

Approximation par projections et simulations de Monte-Carlo des équations différentielles stochastiques rétrogrades.

Jean-Philippe Lemor

▶ **To cite this version:**

Jean-Philippe Lemor. Approximation par projections et simulations de Monte-Carlo des équations différentielles stochastiques rétrogrades.. Mathématiques [math]. Ecole Polytechnique X, 2005. Français. NNT : . pastel-00001396

HAL Id: pastel-00001396
https://pastel.hal.science/pastel-00001396

Submitted on 27 Jul 2010

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.

Thèse présentée pour obtenir le titre de

DOCTEUR DE L'ECOLE POLYTECHNIQUE

spécialité : Mathématiques Appliquées

par

Jean-Philippe Lemor

Approximation par projections et simulations de Monte-Carlo des équations différentielles stochastiques rétrogrades

Thèse soutenue le 13 juin 2005 devant le jury composé de

Jean-Philippe BOUCHAUD	Président du Jury
Nicole EL KAROUI	Directrice de thèse
Emmanuel GOBET	Directeur de thèse
Saïd HAMADENE	Rapporteur
Nadia OUDJANE	
Gilles PAGES	
Philip PROTTER	Rapporteur
Xavier WARIN	Directeur de thèse

Résumé

Cette thèse traite de l'approximation des équations différentielles stochastiques rétrogrades (EDSR) par projections et simulations de Monte-Carlo. Les EDSR intervenant dans de nombreux problèmes de modélisation, en particulier en mathématiques financières pour la valorisation et la couverture de produits dérivés, il apparait important de disposer d'algorithmes souples et efficaces pour les approcher.

Dans une première partie, nous proposons un premier algorithme dont les trois paramètres sont un pas de discrétisation en temps h, des bases de fonctions et un nombre M de simulations de Monte-Carlo. Nous établissons la convergence de cet algorithme ainsi que des bornes d'erreur en fonction de ces trois paramètres.

Les bornes d'erreur établies ne permettant pas de faire varier simultanément les trois paramètres dans le but d'obtenir une précision arbitrairement petite, nous modifions ce premier algorithme dans une deuxième partie. Nous analysons la convergence de ce nouvel algorithme à l'aide de techniques de statistique non-paramétrique et montrons les bornes d'erreurs désirées.

Dans une troisième partie, nous généralisons cet algorithme au cas des EDSR réfléchies qui permettent de modéliser par exemple le problème de réplication d'une option américaine en finance. Nous montrons également que dans certains cas l'on peut obtenir, en utilisant des techniques du calcul de Malliavin, des vitesses de convergence plus fines que celles établies dans les deux premières parties.

Enfin, dans une partie numérique, nous proposons des choix de bases de fonctions explicites afin de tester notre algorithme dans diverses situations. Tout d'abord nous testons la variation simultanée des trois paramètres de l'algorithme afin de valider empiriquement les bornes d'erreur établies dans la deuxième partie. Nous testons ensuite notre algorithme sur des exemples financiers réalistes et proposons deux nouvelles méthodes de discrétisation pour le cas des EDSR réfléchies qui complètent celle présentée dans la troisième partie.

Mots clefs : équations différentielles stochastiques rétrogrades, mathématiques financières, méthodes de Monte-Carlo, statistique non-paramétrique, régression linéaire, bases de fonctions.

Abstract

This thesis deals with the approximation of backward stochastic differential equations (BSDE) using regression-based Monte-Carlo methods. As BSDE are a useful tool especially in financial mathematics, it is an important topic to define efficient algorithms for their approximation.

In a first part, we propose an algorithm with three kind of parameters : a time step h, functions bases and M Monte-Carlo simulations. Then, we derive error bounds for this algorithm.

As these error bounds do not allow us to move together all the algorithm's parameters to achieve a given accuracy, we slightly modify our first algorithm in a second part. We thus derive new error bounds, using techniques of non-parametric statistics.

In a third part, we generalize our algorithm to the case of reflected BSDE which arise in the pricing of American options. We also show that in particular cases, it is possible to derive better error bounds, by using Malliavin calculus techniques.

In the numerical part, we propose explicit choices of basis functions in order to test our algorithm in various cases. First, we test the simultaneous variation of the algorithm's parameters to empirically validate the error bounds derived in the second part. Then, we test our algorithm on real financial examples and propose two others discretization methods for reflected BSDE and empirically compare them to the one proposed in the third part.

Keywords : Backward stochastic differential equations, financial mathematics, Monte-Carlo methods, non-parametric statistics, linear regression, function basis.

Remerciements

Je voulais tout d'abord remercier Emmanuel Gobet pour son encadrement, sa patience, sa compréhension, ses remarques toujours pertinentes, ainsi que pour toutes les heures qu'il a bien voulu me consacrer, séances de travail ou relecture de notes qui ont permis de débusquer des erreurs jusqu'au dernier moment. Cette thèse lui doit beaucoup.

Je remercie également Nicole El Karoui d'avoir bien voulu m'accueillir dans son laboratoire ainsi que pour ses conseils décisifs sur l'orientation de la thèse vers les équations différentielles stochastiques rétrogrades.

J'ai beaucoup apprécié de travailler à EDF avec Xavier Warin, que je remercie chaleureusement de sa confiance et de son ouverture d'esprit qui m'ont permis de réaliser cette thèse avec beaucoup de liberté, ainsi que pour son intuition numérique et les librairies $C++$ qui m'ont beaucoup servi dans la partie numérique.

Je remercie chaleureusement les rapporteurs Philip Protter et Saïd Hamadène d'avoir accepté de relire cette thèse. Je suis très touché que Jean-Philippe Bouchaud, Nadia Oudjane et Dominique Picard aient accepté de faire partie du Jury.

Merci à l'équipe de la bibliothèque Mathématiques-Recherche de Chevaleret. J'ai beaucoup aimé rédiger une partie de cette thèse dans l'ambiance stimulante de cette bibliothèque.

Merci aux membres du CMAP (un grand merci à François Jouve pour le logiciel xgraphic avec lequel les graphes de la partie numérique ont été réalisés et à Sylvain Ferrand pour toutes ses réponses à mes questions de néophyte), laboratoire que j'ai été très heureux de fréquenter durant trois ans, aux thésards (Nicolas, Patrice, Julien, Asma, Claudia, Moeiz, Peter, Katia, François...) et participants du groupe de travail finance.

Merci également aux ingénieurs d'I23 puis de R33 (Vincent, Jean-Philippe...) pour leurs encouragements et les discussions que nous avons pu partager.

Tous mes remerciements à ma famille et mes amis pour leur soutien tout au long de cette thèse.

Enfin, merci Isabelle : sans toi cette thèse n'aurait jamais commencé et sans ta patience et ton soutien, elle n'aurait jamais abouti. Merci.

Table des matières

Introduction **13**

I Un premier algorithme pour la résolution des équations différentielles stochastiques rétrogrades **33**

1 Cadre, hypothèses et algorithme **37**
1.1 Cadre et hypothèses 37
1.2 Description de l'algorithme 38
1.2.1 Notations et hypothèses supplémentaires 38
1.2.2 L'algorithme 40

2 Etape 1 : discrétisation en temps **43**
2.1 Discrétisation en temps de X 43
2.2 Discrétisation de l'EDSR 44
2.3 Erreur de discrétisation 45
2.4 Propriétés de l'équation rétrograde discrète 49

3 Etape 2 : approximation de fonctions **51**
3.1 Définition de l'approximation 51
3.2 Erreur commise à l'étape 2 52

4 Etape 3 : méthode de Monte-Carlo **59**
4.1 Erreur commise à l'étape 3 59
4.2 Commentaires sur les majorations d'erreurs obtenues 67

II Algorithme modifié pour les équations différentielles stochastiques rétrogrades **69**

1 Description de l'algorithme **73**
1.1 Nouvelle discrétisation en temps de l'EDSR 73

1.2 Description et notations . . . 74
1.3 Localisation de l'équation rétrograde . . . 77
1.3.1 Bornitude induite par la localisation . . . 77
1.3.2 Erreur induite par la localisation . . . 79

2 Analyse de l'algorithme 81
2.1 Notations propres à l'analyse de l'algorithme . . . 81
2.1.1 Espérances conditionnelles . . . 81
2.1.2 Coefficients de projection . . . 82
2.1.3 Normes . . . 83
2.1.4 Fonctions de base . . . 83
2.1.5 Evènements . . . 83
2.1.6 Erreurs . . . 83
2.2 Propagation de l'erreur . . . 84
2.3 Majoration des termes d'erreurs . . . 89
2.3.1 Majoration de $\mathbf{T_{1,k,M}}$ . . . 90
2.3.2 Majoration de $\mathbf{T_{2,k,M}}$ . . . 91
2.3.3 Majoration de $\mathbf{T_{3,k,M}}$ . . . 92
2.3.4 Majoration de $\mathbf{T_{4,k,l,M}}$ . . . 93
2.3.5 Majoration de $\mathbf{T_{5,k,l,M}}$ . . . 94
2.3.6 Majoration de $\mathbb{P}([\mathbf{A_k^M}]^\mathbf{c})$ . . . 95

3 Majoration de l'erreur commise dans l'algorithme 99
3.1 Majoration de l'erreur pour la loi de $P_{t_k}^N$ . . . 99
3.2 Erreur sur $z^{N,R}$ . . . 103
3.3 Utilisation des majorations obtenues . . . 106
3.3.1 Estimation des nombres de couverture . . . 106
3.3.2 Variation simultanée des paramètres . . . 107
3.3.3 Erreur et complexité . . . 108

4 Extensions 113
4.1 Driver nul . . . 113
4.2 Analyse de l'algorithme sans incréments browniens supplémentaires . . . 116

III Des extensions 123

1 Vitesse de convergence dans le cas driver linéaire 125
1.1 Cadre et hypothèses . . . 125
1.2 Vitesse de convergence . . . 126

1.2.1 Résultat principal . . . 126
1.2.2 Des résultats intermédiaires . . . 127
1.2.3 Preuve du théorème III.1 . . . 130

2 Algorithme pour les équations différentielles stochastiques rétrogrades réfléchies 135
2.1 Le cadre et les hypothèses . . . 135
2.2 L'algorithme . . . 136
2.3 Erreur de localisation . . . 137
2.4 Erreur dûe aux simulations et projections . . . 138
2.5 Erreur de discrétisation . . . 139
2.5.1 Régularité $C^{1,2}$ . . . 140
2.5.2 Relaxation des hypothèses sur l'obstacle . . . 141
2.6 Conclusion . . . 144

IV Résultats numériques 147

1 Rappel des chapitres précédents 149

2 Choix des fonctions de base 151
2.1 Hypercubes . . . 151
2.2 Cellules de Voronoï . . . 152
2.3 Polynômes globaux . . . 152

3 Comparaison des approches avec ou sans incréments browniens auxiliaires 153

4 Variation simultanée des paramètres 157
4.1 Estimation de l'erreur d'approximation pour la base **HC** . . . 157
4.2 Tests sur une option européenne avec différentiel de taux . . . 160
4.2.1 Tests avec la base **HC** . . . 160
4.2.2 Tests avec la base **HC(1,0)** . . . 166
4.2.3 Tracé des fonctions de prix . . . 167
4.3 Tests sur une EDSR réfléchie . . . 170

5 Comparaison des algorithmes des parties I et II 175

6 Options américaines 183
6.1 Tests à coefficient fixés . . . 185
6.1.1 Méthode de pénalisation . . . 185

6.1.2 Méthode de régularisation 187
6.2 Tests à coefficients variables 189
6.2.1 Méthode de pénalisation 189
6.2.2 Méthode de régularisation 190
6.3 Liens avec la méthode de Longstaff-Schwartz [LS01] 191

7 Exemples financiers **193**
7.1 Modèle factoriel sur le marché de l'électricité 193
7.2 Combinaison de calls avec différentiel de taux d'intérêt 194
7.3 Options asiatiques 196
7.4 Options américaines et influence de la dimension 197

Annexes **203**
A.1 Equations différentielles stochastiques rétrogrades 203
A.1.1 Existence et unicité 203
A.1.2 Propriétés fines des EDSR 204
A.2 Equations différentielles stochastiques rétrogrades réfléchies 205
A.3 Inégalités classiques 209
A.4 Régression non-paramétrique 209
A.4.1 Nombres de couverture 209
A.4.2 Coefficients d'éclatement et dimension VC 210
A.4.3 Vitesse de convergence pour l'estimation moindres carrés de la fonction de régression 212
A.4.4 Inégalité de Hoeffding 212
A.5 Analyse matricielle 213

Introduction

Contexte de la thèse

Cadre général

Cette thèse a été effectuée dans le cadre d'un contrat Cifre chez Electricité de France, d'abord dans le département SINETICS (SImulation en NEutronique Technologie de l'Information et Calcul Scientifique) puis dans le département OSIRIS (Optimisation SImulation RIsque et Statistiques) de la division Recherche et Développement. Sur le plan universitaire, la thèse a été encadrée au Centre de Mathématiques Appliquées de l'Ecole Polytechnique, par les Professeurs Nicole El Karoui et Emmanuel Gobet. Xavier Warin, Ingénieur senior à OSIRIS, a assuré le suivi de la thèse côté EDF.
Le but initial de cette thèse était de proposer et d'analyser des méthodes numériques de valorisation d'options par optimisation de bilan financier.
Avec l'ouverture des marchés de l'électricité, EDF est amenée à valoriser les anciennes clauses clients et à proposer des produits de type optionnel de plus en plus complexes à ses clients. OSIRIS a pour mission de développer des outils de valorisation (*pricing*) à destination de la direction commerciale et des outils de gestion du risque à destination de l'ensemble des entités de EDF exposées au risque de variation du cours de l'électricité ou du combustible (gaz, charbon, fioul...).
Rappelons brièvement le principe d'un contrat optionnel (voir par exemple [KS98] pour de plus amples informations). Supposons qu'une entreprise se fournisse en électricité à un prix dépendant du prix spot horaire de l'électricité côté sur un marché donné. La rentabilité de l'entreprise dépendant de ses coûts de production, elle souhaite se garantir contre une hausse du prix spot de l'électricité et va souscrire un contrat financier appelé option qui va lui payer une certaine somme d'argent si le prix spot de l'électricité dépasse un certain plafond fixé dans le contrat. Ainsi, si le prix spot augmente trop, les coûts de production de l'entreprise vont augmenter mais ils seront équilibrés par la somme reçue grâce à l'option achetée. L'exemple le plus classique est celui d'une option d'achat (*call*) de maturité T et de strike K. L'entreprise qui achète ce contrat se garantit de pouvoir toucher la somme $(S_T - K)_+$ en T : l'entreprise touche une somme d'argent dès que le

prix spot de l'électricité en T, S_T, dépasse le plafond K.
Les sous-jacents des contrats optionnels sont (pour les marchés électriques) le prix spot de l'électricité et les prix forward associés à diverses maturités. Ces sous-jacents présentent de fortes différences avec les marchés financiers classiques comme les marchés actions ou taux d'intérêt. Une des différences est que l'on observe de manière claire des pics (souvent importants) dans l'évolution du prix spot de l'électricité. Pour tenir compte de ces pics ainsi que de phénomènes de retour à la moyenne, plusieurs modèles ont été suggérés : on peut consulter par exemple [CS00] et [LS02] pour des propositions de modélisation du prix spot à l'aide de processus d'Ornstein-Uhlenbeck. Barlow propose dans [Bar02], en s'appuyant sur des arguments d'équilibre offre-demande, de modéliser le prix spot comme une fonction non-linéaire d'un processus d'Ornstein-Uhlenbeck. Dans [BEHN03], les auteurs proposent de modéliser l'évolution du prix spot à l'aide de processus de Lévy. Enfin, on peut consulter par exemple [Den98] pour l'étude de modèles à changement de régime.
L'introduction éventuelle de sauts pose de gros problèmes pour valoriser des contrats optionnels. En effet, la méthodologie définie par Black et Scholes dans leur article fondateur [BS73] permet de valoriser une option écrite sur une action dont l'évolution aléatoire est générée par un mouvement brownien et ne présente donc pas de sauts. Les auteurs proposent alors de donner un prix juste à l'option en construisant un portefeuille autofinançant (i.e. dont les variations sont expliquées totalement par les fluctuations du cours de l'action) qui vaut de manière certaine le payoff de l'option à sa maturité T. Les auteurs, pour déterminer un juste prix à l'option, se placent donc dans la position du vendeur, qui s'est engagé à verser un flux (*payoff*), aléatoire vu de la date 0, en T, maturité de l'option, et qui doit couvrir le risque que le payoff soit positif en T, i.e. le risque de devoir verser de l'argent à l'acheteur de l'option. Ainsi, pour une option call le payoff vaut $(S_T - K)_+$ où S_t désigne le cours de l'action à l'instant t, et le vendeur va fixer le prix V_0 auquel il va vendre l'option de telle sorte qu'il puisse ensuite trouver une stratégie de vente et d'achat de l'action S pour assurer

$$V_0 + \int \phi dS = (S_T - K)_+ \tag{0.1}$$

où $\int \phi dS$ représente les fluctuations de son portefeuille jusqu'à la maturité de l'option, qui sont expliquées par une stratégie d'achat et de vente (représentée par ϕ) et par les fluctuations de l'action S (ici, la rémunération de l'argent placé est négligée). Il n'est pas du tout évident qu'il existe un prix V_0 et une stratégie ϕ qui satisfont la parfaite réplication du payoff de l'option. Les conditions d'existence sont une modélisation de l'évolution de S à l'aide du mouvement brownien, l'absence de contraintes sur la stratégie ϕ (comme par exemple interdire d'acheter S à découvert) et la possibilité d'acheter et vendre S en temps continu.
Lorsque l'option est écrite sur un sous-jacent dont l'évolution peut présenter des discontinuités, cette méthodologie s'effondre et on ne pourra pas assurer (0.1), i.e. le vendeur

ne pourra pas annuler tout le risque qu'il a pris en vendant le contrat optionnel. Il n'est néanmoins pas contraint à la passivité et va essayer de trouver une stratégie qui minimise son risque mesuré sous la forme d'un bilan financier. Il ne va donc plus résoudre (0.1) mais un problème du type

$$\inf_{\phi,V_0} \mathcal{R}\Big(V_0 + \int \phi dS - (S_T - K)_+\Big)$$

où $V_0 + \int \phi dS - (S_T - K)_+$ est donc le bilan financier qui mesure l'écart entre l'objectif du vendeur $(S_T - K)_+$ et sa gestion de portefeuille $V_0 + \int \phi dS$ et où $\mathcal{R}$ est une mesure du risque associée au bilan financier. On comprend mieux le contexte financier de la thèse : EDF intervient sur des marchés où elle ne peut annuler tout le risque résultant de la vente d'un contrat optionnel et veut trouver des méthodes permettant de minimiser le risque lié au bilan financier entre le payoff de l'option et la stratégie de gestion.

Valorisation en marché incomplet

Ce type de situation est connue sous le terme de valorisation en marché incomplet car on ne peut justement pas répliquer parfaitement le payoff d'une option. De nombreux auteurs ont étudié des critères de valorisation dans un tel cas. Une première voie a été celle de la surréplication. Si le vendeur peut trouver un prix V_0 et une stratégie ϕ tels que

$$V_0 + \int \phi dS \geq (S_T - K)_+ \, p.s.$$

alors, le vendeur est sûr en T de pouvoir remplir son engagement de livrer $(S_T - K)_+$. Ce type de valorisation a été étudié par exemple dans [EKQ95]. Le problème est que les prix fixés de cette manière sont trop élevés par rapport aux niveaux de prix d'option observés sur les marchés financiers. Une généralisation de la surréplication a été considérée dans [FL99] où le vendeur cherche cette fois à maximiser la probabilité que $V_0 + \int \phi dS \geq (S_T - K)_+$.
D'autres critères ont été développés comme des critères moindres carrés (voir par exemple [Sch94]) où le vendeur va chercher à minimiser

$$\inf_{\phi,V_0} \mathbb{E}\big(V_0 + \int \phi dS - (S_T - K)_+\big)^2.$$

Comme ces critères considèrent de la même manière le cas où la valeur finale du portefeuille de couverture est supérieure au payoff et le cas où elle est inférieure, ces critères ont été généralisés (voir par exemple [Pha00]) en considérant la norme $\mathbf{L}^p$ (avec $p > 2$) des pertes issues d'une sous-réplication du payoff. Ceci permet de tenir compte de la disymétrie du

problème et de pondérer davantage les pertes importantes. Le vendeur cherche alors à minimiser

$$\inf_{\phi,V_0} \mathbb{E}\big|\{(S_T-K)_+ - V_0 - \int \phi dS\}^+\big|^p.$$

On peut également citer la valorisation par prix d'indifférence. Le vendeur est confronté à la situation suivante : soit il vend l'option et touche la prime V_0 qu'il gère pour obtenir un portefeuille final V_T, soit il ne fait rien. Supposons que le vendeur ait une fonction d'utilité U. Dans un cas son bilan financier en T est $V_T-(S_T-K)_+$ et dans l'autre 0. Le vendeur va évaluer ces deux cas en comparant $\mathbb{E}\big(U(V_T-(S_T-K)_+)\big)$ et $\mathbb{E}\big(U(0)\big)$. Il choisira bien sûr de vendre l'option si la première quantité est supérieure à la seconde. Le prix d'indifférence de l'option est alors défini comme le plus petit montant V_0 tel que le vendeur ait intérêt à vendre l'option. Si la fonction d'utilité du vendeur est la fonction exponentielle, ce problème a été étudié par exemple dans [REK00].
Enfin, une autre possibilité est de ne plus chercher à minimiser un bilan financier global comme $V_0+\int \phi dS-(S_T-K)_+$ qui reflète le résultat de la stratégie de gestion du vendeur sur toute la durée de vie de l'option, mais une suite de bilans financiers locaux. Cette idée a été développée dans [FS91] puis approfondie dans [Sch95] (ou [Sch01] pour une revue de résultats). C'est de cette dernière approche que cette thèse est la plus proche et nous en décrivons l'idée. Le vendeur va déterminer une série de N dates $(t_k)_{0\leq k\leq N-1}$ telles que $0=t_0<t_1<\cdots<t_{N-1}<T$ où le vendeur s'autorise à changer sa stratégie, i.e. à modifier les positions en action S qu'il détient. Le vendeur va raisonner de manière rétrograde en temps. En effet, il voudrait que son portefeuille en T soit égal à $(S_T-K)_+$. Son portefeuille à l'instant t_{N-1}, dernier instant de rebalancement avant T, est déterminé par sa valeur $V_{t_{N-1}}$ ainsi que par le montant que le vendeur va investir en action S que l'on note $\phi_{t_{N-1}}$. Une fois ces deux quantités déterminées, son portefeuille en T vaudra $V_T=V_{t_{N-1}}+\phi_{t_{N-1}}\frac{S_T-S_{t_{N-1}}}{S_{t_{N-1}}}$ (toujours en supposant un taux d'intérêt nul pour simplifier). L'information dont le vendeur dispose pour faire son choix de portefeuille est modélisée par une filtration $(\mathcal{F}_t)_{0\leq t\leq T}$. Pour traduire le fait que le vendeur ne peut anticiper sur l'évolution future de S pour prendre ses décisions d'investissement, on impose que $V_{t_{N-1}}$ et $\phi_{t_{N-1}}$ soient des variables aléatoires mesurables par rapport à $\mathcal{F}_{t_{N-1}}$. Comme le vendeur souhaite minimiser l'écart entre V_T et $(S_T-K)_+$, son problème en t_{N-1} va être de trouver deux variables aléatoires $(V_{t_{N-1}},\phi_{t_{N-1}})$ qui minimisent l'écart $V_{t_{N-1}}+\phi_{t_{N-1}}\frac{S_T-S_{t_{N-1}}}{S_{t_{N-1}}}-(S_T-K)_+$. Le critère quadratique est adopté dans cette approche comme mesure d'écart. Ainsi le vendeur en t_{N-1} cherche à minimiser

$$\mathbb{E}\left(V+\phi\frac{S_T-S_{t_{N-1}}}{S_{t_{N-1}}}-(S_T-K)_+\right)^2 \tag{0.2}$$

parmi les variables aléatoires de carré intégrable et $\mathcal{F}_{t_{N-1}}$-mesurables. Puis, en t_{N-2}, le vendeur sait que si son portefeuille en t_{N-1} vaut $V_{t_{N-1}}$, alors il aura constitué le portefeuille

optimal qui lui permettra alors de minimiser son risque sur la période $[t_{N-1}, T]$. Même problème qu'en t_{N-1} : il ne pourra déterminer un portefeuille sur l'intervalle $[t_{N-2}, t_{N-1}]$ qui égalise $V_{t_{N-1}}$ en t_{N-1}. Ainsi, il va déterminer la valeur de son portefeuille $V_{t_{N-2}}$ en t_{N-2} et sa stratégie d'investissement $\phi_{t_{N-2}}$ de telle sorte qu'ils minimisent

$$\mathbb{E}\left(V + \phi\frac{S_{t_{N-1}} - S_{t_{N-2}}}{S_{t_{N-2}}} - V_{t_{N-1}}\right)^2 \tag{0.3}$$

parmi les variables aléatoires de carré intégrable et $\mathcal{F}_{t_{N-2}}$-mesurables. Le vendeur itère le raisonnement jusqu'à l'instant 0 où il détermine (V_0, ϕ_0). V_0 peut alors être considéré comme la valeur de l'option.

Méthode numérique pour la minimisation du risque quadratique local

Cette dernière approche ayant retenu notre attention pour tenter de valoriser des options en marché incomplet, il restait à savoir comment mettre en place numériquement la méthode. Or, EDF avait justement repéré l'article [PBS01] qui proposait de résoudre les problèmes (0.2) et (0.3) par une méthode de Monte-Carlo. Le principe de la méthode est le suivant :

- On simule M trajectoires d'évolution de S que l'on note $(S^m)_{1\leq m\leq M}$.
- On cherche à résoudre (0.2). Les inconnues sont les variables aléatoires $V_{t_{N-1}}$ et $\phi_{t_{N-1}}$. On comprend aisément que l'espace des inconnues est trop vaste et que l'on va devoir chercher une approximation de $V_{t_{N-1}}$ et $\phi_{t_{N-1}}$ dans un espace de dimension finie. Tout d'abord, on va supposer le problème markovien au sens où on va chercher $V_{t_{N-1}}$ et $\phi_{t_{N-1}}$ comme des fonctions déterministes de $S_{t_{N-1}}$. L'espace de telles fonctions est encore trop vaste et on choisit donc deux bases de fonctions réelles $p_{0,N-1}(\cdot)$ et $p_{1,N-1}(\cdot)$ On va chercher à déterminer des approximations de $V_{t_{N-1}}$ et $\phi_{t_{N-1}}$ par l'intermédiaire de coefficients $\alpha_{0,N-1}$ et $\alpha_{1,N-1}$. Une fois ces coefficients déterminés on prendra $V_{t_{N-1}} \approx \alpha_{0,N-1}.p_{0,N-1}(S_{t_{N-1}})$ et $\phi_{t_{N-1}} \approx \alpha_{1,N-1}.p_{1,N-1}(S_{t_{N-1}})$. $(\alpha_{0,N-1}, \alpha_{1,N-1})$ sont déterminés comme les minimiseurs de la quantité

$$\frac{1}{M}\sum_{m=1}^{M}\left((S_T^m - K)_+ - \alpha_0.p_{0,N-1}(S_{t_{N-1}}^m) - \alpha_1.p_{1,N-1}(S_{t_{N-1}}^m)\frac{S_T^m - S_{t_{N-1}}^m}{S_{t_{N-1}}^m}\right)^2. \tag{0.4}$$

  On voit facilement l'analogie avec (0.2) : on a remplacé l'espérance par une moyenne sur M trajectoires de Monte-Carlo de S et on a remplacé l'espace des variables aléatoires $\mathcal{F}_{t_{N-1}}$-mesurables par les espaces vectoriels engendrés par $p_{0,N-1}(S_{t_{N-1}})$ et $p_{1,N-1}(S_{t_{N-1}})$.

– Une fois $(\alpha_{0,N-1}, \alpha_{1,N-1})$ déterminés, on détermine $(\alpha_{0,N-2}, \alpha_{1,N-2})$ comme les minimiseurs de

$$\frac{1}{M}\sum_{m=1}^{M}\left(\alpha_{0,N-1}.p_{0,N-1}(S^m_{t_{N-1}}) - \alpha_0.p_{0,N-2}(S^m_{t_{N-2}}) - \alpha_1.p_{1,N-2}(S^m_{t_{N-2}})\frac{S^m_{t_{N-1}} - S^m_{t_{N-2}}}{S^m_{t_{N-2}}}\right)^2 \tag{0.5}$$

soit l'analogue de (0.3).

Cette méthode proposée dans [PBS01] a beaucoup d'analogies avec [LS01] où pour valoriser une option bermudéenne, les auteurs approchent la valeur de l'option par une régression linéaire, ce qui leur permet de trouver une approximation de la stratégie d'arrêt optimale. La différence entre les deux méthodes numériques est que dans [PBS01], on cherche en même temps le prix de l'option et sa couverture. En outre, dans [LS01], les trajectoires de Monte-Carlo de S sont simulées sous la probabilité risque-neutre, ce qui n'est pas forcément le cas dans [PBS01] où Potters, Bouchaud et Sestovic en profitent pour proposer de remplacer les trajectoires Monte-Carlo de S par des historiques de cours d'action générant ainsi une méthode de valorisation sans avoir à estimer de modèle d'évolution pour S.

Cependant, il n'y a dans [PBS01] aucune analyse de convergence en fonction des paramètres de l'algorithme qui sont :

– le nombre de dates de rebalancement du portefeuille $\mathbf{N}$

– le nombre de trajectoires Monte-Carlo $\mathbf{M}$

– le choix des bases de fonctions $(\mathbf{p_{0,k}}, \mathbf{p_{1,k}})$ à chaque date de rebalancement $(t_k)_{0\leq k\leq N-1}$.

Dans un cadre de marché complet où l'évolution de S est modélisée à partir d'un mouvement brownien, on s'attend à ce que le prix et la couverture donnés par l'algorithme convergent vers le prix et la couverture classiques de l'option. Le problème est de savoir si l'algorithme converge bien, comment faire tendre N et M vers l'infini et le nombre de fonctions de base vers l'infini en même temps, et enfin de déterminer les vitesses de convergence.

Dans cette thèse, nous nous sommes restreint à ce cadre simple de marché complet, de manière à bien comprendre l'influence des différents paramètres de la méthode. Une fois la méthode bien comprise dans le cadre marché complet, il est ensuite beaucoup plus facile d'étendre l'étude réalisée dans cette thèse au cas où l'évolution de S peut présenter des discontinuités, qui est le cas intéressant dans le contexte des marchés électriques.

Les techniques de preuve et l'aspect rétrograde de l'algorithme nous ont amenés à faire le lien avec les équations différentielles stochastiques rétrogrades, sur les conseils de N. El Karoui, et à oublier quelque peu l'aspect minimisation du risque quadratique local et valorisation en marché incomplet.

En effet, il est facile comme on le verra plus tard dans la thèse de généraliser l'algorithme (0.4-0.5) au cas des équations différentielles stochastiques rétrogrades (EDSR) qui ont de

nombreuses applications en mathématiques financières comme nous allons le constater tout de suite.

Equations différentielles stochastiques rétrogrades

EDSR et réplication d'options

Résoudre une EDSR c'est trouver un couple de processus adaptés $(Y, Z)_{0\leq t\leq T}$ qui vérifie les conditions :

$$\begin{cases} Y_T = \xi, \\ -dY_t = f(t, Y_t, Z_t)dt - Z_t dW_t. \end{cases} \tag{0.6}$$

Dans cette thèse, on étudiera le cas particulier de (0.6) où la condition terminale est une fonctionnelle de la trajectoire de la solution $(X_t)_{0\leq t\leq T}$ d'une EDS . Ceci donne :

$$\begin{cases} Y_T = \Phi(X), \\ -dY_t = f(t, X_t, Y_t, Z_t)dt - Z_t dW_t, \\ dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t. \end{cases} \tag{0.7}$$

On renvoie à l'article introducteur [PP90] ou à [EKPQ97] pour les bonnes hypothèses sous lesquelles étudier l'existence et l'unicité de telles équations, que nous rappelons brièvement en annexes. Nous allons nous efforcer dans ce chapitre, en nous appuyant sur [EKPQ97], de faire le lien avec des applications en mathématiques financières et en particulier avec le problème de réplication d'options.
Considérons le cas du vendeur d'un call de payoff $(S_T - K)_+$. Comme on l'a déjà évoqué précédemment, le vendeur va fixer le prix de l'option (qu'on note ici Y_0) qu'il va demander à l'acheteur de telle sorte qu'il puisse à partir de son capital initial Y_0 trouver une stratégie d'investissement (i.e. une stratégie d'achat et de vente de l'action S) telle qu'il atteigne de manière certaine le payoff $\xi = (S_T - K)_+$ en T. Ainsi le vendeur livre à l'acheteur la valeur finale du portefeuille qu'il a géré pendant la durée de vie de l'option et a complètement couvert son risque.
On voit donc qu'en notant Y_t la valeur du portefeuille du vendeur, celui-ci aimerait réaliser $Y_T = \xi$. On voit déjà apparaitre le premier lien avec (0.6).
De plus, le vendeur est contraint de construire un portefeuille autofinancé, i.e. dont les variations sont complètement expliquées par l'évolution de S. En particulier, il ne peut pas rajouter d'argent au cours de la gestion. Ceci se traduit mathématiquement par le fait qu'entre les instants t et $t + dt$, la variation de valeur du portefeuille est déterminée

par la variation de S, dS_t, ainsi que par le montant δ_t investi dans l'action S :

$$\underbrace{dY_t}_{\text{variation de valeur du portefeuille}} = \overbrace{\frac{\delta_t}{S_t}}^{\text{nombre d'actions détenues en t}} \overbrace{dS_t}^{\text{variation du cours de l'action}} + \underbrace{(Y_t - \delta_t)}_{\text{part investie en compte d'épargne}} r dt. \tag{0.8}$$

Si à présent on suppose un modèle d'évolution pour S comme par exemple le modèle Black-Scholes

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

et que l'on reporte dans l'équation (0.8) on obtient

$$dY_t = [rY_t + \delta_t(\mu - r)]dt + \delta_t \sigma dW_t$$

ce qui en posant $Z_t = \delta_t \sigma$ et en multipliant par -1 donne

$$-dY_t = -[rY_t + Z_t \frac{\mu - r}{\sigma}]dt - Z_t dW_t.$$

En posant $f(t, y, z) = -(ry + z\frac{\mu - r}{\sigma})$, on reconnait alors que le portefeuille du vendeur est contraint à satisfaire (0.6). On en déduit donc que pour répliquer le payoff d'une option, il suffit de savoir résoudre l'EDSR associée.

EDSR et contraintes de portefeuille

Supposons que le payoff de l'option dépende de 2 actions S^1 et S^2 mais que le vendeur, pour réaliser un portefeuille de couverture, ne peut vendre et acheter qu'une des deux actions, par exemple S^1. On imagine aisément qu'il ne parviendra pas en général à répliquer l'option et on se retrouve donc dans le cadre d'un marché incomplet. Supposons que la dynamique de ces deux actions S^1 et S^2 soit :

$$\begin{pmatrix} \frac{dS_t^1}{S_t^1} \\ \frac{dS_t^2}{S_t^2} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} dt + \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix}$$

où σ_i $(i = 1, 2)$ désigne la i-ème ligne de la matrice de volatilité. Comme on l'a déjà évoqué, il peut alors essayer de trouver la plus petite valeur de portefeuille pour laquelle il va être certain de surrépliquer le payoff de l'option. Ceci se traduit par trouver le plus petit x tel qu'il puisse construire un portefeuille Y de valeur initiale x tel que

$$\begin{cases} Y_T = \xi, \\ dY_t = [rY_t + \delta_t(\mu_1 - r)]dt + \delta_t \sigma_1 dW_t - dC_t. \end{cases} \tag{0.9}$$

où δ_t est le montant investi à l'instant t dans l'action S^1. C est ici un processus de consommation croissant. Sa présence est le signe que l'on va surrépliquer l'option et donc que pour assurer $Y_T = \xi$, il faut autoriser une consommation de la richesse générée par le portefeuille. (0.9) n'est pas une EDSR à cause de la présence du processus de consommation mais on peut néanmoins montrer que la solution du problème de surréplication est la limite croissante d'une suite de solutions d'EDSR.
Pour résoudre le problème d'incomplétude du marché, le vendeur peut également décider qu'il est trop fort de chercher une stratégie surrépliquant le payoff de l'option, et qu'il va décider de répliquer le payoff de l'option non pas par une stratégie autofinancée (ce n'est pas possible sinon on serait en marché complet) mais par une stratégie où il s'autorise à ajouter et retirer de l'argent du portefeuille. Dans notre problème à 2 actions, il va donc chercher à résoudre :

$$\begin{cases} Y_T = \xi, \\ dY_t = [rY_t + \delta_t(\mu_1 - r)]dt + \delta_t\sigma_1 dW_t + d\phi_t. \end{cases} \tag{0.10}$$

Bien sûr, le problème posé comme cela est facile à résoudre, il suffit de ne rien faire sur l'intervalle $[0, T[$ et en T de prendre $\phi_T = \xi - Y_0$. On va donc imposer que le processus ϕ soit une martingale orthogonale à $\sigma_1 W_t$. En effet, on sait qu'on ne pourra pas résoudre le problème en imposant $\phi = 0$ (cas marché complet) mais on va tout de même imposer que l'on ne rajoute pas d'argent en moyenne ce que traduit la condition sur ϕ. (0.10) n'est pas exactement une EDSR au sens de (0.6) mais on se réfère à [EKPQ97] pour montrer qu'on peut généraliser la notion d'EDSR pour englober (0.10). De plus, on peut montrer que (0.10) est en fait la même approche que la minimisation du risque quadratique local et nous nous référons une fois encore à [EKPQ97] pour des développements sur ce point.

