}{K})|}{2\alpha}} \left(\frac{S_T^2}{S_t^2} e^{-\frac{|\ln(\frac{B_t}{K})|}{\alpha}} - \frac{S_T^1}{S_t^1}\right)^+$$

4.12. Extension : processus de Lévy

On peut donner une valorisation et une couverture en moyenne variance pour des dérivés sur la variance réalisée en étant robuste sur le subordinateur du processus de Lévy. On étend maintenant les résultats précédents en remplaçant le mouvement Brownien par un processus de Levy L_t , la variance réalisée par un subordinateur indépendant V du processus de Lévy et l'actif risqué $S_t = S_0 e^{L_{V_t}}$ par un processus exponentielle de Lévy avec changement de temps.

Soit L un processus de Levy sur $\mathbb{R}$ défini à l'aide du triplet caractéristique (A, ν, γ) (voir ([**Ber98**])

$$\begin{aligned} L_t &= \gamma t + \sigma W_t + \int_{\mathbb{R}} x\tilde{\mu}(dx, [0,t]), \\ \mathbb{E}\left[e^{pL_t}\right] &= e^{t\psi(p)}, \\ \psi(p) &= \frac{Ap^2}{2} + \gamma p + \int_{\mathbb{R}} \left(e^{px} - 1 - px1_{|x|\leq 1}\right)\nu(dx), \end{aligned}$$

où W est un mouvement Brownien et $\tilde{\mu}$ est une mesure de Poisson compensée $\tilde{\mu}(ds, dt) = \mu(ds, dt) - \nu(dt)$ indépendante de W. On suppose que $|\psi(p)| < +\infty$ pour tout $p \in \mathbb{C}$.

On considère le processus de Lévy avec changement de temps

$$X_t = L_{V_t}$$

où V_t est un subordinateur continu et indépendant de L. On note $L_{t,t'} = L_{t'} - L_t$ (idem pour X et V).

L'actif risqué est maintenant défini par l'exponentielle du processus de Levy

$$\begin{aligned} S_t &= S_0 \exp(X_{0,t}) \\ &= S_0 \exp(L_{V_t}). \end{aligned}$$

On cherche à valoriser une option ξ_T telle que

$$h_t(v) = \mathbb{E}\left[\xi_T | \mathcal{F}_t, V_{t,T} = v\right]$$

en notant $\mathcal{F}_t = \sigma\left(X_s, s \leq t\right)$ et $\mathcal{F}_t^L = \sigma\left(L_s, s \leq t\right)$ (ie $\mathcal{F}_t = \mathcal{F}_{V_t}^L$). Comme dans les parties précédentes, on va chercher à valoriser et à couvrir ξ_T ou $h_t(V_{t,T})$ à l'aide d'option européenne sur $X_{t,T}$ (ou bien de manière équivalente sur $e^{X_{t,T}}$) de maturité T.

Généralement, l'option à valoriser sera de la forme $\xi_T = H_t(X_{t,T}, V_{t,T}, Z)$ avec H_t est $\mathcal{F}_t$-mesurable et Z une variable aléatoire indépendante de $\sigma(\mathcal{F}_t, X_{t,T}, V_{t,T})$. C'est bien si le cas $\xi_T = H(X_T, L^1_{V_T}, L^2_T)$ où L^1 et L^2 sont aussi des processus de Lévy indépendants L à valeurs dans $\mathbb{R}^{d_1}$ et $\mathbb{R}^{d_2}$. Par exemple, pour une option sur la variance réalisée $L^1_{V_T} = [X]_T$ ou sur un fond de la forme $F_T = f(L_{V_T}, L^1_{V_T}, L^2_T)$.

On suppose, comme dans la partie précédente, que les taux d'intérêt sont nuls et l'actif risqué est martingale. D'où e^{L_t} est une martingale et $\psi(1) = 0$.

On remarque que ψ est convexe et $\psi(0) = \psi(1) = 0$. On suppose que S n'est pas constant p.s.. On a $\mathbb{P}\left(L_t > 0\right) > 0$ et $\mathbb{P}\left(L_t < 0\right) > 0$, ce qui implique $\lim_{p\to\pm\infty} \psi(p) = +\infty$. On a $\psi(0) = 0$. On suppose $\psi'(0) \neq 0$, i.e. $\min_p \psi(p) < 0$.

On peut alors définir les bijections $\phi_+ : \mathbb{R}^+ \to \mathbb{R}^+$ et $\phi_- : \mathbb{R}^+ \to \mathbb{R}^+$ strictement croissantes définies par :

$$\psi\left(1+\phi_+(y)\right) = \psi\left(-\phi_-(y)\right) = y, \quad \forall y \geq \min_p \psi(p).$$

On a alors l'égalité

$$e^{yV_t} = \mathbb{E}\left[e^{(p_+ + \phi_+(y))L_{V_t}} | V_t\right] = \mathbb{E}\left[e^{(p_- - \phi_-(y))L_{V_t}} | V_t\right] \quad \forall y \in \mathbb{R}^+ \tag{4.12.1}$$

REMARK 4.12.1. $\diamond$ Dans le cas continu, $X_t = \int_0^t \sigma_s dW_s - \frac{1}{2}\int_0^t \sigma_s^2 ds = B_{V_t} - \frac{V_t}{2}$ avec $V_t = \int_0^t \sigma_s^2 ds$ et B un mouvement Brownien standard, on a $\psi(p) = \frac{p^2-p}{2}$. D'où

$$\phi_\pm(y) = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2y}, \quad \forall y \in \mathbb{R}^+.$$

Supposons $h_t(v) = \int_{\mathbb{R}^+} e^{yv} m_t(dy)$ où m est une mesure vérifiant $\int_{\mathbb{R}^+} e^{yv}|m_t|(dy) < +\infty \ \forall v \geq 0$. Soit $\theta : \mathbb{R}^+ \to \mathbb{R}$ tel que $\int_{\mathbb{R}^+} e^{yv}\,|\theta(y)|\,|m_t|(dy) < +\infty \ \forall v \geq 0$ (par exemple avec θ borné). En prenant

$$P_t(x) = \int_{\mathbb{R}^+} \left\{\theta(y) e^{(1+\phi_+(y))x} + (1-\theta(y))\, e^{(-\phi^-(y))x}\right\} m_t(dy),$$

alors l'égalité (4.12.1) implique

$$\begin{aligned}
\mathbb{E}\left[|P_t(L_{V_t,V_T})|\,|\mathcal{F}_t, V_{t,T}\right] &\leq \mathbb{E}\left[\int_{\mathbb{R}^+} \left\{|\theta(y)|\, e^{(1+\phi_+(y))L_{V_t,V_T}} + |1-\theta(y)|\, e^{(-\phi^-(y))L_{V_t,V_T}}\right\} |m_t|\,(dy) | \mathcal{F}_t, V_{t,T}\right] \\
&\leq \int_{\mathbb{R}^+} (1 + 2\,|\theta(y)|) e^{yV_{t,T}}\,|m_t|\,(dy) \\
&< +\infty,
\end{aligned}$$

et donne l'inégalité

$$\begin{aligned}
\mathbb{E}\left[P_t(L_{V_t,V_T})|\mathcal{F}_t, V_{t,T}\right] &= \mathbb{E}\left[\int_{\mathbb{R}^+} \left\{\theta(y) e^{(1+\phi_+(y))L_{V_t,V_T}} + (1-\theta(y)) e^{(-\phi_-(y))L_{V_t,V_T}}\right\} m_t(dy) | \mathcal{F}_t, V_{t,T}\right] \\
&= \int_{\mathbb{R}^+} e^{yV_{t,T}} m_t(dy) \\
&= h_t(V_{t,T}).
\end{aligned}$$

Ainsi, l'option européenne $P_t(X_{t,T})$ permet de valoriser l'option ξ_T :

$$\mathbb{E}\left[P_t(L_{V_t,V_T})|\mathcal{F}_t\right] = \mathbb{E}\left[\xi_T|\mathcal{F}_t\right].$$

On a un degré de liberté avec le paramètre θ. On peut l'utiliser soit pour être robuste par rapport à un petit changement de modèle, soit pour une couverture en moyenne variance.

4.12.1. Valorisation robuste. Supposons $\xi_T = H_t(X_{t,T}, V_{t,T}, Z)$ avec H_t est $\mathcal{F}_t$-mesurable et Z une variable aléatoire indépendante de $\sigma(\mathcal{F}_t, X_{t,T}, V_{t,T})$.

On cherche une option européenne $P_t(L_{V_T} - L_{V_t})$ telle que

$$\begin{aligned} \mathbb{E}\left[P_t(X_{t,T})|\mathcal{F}_t, V_{t,T}\right] &= \mathbb{E}\left[\xi_T|\mathcal{F}_t, V_{t,T}\right], \\ \mathbb{E}\left[P_t\left(X_{t,T}\right)\frac{X_{t,T}}{V_{t,T}}|\mathcal{F}_t, V_{t,T}\right] &= \mathbb{E}\left[\xi_T\frac{X_{t,T}}{V_{t,T}}|\mathcal{F}_t, V_{t,T}\right]. \end{aligned} \tag{4.12.2}$$

Posons

$$\begin{aligned} h_t(v) &= \mathbb{E}\left[\xi_T|\mathcal{F}_t, V_{t,T} = v\right], \\ g_t(v) &= \mathbb{E}\left[\xi_T\frac{X_{t,T}}{V_{t,T}}|\mathcal{F}_t, V_{t,T} = v\right]. \end{aligned}$$

On suppose que

$$\begin{aligned} h_t(v) &= \int_{\mathbb{R}_+} e^{yv} m_{h_t}(dy), \\ g_t(v) &= \int_{\mathbb{R}^+} e^{yv} m_{g_t}(dy), \end{aligned}$$

où m_{h_t} et m_{g_t} sont des mesures vérifant

$$\int_{\mathbb{R}_+} e^{yv}(|m_{h_t}| + |m_{g_t}|)(dy) < +\infty.$$

On construit de manière linéaire $P_t = P_v(h_t, g_t) = P_v(h_t, 0) + P_v(0, g_t)$

$$\begin{aligned} P_v(h_t, 0)(x) &= \int_{\mathbb{R}_+} \left\{\theta_0(y)e^{(1+\phi_+(y))x} + (1 - \theta_0(y))\, e^{-\phi_-(y)x}\right\} m_t(dy), \\ P_v(0, g_t)(x) &= \int_{\mathbb{R}_+} \left\{e^{(1+\phi_+(y))x} - e^{-\phi_-(y)x}\right\} \theta_1(y) m_t(dy), \end{aligned}$$

avec

$$\begin{aligned} \theta_0(y) &= \frac{\phi'_+(y)}{\phi'_+(y) + \phi'_-(y)} \\ &= \frac{\psi'(p_{\min} - \phi^-(y))}{\psi'(p_{\min} - \phi^-(y)) - \psi'(p_{\min} + \phi^+(y))}, \\ \theta_1(y) &= \frac{\phi'_+(y)\phi'_-(y)}{\phi'_+(y) + \phi'_-(y)} \\ &= \frac{1}{\psi'(1 + \phi^+(y)) - \psi'(-\phi^-(y))}. \end{aligned}$$

PROPOSITION 4.12.2. *On a*

$$\begin{aligned} \mathbb{E}\left[P_v(h_t, 0)(X_{t,T})|\mathcal{F}_t, V_{t,T} = v\right] &= h_t(v), \\ \mathbb{E}\left[P_v(0, g_t)\left(X_{t,T}\right)X_{t,T}|\mathcal{F}_t, V_{t,T} = v\right] &= g_t(v), \\ \mathbb{E}\left[P_v(h_t, 0)(X_{t,T})X_{t,T}|\mathcal{F}_t, V_{t,T}\right] = \mathbb{E}\left[P_v(0, g_t)\left(X_{t,T}\right)|\mathcal{F}_t, V_{t,T}\right] &= 0. \end{aligned}$$

Donc $P_v(h_t, g_t)$ vérifie (4.12.2).

Comme pour l'erreur sur la corrélation dans le cas continu (proposition (4.8.7)), on va obtenir une erreur à l'ordre 2 avec un choc sur la loi de X . On pose $X^\epsilon_{t,T} = L^\epsilon_{V_T} - L^\epsilon_{V_t}$ où L^ϵ est un processus de Lévy dont la loi est définie par

$$\psi_\epsilon(p) = \ln \frac{\mathbb{E}\left[e^{pL^\epsilon_t}\right]}{t} = \psi(p+\epsilon), \quad \forall p \in \mathbb{C}.$$

PROPOSITION 4.12.3. *En posant $X^\epsilon_{t,T} = X_{t,T} + \epsilon V_{t,T}$ et $\xi^\epsilon_T = H_t(X^\epsilon_{t,T}, V_{t,T}, Z)$, on a l'approximation au deuxième ordre en ϵ au voisinage de 0 :*

$$\mathbb{E}\left[\xi^\epsilon_T | \mathcal{F}_t, V_{t,T}\right] = \mathbb{E}\left[P_v(h_t, g_t)(X^\epsilon_{t,T}) | \mathcal{F}_t, V_{t,T}\right] + O(\epsilon^2).$$

4.12.2. Couverture en variance moyenne. En marché complet, le risque s'efface entièrement à l'aide d'une stratégie de couverture par un portefeuille de réplication. Cependant, en marché incomplet, il faut rajouter un critère supplémentaire pour déterminer une stratégie de couverture raisonnable (la couverture ne sera que partielle). Nous allons nous intéresser à une approche courante : la couverture en variance-moyenne. Il s'agit de minimiser l'erreur de couverture quadratique sur l'ensemble des stratégies admissibles :

$$\min_{Y \in \mathcal{A}} \mathbb{E}\left[(\xi_T - Y_T)^2\right]$$

où ξ_T est le payoff de l'option et $\mathcal{A}$ est l'ensemble des stratégies admissibles. Habituellement, les stratégies admissibles sont des portefeuilles qui traitent uniquement sur l'actif risqué S (voir par exemple [**Pha00**] et [**Sch01**]). Ici, l'ensemble $\mathcal{A}$ des stratégies admissibles est défini comme l'ensemble des portefeuilles qui traitent dynamiquement des options européennes sur l'actif risqué S de maturité T.