EDSR et imperfections de marché

Le vendeur peut avoir également à subir des imperfections de marché qui ne sont pas prévues dans la méthodologie Black-Scholes en marché complet. Par exemple, pour construire son portefeuille de réplication, il peut subir la contrainte d'emprunter de l'argent à un taux R supérieur au taux r auquel est rémunéré l'argent qu'il peut placer sur un compte d'épargne (voir par exemple [Ber95]). Ainsi, si S suit le modèle Black-Scholes, la dynamique de son portefeuille de couverture sera :

$$dY_t = [rY_t + \delta_t(\mu - r)]dt + \delta_t\sigma dW_t - (R - r)(Y_t - \delta_t)^- dt \tag{0.11}$$

où le terme $(Y_t - \delta_t)^-$ représente le montant emprunté à la banque. En posant $f(t, y, z) = -ry - z\frac{\mu - r}{\sigma} + (R - r)(y - \frac{z}{\sigma})^-$, on constate une fois encore que le problème de réplication avec deux taux d'intérêt différents est équivalent à résoudre une EDSR.

EDSR et utilité récursive

Les EDSR ont aussi un rapport avec la modélisation de l'utilité d'un agent qui peut consommer sur une durée $[0, T]$. On suppose que l'agent touche une prime Y en T et qu'à un instant t, son utilité Y_t dépend de sa consommation c_t et de son utilité future (correspondant à ses consommations à venir). Ainsi, on obtient une forme récursive pour l'utilité qui satisfait $Y_T = Y$ et

$$-dY_t = f(c_t, Y_t)dt.$$

Si on généralise cela à un cas incertain, on arrive finalement à ce que l'utilité de l'agent $(Y_t)_{0\leq t\leq T}$ satisfait une EDSR. On peut consulter [DE92b] ou [DE92a] à ce propos.

EDSR et équations aux dérivées partielles semilinéaires

Les EDSR ont également des liens avec les équations aux dérivées partielles semilinéaires, de la même façon que la solution d'un problème de Cauchy peut se représenter à l'aide de la solution d'une équation différentielle stochastique progressive, via les formules de Feynman Kac. En particulier, on pourra consulter par exemple [PP92] ou [Par98] pour voir comment on peut représenter à l'aide d'une EDSR la solution de l'équation aux dérivées partielles

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) + L_t u(t,x) + f(t,x,u(t,x),(\partial u\sigma)(t,x)) = 0, \\ u(T,x) = g(x) \end{cases}$$

où L_t est l'opérateur défini par

$$L_t = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(t,x)\frac{\partial}{\partial x_i}.$$

En conclusion de ce chapitre, on voit que savoir résoudre une EDSR a de nombreuses implications en mathématiques financières et qu'il est donc particulièrement important d'avoir des méthodes numériques efficaces pour leur résolution.

Résultats obtenus

Cette thèse est consacrée à l'analyse d'un algorithme (et de ses variantes) de résolution d'EDSR. Plusieurs méthodes d'approximation ont déjà été proposées dans la littérature. Tout d'abord, dans [Zha01] puis dans [Zha04], Jianfeng Zhang établit des propriétés fines de solutions d'EDSR, et en particulier étudie la régularité des trajectoires de Z ainsi que sa régularité $\mathbf{L}^2$. Pour une définition précise, nous renvoyons à [Zha04]. Dans cette thèse,

cela nous permettra dans la partie I de quantifier la vitesse de convergence de l'approximation de $\int_0^T Z_t dW_t$ par des intégrales simples. Dans [MZ05], Ma et Zhang étendent ces résultats au cas des équations différentielles stochastiques rétrogrades réfléchies (EDSRR). Nous utiliserons également ces résultats dans la partie III où nous étendons l'algorithme proposé au cas des EDSRR. En plus de ces études théoriques, Zhang propose un schéma numérique d'approximation dans [Zha04]. Ce schéma (qui suppose une discrétisation de $[0,T]$ à l'aide de dates $(t_k)_{0\leq k\leq N}$) est rétrograde en partant de la condition terminale de l'EDSR, puis les approximations de Y et Z à l'instant t_k sont construites à partir de leurs approximations à l'instant t_{k+1}. Le problème est que ce schéma reste théorique car ces approximations sont définies par des espérances, et l'auteur ne donne pas d'explications sur la méthode employée pour les calculer. Nous verrons dans cette thèse qu'il n'est pas facile de comprendre comment calculer itérativement ces espérances et comment se propage l'erreur commise dans l'algorithme.

Dans [Bal97], Bally présente un schéma de discrétisation en temps de (0.6) basé sur les temps de sauts d'un processus de Poisson. Cette astuce lui permet de ne pas utiliser la régularité inconnue de Z et de déduire une vitesse de convergence par rapport à l'intensité du processus de Poisson. Cependant, le schéma numérique, pour être implémentable, nécessite le calcul d'intégrales de grande dimension, et ce n'est pas traité dans [Bal97].

Des schémas numériques totalement implémentables ont néanmoins été proposés. Ainsi, dans [DMP96], Douglas, Ma et Protter proposent un schéma numérique pour une généralisation des EDSR, les équations différentielles stochastiques progressives-rétrogrades, basé sur le four-step scheme ([MPY94]) qui suppose des hypothèses restrictives sur la régularité des coefficients. Par exemple, si la condition terminale ξ de (0.6) est de la forme $g(X_T)$ où X_T est la valeur terminale d'une EDS, alors g doit être de classe C^4 plus une régularité Hölder sur la dérivée quatrième. De plus, leur schéma numérique nécessite de résoudre une équation aux dérivées partielles quasilinéaire.

Dans [Che97], Chevance propose pour approcher (0.7) un algorithme entièrement implémentable, basé sur des techniques de quantification aléatoire. Une borne pour l'erreur sur Y commise dans l'algorithme est établie mais sous des hypothèses de régularité très forte sur ϕ et sur le driver f. En outre, l'erreur sur Z n'est pas analysée.

Dans [BDM01], Briand, Delyon et Mémin proposent un schéma d'approximation de (0.6) basé sur une approximation du mouvement brownien par une somme de variables aléatoires de Bernoulli. Ceci donne un schéma numérique facilement implémentable dont les auteurs prouvent la convergence. Cependant aucune vitesse de convergence n'est obtenue dans [BDM01]. Dans [MPSMT02], les auteurs étudient également la convergence faible de la solution d'une EDSR où le mouvement brownien a été remplacé par une marche aléatoire vers la solution de (0.6), mais dans le cas où f est indépendant de Z (ou en dépend linéairement).

Dans [BP03], Bally et Pagès proposent un schéma numérique pour les EDSR réfléchies

du type (0.7) basé sur des techniques de quantification. L'idée est de remplacer X par une chaine de Markov à nombre d'états finis qui sont les points d'une grille de quantification optimale. Une fois estimées les probabilités de transition sur cette grille, les auteurs peuvent implémenter facilement une équation de programmation dynamique pour résoudre l'EDSRR.
Dans [DM04], Delarue et Menozzi proposent une méthode de résolution des équations différentielles stochastiques progressives-rétrogrades, basée sur des techniques de quantification à l'aide de grilles déterministes. Les auteurs analysent complètement l'influence des paramètres de leur méthode.
Enfin, dans [BT04], Bouchard et Touzi proposent une méthode d'approximation issue d'une discrétisation en temps de (0.7). Une fois cette discrétisation en temps effectuée, les auteurs remarquent que la solution de cette approximation en temps se représente à chaque instant de discrétisation sous la forme d'espérances conditionnelles. Ils utilisent alors des techniques du calcul de Malliavin pour approcher ces espérances conditionnelles (voir également [BET04]). Les auteurs analysent la propagation de l'erreur sur Y mais pas sur Z. En outre, le schéma numérique semble difficile et coûteux à mettre en place en dehors de la dynamique Black-Scholes pour X.

Voyons à présent le type d'algorithme que nous étudions ici. Vu le point de départ de la thèse exposé dans le chapitre , on ne sera pas surpris d'apprendre que :
- cet algorithme est basé sur une discrétisation en temps de l'intervalle $[0,T]$ à l'aide de N dates $(t_k = kh)_{0\leq k\leq N-1}$ où $h = \frac{T}{N}$ désigne le pas de temps de l'algorithme ;
- l'algorithme est rétrograde en partant de la date T ;
- à chaque instant t_k, on approche la solution de (0.7) en résolvant un problème de régression du type (0.5), basé sur M trajectoires de Monte-Carlo. Ce problème de régression permet de déterminer des coefficients $\alpha_{0,k}$ et $\alpha_{1,k}$ et $Y_{t_k} \approx \alpha_{0,k}.p_{0,k}(X_{t_k})$, $Z_{t_k} \approx \alpha_{1,k}.p_{1,k}(X_{t_k})$ où $p_{0,k}$ et $p_{1,k}$ sont des bases de fonctions.

Nous allons décrire un peu plus en détail (en renvoyant aux parties I et II pour une description complète) les deux algorithmes proposés et étudiés dans cette thèse. Nous nous plaçons dans le cadre simplifié où la condition terminale de (0.7) est de la forme $\phi(X_T)$ et renvoyons à la partie I pour voir comment procéder dans le cas où la condition terminale est une fonctionnelle de toute la trajectoire de X. Nous supposons également que W est un mouvement brownien de dimension 1. Dans les deux algorithmes, nous approchons X par son schéma d'Euler que nous notons X^N et dont nous simulons M trajectoires Monte-Carlo aux dates de discrétisation $(t_k)_{0\leq k\leq N}$ notées $(X_{t_k}^{N,m})_{0\leq k\leq N, 1\leq m\leq M}$. Ainsi, on a

$$X_{t_{k+1}}^{N,m} = X_{t_k}^{N,m} + b(t_k, X_{t_k}^{N,m})h + \sigma(t_k, X_{t_k}^{N,m})\Delta W_k^m$$

où $\Delta W_k = W_{t_{k+1}} - W_{t_k}$ et $(\Delta W_k^m)_{1\leq m\leq M}$ sont des simulations indépendantes de ΔW_k.

Dans la partie I, l'algorithme se présente sous la forme suivante :

→ **Initialisation**. On définit le vecteur de $\mathbb{R}^M$ $(Y_{t_N}^{N,I,I,M,m})_{1\leq m\leq M}$ par $Y_{t_N}^{N,I,I,M,m} = \phi(X_{t_N}^{N,m})$.

→ **Itération (rétrograde en temps) à l'instant $t_k < T$.** Rappelons que l'on a choisi à l'instant t_k deux bases de fonctions $p_{0,k}$ et $p_{1,k}$. On définit alors les vecteurs $(\alpha_{l,k}^M)_{0\leq l\leq 1}$ comme les minimiseurs en (α_0, α_1) de la quantité

$$\frac{1}{M}\sum_{m=1}^{M}\{\big(Y_{t_{k+1}}^{N,I,I,M,m} + hf(t_k, X_{t_k}^{N,m}, \alpha_0.p_{0,k}(X_{t_k}^{N,m}), \alpha_1.p_{1,k}(X_{t_k}^{N,m})) \\ - \alpha_0.p_{0,k}(X_{t_k}^{N,m}) - \alpha_1.p_{1,k}(X_{t_k}^{N,m})\Delta W_k^m\big)^2\}.$$

Enfin, on définit $(Y_{t_k}^{N,I,I,M,m})_{1\leq m\leq M}$ par $Y_{t_k}^{N,I,I,M,m} = \alpha_{0,k}^M.p_{0,k}(X_{t_k}^{N,m})$.

Notre approximation de (Y_{t_k}, Z_{t_k}) est alors $\big(\alpha_{0,k}^M.p_{0,k}(X_{t_k}^N), \alpha_{1,k}^M.p_{1,k}(X_{t_k}^N)\big)$. On peut remarquer qu'à l'instant t_k $(\alpha_{0,k}^M, \alpha_{1,k}^M)$ est défini comme un point fixe et un algorithme réellement implémentable nécessite de préciser comment on approche ce point fixe. L'algorithme exposé de manière complète dans la partie I est basé sur I itérations de Picard à chaque instant t_k, ce qui explique la présence de l'indice I dans notre approximation $Y_{t_k}^{N,I,I,M}$. L'algorithme nécessite également d'utiliser des fonctions de troncature et nous renvoyons à la partie I pour leur description.

L'algorithme de la partie II se présente également sous une forme rétrograde :

→ **Initialisation**. On pose $\alpha_{0,N}^M.p_{0,N}(\cdot) = \phi(\cdot)$.

→ **Itération (rétrograde en temps) à l'instant $t_k < T$.** On simule $(\Delta\tilde{W}_k^m)_{1\leq m\leq M}$, une copie indépendante de $(\Delta W_k^m)_{1\leq m\leq M}$. On en déduit $(\tilde{X}_{t_{k+1}}^{N,m})_{1\leq m\leq M}$, une copie de $(X_{t_{k+1}}^{N,m})_{1\leq m\leq M}$ définie par

$$\tilde{X}_{t_{k+1}}^{N,m} = X_{t_k}^{N,m} + b(t_k, X_{t_k}^{N,m})h + \sigma(t_k, X_{t_k}^{N,m})\Delta\tilde{W}_k^m.$$

On définit alors le vecteur $\alpha_{1,k}^M$ comme le minimiseur en α_1 de

$$\frac{1}{M}\sum_{m=1}^{M}\big(\alpha_{0,k+1}^M.p_{0,k+1}(\tilde{X}_{t_{k+1}}^{N,m})\frac{\Delta\tilde{W}_k^m}{h} - \alpha_1.p_{1,k}(X_{t_k}^{N,m})\big)^2$$

puis le vecteur $\alpha_{0,k}^M$ comme le minimiseur en α_0 de

$$\frac{1}{M}\sum_{m=1}^{M}\{\big(\alpha_{0,k+1}^M.p_{0,k+1}(\tilde{X}_{t_{k+1}}^{N,m}) + hf(t_k, X_{t_k}^{N,m}, \alpha_{0,k+1}^M.p_{0,k+1}(\tilde{X}_{t_{k+1}}^{N,m}), \alpha_{1,k}^M.p_{1,k}(X_{t_k}^{N,m})) \\ - \alpha_0.p_{0,k}(X_{t_k}^{N,m})\big)^2\}.$$

Notre approximation de (Y_{t_k}, Z_{t_k}) est $(\alpha^M_{0,k}.p_{0,k}(X^N_{t_k}), \alpha^M_{1,k}.p_{1,k}(X^N_{t_k}))$. Ajoutons que l'implémentation de cet algorithme nécessite d'utiliser une localisation de l'EDSR (0.7) ainsi que des fonctions de troncature. Nous renvoyons à la partie II pour plus de précisions.

Les paramètres de ces deux algorithmes sont le pas de temps $h = \frac{T}{N}$, le choix des bases de fonctions et le nombre de trajectoires Monte-Carlo M. Dans la partie I, nous analysons progressivement les trois types d'erreurs pour le premier algorithme. L'erreur dûe à h est bornée dans le Théorème I.1 qui nous dit, en nous appuyant sur des résultats de Zhang ([Zha04]), que cette erreur est d'ordre $\sqrt{h}$ dans le cas où la condition terminale $\Phi(X)$ de (0.7) est une fonctionnelle Lipschitz de la trajectoire de X et sous des hypothèses standard sur b, σ et f (voir hypothèses 0.1, I.1 et I.2). Puis, l'erreur dûe au choix des bases de fonctions est analysée dans le Théorème I.2 où l'on voit que cette erreur est majorée par la somme d'erreurs de projections aux instants de discrétisation : c'est le cas le plus favorable que l'on pouvait espérer car la propagation de l'erreur dûe aux fonctions de bases est seulement additive et n'explose pas au fur et à mesure de l'itération rétrograde de l'algorithme. Enfin, nous analysons l'erreur dûe aux simulations à travers les théorèmes I.3 et I.4 qui sont de nature différentes : le théorème I.4 établit un théorème la limite centrale lorsque M tend vers l'infini alors que le théorème I.3 est un résultat non asymptotique qui borne l'erreur dûe aux simulations en fonction de M et des moments des fonctions de base utilisées dans l'algorithme. L'association de ces résultats fournit un contrôle de l'erreur globale commise dans l'algorithme.
Dans la partie II, nous analysons le deuxième algorithme qui nous permet de retrouver les résultats des théorèmes I.1 et I.2 et d'améliorer grandement les résultats du théorème I.3 en nous fournissant des bornes ne dépendant plus des moments des fonctions de base mais uniquement de leur nombre. Ceci permet de faire varier **en même temps** les paramètres N, M et le nombre de fonctions de base pour atteindre une précision désirée.
Dans la partie III, nous présentons des résultats complémentaires à ceux exposés dans les parties I et II. Nous montrons tout d'abord dans le chapitre 1 comment dans le cas où f est linéaire en (y, z) les vitesses de convergence de la discrétisation en temps données dans le théorème I.1 peuvent être améliorées. On montre en effet que l'erreur d'approximation sur Y est alors d'ordre h et non plus $\sqrt{h}$. Puis, dans le chapitre 2, nous généralisons les résultats de la partie II au cas d'une EDSR réfléchie.
Ceci constitue donc un ensemble de nouveaux résultats permettant de traiter le cas de conditions terminales path-dependent et donnant des estimations sur Y et Z, en fonction des paramètres N, M et des fonctions de base.
Enfin dans la partie numérique, nous testons tout d'abord les résultats obtenus dans la partie II et vérifions numériquement l'existence de règles empiriques pour faire varier simultanément les paramètres N, M et le nombre de fonctions de base. Si ces règles ne sont pas vérifiées, l'algorithme ne converge pas. En outre, nous montrons que les résultats observés numériquement sont souvent plus optimistes que ceux donnés dans la partie II.

Nous testons également les résultats obtenus dans la partie III pour le cas d'options américaines en dimension 1. Nous en profitons pour tester numériquement deux algorithmes alternatifs à celui de la partie III. Le premier est basé sur une méthode de pénalisation, utilisée dans [EKKP$^+$97] pour montrer l'existence et l'unicité de la solution à une EDSRR et le second est basé sur une méthode de régularisation proposée dans [BCFEK02]. L'utilisation numérique de ces idées est à notre connaissance nouvelle et les résultats obtenus numériquement montrent leur intérêt. Enfin, dans le chapitre 7, nous abandonnons la phase de tests systématiques des chapitres précédents pour traiter des cas intéressants du point de vue financier. Tout d'abord, nous délaissons le modèle Black-Scholes pour X pour traiter un cas de valorisation d'options sur le marché électrique avec le modèle factoriel pour les contrats forward. Puis nous valorisons une option dans le cas d'un différentiel de taux d'intérêt et vérifions que la non linéarité de f induite par les taux d'intérêt différents a un réel impact sur le prix de l'option. Nous traitons également le cas d'une option asiatique pour montrer comment utiliser l'algorithme pour valoriser une option path-dependent et enfin le cas d'une option américaine en dimension 3 pour montrer l'impact de la dimension dans l'algorithme.

Cet algorithme nous semble présenter des avantages par rapport aux approches proposées précédemment dans la littérature.
On voit tout d'abord que cet algorithme se rapproche de [BT04] et [Zha04] au sens où on commence par discrétiser en temps (0.7) puis on approche la solution de cette discrétisation par un algorithme rétrogade. Notre algorithme se différencie alors par la manière dont on définit notre approximation à chaque instant de discrétisation. Nous cherchons en effet une approximation de (Y_{t_k}, Z_{t_k}) dans les espaces vectoriels engendrés par des bases de fonctions $p_{0,k}$ et $p_{1,k}$. Les coefficients de notre approximation sur ces bases sont trouvés en résolvant un simple problème de moindres carrés basé sur M trajectoires Monte-Carlo. Notre algorithme se distingue alors de [Zha04] car il est réellement implémentable et de [BT04] par sa simplicité d'utilisation (pour des modélisation de X très générales) et sa faible complexité. En effet, la mise en place de [BT04] nécessite de simuler M trajectoires Monte-Carlo de X à chaque instant de discrétisation soit au total NM trajectoires. Notre approche utilise les mêmes M trajectoires Monte-Carlo à chaque pas de temps. En outre, nous bornons l'erreur commise non seulement sur Y mais également sur Z et notre approche se généralise au cas d'une EDSR avec un aléa martingale non-brownien.
Par rapport à [BDM01] et [MPSMT02], l'avantage principal de notre méthode est qu'elle permet d'obtenir une vitesse de convergence pour l'algorithme, qui résulte du fait que nous ne remplaçons pas le mouvement brownien par une marche aléatoire.
Enfin, nous n'avons pas besoin de grilles de quantification contrairement à [BP03] et [DM04], ce qui n'est pas un avantage décisif dans le cas où l'aléa est brownien car des grilles de quantification optimale sont disponibles dans ce cas, mais le devient dans le

cas où l'aléa n'est plus brownien. En outre, les méthodes de quantification peuvent être vues comme des cas particuliers de notre méthode en utilisant des fonctions de base $p_{0,k}$ et $p_{1,k}$ qui sont les indicatrices des cellules de Voronoi d'une grille de quantification optimale. Cependant, notre méthode ne restreint pas du tout les bases de fonctions à avoir cette forme et ceci a un double intérêt. D'une part, même si l'on utilise des bases locales pour définir $p_{0,k}$ et $p_{1,k}$ dans l'esprit des cellules de Voronoï, l'ajout de polynômes sur ces supports locaux, ce qui est possible dans notre méthode mais pas dans [BP03] ou [DM04], réduit beaucoup le nombre de cellules à utiliser dans l'algorithme pour une précision donnée (voir partie IV). D'autre part, on peut utiliser dans les bases $p_{0,k}$ et $p_{1,k}$ des fonctions globales dont on pense qu'elles vont bien approcher les vraies fonctions de prix et de couverture de l'option. Ceci est particulièrement utile en grande dimension pour obtenir de bonnes approximations du prix et de la couverture avec un faible coût de calcul.

Ce type d'algorithme, basé sur des régressions itérées sur des bases de fonctions, que nous proposons et analysons ici, n'est pas nouveau. Comme nous l'avons déjà indiqué précédemment, une méthode analogue a déjà été proposée pour le cas d'options américaines (driver f indépendant de Z) dans [LS01] et un théorème de la limite centrale pour cette méthode a été montré dans [CLP02]. De plus, l'idée que l'on peut contrebalancer le problème de la grande dimension de X par un bon choix des bases de fonctions en fait un algorithme populaire dans le domaine de l'apprentissage (statistical learning). En effet, dans des problèmes de grande dimension, on ne peut se permettre de réaliser des pavages fins de l'espace de la variable d'état X (comme dans les méthodes de quantification par exemple), et l'on a recours à un choix de bonnes fonctions de base. On trouve dans [TVR99] ou [Egl05] des analyses d'algorithmes proches de celui que nous analysons ici. **Cependant, c'est la première fois, à notre connaissance, que sont établies des bornes d'erreurs en fonction de tous les paramètres de l'algorithme (ce qui permet de les faire varier simultanément). En outre, c'est la première fois que sont discutés explicitement différents choix de fonctions de base qui permettent de retrouver numériquement les règles de variation simultanée.** Cette analyse nous parait fondamentale car il n'est pas du tout intuitif de savoir comment faire varier ces paramètres pour obtenir une précision désirée : un mauvais choix de variation simultanée conduit à une absence de convergence de l'algorithme ou à une convergence vers une valeur erronée. A présent que cette analyse sur l'influence des paramètres N, M et choix des fonctions de base dans l'erreur commise dans l'algorithme est effectuée, cet algorithme nous parait être très prometteur (surtout en grande dimension) de par sa flexibilité et sa faible complexité.
La thèse est présentée en deux parties : une partie théorique de la page 33 à 145 analyse mathématiquement les algorithmes, une partie numérique les testant sur des cas standards puis sur des exemples plus élaborés.

Notations

Variable d'état

Dans toute cette thèse, nous considérons des équations différentielles stochastiques rétrogrades sur un intervalle $[0,T]$ où la condition terminale et le driver s'expriment en fonction d'un vecteur de variables d'état de dimension d que nous noterons toujours par X. Dans tous les chapitres, nous supposerons que X est l'unique solution forte de l'équation différentielle stochastique

$$X_{i,t} = X_{i,0} + \int_0^t b_i(s, X_s)ds + \int_0^t \sigma_i(s, X_s)dW_s,\ 1 \leq i \leq d \tag{0.12}$$

où W est un mouvement brownien standard de dimension q et où σ_i désigne la i-ème ligne de la matrice σ de taille $d \times q$. Nous ferons dans tous les chapitres l'hypothèse :

Hypothèse 0.1 *$\forall i,\ 1 \leq i \leq d$, les fonctions $b_i(t,x)$ et $\sigma_i(t,x)$ sont uniformément Lipschitz par rapport à $(t,x) \in [0,T] \times \mathbb{R}^d$.*

Lorsque des hypothèses supplémentaires seront nécessaires sur les coefficients b et σ nous le mentionnerons en début de chapitre.
Donnons un exemple de ce que sera X dans les applications en finance. On fixe $d = q = 1$ et on considère une action de prix S_t et de dynamique le modèle de Black-Scholes

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t.$$

En appliquant la formule d'Itô, $U_t = \log(S_t)$ suit quant à lui la dynamique

$$dU_t = (\mu - \frac{\sigma^2}{2})dt + \sigma dW_t.$$

Ainsi, avec ce choix de modélisation pour S, U et S vérifient l'hypothèse 0.1. Dans les parties I et II, on verra que les résultats présentés sont valables sous la seule hypothèse 0.1, sans supposer d'uniforme ellipticité sur la matrice $\sigma\sigma^*$. Ainsi pour les parties I et II, on pourra appliquer l'algorithme d'approximation d'EDSR soit avec U comme variable d'état soit avec S. C'est la raison pour laquelle nous restons volontairement vagues sur la signification de X.

Constantes

Dans toute la thèse, pour ne pas alourdir les calculs, on note par C une constante qui ne dépend pas des paramètres de l'algorithme N, M, fonctions de base ou vecteur de localisation de l'EDSR (voir partie II). C peut dépendre de T, des coefficients b et σ de la diffusion X ou des constantes de Lipschitz de la condition terminale et du driver f. La valeur de C peut changer (ce qui sera fréquemment le cas) de ligne en ligne.

Partie théorique

Première partie

Un premier algorithme pour la résolution des équations différentielles stochastiques rétrogrades

Nous allons dans cette partie proposer un algorithme pour la résolution d'équations différentielles stochastiques rétrogrades (EDSR). Cette partie a donné lieu à un article [GLW05] à paraitre dans Annals of Applied Probability.

Chapitre 1

Cadre, hypothèses et algorithme

1.1 Cadre et hypothèses

Nous considérons dans cette partie l'équation rétrograde suivante :

$$Y_t = \Phi(\mathbf{X}) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s \tag{I.1.1}$$

où $f : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^q \to \mathbb{R}$ et où on rappelle que X est la solution de l'équation différentielle stochastique en dimension d :

$$X_{i,t} = X_{i,0} + \int_0^t b_i(s, X_s)ds + \int_0^t \sigma_i(s, X_s)dW_s,\ 1 \leq i \leq d, \tag{I.1.2}$$

où σ_i désigne la i-ème ligne de la matrice σ de taille $d \times q$. Nous renvoyons à l'annexe A.1 pour des rappels de résultats sur les EDSR. On pourra remarquer en comparant avec (A.1) que cela revient à considérer un driver f où la dépendance en X traduit explicitement la dépendance en ω. Dans cette partie, on supposera que l'hypothèse 0.1 est vérifiée. En outre, on suppose :

Hypothèse I.1 *Le driver f satisfait la propriété :*

$$|f(t_2, x_2, y_2, z_2) - f(t_1, x_1, y_1, z_1)| \leq C_f(|t_2 - t_1|^{1/2} + |x_2 - x_1| + |y_2 - y_1| + |z_2 - z_1|)$$

pour tout $(t_1, x_1, y_1, z_1), (t_2, x_2, y_2, z_2) \in [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^q$. De plus, $\sup_{0 \leq t \leq T} |f(t, 0, 0, 0)| < \infty$.

Hypothèse I.2 *La condition terminale Φ est $\mathbf{L}^\infty$-Lipschitz, i.e. pour toutes fonctions continues $\mathbf{x^1}$ et $\mathbf{x^2}$ on a*

$$|\Phi(\mathbf{x^1}) - \Phi(\mathbf{x^2})| \leq C \sup_{t \in [0,T]} |x_t^1 - x_t^2|.$$

De plus, $|\Phi(\mathbf{0})| < \infty$ où $\mathbf{0}$ est la fonction constante égale à 0 sur $[0,T]$.

L'hypothèse 0.1 assure qu'il existe une unique solution forte à (I.1.2). De plus, il n'est pas difficile de voir que le couple $(f, \Phi(\mathbf{X}))$ est un couple de paramètres standard (voir annexe A.1) sous les hypothèses I.1 et I.2. Ainsi, le théorème A.1 permet de conclure à l'existence et l'unicité d'une solution à (I.1.1).

1.2 Description de l'algorithme

Comme nous l'avons expliqué dans l'introduction, notre but est de fournir une approximation du couple (Y, Z) solution de (I.1.1). Avant de décrire notre algorithme, nous allons définir quelques notations qui sont utiles à son exposé :

1.2.1 Notations et hypothèses supplémentaires

Discrétisation en temps

On suppose que l'on a fixé un pas de temps $h = \frac{T}{N}$ (supposé plus petit que 1) et des dates de discrétisation $t_k = kh$, $0 \leq k \leq N$. A ces dates de discrétisation, on approche X_{t_k} par $X^N_{t_k}$ son schéma d'Euler.

Approximation de la condition terminale

La condition terminale $\Phi(\mathbf{X})$ est approchée par $\Phi^N(P^N_{t_N})$, où Φ^N est une fonction déterministe et $(P^N_{t_k})_{0\leq k\leq N}$ est une chaîne de Markov de dimension $\mathbf{d'} \geq \mathbf{d}$ dont les d premières composantes sont données par celles de $(X^N_{t_k})_{0\leq k\leq N}$. En d'autres termes, on ajoute éventuellement d'autres variables d'état pour rendre markovienne la dynamique implicite de la condition terminale. On suppose que $P^N_{t_k}$ est $\mathcal{F}_{t_k}$-mesurable et que $\mathbb{E}[\Phi^N(P^N_{t_N})]^2 < \infty$. Bien sûr cette approximation dépend fortement du type de la condition terminale et son impact sera mesuré par l'erreur $\mathbb{E}|\Phi(\mathbf{X}) - \Phi^N(P^N_{t_N})|^2$ (voir théorème I.1 dans le prochain chapitre). Des exemples de fonctions Φ^N seront également donnés dans la section suivante. Une autre hypothèse est requise pour montrer qu'une certaine EDSR à temps discret $(Y^N_{t_k})_k$ peut être représentée comme une fonction Lipschitz $y^N_k(\cdot)$ de $P^N_{t_k}$ (voir proposition I.2 dans le prochain chapitre). Cette propriété sera surtout utilisée dans la partie IV pour quantifier explicitement des erreurs d'approximation.

Hypothèse I.3 *La fonction $\Phi^N(\cdot)$ est Lipschitz (uniformément en N) et $\sup_N |\Phi^N(\mathbf{0})| < \infty$. De plus, $\mathbb{E}|P^{N,k_0,x}_{t_N} - P^{N,k_0,x'}_{t_N}|^2 + \mathbb{E}|P^{N,k_0,x}_{t_{k_0+1}} - P^{N,k_0,x'}_{t_{k_0+1}}|^2 \leq C|x - x'|^2$ uniformément en k_0 et N.*

Ici, $(P^{N,k_0,x}_{t_k})_k$ représente la chaîne de Markov $(P^N_{t_k})_k$ partant de $P^N_{t_{k_0}} = x$. De plus, comme nous traitons des propriétés de flot de $(P^N_{t_k})_k$, nous utilisons la représentation

standard de cette chaîne de Markov comme une suite récurrente aléatoire de la forme $P^N_{t_k} = F^N_k(U_k, P^N_{t_{k-1}})$ où $(F^N_k)_k$ sont des fonctions mesurables et $(U_k)_k$ sont des variables aléatoires i.i.d.

Projection sur des bases de fonctions

- La projection $\mathbf{L}_2(\Omega, \mathbb{P})$ d'une variable aléatoire U sur une famille finie $\phi = [\phi_1, \cdots, \phi_n]^*$ (considérée comme un vecteur aléatoire) est notée $\mathcal{P}_\phi(U)$. On désigne alors par $\mathcal{R}_\phi(U) = U - \mathcal{P}_\phi(U)$ l'erreur de projection.
- A chaque instant t_k, pour approcher respectivement Y_{t_k} et Z_{l,t_k} (Z_{l,t_k} est la l-ème composante de Z_{t_k}, $1 \leq l \leq q$), on va utiliser respectivement les bases de fonctions $p_{0,k}(P^N_{t_k})$ et $p_{l,k}(P^N_{t_k})$ $(1 \leq l \leq q)$, que l'on écrira aussi $p_{0,k}$ et $p_{l,k}$ $(1 \leq l \leq q)$ pour simplifier. Dans toute la suite, de manière à simplifier les écritures, on fera la confusion entre base de fonctions $\big(p_{l,k}(\cdot)\big)$ et base aléatoire $\big(p_{l,k}(P^N_{t_k})\big)$. Des choix explicites de $p_{0,k}$ et $(p_{l,k})_{1\leq l\leq q}$ seront donnés dans la partie IV. Les coefficients de projection seront notés $\alpha_{0,k}, \alpha_{1,k}, \cdots, \alpha_{q,k}$ (vus comme des vecteurs colonnes). On suppose que $\mathbb{E}|p_{l,k}|^2 < \infty$ $(0 \leq l \leq q)$ et sans perte de généralité que $\mathbb{E}(p_{l,k}p^*_{l,k})$ est inversible, ce qui assure l'unicité des coefficients de la projection $\mathcal{P}_{p_{l,k}}$ $(0 \leq l \leq q)$.
- Pour simplifier, on écrit $f_k(\alpha_{0,k}, \cdots, \alpha_{q,k})$ ou $f_k(\alpha_k)$ pour $f(t_k, X^N_{t_k}, \alpha_{0,k} \cdot p_{0,k}, \cdots, \alpha_{q,k} \cdot p_{q,k})$.
- Toujours pour simplifier, on note $\mathbb{E}_k(.) = \mathbb{E}(.|\mathcal{F}_{t_k})$ et $\Delta W_k = W_{t_{k+1}} - W_{t_k}$ ($\Delta W_{l,k}$ composante par composante).
- On définit le vector (colonne) v_k par $[v_k]^* = ({p_{0,k}}^*, {p_{1,k}}^* \frac{\Delta W_{1,k}}{\sqrt{h}}, \cdots, {p_{q,k}}^* \frac{\Delta W_{q,k}}{\sqrt{h}})$.
- Pour un vecteur x, $|x|$ désigne sa norme euclidienne. La dimension relative est implicite. Pour un entier M et $x \in \mathbb{R}^M$, on définit $|x|^2_M = \frac{1}{M}\sum_{m=1}^M |x_m|^2$. Pour une famille de coefficients de projection $\alpha = (\alpha_0, \cdots, \alpha_q)$, on pose $|\alpha| = \max_{0\leq l\leq q} |\alpha_l|$ (les dimensions des α_l peuvent être différentes). Pour une famille de fonctions de base à un instant t_k, $|p_k|$ est défini de la même manière.
- Pour une matrice symétrique réelle A, $\|A\|$ et $\|A\|_F$ sont respectivement le maximum de la valeur absolue de ses valeurs propres et sa norme de Frobenius (définie par $\|A\|^2_F = \sum_{i,j} a^2_{i,j}$).