On regarde maintenant des dérivés sous la forme ξ_T vérifiant

$$\xi_T = h_t(V_{t,T}) + \int_{V_t}^{V_T} \delta_s\left(V_{s,T}\right) \sigma dW_s + \int_{\mathbb{R}} U_s(x, V_{s,T}) \tilde{\mu}(dx, ds) + M_{V_T} - M_{V_t}.$$

où M est une martingale vérifiant $[dM_s, dL_s] = 0$ et

$$\begin{aligned}
h_t(v) &= \mathbb{E}\left[\xi_T | \mathcal{F}_{V_t}, V_{t,T} = v\right], \\
\delta_t(v) &= \mathbb{E}\left[\frac{D_t \xi_T}{\sigma} | \mathcal{F}_{V_t}, V_{t,T} = v\right], \\
U_t(x,v) &= \mathbb{E}\left[\xi_T | \mathcal{F}_{V_t^-}, V_{t,T} = v, \Delta X_t = x\right] - \mathbb{E}\left[\xi_T | \mathcal{F}_{V_t^-}, V_{t,T} = v\right].
\end{aligned}$$

où D_t est la dérivée de Malliavin par rapport W (cf ([**Nua95**])).

On suppose que l'on a les décompositions à l'aide des fonctions exponentielles suivantes :

$$\begin{aligned}
h_t(v) &= \int_{\mathbb{R}_+} e^{yv} \alpha^h_t(y) m(dy), \\
\delta_t(v) &= \int_{\mathbb{R}^+} e^{yv} \alpha^\delta_t(y) m(dy), \\
U_t(x,v) &= \int_{\mathbb{R}^+} e^{yv} \alpha^x_t(y) m(dy),
\end{aligned}$$

où m est une mesure vérifiant

$$\int_{\mathbb{R}_+} e^{yv}\{|\alpha_t^h(y)| + |\alpha_t^\delta| + \int_{\mathbb{R}} |\alpha_t^x(y)|\nu(dx)\}|m|(dy) < +\infty, \quad \forall v > 0, \forall t \in [0,T].$$

On considère

$$P_c(h_t, \delta_t, U_t)(x) = \int_{\mathbb{R}_+} \left\{\theta_t(y)e^{(1+\phi^+(y))x} + \left(\alpha_t^h(y) - \theta_t(y)\right) e^{-\phi^-(y)x}\right\} m_t(dy)$$

avec

$$\begin{aligned}\theta_t(y) &= (1 + \phi_+(y) + \phi_-(y))\,(\alpha_{\delta_t}(y) + \alpha_{h_t}(y)\phi_-(y))\,\sigma^2/A(y) \\ &\quad + \int_{\mathbb{R}} \left(\alpha_{U_t}^x(y) - \alpha_{h_t}(y)\left(e^{-\phi_-(y)x} - 1\right)\right)\left(e^{(1+\phi_+(y))x} - e^{-\phi_-(y)x}\right)\nu(dx)/A(y), \\ A(y) &= |1 + \phi_+(y) + \phi_-(y)|^2\,\sigma^2 + \int_{\mathbb{R}} \left|e^{(1+\phi_+(y))x} - e^{-\phi_-(y)x}\right|^2 \nu(dx).\end{aligned}$$

La proposition suivante nous indique que notre couverture robuste est déterminée par la position dynamique à la date t par l'option européenne de payoff $P_t(X_{t,T})$.

PROPOSITION 4.12.4. *La réplication dynamique avec une position sur $P_c(h_t, \delta_t, U_t)(X_{t,T})$ à la date t couvre en moyenne variance (c'est-à-dire minimise l'erreur quadratique de réplication avec une couverture dynamique sur l'ensemble des options européennes de maturité T).*

En outre, l'option $P_c(h_t, \delta_t, U_t)(X_{t,T})$ permet de valoriser ξ_T :

$$\mathbb{E}\left[P_c(h_t, \delta_t, U_t)(X_{t,T})|\mathcal{F}_t, V_{t,T} = v\right] = \mathbb{E}\left[\xi_T|\mathcal{F}_t, V_{t,T} = v\right], \quad \forall v \geq 0.$$

4.12.3. Marché complet. On se place dans le cadre d'un marché complet, c'est-dire que le processus de Lévy L est

◇ soit un mouvement Brownien (dans le cas continu) et V_t représente la variance inconnue

◇ soit un processus de Poisson et V_t représente l'intensité inconnue.

Comme dans la première partie, on va s'intéresser aux options de la forme $\xi_T = H(X_T, V_T) = H_t(X_{t,T}, V_{t,T})$ (avec H_t $\mathcal{F}_t$-mesurable) par souci de simplication (on peut facilement étendre le résultat au cas où $\xi_T = H_t(X_{t,T}, V_{t,T}, Z)$ avec Z indépendant de (X_T, V_T) en prenant l'espérance conditionnelle).

4.12.3.1. *Cas continu.* Dans le cas continu, on retrouve la stratégie et les résultats de la première partie.

PROPOSITION 4.12.5. *On suppose que $X_t = B_{V_t} - \frac{V_t}{2}$ où B est un mouvement Brownien standard indépendant de la variance $V_t = \int_0^t \sigma_s^2 ds$. Le payoff de couverture robuste en moyenne variance de $\xi_T = H(X_T, V_T)$ et celui pour sa valorisation robuste sont les mêmes : $P_v(h_t, g_t) = P_c(h_t, \delta_t, U_t)$. Avec la position dynamique sur $P_t(X_{t,T})$, on obtient la valorisation et la couverture pour toutes variances $V_{t,T}$ indépendantes de B.*

4.12.3.2. *Cas discontinu.*

PROPOSITION 4.12.6. *On suppose que $X_t = aN_{V_t} - e^a V_t$ où N est un processus de Poisson standard indépendant de l'intensité $V_t = \int_0^t \lambda_s ds$. Le payoff de couverture robuste en moyenne variance et celui pour la valorisation robuste sont les mêmes $P_v(h_t, g_t) = P_c(h_t, \delta_t, U_t)$. Avec une position sur $P_t(X_{t,T})$, on obtient la valorisation et la couverture pour toutes intensités $V_{t,T}$ indépendantes de N.*

4.12.4. Discrétisation du payoff. Ici, nous cherchons à décomposer la fonction $f_t : \mathbb{R}^+ \to \mathbb{R}$ sous la forme $v \in \mathbb{R}^+ \to \sum_{k\geq 0} a_k e^{kv}$ pour $f_t(v) = h_t(v), g_t(v), \delta_t(v)$ ou $U_t(x, v)$. Pour cela, on va faire une approximation uniforme comme dans la proposition (4.9.3) (une approximation par régression linéaire est aussi possible comme dans la proposition (4.9.4)).

PROPOSITION 4.12.7. *On suppose que f_t est continue et que $V_{t,T}$ bornée par m, on a $1 \leq e^{V_{t,T}} \leq 1 + M = e^m$. En notant*

$$b_{n,k}^{f_t} = \sum_{i=0}^{n} f_t(\ln(1 + \frac{iM}{n}))(-1)^{k+i} \sum_{j=i}^{n} \frac{1_{j\geq k}}{M^j} \binom{n}{j} \binom{j}{k} \binom{j}{i},$$

$B_n^{f_t} = \sum_{k=0}^{n} b_{n,k}^{f_t} e^{kv}$ *converge uniformément sur* $[0, m]$.

On approxime alors $P_v(h_t, g_t)$ et $P_c(h_t, \delta_t, U_t)$ par $P_v(B_n^{h_t}, B_n^{g_t})$ et par $P_c(B_n^{h_t}, B_n^{\delta_t}, B_n^{U_t})$.

4.13. Preuves

Preuve proposition (4.8.6).

DÉMONSTRATION. Supposons qu'il existe $g \in \mathbb{H}_t(\mathbb{R})$ vérifiant : $\forall v > 0$

$$\begin{aligned}\mathbb{E}\left[g\left(X_{t,T}\right)|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right] &= h_t(v),\\ \frac{\partial}{\partial x}\mathbb{E}\left[g\left(X_{t,T}\right)|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right] &= \delta_t(v).\end{aligned}$$

Comme $X_{t,T} = \sqrt{v}N - \frac{v}{2}$ avec $N \sim \mathcal{N}(0;1)$, on a $\forall v \geq 0$

$$\begin{aligned}\frac{\partial}{\partial x}\mathbb{E}\left[g\left(X_{t,T}\right)|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right] &= \mathbb{E}\left[g'\left(X_{t,T}\right)|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right]\\ &= \frac{1}{v}\mathbb{E}\left[g\left(X_{t,T}\right)\left(X_{t,T} + \frac{\langle X\rangle_{t,T}}{2}\right)|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right],\end{aligned}$$

et

$$\begin{aligned}\int_{\mathbb{R}} g\left(\sqrt{v}x' - \frac{v}{2}\right)e^{-\frac{x'^2}{2}}\frac{dx'}{\sqrt{2\pi}} &= h_t(v),\\ \int_{\mathbb{R}} g'\left(\sqrt{v}x' - \frac{v}{2}\right)e^{-\frac{x'^2}{2}}\frac{dx'}{\sqrt{2\pi}} &= \delta_t(v).\end{aligned}$$

A l'aide du changement de variable $x = x' - \frac{\sqrt{v}}{2}$, on a :

$$\begin{aligned}\int_{\mathbb{R}} g\left(\sqrt{v}x\right)e^{-\frac{\sqrt{v}x}{2}}e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}} &= e^{\frac{v}{8}}h_t(v),\\ \int_{\mathbb{R}} g'\left(\sqrt{v}x\right)e^{-\frac{\sqrt{v}x}{2}}e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}} &= e^{\frac{v}{8}}\delta_t(v).\end{aligned}$$

On fait une intégration par partie à la deuxième égalité :

$$\int_{\mathbb{R}} g'\left(\sqrt{v}x\right)e^{-\frac{\sqrt{v}x}{2}}e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}} = \int_{\mathbb{R}} g\left(\sqrt{v}x\right)e^{-\frac{\sqrt{v}x}{2}}e^{-\frac{x^2}{2}}\left(\frac{x}{\sqrt{v}} + \frac{1}{2}\right)\frac{dx}{\sqrt{2\pi}}.$$

D'où

$$\begin{aligned}\int_{\mathbb{R}} g\left(\sqrt{v}x\right)e^{-\frac{\sqrt{v}x}{2}}e^{-\frac{x^2}{2}}dx &= \sqrt{2\pi}e^{\frac{v}{8}}h_t(v),\\ \int_{\mathbb{R}} \sqrt{v}x g\left(\sqrt{v}x\right)e^{-\frac{\sqrt{v}x}{2}}e^{-\frac{x^2}{2}}dx &= \sqrt{2\pi}ve^{\frac{v}{8}}\delta_t(v) - \frac{\sqrt{2\pi}}{2}ve^{\frac{v}{8}}h_t(v).\end{aligned}$$

En décomposant $g(x)e^{-\frac{x}{2}}$ par parité, on obtient les fonctions $G_i, G_p : \mathbb{R}^+ \to \mathbb{R}$:

$$g(x)e^{-\frac{x}{2}} = |x|\,G_p(x^2) + \operatorname{sign}(x)G_i(x^2).$$

On peut avoir $G_p(0) = \infty$. Or,

$$\begin{aligned}\mathbb{E}\left[|g(X_{t,T})|\,|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right] &= \int_{\mathbb{R}} \left|\sqrt{v}\,|x|\,G_p(vx^2) + \operatorname{sign}(x)G_i(vx^2)\right|e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}\\ &= \int_{\mathbb{R}^+} \left|\sqrt{v}xG_p(vx^2) + G_i(vx^2)\right|e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}\\ &\quad + \int_{\mathbb{R}^+} \left|\sqrt{v}xG_p(vx^2) - G_i(vx^2)\right|e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}\end{aligned}$$

et

$$\mathbb{E}\left[|X_{t,T}g(X_{t,T})|\,|\mathcal{F}_t, \langle X\rangle_{t,T} = v\right] = \int_{\mathbb{R}} \left|vx^2G_p(vx^2) + vxG_i(vx^2)\right| e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}$$
$$= \int_{\mathbb{R}^+} \left|vx^2G_p(vx^2) + vxG_i(vx^2)\right| e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}$$
$$+ \int_{\mathbb{R}^+} \left|vx^2G_p(vx^2) - vxG_i(vx^2)\right| e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}$$

On en déduit

$$\sup\left\{\left||x|\,G_p(x^2)e^{\frac{x}{2}}\right|_{\mathbb{H}_t}, \left|G_i(x^2)e^{\frac{x}{2}}\right|_{\mathbb{H}_t}\right\} \le |g|_{\mathbb{H}_t}$$
$$|g|_{\mathbb{H}_t} \le \left||x|\,G_p(x^2)e^{\frac{x}{2}}\right|_{\mathbb{H}_t} + \left|G_i(x^2)e^{\frac{x}{2}}\right|_{\mathbb{H}_t}$$

d'où une équivalence entre $g \in \mathbb{H}_t$ et $|x|\,G_p(x^2)e^{\frac{x}{2}}, \mathrm{sign}(x)G_i(x^2)e^{\frac{x}{2}} \in \mathbb{H}_t$.
La symétrie de la gaussienne et la parité permettent d'annuler les termes suivants

$$\int_{\mathbb{R}} \mathrm{sign}(\sqrt{v}x)G_i(vx^2)e^{-\frac{x^2}{2}}dx = \int_{\mathbb{R}} \sqrt{v}x\left|\sqrt{v}x\right| G_p(vx^2)e^{-\frac{x^2}{2}}dx = 0$$

d'où

$$\int_{\mathbb{R}} \left|\sqrt{v}x\right| G_p(vx^2)e^{-\frac{x^2}{2}}dx = \sqrt{2\pi}e^{\frac{v}{8}}h_t(v)$$
$$\int_{\mathbb{R}} \left|\sqrt{v}x\right| G_i(vx^2)e^{-\frac{x^2}{2}}dx = \sqrt{2\pi}ve^{\frac{v}{8}}\delta_t(v) - \frac{\sqrt{2\pi}}{2}ve^{\frac{v}{8}}h_t(v)$$

Par symétrie, on a :

$$2\int_0^{+\infty} \sqrt{v}xG_p(vx^2)e^{-\frac{x^2}{2}}dx = \sqrt{2\pi}e^{\frac{v}{8}}h_t(v),$$
$$2\int_0^{+\infty} \sqrt{v}xG_i(vx^2)e^{-\frac{x^2}{2}}dx = \sqrt{2\pi}ve^{\frac{v}{8}}\delta_t(v) - \frac{\sqrt{2\pi}}{2}ve^{\frac{v}{8}}h_t(v).$$

En faisant le changement de variable $t = vx^2$ et avec $\lambda = \frac{1}{2v}$, on obtient :

$$\mathcal{L}G_p(\lambda) = \int_0^{+\infty} G_p(t)e^{-\lambda t}dt = \sqrt{\pi}e^{\frac{1}{16\lambda}}h_t\left(\frac{1}{2\lambda}\right),$$
$$\mathcal{L}G_i(\lambda) = \int_0^{+\infty} G_i(t)e^{-\lambda t}dt = \frac{\sqrt{\pi}}{2}\lambda^{-\frac{3}{2}}e^{\frac{1}{16\lambda}}\delta_t\left(\frac{1}{2\lambda}\right) - \frac{\sqrt{\pi}}{4}\lambda^{-\frac{3}{2}}e^{\frac{1}{16\lambda}}h_t\left(\frac{1}{2\lambda}\right).$$

et

$$g(x) = P(h_t, 0)(x) + P(0, \delta_t)(x)$$

avec $P(h_t,0)(x) = |x|\,G_p(x^2)e^{\frac{x}{2}} \in \mathbb{H}_t$ et $P(0,\delta_t)(x) = \mathrm{sign}(x)G_i(x^2)e^{\frac{x}{2}} \in \mathbb{H}_t$.
Réciproquement, si $P(h_t, \delta_t) \in \mathbb{H}_t$, alors $P(h_t, \delta_t)$ vérifie (4.8.1). □

Preuve de proposition (4.8.7).