Simulations

Dans la suite, M simulations indépendates de $(P^N_{t_k})_{0\leq k\leq N}$, $(\Delta W_k)_{0\leq k\leq N-1}$ seront utilisées. On les note $((P^{N,m}_{t_k})_{0\leq k\leq N})_{1\leq m\leq M}$, $((\Delta W^m_k)_{0\leq k\leq N-1})_{1\leq m\leq M}$.

- Les valeurs des fonctions de base évaluées le long des trajectoires simulées sont notées $(p^m_{l,k} = p_l(P^{N,m}_{t_k}))_{0\leq l\leq q, 0\leq k\leq N-1, 1\leq m\leq M}$.

– De manière analogue à $f_k(\alpha_{0,k},\cdots,\alpha_{q,k})$ ou $f_k(\alpha_k)$, on note $f_k^m(\alpha_{0,k},\cdots,\alpha_{q,k})$ ou $f_k^m(\alpha_k)$ pour $f(t_k, X_{t_k}^{N,m}, \alpha_{0,k}\cdot p_{0,k}^m,\cdots,\alpha_{q,k}\cdot p_{q,k}^m)$.

On définit

– le vecteur colonne v_k^m par $[v_k^m]^* = (p_{0,k}^{m\,*}, p_{1,k}^{m\,*}\frac{\Delta W_{1,k}^m}{\sqrt{h}},\cdots,p_{q,k}^{m\,*}\frac{\Delta W_{q,k}^m}{\sqrt{h}})$;

– la matrice $V_k^M = \frac{1}{M}\sum_{m=1}^M v_k^m[v_k^m]^*$;

– les matrices $P_{l,k}^M = \frac{1}{M}\sum_{m=1}^M p_{l,k}^m[p_{l,k}^m]^*$ $(0\le l\le q)$.

Troncatures

Pour assurer la stabilité de l'algorithme, on utilise des techniques de seuillage, basées sur les notations suivantes :

– Dans la proposition I.3 à venir, en se basant sur des estimées à priori, on construit explicitement des fonctions à valeurs réelles $(\rho_{l,k}^N)_{0\le l\le q,0\le k\le N-1}$ minorées par 1. On pose alors $\rho_k^N(P_{t_k}^N) = [\rho_{0,k}^N(P_{t_k}^N),\cdots,\rho_{q,k}^N(P_{t_k}^N)]^*$.

– Associées à ces estimées, on définit des fonctions (aléatoires) de troncature $\hat\rho_{l,k}^N(x) = \rho_{l,k}^N(P_{t_k}^N)\xi(x/\rho_{l,k}^N(P_{t_k}^N))$ et $\hat\rho_{l,k}^{N,m}(x) = \rho_{l,k}^N(P_{t_k}^{N,m})\xi(x/\rho_{l,k}^N(P_{t_k}^{N,m}))$ où $\xi:\mathbb{R}\mapsto\mathbb{R}$ est une fonction de classe C_b^2, telle que $\xi(x)=x$ pour $|x|\le 3/2$, $|\xi|_\infty\le 2$ et $|\xi'|_\infty\le 1$.

1.2.2 L'algorithme

On peut à présent exposer notre algorithme pour approcher (I.1.1-I.1.2). Notre procédure combine un schéma rétrograde en temps (de l'instant $t_N = T$ à l'instant $t_0 = 0$), un argument de point fixe (en utilisant $i = 1,\cdots,I$ itérations de Picard) et la résolution de problèmes de moindres carrés basés sur les M trajectoires simulées (en utilisant bien sûr les bases de fonctions choisies précédemment). L'approximation construite par l'algorithme est notée $(Y^{N,I,I,M}, Z^{N,I,I,M})$ pour montrer que cette approximation dépend de la discrétisation en temps, du nombre d'itérations de Picard (la présence du double I est expliquée dans le chapitre 3) et du nombre de trajectoires Monte-Carlo M. On a omis la dépendance dans les fonctions de base pour ne pas alourdir davantage des notations déjà peu conviviales ! Voyons comment se construit l'approximation $(Y^{N,I,I,M}, Z^{N,I,I,M})$.

→ **Initialisation.** L'algorithme est initialisé avec $Y_{t_N}^{N,i,I,M} = \Phi^N(P_{t_N}^N)$ (indépendamment de i et I). Puis, la solution $(Y_{t_k}, Z_{1,t_k},\cdots,Z_{q,t_k})$ à un instant t_k est représentée via des coefficients de projection $(\alpha_{l,k}^{i,I,M})_{0\le l\le q}$ par

$$Y_{t_k}^{N,i,I,M} = \hat\rho_{0,k}^N(\alpha_{0,k}^{i,I,M}\cdot p_{0,k}),\quad \sqrt{h}\,Z_{l,t_k}^{N,i,I,M} = \hat\rho_{l,k}^N(\sqrt{h}\,\alpha_{l,k}^{i,I,M}\cdot p_{l,k})$$

($\hat\rho_{0,k}^N$ et $\hat\rho_{l,k}^N$ sont les troncatures introduites auparavant). Expliquons à présent comme ces coefficients sont calculés en utilisant les réalisations indépendantes $((P_{t_k}^{N,m})_{0\le k\le N})_{1\le m\le M}$, $((\Delta W_k^m)_{0\le k\le N-1})_{1\le m\le M}$.

→ **Itération (rétrograde en temps) à l'instant $t_k < T$.** On suppose que l'on a construit une approximation $Y_{t_{k+1}}^{N,I,I,M} := \hat{\rho}_{0,k+1}^{N}(\alpha_{0,k+1}^{I,I,M} \cdot p_{0,k+1})$, et on note $Y_{t_{k+1}}^{N,I,I,M,m} = \hat{\rho}_{0,k+1}^{N,m}(\alpha_{0,k+1}^{I,I,M} \cdot p_{0,k+1}^{m})$ son évaluation pour la m-ème simulation.

→ Pour l'initialisation $i = 0$ des itérations de Picard, on pose $Y_{t_k}^{N,0,I,M} = 0$ et $Z_{t_k}^{N,0,I,M} = 0$, i.e. $\alpha_{l,k}^{0,I,M} = 0$ $(0 \leq l \leq q)$.

→ Pour $i = 1, \cdots, I$, les coefficients $\alpha_k^{i,I,M} = (\alpha_{l,k}^{i,I,M})_{0 \leq l \leq q}$ sont itérativement obtenus comme l'argmin en $(\alpha_0, \cdots, \alpha_q)$ de la quantité

$$\frac{1}{M} \sum_{m=1}^{M} \left(Y_{t_{k+1}}^{N,I,I,M,m} - \alpha_0 \cdot p_{0,k}^{m} + h f_k^m(\alpha_k^{i-1,I,M}) - \sum_{l=1}^{q} \alpha_l \cdot p_{l,k}^{m} \, \Delta W_{l,k}^{m} \right)^2. \qquad \text{(I.1.3)}$$

Si le problème de moindres carrés ci-dessus a plusieurs solutions (i.e. la matrice de régression empirique n'est pas inversible, ce qui arrive avec faible probabilité lorsque M devient grand), on peut par exemple prendre la solution de norme minimale. Sur ce point, on peut se référer à [GVL96] ou à quelques rappels effectués dans l'annexe A.5. En fait, ce choix est arbitraire et n'aura pas d'incidence sur l'analyse à venir.

Les paramètres de convergence de cet algorithme sont le pas de temps h ($h \to 0$), les bases de fonctions, le nombre de simulations M ($M \to +\infty$). L'impact de chaque paramètre est analysé dans les trois prochains chapitres, correspondant à trois grandes étapes : discrétisation en temps de l'EDSR, projections sur des bases de fonctions dans $\mathbf{L}_2(\Omega, \mathbb{P})$, projections empiriques en utilisant les trajectoires simulées. Une estimation de l'erreur globale commise dans l'algorithme décrit ci-dessus découlera directement de la combinaison des théorèmes I.1-I.2-I.3. Nous verrons également qu'il est suffisant d'effectuer $I = 3$ itérations de Picard (voir théorème I.3).

L'intuition qui se cache derrière cette suite de problèmes de moindres carrés (I.1.3) est simple. Son but est d'imiter ce qui pourrait être fait idéalement avec un nombre infini de simulations, d'itérations de Picard et de fonctions de base lorsque

$$(Y_{t_k}^N, Z_{t_k}^N) = \arg \inf_{(Y,Z) \in \mathbf{L}_2(\mathcal{F}_{t_k})} \mathbb{E}(Y_{t_{k+1}}^N - Y + h f(t_k, X_{t_k}^N, Y, Z) - Z \Delta W_k)^2$$

où $\mathbf{L}_2(\mathcal{F}_{t_k})$ représente les variables aléatoires (éventuellement multidimensionnelles) de carré intégrable et $\mathcal{F}_{t_k}$-mesurables. Ce cas idéal est une approximation de l'EDSR (I.1.1) qui s'écrit

$$Y_{t_{k+1}} + \int_{t_k}^{t_{k+1}} f(s, X_s, Y_s, Z_s) ds = Y_{t_k} + \int_{t_k}^{t_{k+1}} Z_s dW_s$$

sur l'intervalle de temps $[t_k, t_{k+1}]$. $(Y_{t_k}^N)_k$ sera interprété comme une EDSR à temps discret (voir chapitre suivant).

Chapitre 2

Etape 1 : discrétisation en temps

Comme indiqué plus haut, nous considérons tout d'abord une discrétisation en temps des équations (I.1.1) et (I.1.2).

2.1 Discrétisation en temps de X

Pour discrétiser l'équation (I.1.2), on utilise un schéma d'Euler. Ce schéma est noté $(X^N_{t_k})_{0\leq k\leq N}$ et l'évolution de la composante i est définie par :

$$\begin{cases} X^N_{i,0} = X_{i,0} \\ X^N_{i,t_{k+1}} = X^N_{i,t_k} + b_i(t_k, X^N_{t_k})h + \sigma_i(t_k, X^N_{t_k})\Delta W_k. \end{cases} \tag{I.2.1}$$

L'hypothèse 0.1 assure la convergence de ce schéma vers X.

Discrétisation de la condition terminale

Nous avons déjà expliqué que pour approcher $\Phi(\mathbf{X})$, on utilisait $\Phi^N(P^N_{t_N})$, où Φ^N est une fonction déterministe et $(P^N_{t_k})_{0\leq k\leq N}$ est une chaîne de Markov de dimension $d' \geq d$, dont les premières composantes sont données par celles de $(X^N_{t_k})_{0\leq k\leq N}$.
En d'autres termes, nous ajoutons éventuellement des variables d'état supplémentaires au vecteur $X^N_{t_k}$ pour rendre le problème markovien malgré la présence d'une condition terminale dépendant de toute la trajectoire de X. Assurer que le problème est markovien est essentiel pour l'étape 2.
Pour clarifier les choses, voyons quelques exemples concrets avec $d = 1$ et $q = 1$:

- **Option vanille.** Le payoff est une fonction déterministe de la valeur finale du sous-jacent : $\Phi(\mathbf{X}) = \phi(X_T)$. Il n'y a pas de problème dans ce cas, on choisit simplement $P^N_{t_k} = X^N_{t_k}$ et $\Phi^N(P^N_{t_N}) = \phi(P^N_{t_N})$. d' est égal à 1. L'hypothèse I.2 assure alors que ϕ est Lipschitz et l'on peut quantifier l'erreur réalisée en approchant la condition terminale : $\mathbb{E}|\Phi^N(P^N_{t_N}) - \Phi(\mathbf{X})|^2 \leq Ch$.

- **Option asiatique.** Le payoff dépend cette fois de toute la trajectoire de X : $\Phi(\mathbf{X}) = \phi(X_T, \int_0^T X_t dt)$. On peut alors choisir $P^N_{t_k} = (X^N_{t_k}, h\sum_{i=0}^{k-1} X^N_{t_i})$ et $\Phi^N(P^N_{t_N}) = \phi(P^N_{t_N})$. d' est égal à 2. Pour des fonctions ϕ usuelles dans ce contexte, l'erreur d'approximation de la condition terminale sera encore $\mathbb{E}|\Phi^N(P^N_{t_N}) - \Phi(\mathbf{X})|^2 \leq Ch$. On peut également considérer des approximations plus précises de la moyenne de $\mathbf{X}$ (par exemple celles décrites dans [LT01]).
- **Option lookback.** $\Phi(\mathbf{X}) = \phi(X_T, \min_{t\in[0,T]} X_t, \max_{t\in[0,T]} X_t)$. On peut cette fois choisir $\Phi^N(P^N_{t_N}) = \phi(P^N_{t_N})$ avec $P^N_{t_k} = (X^N_{t_k}, \min_{i\leq k} X^N_{t_i}, \max_{i\leq k} X^N_{t_i})$. d' est égal à 3. En général, l'erreur d'approximation de la condition terminale $\mathbb{E}|\Phi^N(P^N_{t_N}) - \Phi(\mathbf{X})|^2$ est légèrement moins bonne que précédemment et se majore désormais par $\sqrt{h\log(1/h)}$ (voir par exemple [Zha04]). On peut néanmoins atteindre la vitesse de convergence $\sqrt{h}$ en considérant les extrema exacts du schéma d'Euler continu (voir par exemple [ABR96]).

Remarquons que pour ces exemples, l'hypothèse I.3 est satisfaite.

2.2 Discrétisation de l'EDSR

On va dans cette section approcher (I.1.1) en définissant une équation rétrograde (Y^N, Z^N) discrète définie aux instants de discrétisation $t_k = kh$, $0 \leq k \leq N$. On initialise naturellement cette équation rétrograde discrète en posant $Y^N_{t_N} = \Phi^N(P^N_{t_N})$. Puis, $(Y^N_{t_k}, Z^N_{t_k})_{0\leq k\leq N-1}$ sont définis par

$$Z^N_{l,t_k} = \frac{1}{h}\mathbb{E}_k(Y^N_{t_{k+1}}\Delta W_{l,k}), \tag{I.2.2}$$

$$Y^N_{t_k} = \mathbb{E}_k(Y^N_{t_{k+1}}) + hf(t_k, X^N_{t_k}, Y^N_{t_k}, Z^N_{t_k}). \tag{I.2.3}$$

Remarquons que $Y^N_{t_k}$ est bien défini dans (I.2.3), car l'application $Y \mapsto \mathbb{E}_k(Y^N_{t_{k+1}}) + hf(t_k, X^N_{t_k}, Y, Z^N_{t_k})$ est une contraction dans $\mathbf{L}^2(\mathcal{F}_{t_k})$ pour h suffisamment petit.
On voit donc qu'à chaque instant de discrétisation t_k, $Y^N_{t_k}$ et $Z^N_{t_k}$ sont définis par des espérances conditionnelles. On peut alors imaginer dans la suite utiliser un procédé d'approximation d'espérance conditionnelle pour approcher $Y^N_{t_k}$ et $Z^N_{t_k}$. Dans cette partie, nous n'utiliserons pas ce point de vue et nous allons plutôt utiliser une autre caractérisation du couple $(Y^N_{t_k}, Z^N_{t_k})$.
En effet, en appliquant l'inégalité de Cauchy-Schwarz conditionnelle à (I.2.2), on déduit que $|Z^N_{l,t_k}| \leq \frac{1}{\sqrt{h}}\sqrt{\mathbb{E}_k(Y^N_{t_{k+1}})^2}$. Il est alors facile par un raisonnement par récurrence de montrer que $Y^N_{t_k}$ et $Z^N_{t_k}$ appartiennent à $\mathbf{L}^2(\mathcal{F}_{t_k})$.
Ayant plongé $(Y^N_{t_k}, Z^N_{t_k})$ dans un cadre $\mathbf{L}^2$, on peut alors remarquer que ce couple peut être défini comme l'unique minimiseur (pour h suffisamment petit) dans $\mathbf{L}^2(\mathcal{F}_{t_k})$ de la quantité

$$\mathbb{E}(Y^N_{t_{k+1}} - Y + hf(t_k, X^N_{t_k}, Y, Z) - Z\Delta W_k)^2. \tag{I.2.4}$$

2.3 Erreur de discrétisation

On peut alors quantifier l'erreur commise entre (Y^N, Z^N) et (Y, Z).

Théorème I.1 *On suppose que les hypothèses 0.1,I.1 et I.2 sont satisfaites. Pour h suffisamment petit, on a :*

$$\max_{0\le k\le N}\mathbb{E}|Y_{t_k}-Y^N_{t_k}|^2+\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\mathbb{E}|Z_t-Z^N_{t_k}|^2dt \le C\big((1+|X_0|^2)h+\mathbb{E}|\Phi(\mathbf{X})-\Phi^N(P^N_{t_N})|^2\big).$$

PREUVE: Notre but est d'utiliser le Lemme de Gronwall discret A.3 et pour cela, de relier l'erreur sur Y à l'instant t_k, $\mathbb{E}|Y_{t_k}-Y^N_{t_k}|^2$, à son équivalent à l'instant t_{k+1} $\mathbb{E}|Y_{t_{k+1}}-Y^N_{t_{k+1}}|^2$. Vu (I.1.1) et (I.2.3), la différence $Y_{t_k}-Y^N_{t_k}$ s'écrit :

$$Y_{t_k}-Y^N_{t_k}=\mathbb{E}_{t_k}(Y_{t_{k+1}}-Y^N_{t_{k+1}})+\mathbb{E}_{t_k}\int_{t_k}^{t_{k+1}}[f(s,X_s,Y_s,Z_s)-f(t_k,X^N_{t_k},Y^N_{t_k},Z^N_{t_k})]ds.$$

En élevant au carré, en utilisant l'inégalité de Young (A.5) et en prenant l'espérance, il vient ($\gamma>0$ sera choisi plus tard) :

$$\begin{aligned}
&\mathbb{E}|Y_{t_k}-Y^N_{t_k}|^2\\
\le&(1+\gamma h)\mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}}-Y^N_{t_{k+1}})|^2)\\
&+(h+\frac{1}{\gamma})\mathbb{E}\int_{t_k}^{t_{k+1}}|f(s,X_s,Y_s,Z_s)-f(t_k,X^N_{t_k},Y^N_{t_k},Z^N_{t_k})|^2ds\\
\le&(1+\gamma h)\mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}}-Y^N_{t_{k+1}})|^2)+C(h+\frac{1}{\gamma})\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-Z^N_{t_k}|^2ds\\
&+C(h+\frac{1}{\gamma})h(h+\sup_{t_k\le s\le t_{k+1}}\mathbb{E}|X_s-X^N_{t_k}|^2+\sup_{t_k\le s\le t_{k+1}}\mathbb{E}|Y_s-Y^N_{t_k}|^2) \qquad \text{(I.2.5)}
\end{aligned}$$

en utilisant pour la dernière inégalité que f est Lipschitz. Considérons l'espace de Hilbert $\mathbb{H}^2_{t_k,t_{k+1}}(\mathbb{R}^q)$ des processus $\phi:[t_k,t_{k+1}]\to\mathbb{R}^q$ tels que $\frac{1}{h}\mathbb{E}\int_{t_k}^{t_{k+1}}|\phi_s|^2ds<\infty$. Il est facile de voir que $\overline{Z}_{t_k}=\frac{1}{h}\mathbb{E}_{t_k}\int_{t_k}^{t_{k+1}}Z_sds$ est le projeté de $(Z_s)_{t_k\le s\le t_{k+1}}$ sur $\mathbf{L}^2(\mathcal{F}_{t_k})\subset\mathbb{H}^2_{t_k,t_{k+1}}(\mathbb{R}^q)$. En utilisant la propriété de Pythagore associée à cette projection, il vient :

$$\begin{aligned}
\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-Z^N_{t_k}|^2ds&=\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-\overline{Z}_{t_k}|^2ds+\mathbb{E}\int_{t_k}^{t_{k+1}}|\overline{Z}_{t_k}-Z^N_{t_k}|^2ds\\
&=\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-\overline{Z}_{t_k}|^2ds+h\mathbb{E}|\overline{Z}_{t_k}-Z^N_{t_k}|^2. \qquad \text{(I.2.6)}
\end{aligned}$$

Remarquons que

$$\overline{Z}_{t_k} = \frac{1}{h}\mathbb{E}_{t_k}(Y_{t_{k+1}}\Delta W_k) + \frac{1}{h}\mathbb{E}_{t_k}(\Delta W_k \int_{t_k}^{t_{k+1}} f(s, X_s, Y_s, Z_s)ds).$$

En comparant avec (I.2.2), il vient :

$$\begin{aligned}\mathbb{E}(|Z_{t_k}^N - \overline{Z}_{t_k}|^2) \leq & \frac{2}{h^2}\sum_{l=1}^{q}\mathbb{E}|\mathbb{E}_{t_k}(\{Y_{t_{k+1}} - Y_{t_{k+1}}^N\}\Delta W_{l,k})|^2 \\ & + \frac{2}{h^2}\mathbb{E}\bigg(|\mathbb{E}_{t_k}(\Delta W_k \int_{t_k}^{t_{k+1}} f(s, X_s, Y_s, Z_s)ds)|^2\bigg). \qquad \text{(I.2.7)}\end{aligned}$$

Remarquons alors que

$$\begin{aligned}|\mathbb{E}_{t_k}(\{Y_{t_{k+1}} - Y_{t_{k+1}}^N\}\Delta W_{l,k})|^2 = & |\mathbb{E}_{t_k}(\{Y_{t_{k+1}} - Y_{t_{k+1}}^N - \mathbb{E}_{t_k}(Y_{t_{k+1}} - Y_{t_{k+1}}^N)\}\Delta W_{l,k})|^2 \\ \leq & h\{\mathbb{E}_{t_k}(|Y_{t_{k+1}} - Y_{t_{k+1}}^N|^2) - |\mathbb{E}_{t_k}(Y_{t_{k+1}} - Y_{t_{k+1}}^N)|^2\}\end{aligned}$$

en appliquant l'inégalité de Cauchy-Schwarz conditionnelle. (I.2.7) devient alors :

$$\begin{aligned}& \mathbb{E}(|Z_{t_k}^N - \overline{Z}_{t_k}|^2) \\ \leq & \frac{C}{h}\{\mathbb{E}(|Y_{t_{k+1}} - Y_{t_{k+1}}^N|^2) - \mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}} - Y_{t_{k+1}}^N)|^2)\} + C\mathbb{E}\int_{t_k}^{t_{k+1}} f(s, X_s, Y_s, Z_s)^2 ds. \qquad \text{(I.2.8)}\end{aligned}$$

Comme nous allons le voir dans un instant, le terme $-\mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}} - Y_{t_{k+1}}^N)|^2)$ dans (I.2.8) joue un rôle important dans le contrôle de l'erreur. En injectant (I.2.6) et (I.2.8) dans (I.2.5), il vient :

$$\begin{aligned}& \mathbb{E}|Y_{t_k} - Y_{t_k}^N|^2 \\ \leq & (1+\gamma h)\mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}} - Y_{t_{k+1}}^N)|^2) + C(h + \frac{1}{\gamma})h(h + \sup_{[t_k, t_{k+1}]}\mathbb{E}|X_s - X_{t_k}^N|^2 + \sup_{[t_k, t_{k+1}]}\mathbb{E}|Y_s - Y_{t_k}^N|^2) \\ & + C(h + \frac{1}{\gamma})\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s - \overline{Z}_{t_k}|^2 ds + C(h + \frac{1}{\gamma})\{\mathbb{E}(|Y_{t_{k+1}} - Y_{t_{k+1}}^N|^2) - \mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}} - Y_{t_{k+1}}^N)|^2)\} \\ & + Ch(h + \frac{1}{\gamma})\mathbb{E}\int_{t_k}^{t_{k+1}} f(s, X_s, Y_s, Z_s)^2 ds.\end{aligned}$$

En choisissant à présent $\gamma = C$, l'expression ci-dessus se simplifie (avec une constante C différente) :

$$\begin{aligned}& \mathbb{E}|Y_{t_k} - Y_{t_k}^N|^2 \\ \leq & (1+Ch)\mathbb{E}(|Y_{t_{k+1}} - Y_{t_{k+1}}^N|^2) + Ch(h + \sup_{t_k \leq s \leq t_{k+1}}\mathbb{E}|X_s - X_{t_k}^N|^2 + \sup_{t_k \leq s \leq t_{k+1}}\mathbb{E}|Y_s - Y_{t_k}^N|^2) \\ & + C\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s - \overline{Z}_{t_k}|^2 ds + Ch\mathbb{E}\int_{t_k}^{t_{k+1}} f(s, X_s, Y_s, Z_s)^2 ds.\end{aligned}$$

En tenant compte de $\mathbb{E}|Y_s - Y^N_{t_k}|^2 \leq 2\mathbb{E}|Y_s - Y_{t_k}|^2 + 2\mathbb{E}|Y_{t_k} - Y^N_{t_k}|^2$, et en supposant h suffisamment petit, l'expression s'arrange encore en :

$$\begin{aligned}
&\mathbb{E}|Y_{t_k} - Y^N_{t_k}|^2 \\
\leq&(1+Ch)\mathbb{E}(|Y_{t_{k+1}} - Y^N_{t_{k+1}}|^2) + Ch(h + \sup_{t_k \leq s \leq t_{k+1}} \mathbb{E}|X_s - X^N_{t_k}|^2 + \sup_{t_k \leq s \leq t_{k+1}} \mathbb{E}|Y_s - Y_{t_k}|^2) \\
&+ C\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s - \overline{Z}_{t_k}|^2 ds + Ch\mathbb{E}\int_{t_k}^{t_{k+1}} f(s, X_s, Y_s, Z_s)^2 ds.
\end{aligned}$$

En appliquant le lemme de Gronwall discret A.3, on obtient donc :

$$\begin{aligned}
&\max_{0\leq k\leq N} \mathbb{E}|Y_{t_k} - Y^N_{t_k}|^2 \\
&\leq C\mathbb{E}|\Phi(\mathbf{X}) - \Phi^N(P^N_{t_N})|^2 + Ch + C\max_k\{\sup_{t_k \leq s \leq t_{k+1}} \mathbb{E}|X_s - X^N_{t_k}|^2 + \sup_{t_k \leq s \leq t_{k+1}} \mathbb{E}|Y_s - Y_{t_k}|^2\} \\
&+ C\mathbb{E}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} |Z_s - \overline{Z}_{t_k}|^2 ds + Ch\mathbb{E}\int_0^T f(s, X_s, Y_s, Z_s)^2 ds.
\end{aligned}$$

Il est alors facile de vérifier que

$$\max_k\{\sup_{t_k \leq s \leq t_{k+1}} \mathbb{E}|X_s - X^N_{t_k}|^2 + \sup_{t_k \leq s \leq t_{k+1}} \mathbb{E}|Y_s - Y_{t_k}|^2\} + h\mathbb{E}\int_0^T f(s, X_s, Y_s, Z_s)^2 ds \leq Ch.$$

Ainsi, on obtient finalement :

$$\max_{0\leq k\leq N} \mathbb{E}|Y_{t_k} - Y^N_{t_k}|^2 \leq C\mathbb{E}|\Phi(\mathbf{X}) - \Phi^N(P^N_{t_N})|^2 + Ch + C\mathbb{E}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} |Z_s - \overline{Z}_{t_k}|^2 ds.$$

Pour borner à présent l'erreur sur Z, on combine (I.2.6) et (I.2.8) pour obtenir

$$\begin{aligned}
&\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s - Z^N_{t_k}|^2 ds \\
=&\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s - \overline{Z}_{t_k}|^2 ds + h\mathbb{E}|\overline{Z}_{t_k} - Z^N_{t_k}|^2 \\
\leq&\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s - \overline{Z}_{t_k}|^2 ds + C\{\mathbb{E}|Y_{t_{k+1}} - Y^N_{t_{k+1}}|^2 - \mathbb{E}|\mathbb{E}_{t_k}(Y_{t_{k+1}} - Y^N_{t_{k+1}})|^2\} \\
&+ Ch\mathbb{E}\int_{t_k}^{t_{k+1}} f(s, X_s, Y_s, Z_s)^2 ds.
\end{aligned}$$

En sommant sur k et en majorant $\mathbb{E}\int_0^T f^2(s, X_s, Y_s, Z_s) ds$, on obtient :

$$\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-Z_{t_k}^N|^2ds\leq\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-\overline{Z}_{t_k}|^2ds+Ch$$
$$+C\sum_{k=0}^{N-1}\{\mathbb{E}|Y_{t_{k+1}}-Y_{t_{k+1}}^N|^2-\mathbb{E}|\mathbb{E}_{t_k}(Y_{t_{k+1}}-Y_{t_{k+1}}^N)|^2\}$$
$$\leq\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-\overline{Z}_{t_k}|^2ds+Ch+C\mathbb{E}|\Phi(\mathbf{X})-\Phi^N(P_{t_N}^N)|^2$$
$$+C\sum_{k=0}^{N-1}\{\mathbb{E}|Y_{t_k}-Y_{t_k}^N|^2-\mathbb{E}|\mathbb{E}_{t_k}(Y_{t_{k+1}}-Y_{t_{k+1}}^N)|^2\}\qquad(\text{I.2.9})$$

en effectuant un changement d'indice dans la dernière somme du membre de droite. En utilisant à présent (I.2.5), on obtient, pour $0\leq k\leq N-1$ et $\gamma>0$:

$$C\{\mathbb{E}|Y_{t_k}-Y_{t_k}^N|^2-\mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}}-Y_{t_{k+1}}^N)|^2)\}$$
$$\leq C\gamma h\mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}}-Y_{t_{k+1}}^N)|^2)+C_1(h+\frac{1}{\gamma})h(h+\sup_{[t_k,t_{k+1}]}\mathbb{E}|X_s-X_{t_k}^N|^2+\sup_{[t_k,t_{k+1}]}\mathbb{E}|Y_s-Y_{t_k}^N|^2)$$
$$+C_1(h+\frac{1}{\gamma})\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-Z_{t_k}^N|^2ds.$$

On prend maintenant $\gamma=3C_1$ ci-dessus, et pour h assez petit, il vient (pour une autre constante C_1) :

$$C\{\mathbb{E}|Y_{t_k}-Y_{t_k}^N|^2-\mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}}-Y_{t_{k+1}}^N)|^2)\}$$
$$\leq C_1h\max_k\mathbb{E}(|Y_{t_k}-Y_{t_k}^N|^2)+C_1h^2+\frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-Z_{t_k}^N|^2ds.$$

En injectant dans (I.2.9), une simplification s'opère pour donner :

$$\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-Z_{t_k}^N|^2ds\leq C\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-\overline{Z}_{t_k}|^2ds+Ch+C\max_k\mathbb{E}(|Y_{t_k}-Y_{t_k}^N|^2).$$

Ainsi, l'erreur sur Y et sur Z sont du même ordre et sont contrôlées d'une part par h et l'erreur d'approximation de la condition terminale, et d'autre part par $\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s-\overline{Z}_{t_k}|^2ds$. Ainsi, le point clé est la régularité $\mathbf{L}^2$ de Z, telle que définie dans [Zha04]. Dans notre cas, sous les hypothèses 0.1, I.1 et I.2, Z est càdlàg (voir remarque A.1). Ainsi, le Théorème A.2 permet d'affirmer que

$$\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_t-Z_{t_k}|^2dt\leq C(1+|X_0|^2)h.$$

$\sum_{k=0}^{N-1} \mathbb{E} \int_{t_k}^{t_{k+1}} |Z_s - \overline{Z}_{t_k}|^2 ds$ étant majoré par le membre de gauche de l'inégalité ci-dessus, ceci conclut la démonstration. □

Remarque I.1 *On a vu dans la démonstration ci-dessus, que l'erreur était contrôlée par* $\mathbb{E}|\Phi(\mathbf{X}) - \Phi^N(P^N_{t_N})|^2$, *un terme d'ordre* h *et le terme* $\sum_{k=0}^{N-1} \mathbb{E} \int_{t_k}^{t_{k+1}} |Z_s - \overline{Z}_{t_k}|^2 ds$. *Ce dernier terme tend toujours vers* 0 *lorsque* N *tend vers l'infini par un théorème de convergence de martingales dans* L^2. *Ainsi, dès que* $\mathbb{E}|\Phi(\mathbf{X}) - \Phi^N(P^N_{t_N})|^2$ *tend vers* 0, *on a convergence de l'approximation* (Y^N, Z^N) *vers* (Y, Z). *L'utilisation de résultats de [Zha04] permet de donner une vitesse de convergence mais même dans les cas où cette vitesse n'est pas calculable, on aura toujours convergence de l'approximation en temps.*

2.4 Propriétés de l'équation rétrograde discrète

Tout d'abord il est facile de voir en utilisant le caractère markovien de P^N que $Y^N_{t_k}$ et $Z^N_{t_k}$ sont des fonctions déterministes de $P^N_{t_k}$:

Proposition I.1 *Les hypothèses 0.1, I.1 et I.2 sont satisfaites. Pour* h *suffisamment petit, on a :*

$$Y^N_{t_k} = y^N_k(P^N_{t_k}),\ Z^N_{l,t_k} = z^N_{l,k}(P^N_{t_k}) \quad \textit{for}\ \ 0 \le k \le N\ \textit{et}\ \ 1 \le l \le q, \tag{I.2.10}$$

où $(y^N_k(\cdot))_k$ *et* $(z^N_{l,k}(\cdot))_{k,l}$ *sont des fonctions mesurables.*

On voit ainsi l'importance d'avoir approché la condition terminale par une fonction déterministe de la valeur finale d'une chaîne de Markov.
En outre, on peut montrer que les fonctions $y^N_k(\cdot)$ et $z^N_{l,k}(\cdot)$ sont Lipschitz. Ceci découle du caractère Lipschitz de la condition terminale et de la chaîne de Markov $P^N_{t_k}$ par rapport à sa condition initiale.

Proposition I.2 *On suppose que les hypothèses 0.1, I.1, I.2 et I.3 sont satisfaites. Pour* h *assez petit, on a*

$$|y^N_{k_0}(x) - y^N_{k_0}(x')| + \sqrt{h}|z^N_{k_0}(x) - z^N_{k_0}(x')| \le C|x - x'| \tag{I.2.11}$$

uniformément en $k_0 \le N - 1$.

PREUVE: En utilisant l'inégalité de Young (A.5) de la même façon que dans la démonstration du Théorème I.1, on peut obtenir

$$\begin{aligned}
&\mathbb{E}|Y_{t_k}^{N,k_0,x} - Y_{t_k}^{N,k_0,x'}|^2 \\
\leq &\frac{(1+\gamma h)}{1-Ch(h+\frac{1}{\gamma})}\mathbb{E}|\mathbb{E}_k(Y_{t_{k+1}}^{N,k_0,x} - Y_{t_{k+1}}^{N,k_0,x'})|^2 + \frac{Ch(h+\frac{1}{\gamma})}{1-Ch(h+\frac{1}{\gamma})}\mathbb{E}|X_{t_k}^{N,k_0,x} - X_{t_k}^{N,k_0,x'}|^2 \\
&+ \frac{C(h+\frac{1}{\gamma})}{1-Ch(h+\frac{1}{\gamma})}\big(\mathbb{E}|Y_{t_{k+1}}^{N,k_0,x} - Y_{t_{k+1}}^{N,k_0,x'}|^2 - \mathbb{E}|\mathbb{E}_k(Y_{t_{k+1}}^{N,k_0,x} - Y_{t_{k+1}}^{N,k_0,x'})|^2\big).
\end{aligned}$$

En choisissant $\gamma = C$ et h suffisamment petit, on obtient (pour une autre constante C) :

$$\mathbb{E}|Y_{t_k}^{N,k_0,x} - Y_{t_k}^{N,k_0,x'}|^2 \leq (1+Ch)\mathbb{E}|Y_{t_{k+1}}^{N,k_0,x} - Y_{t_{k+1}}^{N,k_0,x'}|^2 + Ch\mathbb{E}|X_{t_k}^{N,k_0,x} - X_{t_k}^{N,k_0,x'}|^2.$$

$\mathbb{E}|X_{t_k}^{N,k_0,x} - X_{t_k}^{N,k_0,x'}|^2$ est borné par $C|x-x'|^2$ (en utilisant les propriétés classiques du flot). Ainsi, en utilisant le Lemme de Gronwall discret A.3 et l'hypothèse I.3, on obtient le résultat pour $y_{k_0}^N(\cdot)$. Le résultat pour $\sqrt{h}z_{k_0}^N(\cdot)$ découle de (I.2.2). □

Dans les cas option vanille, lookback ou asiatique décrits plus haut, l'hypothèse I.3 est vérifiée. Ainsi, on doit à chaque pas de temps approcher des fonctions Lipschitz.