Démonstration. On a

$$S_T^{(\rho)} = S_t^{(\rho)} \exp\left(\sqrt{1-\rho^2}\int_t^T \sigma_s dW_s^1 + \rho\int_t^T \sigma_s dW_s^2 - \frac{\langle X\rangle_{t,T}}{2}\right)$$
$$= S_t^{(\rho)}e^{X_{t,T}} \exp\left((\sqrt{1-\rho^2}-1)\int_t^T \sigma_s dW_s^1 + \rho\int_t^T \sigma_s dW_s^2\right)$$

où W^1 et W^2 sont des mouvements Brownien $\mathcal{F}_t$-adaptés, σ et W^2 sont adaptés à une filtration $\mathcal{H}_t \subseteq \mathcal{F}_t$ avec $\mathcal{H}_t$ et $\mathcal{F}_t^{W^1}$ indépendants et où $B \sim \mathcal{N}(0,1)$ est indépendant de $\mathcal{F}_t$ et de $\mathcal{H}_T$. On en déduit que

$$\begin{aligned}\mathbb{E}_t\left[H\left(\ln S_T^{(\rho)}, \langle X\rangle_T\right)\right] &= \mathbb{E}_t\left[\mathbb{E}\left[H\left(\ln S_t^{(\rho)} + X_{t,T}, \langle X\rangle_T\right)|\mathcal{F}_t, \mathcal{H}_T\right]|\mathcal{F}_t\right] \\ &\quad + \mathbb{E}_t\left[\rho\int_t^T \sigma_s dW_s^2 \mathbb{E}\left[\frac{\partial}{\partial x}H\left(\ln S_t^{(\rho)} + X_{t,T}, \langle X\rangle_T\right)|\mathcal{F}_t, \mathcal{H}_T\right]\right] + O(\rho^2) \\ &= \mathbb{E}_t\left[h_t(\langle X\rangle_{t,T})|\mathcal{F}_t\right] + \rho\mathbb{E}_t\left[\delta_t(\langle X\rangle_{t,T})\int_t^T \sigma_s dW_s^2|\mathcal{F}_t\right] + O(\rho^2)\end{aligned}$$

et

$$\begin{aligned}\mathbb{E}_t\left[P_t\left(\ln(\frac{S_T^{(\rho)}}{S_t^{(\rho)}})\right)\right] &= \mathbb{E}_t\left[\mathbb{E}\left[P_t(X_{t,T}) + \frac{\partial}{\partial x}P_t(X_{t,T})\rho\int_t^T \sigma_s dW_s^2|\mathcal{F}_t, \mathcal{H}_T\right]|\mathcal{F}_t\right] + O(\rho^2) \\ &= \mathbb{E}_t\left[h_t(\langle X\rangle_{t,T})|\mathcal{F}_t\right] + \rho\mathbb{E}_t\left[\delta_t(\langle X\rangle_{t,T})\int_t^T \sigma_s dW_s^2|\mathcal{F}_t\right] + O(\rho^2).\end{aligned}$$

ce qui permet d'obtenir le résultat souhaité :

$$\mathbb{E}_t\left[P_t\left(\ln(\frac{S_T^{(\rho)}}{S_t^{(\rho)}})\right)\right] = \mathbb{E}_t\left[H\left(\ln S_T^{(\rho)}, \langle X\rangle_T\right)\right] + O(\rho^2).$$

□

Preuve du lemme (4.9.1).

Démonstration. Soit $\alpha > -\frac{1}{2}$ et $\alpha' > -\frac{3}{2}$. Considérons $\forall v > 0$

$$\begin{aligned}h_t(v) &= v^\alpha e^{\beta v - \frac{\theta}{v}}, \\ \delta_t(v) &= v^{\alpha'} e^{\beta v - \frac{\theta}{v}}.\end{aligned}$$

Soit $x \in \mathbb{R}$. D'après la proposition (4.8.6), on a

$$\begin{aligned}P(h_t, 0)(x) &= \sqrt{\pi}e^{\frac{x}{2}}|x|\,\mathcal{L}^{-1}\left(\frac{e^{\frac{1}{16\lambda}}}{\lambda^{\frac{1}{2}}}h_t\left(\frac{1}{2\lambda}\right)\right)(x^2) - \frac{\sqrt{\pi}}{4}e^{\frac{x}{2}}\mathrm{sign}(x)\mathcal{L}^{-1}\left(\frac{e^{\frac{1}{16\lambda}}}{\lambda^{\frac{3}{2}}}h_t\left(\frac{1}{2\lambda}\right)\right)(x^2) \\ &= \frac{\sqrt{\pi}}{2^\alpha}e^{\frac{x}{2}}|x|\,\mathcal{L}^{-1}\left(\frac{e^{\left(\frac{\beta}{2}+\frac{1}{16}\right)\frac{1}{\lambda}-2\theta\lambda}}{\lambda^{\alpha+\frac{1}{2}}}\right)(x^2) - \frac{\sqrt{\pi}}{2^{\alpha+2}}e^{\frac{x}{2}}\mathrm{sign}(x)\mathcal{L}^{-1}\left(\frac{e^{\left(\frac{\beta}{2}+\frac{1}{16}\right)\frac{1}{\lambda}-2\theta\lambda}}{\lambda^{\alpha+\frac{3}{2}}}\right)(x^2).\end{aligned}$$

Une propriété de la fonction de Bessel modifiée de première espèce I_ν (voir [**AS72**])

$$\mathcal{L}\left(t \to \left(\frac{t}{a}\right)^{\frac{\nu-1}{2}} I_{\nu-1}(2\sqrt{at})\right)(p) = \frac{e^{\frac{a}{p}}}{p^\nu}, \quad \nu > 0.$$

et la propriété de translation de la transformée de Laplace $\mathcal{L}(f(t)^{-at}) = \mathcal{L}f(.+a)$ impliquent

$$P(h_t,0)(x) = \sqrt{\frac{\pi}{2}} e^{\frac{x}{2}} |x| \left(\frac{x^2-2\theta}{2\beta+\frac{1}{4}}\right)^{\frac{\alpha-\frac{1}{2}}{2}} I_{\alpha-\frac{1}{2}}\left(\sqrt{(2\beta+\frac{1}{4})(x^2-2\theta)}\right) 1_{x^2\geq 2\theta}.$$
$$- \frac{\sqrt{\pi}}{2\sqrt{2}} e^{\frac{x}{2}} \text{sign}(x) \left(\frac{x^2-2\theta}{2\beta+\frac{1}{4}}\right)^{\frac{\alpha+\frac{1}{2}}{2}} I_{\alpha+\frac{1}{2}}\left(\sqrt{(2\beta+\frac{1}{4})(x^2-2\theta)}\right) 1_{x^2\geq 2\theta}.$$

De la même manière, on obtient

$$\begin{aligned} P(0,\delta_t)(x) &= \frac{\sqrt{\pi}}{2} e^{\frac{x}{2}} \text{sign}(x) \mathcal{L}^{-1}\left(\frac{e^{\frac{1}{16\lambda}}}{\lambda^{\frac{3}{2}}} \delta_t\left(\frac{1}{2\lambda}\right)\right)(x^2) \\ &= \sqrt{\frac{\pi}{2}} e^{\frac{x}{2}} \text{sign}(x) \left(\frac{x^2-2\theta}{2\beta+\frac{1}{4}}\right)^{\frac{\alpha'+\frac{1}{2}}{2}} I_{\alpha'+\frac{1}{2}}\left(\sqrt{(2\beta+\frac{1}{4})(x^2-2\theta)}\right) 1_{x^2\geq 2\theta} \end{aligned}$$

Pour $\theta = 0$ et $\beta = -\frac{1}{8}$, en utilisant un équivalent au voisinage de 0 de la fonction de Bessel modifiée de première espèce I_ν

$$\frac{I_\nu(z)}{\left(\frac{z}{2}\right)^\nu} = \frac{1}{\Gamma(\nu+1)} + O(z),$$

on obtient

$$P(v^\alpha e^{-\frac{v}{8}},0)(x) = \frac{\sqrt{\pi}}{2^\alpha} e^{\frac{x}{2}} |x|^{2\alpha} \left(\frac{1}{\Gamma(\alpha-\frac{1}{2})} - \frac{x}{4\Gamma(\alpha+\frac{1}{2})}\right),$$
$$P(0,v^{\alpha'} e^{-\frac{v}{8}})(x) = \frac{\sqrt{\pi}}{2^{\alpha+1}\Gamma(\alpha+\frac{1}{2})} e^{\frac{x}{2}} x|x|^{2\alpha'}.$$

□

Preuve de la proposition (4.9.2).

Démonstration. Soit $H(X_T, \langle X\rangle_T) = e^{\lambda\langle X\rangle_T + \mu X_T}$, alors en notant $B \sim \mathcal{N}(0,1)$

$$\begin{aligned} h_t(v) &= \mathbb{E}\left[e^{\lambda\langle X\rangle_T+\mu X_T} | \mathcal{F}_t, \langle X\rangle_{t,T} = v\right] \\ &= e^{\lambda\langle X\rangle_t+\mu X_t} \mathbb{E}\left[e^{\mu\sqrt{v}B + (\lambda-\frac{\mu}{2}\lambda)v}\right] \\ &= e^{\lambda\langle X\rangle_t+\mu X_t} e^{(\lambda+\frac{\mu^2-\mu}{2})v}, \\ \delta_t(v) &= \frac{\partial}{\partial x}\mathbb{E}\left[e^{\lambda\langle X\rangle_T+\mu X_T} | \mathcal{F}_t, \langle X\rangle_{t,T} = v\right] \\ &= \mu\mathbb{E}\left[e^{\lambda\langle X\rangle_T+\mu X_T} | \mathcal{F}_t, \langle X\rangle_{t,T} = v\right] \\ &= \mu e^{\lambda\langle X\rangle_t+\mu X_t} e^{(\lambda+\frac{\mu^2-\mu}{2})v}. \end{aligned}$$

En appliquant le lemme (4.9.1) à la fonction $v \mapsto v^{\alpha} e^{\beta v - \frac{\theta}{v}}$ (avec $\alpha = \theta = 0$, $\beta = \lambda + \frac{\mu^2 - \mu}{2}$), on obtient

$$P(e^{\beta v}, 0)(x) = \sqrt{\frac{\pi}{2}} e^{\frac{x}{2}} \sqrt{|x|} \left(2\beta + \frac{1}{4}\right)^{\frac{1}{4}} I_{-\frac{1}{2}}\left(|x| \sqrt{2\beta + \frac{1}{4}}\right) - \frac{\sqrt{\pi}}{2\sqrt{2}} e^{\frac{x}{2}} \operatorname{sign}(x) \sqrt{|x|} \left(2\beta + \frac{1}{4}\right)^{-\frac{1}{4}} I_{\frac{1}{2}}\left(|x| \sqrt{2\beta + \frac{1}{4}}\right),$$

$$P(0, e^{\beta v})(x) = \sqrt{\frac{\pi}{2}} e^{\frac{x}{2}} \operatorname{sign}(x) \sqrt{|x|} \left(2\beta + \frac{1}{4}\right)^{-\frac{1}{4}} I_{\frac{1}{2}}\left(|x| \sqrt{2\beta + \frac{1}{4}}\right).$$

Or,

$$I_{-\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi}} \cosh(x) \quad \text{et} \quad I_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi}} \sinh(x),$$

d'où

$$P(h_t, 0)(x) = e^{\lambda \langle X \rangle_t + \mu X_t} \left\{ e^{\frac{x}{2}} \cosh\left(xp(\mu, \lambda)\right) - \frac{e^{\frac{x}{2}} \sinh\left(xp(\mu, \lambda)\right)}{2p(\mu, \lambda)} \right\},$$

$$P(0, \delta_t)(x) = e^{\lambda \langle X \rangle_t + \mu X_t} \mu e^{\frac{x}{2}} \frac{\sinh\left(xp(\mu, \lambda)\right)}{p(\mu, \lambda)},$$

$$\begin{aligned} P_t(x) &= e^{\lambda \langle X \rangle_t + \mu X_t} e^{\frac{x}{2}} \left\{ \cosh\left(xp(\mu, \lambda)\right) + \frac{\mu - \frac{1}{2}}{p(\mu, \lambda)} \sinh\left(xp(\mu, \lambda)\right) \right\} \\ &= e^{\lambda \langle X \rangle_t + \mu X_t} \left\{ \left(\frac{1}{2} + \theta(\mu, \lambda)\right) e^{\left(\frac{1}{2} + p(\mu, \lambda)\right)x} + \left(\frac{1}{2} - \theta(\mu, \lambda)\right) e^{\left(\frac{1}{2} - p(\mu, \lambda)\right)x} \right\}, \end{aligned}$$

avec $p(\mu, \lambda) = \sqrt{\frac{1}{4} + 2\lambda + \mu^2 - \mu} = \sqrt{2\beta + \frac{1}{4}}$ et $\theta(\mu, \lambda) = \frac{\mu - \frac{1}{2}}{2p(\mu, \lambda)}$. □

Preuve de la proposition (4.9.3).

DÉMONSTRATION. Les approximations de $v \to e^{\frac{v}{8}} h_t(v)$ et de $v \to e^{\frac{v}{8}} \delta_t(v)$ par les n-ième polynômes de Bernstein $B_n^{h_t}(v) = \sum_{k=0}^n b_{n,k}^{h_t} v^k$ et $B_n^{\delta_t}(v) = \sum_{k=0}^n b_{n,k}^{\delta_t} v^k$ impliquent que $h_t(v) = \lim_{n \to +\infty} B_n^{h_t}(v) e^{-\frac{v}{8}}$ et $\delta_t(v) = \lim_{n \to +\infty} B_n^{\delta_t}(v) e^{-\frac{v}{8}}$ uniformément sur $[0, M]$.

Or, d'après le lemme (4.9.1),

$$P(v^k e^{-\frac{v}{8}}, 0)(x) = \frac{\sqrt{\pi}}{2^k} e^{\frac{x}{2}} x^{2k} \left(\frac{1}{\Gamma(k - \frac{1}{2})} - \frac{x}{4\Gamma(k + \frac{1}{2})}\right),$$

$$P(0, v^k e^{-\frac{v}{8}})(x) = \frac{\sqrt{\pi}}{2^{k+1} \Gamma(k + \frac{1}{2})} e^{\frac{x}{2}} x^{2k+1}.$$

Avec

$$\begin{aligned} P_{n,t}(x) &= \sqrt{\pi} e^{\frac{x}{2}} \sum_{k=0}^n \left(b_{n,k}^{h_t} \frac{x^{2k}}{2^k} \{ \frac{1}{\Gamma(k - \frac{1}{2})} - \frac{x}{4\Gamma(k + \frac{1}{2})} \} + b_{n,k}^{\delta_t} \frac{x^{2k+1}}{2^{k+1} \Gamma(k + \frac{1}{2})} \right) \\ &= P\left(B_n^{h_t}(v) e^{-\frac{v}{8}}, B_n^{\delta_t}(v) e^{-\frac{v}{8}}\right), \end{aligned}$$

on obtient alors

$$\lim_{n\to+\infty}\mathbb{E}\left[P_{n,t}(X_{t,T})|\mathcal{F}_t,\langle X\rangle_{t,T}\right]=\lim_{n\to+\infty}B_n^{h_t}(\langle X\rangle_{t,T})e^{-\frac{\langle X\rangle_{t,T}}{8}}=h_t(\langle X\rangle_{t,T}),$$
$$\lim_{n\to+\infty}\frac{\partial}{\partial x}\mathbb{E}\left[P_{n,t}(X_{t,T})|\mathcal{F}_t,\langle X\rangle_{t,T}\right]=\lim_{n\to+\infty}B_n^{\delta_t}(\langle X\rangle_{t,T})e^{-\frac{\langle X\rangle_{t,T}}{8}}=\delta_t(\langle X\rangle_{t,T}).$$

□

Preuve de la proposition (4.9.4).

DÉMONSTRATION. L'espace des polynômes est dense dans $[0,M]$ dans $\mathcal{C}([0,M])$ avec la norme uniforme et par conséquent avec la norme $\mathbb{L}^2(\mu)$. En notant $A_n^{h_t}(v)=\sum_{k=0}^n a_{n,k}^{h_t}v^k$ et $A_n^{\delta_t}(v)=\sum_{k=0}^n a_{n,k}^{\delta_t}v^k$, on a alors $h_t(v)e^{\frac{v}{8}}=\lim_{n\to+\infty}A_n^{h_t}(v)$ et $\delta_t(v)e^{\frac{v}{8}}=\lim_{n\to+\infty}A_n^{\delta_t}(v)$ sous la norme $\mathbb{L}^2(\mu)$.

Or, d'après le lemme (4.9.1),

$$P(v^ke^{-\frac{v}{8}},0)(x)=\frac{\sqrt{\pi}}{2^k}e^{\frac{x}{2}}x^{2k}\left(\frac{1}{\Gamma(k-\frac{1}{2})}-\frac{x}{4\Gamma(k+\frac{1}{2})}\right),$$
$$P(0,v^ke^{-\frac{v}{8}})(x)=\frac{\sqrt{\pi}}{2^{k+1}\Gamma(k+\frac{1}{2})}e^{\frac{x}{2}}x^{2k+1}.$$

Donc, avec

$$\begin{aligned}P_{n,t}(x)&=\sqrt{\pi}e^{\frac{x}{2}}\sum_{k=0}^n\left(a_{n,k}^{h_t}\frac{x^{2k}}{2^k}\{\frac{1}{\Gamma(k-\frac{1}{2})}-\frac{x}{4\Gamma(k+\frac{1}{2})}\}+a_{n,k}^{\delta_t}\frac{x^{2k+1}}{2^{k+1}\Gamma(k+\frac{1}{2})}\right)\\&=P\left(A_n^{h_t}(v)e^{-\frac{v}{8}},A_n^{\delta_t}(v)e^{-\frac{v}{8}}\right),\end{aligned}$$

on obtient

$$\lim_{n\to+\infty}\mathbb{E}_t\left[P_{n,t}(X_{t,T})\right]=\lim_{n\to+\infty}\mathbb{E}_t\left[A_n^{h_t}(\langle X\rangle_{t,T})e^{-\frac{\langle X\rangle_{t,T}}{8}}\right]=\mathbb{E}_t\left[h_t(\langle X\rangle_{t,T})\right],$$
$$\lim_{n\to+\infty}\frac{\partial}{\partial x}\mathbb{E}_t\left[P_{n,t}(X_{t,T})\right]=\lim_{n\to+\infty}\mathbb{E}_t\left[A_n^{\delta_t}(\langle X\rangle_{t,T})e^{-\frac{\langle X\rangle_{t,T}}{8}}\right]=\mathbb{E}_t\left[\delta_t(\langle X\rangle_{t,T})\right].$$

□

Preuve de la proposition (4.10.1).

DÉMONSTRATION. Pour le swap de volatilité $\sqrt{\langle X\rangle_T}$, on a

$$\begin{aligned}h_t(v)&=\sqrt{\langle X\rangle_t+v},\quad\forall v\geq 0,\\\delta_t(v)&=0.\end{aligned}$$

Le lemme (4.9.1) appliqué $v\in\mathbb{R}_+^*\to\sqrt{v}=v^\alpha e^{\beta v-\frac{\theta}{v}}$ ($\alpha=\frac{1}{2}$ et $\beta=\theta=0$) donne à la date $t=0$ ($\langle X\rangle_t=0$) :

$$P_0(x)=\sqrt{\frac{\pi}{2}}e^{\frac{x}{2}}|x|\left(I_0\left(\frac{x}{2}\right)-I_1\left(\frac{x}{2}\right)\right).$$

A la date t, l'identité

$$\sqrt{q}=\frac{1}{2\sqrt{\pi}}\int_0^{+\infty}\frac{1-e^{-sq}}{s^{\frac{3}{2}}}ds$$

permet d'obtenir la décomposition

$$\begin{aligned}\sqrt{\langle X\rangle_t + v} &= v^{\frac{1}{2}}\sqrt{\frac{1}{v}\langle X\rangle_t + 1} \\ &= \frac{1}{2\sqrt{\pi}}\int_0^{+\infty} \frac{v^{\frac{1}{2}} - e^{-s}v^{\frac{1}{2}}e^{-s\langle X\rangle_t\frac{1}{v}}}{s^{\frac{3}{2}}}ds.\end{aligned}$$

Il suffit alors d'appliquer le lemme (4.9.1) à $h(v) = v^{\frac{1}{2}}$ et $h(v) = v^{\frac{1}{2}}e^{-s\langle X\rangle_t\frac{1}{v}}$, puis d'intégrer. Finalement on obtient :

$$\begin{aligned}P(h_t,0)(x) &= \frac{e^{\frac{x}{2}}|x|}{2\sqrt{2}}\int_0^{+\infty}\left\{I_0\left(\frac{|x|}{2}\right) - e^{-s}I_0\left(\frac{\sqrt{x^2 - s\langle X\rangle_t}}{2}\right)1_{x^2\geq s\langle X\rangle_t}\right\}\frac{ds}{s^{\frac{3}{2}}} \\ &- \frac{e^{\frac{x}{2}}}{2\sqrt{2}}\text{sign}(x)\int_0^{+\infty}\left\{|x|\,I_1\left(\frac{|x|}{2}\right) - e^{-s}I_1\left(\frac{\sqrt{x^2 - s\langle X\rangle_t}}{2}\right)\sqrt{x^2 - s\langle X\rangle_t}1_{x^2\geq s\langle X\rangle_t}\right\}\frac{ds}{s^{\frac{3}{2}}}.\end{aligned}$$

Pour $a \geq 0$, on a

$$\int_0^{+\infty}\frac{1 - e^{-s}1_{s\leq a^2}}{s^{\frac{3}{2}}}ds = 2\sqrt{\pi} + 2\frac{e^{-a^2}}{a} - 4\int_a^{\infty}e^{-\frac{y^2}{2}}dy.$$

On note $\text{erfc}(a) = \frac{2}{\sqrt{\pi}}\int_a^{\infty}e^{-\frac{y^2}{2}}dy$. La formule et le changement de variable $u^2 = s\langle X\rangle_t$ permettent de réécrire l'égalité sous la forme :

$$\begin{aligned}P(h_t,0)(x) &= \sqrt{\frac{\pi}{2}}|x|\,e^{\frac{x}{2}}\left(I_0\left(\frac{x}{2}\right) - I_1\left(\frac{x}{2}\right)\right)\left\{1 + \frac{\sqrt{\langle X\rangle_t}}{\sqrt{\pi}|x|}e^{-\frac{x^2}{\langle X\rangle_t}} - \text{erfc}\left(\frac{|x|}{\sqrt{\langle X\rangle_t}}\right)\right\} \\ &+\sqrt{\langle X\rangle_t}\frac{e^{\frac{x}{2}}|x|}{\sqrt{2}}\int_0^{|x|}e^{-\frac{u^2}{\langle X\rangle_t}}\left\{I_0\left(\frac{|x|}{2}\right) - I_0\left(\frac{\sqrt{x^2-u^2}}{2}\right)\right\}\frac{du}{u^2} \\ &-\sqrt{\langle X\rangle_t}\frac{e^{\frac{x}{2}}}{\sqrt{2}}\text{sign}(x)\int_0^{|x|}e^{-\frac{u^2}{\langle X\rangle_t}}\left\{|x|\,I_1\left(\frac{|x|}{2}\right) - I_1\left(\frac{\sqrt{x^2-u^2}}{2}\right)\sqrt{x^2-u^2}\right\}\frac{du}{u^2}.\end{aligned}$$

La formule et le changement de variable $u = \sqrt{1 - s\frac{\langle X\rangle_t}{|x|^2}}$ permettent de réécrire l'égalité sous la forme :

$$\begin{aligned}P(h_t,0)(x) &= \sqrt{\frac{\pi}{2}}|x|\,e^{\frac{x}{2}}\left(I_0\left(\frac{x}{2}\right) - I_1\left(\frac{x}{2}\right)\right)\left\{1 + \frac{\sqrt{\langle X\rangle_t}}{\sqrt{\pi}|x|}e^{-\frac{x^2}{\langle X\rangle_t}} - \text{erfc}\left(\frac{|x|}{\sqrt{\langle X\rangle_t}}\right)\right\} \\ &- \sqrt{\langle X\rangle_t}\frac{e^{\frac{x}{2}}}{\sqrt{2}}\int_0^1 e^{\frac{x^2}{\langle X\rangle_t}(u^2-1)}\left\{\left(I_0\left(\frac{x}{2}\right) - I_0\left(\frac{ux}{2}\right)\right) - \left(I_1\left(\frac{x}{2}\right) - I_1\left(\frac{ux}{2}\right)u\right)\right\}\frac{udu}{(1-u^2)^{\frac{3}{2}}}.\end{aligned}$$

□

Preuve de la proposition (4.10.2).