Chapitre 3

Etape 2 : approximation de fonctions

A l'étape 1, on a vu que résoudre l'EDSR définie par (I.2.2) et (I.2.3), revient, grâce à la proposition I.1, à déterminer les fonctions $y_k^N(\cdot)$ et $z_{l,k}^N(\cdot)$ pour tout k et tout l. On va approcher ces fonctions à l'aide de projections sur des espaces vectoriels de dimension finie. Plus particulièrement, à chaque instant t_k, $0 \leq k \leq N-1$, on va chercher une approximation de $y_k^N(\cdot)$ dans l'espace vectoriel engendré par une base de fonctions notée $p_{0,k}(\cdot)$ et, pour chaque l, une approximation de $z_{l,k}^N(\cdot)$ dans l'espace vectoriel engendré par une base de fonctions notée $p_{l,k}(\cdot)$.
A chaque instant de discrétisation t_k, on a vu que $Y_{t_k}^N$ était défini par un point fixe (voir (I.2.3)). Dans cette deuxième étape, nous allons nous rapprocher d'une approximation réellement calculable du couple solution (Y, Z), et nous allons donc devoir spécifier comment régler le problème du point fixe à chaque instant de discrétisation. Pour cela, on ne va effectuer qu'un nombre fini d'itérations de Picard noté I à chaque instant de discrétisation t_k.

3.1 Définition de l'approximation

On va définir une approximation de (Y^N, Z^N) de manière rétrograde. On note cette approximation $(Y^{N,I,I}, Z^{N,I,I})$. Le premier I rappelle qu'à un instant de discrétisation t_k, on effectue I itérations de Picard pour définir $Y_{t_k}^{N,I,I}$ et le second I rappelle qu'avant de définir $Y_{t_k}^{N,I,I}$, on a défini des approximations $(Y_{t_j}^{N,I,I})_{j\geq k+1}$ en effectuant à chaque fois I itérations de Picard. La notation $Y_{t_k}^{N,i,\overbrace{I,\cdots,I}^{N-k-1 \text{ fois}}}$ serait plus juste mais aussi beaucoup plus lourde.
L'approximation est initialisée naturellement en supposant qu'on a une seule fonction de base $p_{0,N}(\cdot) = \Phi^N(\cdot)$ à l'instant t_N et en posant $\alpha_{0,N}^{I,I} = 1$.
Puis, à un instant t_k, pour $0 \leq k \leq N-1$:

– On initialise $\alpha_{l,k}^{0,I} = 0$ pour $0 \leq l \leq q$
– Pour $0 \leq i \leq I-1$, on résout :

$$\begin{aligned}&(\alpha_{0,k}^{i,I}, \cdots, \alpha_{q,k}^{i,I})\\ &= \arg \inf_{(\alpha_0,\cdots,\alpha_q)} \mathbb{E}(|\alpha_{0,k+1}^{I,I}.p_{0,k+1} + hf_k(\alpha_{0,k}^{i-1,I}, \cdots, \alpha_{q,k}^{i-1,I}) - \alpha_0.p_{0,k} - \sum_{l=1}^{q} \alpha_l.p_{l,k}\Delta W_{l,k}|^2).\end{aligned} \tag{I.3.1}$$

– Notre approximation $Y_{t_k}^{N,I,I}$ est alors définie par $Y_{t_k}^{N,I,I} = \alpha_{0,k}^{I,I}.p_{0,k}$ et $Z_{l,t_k}^{N,I,I}$ par $Z_{l,t_k}^{N,I,I} = \alpha_{l,k}^{I,I}.p_{l,k}$.

On peut remarquer l'analogie avec (I.2.4). Comme nous l'avons déjà dit, nous utilisons dans cette partie cette caractérisation de (Y^N, Z^N) pour définir notre approximation. Nous verrons dans la partie II comment utiliser l'autre caractérisation de (Y^N, Z^N) comme espérance conditionnelle.

3.2 Erreur commise à l'étape 2

L'erreur entre (Y^N, Z^N) et l'approximation construite à l'étape 2 $(Y^{N,I,I}, Z^{N,I,I})$ est bornée dans le théorème suivant :

Théorème I.2 *On suppose que les hypothèses 0.1, I.1 et I.2 sont satisfaites. Pour h suffisamment petit, on a*

$$\begin{aligned}&\max_{0\leq k\leq N} \mathbb{E}|Y_{t_k}^{N,I,I} - Y_{t_k}^N|^2 + h\sum_{k=0}^{N-1}\sum_{l=1}^{q} \mathbb{E}|Z_{l,t_k}^{N,I,I} - Z_{l,t_k}^N|^2\\ &\leq Ch^{2I-2}\big[1 + |X_0|^2 + \mathbb{E}|\Phi^N(P_{t_N}^N)|^2\big]\\ &\quad + C\sum_{k=0}^{N-1} \mathbb{E}|\mathcal{R}_{p_{0,k}}(Y_{t_k}^N)|^2 + Ch\sum_{k=0}^{N-1}\sum_{l=1}^{q} \mathbb{E}|\mathcal{R}_{p_{l,k}}(Z_{l,t_k}^N)|^2.\end{aligned}$$

Ce théorème montre que les erreurs de projection sur les bases $p_{l,k}$, $0 \leq l \leq q$ se cumulent au fur et à mesure des itérations rétrogrades. Le point important est que ces erreurs ne font que se sommer avec une constante C qui n'explose pas lorsque N tend vers l'infini. Ces estimations améliorent celles du théorème 4.1 dans [BT04] pour deux raisons. Tout d'abord, on borne ici non seulement l'erreur commise sur Y^N mais également sur Z^N. D'autre part, dans le théorème 4.1, plongé dans notre contexte, l'erreur est analysée en termes de $\mathbb{E}|\mathcal{R}_{p_{0,k}}(Y_{t_k}^{N,I,I})|^2$ et $\mathbb{E}|\mathcal{R}_{p_{l,k}}(Z_{l,t_k}^{N,I,I})|^2$: ainsi, l'influence du choix des bases de fonctions n'est pas claire puisqu'elle est cachée non seulement dans les résidus de projection $\mathcal{R}_{p_k}$ mais également dans les variables aléatoires $\alpha_{0,k}^{I,I}.p_{0,k}$ et $\alpha_{l,k}^{I,I}.p_{l,k}$. Notre théorème

permet d'analyser directement l'influence des fonctions de base.
Concernant le nombre d'itérations de Picard I réalisées à chaque pas de temps, on voit qu'il est suffisant que I soit égal à 2 pour que le terme $h^{2I-2}\big[1+|X_0|^2+\mathbb{E}|\Phi^N(P^N_{t_N})|^2\big]$ devienne négligeable par rapport aux erreurs commises par ailleurs.

PREUVE: L'idée générale de la preuve est assez semblable à celle du théorème I.1. Les différences viennent de l'utilisation de la propriété de contraction de la projection sur $(p_{l,k})$ et de la nécessité de se ramener en chaque instant t_k au point fixe de (I.3.1).
On introduit tout d'abord une notation utile par la suite : $\mathcal{A}^N(X_0) = 1+|X_0|^2+\mathbb{E}|\Phi^N(P^N_{t_N})|^2$. Le problème de moindres carrés (I.3.1)

$$(\alpha^{i,I}_{0,k},\cdots,\alpha^{i,I}_{q,k}) =$$
$$\arg\inf_{(\alpha_0,\cdots,\alpha_q)}\mathbb{E}(|\alpha^{I,I}_{0,k+1}.p_{0,k+1}+hf_k(\alpha^{i-1,I}_{0,k},\cdots,\alpha^{i-1,I}_{q,k})-\alpha_0.p_{0,k}-\sum_{l=1}^{q}\alpha_l.p_{l,k}\Delta W_{l,k}|^2)$$

se réécrit

$$\theta^{i,I}_k = \arg\inf_{\theta}\mathbb{E}(Y^{N,I,I}_{t_{k+1}}+hf_k(\alpha^{i-1,I}_k)-\theta\cdot v_k)^2$$

en effectuant le changement de variables $\theta^* = (\alpha^*_0,\sqrt{h}\alpha^*_1,\cdots,\sqrt{h}\alpha^*_q)$. Comme solution d'un problème de moindres carrés, $\theta^{i,I}_k$ vérifie :

$$\mathbb{E}(v_kv_k^*)\theta^{i,I}_k = \mathbb{E}(v_k\{Y^{N,I,I}_{t_{k+1}}+hf_k(\alpha^{i-1,I}_k)\}).$$

ΔW_k étant centré et indépendant de $(p_{l,k})_{0\le l\le q}$, il est facile de voir que la matrice $\mathbb{E}(v_kv_k^*)$ prend la forme bloc-diagonale suivante :

$$\mathbb{E}(v_kv_k^*) = \begin{pmatrix}\mathbb{E}(p_{0,k}p^*_{0,k}) & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \mathbb{E}(p_{q,k}p^*_{q,k})\end{pmatrix}$$

Ainsi, on obtient

$$\begin{cases}\mathbb{E}(p_{0,k}p^*_{0,k})\alpha^{i,I}_{0,k} = \mathbb{E}(p_{0,k}\{Y^{N,I,I}_{t_{k+1}}+hf_k(\alpha^{i-1,I}_k)\}),\\ \mathbb{E}(p_{l,k}p^*_{l,k})\alpha^{i,I}_{l,k} = \frac{1}{h}\mathbb{E}(p_{l,k}Y^{N,I,I}_{t_{k+1}}\Delta W_{l,k}),\ \text{pour } 1\le l\le q,\end{cases}$$

ce qui équivaut, vu la définition de $Y^{N,i,I}_{t_k}$ et $Z^{N,i,I}_{l,t_k}$ à :

$$Z^{N,i,I}_{l,t_k} = \frac{1}{h}\mathcal{P}_{p_{l,k}}\big(Y^{N,I,I}_{t_{k+1}}\Delta W_{l,k}\big), \tag{I.3.2}$$
$$Y^{N,i,I}_{t_k} = \mathcal{P}_{p_{0,k}}\big(Y^{N,I,I}_{t_{k+1}}+hf(t_k,X^N_{t_k},Y^{N,i-1,I}_{t_k},Z^{N,i-1,I}_{t_k})\big). \tag{I.3.3}$$

La preuve du théorème I.2 peut être divisée en plusieurs étapes.

Etape 1 : une borne préliminaire pour $\mathbb{E}|Z^{N,i,I}_{l,t_k}|^2$. Remarquons que vu (I.3.2), $Z^{N,i,I}_{l,t_k}$ est constant pour $i \geq 1$. En outre, une application de l'inégalité de Cauchy-Schwarz donne $|\mathbb{E}_k(Y^{N,I,I}_{t_{k+1}}\Delta W_{l,k})|^2 = |\mathbb{E}_k([Y^{N,I,I}_{t_{k+1}} - \mathbb{E}_k(Y^{N,I,I}_{t_{k+1}})]\Delta W_{l,k})|^2 \leq h\big(\mathbb{E}_k[Y^{N,I,I}_{t_{k+1}}]^2 - [\mathbb{E}_k(Y^{N,I,I}_{t_{k+1}})]^2\big)$. Puisque $(p_{l,k})_l$ est $\mathcal{F}_{t_k}$-mesurable et en utilisant la propriété de contraction de l'opérateur de projection, il vient :

$$\begin{aligned}\mathbb{E}|Z^{N,i,I}_{l,t_k}|^2 &= \frac{1}{h^2}\mathbb{E}\big[\mathcal{P}_{p_{l,k}}\big(\mathbb{E}_k[Y^{N,I,I}_{t_{k+1}}\Delta W_{l,k}]\big)\big]^2 \leq \frac{1}{h^2}\mathbb{E}\big(\mathbb{E}_k[Y^{N,I,I}_{t_{k+1}}\Delta W_{l,k}]\big)^2 \\ &\leq \frac{1}{h}\big(\mathbb{E}[Y^{N,I,I}_{t_{k+1}}]^2 - \mathbb{E}[\mathbb{E}_k(Y^{N,I,I}_{t_{k+1}})]^2\big). \qquad \text{(I.3.4)}\end{aligned}$$

Comme on le verra par la suite, le terme $\mathbb{E}[\mathbb{E}_k(Y^{N,I,I}_{t_{k+1}})]^2$ dans (I.3.4) joue un rôle crucial pour que les estimations à venir ne soient pas explosives lorsque h tend vers 0.

Etape 2 : bornes $\mathbf{L}^2$ pour $Y^{N,i,I}_{t_k}$ et $\sqrt{h}Z^{N,i,I}_{l,t_k}$. Il est facile de montrer que les variables aléatoires $Y^{N,i,I}_{t_k}$ et $\sqrt{h}Z^{N,i,I}_{l,t_k}$ sont de carré intégrable. Nous voulons en fait montrer un contrôle de leurs normes $\mathbf{L}^2$ uniforme en i, I, k, N. A cette fin, nous notons $\chi^{N,I}_k : Y \in \mathbf{L}^2(\mathcal{F}_{t_k}) \mapsto \mathcal{P}_{p_{0,k}}\big(Y^{N,I,I}_{t_{k+1}} + hf(t_k, X^N_{t_k}, Y, Z^{N,i-1,I}_{t_k})\big) \in \mathbf{L}^2(\mathcal{F}_{t_k})$. Clairement, $\mathbb{E}|\chi^{N,I}_k(Y_2) - \chi^{N,I}_k(Y_1)|^2 \leq (C_f h)^2\mathbb{E}|Y_2 - Y_1|^2$ où C_f est la constante de Lipschitz de f. Par conséquent, pour h suffisamment petit, l'application $\chi^{N,I}_k$ est contractante et a un unique point fixe $Y^{N,\infty,I}_{t_k} \in \mathbf{L}^2(\mathcal{F}_{t_k})$ (rappelons que $Z^{N,i,I}_{l,t_k}$ ne dépend pas de $i \geq 1$). On a :

$$Y^{N,\infty,I}_{t_k} = \mathcal{P}_{p_{0,k}}\big(Y^{N,I,I}_{t_{k+1}} + hf(t_k, X^N_{t_k}, Y^{N,\infty,I}_{t_k}, Z^{N,I,I}_{t_k})\big), \qquad \text{(I.3.5)}$$

$$\mathbb{E}|Y^{N,\infty,I}_{t_k} - Y^{N,i,I}_{t_k}|^2 \leq (C_f h)^{2i}\mathbb{E}|Y^{N,\infty,I}_{t_k}|^2 \qquad \text{(I.3.6)}$$

puisque $Y^{N,0,I}_{t_k} = 0$. Ainsi, l'inégalité de Young permet d'écrire pour $i \geq 1$

$$\begin{aligned}\mathbb{E}|Y^{N,i,I}_{t_k}|^2 &\leq (1 + \frac{1}{h})\mathbb{E}|Y^{N,\infty,I}_{t_k} - Y^{N,i,I}_{t_k}|^2 + (1+h)\mathbb{E}|Y^{N,\infty,I}_{t_k}|^2 \\ &\leq (1 + Ch)\mathbb{E}|Y^{N,\infty,I}_{t_k}|^2. \qquad \text{(I.3.7)}\end{aligned}$$

L'inégalité ci-dessus est également vraie pour $i = 0$ car $Y^{N,0,I}_{t_k} = 0$. Nous pouvons à présent estimer $\mathbb{E}|Y^{N,\infty,I}_{t_k}|^2$ à partir de (I.3.5). En combinant l'inégalité de Young (le paramètre γ sera choisi plus tard), l'égalité $\mathcal{P}_{p_{0,k}}(Y^{N,I,I}_{t_{k+1}}) = \mathcal{P}_{p_{0,k}}(\mathbb{E}_k[Y^{N,I,I}_{t_{k+1}}])$, la contraction de $\mathcal{P}_{p_{0,k}}$, la propriété de Lipschitz de f, nous obtenons :

$$\begin{aligned}\mathbb{E}|Y^{N,\infty,I}_{t_k}|^2 &\leq (1+\gamma h)\mathbb{E}|\mathbb{E}_k[Y^{N,I,I}_{t_{k+1}}]|^2 \\ &\quad + Ch(h + \frac{1}{\gamma})\big[\mathbb{E}f_k^2(0,\cdots,0) + \mathbb{E}|Y^{N,\infty,I}_{t_k}|^2 + \mathbb{E}|Z^{N,I,I}_{t_k}|^2\big]. \qquad \text{(I.3.8)}\end{aligned}$$

En réunissant les termes $\mathbb{E}|Y^{N,\infty,I}_{t_k}|^2$, puis en utilisant (I.3.4) et la majoration $\mathbb{E}f_k^2(0,\cdots,0) \leq C(1+|X_0|^2)$, il vient

$$\begin{aligned}\mathbb{E}|Y^{N,\infty,I}_{t_k}|^2 \leq & \frac{(1+\gamma h)}{1-Ch(h+\frac{1}{\gamma})}\mathbb{E}|\mathbb{E}_k[Y^{N,I,I}_{t_{k+1}}]|^2 + \frac{Ch(h+\frac{1}{\gamma})}{1-Ch(h+\frac{1}{\gamma})}\big[1+|X_0|^2\big] \\ & + \frac{C(h+\frac{1}{\gamma})}{1-Ch(h+\frac{1}{\gamma})}\big(\mathbb{E}|Y^{N,I,I}_{t_{k+1}}|^2 - \mathbb{E}|\mathbb{E}_k[Y^{N,I,I}_{t_{k+1}}]|^2\big) \qquad \text{(I.3.9)}\end{aligned}$$

pourvu que h soit suffisamment petit. En choisissant à présent $\gamma = C$, on obtient

$$\begin{aligned}\mathbb{E}|Y^{N,\infty,I}_{t_k}|^2 &\leq Ch\big[1+|X_0|^2\big] + (1+Ch)\mathbb{E}|Y^{N,I,I}_{t_{k+1}}|^2 + Ch\mathbb{E}|\mathbb{E}_k[Y^{N,I,I}_{t_{k+1}}]|^2 \\ &\leq Ch\big[1+|X_0|^2\big] + (1+2Ch)\mathbb{E}|Y^{N,I,I}_{t_{k+1}}|^2 \qquad \text{(I.3.10)}\end{aligned}$$

avec une nouvelle constante C. En injectant cette estimation dans (I.3.7) avec $i = I$, on obtient $\mathbb{E}|Y^{N,I,I}_{t_k}|^2 \leq Ch\big[1+|X_0|^2\big] + (1+Ch)\,\mathbb{E}|Y^{N,I,I}_{t_{k+1}}|^2$. Ainsi, en appliquant le lemme de Gronwall discret, on obtient la majoration $\sup_{0\leq k\leq N}\mathbb{E}|Y^{N,I,I}_{t_k}|^2 \leq C\mathcal{A}^N(X_0)$. Ceci associé à (I.3.10), (I.3.7) et (I.3.4) procure finalement l'estimation uniforme pour $\mathbb{E}|Y^{N,i,I}_{t_k}|^2$ et $\mathbb{E}|Z^{N,i,I}_{l,t_k}|^2$:

$$\sup_{I\geq 1}\sup_{i\geq 0}\sup_{0\leq k\leq N}(\mathbb{E}|Y^{N,i,I}_{t_k}|^2 + h\mathbb{E}|Z^{N,i,I}_{l,t_k}|^2) \leq C\mathcal{A}^N(X_0). \qquad \text{(I.3.11)}$$

Etape 3 : une majoration pour $\eta^{N,I}_k = \mathbb{E}|Y^{N,I,I}_{t_k} - Y^N_{t_k}|^2$.
Remarquons que $\eta^{N,I}_N = 0$. Nous voulons montrer la relation suivante pour $0 \leq k < N$:

$$\begin{aligned}\eta^{N,I}_k \leq & (1+Ch)\eta^{N,I}_{k+1} + Ch^{2I-1}\mathcal{A}^N(X_0) \\ & + C\mathbb{E}|\mathcal{R}_{p_{0,k}}(Y^N_{t_k})|^2 + Ch\sum_{l=1}^{q}\mathbb{E}|\mathcal{R}_{p_{l,k}}(Z^N_{l,t_k})|^2. \qquad \text{(I.3.12)}\end{aligned}$$

On pourra alors directement déduire de cette relation l'estimation sur $\max_{0\leq k\leq N}\mathbb{E}|Y^{N,I,I}_{t_k} - Y^N_{t_k}|^2$ donnée dans le théorème I.2. En reprenant les mêmes arguments qui nous ont conduit à (I.3.7) et en utilisant l'estimation (I.3.11), on obtient

$$\begin{aligned}\eta^{N,I}_k &\leq Ch^{2I-1}\mathcal{A}^N(X_0) + (1+h)\mathbb{E}|Y^{N,\infty,I}_{t_k} - Y^N_{t_k}|^2 \\ &= Ch^{2I-1}\mathcal{A}^N(X_0) + (1+h)\mathbb{E}|\mathcal{R}_{p_{0,k}}(Y^N_{t_k})|^2 + (1+h)\mathbb{E}|Y^{N,\infty,I}_{t_k} - \mathcal{P}_{p_{0,k}}(Y^N_{t_k})|^2 \qquad \text{(I.3.13)}\end{aligned}$$

où nous avons utilisé à la dernière égalité la propriété de Pythagore associée à $\mathcal{P}_{p_{0,k}}$:

$$\mathbb{E}|Y^{N,\infty,I}_{t_k} - Y^N_{t_k}|^2 = \mathbb{E}|\mathcal{R}_{p_{0,k}}(Y^N_{t_k})|^2 + \mathbb{E}|Y^{N,\infty,I}_{t_k} - \mathcal{P}_{p_{0,k}}(Y^N_{t_k})|^2. \qquad \text{(I.3.14)}$$

De plus, avec les mêmes techniques que pour (I.3.4) et (I.3.8), on peut montrer

$$\mathbb{E}|Z_{t_k}^{N,I,I} - Z_{t_k}^N|^2 = \sum_{l=1}^q \mathbb{E}|\mathcal{R}_{p_{l,k}}(Z_{l,t_k}^N)|^2 + \sum_{l=1}^q \mathbb{E}|Z_{l,t_k}^{N,I,I} - \mathcal{P}_{p_{l,k}}(Z_{l,t_k}^N)|^2$$

$$\leq \sum_{l=1}^q \mathbb{E}|\mathcal{R}_{p_{l,k}}(Z_{l,t_k}^N)|^2 + \frac{d}{h}\big(\mathbb{E}[Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N]^2 - \mathbb{E}[\mathbb{E}_k(Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N)]^2\big), \tag{I.3.15}$$

$$\mathbb{E}|Y_{t_k}^{N,\infty,I} - \mathcal{P}_{p_{0,k}}(Y_{t_k}^N)|^2 \leq (1+\gamma h)\mathbb{E}|\mathbb{E}_k[Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N]|^2$$
$$+ Ch(h + \frac{1}{\gamma})\big[\mathbb{E}|Y_{t_k}^{N,\infty,I} - Y_{t_k}^N|^2 + \mathbb{E}|Z_{t_k}^{N,I,I} - Z_{t_k}^N|^2\big]. \tag{I.3.16}$$

En injectant (I.3.15) dans (I.3.16), en choisissant $\gamma = Cd$ et en utilisant (I.3.14), on obtient alors

$$(1-Ch)\mathbb{E}|Y_{t_k}^{N,\infty,I} - \mathcal{P}_{p_{0,k}}(Y_{t_k}^N)|^2 \leq (1+Ch)\eta_{k+1}^{N,I}$$
$$+ Ch\sum_{l=1}^q \mathbb{E}|\mathcal{R}_{p_{l,k}}(Z_{l,t_k}^N)|^2 + Ch\mathbb{E}|\mathcal{R}_{p_{0,k}}(Y_{t_k}^N)|^2. \tag{I.3.17}$$

On complète alors la preuve de (I.3.12) en injectant cette dernière inégalité dans (I.3.13).
Etape 4 : Majoration de $\zeta^N = h\sum_{k=0}^{N-1}\mathbb{E}|Z_{t_k}^{N,I,I} - Z_{t_k}^N|^2$.
Nous voulons montrer

$$\zeta^N \leq Ch^{2I-2}\mathcal{A}^N(X_0) + Ch\sum_{k=0}^{N-1}\sum_{l=1}^q \mathbb{E}|\mathcal{R}_{p_{l,k}}(Z_{l,t_k}^N)|^2$$
$$+ C\sum_{k=0}^{N-1}\mathbb{E}|\mathcal{R}_{p_{0,k}}(Y_{t_k}^N)|^2 + C\max_{0\leq k\leq N-1}\eta_k^{N,I}. \tag{I.3.18}$$

Vu (I.3.15), on a $\zeta^N \leq h\sum_{k=0}^{N-1}\sum_{l=1}^q \mathbb{E}|\mathcal{R}_{p_{l,k}}(Z_{l,t_k}^N)|^2 + d\sum_{k=0}^{N-1}\big(\mathbb{E}[Y_{t_k}^{N,I,I} - Y_{t_k}^N]^2 - \mathbb{E}[\mathbb{E}_k(Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N)]^2\big)$. A partir de (I.3.13) et (I.3.16), on obtient

$$\mathbb{E}|Y_{t_k}^{N,I,I} - Y_{t_k}^N|^2 - \mathbb{E}[\mathbb{E}_k(Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N)]^2 \leq Ch^{2I-1}\mathcal{A}^N(X_0)$$
$$+ C\mathbb{E}|\mathcal{R}_{p_{0,k}}(Y_{t_k}^N)|^2 + [(1+h)(1+\gamma h) - 1]\mathbb{E}|\mathbb{E}_k[Y_{t_{k+1}}^{N,I,I} - Y_{t_{k+1}}^N]|^2$$
$$+ Ch(h + \frac{1}{\gamma})\big[\mathbb{E}|Y_{t_k}^{N,\infty,I} - Y_{t_k}^N|^2 + \mathbb{E}|Z_{t_k}^{N,I,I} - Z_{t_k}^N|^2\big].$$

En choisissant maintenant $\gamma = 4Cd$ et h suffisamment petit pour que $dC(h + \frac{1}{\gamma}) \leq \frac{1}{2}$, nous avons alors prouvé

$$\zeta^N \leq Ch^{2I-2}\mathcal{A}^N(X_0) + Ch\sum_{k=0}^{N-1}\sum_{l=1}^q \mathbb{E}|\mathcal{R}_{p_{l,k}}(Z_{l,t_k}^N)|^2 + C\sum_{k=0}^{N-1}\mathbb{E}|\mathcal{R}_{p_{0,k}}(Y_{t_k}^N)|^2$$
$$+ C\max_{0\leq k\leq N-1}\eta_k^{N,I} + \frac{1}{2}h\sum_{k=0}^{N-1}\mathbb{E}|Y_{t_k}^{N,\infty,I} - Y_{t_k}^N|^2 + \frac{1}{2}\zeta^N.$$

En utilisant alors (I.3.14) et (I.3.17) pour estimer $\mathbb{E}|Y_{t_k}^{N,\infty,I} - Y_{t_k}^N|^2$, on obtient finalement (I.3.18). Ceci complète la preuve du théorème I.2.

□

Chapitre 4

Etape 3 : méthode de Monte-Carlo

Cette étape est très analogue à l'étape 2, à ceci près que l'on remplace dans la suite de problèmes de moindres carrés, $\mathbb{E}$ par une moyenne empirique basée sur M simulations indépendantes de $(P^N_{t_k})_{0\leq k\leq N}, (\Delta W_k)_{0\leq k\leq N-1}$. Ceci conduit à l'algorithme présenté dans la sous-section 1.2.2. Dans cette procédure, des fonctions de troncature $\hat{\rho}^N_{l,k}$ et $\hat{\rho}^{N,m}_{l,k}$ sont utilisées. Nous devons à présent les spécifier.
Ces troncatures proviennent d'estimées a priori sur $Y^{N,i,I}_{t_k}, Z^{N,i,I}_{l,t_k}$ et il est utile de forcer nos candidats $Y^{N,i,I,M,m}_{t_k}, Z^{N,i,I,M,m}_{l,t_k}$ à satisfaire les mêmes estimées. Ces estimées a priori sont données dans le résultat suivant (prouvé un peu plus tard)

Proposition I.3 *Sous les hypothèses 0.1, I.1 et I.2, pour une constante C_0 suffisamment grande, la suite de fonctions $(\rho^N_{l,k}(\cdot) = \max(1, C_0|p_{l,k}(\cdot)|) : 0 \leq l \leq q, 0 \leq k \leq N-1)$ est telle que*

$$|Y^{N,i,I}_{t_k}| \leq \rho^N_{0,k}(P^N_{t_k}), \quad \sqrt{h}|Z^{N,i,I}_{l,t_k}| \leq \rho^N_{l,k}(P^N_{t_k}), p.s.,$$

pour $i \geq 0$, $I \geq 0$ et $0 \leq k \leq N-1$.

Avec les notations de la sous-section 1.2.1, la définition des fonctions (aléatoires) de troncature $\hat{\rho}^N_{l,k}$ (resp. $\hat{\rho}^{N,m}_{l,k}$) en découle. Elles sont telles que :
- elles laissent invariant $\alpha^{i,I}_{0,k} \cdot p_{0,k} = Y^{N,i,I}_{t_k}$ si $l = 0$ ou $\sqrt{h}\alpha^{i,I}_{l,k} \cdot p_{l,k} = \sqrt{h}Z^{N,i,I}_{l,t_k}$ si $l \geq 1$ (resp. $\alpha^{i,I}_{0,k} \cdot p^m_{0,k}$ si $l = 0$ ou $\sqrt{h}\alpha^{i,I}_{l,k} \cdot p^m_{l,k}$ si $l \geq 1$) ;
- elles sont bornées par $2\rho^N_{l,k}(P^N_{t_k})$ (resp. $2\rho^N_{l,k}(P^{N,m}_{t_k})$) ;
- leur dérivée première est bornée par 1 ;
- leur dérivée seconde est bornée uniformément en N, l, k, m.

4.1 Erreur commise à l'étape 3

A présent, nous allons quantifier l'erreur entre $(Y^{N,I,I,M}, Z^{N,I,I,M})$ et $(Y^{N,I,I}, Z^{N,I,I})$, en fonction du nombre de simulations M, des bases de fonctions et du pas de temps h.

L'analyse est ici plus compliquée que dans [BT04] puisque tous les opérateurs de régression sont corrélés par le même ensemble de trajectoires simulées. Pour obtenir des estimations plus pratiques, on fera l'hypothèse que chaque base de fonction $p_{l,k}$ est orthonormale pour la loi de $P^N_{t_k}$ ($\mathbb{E}(p_{l,k}p^*_{l,k}) = \mathrm{Id}$). Bien sûr, cette hypothèse n'affecte pas le schéma numérique, puisque la projection sur une base de fonctions est invariante par une transformation linéaire de la base.
En outre, on définit l'évènement

$$\mathbf{A}^M_k = \{\forall\ j \in \{k, \cdots, N-1\} : \|V^M_j - \mathrm{Id}\| \leq h, \|P^M_{0,j} - \mathrm{Id}\| \leq h \\ \text{et } \|P^M_{l,j} - \mathrm{Id}\| \leq 1 \text{ pour } 1 \leq l \leq q\} \tag{I.4.1}$$

(voir les notations de la sous-section 1.2.1 pour la définition des matrices V^M_j et $P^M_{l,j}$). Sous l'hypothèse d'orthornormalité pour chaque base $p_{l,k}$, les matrices $(V^M_k)_{0\leq k\leq N-1}$, $(P^M_{l,k})_{0\leq l\leq q, 0\leq k\leq N-1}$ convergent vers l'identité avec probabilité 1 lorsque $M \to \infty$. Ainsi, $\lim_{M\to\infty} \mathbb{P}(\mathbf{A}^M_k) = 1$. Nous donnons à présent notre résultat principal quant à l'influence du nombre de simulations.

Théorème I.3 *On suppose que les hypothèses 0.1, I.1 et I.2 sont satisfaites, que $I \geq 3$, que chaque base de fonctions $p_{l,k}$ est orthonormale et que $\mathbb{E}|p_{l,k}|^4 < \infty$ pour tout k, l. Pour h suffisamment petit, on a pour tout $0 \leq k \leq N-1$*

$$\begin{aligned}
&\mathbb{E}|Y^{N,I,I}_{t_k} - Y^{N,I,I,M}_{t_k}|^2 + h\sum_{j=k}^{N-1} \mathbb{E}|Z^{I,I}_{t_j} - Z^{I,I,M}_{t_j}|^2 \\
&\leq 9\sum_{j=k}^{N-1} \mathbb{E}(|\rho^N_j(P^N_{t_j})|^2 \mathbf{1}_{[\mathbf{A}^M_k]^c}) + Ch^{I-1}\sum_{j=k}^{N-1} \left[1 + |X_0|^2 + \mathbb{E}|\rho^N_j(P^N_{t_j})|^2\right] \\
&+ \frac{C}{hM}\sum_{j=k}^{N-1} \Bigg(\mathbb{E}\|v_j v^*_j - \mathrm{Id}\|^2_F\ \mathbb{E}|\rho^N_j(P^N_{t_j})|^2 + \mathbb{E}(|v_j|^2|p_{0,j+1}|^2)\mathbb{E}|\rho^N_{0,j}(P^N_{t_j})|^2 \\
&+ h^2\mathbb{E}\big[|v_j|^2(1 + |X^N_{t_j}|^2 + |p_{0,j}|^2\mathbb{E}|\rho^N_{0,j}(P^N_{t_j})|^2 + \frac{1}{h}\sum_{l=1}^{q} |p_{l,j}|^2\mathbb{E}|\rho^N_{l,j}(P^N_{t_j})|^2)\big]\Bigg).
\end{aligned}$$

Nous avons mentionné plus haut que le terme impliquant $[\mathbf{A}^M_k]^c$ converge vers 0 lorsque $M \to \infty$ mais nous n'avons pas plus explicité d'estimations concernant ce terme dans l'énoncé du théorème I.3. En effet, l'énoncé d'une borne optimale pour ce terme dépend d'hypothèses supplémentaires sur les moments des fonctions de base. Par exemple, si $\rho^N_j(P^N_{t_j})$ a des moments d'ordre supérieur à 2, il vient :
$\mathbb{P}([\mathbf{A}^M_k]^c) \leq \sum_{j=k}^{N-1} \big[\mathbb{P}(\|V^M_j - \mathrm{Id}\| > h) + \mathbb{P}(\|P^M_{0,j} - \mathrm{Id}\| > h) + \sum_{l=1}^{q} \mathbb{P}(\|P^M_{l,j} - \mathrm{Id}\| > 1)\big]$. Nous avons $\mathbb{P}(\|V^M_k - \mathrm{Id}\| > h) \leq h^{-2}\mathbb{E}\|V^M_k - \mathrm{Id}\|^2 \leq h^{-2}\mathbb{E}\|V^M_k - \mathrm{Id}\|^2_F = (Mh^2)^{-1}\mathbb{E}\|v_k v^*_k - \mathrm{Id}\|^2_F$. Remarquons que le théorème I.3 procure une borne d'erreur non-asymptotique. Associé aux théorèmes I.1 et I.2, il nous permet (en théorie) de choisir optimalement le pas de

temps h, les bases de fonctions et le nombre de simulations pour atteindre une précision donnée. Nous reviendrons un peu plus tard sur ce choix conjoint de paramètres.
Il est également possible de montrer un résultat asymptotique de type théorème de la limite centrale.