DÉMONSTRATION. Pour le ratio de Sharpe $\frac{e^{X_T}-1}{\sqrt{\langle X\rangle_T}}$, on a

$$\begin{aligned} h_t(v) &= \mathbb{E}\left[\frac{e^{X_T}-1}{\sqrt{\langle X\rangle_T}}|\mathcal{F}_t, \langle X\rangle_{t,T}=v\right] \\ &= \frac{e^{X_t}-1}{\sqrt{\langle X\rangle_t+v}}, \\ \delta_t(v) &= \frac{\partial}{\partial x}\mathbb{E}\left[\frac{e^{X_T}-1}{\sqrt{\langle X\rangle_T}}|\mathcal{F}_t, \langle X\rangle_{t,T}=v\right] \\ &= \frac{e^{X_t}}{\sqrt{\langle X\rangle_t+v}}. \end{aligned}$$

Initialement $(t=0)$, on a

$$h_0(v)=0 \quad \text{et} \quad \delta_0(v)=v^{-\frac{1}{2}}.$$

D'où $P(h_0,0)=0$. Le lemme (4.9.1) implique

$$P(0,\delta_0)(x)=\sqrt{\frac{\pi}{2}}e^{\frac{x}{2}}\text{sign}(x)I_0\left(\frac{x}{2}\right)$$

A la date t avec $\langle X\rangle_t>0$, en utilisant l'identité

$$q^{-r}=\frac{r}{\Gamma(1-r)}\int_0^\infty e^{-s^{\frac{1}{r}}q}ds \quad r>0, q>0$$

avec $r=\frac{1}{2}$, on obtient :

$$\begin{aligned} (\langle X\rangle_t+v)^{-\frac{1}{2}} &= v^{-\frac{1}{2}}\left(1+\frac{\langle X\rangle_t}{v}\right)^{-\frac{1}{2}} \\ &= \frac{1}{2\sqrt{\pi}}\int_0^\infty e^{-\sqrt{s}}\frac{e^{-\frac{\sqrt{s}\langle X\rangle_t}{v}}}{\sqrt{v}}ds. \end{aligned}$$

Il suffit alors d'appliquer le lemme (4.9.1) avec $h(v)=\delta(v)=\frac{e^{-\frac{\sqrt{s}\langle X\rangle_t}{v}}}{\sqrt{v}}$, puis d'intégrer et de faire le changement de variable $u=\sqrt{1-\sqrt{s}\frac{\langle X\rangle_t}{x^2}}$.

$$\begin{aligned} P(h_t,0)(x) &= \left(e^{X_t}-1\right)2^{-\frac{5}{2}}e^{\frac{x}{2}}\int_0^{\frac{x^4}{\langle X\rangle_t^2}}e^{-\sqrt{s}}\{\frac{I_1\left(\frac{|x|}{2}\sqrt{1-\sqrt{s}\frac{\langle X\rangle_t}{x^2}}\right)}{\sqrt{1-\sqrt{s}\frac{\langle X\rangle_t}{x^2}}}-\text{sign}(x)I_0\left(\frac{|x|}{2}\sqrt{1-\sqrt{s}\frac{\langle X\rangle_t}{x^2}}\right)\}ds \\ &= \frac{\left(e^{X_t}-1\right)}{\sqrt{2}\langle X\rangle_t^2}x^2e^{\frac{x}{2}}\int_0^1 e^{\frac{(u^2-1)\langle X\rangle_t}{x^2}}\left\{I_1\left(\frac{u|x|}{2}\right)-\text{sign}(x)I_0\left(\frac{u|x|}{2}\right)u\right\}(1-u^2)du \end{aligned}$$

et

$$\begin{aligned} P(0,\delta_t)(x) &= e^{X_t}\frac{e^{\frac{x}{2}}\text{sign}(x)}{2\sqrt{2}}\int_0^{\frac{x^4}{\langle X\rangle_t^2}}e^{-\sqrt{s}}I_0\left(\frac{|x|}{2}\sqrt{1-\sqrt{s}\frac{\langle X\rangle_t}{x^2}}\right)ds \\ &= \sqrt{2}e^{X_t}\frac{x^2e^{\frac{x}{2}}\text{sign}(x)}{\langle X\rangle_t^2}\int_0^1 e^{\frac{(u^2-1)\langle X\rangle_t}{x^2}}I_0\left(\frac{u|x|}{2}\right)u(1-u^2)du. \end{aligned}$$

□

Preuve de la proposition (4.11.1).

DÉMONSTRATION. Le raisonnement est le même que celui de la preuve de la proposition (4.8.6). Il suffit juste de remplacer $X_{t,T} = \int_t^T \sigma_s dW_s^S - \frac{1}{2}\int_t^T \sigma_s^2 ds$ par $X_{t,T} = \int_t^T \sigma_s dW_s^S - \frac{b_S}{2}\int_t^T \sigma_s^2 ds$. Alors avec

$$\begin{aligned}
P(h_t,0)(\lambda) &= \sqrt{\pi}\frac{e^{\frac{b_S^2}{16}\frac{1}{\lambda}}}{\sqrt{\lambda}} h_t\left(\frac{1}{2\lambda}\right),\\
P(0,\delta_t)(\lambda) &= \frac{\sqrt{\pi}}{2}\lambda^{-\frac{3}{2}} e^{\frac{b_S^2}{16}\frac{1}{\lambda}}\delta_t\left(\frac{1}{2\lambda}\right) - \frac{\sqrt{\pi}b_S}{4}\lambda^{-\frac{3}{2}} e^{\frac{b_S^2}{16}\frac{1}{\lambda}} h_t\left(\frac{1}{2\lambda}\right).
\end{aligned}$$

$P_t(X_{t,T}) = P(h_t,0)(X_{t,T}) + P(0,\delta_t)(X_{t,T})$ vérifie :

$$\begin{aligned}
\mathbb{E}\left[P_t(X_{t,T})|\mathcal{F}_t, V_{t,T} = v\right] &= h_t(v),\\
\frac{\partial}{\partial x}\mathbb{E}\left[P_t(X_{t,T})|\mathcal{F}_t, V_{t,T} = v\right] &= \delta_t(v).
\end{aligned}$$

□

Preuve de la proposition (4.11.3).

DÉMONSTRATION. On a

$$\begin{aligned}
h_t(v) &= \mathbb{E}_t\left[H\left(F_T\right)|\mathcal{F}_t, V_{t,T} = v\right]\\
&= \mathbb{E}\left[P_{t,\xi}\left(\sqrt{v}N - \frac{b_F\beta}{2}v\right)|\mathcal{F}_t\right],\\
\delta_t\left(v\right) &= \mathbb{E}\left[\frac{D_t^W H(F_T)}{\sigma_t}|\mathcal{F}_t, V_{t,T} = v\right]\\
&= \rho\mathbb{E}\left[P'_{t,\xi}\left(\sqrt{v}N - \frac{b_F\beta}{2}v\right)|\mathcal{F}_t\right]
\end{aligned}$$

avec $N \sim \mathcal{N}(0,1)$ indépendant de $\mathcal{F}_t$. D'où

$$\begin{aligned}
h_t(v) &= \int_{\mathbb{R}^+}\left\{H_{t,\xi}\left(y\right)e^{-\frac{b_F\beta}{2}y} + H_{t,\xi}\left(-y\right)e^{\frac{b_F\beta}{2}y}\right\}e^{-\frac{y^2}{2v}-\frac{b_F^2\beta^2}{8}v}\frac{dy}{\sqrt{2\pi v}},\\
\delta_t\left(v\right) &= \rho\int_{\mathbb{R}^+}\left\{H'_{t,\xi}\left(y\right)e^{-\frac{b_F\beta}{2}y} + H'_{t,\xi}\left(-y\right)e^{\frac{b_F\beta}{2}y}\right\}e^{-\frac{y^2}{2v}-\frac{b_F^2\beta^2}{8}v}\frac{dy}{\sqrt{2\pi v}}.
\end{aligned}$$

D'après la proposition (4.11.1), on a

$$P_t(x) = e^{\frac{b_S}{2}x}\left|x\right|G_p(x^2) + e^{\frac{b_S}{2}x}\mathrm{sign}(x)G_i(x^2)$$

avec

$$\begin{aligned}
G_p &= \mathcal{L}^{-1}\left(\sqrt{\pi}\frac{e^{\frac{b_S^2}{16}\frac{1}{\lambda}}}{\sqrt{\lambda}}h_t\left(\frac{1}{2\lambda}\right)\right)\\
&= \int_{\mathbb{R}^+}\left\{H_{t,\xi}(y)\,e^{-\frac{b_F\beta}{2}y}+PH_{t,\xi}(-y)\,e^{\frac{b_F\beta}{2}y}\right\}\mathcal{L}^{-1}\left(e^{-y^2\lambda-\frac{b_F^2\beta^2-b_S^2}{16\lambda}}\right)dy,\\
G_i &= \mathcal{L}^{-1}\left(\frac{\sqrt{\pi}}{2}\lambda^{-\frac{3}{2}}e^{\frac{b_S^2}{16}\frac{1}{\lambda}}\delta_t\left(\frac{1}{2\lambda}\right)-\frac{\sqrt{\pi}b_S}{4}\lambda^{-\frac{3}{2}}e^{\frac{b_S^2}{16}\frac{1}{\lambda}}h_t\left(\frac{1}{2\lambda}\right)\right)\\
&= \int_{\mathbb{R}^+}\left\{\left(\rho H'_{t,\xi}(y)-\frac{b_SP_{t,\xi}(y)}{2}\right)e^{-\frac{b_F\beta}{2}y}+\left(\rho H'_{t,\xi}(-y)-\frac{b_SP_{t,\xi}(-y)}{2}\right)e^{\frac{b_F\beta}{2}y}\right\}\\
&\quad\times\mathcal{L}^{-1}\left(\frac{e^{-y^2\lambda-\frac{b_F^2\beta^2-b_S^2}{16\lambda}}}{\lambda}\right)\frac{dy}{2}.
\end{aligned}$$

Or, des propriétés de la transformation de Laplace de la fonction de Bessel de première espèce J_ν

$$\begin{aligned}
\mathcal{L}^{-1}\left(\frac{e^{-\frac{a}{\lambda}-b\lambda}}{\lambda}\right)(t) &= J_0(2\sqrt{a(t-b)})1_{t\geq b},\\
\mathcal{L}^{-1}\left(e^{-\frac{a}{\lambda}-b\lambda}\right)(t) &= \delta(t-b)-\sqrt{\frac{t-b}{a}}J_1(2\sqrt{a(t-b)})1_{t\geq b}.
\end{aligned}$$

permettent d'obtenir

$$\begin{aligned}
G_p(x^2) &= \frac{P_{t,\xi}(|x|)\,e^{-\frac{b_F\beta}{2}|x|}+P_{t,\xi}(-|x|)\,e^{\frac{b_F\beta}{2}|x|}}{2|x|}\\
&\quad-\int_{[0,|x|]}\left\{P_{t,\xi}(y)\,e^{-\frac{b_F\beta}{2}y}+P_{t,\xi}(-y)\,e^{\frac{b_F\beta}{2}y}\right\}\sqrt{x^2-y^2}J_1\left(2\mu_\beta\sqrt{x^2-y^2}\right)\frac{dy}{\mu_\beta},\\
G_i(x^2) &= \int_{[0,|x|]}\left\{\left(\rho P'_{t,\xi}(y)-\frac{b_SP_{t,\xi}(y)}{2}\right)e^{-\frac{b_F\beta}{2}y}+\left(\rho P'_{t,\xi}(-y)-\frac{b_SP_{t,\xi}(-y)}{2}\right)e^{\frac{b_F\beta}{2}y}\right\}\\
&\quad\times J_0\left(2\mu_\beta\sqrt{x^2-y^2}\right)\frac{dy}{2}
\end{aligned}$$

avec $\mu_\beta=\frac{\sqrt{b_F^2\beta^2-b_S^2}}{4}$. Finalement, on a le payoff de couverture ρ-neutre

$$\begin{aligned}
&P_t(x)=e^{\frac{b_S}{2}x}\frac{P_{t,\xi}(|x|)\,e^{-\frac{b_F\beta}{2}|x|}+P_{t,\xi}(-|x|)\,e^{\frac{b_F\beta}{2}|x|}}{2}\\
&-e^{\frac{b_S}{2}x}|x|\int_{[0,|x|]}\left\{P_{t,\xi}(y)\,e^{-\frac{b_F\beta}{2}y}+P_{t,\xi}(-y)\,e^{\frac{b_F\beta}{2}y}\right\}\sqrt{x^2-y^2}J_1\left(2\mu_\beta\sqrt{x^2-y^2}\right)\frac{dy}{\mu_\beta}\\
&+e^{\frac{b_S}{2}x}\mathrm{sign}(x)\int_{[0,|x|]}\left\{\left(\rho P'_{t,\xi}(y)-\frac{b_SP_{t,\xi}(y)}{2}\right)e^{-\frac{b_F\beta}{2}y}+\left(\rho P'_{t,\xi}(-y)-\frac{b_SP_{t,\xi}(-y)}{2}\right)e^{\frac{b_F\beta}{2}y}\right\}\\
&\quad\times J_0\left(2\mu_\beta\sqrt{x^2-y^2}\right)\frac{dy}{2}.
\end{aligned}$$

□

Preuve du corollaire (4.11.4).

DÉMONSTRATION. Le but ici est d'obtenir la réplication de l'option $(F_T - K)^+$. On va chercher le payoff de couverture robuste selon le signe de $F_t - K$.

Notons $k = \frac{a_F + b_K K}{a_F + b_F F_t}$.

Supposons $F_t \leq K$ (i.e. $k \geq 1$).

Il suffit de prendre

$$H_{t,\xi}(x) = \left(F_t + \frac{a_F}{b_F}\right)\left(e^{b_F \beta x} - k\right)^+.$$

Alors la proposition (4.11.3) appliquée avec $P_{t,\xi}$ donne le payoff de couverture $P_t(X_{t,T}) = \tilde{P}_t(S_T)$

$$\begin{aligned}
P_t(x) &= \left(F_t + \frac{a_F}{b_F}\right)\left(e^{b_F\beta|x|} - k\right)^+ \frac{e^{\frac{b_S}{2}x - \frac{b_F\beta}{2}|x|}}{2} \\
&- \left(F_t + \frac{a_F}{b_F}\right) e^{\frac{b_S}{2}x} |x| \int_{[0,|x|]} \left(e^{b_F\beta|y|} - k\right)^+ \sqrt{x^2 - y^2} J_1\left(2\mu_\beta \sqrt{x^2 - y^2}\right) \frac{dy}{\mu_\beta} \\
&+ \left(F_t + \frac{a_F}{b_F}\right) e^{\frac{b_S}{2}x} \text{sign}(x) \int_{[0,|x|]} 1_{y \geq \frac{\ln(k)}{b_F\beta}} \left(\left(\rho b_F \beta - \frac{b_S}{2}\right) e^{b_F\beta y} + \frac{b_S k}{2}\right) e^{-\frac{b_F\beta}{2}y} J_0\left(2\mu_\beta\sqrt{x^2 - y^2}\right) \frac{dy}{2} \\
\tilde{P}_t(x) &= H_t\left(\frac{1}{b_S}\ln\left(\frac{a_S + b_S x}{a_S + b_S S_t}\right)\right).
\end{aligned}$$

avec $\mu_\beta = \frac{\sqrt{b_F^2\beta^2 - b_S^2}}{4}$. Or, le fonction $\tilde{P}_t$ vérifie :

$$\begin{aligned}
\tilde{P}_t(S) &= 0 \quad \forall x \in [K_t^p, K_t^c], \\
\tilde{P}_t'\left(K_t^{c+}\right) &= \frac{a_F + b_F F_t}{a_s + b_s S_t}\frac{\beta}{2}(1+\rho)\, k^{-\frac{b_S}{2b_F\beta} + \frac{1}{2}}, \\
\tilde{P}_t'\left(K_t^{p-}\right) &= -\frac{a_F + b_F F_t}{a_s + b_s S_t}\frac{\beta}{2}(1-\rho)\, k^{\frac{b_S}{2b_F\beta} + \frac{1}{2}},
\end{aligned}$$

avec

$$\begin{aligned}
K_t^c &= \left(\frac{a_S}{b_S} + S_t\right) k^{\frac{b_S}{b_F\beta}} - \frac{a_S}{b_S} \\
&= \left(\frac{a_S}{b_S} + S_t\right) e^{\frac{b_S}{b_F\beta}\left|\ln\frac{L_F(K)}{L_F(F_t)}\right|} - \frac{a_S}{b_S}, \\
K_t^p &= \left(\frac{a_S}{b_S} + S_t\right) k^{-\frac{b_S}{b_F\beta}} - \frac{a_S}{b_S} \\
&= \left(\frac{a_S}{b_S} + S_t\right) e^{-\frac{b_S}{b_F\beta}\left|\ln\frac{L_F(K)}{L_F(F_t)}\right|} - \frac{a_S}{b_S}.
\end{aligned}$$

On peut approximer le payoff par :

$$P_t(X_{t,T}) = \tilde{P}_t(S_T) \simeq \tilde{P}_t'\left(K_t^{c,+}\right)(S_T - k_t^c)^+ - \tilde{H}_t'\left(K_t^{p-}\right)(K_t^p - S_T)^+.$$

□

Supposons $F_t \geq K$ (i.e. $k \leq 1$). On décompose le call sur fond sous la forme $(F_T - K)^+ = (F_T - K) + (K - F_T)^+$.