Théorème I.4 *On suppose que les hypothèses 0.1, I.1 et I.2 sont satisfaites, que le driver f est de classe C^1 par rapport à (y,z) avec des dérivées bornées uniformément Höldériennes, et que $\mathbb{E}|p_{l,k}|^{2+\varepsilon} < \infty$ pour tout k,l ($\varepsilon > 0$). Alors, le vecteur $[\sqrt{M}(\alpha_k^{i,I,M} - \alpha_k^{i,I})]_{i\leq I, k\leq N-1}$ converge faiblement vers un vecteur gaussien centré lorsque M tend vers l'infini.*

Preuve de la proposition I.3. Vu la proposition I.1, il est tentant d'utiliser la propriété de Markov et d'affirmer que la proposition I.3 résulte de (I.3.11) écrite avec les espérances conditionnelles $\mathbb{E}_k$. Mais cet argument ne marche pas car la loi utilisée dans la projection n'est pas la loi conditionnelle $\mathbb{E}_k$ mais $\mathbb{E}_0$. Le bon argument pourrait être le suivant. On écrit $Y_{t_k}^{N,i,I} = \alpha_{0,k}^{i,I} \cdot p_{0,k}(P_{t_k}^N)$. D'une part, par (I.3.11) on a $C\mathcal{A}^N(X_0) \geq \mathbb{E}|Y_{t_k}^{N,i,I}|^2 = \alpha_{0,k}^{i,I} \cdot \mathbb{E}[p_{0,k}p_{0,k}^*]\alpha_{0,k}^{i,I} \geq |\alpha_{0,k}^{i,I}|^2 \lambda_{\min}(\mathbb{E}[p_{0,k}p_{0,k}^*])$. D'autre part, $|Y_{t_k}^{N,i,I}| \leq |\alpha_{0,k}^{i,I}||p_{0,k}(P_{t_k}^N)| \leq |p_{0,k}|\sqrt{C\mathcal{A}^N(X_0)/\lambda_{\min}(\mathbb{E}[p_{0,k}p_{0,k}^*])}$.
Ainsi, on peut prendre $\rho_{0,k}^N(x) = \max(1, |p_{0,k}(x)|\sqrt{C\mathcal{A}^N(X_0)/\lambda_{\min}(\mathbb{E}[p_{0,k}p_{0,k}^*])})$. De même pour $\sqrt{h}|Z_{l,t_k}^{N,i,I}|$, on a $\rho_{l,k}^N(x) = \max(1, |p_{l,k}(x)|\sqrt{C\mathcal{A}^N(X_0)/\lambda_{\min}(\mathbb{E}[p_{l,k}p_{l,k}^*])})$. Remarquons que si $p_{l,k}$ est une base orthonormale, on a $\lambda_{\min}(\mathbb{E}[p_{l,k}p_{l,k}^*]) = 1$ et les bornes supérieures précédentes se simplifient. □

Preuve du théorème I.3. Dans la suite, on pose

$$\mathcal{A}_k^{N,M} = \frac{1}{M}\sum_{m=1}^M |\rho_{0,k}^N(P_{t_k}^{N,m})|^2, \quad \mathcal{B}_k^{N,M} = \frac{1}{M}\sum_{m=1}^M |f_k^m(0,\cdots,0)|^2.$$

De manière évidente, on a $\mathbb{E}(\mathcal{A}_k^{N,M}) = \mathbb{E}|\rho_{0,k}^N(P_{t_k}^N)|^2$ et $\mathbb{E}(\mathcal{B}_k^{N,M}) \leq C(1+|X_0|^2)$. A présent, nous rappelons la propriété standard de contraction dans le cas d'un problème de moindres carrés dans $\mathbb{R}^M$, analogue au cas $\mathbf{L}_2(\Omega,\mathbb{P})$. On considère une suite de réels $(x^m)_{1\leq m\leq M}$ et une suite $(v^m)_{1\leq m\leq M}$ de vecteurs de $\mathbb{R}^n$, associée à la matrice $V^M = \frac{1}{M}\sum_{m=1}^M v^m[v^m]^*$ qui est supposée inversible ($\lambda_{\min}(V^M) > 0$). Alors, l'unique vecteur de $\mathbb{R}^n$ $\theta_x = \arg\inf_\theta |x - \theta\cdot v|_M^2$ est donné par

$$\theta_x = \frac{[V^M]^{-1}}{M}\sum_{m=1}^M v^m x^m. \tag{I.4.2}$$

L'application $x \mapsto \theta_x$ est linéaire et de plus, on a l'inégalité

$$\lambda_{\min}(V^M)|\theta_x|^2 \leq |\theta_x.v|_M^2 \leq |x|_M^2. \tag{I.4.3}$$

Pour la suite des calculs, il sera plus pratique d'utiliser

$$(\theta_k^{i,I,M})^* = \big((\alpha_{0,k}^{i,I,M})^*, \sqrt{h}(\alpha_{1,k}^{i,I,M})^*, \cdots, \sqrt{h}(\alpha_{q,k}^{i,I,M})^*\big)$$

à la place de $\alpha_k^{i,I,M}$. Alors, les itérations de Picard de (I.1.3) peuvent se réécrire

$$\theta_k^{i+1,I,M} = \arg\inf_\theta \frac{1}{M}\sum_{m=1}^{M}\big(\hat{\rho}_{0,k+1}^{N,m}(\alpha_{0,k+1}^{I,I,M}.p_{0,k+1}^m) + hf_k^m(\alpha_k^{i,I,M}) - \theta.v_k^m\big)^2. \tag{I.4.4}$$

En introduisant l'évènement $\mathbf{A}_k^M$, en prenant en compte la propriété de Lipschitz des fonctions $\hat{\rho}_{l,k}^N$ et en utilisant l'orthonormalité de $p_{l,k}$, on obtient

$$\begin{aligned}\mathbb{E}|Y_{t_k}^{N,I,I} - Y_{t_k}^{N,I,I,M}|^2 + h\sum_{j=k}^{N-1}\mathbb{E}|Z_{t_j}^{N,I,I} - Z_{t_j}^{N,I,I,M}|^2 \le 9\sum_{j=k}^{N-1}\mathbb{E}(|\rho_j^N(P_{t_j}^N)|^2\mathbf{1}_{[\mathbf{A}_k^M]^c})\\ + \mathbb{E}(\mathbf{1}_{\mathbf{A}_k^M}|\alpha_{0,k}^{I,I,M} - \alpha_{0,k}^{I,I}|^2) + h\sum_{j=k}^{N-1}\sum_{l=1}^{q}\mathbb{E}(\mathbf{1}_{\mathbf{A}_k^M}|\alpha_{l,j}^{I,I,M} - \alpha_{l,j}^{I,I}|^2).\end{aligned} \tag{I.4.5}$$

Pour établir le théorème I.3, on va estimer $|\theta_k^{I,I,M} - \theta_k^{I,I}|^2$ sur l'évènement $\mathbf{A}_k^M$. Ceci est réalisé en plusieurs étapes.

Etape 1 : propriétés de contraction relatives à la suite $(\theta_k^{i,I,M})_{i\ge 0}$. Elles sont résumées dans le lemme suivant.

Lemme I.1 *Pour h suffisamment petit, sur $\mathbf{A}_k^M$, les propriétés suivantes sont vraies.*
a) $|\theta_k^{i+1,I,M} - \theta_k^{i,I,M}|^2 \le Ch|\theta_k^{i,I,M} - \theta_k^{i-1,I,M}|^2$.
b) Il existe un unique vecteur $\theta_k^{\infty,I,M}$ tel que

$$\theta_k^{\infty,I,M} = \arg\inf_\theta \frac{1}{M}\sum_{m=1}^{M}\big(\hat{\rho}_{0,k+1}^{N,m}(\alpha_{0,k+1}^{I,I,M}.p_{0,k+1}^m) + hf_k^m(\alpha_k^{\infty,I,M}) - \theta.v_k^m\big)^2.$$

c) On a $|\theta_k^{\infty,I,M} - \theta_k^{I,I,M}|^2 \le [Ch]^I|\theta_k^{\infty,I,M}|^2$.

Preuve. Montrons a). Puisque $1 - h \le \lambda_{\min}(V_k^M)$ et $\lambda_{\max}(P_{l,k}^M) \le 2$ $(0 \le l \le q)$ sur $\mathbf{A}_k^M$, au vu de (I.4.3) on obtient que $(1-h)|\theta_k^{i+1,I,M} - \theta_k^{i,I,M}|^2$ est borné par

$$\begin{aligned}&\frac{h^2}{M}\sum_{m=1}^{M}\big(f_k^m(\alpha_k^{i,I,M}) - f_k^m(\alpha_k^{i-1,I,M})\big)^2\\ \le& Ch^2\sum_{l=0}^{q}|\alpha_{l,k}^{i,I,M} - \alpha_{l,k}^{i-1,I,M}|^2\lambda_{\max}(P_{l,k}^M) \le Ch|\theta_k^{i,I,M} - \theta_k^{i-1,I,M}|^2.\end{aligned}$$

A présent, a) et b) sont clairs. Pour c), on applique a) en se rappelant que $\theta_k^{0,I,M} = 0$. □

Etape 2 : bornes pour $|\theta_k^{i,I,M}|$ sur l'évènement $\mathbf{A}_k^M$. En effet, nous voulons montrer que

$$|\theta_k^{i,I,M}|^2 \leq C\big(\mathcal{A}_{k+1}^{N,M} + h\mathcal{B}_k^{N,M}\big) \text{ sur } \mathbf{A}_k^M. \tag{I.4.6}$$

Tout d'abord, nous considérons $i = \infty$. Comme dans la preuve du lemme I.1, on obtient

$$\begin{aligned}(1-h)|\theta_k^{\infty,I,M}|^2 &\leq \frac{1}{M}\sum_{m=1}^{M}\big[\hat{\rho}_{0,k+1}^{N,m}(\alpha_{0,k+1}^{I,I,M}.p_{0,k+1}^m) + hf_k^m(\alpha_k^{\infty,I,M})\big]^2 \\ &\leq (1+\gamma h)\mathcal{A}_{k+1}^{N,M} + Ch(h+\frac{1}{\gamma})\big(\mathcal{B}_k^{N,M} + \sum_{l=0}^{q}|\alpha_{l,k}^{\infty,I,M}|^2\lambda_{\max}(P_{l,k}^M)\big).\end{aligned}$$

On prend $\gamma = 8C$ et h suffisamment petit pour assurer $2C(h+\frac{1}{\gamma})(1+h) \leq \frac{1}{2}(1-h)$. Il s'ensuit $|\theta_k^{\infty,I,M}|^2 \leq C(\mathcal{A}_{k+1}^{N,M} + h\mathcal{B}_k^{N,M})$, montrant que (I.4.6) est valable pour $i = \infty$. Le lemme I.1-c) conduit aux bornes attendues pour d'autres valeurs de i.

Etape 3 : on rappelle des bornes pour $\theta^{i,I}$. En utilisant la proposition I.3 et au vu de (I.3.2-I.3.6), on a pour $i \geq 1$

$$|\theta_{l,k}^{i,I}|^2 \leq \mathbb{E}|\rho_{l,k}^N(P_{t_k}^N)|^2,\ 0 \leq l \leq q; \quad |\theta_k^{\infty,I} - \theta_k^{i,I}|^2 \leq (C_f h)^{2i}\mathbb{E}|\rho_{0,k}^N(P_{t_k}^N)|^2. \tag{I.4.7}$$

Rappelons aussi l'expression suivante de $\theta_k^{\infty,I}$, dérivée de (I.3.2-I.3.5) et de l'orthonormalité de chaque base $p_{l,k}$:

$$\theta_k^{\infty,I} = \mathbb{E}\big(v_k[\alpha_{0,k+1}^{I,I}.p_{0,k+1} + hf_k(\alpha_k^{\infty,I})]\big). \tag{I.4.8}$$

Etape 4 : Décomposition de la quantité $\mathbb{E}(\mathbf{1}_{\mathbf{A}_k^M}|\theta_k^{I,I,M} - \theta_k^{I,I}|^2)$. A cause du lemme I.1, sur $\mathbf{A}_k^M$ on obtient $|\theta_k^{\infty,I,M} - \theta_k^{I,I,M}|^2 \leq Ch^I|\theta_k^{\infty,I,M}|^2 \leq Ch^I|\theta_k^{\infty,I}|^2 + Ch^I|\theta_k^{\infty,I,M} - \theta_k^{\infty,I}|^2$. Ainsi, en utilisant (I.4.7), $\mathbb{E}(1_{\mathbf{A}_k^M}|\theta_k^{I,I,M} - \theta_k^{I,I}|^2)$ est borné par

$$\begin{aligned}&(1+h)\mathbb{E}(1_{\mathbf{A}_k^M}|\theta_k^{\infty,I,M} - \theta_k^{\infty,I}|^2) \\ &\quad + 2(1+\frac{1}{h})\{\mathbb{E}(1_{\mathbf{A}_k^M}|\theta_k^{I,I,M} - \theta_k^{\infty,I,M}|^2) + |\theta_k^{I,I} - \theta_k^{\infty,I}|^2\} \\ &\leq (1+Ch)\mathbb{E}(1_{\mathbf{A}_k^M}|\theta_k^{\infty,I,M} - \theta_k^{\infty,I}|^2) + Ch^{I-1}\mathbb{E}|\rho_k^N(P_{t_k}^N)|^2\end{aligned} \tag{I.4.9}$$

en tenant compte de $I \geq 3$. Sur $\mathbf{A}_k^M$, V_k^M est inversible et on peut poser

$$B_1 = (\mathrm{Id} - (V_k^M)^{-1})\theta_k^{\infty,I}$$

$$B_2 = (V_k^M)^{-1}\big[\mathbb{E}(v_k\hat{\rho}_{0,k+1}^N(\alpha_{0,k+1}^{I,I}\cdot p_{0,k+1})) - \frac{1}{M}\sum_{m=1}^{M}v_k^m\hat{\rho}_{0,k+1}^{N,m}(\alpha_{0,k+1}^{I,I}\cdot p_{0,k+1}^m)\big],$$

$$B_3 = (V_k^M)^{-1}h\big[\mathbb{E}(v_k f_k(\alpha_k^{\infty,I})) - \frac{1}{M}\sum_{m=1}^{M}v_k^m f_k^m(\alpha_k^{\infty,I})\big],$$

$$\begin{aligned}B_4 = \frac{(V_k^M)^{-1}}{M}\sum_{m=1}^{M}v_k^m\big[&\hat{\rho}_{0,k+1}^{N,m}(\alpha_{0,k+1}^{I,I}\cdot p_{0,k+1}^m) - \hat{\rho}_{0,k+1}^{N,m}(\alpha_{0,k+1}^{I,I,M}\cdot p_{0,k+1}^m) \\ &+ h(f_k^m(\alpha_k^{\infty,I}) - f_k^m(\alpha_k^{\infty,I,M}))\big].\end{aligned}$$

Ainsi, par (I.4.2-I.4.8) on peut décomposer $\theta_k^{\infty,I} - \theta_k^{\infty,I,M} = B_1 + B_2 + B_3 + B_4$, ce qui donne sur $\mathbf{A}_k^M$

$$|\theta_k^{\infty,I} - \theta_k^{\infty,I,M}|^2 \leq 3(1+\frac{1}{h})(|B_1|^2 + |B_2|^2 + |B_3|^2) + (1+h)|B_4|^2. \tag{I.4.10}$$

Etape 5 : estimations séparées de B_1, B_2, B_3, B_4 sur $\mathbf{A}_k^M$. On rappelle le résultat classique ([GVL96]) : Si $\|\mathrm{Id} - F\| < 1$, $F^{-1} - \mathrm{Id} = \sum_{k=1}^{\infty}[\mathrm{Id} - F]^k$ et $\|\mathrm{Id} - F^{-1}\| \leq \frac{\|F-\mathrm{Id}\|}{1-\|F-\mathrm{Id}\|}$. Par conséquent, pour $F = V_k^M$ on obtient $\mathbb{E}(\mathbf{1}_{\mathbf{A}_k^M}\|\mathrm{Id} - (V_k^M)^{-1}\|^2) \leq (1-h)^{-2}\mathbb{E}\|\mathrm{Id} - V_k^M\|^2 \leq (1-h)^{-2}\mathbb{E}\|V_k^M - \mathrm{Id}\|_F^2 = (M(1-h)^2)^{-1}\mathbb{E}\|v_k v_k^* - \mathrm{Id}\|_F^2$. Ainsi, on a

$$\mathbb{E}(|B_1|^2 1_{\mathbf{A}_k^M}) \leq \frac{C}{M}\mathbb{E}\|v_k v_k^* - \mathrm{Id}\|_F^2 \; \mathbb{E}|\rho_k^N(P_{t_k}^N)|^2.$$

Puisque sur $\mathbf{A}_k^M$, $\|(V_k^M)^{-1}\| \leq 2$, il s'ensuit

$$\mathbb{E}(|B_2|^2 1_{\mathbf{A}_k^M}) \leq \frac{C}{M}\mathbb{E}(|v_k|^2|p_{0,k+1}|^2)\mathbb{E}|\rho_{0,k}^N(P_{t_k}^N)|^2,$$

$$\mathbb{E}(|B_3|^2 1_{\mathbf{A}_k^M}) \leq \frac{Ch^2}{M}\mathbb{E}\big[|v_k|^2(1 + |X_{t_k}^N|^2 + |p_{0,k}|^2\mathbb{E}|\rho_{0,k}^N(P_{t_k}^N)|^2 + \frac{1}{h}\sum_{l=1}^{q}|p_{l,k}|^2\mathbb{E}|\rho_{l,k}^N(P_{t_k}^N)|^2)\big].$$

Comme dans la preuve du lemme I.1 et en utilisant $\|P_{0,k+1}^M\| \leq 1 + h$ sur $\mathbf{A}_k^M$, on obtient facilement

$$(1-h)|B_4|^2 \leq (1+h)(1+\gamma h)|\alpha_{0,k+1}^{I,I} - \alpha_{0,k+1}^{I,I,M}|^2 + Ch(h + \frac{1}{\gamma})\sum_{l=0}^{q}|\alpha_{l,k}^{\infty,I} - \alpha_{l,k}^{\infty,I,M}|^2.$$

Etape 6 : estimations finales.
Soit $\epsilon_k = \mathbb{E}\|v_k v_k^* - \mathrm{Id}\|_F^2 \; \mathbb{E}|\rho_k^N(P_{t_k}^N)|^2 + \mathbb{E}(|v_k|^2|p_{0,k+1}|^2)\mathbb{E}|\rho_{0,k}^N(P_{t_k}^N)|^2 + h^2\mathbb{E}\big[|v_k|^2(1+|X_{t_k}^N|^2 + |p_{0,k}|^2\mathbb{E}|\rho_{0,k}^N(P_{t_k}^N)|^2 + \frac{1}{h}\sum_{l=1}^{q}|p_{l,k}|^2\mathbb{E}|\rho_{l,k}^N(P_{t_k}^N)|^2)\big]$.
On injecte les estimées ci-dessus de B_1, B_2, B_3, B_4 dans (I.4.10), on choisit $\gamma = 3C$ et h proche de 0 pour assurer que $Ch + \frac{C}{\gamma} \leq \frac{1}{2}$; après simplifications, on obtient

$$\mathbb{E}(1_{\mathbf{A}_k^M}|\theta_k^{\infty,I,M} - \theta_k^{\infty,I}|^2) \leq C\frac{\epsilon_k}{hM} + (1+Ch)\mathbb{E}(1_{\mathbf{A}_k^M}|\alpha_{0,k+1}^{I,I} - \alpha_{0,k+1}^{I,I,M}|^2).$$

Mais au vu du lemme I.1-c) et des estimées (I.4.7-I.4.6), on obtient $\mathbb{E}(1_{\mathbf{A}_k^M}|\alpha_{0,k+1}^{I,I} - \alpha_{0,k+1}^{I,I,M}|^2) \leq (1+h)\mathbb{E}(1_{\mathbf{A}_k^M}|\alpha_{0,k+1}^{\infty,I} - \alpha_{0,k+1}^{\infty,I,M}|^2) + Ch^{I-1}\big(1 + |X_0|^2 + \mathbb{E}|\rho_{0,k+1}^N(P_{t_{k+1}}^N)|^2 + \mathbb{E}|\rho_{0,k+2}^N(P_{t_{k+2}}^N)|^2\big)$. Finalement, nous avons montré

$$\mathbb{E}(1_{\mathbf{A}_k^M}|\theta_k^{\infty,I,M} - \theta_k^{\infty,I}|^2) \leq C\frac{\epsilon_k}{hM} + Ch^{I-1}\big(1 + |X_0|^2 + \mathbb{E}|\rho_{0,k+1}^N(P_{t_{k+1}}^N)|^2 + \mathbb{E}|\rho_{0,k+2}^N(P_{t_{k+2}}^N)|^2\big) + (1+Ch)\mathbb{E}(1_{\mathbf{A}_k^M}|\alpha_{0,k+1}^{\infty,I,M} - \alpha_{0,k+1}^{\infty,I}|^2).$$

En utilisant l'argument de contraction comme dans (I.4.9), l'indice ∞ peut être remplacé par I, sans changer l'inégalité (avec une constante C potentiellement différente). Ceci peut être écrit

$$\begin{aligned}
&\mathbb{E}(1_{\mathbf{A}_k^M}|\alpha_{0,k}^{I,I,M} - \alpha_{0,k}^{I,I}|^2) + h\sum_{l=1}^{q}\mathbb{E}(1_{\mathbf{A}_k^M}|\alpha_{l,k}^{I,I,M} - \alpha_{l,k}^{I,I}|^2) \\
&\leq C\frac{\epsilon_k}{hM} + Ch^{I-1}\big(1 + |X_0|^2 + \mathbb{E}|\rho_{0,k+1}^N(P_{t_{k+1}}^N)|^2 + \mathbb{E}|\rho_{0,k+2}^N(P_{t_{k+2}}^N)|^2\big) \\
&\quad + (1+Ch)\mathbb{E}(1_{\mathbf{A}_k^M}|\alpha_{0,k+1}^{I,I,M} - \alpha_{0,k+1}^{I,I}|^2).
\end{aligned}$$

On complète la preuve en utilisant le lemme de Gronwall discret A.3. □

Remarque I.2 *On peut remarquer que les puissances de h sont plus petites ici que dans le théorème I.2, ce qui oblige à prendre $I \geq 3$ au lieu de $I \geq 2$ avant. En effet, on ne peut pas utiliser d'espérances conditionnelles par rapport aux simulations comme nous fimes pour* (I.3.4) *par exemple.*
Notons que dans la preuve ci-dessus, on a seulement utilisé la propriété de Lipschitz des fonctions de troncature $\hat{\rho}_{l,k}^N$ and $\hat{\rho}_{l,k}^{N,m}$.

Preuve du théorème I.4. Les arguments sont standards et il y a essentiellement des difficultés de notations. Les dérivées premières de f par rapport à y et z_l sont respectivement notées $\partial_0 f$ et $\partial_l f$. Le paramètre $\beta \in]0,1]$ représente leur indice de Hölder. Supposons sans perte de généralité que $\varepsilon < \beta$ et que chaque base de fonctions $p_{l,k}$ est orthonormale. Pour $k < N-1$, on définit les quantités

$$\begin{aligned}
A_{l,k}^M(\alpha) =& \frac{1}{M}\sum_{m=1}^{M} v_k^m \partial_l f(t_k, X_{t_k}^{N,m}, \alpha_0 \cdot p_{0,k}^m, \cdots, \alpha_q \cdot p_{q,k}^m)[p_{l,k}^m]^*, \\
B_k^M =& \frac{1}{M}\sum_{m=1}^{M} v_k^m [p_{0,k+1}^m]^*, \qquad D_k^M = \sqrt{M}(\mathrm{Id} - V_k^M), \\
C_k^M(\alpha) =& \sum_{m=1}^{M} \frac{\big\{v_k^m[\alpha_{0,k+1}^{I,I}\cdot p_{0,k+1}^m + hf_k^m(\alpha)] - \mathbb{E}\big(v_k[\alpha_{0,k+1}^{I,I}\cdot p_{0,k+1} + hf_k(\alpha)]\big)\big\}}{\sqrt{M}}.
\end{aligned}$$

Pour $k = N-1$, on pose $B_k^M = 0$ et dans $C_k^M(\alpha)$, les termes $\alpha_{0,k+1}^{I,I}\cdot p_{0,k+1}^m$ et $\alpha_{0,k+1}^{I,I}\cdot p_{0,k+1}$ sont à remplacer respectivement par $\Phi^N(P_{t_N}^{N,m})$ et $\Phi^N(P_{t_N}^N)$. Les définitions de $A_{l,k}^M(\alpha)$ et D_k^M sont toujours valides. Pour simplifier, on écrit $X^M \xrightarrow{w}$ si la suite (potentiellement à valeurs vectorielle ou matricielle) $(X^M)_M$ converge faiblement vers une variable aléatoire gaussienne centrée, lorsque M tend vers l'infini. Pour la convergence en probabilité vers

une constante, on note $X^M \xrightarrow{\mathbb{P}}$. Puisque les simulations sont indépendantes, les convergences suivantes sont vraies :

$$\begin{aligned}(A^M_{l,k}(\alpha^{i,I}_k), B^M_k, V^M_k)_{i\le I-1, l\le q, k\le N-1} &\xrightarrow{\mathbb{P}},\\ \mathcal{G}^M = (C^M_k(\alpha^{i,I}_k), D^M_k)_{i\le I-1, l\le q, k\le N-1} &\xrightarrow{w}.\end{aligned} \tag{I.4.11}$$

Remarquons que $\lim_{M\to\infty} V^M_k \overset{p.s.}{=} \mathrm{Id}$ est inversible. En linéarisant les fonctions f et $\hat\rho^{N,m}_{0,k+1}$ dans les expressions de $\theta^{i,I}_k = \mathbb{E}(v_k[\alpha^{I,I}_{0,k+1}.p_{0,k+1} + hf_k(\alpha^{i-1,I}_{0,k}, \cdots, \alpha^{i-1,I}_{q,k})])$ et $\theta^{i,I,M}_k$ données par (I.4.2) on obtient

$$\begin{aligned}&|V^M_k\sqrt{M}(\theta^{i,I,M}_k - \theta^{i,I}_k) - D^M_k\theta^{i,I}_k - C^M_k(\alpha^{i-1,I}_k)\\ &- B^M_k\sqrt{M}(\alpha^{I,I,M}_{0,k+1} - \alpha^{I,I}_{0,k+1}) - h\sum_{l=0}^{q} A^M_{l,k}(\alpha^{i-1,I}_k)\sqrt{M}(\alpha^{i-1,I,M}_{l,k} - \alpha^{i-1,I}_{l,k})|\\ \le\ &\mathbf{1}_{k<N-1}\frac{C}{\sqrt{M}}|\alpha^{I,I,M}_{0,k+1} - \alpha^{I,I}_{0,k+1}|^2\sum_{m=1}^{M}|v^m_k||p^m_{0,k+1}|^2\\ &+ \frac{C}{\sqrt{M}}|\alpha^{i-1,I,M}_k - \alpha^{i-1,I}_k|^{1+\beta}\sum_{m=1}^{M}|v^m_k||p^m_k|^{1+\beta}.\end{aligned} \tag{I.4.12}$$

Pour obtenir le théorème I.4, on montre par récurrence sur k que $([\sqrt{M}(\theta^{i,I,M}_j - \theta^{i,I}_j)]_{j\ge k, i\le I}, \mathcal{G}^M) \xrightarrow{w}$. Rappelons que $\theta^{0,I,M}_j = \theta^{0,I}_j = 0$ pour tout j.
Considérons d'abord $k = N-1$, pour lequel $B^M_k = 0$, et $i = 1$. Au vu de (I.4.11-I.4.12), clairement $([\sqrt{M}(\theta^{i,I,M}_{N-1} - \theta^{i,I}_{N-1})]_{i\le\mathbf{1}}, \mathcal{G}^M) \xrightarrow{w}$. Pour $i = 2$, on peut invoquer le même argument en utilisant (I.4.11-I.4.12) et on obtient $([\sqrt{M}(\theta^{i,I,M}_{N-1} - \theta^{i,I}_{N-1})]_{i\le\mathbf{2}}, \mathcal{G}^M) \xrightarrow{w}$ pourvu que la borne supérieure dans (I.4.12) converge vers 0 en probabilité. Pour prouver cela, on pose $\mathcal{M}^M = M^{-1-\beta/2}\sum_{m=1}^{M}|v^m_{N-1}||p^m_{N-1}|^{1+\beta}$ et on écrit

$$\frac{1}{\sqrt{M}}|\alpha^{1,I,M}_{N-1} - \alpha^{1,I}_{N-1}|^{1+\beta}\sum_{m=1}^{M}|v^m_{N-1}||p^m_{N-1}|^{1+\beta} = |\sqrt{M}(\alpha^{1,I,M}_{N-1} - \alpha^{1,I}_{N-1})|^{1+\beta}\mathcal{M}^M.$$

Puisque $[\sqrt{M}(\alpha^{1,I,M}_{N-1} - \alpha^{1,I}_{N-1})]_M$ est tendue, notre assertion est vraie si $\mathcal{M}^M$ converge vers 0 lorsque $M\to\infty$. Remarquons que $|v_{N-1}||p_{N-1}|^{1+\beta} \in \mathbf{L}_{\frac{2+\varepsilon}{2+\beta}(\mathbb{P})}$. Ainsi, la loi forte des grands nombres, dans le cas de variables aléatoires i.i.d. avec moyenne infinie, conduit à $\sum_{m=1}^{M}|v^m_{N-1}||p^m_{N-1}|^{1+\beta} = O(M^{\frac{2+\beta}{2+\varepsilon}+r})$ p.s. pour tout $r>0$. Par conséquent, en choisissant r suffisamment petit, il vient $\mathcal{M}^M \to 0$ p.s.
En répétant cet argument, on obtient $([\sqrt{M}(\theta^{i,I,M}_{N-1} - \theta^{i,I}_{N-1})]_{i\le\mathbf{I}}, \mathcal{G}^M) \xrightarrow{w}$.
Pour l'induction lorsque $k < N-1$, on applique les mêmes techniques. Il y a une contribution additionnelle dûe à B^M_k, qui peut être traitée comme avant. □

4.2 Commentaires sur les majorations d'erreurs obtenues

Le théorème I.1 reflète l'influence de la discrétisation en temps alors que le théorème I.2 quantifie dans l'erreur commise par l'algorithme, l'influence du type et du nombre de fonctions de base. En vue d'obtenir une convergence de $(Y^{N,I,I}, Z^{N,I,I})$ vers (Y, Z), il convient donc de faire tendre N vers l'infini puis d'ajuster le type et le nombre de fonctions de base de telle sorte que non seulement l'erreur d'approximation à l'instant t_k

$$\sum_{l=1}^{q} \mathbb{E}|\mathcal{R}_{p_{l,k}}(\sqrt{h}Z^N_{l,t_k})|^2 + \mathbb{E}|\mathcal{R}_{p_{0,k}}(Y^N_{t_k})|^2$$

tende vers 0 mais également suffisamment vite pour que

$$\sum_{k=0}^{N-1} \left(\sum_{l=1}^{q} \mathbb{E}|\mathcal{R}_{p_{l,k}}(\sqrt{h}Z^N_{l,t_k})|^2 + \mathbb{E}|\mathcal{R}_{p_{0,k}}(Y^N_{t_k})|^2 \right) \tag{I.4.13}$$

tende vers 0. Ceci sera possible car nous avons des informations sur $\left(y^N_k(\cdot), \sqrt{h}z^N_k(\cdot)\right)$ qui sont exposées dans la proposition I.2. En particulier, pour tirer parti du caractère Lipschitz de ces fonctions inconnues, on utilisera entre autres dans la partie numérique, des bases locales dont les supports seront des petits hypercubes d'arête δ. Bien sûr, on fera tendre δ vers 0 de manière à garantir que la somme (I.4.13) tend vers 0 lorsque N tend vers l'infini.

Reste l'erreur Monte-Carlo. En effet, pour assurer que l'erreur entre $(Y^{N,I,I,M}, Z^{N,I,I,M})$ et $(Y^{N,I,I}, Z^{N,I,I})$ tend vers 0 alors que N tend vers l'infini et δ vers 0, il nous faut augmenter le nombre de simulations M. En effet, le majorant de l'erreur dans le théorème I.3 est clairement croissant en N mais également croissant dans le nombre de fonctions de base utilisées à cause (entre autres) des termes $\mathbb{E}\|v_j v_j^* - \mathrm{Id}\|_F^2$ et $\mathbb{E}(|v_j|^2|p_{0,j+1}|^2)$. En effet, prenons par exemple un élément de la base de petits hypercubes qu'on utilisera dans la partie numérique à un instant donné t_k et fixons $d = q = d' = 1$. Cet élément s'exprimera sous la forme $\mathbf{1}_{[a,b]}(x)$ où $b - a = \delta$. Or, le contrôle de l'erreur dans le théorème I.3 se fait à l'aide de la base v_k, orthonormalisation pour la loi de $P^N_{t_k}$ de la base d'hypercubes. Dans notre exemple, la fonction orthonormale correspondant à $\mathbf{1}_{[a,b]}(x)$ sera $\frac{1}{\sqrt{\mathbb{P}(P^N_{t_k} \in [a,b])}} \mathbf{1}_{[a,b]}(x)$. Comme les termes $\mathbb{E}\|v_j v_j^* - \mathrm{Id}\|_F^2$ font intervenir des puissances plus grandes que 2 de cette fonction orthonormale, on va obtenir un contrôle de l'erreur dans le théorème I.3 à l'aide de puissances de $\frac{1}{\mathbb{P}(P^N_{t_k} \in [a,b])}$. On imagine rapidement le problème que cela peut poser lorsque $\delta \to 0$ et que l'on doit estimer des minorants de la probabilité que $P^N_{t_k}$ se trouve dans un petit intervalle de longueur δ.

De manière plus générale, si on sort de cet exemple de bases d'indicatrices de petits

hypercubes, il sera très difficile d'obtenir de manière explicite une évaluation du membre de droite dans le théorème I.3 et donc très difficile de savoir comment faire varier M en fonction du nombre de fonctions de base et de N pour assurer une convergence de l'algorithme.
Un besoin de majorations plus explicites se fait donc clairement sentir et justifie la partie suivante.

Deuxième partie

Algorithme modifié pour les équations différentielles stochastiques rétrogrades

Nous avons exploité dans la partie I la propriété de $(Y_{t_k}^N, Z_{t_k}^N)$ d'être les minimiseurs dans $\mathbf{L}^2$ d'une certaine quantité (voir l'équation (I.2.4)). Nous allons à présent exploiter leur représentation sous forme d'espérance conditionnelle (I.2.2-I.2.3) pour pouvoir affiner les résultats de convergence présentés dans le théorème I.3 et surtout obtenir des majorations d'erreur qui nous permettent de faire varier simultanément tous les paramètres de l'algorithme (pas de temps, type et nombre de fonctions de base, nombre de simulations) pour obtenir sa convergence. On a en effet constaté que les estimations données dans le théorème I.3 impliquaient les moments d'ordre supérieur à 2 des fonctions de base orthonormées pour la loi de P^N et que ces moments étaient difficilement calculables en toute généralité.
Nous allons ici nous efforcer d'obtenir des estimations faisant intervenir seulement le nombre de fonctions de base utilisées pour approcher (Y_{t_k}, Z_{t_k}) et non leurs moments.
Nous allons dans cette partie modifier la discrétisation en temps que nous avons effectuée dans la partie I et verrons que nous devrons localiser cette EDSR discrète de manière à assurer que les solutions de l'EDSR sont bornées. Puis, nous décrirons l'algorithme pour approcher cette EDSR discrète localisée. Nous établirons alors des bornes pour l'erreur commise dans l'algorithme et enfin nous discuterons de l'utilisation en pratique de ces bornes.
Cette partie est l'objet d'un article en cours de rédaction.