La couverture robuste (en moyenne variance) de $F_T - K$ est donnée par $(F_t - K) + \rho\beta \frac{a_F + b_F F_t}{a_S + b_S S_t} (S_T - S_t)$. En effet, il suffit d'appliquer la formule d'inversion (avec la proposition (4.11.3)) pour $h_t(v) = F_t - K$ et $\delta_t(v) = \rho b_F \beta \left(\frac{a_F}{b_F} + F_t \right)$.

DÉMONSTRATION. On cherche à couvrir la partie $(K - F_T)^+$. Avec

$$H^p_{t,\xi}(x) = \left(F_t + \frac{a_F}{b_F} \right) \left(k - e^{b_F \beta x} \right)^+$$

et en appliquant la proposition (4.11.3) avec $P^p_{t,\xi}(x)$, la couverture robuste dynamique de $(K - F_T)^+$ est alors déterminée à la date t par la position dynamique avec une option de payoff donnée par $H^p_t(X_{t,T}) = \tilde{H}^p_t(S_T)$. D'où

$$\begin{aligned}
H^p_t(x) &= \left(F_t + \frac{a_F}{b_F} \right) \left(k - e^{-b_F \beta |x|} \right)^+ \frac{e^{\frac{b_S}{2} x + \frac{b_F \beta}{2} |x|}}{2} \\
&- \left(F_t + \frac{a_F}{b_F} \right) e^{\frac{b_S}{2} x} |x| \int_{[0,|x|]} \left(k - e^{-b_F \beta |y|} \right)^+ \sqrt{x^2 - y^2} J_1 \left(2\mu_\beta \sqrt{x^2 - y^2} \right) \frac{dy}{\mu_\beta} \\
&+ \left(F_t + \frac{a_F}{b_F} \right) e^{\frac{b_S}{2} x} \text{sign}(x) \int_{[0,|x|]} 1_{y \geq -\frac{\ln(k)}{b_F \beta}} \left(\left(\frac{b_S}{2} - \rho b_F \beta \right) e^{-b_F \beta y} - \frac{k b_S}{2} \right) e^{\frac{b_F \beta}{2} y} J_0 \left(2\mu_\beta \sqrt{x^2 - y^2} \right) \frac{dy}{2} \\
\tilde{H}^p_t(S) &= H^p_t \left(\frac{1}{b_S} \ln \left(\frac{a_S + b_S S_T}{a_S + b_S S_t} \right) \right).
\end{aligned}$$

Or,

$$\begin{aligned}
\tilde{H}^p_t(S) &= 0 \quad \forall x \in [K^p_t, K^c_t], \\
\tilde{H}^{p\prime}_t \left(K^{c+}_t \right) &= (1 - \rho) \frac{\beta}{2} \frac{a_F + b_F F_t}{a_s + b_s S_t} k^{\frac{b_S}{2 b_F \beta} + \frac{1}{2}}, \\
\tilde{H}^{p\prime}_t \left(K^{p-}_t \right) &= -(1 + \rho) \frac{\beta}{2} \frac{a_F + b_F F_t}{a_s + b_s S_t} \frac{\beta}{2} k^{-\frac{b_S}{2 b_F \beta} + \frac{1}{2}},
\end{aligned}$$

$$\begin{aligned}
K^c_t &= \left(\frac{a_S}{b_S} + S_t \right) k^{-\frac{b_S}{b_F \beta}} - \frac{a_S}{b_S} \\
&= \left(\frac{a_S}{b_S} + S_t \right) e^{\frac{b_S}{b_F \beta} \left| \ln \frac{L_F(K)}{L_F(F_t)} \right|} - \frac{a_S}{b_S}, \\
K^p_t &= \left(\frac{a_S}{b_S} + S_t \right) k^{\frac{b_S}{b_F \beta}} - \frac{a_S}{b_S} \\
&= \left(\frac{a_S}{b_S} + S_t \right) e^{-\frac{b_S}{b_F \beta} \left| \ln \frac{L_F(K)}{L_F(F_t)} \right|} - \frac{a_S}{b_S}.
\end{aligned}$$

On peut approximer le payoff par :

$$\tilde{H}_t(S_T) \simeq (F_t - K) + \rho\beta \frac{L_F(F_t)}{L_S(S_t)} (S_T - S_t) + \tilde{H}^{p\prime}_t \left(K^{c+}_t \right) (S_T - K^c_t)^+ - \tilde{H}^{p\prime}_t \left(K^{p-}_t \right) (K^p_t - S_T)^+ .$$

□

Preuve de la proposition (4.12.2).

DÉMONSTRATION. Comme

$$\mathbb{E} \left[e^{(1 + \phi_+(y)) X_{t,T}} | V_{t,T} \right] = \mathbb{E} \left[e^{-\phi^-(y) X_{t,T}} | V_{t,T} \right] = e^{y V_{t,T}}$$

et

$$\begin{aligned}\mathbb{E}\left[\phi'_+(y)X_{t,T}e^{1+\phi_+(y))X_{t,T}}|V_{t,T}\right] &= \frac{\partial}{\partial y}\mathbb{E}\left[e^{(1+\phi_+(y))X_{t,T}}|V_{t,T}\right]\\ &= \frac{\partial}{\partial y}e^{yV_{t,T}}\\ &= V_{t,T}e^{yV_{t,T}},\\ \mathbb{E}\left[\phi'_-(y)X_{t,T}e^{-\phi_-(y)X_{t,T}}|V_{t,T}\right] &= -V_{t,T}e^{yV_{t,T}},\end{aligned}$$

on obtient les inégalités suivantes

$$\begin{aligned}\mathbb{E}\left[\theta_0(y)e^{(1+\phi_+(y))X_{t,T}}+(1-\theta_0(y))\,e^{-\phi_-(y)X_{t,T}}|V_{t,T}\right] &= e^{yV_{t,T}},\\ \mathbb{E}\left[\theta_1(y)\left(e^{(1+\phi_+(y))X_{t,T}}-e^{-\phi_-(y)X_{t,T}}\right)|V_{t,T}\right] &= 0\end{aligned}$$

et

$$\begin{aligned}&\mathbb{E}\left[\left\{\theta_0(y)e^{(1+\phi_+(y))X_{t,T}}+(1-\theta_0(y))\,e^{-\phi_-(y)X_{t,T}}\right\}\frac{X_{t,T}}{V_{t,T}}|V_{t,T}\right]\\ &= e^{yV_{t,T}}\left\{\frac{\theta_0(y)}{\phi'_+(y)}-\frac{1-\theta_0(y)}{\phi'_-(y)}\right\}\\ &= 0,\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left[\theta_1(y)\left(e^{(1+\phi_+(y))X_{t,T}}-e^{-\phi_-(y)X_{t,T}}\right)\frac{X_{t,T}}{V_{t,T}}|V_{t,T}\right] &= e^{yV_{t,T}}\theta_1(y)\left\{\frac{1}{\phi'_+(y)}+\frac{1}{\phi'_-(y)}\right\}\\ &= e^{yV_{t,T}}.\end{aligned}$$

Comme $\theta_0(y)\in[0;1]$ (car $\phi'_\pm\geq 0$) est borné, on peut appliquer le théorème de Fubini ce qui donne

$$\begin{aligned}&\mathbb{E}\left[P(h_t,0)(X_{t,T})|\mathcal{F}_t,V_{t,T}=v\right]\\ &= \mathbb{E}\left[\int_{\mathbb{R}_+}\left\{\theta_0(y)e^{(1+\phi_+(y))X_{t,T}}+(1-\theta_0(y))\,e^{-\phi^-(y)X_{t,T}}\right\}m_t(dy)|\mathcal{F}_t,V_{t,T}=v\right]\\ &= \int_{\mathbb{R}_+}\left\{\theta_0(y)\mathbb{E}\left[e^{(1+\phi_+(y))L_v}\right]+(1-\theta_0(y))\,\mathbb{E}\left[e^{-\phi_-(y)L_v}\right]\right\}m_t(dy)\\ &= \int_{\mathbb{R}_+}e^{yv}m_t(dy)\\ &= h_t(v)\end{aligned}$$

et de même

$$\mathbb{E}\left[P(h_t,0)(X_{t,T})\frac{X_{t,T}}{V_{t,T}}|\mathcal{F}_t,V_{t,T}=v\right]=0.$$

Or,

$$|\theta_1(y)|\leq\sup\left\{\phi'_+(0),\phi'_-(y)\right\},\quad\forall y\geq 0,$$

donc θ_1 est borné et on peut appliquer Fubini permettant d'obtenir

$$\mathbb{E}\left[P(0,g_t)(X_{t,T})|\mathcal{F}_t, V_{t,T}=v\right]=0,$$
$$\mathbb{E}\left[P(0,g_t)(X_{t,T})\frac{X_{t,T}}{V_{t,T}}|\mathcal{F}_t, V_{t,T}=v\right]=g_t(v).$$

□

Preuve de la proposition (4.12.3).

Démonstration. Pour $a,b\in\mathbb{R}$, on a :

$$\begin{aligned}\mathbb{E}\left[e^{aX_{t,T}+bV_{t,T}}|V_{t,T}\right]&=e^{(\psi(a)+b)V_{t,T}},\\ \mathbb{E}\left[X_{t,T}e^{aX_{t,T}+bV_{t,T}}|V_{t,T}\right]&=\frac{\partial}{\partial a}\mathbb{E}\left[e^{aX_{t,T}+bV_{t,T}}|V_{t,T}\right]\\ &=\psi'(a)e^{(\psi(a)+b)V_{t,T}}.\end{aligned}$$

D'où

$$\begin{aligned}\mathbb{E}\left[e^{aX^\epsilon_{t,T}+bV_{t,T}}|V_{t,T}\right]&=e^{(\psi(a+\epsilon)+\beta)V_{t,T}}\\ &=e^{(\psi(a)+b)V_{t,T}}+\epsilon\psi'(a)e^{(\psi(a)+b)V_{t,T}}+O(\epsilon^2)\\ &=\mathbb{E}\left[e^{aX_{t,T}+bV_{t,T}}|V_{t,T}\right]+\epsilon\mathbb{E}\left[X_{t,T}e^{aX_{t,T}+bV_{t,T}}|V_{t,T}\right]+O(\epsilon^2).\end{aligned}$$

Soit Z variable aléatoire indépendante de $(X^\epsilon_{t,T},V_{t,T})$. Par densité, il peut étendre ce résultat pour $f(X^\epsilon_{t,T},V_{t,T},Z)=e^{\alpha X^\epsilon_{t,T}+\beta V_{t,T}+\theta Z}$ avec $\alpha,\beta,\theta\in\mathbb{C}$ (il suffit d'appliquer le résultat avec l'espérance conditionnelle par rapport à Z). On en déduit que pour toute fonction f

$$\mathbb{E}\left[f(X^\epsilon_{t,T},V_{t,T},Z)|V_{t,T}\right]=\mathbb{E}\left[f(X_{t,T},V_{t,T},Z)|V_{t,T}\right]+\epsilon\mathbb{E}\left[X_{t,T}f(X_{t,T},V_{t,T},Z)|V_{t,T}\right]+O(\epsilon^2).$$

On obtient alors le résultat souhaité en appliquant cette égalité à ξ^ϵ_T et $P_t(X^\epsilon_{t,T})$:

$$\begin{aligned}\mathbb{E}\left[\xi^\epsilon_T|\mathcal{F}_t,V_{t,T}\right]&=\mathbb{E}\left[\xi_T|\mathcal{F}_t,V_{t,T}\right]+\epsilon\mathbb{E}\left[X_{t,T}\xi_T|\mathcal{F}_t,V_{t,T}\right]+O(\epsilon^2)\\ &=h_t(V_{t,T})+\epsilon V_{t,T}g_t(V_{t,T})+O(\epsilon^2)\\ &=\mathbb{E}\left[P_t(X_{t,T})|\mathcal{F}_t,V_{t,T}\right]+\epsilon V_{t,T}\mathbb{E}\left[X_{t,T}P_t(X_{t,T})|\mathcal{F}_t,V_{t,T}\right]+O(\epsilon^2)\\ &=\mathbb{E}\left[P_t(X^\epsilon_{t,T})|\mathcal{F}_t,V_{t,T}\right]+O(\epsilon^2).\end{aligned}$$

□

Preuve de la proposition (4.12.4). A la date t, on regarde la couverture donnée par une position dynamique donnée par l'option $P^\theta_t(X_{t,T})$ (que l'on revend juste après et en achetant pour l'option $P^\theta_{t+\epsilon}(X_{t+\epsilon,T})$ en $t+\epsilon$ avec $\epsilon\mapsto 0^+$). On note dp^θ_t la variation instantanée du portefeuille à l'instant t et $dp^\theta_t(v)$ cette variation

conditionnellement à $V_{t,T} = v$:

$$\begin{aligned}
dp_t^\theta &= \mathbb{E}_t\left[(P_{t-}^\theta)'(X_{t,T})\right]\sigma dW_t + \int_{\mathbb{R}} \mathbb{E}_t\left[P_{t-}^\theta(x+X_{t,T}) - P_{t-}^\theta(X_{t,T})\right]\mu(dx,dt) \\
&\quad + \int_{\mathbb{R}} \mathbb{E}_t\left[P_{t-}(x+X_{t,T}) - P_{t-}(X_{t,T})\right]\nu(dx)dt, \\
dp_t^\theta(v) &= \mathbb{E}_t\left[(P_{t-}^\theta)'(X_{t,T})|V_{t,T}=v\right]\sigma dW_t + \int_{\mathbb{R}} \mathbb{E}_t\left[P_{t-}^\theta(x+X_{t,T}) - P_{t-}^\theta(X_{t,T})|V_{t,T}=v\right]\mu(dx,dt) \\
&\quad + \int_{\mathbb{R}} \mathbb{E}_t\left[P_{t-}(x+X_{t,T}) - P_{T_u-}(X_{t,t})|V_{t,T}=v\right]\nu(dx)dt.
\end{aligned}$$

L'erreur quadratique de couverture vaut alors

$$\begin{aligned}
&\mathbb{E}_t\left[\left|\xi_T - \mathbb{E}_t\left[P_t^\theta(X_{t,T})\right] - \int_t^T dp_s^\theta\right|^2\right] \\
&= \left|\mathbb{E}_t\left[\xi_T - P_t^\theta(X_{t,T})\right]\right|^2 + \mathbb{E}_t\left[\int_t^T \left|\delta_s(V_{s,T}) - \mathbb{E}_s\left[(P_s^\theta)'(X_{s,T})\right]\right|^2 d\left\langle X\right\rangle_s\right] \\
&+\mathbb{E}_t\left[\sum_{s\in(t,T)}\left|U_s(\Delta X_s, V_{s,T}) - \mathbb{E}\left[P_{s-}^\theta(x+X_{s,T}) - P_{s-}^\theta(X_{s,T})|\mathcal{F}_{s^-}, \Delta X_s = x\right]\right|^2\right] \\
&=\mathbb{E}_t\left|h_t(V_{t,T}) - \mathbb{E}_t\left[P_t(X_{t,T})\right]\right|^2 + \mathbb{E}_t\int_{V_t}^{V_T}\left|\delta_{T_u}(V_T-u) - \mathbb{E}_s\left[P'_{T_u}(X_{T_u,T})\right]\right|^2\sigma^2 du \\
&+\mathbb{E}_t\int_{V_t}^{V_T}\int_{\mathbb{R}}\left|U_{T_u}(x, V_T-u) - \mathbb{E}_{T_u^-}\left[P_{T_u-}(x+X_{T_u,T}) - P_{T_u-}(X_{T_u,T})\right]\right|^2\nu(dx)du
\end{aligned}$$

avec le changement de temps $T_{V_u} = u$. On considère P_t^θ sous la forme

$$P_t^\theta(x) = \int_{\mathbb{R}_+}\left\{\theta_t^+(y)e^{(1+\phi_+(y))x} + \theta_t^-(y)e^{-\phi_-(y)x}\right\}m_t(dy) \tag{4.13.1}$$

ce qui donne

$$\begin{aligned}
dp_t^\theta(v) &= \int_{\mathbb{R}_+}\left\{\theta_t^+(y)\left(1+\phi_+(y)\right) + \theta_t^-(y)\left(-\phi_-(y)\right)\right\}e^{yv}m(dy)\sigma dW_{V_t} \\
&\quad + \int_{\mathbb{R}}\int_{\mathbb{R}^+}\left\{\theta_t^+(y)\left(e^{(1+\phi_+(y))x}-1\right) + \theta_t^-(y)\left(e^{-\phi_-(y)x}-1\right)\right\}m(dy)\tilde{\mu}(dx,dV_t)
\end{aligned}$$

et

$$\begin{aligned}
&\mathbb{E}_t\left[\left|\xi_T - \mathbb{E}_t\left[P_t^\theta(X_{t,T})\right] - \int_t^T dp_s^\theta\right|^2\right] \\
&= \sum_{t\le s\le T}\mathbb{E}_t\left[\left|\mathbb{E}_s\left[\xi_T - P_s^\theta(X_{s,T})\right]\right|^2\right] + \mathbb{E}_t\left[\int_t^T\int_{\mathbb{R}_+} R_s^\theta(y)e^{2yV_{s,T}}m^2(dy)ds\right]
\end{aligned}$$

avec

$$\begin{aligned}
R_s^\theta(y) &= \left|\alpha_{\delta_s}(y) - \left\{\theta_s^+(y)\left(1+\phi_+(y)\right) + \theta_s^-(y)\left(-\phi_-(y)\right)\right\}\right|^2\sigma^2 \\
&\quad + \int_{\mathbb{R}}\left|\alpha_{U_s}^x(y) - \left\{\theta_s^+(y)\left(e^{(1+\phi_+(y))x}-1\right) + \theta_t^-(y)\left(e^{-\phi_-(y)x}-1\right)\right\}\right|^2\nu(dx).
\end{aligned}$$

Le minimum

$$\inf_{\theta} \mathbb{E}_t \left[\left| \xi_T - \mathbb{E}_t \left[P_t^{\theta}(X_{t,T}) \right] - \int_t^T dp_s^{\theta} \right|^2 \right]$$

est atteint en θ^* minimisant

$$R_s^{\theta^*} = \inf_{\theta} R_s^{\theta}, \quad \forall s \in [0,T]$$

avec la contrainte

$$\mathbb{E}_t \left[P_t^{\theta^*}(X_{t,T}) | V_{t,T} = v \right] = \mathbb{E}_t \left[\xi_T | V_{t,T} = v \right], \quad \forall t \in [0,T], \forall v \geq 0. \tag{4.13.2}$$

En écrivant P_t sous la forme exponentielle (4.13.1), la contrainte (4.13.2) est équivalente à

$$\theta_s^{*,+}(y) + \theta_s^{*,-}(y) = \alpha_{h_s}(y), \quad \forall s \in [0,T], \forall y \in \mathbb{R}^+. \tag{4.13.3}$$

La minimisation sur R_s sous la contrainte (4.13.3) revient alors à minimiser un polynôme d'ordre 2 et le minimum est atteint en

$$\begin{aligned}
\theta_t^{+,*}(y) &= (1 + \phi_+(y) + \phi_-(y)) \left(\alpha_{\delta_t}(y) + \alpha_{h_t}(y)\phi_-(y) \right) \sigma^2 / A(y) \\
&\quad + \int_{\mathbb{R}} \left(\alpha_{U_t}^x(y) - \alpha_{h_t}(y) \left(e^{-\phi_-(y)x} - 1 \right) \right) \left(e^{(1+\phi_+(y))x} - e^{-\phi_-(y)x} \right) \nu(dx) / A(y), \\
\theta_t^{-,*}(y) &= \alpha_{h_t}(y) - \theta_t^{+,*}(y) \\
A(y) &= |1 + \phi_+(y) + \phi_-(y)|^2 \sigma^2 + \int_{\mathbb{R}} \left| e^{(p_+ + \phi_+(y))x} - e^{-\phi_-(y)x} \right|^2 \nu(dx).
\end{aligned}$$

Preuve de la proposition (4.12.5).

DÉMONSTRATION. Dans le cas continu complet avec $X_t = B_{V_t} - \frac{V_t}{2}$, la couverture en variance moyenne est en fait une couverture complète, c'est-à-dire l'erreur de couverture quadratique est nulle d'où

$$\mathbb{E}\left[P_c'(h_t, \delta_t, U_t)(X_{t,T}) | \mathcal{F}_t, V_{t,T} \right] = \delta_t(V_{t,T}) = \mathbb{E}\left[\frac{\partial}{\partial x} H(X_T, V_T) | \mathcal{F}_t, V_{t,T} \right].$$

En remarquant que $\forall f : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$

$$\mathbb{E}\left[\frac{X_{t,T}}{V_{t,T}} f(X_{t,T}, V_{t,T}) | V_{t,T} \right] = \mathbb{E}\left[f'(X_{t,T}, V_{t,T}) | V_{t,T} \right] - \frac{1}{2} \mathbb{E}\left[f(X_{t,T}, V_{t,T}) | V_{t,T} \right],$$

on obtient avec $f(X_{t,T}, V_{t,T}) = H_t(X_{t,T}, V_{t,T})$

$$g_t(v) = \delta_t(v) - \frac{1}{2} h_t(v), \quad \forall v \geq 0,$$

puis avec $f(X_{t,T}, V_{t,T}) = P_v(h_t, g_t)(X_{t,T}), P_c(h_t, \delta_t, U_t)$

$$\begin{aligned}
\delta_t(V_{t,T}) &= \mathbb{E}\left[P_v'(h_t, g_t)(X_{t,T})|\mathcal{F}_t, V_{t,T}\right] \\
&= \mathbb{E}\left[P_c'(h_t, \delta_t, U_t)(X_{t,T})|\mathcal{F}_t, V_{t,T}\right], \\
g_t(V_{t,T}) &= \mathbb{E}\left[\frac{X_{t,T}}{V_{t,T}} P_v'(h_t, g_t)(X_{t,T})|\mathcal{F}_t, V_{t,T}\right] \\
&= \mathbb{E}\left[\frac{X_{t,T}}{V_{t,T}} P_c(h_t, \delta_t, U_t)(X_{t,T})|\mathcal{F}_t, V_{t,T}\right], \\
h_t(V_{t,T}) &= \mathbb{E}\left[P_v(h_t, g_t)(X_{t,T})|\mathcal{F}_t, V_{t,T}\right] \\
&= \mathbb{E}\left[P_c(h_t, \delta_t, U_t)(X_{t,T})|\mathcal{F}_t, V_{t,T}\right].
\end{aligned}$$

D'où P_v et P_c vérifient (4.8.1). On conclut par unicité dans $\mathbb{H}_t(\mathbb{R})$ (d'après la proposition (4.8.6)) : $P_v(h_t, g_t) = P_c(h_t, \delta_t, U_t)$. □

Preuve de la proposition (4.12.6).

Démonstration. Le raisonnement est identique à celui de la preuve de la proposition (4.12.5) dans cadre Brownien. Dans le cas complet avec un processus de Poisson $X_t = aN_{V_t} - e^a V_t$, la couverture en variance moyenne est en fait une couverture complète, avec une erreur de couverture quadratique nulle d'où

$$\mathbb{E}\left[P_c(h_t, \delta_t, U_t)(X_{t,T} + a)|\mathcal{F}_t, V_{t,T}\right] = U_t(a, V_{t,T}) = \mathbb{E}\left[H(a + X_T, V_T)|\mathcal{F}_t, V_{t,T}\right],$$
$$\mathbb{E}\left[P_c'(h_t, \delta_t, U_t)(X_{t,T})|\mathcal{F}_t, V_{t,T}\right] = \delta_t(V_{t,T}).$$

En remarquant que $\forall f : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$

$$\mathbb{E}\left[X_{t,T} f(X_{t,T}, V_{t,T})|V_{t,T}\right] = V_{t,T}\mathbb{E}\left[f(a + X_{t,T}, V_{t,T})|V_{t,T}\right] - e^a V_{t,T}\mathbb{E}\left[f(X_{t,T}, V_{t,T})|V_{t,T}\right]$$

et donc que

$$g_t(v) = U_t(a, v) - e^a h_t(v), \quad \forall v \geq 0,$$

on en déduit comme dans la preuve de la proposition (4.12.5) que $P_v(h_t, g_t) = P_c(h_t, \delta_t, U_t)$. □

Preuve de la proposition (4.12.7). On pose $F(s) = f_t(\ln(1 + sM))$ pour $s \in [0, 1]$. En posant

$$a_{n,k} = \sum_{j=0}^{k} F(\frac{j}{n}) \binom{n}{k} \binom{k}{j} (-1)^{k-j},$$

l'appoximation de Bernstein $A_n(s) = \sum_{k=0}^{n} a_{n,k} s^k$ converge uniformément vers F sur $[0; M]$. Le changement de variable $s = \frac{e^v - 1}{M}$ permet de conclure : avec

$$b_{n,k}^{f_t} = \sum_{i=0}^{n} f_t(\ln(1 + \frac{iM}{n}))(-1)^{k+i} \sum_{j=i}^{n} \frac{1_{j \geq k}}{M^j} \binom{n}{j} \binom{j}{k} \binom{j}{i},$$

$B_n(v) = \sum b_{n,k}^{f_t} e^{kv} = A_n\left(\frac{e^v - 1}{M}\right)$ converge uniformément vers f_t sur $[0, m]$.

Bibliographie

[ALP95] M. Avellaneda, A. Levy, and A. Paras, *Pricing and hedging derivative securities in markets with uncertain volatilities*, Applied Mathematical Finance **2** (1995).

[AS72] M. Abramowitz and I. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables, 9th printing*, New York : Dover, 9ème édition, 1972.

[BBP97] G. Barles, R. Buckdahn, and E. Pardoux, *Bsdes and integral-partial differential equations*, Stochastics and Stochastics Reports **60** (1997), 57–83.

[BC08] B. Bouchard and J.F. Chassagneux, *Discrete time approximation for continuously and discretely reflected bsde's*, Stochastic Processes and their Applications **118** (2008), 2269–2293.

[BDM01] P. Briand, B. Delyon, and J. Ménin, *Donsker-type theorem for bsdes*, Electron. Comm. Probab **6** (2001), 1–14.

[BE08] B. Bouchard and R. Elie, *Discrete time approximation of decoupled forward-backward sde with jumps*, Stochastic Processes and their Applications **118 (1)** (2008), 53–75.