Chapitre 1

Description de l'algorithme

1.1 Nouvelle discrétisation en temps de l'EDSR

La manière dont nous avons mené l'analyse de l'algorithme de la partie I nécessitait un schéma implicite pour définir Y^N. Ici, cela ne sera pas nécessaire. Pour cette raison, nous redéfinissons ici (Y^N, Z^N) de cette manière :

$$Y^N_{t_k} = \mathbb{E}_{t_k}(Y^N_{t_{k+1}}) + h\mathbb{E}_{t_k} f(t_k, X^N_{t_k}, Y^N_{t_{k+1}}, Z^N_{t_k}), \tag{II.1.1}$$

$$hZ^N_{t_k} = \mathbb{E}_{t_k}(Y^N_{t_{k+1}} \Delta W_k). \tag{II.1.2}$$

avec $Y^N_{t_N} = \phi^N(P^N_{t_N})$.
La seule différence avec (I.2.2) et (I.2.3) est que l'on remplace, dans le driver f, $Y^N_{t_k}$ par $Y^N_{t_{k+1}}$ produisant ainsi un schéma explicite. Par contre, on conserve $Z^N_{t_k}$ dans le driver f et non pas $Z^N_{t_{k+1}}$: **ce terme est en effet fondamental pour rendre toutes les estimations non explosives par rapport à N**. En outre, cette présence assure dans un cadre très général la validité de la remarque I.1.
Il est alors facile de montrer que la conclusion du théorème I.1 reste vraie avec ce schéma.
Pour des raisons qui apparaitront plus loin, nous allons à présent localiser l'EDSR discrète (II.1.1-II.1.2). Pour cela, on introduit le vecteur $R^T = (R_i)_{1 \leq i \leq d'}$. Ses coefficients vont correspondre à des seuils de troncature sur la condition terminale ϕ^N et sur le driver f (soft-thresholding en statistique). Ces troncatures vont permettre d'assurer que **la solution de l'équation rétrograde localisée est bornée**.
On introduit en particulier :

$$\begin{aligned} f^R(t,x,y,z) &= f(t, -R_1 \vee x_1 \wedge R_1, \cdots, -R_d \vee x_d \wedge R_d, y, z) \\ \phi^{N,R}(x) &= \phi^N(-R_1 \vee x_1 \wedge R_1, \cdots, -R_{d'} \vee x_{d'} \wedge R_{d'}). \end{aligned} \tag{II.1.3}$$

On désigne par $\|\phi^{\mathbf{N},\mathbf{R}}\|_\infty$ et $\|\mathbf{f}^{\mathbf{R}}\|_\infty$ les quantités $\sup_x |\phi^{N,R}(x)|$ et $\sup_t \sup_x |f^R(t,x,0,0)|$ qui interviendront dans les bornes d'erreur qu'on obtiendra plus loin. A partir de ces

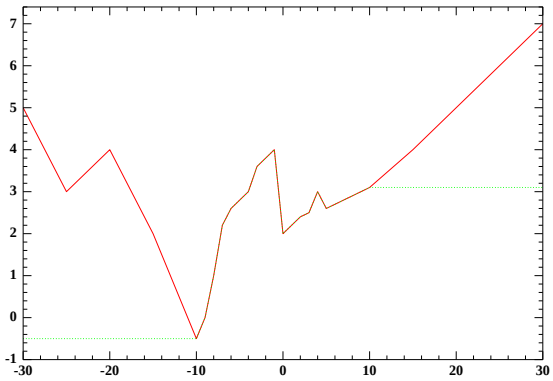


FIG. 1.1 – $\phi^{N,R}$ et ϕ^N pour $R_1 = 10$

nouveaux driver et condition terminale, on définit l'équation rétrograde discrète suivante :

$$Y^{N,R}_{t_k} = \mathbb{E}_{t_k}(Y^{N,R}_{t_{k+1}}) + h\mathbb{E}_{t_k} f^R(t_k, X^N_{t_k}, Y^{N,R}_{t_{k+1}}, Z^{N,R}_{t_k}), \tag{II.1.4}$$

$$hZ^{N,R}_{t_k} = \mathbb{E}_{t_k}(Y^{N,R}_{t_{k+1}} \Delta W_k), \tag{II.1.5}$$

avec $Y^{N,R}_{t_N} = \phi^{N,R}(P^N_{t_N})$.

1.2 Description et notations

Voyons à présent comment approcher $(Y^{N,R}, Z^{N,R})$. Ceci va se faire en évaluant des fonctions $y^{N,R,M}_k(\cdot)$, $(z^{N,R,M}_{l,k}(\cdot))_{1\leq l\leq q}$ et en approchant $Y^{N,R}_{t_k}$, $(Z^{N,R}_{l,t_k})_{1\leq l\leq q}$ par $y^{N,R,M}_k(P^N_{t_k})$, $(z^{N,R,M}_{l,k}(P^N_{t_k}))_{1\leq l\leq q}$.
Comme dans la partie I, les fonctions $y^{N,R,M}_k(\cdot)$, $(z^{N,R,M}_{l,k}(\cdot))_{1\leq l\leq q}$ sont définies à partir de leurs coefficients $(\alpha^M_{l,k})_{0\leq l\leq q}$ sur des bases de fonctions $(p_{l,k})_{0\leq l\leq q}$.
De la même manière que dans la partie I, on simule M trajectoires de $(P^N, \Delta W)$ aux instants de discrétisation $(t_k)_{0\leq k\leq N}$. En fait, pour être plus précis, on va plutôt supposer qu'on simule M trajectoires de $(P^N_{t_k}, \Delta W_{k,k+1})_k$ où $\Delta W_{k,k+1}$ est un vecteur aléatoire dont les premières composantes sont constituées des composantes de l'incrément brownien ΔW_k. En effet si $P^N = X^N$, il est suffisant de connaitre, à l'instant t_k, ΔW_k pour simuler $P^N_{t_{k+1}}$. Mais si par exemple $(P^N_{t_k})^* = \left((X^N_{t_k})^*, \max_{0\leq i\leq k} X^{N,1}_{t_i}\right)$ en notant $X^{N,1}$ la première composante de X^N, on peut avoir recours, pour simuler P^N de manière plus efficiente, à d'autres variables aléatoires (comme par exemple, la simulation du sup d'un pont brownien sur l'intervalle de temps $[t_k, t_{k+1}]$). Ainsi, pour tenir compte de ces variables aléatoires additionnelles, on suppose qu'on simule à chaque instant t_k M réalisations indépendantes

d'un vecteur aléatoire $(\Delta W^m_{k,k+1})_{1\leq m\leq M}$ (vecteur qui contient au minimum l'incrément brownien ΔW^m_k).
L'algorithme que l'on analyse dans cette partie nécessite la simulation de davantage de variables aléatoires que dans la partie I (voir cependant le dernier chapitre de cette partie pour une relaxation de cette hypothèse). En particulier, on simule **deux fois plus de variables aléatoires** à chaque instant t_k. Ainsi, on suppose qu'à chaque instant t_k, on simule $\mathbf{M}$ **réalisations** $\mathbf{(\Delta\tilde{W}^m_{k,k+1})_{1\leq m\leq M}}$**, de même loi que** $\mathbf{(\Delta W^m_{k,k+1})_{1\leq m\leq M}}$ **(avec indépendance des variables** $\mathbf{(\Delta\tilde{W}^m_{i,i+1}, \Delta W^m_{i,i+1} : 1\leq i\leq N-1, 1\leq m\leq M)}$ **)**. On va utiliser ces variables aléatoires pour construire $\mathbf{M}$ **réalisations de** $\mathbf{P^N_{t_{k+1}}}$ **notées** $\mathbf{(\tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}}$ en notant $\tilde{P}^{N,m}_{t_{k+1}} = T_k(P^{N,m}_{t_k}, \Delta\tilde{W}^m_{k,k+1})$ où T_k désigne la transition de l'instant t_k à l'instant t_{k+1} de la chaine de Markov P^N.

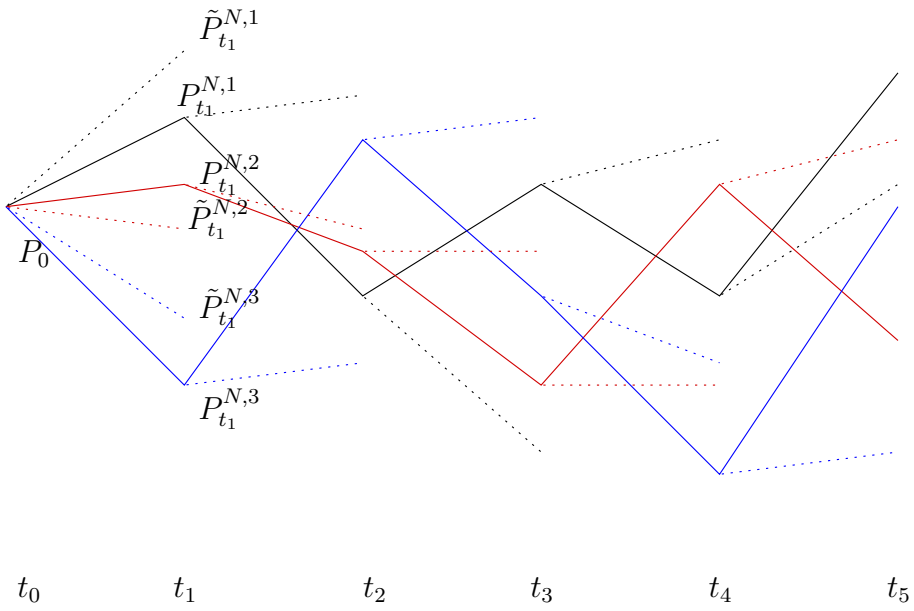


FIG. 1.2 – Simulations d'incréments browniens auxiliaires

On choisit deux réels positifs $\mathbf{C_y(R)}$ et $\mathbf{C_z(R)}$ qui sont reliés à la localisation décrite dans le chapitre précédent et qui seront explicités par la suite.
A partir d'une fonction réelle ψ, on définit les fonctions $[\psi]_y$ et $[\psi]_z$ par :

$$[\psi]_y(\cdot) = -C_y(R) \vee (\psi(\cdot) \wedge C_y(R)),$$
$$[\psi]_z(\cdot) = -C_z(R) \vee (\psi(\cdot) \wedge C_z(R)).$$

L'algorithme de cette partie se présente toujours sous une forme rétrograde :
→ On initialise pour $k = N$ par $y^{N,R,M}_N(\cdot) = \phi^{N,R}(\cdot)$.
→ A un instant de discrétisation t_k, $(\alpha^M_{l,k})_{1\leq l\leq q}$ est défini comme le minimiseur (voir annexe A.5 pour le choix du minimiseur en cas de matrice de régression dégénérée) du

problème de moindres carrés :

$$\inf_{\alpha_l} \frac{1}{M} \sum_{m=1}^{M} |y_{k+1}^{N,R,M}(\tilde{P}_{t_{k+1}}^{N,m}) \frac{\Delta \tilde{W}_{l,k}^{m}}{h} - \alpha_l . p_{l,k}^{m}|^2. \tag{II.1.6}$$

On définit alors $\mathbf{z_{l,k}^{N,R,M}}(\cdot)$ par $\mathbf{z_{l,k}^{N,R,M}}(\cdot) = [\alpha_{\mathbf{l,k}}^{\mathbf{M}} . \mathbf{p_{l,k}}]_{\mathbf{z}}(\cdot)$. Puis $\alpha_{0,k}^{M}$ est défini comme le minimiseur du problème de moindres carrés :

$$\inf_{\alpha_0} \frac{1}{M} \sum_{m=1}^{M} |y_{k+1}^{N,R,M}(\tilde{P}_{t_{k+1}}^{N,m}) + hf(t_k, X_{t_k}^{N,m}, y_{k+1}^{N,R,M}(\tilde{P}_{t_{k+1}}^{N,m}), z_{l,k}^{N,R,M}(P_{t_k}^{N,m})) - \alpha_0 . p_{0,k}^{m}|^2. \tag{II.1.7}$$

On définit alors $\mathbf{y_k^{N,R,M}}(\cdot)$ par $\mathbf{y_k^{N,R,M}}(\cdot) = [\alpha_{\mathbf{0,k}}^{\mathbf{M}} . \mathbf{p_{0,k}}]_{\mathbf{y}}(\cdot)$.

$\rightarrow$ On itère jusqu'à l'instant t_0.

Si on compare avec l'algorithme décrit dans la partie I, on peut répertorier les différences suivantes :

- A chaque instant t_k, le problème de moindres carrés (I.1.3) fait intervenir les variables aléatoires $(P_{t_k}^{N,m}, P_{t_{k+1}}^{N,m}, \Delta W_k^m)_{1 \leq m \leq M}$ alors que les problèmes de moindres carrés (II.1.6-II.1.7) font intervenir les variables aléatoires $(P_{t_k}^{N,m}, \tilde{P}_{t_{k+1}}^{N,m}, \Delta \tilde{W}_k^m)_{1 \leq m \leq M}$.
- L'algorithme de cette partie ne nécessite pas d'itérations de Picard car on utilise dans l'équation (II.1.4) $Y_{t_{k+1}}^{N,R}$ dans le driver f.
- Dans l'algorithme de la partie I, on trouvait en même temps tous les coefficients $(\alpha_{l,k}^{i,I,M})_{0 \leq l \leq q}$ en résolvant un seul problème de moindres carrés (I.1.3). Pour des problèmes de quasi-colinéarité de colonnes de la matrice de régression et pour des problèmes de complexité de la résolution d'un problème de moindres carrés, résoudre $q+1$ petits problèmes de moindres carrés tels que (II.1.6-II.1.7) apparait plus facile que de résoudre un seul gros problème tel que (I.1.3). Enfin, notons que les matrices de régression de (II.1.6) pour des l différents seront bien souvent les mêmes ou quasiment les mêmes pour les choix de bases de fonctions que nous considérons dans la partie numérique. Ainsi, l'algorithme de cette partie semble plus agréable à mettre en place numériquement. Cependant, l'algorithme de la partie I semble donner des résultats avec des variances plus petites (voir partie numérique).
- Dans les deux algorithmes, on utilise des troncatures. La raison est assez différente dans les deux cas. Dans la partie I, les troncatures ont pour but de nous procurer des contrôles sur l'évènement $[\mathbf{A}_k^M]^c$ pour lequel les matrices de régression sont éloignées de leur limite lorsque $M \rightarrow \infty$. Dans l'algorithme de cette partie, la troncature (ou plutôt la localisation) intervient pour assurer que la solution de l'EDSR discrète définie par (II.1.4-II.1.5) est bornée et pour assurer qu'une certaine classe de fonctions est bornée. Ce dernier point sera expliqué plus loin.

– Enfin, les deux algorithmes correspondent à deux approches différentes du problème. La minimisation globale (I.1.3) cherche à approcher la solution du même problème d'optimisation (I.3.1) en remplaçant l'espérance par une moyenne empirique sur M simulations de Monte-Carlo. En ce sens, il est assez proche de problèmes d'optimisation stochastique (voir entre autres [DW88]) bien que dans notre cas, le problème soit particulièrement simple du fait du critère moindres carrés. L'algorithme de cette partie tire plutôt avantage de l'expression de (Y^N, Z^N) sous forme d'espérances conditionnelles et utilise des techniques classiques d'approximations d'espérances conditionnelles (voir par exemple [GKKW02]).

1.3 Localisation de l'équation rétrograde

Le couple $(Y^N_{t_k}, Z^N_{t_k})$ défini par (II.1.1) et (II.1.2) n'est pas borné. L'intérêt d'avoir localisé l'équation rétrograde est que le couple $(Y^{N,R}_{t_k}, Z^{N,R}_{t_k})$ défini par (II.1.4) et (II.1.5) est lui borné. C'est ce que nous allons voir ci-dessous.

1.3.1 Bornitude induite par la localisation

Vérifions donc que $(Y^{N,R}, Z^{N,R})$ est borné. Pour cela on va démontrer la proposition suivante :

Proposition II.1 *Pour $h \leq 1$, $\forall k : 0 \leq k \leq N$, $\forall l : 1 \leq l \leq q$, $|Y^{N,R}_{t_k}| \leq C_y(R)$ et $|Z^{N,R}_{l,t_k}| \leq \frac{C_y(R)}{\sqrt{h}} = C_z(R)$ p.s. où :*

$$C_y(R) = \exp(\{2\gamma^* + \frac{1+\gamma^*}{q}\}T)\{\|\phi^{N,R}\|^2_\infty + 2T(1+\frac{1}{\gamma^*})\|f^R\|^2_\infty\} \tag{II.1.8}$$

avec $\gamma^* = 4qC_f^2$.

PREUVE: Notre but est d'appliquer le Lemme de Gronwall discret A.3 et donc de relier $|Y^{N,R}_{t_k}|$ à $|Y^{N,R}_{t_{k+1}}|$.
Pour cela, appliquons l'inégalité de Young (A.5) à l'équation (II.1.4). Il vient, pour tout $\gamma > 0$:

$$|Y^{N,R}_{t_k}|^2 \leq (1+\gamma h)|\mathbb{E}_{t_k}(Y^{N,R}_{t_{k+1}})|^2 + 2(h^2+\frac{h}{\gamma})|f^R(t_k, X^N_{t_k}, 0, 0)|^2 + 2C_f^2(h^2+\frac{h}{\gamma})\mathbb{E}_{t_k}\{|Y^{N,R}_{t_{k+1}}| + |Z^{N,R}_{t_k}|\}^2,$$

soit :

$$|Y_{t_k}^{N,R}|^2 \leq (1+\gamma h)|\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,R})|^2 + 2(h^2+\frac{h}{\gamma})|f^R(t_k, X_{t_k}^N, 0, 0)|^2 + 4C_f^2(h^2+\frac{h}{\gamma})\mathbb{E}_{t_k}|Y_{t_{k+1}}^{N,R}|^2 \tag{II.1.9}$$
$$+ 4C_f^2(h^2+\frac{h}{\gamma})\sum_{l=1}^{q}|Z_{l,t_k}^{N,R}|^2.$$

En élevant au carré l'équation (II.1.5) et en appliquant l'inégalité de Cauchy-Schwarz, on obtient l'inégalité :

$$h|Z_{l,t_k}^{N,R}|^2 \leq \{\mathbb{E}_{t_k}(|Y_{t_{k+1}}^{N,R}|^2) - |\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,R})|^2\}. \tag{II.1.10}$$

En injectant cette inégalité dans (II.1.9), on obtient :

$$|Y_{t_k}^{N,R}|^2 \leq (1+\gamma h)|\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,R})|^2 + 2(h^2+\frac{h}{\gamma})|f^R(t_k, X_{t_k}^N, 0, 0)|^2 + 4C_f^2(h^2+\frac{h}{\gamma})\mathbb{E}_{t_k}|Y_{t_{k+1}}^{N,R}|^2$$
$$+ 4C_f^2 q(h+\frac{1}{\gamma})\{\mathbb{E}_{t_k}(|Y_{t_{k+1}}^{N,R}|^2) - |\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,R})|^2\}. \tag{II.1.11}$$

Ainsi, en choisissant $\gamma = \gamma^* = 4qC_f^2$ dans (II.1.11), il vient :

$$|Y_{t_k}^{N,R}|^2 \leq (1+\{2\gamma^*+\frac{1}{q}\}h+\frac{\gamma^*}{q}h^2)\mathbb{E}_{t_k}(|Y_{t_{k+1}}^{N,R}|^2) + 2(h^2+\frac{h}{\gamma^*})|f^R(t_k, X_{t_k}^N, 0, 0)|^2.$$

Pour $h \leq 1$, on obtient donc :

$$|Y_{t_k}^{N,R}|^2 \leq (1+\{2\gamma^*+\frac{1+\gamma^*}{q}\}h)\mathbb{E}_{t_k}(|Y_{t_{k+1}}^{N,R}|^2) + 2h(1+\frac{1}{\gamma^*})|f^R(t_k, X_{t_k}^N, 0, 0)|^2.$$

On peut à présent appliquer le lemme de Gronwall discret A.3 pour obtenir :

$$|Y_{t_k}^{N,R}|^2 \leq \exp(\{2\gamma^*+\frac{1+\gamma^*}{q}\}T)\{\mathbb{E}_{t_k}(|\phi^{N,R}(P_{t_N}^N)|^2)+2h(1+\frac{1}{\gamma^*})\mathbb{E}_{t_k}\sum_{i=k}^{N-1}|f^R(t_i, X_{t_i}^N, 0, 0)|^2\}.$$

On utilise à présent que $|\phi^{N,R}(x)| \leq \|\phi^{N,R}\|_\infty$ et que $|f^R(t_i, x, 0, 0)| \leq \|f^R\|_\infty$ pour avoir finalement :

$$|Y_{t_k}^{N,R}|^2 \leq \exp(\{2\gamma^*+\frac{1+\gamma^*}{q}\}T)\{\|\phi^{N,R}\|_\infty^2 + 2T(1+\frac{1}{\gamma^*})\|f^R\|_\infty^2\}.$$

On déduit alors facilement de (II.1.10) que $|Z_{l,t_k}^{N,R}| \leq \frac{C_y(R)}{\sqrt{h}} = C_z(R)$.

□

1.3.2 Erreur induite par la localisation

Voyons à présent quelle est l'erreur induite par cette localisation. On va établir la proposition suivante :

Proposition II.2 *Pour h suffisamment petit*

$$\max_{0\leq k\leq N}\mathbb{E}|Y_{t_k}^{N,R}-Y_{t_k}^{N}|^2+h\mathbb{E}\sum_{k=0}^{N-1}|Z_{t_k}^{N}-Z_{t_k}^{N,R}|^2$$
$$\leq C\mathbb{E}(|\phi^N(P_{t_N}^N)-\phi^{N,R}(P_{t_N}^N)|^2)+Ch\mathbb{E}\sum_{k=0}^{N-1}|f(t_k,X_{t_k}^N,Y_{t_{k+1}}^N,Z_{t_k}^N)-f^R(t_k,X_{t_k}^N,Y_{t_{k+1}}^N,Z_{t_k}^N)|^2.$$

PREUVE: En appliquant l'inégalité de Young (A.5) à la différence des équations (II.1.1) et (II.1.4), il vient :

$$\begin{aligned}|Y_{t_k}^{N,R}-Y_{t_k}^{N}|^2\leq&(1+\gamma h)|\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,R}-Y_{t_{k+1}}^{N})|^2+Ch^2(1+\frac{1}{\gamma h})\\&\Big(\mathbb{E}_{t_k}|f(t_k,X_{t_k}^N,Y_{t_{k+1}}^N,Z_{t_k}^N)-f^R(t_k,X_{t_k}^N,Y_{t_{k+1}}^N,Z_{t_k}^N)|^2\\&+\mathbb{E}_{t_k}|Y_{t_{k+1}}^N-Y_{t_{k+1}}^{N,R}|^2+|Z_{t_k}^N-Z_{t_k}^{N,R}|^2\Big).\end{aligned}\tag{II.1.12}$$

De la même façon que pour établir (II.1.10), on obtient pour la différence des équations (II.1.2) et (II.1.5) :

$$|Z_{t_k}^N-Z_{t_k}^{N,R}|^2\leq\frac{C}{h}\{\mathbb{E}_{t_k}(|Y_{t_{k+1}}^N-Y_{t_{k+1}}^{N,R}|^2)-|\mathbb{E}_{t_k}(Y_{t_{k+1}}^N-Y_{t_{k+1}}^{N,R})|^2\}.\tag{II.1.13}$$

Ainsi, en ajustant γ, il vient :

$$\begin{aligned}|Y_{t_k}^{N,R}-Y_{t_k}^{N}|^2\leq&(1+Ch)\mathbb{E}_{t_k}(|Y_{t_{k+1}}^N-Y_{t_{k+1}}^{N,R}|^2)\\&+Ch\mathbb{E}_{t_k}|f(t_k,X_{t_k}^N,Y_{t_{k+1}}^N,Z_{t_k}^N)-f^R(t_k,X_{t_k}^N,Y_{t_{k+1}}^N,Z_{t_k}^N)|^2.\end{aligned}$$

En appliquant le Lemme de Gronwall discret A.3, il vient alors :

$$\begin{aligned}\mathbb{E}|Y_{t_k}^{N,R}-Y_{t_k}^{N}|^2\leq&C\mathbb{E}(|\phi^N(P_{t_N}^N)-\phi^{N,R}(P_{t_N}^N)|^2)\\&+Ch\mathbb{E}\sum_{i=k}^{N-1}|f(t_i,X_{t_i}^N,Y_{t_{i+1}}^N,Z_{t_i}^N)-f^R(t_i,X_{t_i}^N,Y_{t_{i+1}}^N,Z_{t_i}^N)|^2.\end{aligned}$$

L'inégalité pour $h\mathbb{E}\sum_{k=0}^{N-1}|Z_{t_k}^N-Z_{t_k}^{N,R}|^2$ se déduit de la même façon que pour le théorème I.1. En effet, vu (II.1.13), on obtient :

$$
\begin{aligned}
&h\mathbb{E}\sum_{k=0}^{N-1}|Z_{t_k}^N - Z_{t_k}^{N,R}|^2 \\
\leq& C\sum_{k=0}^{N-1}\left[\mathbb{E}|Y_{t_{k+1}}^N - Y_{t_{k+1}}^{N,R}|^2 - \mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}}^N - Y_{t_{k+1}}^{N,R})|^2)\right] \\
\leq& C\mathbb{E}|\phi^N(P_{t_N}^N) - \phi^{N,R}(P_{t_N}^N)|^2 + C\sum_{k=0}^{N-1}\left[\mathbb{E}|Y_{t_k}^N - Y_{t_k}^{N,R}|^2 - \mathbb{E}(|\mathbb{E}_{t_k}(Y_{t_{k+1}}^N - Y_{t_{k+1}}^{N,R})|^2)\right] \\
\leq& C\mathbb{E}|\phi^N(P_{t_N}^N) - \phi^{N,R}(P_{t_N}^N)|^2 + C\{\gamma h + Ch^2(1+\frac{1}{\gamma h})\}\sum_{k=0}^{N-1}\mathbb{E}|Y_{t_{k+1}}^N - Y_{t_{k+1}}^{N,R}|^2 \\
&+ Ch^2(1+\frac{1}{\gamma h})\mathbb{E}\sum_{k=0}^{N-1}|f(t_k, X_{t_k}^N, Y_{t_{k+1}}^N, Z_{t_k}^N) - f^R(t_k, X_{t_k}^N, Y_{t_{k+1}}^N, Z_{t_k}^N)|^2 \\
&+ Ch^2(1+\frac{1}{\gamma h})\mathbb{E}\sum_{k=0}^{N-1}|Z_{t_k}^N - Z_{t_k}^{N,R}|^2
\end{aligned}
$$

en utilisant (II.1.12) pour la dernière inégalité. En ajustant à présent $\gamma = 2C$ et pour h suffisamment petit, l'inégalité de la proposition se déduit facilement.

□

L'intérêt est que l'on peut à présent facilement borner le majorant de la proposition ci-dessus en utilisant des majorations de queues de distribution de X^N et P^N. Ainsi, en résumé, on voit qu'il est possible de borner la solution de l'équation rétrograde en localisant ϕ^N et f. Remarquons alors que de la même manière que dans la partie I, $Y_{t_k}^{N,R} = y_k^{N,R}(P_{t_k}^N)$ et $Z_{l,t_k}^{N,R} = z_{l,k}^{N,R}(P_{t_k}^N)$ pour des fonctions mesurables $y_k^{N,R}(\cdot)$ et $z_{l,k}^{N,R}(\cdot)$, $1 \leq l \leq q$. L'intérêt d'utiliser le soft-thresholding est qu'il permet de conserver des fonctions $y_k^{N,R}(\cdot)$ et $z_{l,k}^{N,R}(\cdot)$ qui sont Lipschitz uniformément en R car les fonctions $\phi^{N,R}$ et f^R sont Lipschitz uniformément en R : en effet, la démonstration de la proposition I.2 page 49 s'applique à l'identique avec des constantes de Lipschitz pour $y_k^{N,R}(\cdot)$ et $\sqrt{h}z_{l,k}^{N,R}(\cdot)$ indépendantes de R.
On peut alors tirer parti de cette propriété pour fixer le type et le nombre de fonctions de base à utiliser.

Chapitre 2

Analyse de l'algorithme

Nous allons dans ce chapitre nous intéresser à l'erreur

$$\max_{0\leq k\leq N}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R,M}(P_{t_k}^{N,m})-y_k^{N,R}(P_{t_k}^{N,m})|^2 \tag{II.2.1}$$

dont nous déduirons une borne pour l'erreur

$$\max_{0\leq k\leq N}\mathbb{E}\int|y_k^{N,R}(x)-y_k^{N,R,M}(x)|^2\mu_k(dx) \tag{II.2.2}$$

où $\mu_k(dx)$ désigne la loi de $P_{t_k}^N$. Puis, nous déduirons des bornes pour les erreurs correspondantes sur Z :

$$h\mathbb{E}\sum_{k=0}^{N-1}\sum_{l=1}^{q}\frac{1}{M}\sum_{m=1}^{M}|z_{l,k}^{N,R}(P_{t_k}^{N,m})-z_{l,k}^{N,R,M}(P_{t_k}^{N,m})|^2,$$
$$h\mathbb{E}\sum_{k=0}^{N-1}\sum_{l=1}^{q}\int|z_{l,k}^{N,R}(\cdot)-z_{l,k}^{N,R,M}(\cdot)|^2\mu_k(dx).$$

Nous allons commencer par introduire un certain nombre de notations qui nous serviront pour l'analyse de la propagation de l'erreur (II.2.1).

2.1 Notations propres à l'analyse de l'algorithme

2.1.1 Espérances conditionnelles

On note $\tilde{\mathcal{F}}_{\mathbf{k+1}}$ la tribu engendrée par les variables aléatoires $(\Delta W_{i,i+1}^m)_{0\leq i\leq N-1,1\leq m\leq M}$ et $(\Delta\tilde{W}_{i,i+1}^m)_{i\geq k+1,1\leq m\leq M}$. On note $\tilde{\mathbb{E}}_{\mathbf{k+1}}=\mathbb{E}(\cdot|\tilde{\mathcal{F}}_{\mathbf{k+1}})$.

2.1.2 Coefficients de projection

On note par souci d'alléger les notations, pour $0 \leq l \leq q$ et $0 \leq k \leq N-1$:

$$\begin{aligned} p_{l,k}^m =& p_{l,k}(P_{t_k}^{N,m}), \\ \tilde{p}_{l,k+1}^m =& p_{l,k+1}(\tilde{P}_{t_{k+1}}^{N,m}), \\ \mathbf{P}_{l,k}^M =& \frac{1}{M}\sum_{m=1}^M p_{l,k}^m (p_{l,k}^m)^*. \end{aligned}$$

On peut alors définir un certain nombre de coefficients de projection à l'instant t_k :

$$\mathbf{P}_{0,k}^M \alpha_{0,k}^M = \frac{1}{M}\sum_{m=1}^M p_{0,k}^m \left([\alpha_{0,k+1}^M . \tilde{p}_{0,k+1}^m]_y + h f^R(t_k, X_{t_k}^{N,m}, [\alpha_{0,k+1}^M . \tilde{p}_{0,k+1}^m]_y, [\alpha_{l,k}^M . p_{l,k}^m]_z) \right), \tag{II.2.3}$$

$$\mathbf{P}_{0,k}^M \alpha_{0,k}^{1,M} = \frac{1}{M}\sum_{m=1}^M p_{0,k}^m [\alpha_{0,k+1}^M . \tilde{p}_{0,k+1}^m]_y, \tag{II.2.4}$$

$$\mathbf{P}_{0,k}^M \alpha_{0,k}^{2,M} = \frac{h}{M}\sum_{m=1}^M p_{0,k}^m f^R(t_k, X_{t_k}^{N,m}, [\alpha_{0,k+1}^M . \tilde{p}_{0,k+1}^m]_y, [\alpha_{l,k}^M . p_{l,k}^m]_z), \tag{II.2.5}$$

$$h\mathbf{P}_{l,k}^M \alpha_{l,k}^M = \frac{1}{M}\sum_{m=1}^M p_{l,k}^m \Delta\tilde{W}_{l,k}^m [\alpha_{0,k+1}^M . \tilde{p}_{0,k+1}^m]_y, \tag{II.2.6}$$

$$\mathbf{P}_{0,k}^M \beta_{0,k}^M = \frac{1}{M}\sum_{m=1}^M p_{0,k}^m \{ y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) + h f^R(t_k, X_{t_k}^{N,m}, y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}), z_{l,k}^{N,R}(P_{t_k}^{N,m})) \}, \tag{II.2.7}$$

$$\mathbf{P}_{0,k}^M \beta_{0,k}^{1,M} = \frac{1}{M}\sum_{m=1}^M p_{0,k}^m y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}), \tag{II.2.8}$$

$$\mathbf{P}_{0,k}^M \beta_{0,k}^{2,M} = \frac{h}{M}\sum_{m=1}^M p_{0,k}^m f^R(t_k, X_{t_k}^{N,m}, y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}), z_{l,k}^{N,R}(P_{t_k}^{N,m})), \tag{II.2.9}$$

$$h\mathbf{P}_{l,k}^M \beta_{l,k}^M = \frac{1}{M}\sum_{m=1}^M p_{l,k}^m \Delta\tilde{W}_{l,k}^m y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}). \tag{II.2.10}$$

Si les équations ci-dessus ne définissent pas les coefficients de projection de manière unique (i.e. si les matrices $\mathbf{P}_{l,k}^M$ sont dégénérées) on choisit de définir les coefficients de projection de la manière expliquée dans l'annexe A.5. On note également pour un vecteur de coefficients de projection α ou β, $\tilde{\mathbb{E}}_{\mathbf{k+1}}(\alpha) = \overline{\alpha}$ et $\tilde{\mathbb{E}}_{\mathbf{k+1}}(\beta) = \overline{\beta}$.
Remarquons que ci-dessus, on adopte la notation $f^R(t,x,y,z_l) = f^R(t,x,y,(z_l)_{1\leq l\leq q})$ de manière à éviter de répéter les q composantes de z. Notons que dans la définition de $\alpha_{l,N-1}^M$, on pose $\alpha_{0,N}.p_{0,N}(\cdot) = \phi^{N,R}(\cdot)$.

2.1.3 Normes

Pour une fonction ψ on définit :

$$\|\psi\|^2_{k,M} = \frac{1}{M}\sum_{m=1}^{M} |\psi(P^{N,m}_{t_k})|^2,$$
$$\|\psi\|^2_{k,\tilde{M}} = \frac{1}{M}\sum_{m=1}^{M} |\psi(\tilde{P}^{N,m}_{t_k})|^2.$$

2.1.4 Fonctions de base

On désigne par $\mathbf{K_{l,k}}$ le cardinal de $p_{l,k}$,$0 \leq l \leq q$. On note $\mathcal{P}_{\mathbf{l,k}}$ l'espace vectoriel de fonctions engendré par la base $p_{l,k}$ et $[\mathcal{P}_{\mathbf{l,k}}]_{\mathbf{y}}$ le même espace de fonctions mais tronquées au seuil $C_y(R)$, i.e. $[\mathcal{P}_{l,k}]_y = \{[\alpha.p_{l,k}(\cdot)]_y, \alpha \in \mathbb{R}^{K_{l,k}}\}$. On note par $\mathbf{K_{l,k}}(\omega)$ $(0 \leq k \leq N-1, 0 \leq l \leq q)$ le rang de la matrice de régression dont les lignes sont $(p_{l,k}(P^{N,m}_{t_k})^*)_{1\leq m\leq M}$.

2.1.5 Evènements

Soit un paramètre réel β tel que $1 < \beta \leq 2$. Ce paramètre réglera la vitesse de convergence en fonction du pas de temps h de l'algorithme. On reviendra plus tard sur l'interprétation de β mais disons dès maintenant que $\beta > 1$ est nécessaire pour que l'algorithme converge (vu les bornes d'erreurs que nous allons obtenir) et que $\beta = 2$ permet de rendre l'erreur de l'algorithme du même ordre de grandeur que l'erreur de discrétisation en temps du théorème I.1.
Une fois β fixé, nous aurons besoin de l'évènement $\mathbf{A^M_k}$ défini par :

$$A^M_k = \{\forall \psi \in [\mathcal{P}_{0,k+1}]_y - y^{N,R}_{k+1}(\cdot) : \|\psi\|_{k+1,\tilde{M}} - \|\psi\|_{k+1,M} \leq h^{\frac{\beta+1}{2}}\}. \tag{II.2.11}$$

2.1.6 Erreurs

Enfin, introduisons les termes d'erreur suivants, qui vont intervenir dans l'analyse de la propagation d'erreur de l'algorithme, et pour lequels on montrera des contrôles en fonction de M, h et du nombre de fonctions de base :

$$T_{1,k,M} = \|y^{N,R}_k(\cdot) - \overline{\beta}^M_{0,k}.p_{0,k}(\cdot)\|^2_{k,M}, \tag{II.2.12}$$
$$T_{2,k,M} = \|\{\alpha^{1,M}_{0,k} - \overline{\alpha}^{1,M}_{0,k}\}.p_{0,k}(\cdot)\|^2_{k,M}, \tag{II.2.13}$$
$$T_{3,k,M} = \|\{\beta^{2,M}_{0,k} - \overline{\beta}^{2,M}_{0,k}\}.p_{0,k}(\cdot)\|^2_{k,M}, \tag{II.2.14}$$
$$T_{4,k,l,M} = \|\overline{\beta}^M_{l,k}.p_{l,k}(\cdot) - z^{N,R}_{l,k}(\cdot)\|^2_{k,M}, \tag{II.2.15}$$
$$T_{5,k,l,M} = \|\{\overline{\alpha}^M_{l,k} - \alpha^M_{l,k}\}.p_{l,k}(\cdot)\|^2_{k,M}. \tag{II.2.16}$$

2.2 Propagation de l'erreur

Pour analyser l'erreur (II.2.1), on va appliquer un lemme de Gronwall discret. Il nous sera utile de disposer de la proposition suivante qui relie l'erreur à l'instant t_k à une erreur équivalente à l'instant t_{k+1}.

Proposition II.3

$$\begin{aligned}\mathbb{E}\|y_k^{N,R}(\cdot) - [\alpha_{0,k}^M.p_{0,k}]_y(\cdot)\|_{k,M}^2 \leq & \mathbb{E}T_{1,k,M} + (1+Ch)\mathbb{E}(T_{2,k,M}) \\ & + (1+Ch)\mathbb{E}\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^M.p_{0,k+1}]_y(\cdot)\|_{k+1,\tilde{M}}^2 \\ & + \frac{C}{h}\mathbb{E}T_{3,k,M} + Ch\mathbb{E}\sum_{l=1}^{q}\{T_{4,k,l,M} + T_{5,k,l,M}\}.\end{aligned}$$

PREUVE: Tout d'abord, puisque $[y_k^{N,R}]_y = y_k^{N,R}$ (proposition II.1), on obtient, en utilisant que $[\cdot]_y$ est 1-Lipschitz

$$\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m}) - [\alpha_{0,k}^M.p_{0,k}^m]_y|^2 \leq \frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m}) - \alpha_{0,k}^M.p_{0,k}^m|^2.$$

Puis, comme (voir (II.1.4))

$$\tilde{\mathbb{E}}_{k+1}\left(y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) + hf^R(t_k, X_{t_k}^{N,m}, y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}), z_{l,k}^{N,R}(P_{t_k}^{N,m}))\right) = y_k^{N,R}(P_{t_k}^{N,m}),$$

$\overline{\beta}_{0,k}^M$ (défini par (II.2.7) et $\overline{\beta}_{0,k}^M = \tilde{\mathbb{E}}_{k+1}(\beta_{0,k}^M)$) est solution du problème de moindres carrés :

$$\inf_{\alpha}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m}) - \alpha.p_{0,k}^m|^2.$$

De cette propriété d'optimalité, on tire en appliquant le théorème de Pythagore :

$$\begin{aligned}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m}) - \alpha_{0,k}^M.p_{0,k}^m|^2 = & \frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m}) - \overline{\beta}_{0,k}^M.p_{0,k}^m|^2 \\ & + \frac{1}{M}\sum_{m=1}^{M}|\{\overline{\beta}_{0,k}^M - \alpha_{0,k}^M\}.p_{0,k}^m|^2 \\ = & T_{1,k,M} + \frac{1}{M}\sum_{m=1}^{M}|\{\overline{\beta}_{0,k}^M - \alpha_{0,k}^M\}.p_{0,k}^m|^2.\end{aligned}$$

Etant donné que $\alpha_{0,k}^M = \alpha_{0,k}^{1,M} + \alpha_{0,k}^{2,M}$ (voir les équations (II.2.3-II.2.4-II.2.5)) et $\beta_{0,k}^M = \beta_{0,k}^{1,M} + \beta_{0,k}^{2,M}$ (voir (II.2.7-II.2.8-II.2.9)), on obtient, $\forall \gamma > 0$, en appliquant l'inégalité de Young (A.5)

$$\begin{aligned}\frac{1}{M}\sum_{m=1}^{M} |\{\overline{\beta}_{0,k}^M - \alpha_{0,k}^M\}.p_{0,k}^m|^2 \leq & \frac{1+\gamma h}{M}\sum_{m=1}^{M} |(\overline{\beta}_{0,k}^{1,M} - \alpha_{0,k}^{1,M}).p_{0,k}^m|^2 \\ & + (1+\frac{1}{\gamma h})\frac{1}{M}\sum_{m=1}^{M} |(\overline{\beta}_{0,k}^{2,M} - \alpha_{0,k}^{2,M}).p_{0,k}^m|^2.\end{aligned} \tag{II.2.17}$$

Traitons tout d'abord le premier terme du membre de droite de (II.2.17). Observons que $\overline{\beta}_{0,k}^{1,M}.p_{0,k}^m$ est $\tilde{\mathcal{F}}_{k+1}$-mesurable. De plus, $\tilde{\mathbb{E}}_{k+1}(\alpha_{0,k}^{1,M}.p_{0,k}^m) = \tilde{\mathbb{E}}_{k+1}(\alpha_{0,k}^{1,M}).p_{0,k}^m = \overline{\alpha}_{0,k}^{1,M}.p_{0,k}^m$. Ainsi, pour tout m, on a l'égalité :

$$\mathbb{E}|\{\overline{\beta}_{0,k}^{1,M} - \alpha_{0,k}^{1,M}\}.p_{0,k}^m|^2 = \mathbb{E}|\{\alpha_{0,k}^{1,M} - \overline{\alpha}_{0,k}^{1,M}\}.p_{0,k}^m|^2 + \mathbb{E}|(\overline{\alpha}_{0,k}^{1,M} - \overline{\beta}_{0,k}^{1,M}).p_{0,k}^m|^2.$$

En sommant sur m, il vient :

$$\begin{aligned}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M} |\{\overline{\beta}_{0,k}^{1,M} - \alpha_{0,k}^{1,M}\}.p_{0,k}^m|^2 = & \mathbb{E}\frac{1}{M}\sum_{m=1}^{M} |\{\alpha_{0,k}^{1,M} - \overline{\alpha}_{0,k}^{1,M}\}.p_{0,k}^m|^2 \\ & + \mathbb{E}\frac{1}{M}\sum_{m=1}^{M} |(\overline{\alpha}_{0,k}^{1,M} - \overline{\beta}_{0,k}^{1,M}).p_{0,k}^m|^2 \\ = & \mathbb{E}(T_{2,k,M}) + \mathbb{E}\frac{1}{M}\sum_{m=1}^{M} |(\overline{\alpha}_{0,k}^{1,M} - \overline{\beta}_{0,k}^{1,M}).p_{0,k}^m|^2.\end{aligned}$$

Ce tout dernier terme se majore en utilisant la contraction de la projection sur $\left((p_{0,k}^m)^*\right)_{1\leq m\leq M}$ de la manière suivante :

$$\frac{1}{M}\sum_{m=1}^{M} |(\overline{\alpha}_{0,k}^{1,M} - \overline{\beta}_{0,k}^{1,M}).p_{0,k}^m|^2 \leq \frac{1}{M}\sum_{m=1}^{M} |\tilde{\mathbb{E}}_{k+1}\{y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) - [\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y\}|^2.$$

Passons à présent au deuxième terme du membre de droite de (II.2.17). On a, en utilisant $(a+b)^2 \leq 2(a^2+b^2)$:

$$\begin{aligned}\frac{1}{M}\sum_{m=1}^{M} |(\overline{\beta}_{0,k}^{2,M} - \alpha_{0,k}^{2,M}).p_{0,k}^m|^2 \leq & \frac{2}{M}\sum_{m=1}^{M} |(\overline{\beta}_{0,k}^{2,M} - \beta_{0,k}^{2,M}).p_{0,k}^m|^2 \\ & + \frac{2}{M}\sum_{m=1}^{M} |(\beta_{0,k}^{2,M} - \alpha_{0,k}^{2,M}).p_{0,k}^m|^2 \\ = & 2T_{3,k,M} + \frac{2}{M}\sum_{m=1}^{M} |(\beta_{0,k}^{2,M} - \alpha_{0,k}^{2,M}).p_{0,k}^m|^2.\end{aligned}$$

On utilise alors à nouveau la contraction de la projection sur $\left((p^m_{0,k})^*\right)_{1\leq m\leq M}$ ainsi que le caractère Lipschitz de f^R (uniformément en R) pour écrire (en comparant (II.2.5) et (II.2.9)) :

$$\begin{aligned}\frac{1}{M}\sum_{m=1}^{M}|(\beta^{2,M}_{0,k}-\alpha^{2,M}_{0,k}).p^m_{0,k}|^2 \leq & \frac{Ch^2}{M}\sum_{m=1}^{M}|y^{N,R}_{k+1}(\tilde{P}^{N,m}_{t_{k+1}})-[\alpha^M_{0,k+1}.\tilde{p}^m_{0,k+1}]_y|^2\\ &+\frac{Ch^2}{M}\sum_{l=1}^{q}\sum_{m=1}^{M}|z^{N,R}_{l,k}(P^{N,m}_{t_k})-[\alpha^M_{l,k}.p^m_{l,k}]_z|^2.\end{aligned}$$

Récapitulons. Nous avons obtenu la majoration :

$$\begin{aligned}&\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y^{N,R}_k(P^{N,m}_{t_k})-y^{N,R,M}_k(P^{N,m}_{t_k})|^2\\ \leq&(1+\gamma h)\frac{1}{M}\mathbb{E}\sum_{m=1}^{M}|\tilde{\mathbb{E}}_{k+1}\{y^{N,R}_{k+1}(\tilde{P}^{N,m}_{t_{k+1}})-[\alpha^M_{0,k+1}.\tilde{p}^m_{0,k+1}]_y\}|^2+\mathbb{E}T_{1,k,M}+(1+\gamma h)\mathbb{E}(T_{2,k,M})\\ &+2(1+\frac{1}{\gamma h})\mathbb{E}T_{3,k,M}+C(h^2+\frac{h}{\gamma})\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y^{N,R}_{k+1}(\tilde{P}^{N,m}_{t_{k+1}})-[\alpha^M_{0,k+1}.\tilde{p}^m_{0,k+1}]_y|^2\\ &+C(h^2+\frac{h}{\gamma})\mathbb{E}\frac{1}{M}\sum_{l=1}^{q}\sum_{m=1}^{M}|z^{N,R}_{l,k}(P^{N,m}_{t_k})-[\alpha^M_{l,k}.p^m_{l,k}]_z|^2. \qquad \text{(II.2.18)}\end{aligned}$$

Nous allons simplifier cette dernière équation en nous intéressant au terme :

$$\mathbb{E}\frac{1}{M}\sum_{l=1}^{q}\sum_{m=1}^{M}|z^{N,R}_{l,k}(P^{N,m}_{t_k})-[\alpha^M_{l,k}.p^m_{l,k}]_z|^2.$$

Remarquons d'abord que ce terme se majore par

$$\mathbb{E}\frac{1}{M}\sum_{l=1}^{q}\sum_{m=1}^{M}|z^{N,R}_{l,k}(P^{N,m}_{t_k})-\alpha^M_{l,k}.p^m_{l,k}|^2$$

en utilisant la proposition II.1 et le fait que $[\cdot]_z$ est 1-Lipschitz. Puis, en utilisant l'inégalité $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$, il vient :

$$\begin{aligned}\frac{1}{M}\sum_{m=1}^{M}|z_{l,k}^{N,R}(P_{t_k}^{N,m})-\alpha_{l,k}^{M}.p_{l,k}^{m}|^2 \leq & \frac{3}{M}\sum_{m=1}^{M}|z_{l,k}^{N,R}(P_{t_k}^{N,m})-\overline{\beta}_{l,k}^{M}.p_{l,k}^{m}|^2 \\ & + \frac{3}{M}\sum_{m=1}^{M}|(\alpha_{l,k}^{M}-\overline{\alpha}_{l,k}^{M}).p_{l,k}^{m}|^2 \\ & + \frac{3}{M}\sum_{m=1}^{M}|(\overline{\beta}_{l,k}^{M}-\overline{\alpha}_{l,k}^{M}).p_{l,k}^{m}|^2 \\ = & 3T_{4,k,l,M}+3T_{5,k,l,M}+\frac{3}{M}\sum_{m=1}^{M}|(\overline{\beta}_{l,k}^{M}-\overline{\alpha}_{l,k}^{M}).p_{l,k}^{m}|^2.\end{aligned} \tag{II.2.19}$$

En utilisant la contraction de la projection sur $\left((p_{l,k}^{m})^*\right)_{1\leq m\leq M}$ et en comparant (II.2.6) et (II.2.10), il vient :

$$\begin{aligned}\frac{h}{M}\sum_{m=1}^{M}|(\overline{\beta}_{l,k}^{M}-\overline{\alpha}_{l,k}^{M}).p_{l,k}^{m}|^2 \leq & \frac{1}{M}\sum_{m=1}^{M}|\tilde{\mathbb{E}}_{k+1}\{(y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})-[\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y)\frac{\Delta\tilde{W}_{l,k}^{m}}{\sqrt{h}}\}|^2 \\ \leq & \frac{1}{M}\sum_{m=1}^{M}\tilde{\mathbb{E}}_{k+1}|y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})-[\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y|^2 \\ & - \frac{1}{M}\sum_{m=1}^{M}|\tilde{\mathbb{E}}_{k+1}\{y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})-[\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y\}|^2\end{aligned} \tag{II.2.20}$$

en utilisant l'inégalité de Cauchy-Schwarz pour la dernière inégalité. C'est cette dernière majoration qui va nous permettre de simplifier (II.2.18). En effet, en injectant cette majoration dans (II.2.18), on peut donc majorer

$$\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m})-y_k^{N,R,M}(P_{t_k}^{N,m})|^2$$

par :

$$\mathbb{E}T_{1,k,M} + (1+\gamma h)\mathbb{E}(T_{2,k,M}) + (1+\gamma h)\frac{1}{M}\mathbb{E}\sum_{m=1}^{M}|\tilde{\mathbb{E}}_{k+1}\{y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) - [\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y\}|^2 \tag{II.2.21}$$

$$+2(1+\frac{1}{\gamma h})\mathbb{E}T_{3,k,M} + C(h^2+\frac{h}{\gamma})\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) - [\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y|^2$$

$$+C(h^2+\frac{h}{\gamma})\mathbb{E}\sum_{l=1}^{q}\{T_{4,k,l,M}+T_{5,k,l,M}\}$$

$$+C(h+\frac{1}{\gamma})\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) - [\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y|^2$$

$$-C(h+\frac{1}{\gamma})\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\tilde{\mathbb{E}}_{k+1}\{y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) - [\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y\}|^2.$$

En choisissant à présent $\gamma = C$ dans l'équation ci-dessus, il vient, pour une constante C différente :

$$\begin{aligned}
&\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m}) - y_k^{N,R,M}(P_{t_k}^{N,m})|^2\\
\leq&(1+Ch)\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\{y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) - [\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y\}|^2\\
&+\mathbb{E}T_{1,k,M} + (1+Ch)\mathbb{E}T_{2,k,M} + \frac{C}{h}\mathbb{E}T_{3,k,M} + Ch\mathbb{E}\sum_{l=1}^{q}\{T_{4,k,l,M}+T_{5,k,l,M}\}.
\end{aligned} \tag{II.2.22}$$

□

Pour appliquer le lemme de Gronwall discret, il faudrait que dans la proposition II.3, on ait $\mathbb{E}\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,M}^2$ à la place de $\mathbb{E}\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,\tilde{M}}^2$. Le passage de l'un à l'autre se fait à l'aide de l'évènement A_k^M et l'on peut alors établir le théorème :

Théorème II.1

$$\begin{aligned}
&\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m}) - y_k^{N,R,M}(P_{t_k}^{N,m})|^2\\
\leq&C\mathbb{E}\sum_{i=k}^{N-1}\{T_{1,i,M}+T_{2,i,M}+h^{-1}T_{3,i,M}\} + Ch^{\beta-1}\\
&+C\frac{C_y(R)^2}{h}\sum_{i=k}^{N-1}\mathbb{P}([A_i^M]^c) + Ch\mathbb{E}\sum_{i=k}^{N-1}\sum_{l=1}^{q}\{T_{4,i,l,M}+T_{5,i,l,M}\}.
\end{aligned}$$

PREUVE: Comme annoncé plus haut, ce théorème est la conséquence directe du lemme de Gronwall discret si on transforme dans la proposition II.3 le terme $\mathbb{E}\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,\tilde{M}}^2$ en $\mathbb{E}\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,M}^2$. Pour cela, on écrit :

$$
\begin{aligned}
&\mathbb{E}\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,\tilde{M}}^2 \\
=&\mathbb{E}\Big(\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,\tilde{M}} - \|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,M} + \\
&\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,M} \Big)^2 \\
\leq&(1+h^{-1})\mathbb{E}\Big(\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,\tilde{M}} - \|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,M} \Big)_+^2 \\
&+ (1+h)\mathbb{E}\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,M}^2.
\end{aligned}
$$

En distinguant à présent suivant que l'on se trouve sur A_k^M ou sur $[A_k^M]^c$ et en utilisant la définition de A_k^M il vient :

$$
\begin{aligned}
&\mathbb{E}\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,\tilde{M}}^2 \\
\leq& Ch^\beta + \frac{C}{h} C_y(R)^2 \mathbb{P}([A_k^M]^c) + (1+h)\mathbb{E}\|y_{k+1}^{N,R}(\cdot) - [\alpha_{0,k+1}^{M}.p_{0,k+1}(\cdot)]_y\|_{k+1,M}^2.
\end{aligned}
$$

Alors, (II.2.22) devient :

$$
\begin{aligned}
&\mathbb{E}\frac{1}{M}\sum_{m=1}^{M} |y_k^{N,R}(P_{t_k}^{N,m}) - y_k^{N,R,M}(P_{t_k}^{N,m})|^2 \qquad \text{(II.2.23)} \\
\leq&\mathbb{E}T_{1,k,M} + (1+Ch)\mathbb{E}(T_{2,k,M}) + Ch^\beta \\
&+ (1+Ch)\mathbb{E}\frac{1}{M}\sum_{m=1}^{M} |y_{k+1}^{N,R}(P_{t_{k+1}}^{N,m}) - y_{k+1}^{N,R,M}(P_{t_{k+1}}^{N,m})|^2 \\
&+ C\frac{C_y(R)^2}{h}\mathbb{P}([A_k^M]^c) + \frac{C}{h}\mathbb{E}T_{3,k,M} + Ch\mathbb{E}\sum_{l=1}^{q}\{T_{4,k,l,M} + T_{5,k,l,M}\}.
\end{aligned}
$$

Une application du lemme de Gronwall permet alors de conclure. □

2.3 Majoration des termes d'erreurs

Le théorème II.1 fait intervenir un certain nombre de termes d'erreurs que nous allons à présent majorer en fonction de M, h et du nombre de fonctions de base.

2.3.1 Majoration de $\mathbf{T_{1,k,M}}$

Lemme II.1 *Rappelons que* $\mathbf{K_{l,k}}(\omega)$ *désigne le rang de la matrice de régression dont les lignes sont* $(p_{l,k}(P^{N,m}_{t_k})^*)_{1\leq m\leq M}$.

$$\mathbb{E}T_{1,k,M} \leq C\frac{C_y(R)^2+\|f^R\|^2_\infty}{M}\mathbb{E}(K_{0,k}(\omega)) + \inf_\alpha \mathbb{E}|\alpha.p_{0,k}(P^N_{t_k}) - y^{N,R}_k(P^N_{t_k})|^2.$$

Remarque II.1 *Dans l'énoncé du lemme ci-dessus, nous aurions pu remplacer* $\mathbb{E}(K_{0,k}(\omega))$ *par* $K_{0,k}$ *étant donné que* $K_{0,k}(\omega)\leq K_{0,k}$. *Cependant, conserver* $\mathbb{E}(K_{0,k}(\omega))$ *a un intérêt crucial car dans bien des cas,* $K_{0,k}(\omega)\ll K_{0,k}$. *Considérons par exemple la base constituée d'indicatrices d'hypercubes d'arête* δ, *que nous avons déjà évoquée dans la section 4.2 de la partie I et que nous décrirons complètement dans la partie numérique. Supposons que nous ayons choisi* δ *très petit, générant ainsi un grand nombre d'hypercubes et donc un nombre* $K_{0,k}$ *très élevé. Alors le terme* $K_{0,k}(\omega)$ *nous indique qu'à un instant donné* t_k, *seul le nombre d'hypercubes contenant des simulations* $P^{N,m}_{t_k}$ *va compter dans la borne d'erreur. En effet, les hypercubes ne contenant pas de simulations vont donner des colonnes nulles dans la matrice de régression et ne compteront pas dans l'estimation du rang* $K_{0,k}(\omega)$ *de cette matrice.*

PREUVE: En utilisant $\tilde{\mathbb{E}}_{k+1}(\beta^M_{0,k}) = \overline{\beta}^M_{0,k}$, il vient

$$\begin{aligned}
\mathbb{E}T_{1,k,M} =& \mathbb{E}\frac{1}{M}\sum_{m=1}^M |y^{N,R}_k(P^{N,m}_{t_k}) - \overline{\beta}^M_{0,k}.p^m_{0,k}|^2\\
=& \mathbb{E}\frac{1}{M}\sum_{m=1}^M |\tilde{\mathbb{E}}_{k+1}\{y^{N,R}_k(P^{N,m}_{t_k}) - \beta^M_{0,k}.p^m_{0,k}\}|^2\\
\leq& \mathbb{E}\frac{1}{M}\sum_{m=1}^M |y^{N,R}_k(P^{N,m}_{t_k}) - \beta^M_{0,k}.p^m_{0,k}|^2.
\end{aligned}$$

On est alors dans le cadre du théorème A.8 avec

$$\begin{cases} Y = y^{N,R}_{k+1}(P^N_{t_{k+1}}) + hf^R(t_k, X^N_{t_k}, y^N_{k+1}(P^N_{t_{k+1}}), z^N_{l,k}(P^N_{t_k})),\\ X = P^N_{t_k}. \end{cases}$$

Remarquons que $\mathbb{E}(Y^2|X=x)$ se majore par (en notant $X^N_{t_k}=x$ si $P^N_{t_k}=x'$) :

$$\begin{aligned}
&2C_y(R)^2 + 2h^2\mathbb{E}\left(|f^R(t_k, X^N_{t_k}, y^{N,R}_{k+1}(P^N_{t_{k+1}}), z^{N,R}_{l,k}(P^N_{t_k}))|^2\middle| P^N_{t_k}=x'\right)\\
\leq& 2C_y(R)^2 + 8C_f^2h^2C_y(R)^2 + 8qhC_f^2C_y(R)^2 + 4h^2|f^R(t_k,x,0,0)|^2\\
\leq& C(C_y(R)^2 + \|f^R\|^2_\infty).
\end{aligned}$$

□

2.3.2 Majoration de $\mathbf{T_{2,k,M}}$

Lemme II.2

$$\mathbb{E}T_{2,k,M} \leq \frac{C_y(R)^2}{M}\mathbb{E}(K_{0,k}(\omega)).$$

PREUVE: Vu la discussion faite dans l'annexe A.5, on sait qu'en notant B la matrice de M lignes $(p_{0,k}^m)^*$, on peut toujours supposer que $\frac{B^*B}{M} = Id$, en considérant que B a un nombre de colonnes égal à $K_{0,k}(\omega)$. Dans ce cas, $\alpha_{0,k}^{1,M}$ est défini par

$$\alpha_{0,k}^{1,M} = \frac{1}{M}B^*V$$

où V désigne le vecteur de $\mathbb{R}^M$ de coordonnées $V_m = [\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y$. On a alors :

$$|p_{0,k}^m.(\alpha_{0,k}^{1,M} - \overline{\alpha}_{0,k}^{1,M})|^2 = (p_{0,k}^m)^*\frac{B^*}{M}\{V - \tilde{\mathbb{E}}_{k+1}(V)\}\{V - \tilde{\mathbb{E}}_{k+1}(V)\}^*\frac{B}{M}p_{0,k}^m,$$

puis :

$$\tilde{\mathbb{E}}_{k+1}|p_{0,k}^m.(\alpha_{0,k}^{1,M} - \overline{\alpha}_{0,k}^{1,M})|^2 = (p_{0,k}^m)^*\frac{B^*}{M}\tilde{\mathbb{E}}_{k+1}\Big(\{V - \tilde{\mathbb{E}}_{k+1}(V)\}\{V - \tilde{\mathbb{E}}_{k+1}(V)\}^*\Big)\frac{B}{M}p_{0,k}^m.$$

Or le terme (m, m') de la matrice $\tilde{\mathbb{E}}_{k+1}\big(\{V - \tilde{\mathbb{E}}_{k+1}(V)\}\{V - \tilde{\mathbb{E}}_{k+1}(V)\}^*\big)$ vaut :

$$\begin{aligned}\tilde{\mathbb{E}}_{k+1}(\{V_m - \tilde{\mathbb{E}}_{k+1}(V_m)\}\{V_{m'} - \tilde{\mathbb{E}}_{k+1}(V_{m'})\}) =& \tilde{\mathbb{E}}_{k+1}([\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y[\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^{m'}]_y) \\ &- \tilde{\mathbb{E}}_{k+1}([\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y)\mathbb{E}_{k+1}^*([\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^{m'}]_y).\end{aligned}$$

Or $[\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y = \psi(\alpha_{0,k+1}^M, P_{t_k}^{N,m}, \Delta\tilde{W}_{k,k+1}^m)$ pour une certaine fonction mesurable ψ. Comme $\alpha_{0,k+1}^M$ et $P_{t_k}^{N,m}$ sont $\tilde{\mathcal{F}}_{k+1}$-mesurables et que pour $m \neq m'$, $\Delta\tilde{W}_{k,k+1}^m$ et $\Delta\tilde{W}_{k,k+1}^{m'}$ sont indépendants de $\tilde{\mathcal{F}}_{k+1}$ et indépendants entre eux, on a :

$$\tilde{\mathbb{E}}_{k+1}([\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y[\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^{m'}]_y) = \tilde{\mathbb{E}}_{k+1}([\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y)\tilde{\mathbb{E}}_{k+1}([\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^{m'}]_y)$$

et donc $\tilde{\mathbb{E}}_{k+1}(\{V_m - \tilde{\mathbb{E}}_{k+1}(V_m)\}\{V_{m'} - \tilde{\mathbb{E}}_{k+1}(V_{m'})\}) = 0$ pour $m \neq m'$. Ainsi les termes non-diagonaux de la matrice $\tilde{\mathbb{E}}_{k+1}\big(\{V - \tilde{\mathbb{E}}_{k+1}(V)\}\{V - \tilde{\mathbb{E}}_{k+1}(V)\}^*\big)$ sont nuls. **C'est vraiment dans l'annulation de ces termes que réside l'intérêt de la simulation des variables** $(\Delta\tilde{W}_{k,k+1}^m)_{1\leq m\leq M}$.
Les termes diagonaux se majorent par $\tilde{\mathbb{E}}_{k+1}([\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y^2) \leq C_y(R)^2$.

Ainsi $\|\tilde{\mathbb{E}}_{k+1}\big(\{V-\tilde{\mathbb{E}}_{k+1}(V)\}\{V-\tilde{\mathbb{E}}_{k+1}(V)\}^*\big)\|_2 \leq C_y(R)^2$ et il vient :

$$\begin{aligned}
&(p_{0,k}^m)^*\frac{B^*}{M}\tilde{\mathbb{E}}_{k+1}\big(\{V-\tilde{\mathbb{E}}_{k+1}(V)\}\{V-\tilde{\mathbb{E}}_{k+1}(V)\}^*\big)\frac{B}{M}p_{0,k}^m\\
\leq&\|\frac{Bp_{0,k}^m}{M}\|_2\|\tilde{\mathbb{E}}_{k+1}\big(\{V-\tilde{\mathbb{E}}_{k+1}(V)\}\{V-\tilde{\mathbb{E}}_{k+1}(V)\}^*\big)\frac{B}{M}p_{0,k}^m\|_2\\
\leq&\|\frac{Bp_{0,k}^m}{M}\|_2^2\|\tilde{\mathbb{E}}_{k+1}\big(\{V-\tilde{\mathbb{E}}_{k+1}(V)\}\{V-\tilde{\mathbb{E}}_{k+1}(V)\}^*\big)\|_2\\
\leq&\frac{C_y(R)^2}{M}(p_{0,k}^m)^*\frac{B^*B}{M}p_{0,k}^m\\
=&\frac{C_y(R)^2}{M}(p_{0,k}^m)^*p_{0,k}^m\\
=&\frac{C_y(R)^2}{M}\sum_{j=1}^{K_{0,k}(\omega)}|p_{0,k,j}^m|^2.
\end{aligned}$$

en utilisant l'hypothèse $\frac{B^*B}{M}=\mathrm{Id}$. En sommant sur m cette relation, il vient :

$$\tilde{\mathbb{E}}_{k+1}\frac{1}{M}\sum_{m=1}^{M}|p_{0,k}^m.(\alpha_{0,k}^{1,M}-\overline{\alpha}_{0,k}^{1,M})|^2\leq\frac{C_y(R)^2}{M}\frac{1}{M}\sum_{m=1}^{M}\sum_{j=1}^{K_{0,k}(\omega)}|p_{0,k,j}^m|^2=\frac{C_y(R)^2K_{0,k}(\omega)}{M}$$

en tenant compte de $\frac{B^*B}{M}=\mathrm{Id}$. Il en découle $\mathbb{E}(T_{2,k,M})\leq\frac{C_y(R)^2}{M}\mathbb{E}(K_{0,k}(\omega))$.

□

2.3.3 Majoration de $\mathbf{T_{3,k,M}}$

Lemme II.3

$$\mathbb{E}T_{3,k,M}\leq C\frac{(C_y(R)^2+\|f^R\|_\infty)h}{M}\mathbb{E}(K_{0,k}(\omega)).$$

PREUVE: Remarquons que

$$\beta_{0,k}^{2,M}-\overline{\beta}_{0,k}^{2,M}=h\frac{B^*}{M}\{V-\tilde{\mathbb{E}}_{k+1}(V)\}$$

avec la même définition de B qu'au lemme précédent, toujours avec l'hypothèse $\frac{B^*B}{M}=\mathrm{Id}$ et en notant V le vecteur de $\mathbb{R}^M$ de coordonnées :

$$V_m=f^R(t_k,X_{t_k}^{N,m},y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}),z_{l,k}^{N,R}(P_{t_k}^{N,m})).$$

Remarquons alors que $f^R(t_k,X_{t_k}^{N,m},y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}),z_{l,k}^{N,R}(P_{t_k}^{N,m}))=\psi(P_{t_k}^{N,m},\Delta\tilde{W}_{k,k+1}^m)$ pour une certaine fonction ψ mesurable. Les mêmes arguments qu'au lemme précédent permettent d'affirmer que $\tilde{\mathbb{E}}_{k+1}\big(\{V_m-\tilde{\mathbb{E}}_{k+1}(V_m)\}\{V_{m'}-\tilde{\mathbb{E}}_{k+1}(V_{m'})\}\big)=0$ pour $m\neq m'$.

Ainsi, comme précédemment, on a :

$$\begin{aligned}&\tilde{\mathbb{E}}_{k+1}\left((p^m_{0,k})^*\frac{B^*}{M}\big(\{V-\tilde{\mathbb{E}}_{k+1}(V)\}\{V-\tilde{\mathbb{E}}_{k+1}(V)\}^*\big)\frac{B}{M}p^m_{0,k}\right)\\ \leq&\frac{\max_m \tilde{\mathbb{E}}_{k+1}(V_m^2)}{M}\sum_{j=1}^{K_{0,k}(\omega)}|p^m_{0,k,j}|^2.\end{aligned}$$

On a alors, pour tout m :

$$\begin{aligned}\tilde{\mathbb{E}}_{k+1}(V_m^2)\leq&\tilde{\mathbb{E}}_{k+1}|f^R(t_k,X^{N,m}_{t_k},y^{N,R}_{k+1}(\tilde{P}^{N,m}_{t_{k+1}}),z^{N,R}_{l,k}(P^{N,m}_{t_k}))|^2\\ \leq&2|f^R(t_k,X^{N,m}_{t_k},0,0)|^2+4C_f^2C_y(R)^2+4C_f^2\frac{q}{h}C_y(R)^2\leq C\frac{C_y(R)^2+\|f^R\|^2_\infty}{h}.\end{aligned}$$

Ainsi on conclut comme au lemme précédent pour obtenir :

$$\mathbb{E}T_{3,k,M}\leq C\frac{(C_y(R)^2+\|f^R\|^2_\infty)h}{M}\mathbb{E}(K_{0,k}(\omega)).$$

□

2.3.4 Majoration de $\mathbf{T_{4,k,l,M}}$

Lemme II.4

$$\mathbb{E}T_{4,k,l,M}\leq\frac{C_y(R)^2}{Mh}\mathbb{E}(K_{l,k}(\omega))+\frac{1}{h}\inf_\alpha\mathbb{E}|\sqrt{h}z^{N,R}_{l,k}(P^N_{t_k})-\alpha.p_{l,k}(P^N_{t_k})|^2.$$

PREUVE: On a :

$$T_{4,k,l,M}=\frac{1}{M}\sum_{m=1}^M|z^{N,R}_{l,k}(P^{N,m}_{t_k})-\overline{\beta}^M_{l,k}.p^m_{l,k}|^2=\frac{1}{Mh}\sum_{m=1}^M|\sqrt{h}z^{N,R}_{l,k}(P^{N,m}_{t_k})-(\sqrt{h})\overline{\beta}^M_{l,k}.p^m_{l,k}|^2.$$

Or, rappelons que $\overline{\beta}^M_{l,k}=\tilde{\mathbb{E}}_{k+1}(\beta^M_{l,k})$ et que $\sqrt{h}z^{N,R}_{l,k}(P^{N,m}_{t_k})$ est $\tilde{\mathcal{F}}_{k+1}$-mesurable. Ainsi :

$$\begin{aligned}&\mathbb{E}\frac{1}{Mh}\sum_{m=1}^M|\sqrt{h}z^{N,R}_{l,k}(P^{N,m}_{t_k})-(\sqrt{h})\overline{\beta}^M_{l,k}.p^m_{l,k}|^2\\ =&\mathbb{E}\frac{1}{Mh}\sum_{m=1}^M|\tilde{\mathbb{E}}_{k+1}\{\sqrt{h}z^{N,R}_{l,k}(P^{N,m}_{t_k})-\sqrt{h}\beta^M_{l,k}.p^m_{l,k}\}|^2 \qquad \text{(II.2.24)}\\ \leq&\mathbb{E}\frac{1}{Mh}\sum_{m=1}^M|\sqrt{h}z^{N,R}_{l,k}(P^{N,m}_{t_k})-\sqrt{h}\beta^M_{l,k}.p^m_{l,k}|^2.\end{aligned}$$

On est alors dans le cadre d'application du théorème A.8 avec $Y = \frac{\Delta W_{l,k}}{\sqrt{h}} y_{k+1}^{N,R}(P_{t_{k+1}}^N)$ et $X = P_{t_k}^N$. Remarquons alors que

$$\mathbb{E}(Y^2|X) \leq C_y(R)^2.$$

Ainsi, en appliquant le théorème A.8, il vient :

$$\mathbb{E}T_{4,k,l,M} \leq \frac{C_y(R)^2}{Mh}\mathbb{E}(K_{l,k}(\omega)) + \frac{1}{h}\inf_{\alpha}\mathbb{E}|\sqrt{h}z_{l,k}^{N,R}(P_{t_k}^N) - \alpha.p_{l,k}(P_{t_k}^N)|^2.$$

□

2.3.5 Majoration de $\mathbf{T_{5,k,l,M}}$

Lemme II.5

$$\mathbb{E}T_{5,k,l,M} \leq \frac{C_y(R)^2}{Mh}\mathbb{E}(K_{l,k}(\omega)).$$

PREUVE: Rappelons que $T_{5,k,l,M} = \frac{1}{M}\sum_{m=1}^{M}|(\overline{\alpha}_{l,k}^M - \alpha_{l,k}^M).p_{l,k}^m|^2$. Désignons cette fois par B la matrice de M lignes $(p_{l,k}^m)^*$ et par V le vecteur de $\mathbb{R}^M$ de coordonnées

$$V_m = \frac{\Delta\tilde{W}_{l,k}^m}{\sqrt{h}}[\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y.$$

On a alors, avec ces notations matricielles :

$$\tilde{\mathbb{E}}_{k+1}|(\overline{\alpha}_{l,k}^M - \alpha_{l,k}^M).p_{l,k}^m|^2 = \frac{1}{h}(p_{l,k}^m)^*\frac{B^*}{M}\tilde{\mathbb{E}}_{k+1}\Big(\{V - \tilde{\mathbb{E}}_{k+1}(V)\}\{V - \tilde{\mathbb{E}}_{k+1}(V)\}^*\Big)\frac{B}{M}p_{l,k}^m.$$