[Bec05] D. Becherer, *Bounded solutions to backward sdes with jumps for utility optimization and indifference hedging*, Preprint, Imperial College London (2005).

[Ben92] A. Bensoussan, *Stochastic control of partially observable systems*, Cambridge University Press, 1992.

[Ber98] J. Bertoin, *Levy process*, Cambridge University, 1998.

[BG97] M. Broadie and P. Glassermann, *A stochastic mesh method for pricing high-dimensional american options*, Working Paper, Columbia University **11** (1997), 1–37.

[Bil99] P. Billinsley, *Convergence of probability measures*, 1999.

[Bis73] J.-M. Bismut, *Conjugate convex functions in optimal stochastic control*, Journal of Mathematical Analysis and Applications **44** (1973), 384–404.

[BMP90] A. Benveniste, M. Métivier, and P. Priouret, *Adaptive algorithms and stochastic approximations*, vol. 2, 1990.

[BP03a] V. Bally and G. Pagès, *Error analysis of the quantization algorithm for obstacle problems*, Stochastic Processes and Their Applications **106** (2003).

[BP03b] ———, *A quantization algorithm for solving discrete time multidimensional optimal stopping problems*, Bernoulli **9** (2003).

[BPP05] V. Bally, G. Pagès, and J. Printemps, *A quantization tree method for pricing and hedging multidimensional american options*, Mathematical Finance **15** (2005).

[BR07] C. Bender and R. Robert, *A forward scheme for backward sdes*, Stochastic processes and their applications **117** (2007), 1793–1812.

[BS73] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy **81** (1973), 637–654.

[BS77] M.J. Brennan and E.S. Schwartz, *The valuation of the american put*, J. of Finance **32** (1977), 449–462.

[BS91] G. Barles and P. Souganidis, *Convergence of approximation schemes for fully nonlinear second-order equations*, Asymptotic Analysis **4** (1991), 271–283.

[BT04] B. Bouchard and N. Touzi, *Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations*, Stochastic Processes and their Applications **111(2)** (2004), 175–206.

[BW82] J. Bucklew and G. Wise, *Multidimensional asymptotic quantization theory with rth power distortion measures*, Information Theory, IEEE Transactions on **28** (1982), 239 – 247.

[Cha09] J.F. Chassagneux, *Discrete-time approximation of doubly reflected bsdes*, Advances in Applied Probability **41** (2009).

[Che97] D. Chevance, *Numerical methods for backward stochastic differential equations*, In C. Rogers and D.T. editors, Newton Institute, Cambridge University Press (1997), 232–244.

[CK96] J. Cvitanic and I. Karatzas, *Backward stochastic differential equations with reflection and dynkin games*, Annals of Probability **24** (1996), 2024–2056.

[CL08] P. Carr and R. Lee, *Robust replication of volatility derivatives*, PRMIA award for Best Paper in Derivatives, MFA 2008 Annual Meeting (2008).

[CM02] P. Carr and D. Madan, *Volatility*, ch. Towards a Theory of Volatilty Trading, Risk Publications, 2002.

[CM08] S. Crepey and A. Matoussi, *Reflected and doubly reflected bsdes with jumps : A priori estimates and comparison*, Annals of Applied Probability **18(5)** (2008), 2041–2069.

[CRR79] J.C. Cox, S. A. Ross, and M. Rubinstein, *Option pricing : a simplified approach*, Journal of Financial Economics **7** (1979), 229–263.

[CST05] P. Cheridito, M. Soner, and N. Touzi, *The multidensional super-replication problem under gamma constraints*, Annales de lInstitut Henri Poincaré, Série C : Analyse Non-Linéaire **22** (2005), 633–666.

[CSTV06] P. Cheridito, M. Soner, N. Touzi, and N. Victoir, *Second-order backward stochastic differential equations and fully nonlinear parabolic pdes*, Communications on Pure and Applied Mathematics **60** (2006), 1081 – 1110.

[CTZ03] T. Chouli, M. Taksar, and X.Y. Zhou, *A diffusion model for optimal dividend distribution with constraints on risk control*, SIAM J. Cont. Optim. **41** (2003), 1946–1979.

[DM06] F. Delarue and S. Menozzi, *A forward-backward stochastic algorithm for quasi-linear pdes*, Ann. Appl. Probab. **16** (2006), 140–184.

[DM08] ———, *An interpolated stochastic algorithm for quasi-linear pdes*, Mathematics of Computation **261-77** (2008), 125–158.

[EL05] A. Eyraud-Loisel, *Backward stochastic differential equations with enlarged filtration. option hedging of an insider trader in a financial market with jumps*, Stochastic Processes and their Applications **115** (2005).

[FG05] P. Fritz and J. Gatheral, *Valuation of volatility derivatives as an inverse problem*, working paper, 2005.

[FK85] T. Fujiwara and H. Kunita, *Stochastic differential equations of jump type and lévy processes in diffeomorphisms group*, Journal of Mathematics of Kyoto University **25(1)** (1985), 71–106.

[FP02] J.C. Fort and G. Pagès, *Asymptotic of optimal quantizers for some scalar distributions*, Journal of Computational and Applied Mathematics **146** (2002), 253–275.

[FTW11] A. Fahim, N. Touzi, and X. Warin, *A probabilistic numerical scheme for fully nonlinear pdes*, Annals of Applied Probability **25** (2011), 1364–1364.

[GG00] A. Gersho and R.M. Gray, *Vector quantization and signal compression*, The Kluwer International Series in Engineering and Computer Science. Springer, 2000.

[GL00] S. Graf and H. Luschgy, *Foundations of quantization for probability distributions*, Springer, 2000.

[GL07] Emmanuel Gobet and Céline Labart, *A sequential Monte-Carlo algorithm for solving BSDEs*, Proc. Appl. Math. Mech. **7** (2007), no. 1, 1081801–1081802, Special Issue : 6th International Congress on Industrial Applied Mathematics (ICIAM07) and GAMM Annual Meeting, Zürich 2007.

[GV02] F. Gozzi and T. Vargiolu, *Surreplication of european multiasset derivatives with bounded stochastic voaltility*, Mathematical Methods of Operations Research **55 (1)** (2002), 69–91.

[Hes93] S. Heston, *A closed-form solution for option with stochastic volatility with applications to bond and currency options*, Review of Financial Studies (1993), no. 6.

[HH06] S. Hamadène and H. Hassani, *Bsdes with two reflecting barriers driven by a brownian motion and an independent poisson noise and related dynkin game*, Electronic Journal of Probabilit **11** (2006), 121–145.

[HN89] S. D. Hodges and A. Neuberger, *Optimal replication of contingent claims under transaction costs*, Review of Futures Markets (1989), no. 8, 222–239.

[JLL90] P. Jaillet, D. Lamberton, and B. Lapeyre, *Variational inequalities and the pricing of american options*, Acta Applicandae Mathematicae **21** (1990), 263–289.

[JLW05] E. Gobet J.P. Lemor and X. Warin, *A regression-based monte-carlo method to solve backward stochastic differential equations*, Annals of appliaed Probability **15(3)** (2005), 2172–2002.

[KB61] R. Kalman and R. Bucy, *New results in linear filtering and prediction theory*, Journal of Basic Engineering **83** (1961), 95–108.

[KD85] H.J. Kushner and P. Dupuis, *Numerical methods for stochastic control problems in continuous time*, Stochastic Modelling and Applied Probability, Spinger-Verlag, 1985.

[KDZ99] M. Kamal K. Demeterfi, E. Derman and J. Zou, *More than you ever want to know about volatility swaps*, Tech. report, Quantitative Research Notes, Goldman Sachs, 1999.

[Kif00] Y. Kifer, *Game options*, Finance and Stochastics **4** (2000), 547–572.

[Kif06] ———, *Error estimates for binomial approximations of game options*, The annals of Applied Probability **16** (2006).

[KKP+97] N. El Karoui, C. Kapoudjan, E. Pardoux, S. Peng, and M.C. Quenez, *Reflected solutions of backward stochastic differential equations and related obstacle problem for pdes*, Ann. Prob. **25** (1997), 702–737.

[KPQ97] N. El Karoui, S. Peng, and M.C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance **7** (1997), 1–71.

[Lel85] H. Leland, *Option pricing and replication with transaction costs*, Journal of Finance **11** (1985).

[Llo82] S. Lloyd, *Least squares quantization in pcm*, IEEE Transactions on Information Theory **28** (1982), no. 2, 129–137.

[LP02] H. Luschgy and G. Pagès, *Functional quantization of gaussian processes*, Journal of Functional Analysis **196** (2002), no. 2, 486–531.

[LP08] ———, *Functional quantization rate and mean pathwise regularity of processes with an application to lévy processes*, Annals of Applied Probability **18** (2008), no. 2, 427–469.

[LR02] P.L. Lions and H. Régnier, *Calcul des prix et des sensibilités d'une option américaine par une méthode de monte carlo*, Tech. report, 2002.

[LS01] F. A. Longstaff and Edouardo Schwartz, *Valuing american options by simulation : a simple least-squares approach*, Review of Financial Studies **14** (2001).

[LT01] B. Lapeyre and E. Temam, *Competitive monte carlo methods for the pricing of asian options*, Journal of Computational Finance **5** (2001).

[Lud09] M. Ludkovski, *A simulation approach to optimal stopping under partial information*, Stochastic Processes and Applications **119(12)** (2009), 2071–2087.

[McN01] J. McNames, *A fast nearest neighbor algorithm based on a principal axis tree*, IEEE Transactions on Pattern Analysis and Machine Intelligence **23** (2001), no. 9, 964–976.

[Mer71] Robert Merton, *Optimum consumption and portfolio rules in a continuous time model*, Journal of Economic Theory **3** (1971), 373–413.

[MPY94] J. Ma, P. Protter, and J. Yong, *Solving forward backward stochastic differential equations explicitly a four step scheme*, Probability Theory Related Fields **98** (1994).

[Neu94] A. Neuberger, *The log contract*, Journal of Protfolio Management **20** (1994), 74–80.

[Nua95] D. Nualart, *The malliavin calculus and related topics*, Springer, 1995.

[Otm08] M. El Otmani, *Approximation scheme for solutions of bsdes with two reflecting barriers*, Mathematical Finance **1** (2008), 60–83.

[Pag97] G. Pagès, *A space vector quantization method for numerical integration*, Journal of Applied and Computional Mathematics **89** (1997), 1–38.

[Pen99] S. Peng, *Monotonic limit theorem of bsde and nonlinear decomposition theorem of doob-meyer's type*, Probability Theory and Realated Fields **113** (1999), 473–499.

[Pha00] H. Pham, *On quadratic hedging in continuous time*, Mathematical Methods of Operations Research **51** (2000).

[Pha09] ———, *Continuous-time stochastic control and applications with financial applications*, Series Stochastic Modeling and Applied Probability, vol 61, Springer, 2009.

[PP90] E. Paroux and S. Peng, *Adapted solution of a backward stochastic dfferential equation*, Systems Control Letters **14** (1990), 55–61.

[PP03] G. Pagès and J. Printems, *Optimal quadratic quantization for numerics : the gaussian case*, Monte Carlo Methods and Applications **9** (2003), no. 2, 135–166.

[PP04] G. Pagès and H. Pham, *An optimal markovian quantization algorithm for multidimensional stochastic problems*, stochastic and Dynamics **4** (2004), 501–545.

[PP05a] ———, *Optimal quantization methods for nonlinear filtering with discrete time observations*, Bernouilli **11** (2005), no. 5, 893–932.

[PP05b] G. Pagès and J. Printems, *Functional quantization for numerics with an application to option pricing*, Monte Carlo Methods and Applications **11** (2005), no. 4, 407–446.

[PRS05] H. Pham, W. Runggaldier, and A. Sellami, *Approximation by quantization of the filter process and applications to optimal stopping problems under partial observation*, Monte Carlo Methods and Applications **11** (2005), 57–82.

[PS88] A.F. Perold and W. F. Sharpe, *Dynamic strategies for asset allocation*, Financial Analyst Journal (1988), 16–27.

[PW10] G. Pagès and B. Wilbertz, *Intrinsic stationarity for vector quantization : Foundation of dual quantization*, Preprint PMA-1393 (2010).

[PX08] S. Peng and M. Xu, *Numerical algorithms for bsdes : convergence and simulations*, preprint (2008).

[RK00] R. Rouge and N. El Karoui, *Pricing via utility maximization and entropy*, Mathematical Finance **10** (2000), 259–276.

[Rog02] L. C. G. Rogers, *Monte carlo valuation of american options*, Mathematical Finance **12** (2002), 271–286.

[Sch01] M. Schweizer, *A guided tour through quadratic hedging approaches*, Option Pricing, Interest Rates and Risk Management, Cambridge University Press (2001), 538–578.

[Sne52] J.L. Snell, *Applications of martingale system theorems*, Trans. Ameri. Math. Soc. **73** (1952), 293–312.

[ST00] M. Soner and N. Touzi, *Super replication under gamma constraints.*, SIAM Journal on Control and Optimization **39** (2000), no. 1, 73–96.

[Ste70] H. Stehfest, *numerical inversion of laplace tranforms*, Commun. ACM 13 (1970).

[TL94] S. Tang and X. Li, *Necessary conditions for optimal control of stochastic systems with random jumps*, SIAM J. Control Optm **32 (5)** (1994), 1447–1475.

[TR99] J. N. Tsitsiklis and B. Van Roy, *optimal stopping of markov processes : Hilbert sapces theory, approximations algorithms and an application to pricing high-dimensional financial derivatives*, IEEE Transactions on Automatic Control **44** (1999), 1840–1851.

[TR00] ———, *Regression methods for pricing complex american-style options*, working papers, MIT (2000), 1–22.

[Xu08] M. Xu, *Numerical algorithms and simulations for reflected backward stochastic differential equations with two continuous barriers*, preprint (2008).

[Zad82] P.L. Zador, *Asymptotic quantization error of continuous signals and the quantization dimension*, IEEE Trans. Infi. Theory, Special issue on Quantization **28** (1982), no. 2, 139–166.

[Zha04] Jianfeng Zhang, *A numerical scheme for bsdes*, Annals of Applied Probability **14** (2004), 459–488.