Remarquons alors que $\frac{\Delta\tilde{W}_{l,k}^m}{\sqrt{h}}[\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y = \psi(\alpha_{0,k+1}^M, P_{t_k}^{N,m}, \Delta\tilde{W}_{k,k+1}^m)$. Ainsi, de la même façon que précédemment, les termes non-diagonaux de $\tilde{\mathbb{E}}_{k+1}\big(\{V-\tilde{\mathbb{E}}_{k+1}(V)\}\{V-\tilde{\mathbb{E}}_{k+1}(V)\}^*\big)$ sont nuls. Les termes diagonaux se majorent par

$$\tilde{\mathbb{E}}_{k+1}(|\frac{\Delta\tilde{W}_{l,k}^m}{\sqrt{h}}|^2[\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y^2) \leq C_y(R)^2.$$

Ainsi, de la même façon que précédemment, on a :

$$\mathbb{E}T_{5,k,l,M} \leq \frac{C_y(R)^2}{Mh}\mathbb{E}(K_{l,k}(\omega)).$$

□

2.3.6 Majoration de $\mathbb{P}([\mathbf{A}_\mathbf{k}^\mathbf{M}]^\mathbf{c})$

Lemme II.6 *Avec les notations de l'annexe A.4*

$$\begin{aligned}&\mathbb{P}([A_k^M]^c)\\ \leq& 2\mathbb{E}\left(\mathcal{N}_2\left(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}}, [\mathcal{P}_{0,k+1}]_y - y_{k+1}^{N,R}(\cdot), (P_{t_{k+1}}^{N,m})_{1\leq m\leq M}, (\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M}\right)\right) \exp\left(-\frac{Mh^{\beta+1}}{144C_y(R)^2}\right).\end{aligned}$$

PREUVE: Le résultat est classique (voir [Pol84]) et nous reprenons par exemple la démonstration donnée dans [GKKW02] (théorème 11.2). Pour simplifier les notations, on désigne par $\mathcal{H}$ la classe de fonctions $[\mathcal{P}_{0,k+1}]_y - y_{k+1}^{N,R}(\cdot)$. Introduisons une suite de variables aléatoires i.i.d. $(U_m)_{1\leq m\leq M}$ valant 1 et -1 avec probabilité $\frac{1}{2}$, indépendantes de toute autre variable aléatoire simulée dans l'algorithme. Définissons :

$$\begin{aligned}Z_m =& \tilde{P}_{t_{k+1}}^{N,m}\\ Z_{M+m} =& P_{t_{k+1}}^{N,m}\end{aligned}$$

si $U_m = 1$ et

$$\begin{aligned}Z_m =& P_{t_{k+1}}^{N,m}\\ Z_{M+m} =& \tilde{P}_{t_{k+1}}^{N,m}\end{aligned}$$

si $U_m = -1$. Alors, on peut écrire :

$$\begin{aligned}\mathbb{P}([A_k^M]^c) =& \mathbb{P}\big(\exists\psi\in\mathcal{H} : \|\psi\|_{k+1,\tilde{M}} - \|\psi\|_{k+1,M} > h^{\frac{\beta+1}{2}}\big)\\ =& \mathbb{P}\Big(\exists\psi\in\mathcal{H} : \sqrt{\frac{1}{M}\sum_{m=1}^M |\psi(\tilde{P}_{t_{k+1}}^{N,m})|^2} - \sqrt{\frac{1}{M}\sum_{m=1}^M |\psi(P_{t_{k+1}}^{N,m})|^2} > h^{\frac{\beta+1}{2}}\Big)\\ =& \mathbb{P}\Big(\exists\psi\in\mathcal{H} : \sqrt{\frac{1}{M}\sum_{m=1}^M |\psi(Z_m)|^2} - \sqrt{\frac{1}{M}\sum_{m=1}^M |\psi(Z_{M+m})|^2} > h^{\frac{\beta+1}{2}}\Big)\end{aligned}$$

car $P_{t_{k+1}}^{N,m}$ et $\tilde{P}_{t_{k+1}}^{N,m}$ ont même loi. On peut se convaincre de la dernière égalité en conditionnant par $(U_m)_m$. Introduisons alors une couverture $\mathcal{G} = \{g_1, \ldots, g_n\}$ de $\mathcal{H}$ au sens où pour tout $\psi\in\mathcal{H}$, il existe $g_j\in\mathcal{G}$ tel que

$$\sqrt{\frac{1}{2M}\sum_{m=1}^M \{|\psi(P_{t_{k+1}}^{N,m}) - g_j(P_{t_{k+1}}^{N,m})|^2 + |\psi(\tilde{P}_{t_{k+1}}^{N,m}) - g_j(\tilde{P}_{t_{k+1}}^{N,m})|^2\}} \leq \frac{1}{3\sqrt{2}}h^{\frac{\beta+1}{2}}.$$

Le cardinal de $\mathcal{G}$ est égal au nombre de couverture $\mathcal{N}_2\left(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}}, \mathcal{H}, (P_{t_{k+1}}^{N,m})_{1\leq m\leq M}, (\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M}\right)$ avec les notations introduites dans l'annexe A.4.1. Remarquons que comme $\mathcal{P}_{0,k+1}$ est un

espace vectoriel de dimension finie (donc de dimension VC finie), ce cardinal est fini (voir [LT91] ou théorème A.6). De plus, les éléments de $\mathcal{G}$ sont bornés par $2C_y(R)$, cela étant le cas pour les éléments de $\mathcal{H}$. Soit $\psi \in \mathcal{H}$ et considérons un élément $g_j \in \mathcal{G}$ qui vérifie la condition de couverture ci-dessus. On a alors, en utilisant l'inégalité triangulaire :

$$\begin{aligned}
&\sqrt{\frac{1}{M}\sum_{m=1}^{M}|\psi(Z_m)|^2} - \sqrt{\frac{1}{M}\sum_{m=1}^{M}|\psi(Z_{M+m})|^2} \\
\leq &\sqrt{\frac{1}{M}\sum_{m=1}^{M}|\psi(Z_m)-g_j(Z_m)|^2} + \sqrt{\frac{1}{M}\sum_{m=1}^{M}|g_j(Z_m)|^2} - \sqrt{\frac{1}{M}\sum_{m=1}^{M}|g_j(Z_{m+M})|^2} \\
&+ \sqrt{\frac{1}{M}\sum_{m=1}^{M}|g_j(Z_{M+m})-\psi(Z_{M+m})|^2}.
\end{aligned}$$

Le premier terme du membre de droite se majore par :

$$\sqrt{2}\sqrt{\frac{1}{2M}\sum_{m=1}^{M}\{|\psi(P^{N,m}_{t_{k+1}})-g_j(P^{N,m}_{t_{k+1}})|^2 + |\psi(\tilde{P}^{N,m}_{t_{k+1}})-g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2\}} \leq \frac{1}{3}h^{\frac{\beta+1}{2}}$$

et la même inégalité est valable pour le dernier terme du membre de droite. Ainsi :

$$\begin{aligned}
&\mathbb{P}\big(\exists\psi\in\mathcal{H} : \sqrt{\frac{1}{M}\sum_{m=1}^{M}|\psi(Z_m)|^2} - \sqrt{\frac{1}{M}\sum_{m=1}^{M}|\psi(Z_{M+m})|^2} > h^{\frac{\beta+1}{2}}\big) \\
\leq &\mathbb{P}\big(\exists g\in\mathcal{G} : \sqrt{\frac{1}{M}\sum_{m=1}^{M}|g(Z_m)|^2} - \sqrt{\frac{1}{M}\sum_{m=1}^{M}|g(Z_{M+m})|^2} > \frac{1}{3}h^{\frac{\beta+1}{2}}\big).
\end{aligned}$$

Conditionnons par $(P^{N,m}_{t_{k+1}}, \tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}$. Notons que la classe $\mathcal{G}$ dépend uniquement de ces variables aléatoires et pas des $(U_m)_m$. On obtient :

$$\begin{aligned}
&\mathbb{P}\left(\exists g\in\mathcal{G} : \sqrt{\frac{1}{M}\sum_{m=1}^{M}|g(Z_m)|^2} - \sqrt{\frac{1}{M}\sum_{m=1}^{M}|g(Z_{M+m})|^2} > \frac{1}{3}h^{\frac{\beta+1}{2}} \middle| (P^{N,m}_{t_{k+1}}, \tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}\right) \\
\leq &\mathcal{N}_2\left(\frac{1}{3\sqrt{2}}h^{\frac{\beta+1}{2}}, \mathcal{H}, (P^{N,m}_{t_{k+1}})_{1\leq m\leq M}, (\tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}\right) \\
&\max_j \mathbb{P}\left(\sqrt{\frac{1}{M}\sum_{m=1}^{M}|g_j(Z_m)|^2} - \sqrt{\frac{1}{M}\sum_{m=1}^{M}|g_j(Z_{M+m})|^2} > \frac{1}{3}h^{\frac{\beta+1}{2}} \middle| (P^{N,m}_{t_{k+1}}, \tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}\right).
\end{aligned} \tag{II.2.25}$$

De manière à appliquer l'inégalité de Hoeffding, remarquons alors que lorsque

$$A = \sqrt{\frac{1}{M}\sum_{m=1}^{M} |g_j(Z_m)|^2} - \sqrt{\frac{1}{M}\sum_{m=1}^{M} |g_j(Z_{M+m})|^2}$$

est posifif, on a :

$$\begin{aligned} A =& \frac{\frac{1}{M}\sum_{m=1}^{M}\{|g_j(Z_m)|^2 - |g_j(Z_{M+m})|^2\}}{\sqrt{\frac{1}{M}\sum_{m=1}^{M}|g_j(Z_m)|^2} + \sqrt{\frac{1}{M}\sum_{m=1}^{M}|g_j(Z_{M+m})|^2}} \\ \leq& \frac{\frac{1}{M}\sum_{m=1}^{M} U_m\{|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\}}{\sqrt{\frac{1}{M}\sum_{m=1}^{M}\{|g_j(P^{N,m}_{t_{k+1}})|^2 + |g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2\}}} \end{aligned}$$

en utilisant pour la dernière inégalité que $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$. Ainsi, en réinjectant dans (II.2.25), on obtient :

$$\begin{aligned} &\mathbb{P}\left(\sqrt{\frac{1}{M}\sum_{m=1}^{M}|g_j(Z_m)|^2} - \sqrt{\frac{1}{M}\sum_{m=1}^{M}|g_j(Z_{M+m})|^2} > \frac{1}{3}h^{\frac{\beta+1}{2}} \middle| (P^{N,m}_{t_{k+1}}, \tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}\right) \\ \leq& \mathbb{P}\left(\frac{\frac{1}{M}\sum_{m=1}^{M} U_m\{|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\}}{\sqrt{\frac{1}{M}\sum_{m=1}^{M}\{|g_j(P^{N,m}_{t_{k+1}})|^2 + |g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2\}}} > \frac{1}{3}h^{\frac{\beta+1}{2}} \middle| (P^{N,m}_{t_{k+1}}, \tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}\right). \end{aligned}$$

Notons $X_m = U_m\{|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\}$. La classe $\mathcal{G}$ ne dépendant pas de $(U_m)_m$, on a $\mathbb{E}(X_m|(P^{N,m}_{t_{k+1}}, \tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}) = 0$. En outre $|X_m| \leq \left||g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\right|$. Ainsi, on peut appliquer l'inégalité de Hoeffding (théorème A.10) et on obtient

$$\begin{aligned} &\mathbb{P}\left(\frac{\frac{1}{M}\sum_{m=1}^{M} U_m\{|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\}}{\sqrt{\frac{1}{M}\sum_{m=1}^{M}\{|g_j(P^{N,m}_{t_{k+1}})|^2 + |g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2\}}} > \frac{1}{3}h^{\frac{\beta+1}{2}} \middle| (P^{N,m}_{t_{k+1}}, \tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}\right) \\ \leq& 2\exp\left(-\frac{Mh^{\beta+1}\frac{1}{M}\sum_{m=1}^{M}\{|g_j(P^{N,m}_{t_{k+1}})|^2 + |g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2\}}{18\frac{1}{M}\sum_{m=1}^{M}\left(|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\right)^2}\right). \end{aligned} \tag{II.2.26}$$

Reste à transformer le dénominateur de la fraction dans l'exponentielle :

$$\begin{aligned} &\frac{1}{M}\sum_{m=1}^{M}\left(|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\right)^2 \\ =& \frac{1}{M}\sum_{m=1}^{M}\left(|g_j(\tilde{P}^{N,m}_{t_{k+1}})| + |g_j(P^{N,m}_{t_{k+1}})|\right)^2\left(|g_j(\tilde{P}^{N,m}_{t_{k+1}})| - |g_j(P^{N,m}_{t_{k+1}})|\right)^2 \\ \leq& \frac{8C_y(R)^2}{M}\sum_{m=1}^{M}\left(|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 + |g_j(P^{N,m}_{t_{k+1}})|^2\right). \end{aligned}$$

Ainsi, en injectant dans (II.2.26), il vient

$$\mathbb{P}\Bigg(\frac{\frac{1}{M}\sum_{m=1}^{M} U_m\{|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\}}{\sqrt{\frac{1}{M}\sum_{m=1}^{M}\{|g_j(P^{N,m}_{t_{k+1}})|^2 + |g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2\}}} > \frac{1}{3}h^{\frac{\beta+1}{2}}\Bigg|(P^{N,m}_{t_{k+1}}, \tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}\Bigg)$$
$$\leq 2\exp\Bigg(-\frac{Mh^{\beta+1}}{144C_y(R)^2}\Bigg).$$

□

Chapitre 3

Majoration de l'erreur commise dans l'algorithme

A l'aide des majorations des termes d'erreurs du chapitre précédent, on obtient finalement le théorème suivant qui borne l'erreur commise dans l'algorithme

Théorème II.2

$$\begin{aligned}
&\max_{0\le k\le N}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m})-y_k^{N,R,M}(P_{t_k}^{N,m})|^2\\
\le& C\frac{C_y(R)^2+\|f^R\|_\infty^2}{M}\sum_{k=0}^{N-1}\max_{0\le l\le q}\mathbb{E}(K_{l,k}(\omega))+C\sum_{k=0}^{N-1}\inf_{\alpha}\mathbb{E}|\alpha.p_{0,k}(P_{t_k}^N)-y_k^{N,R}(P_{t_k}^N)|^2\\
&+C\sum_{k=0}^{N-1}\sum_{l=1}^{q}\inf_{\alpha}\mathbb{E}|\alpha.p_{l,k}(P_{t_k}^N)-\sqrt{h}z_{l,k}^{N,R}(P_{t_k}^N)|^2+Ch^{\beta-1}\\
&+\frac{C}{h}C_y(R)^2e^{-\frac{Mh^{\beta+1}}{144C_y(R)^2}}\sum_{k=0}^{N-1}\mathbb{E}\left(\mathcal{N}_2\big(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}},[\mathcal{P}_{0,k+1}]_y-y_{k+1}^{N,R},(P_{t_{k+1}}^{N,m})_{1\le m\le M},(\tilde{P}_{t_{k+1}}^{N,m})_{1\le m\le M}\big)\right).
\end{aligned}$$

On vérifie donc que β doit être supérieur à 1 pour assurer la convergence de l'algorithme lorsque $h\to 0$ et que $\beta>2$ est inutile car le terme $h^{\beta-1}$ deviendrait alors négligeable devant h qui intervient dans la majoration de l'erreur de discrétisation en temps (voir théorème I.1). On reviendra plus tard dans ce chapitre sur l'estimation du nombre de couverture $\mathcal{N}_2$ qui intervient ci-dessus mais pour l'instant, on va s'efforcer de déduire du théorème II.2 des bornes pour l'erreur (II.2.2) puis pour les erreurs sur $Z^{N,R}$.

3.1 Majoration de l'erreur pour la loi de $P_{t_k}^N$

On va établir le théorème

Théorème II.3

$$
\begin{aligned}
&\max_{0\le k\le N}\mathbb{E}\int |y_k^{N,R}(x)-y_k^{N,R,M}(x)|^2\mu_k(dx)\\
\le& Ch^{\beta-1}+C\frac{C_y(R)^2+\|f^R\|_\infty^2}{M}\sum_{k=0}^{N-1}\max_{0\le l\le q}\mathbb{E}(K_{l,k}(\omega))+C\sum_{k=0}^{N-1}\inf_{\alpha}\mathbb{E}\big(|y_k^{N,R}(P_{t_k}^N)-\alpha.p_{0,k}(P_{t_k}^N)|^2\big)\\
&+C\sum_{k=0}^{N-1}\sum_{l=1}^{q}\inf_{\alpha}\mathbb{E}\big(|\alpha.p_{l,k}(P_{t_k}^N)-\sqrt{h}z_{l,k}^{N,R}(P_{t_k}^N)|^2\big)+C\frac{C_y(R)^2\max_{0\le k\le N-1}K_{0,k}}{M}\log(M)\\
&+\frac{C}{h}C_y(R)^2e^{-\frac{Mh^{\beta+1}}{144C_y(R)^2}}\sum_{k=0}^{N-1}\mathbb{E}\bigg(\mathcal{N}_2\big(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}},[\mathcal{P}_{0,k+1}]_y-y_{k+1}^{N,R},(P_{t_{k+1}}^{N,m})_{1\le m\le M},(\tilde{P}_{t_{k+1}}^{N,m})_{1\le m\le M}\big)\bigg).
\end{aligned}
$$

Avant de démontrer ce théorème, on voit qu'on obtient la même borne que pour l'erreur (II.2.1) à un terme additionnel près $C\frac{C_y(R)^2\max_{0\le k\le N-1}K_{0,k}}{M}\log(M)$.

PREUVE: Nous adoptons pour cette preuve la notation

$$\|\psi\|_k^2=\int|\psi(x)|^2\mu_k(dx)$$

pour une fonction ψ mesurable bornée. L'erreur $\mathbb{E}\int|y_k^{N,R}(x)-y_k^{N,R,M}(x)|^2\mu_k(dx)$ se réécrit alors

$$
\begin{aligned}
&\mathbb{E}\|y_k^{N,R}-y_k^{N,R,M}\|_k^2\\
=&\mathbb{E}\bigg(\|y_k^{N,R}-y_k^{N,R,M}\|_k-2\|y_k^{N,R}-y_k^{N,R,M}\|_{k,M}+2\|y_k^{N,R}-y_k^{N,R,M}\|_{k,M}\bigg)^2\\
\le&\mathbb{E}\bigg(\max\{\|y_k^{N,R}-y_k^{N,R,M}\|_k-2\|y_k^{N,R}-y_k^{N,R,M}\|_{k,M},0\}+2\|y_k^{N,R}-y_k^{N,R,M}\|_{k,M}\bigg)^2\\
\le&\mathbb{E}(T_1)+\mathbb{E}(T_2)
\end{aligned}
$$

avec $T_1=2\max\{\|y_k^{N,R}-y_k^{N,R,M}\|_k-2\|y_k^{N,R}-y_k^{N,R,M}\|_{k,M},0\}^2$ et $T_2=8\|y_k^{N,R}-y_k^{N,R,M}\|_{k,M}^2$. En appliquant le théorème II.2, il est facile de borner $\mathbb{E}(T_2)$ étant donné qu'on reconnait l'erreur (II.2.1). Pour majorer $\mathbb{E}(T_1)$, on adapte la preuve du théorème 11.3 dans [GKKW02] en écrivant :

$$
\begin{aligned}
\mathbb{P}(T_1>u)=&\mathbb{P}\bigg(2\max\{\|y_k^{N,R}-y_k^{N,R,M}\|_k-2\|y_k^{N,R}-y_k^{N,R,M}\|_{k,M},0\}^2>u\bigg)\\
\le&\mathbb{P}\bigg(\exists\psi\in[\mathcal{P}_{0,k}]_y-y_k^{N,R}:\|\psi\|_k-2\|\psi\|_{k,M}>\sqrt{\frac{u}{2}}\bigg)\\
\le&3\mathbb{E}\mathcal{N}_2(\frac{\sqrt{u}}{24},[\mathcal{P}_{0,k}]_y-y_k^{N,R},P_{t_k}^{N,1:2M})\exp(-\frac{Mu}{576(2C_y(R))^2}) \qquad \text{(II.3.1)}
\end{aligned}
$$

en utilisant le théorème A.9, la définition de $\mathcal{N}_2$ (définition A.1) et la notation $P_{t_k}^{N,1:2M} = (P_{t_k}^{N,1}, \cdots, P_{t_k}^{N,M}, P_{t_k}^{N,M+1}, \cdots, P_{t_k}^{N,2M})$ où $(P_{t_k}^{N,M+m})_{1\leq m\leq M}$ est une copie indépendante de $(P_{t_k}^{N,m})_{1\leq m\leq M}$. On a à présent :

$$\begin{aligned}\mathcal{N}_2(\frac{\sqrt{u}}{24}, [\mathcal{P}_{0,k}]_y - y_k^{N,R}, P_{t_k}^{N,1:2M}) =& \mathcal{N}_2(\frac{\sqrt{u}}{24}, [\mathcal{P}_{0,k}]_y, P_{t_k}^{N,1:2M}) \\ =& \mathcal{N}_2(\frac{\sqrt{u}}{24}, [\mathcal{P}_{0,k}]_y + C_y(R), P_{t_k}^{N,1:2M}).\end{aligned}$$

$[\mathcal{P}_{0,k}]_y + C_y(R)$ est alors une classe de fonctions positives et bornées par $2C_y(R)$. On peut donc appliquer le théorème A.6 (voir ce théorème pour la notation $V_{\mathcal{G}^+}$ pour une classe de fonctions $\mathcal{G}$) pour écrire que :

$$\begin{aligned}\mathcal{N}_2(\frac{\sqrt{u}}{24}, [\mathcal{P}_{0,k}]_y + C_y(R), P_{t_k}^{N,1:2M}) \leq& 3\left(\frac{2e(2C_y(R))^2}{\frac{u}{576}} \log\{\frac{3e(2C_y(R))^2}{\frac{u}{576}}\}\right)^{V_{([\mathcal{P}_{0,k}]_y + C_y(R))^+}} \\ \leq& 3\left(\frac{1728e(2C_y(R))^2}{u}\right)^{2V_{([\mathcal{P}_{0,k}]_y + C_y(R))^+}}\end{aligned}$$

en utilisant que $\log(x) \leq x$.
Nous avons alors :

$$V_{([\mathcal{P}_{0,k}]_y + C_y(R))^+} = V_{([\mathcal{P}_{0,k}]_y)^+}.$$

En effet, notons $n = V_{([\mathcal{P}_{0,k}]_y + C_y(R))^+}$ et considérons un sous-ensemble de $\mathbb{R}^{d'+1}$ associé

$$\{(z_1, t_1), \ldots, (z_n, t_n)\}.$$

Par définition de $V_{([\mathcal{P}_{0,k}]_y + C_y(R))^+}$, pour tout sous-ensemble I de $\{1, \ldots, n\}$, on a qu'il existe $g \in [\mathcal{P}_{0,k}]_y + C_y(R)$ tel que $g(z_i) \geq t_i$ pour $i \in I$ et $g(z_i) < t_i$ sinon. On a alors juste à considérer le sous-ensemble $\{(z_1, t_1 - C_y(R)), \ldots, (z_n, t_n - C_y(R))\}$ pour conclure que $V_{([\mathcal{P}_{0,k}]_y + C_y(R))^+} \leq V_{([\mathcal{P}_{0,k}]_y)^+}$. L'inégalité dans l'autre sens s'établit de la même façon.
De même, il est facile de montrer que

$$V_{([\mathcal{P}_{0,k}]_y)^+} \leq V_{(\mathcal{P}_{0,k})^+}.$$

En effet, notons $n = V_{([\mathcal{P}_{0,k}]_y)^+}$ et considérons un sous-ensemble de $\mathbb{R}^{d'+1}$ associé $\{(z_1, t_1), \cdots, (z_n, t_n)\}$, qui est éclaté par $([\mathcal{P}_{0,k}]_y)^+$. Par définition, pour tout sous-ensemble I de $\{1, \ldots, n\}$, il existe $g \in \mathcal{P}_{0,k}$ tel que $[g]_y(z_i) < t_i$ pour $i \in I$ et $[g]_y(z_i) \geq t_i$ pour $i \notin I$. Il nous suffit de montrer que pour $i \in I$ $g(z_i) \leq [g]_y(z_i)$ et pour $i \notin I$ $g(z_i) \geq [g]_y(z_i)$ pour exhiber un élément de $(\mathcal{P}_{0,k})^+$ qui réalise le sous-ensemble I.
Supposons qu'il existe $i \in I$ tel que $g(z_i) > [g]_y(z_i)$. Alors nécessairement $g(z_i) > C_y(R)$ et donc $[g]_y(z_i) = C_y(R)$. Ainsi on devrait avoir $t_i > C_y(R)$ et il ne sera pas possible de trouver un élément de $[\mathcal{P}_{0,k}]_y$ qui réalisera le complémentaire de i dans $\{1, \ldots, n\}$. En effet, si un tel élément g' existait, ceci nécessiterait $[g']_y(z_i) \geq t_i > C_y(R)$ ce qui est impossible.

Ainsi, $g(z_i) \leq [g]_y(z_i)$ pour $i \in I$.
D'autre part, supposons qu'il existe $i \notin I$ tel que $g(z_i) < [g]_y(z_i)$. Alors nécessairement $g(z_i) < -C_y(R)$ et donc $[g]_y(z_i) = -C_y(R)$. Ainsi, $t_i \leq -C_y(R)$ et il ne sera pas possible de réaliser le sous-ensemble $\{(z_i, t_i)\}$. En effet, ceci nécessiterait d'exhiber un élément g' de $\mathcal{P}_{0,k}$ tel que $[g']_y(z_i) < t_i \leq -C_y(R)$ ce qui est impossible. Ainsi pour $i \notin I$ $g(z_i) \geq [g]_y(z_i)$. On a donc exhibé un élément de $\mathcal{P}_{0,k}$ tel que $g(z_i) < t_i$ pour $i \in I$ et $g(z_i) \geq t_i$ pour $i \notin I$. Ainsi, $V_{([\mathcal{P}_{0,k}]_y)^+} \leq V_{(\mathcal{P}_{0,k})^+}$.
Comme le mentionne la remarque du haut de la page 152 dans [GKKW02], on a

$$\begin{aligned}\mathcal{P}_{0,k}^+ =& \big\{\{(z,t) \in \mathbb{R}^{d'} \times \mathbb{R} : t \leq g(z)\} : g \in \mathcal{P}_{0,k}\big\} \\ \subset& \big\{\{(z,t) \in \mathbb{R}^{d'} \times \mathbb{R} : \alpha.t + g(z) \geq 0\} : g \in \mathcal{P}_{0,k}, \alpha \in \mathbb{R}\big\}.\end{aligned}$$

On applique alors le théorème A.7 pour conclure que

$$V_{(\mathcal{P}_{0,k})^+} \leq K_{0,k} + 1.$$

On en déduit finalement l'inégalité :

$$\mathcal{N}_2(\frac{\sqrt{u}}{24}, [\mathcal{P}_{0,k}]_y + C_y(R), P_{t_k}^{N,1:2M}) \leq 3\bigg(\frac{1728e(2C_y(R))^2}{u}\bigg)^{2(K_{0,k}+1)}.$$

En injectant dans (II.3.1), il vient

$$\begin{aligned}\mathbb{P}(T_1 > u) \leq& 9\big(\frac{1728e(2C_y(R))^2}{u}\big)^{2(K_{0,k}+1)} \exp\big(-\frac{Mu}{576(2C_y(R))^2}\big) \\ \leq& 9(12eM)^{2(K_{0,k}+1)} \exp\big(-\frac{Mu}{2304C_y(R)^2}\big)\end{aligned}$$

pour $u > \frac{576C_y(R)^2}{M}$. On peut alors se servir de cette dernière inégalité pour majorer $\mathbb{E}(T_1)$. En effet :

$$\begin{aligned}\mathbb{E}(T_1) =& \int_0^\infty \mathbb{P}(T_1 > u)du \\ \leq& v + \int_v^\infty \mathbb{P}(T_1 > u)du \\ \leq& v + 9(12eM)^{2(K_{0,k}+1)} \int_v^\infty \exp\big(-\frac{Mu}{2304C_y(R)^2}\big)du \\ =& v + 9(12eM)^{2(K_{0,k}+1)}.\frac{2304C_y(R)^2}{M} \exp\big(-\frac{Mv}{2304C_y(R)^2}\big) \\ =& \phi(v)\end{aligned}$$

pour $v > \frac{576C_y(R)^2}{M}$. Le minimum de la fonction ϕ est atteint pour

$$v = \frac{2304C_y(R)^2}{M} \log\{9(12eM)^{2(K_{0,k}+1)}\}.$$

Pour cette valeur, on obtient :

$$\mathbb{E}(T_1) \leq C\frac{C_y(R)^2 K_{0,k}}{M}\log(M).$$

□

3.2 Erreur sur $z^{N,R}$

A présent que nous avons déterminé la vitesse de convergence de l'erreur sur $y^{N,R}$, il est facile de montrer la vitesse de convergence de l'erreur sur $z^{N,R}$. On peut en particulier établir

Théorème II.4

$$\begin{aligned}
&h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}\sum_{l=1}^{q}|z_{l,k}^{N,R}(P_{t_k}^{N,m}) - z_{l,k}^{N,R,M}(P_{t_k}^{N,m})|^2 \\
\leq & Ch^{\beta-1} + C\frac{C_y(R)^2 + \|f^R\|_\infty^2}{M}\sum_{k=0}^{N-1}\max_{0\leq l\leq q}\mathbb{E}(K_{l,k}(\omega)) + C\sum_{k=0}^{N-1}\inf_{\alpha}\mathbb{E}\big(|y_k^{N,R}(P_{t_k}^N) - \alpha.p_{0,k}(P_{t_k}^N)|^2\big) \\
&+ C\sum_{k=0}^{N-1}\sum_{l=1}^{q}\inf_{\alpha}\mathbb{E}\big(|\alpha.p_{l,k}(P_{t_k}^N) - \sqrt{h}z_{l,k}^{N,R}(P_{t_k}^N)|^2\big) \\
&+ \frac{C}{h}C_y(R)^2 e^{-\frac{Mh^{\beta+1}}{144C_y(R)^2}}\sum_{k=0}^{N-1}\mathbb{E}\bigg(\mathcal{N}_2\big(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}}, [\mathcal{P}_{0,k+1}]_y - y_{k+1}^{N,R}, (P_{t_{k+1}}^{N,m})_{1\leq m\leq M}, (\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M}\big)\bigg).
\end{aligned}$$

En particulier, on remarque que l'erreur sur $z^{N,R}$

$$h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}\sum_{l=1}^{q}|z_{l,k}^{N,R}(P_{t_k}^{N,m}) - z_{l,k}^{N,R,M}(P_{t_k}^{N,m})|^2$$

est du même ordre que l'erreur sur $y^{N,R}$

$$\max_{0\leq k\leq N}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R,M}(P_{t_k}^{N,m}) - y_k^{N,R}(P_{t_k}^{N,m})|^2,$$

ce qui est assez naturel pour des normes d'EDSR.

PREUVE: (II.2.19-II.2.20) nous donnaient

$$
\begin{aligned}
&h\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}\sum_{l=1}^{q}|z_{l,k}^{N,R}(P_{t_k}^{N,m})-z_{l,k}^{N,R,M}(P_{t_k}^{N,m})|^2\\
\leq& Ch\mathbb{E}\sum_{l=1}^{q}\{T_{4,k,l,M}+T_{5,k,l,M}\}\\
&+C\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}\{|y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})-y_{k+1}^{N,R,M}(\tilde{P}_{t_{k+1}}^{N,m})|^2-|\tilde{\mathbb{E}}_{k+1}\{y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})-y_{k+1}^{N,R,M}(\tilde{P}_{t_{k+1}}^{N,m})\}|^2\}
\end{aligned}
$$

ce qui, en sommant sur k permet d'écrire

$$
\begin{aligned}
&h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}\sum_{l=1}^{q}|z_{l,k}^{N,R}(P_{t_k}^{N,m})-z_{l,k}^{N,R,M}(P_{t_k}^{N,m})|^2\\
\leq& Ch\mathbb{E}\sum_{k=0}^{N-1}\sum_{l=1}^{q}\{T_{4,k,l,M}+T_{5,k,l,M}\}\\
&+C\sum_{k=0}^{N-1}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}\{|y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})-y_{k+1}^{N,R,M}(\tilde{P}_{t_{k+1}}^{N,m})|^2-|\tilde{\mathbb{E}}_{k+1}\{y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})-y_{k+1}^{N,R,M}(\tilde{P}_{t_{k+1}}^{N,m})\}|^2\}.
\end{aligned}
$$

En effectuant un changement d'indice dans le dernier terme du membre de droite, on obtient :

$$
\begin{aligned}
&h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}\sum_{l=1}^{q}|z_{l,k}^{N,R}(P_{t_k}^{N,m})-z_{l,k}^{N,R,M}(P_{t_k}^{N,m})|^2\\
\leq& Ch\mathbb{E}\sum_{k=0}^{N-1}\sum_{l=1}^{q}\{T_{4,k,l,M}+T_{5,k,l,M}\}\\
&+C\sum_{k=0}^{N-1}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}\{|y_{k}^{N,R}(\tilde{P}_{t_{k}}^{N,m})-y_{k}^{N,R,M}(\tilde{P}_{t_{k}}^{N,m})|^2-|\tilde{\mathbb{E}}_{k+1}\{y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})-y_{k+1}^{N,R,M}(\tilde{P}_{t_{k+1}}^{N,m})\}|^2\}.
\end{aligned}
\tag{II.3.2}
$$

On veut alors utiliser (II.2.18) qui borne

$$
\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}\{|y_{k}^{N,R}(P_{t_k}^{N,m})-y_{k}^{N,R,M}(P_{t_k}^{N,m})|^2-|\tilde{\mathbb{E}}_{k+1}\{y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})-y_{k+1}^{N,R,M}(\tilde{P}_{t_{k+1}}^{N,m})\}|^2\}.
$$

Il nous faut donc une fois encore introduire les évènements $(A_k^M)_k$. En effet, pour $k \geq 1$:

$$
\begin{aligned}
&\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(\tilde{P}_{t_k}^{N,m}) - y_k^{N,R,M}(\tilde{P}_{t_k}^{N,m})|^2 \\
=&\mathbb{E}\|y_k^{N,R} - y_k^{N,R,M}\|_{k,\tilde{M}}^2 \\
\leq&(1+h^{-1})\mathbb{E}\left(\|y_k^{N,R} - y_k^{N,R,M}\|_{k,\tilde{M}} - \|y_k^{N,R} - y_k^{N,R,M}\|_{k,M}\right)_+^2 + (1+h)\mathbb{E}\|y_k^{N,R} - y_k^{N,R,M}\|_{k,M}^2 \\
\leq&Ch^{\beta} + \frac{C}{h}C_y(R)^2\mathbb{P}([A_{k-1}^M]^c) + (1+h)\mathbb{E}\|y_k^{N,R} - y_k^{N,R,M}\|_{k,M}^2. \qquad \text{(II.3.3)}
\end{aligned}
$$

En injectant cette dernière inégalité dans (II.3.2) et en utilisant (II.2.18), on obtient finalement, pour tout $\gamma > 0$:

$$
\begin{aligned}
&h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}\sum_{l=1}^{q}|z_{l,k}^{N,R}(P_{t_k}^{N,m}) - z_{l,k}^{N,R,M}(P_{t_k}^{N,m})|^2 \\
\leq&Ch\mathbb{E}\sum_{k=0}^{N-1}\sum_{l=1}^{q}\{T_{4,k,l,M} + T_{5,k,l,M}\} + Ch^{\beta-1} + \frac{C}{h}C_y(R)^2\sum_{k=0}^{N-1}\mathbb{P}([A_k^M]^c) \\
&+ C\max_{0\leq k\leq N}\mathbb{E}\|y_k^{N,R} - y_k^{N,R,M}\|_{k,M}^2 + C(1+\gamma h)\sum_{k=0}^{N-1}\mathbb{E}(T_{2,k,M}) \\
&+ C\gamma h\sum_{k=0}^{N-1}\frac{1}{M}\mathbb{E}\sum_{m=1}^{M}|\tilde{\mathbb{E}}_{k+1}\{y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) - [\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y\}|^2 + C\sum_{k=0}^{N-1}\mathbb{E}T_{1,k,M} \\
&+ C(1+\frac{1}{\gamma h})\sum_{k=0}^{N-1}\mathbb{E}T_{3,k,M} + C(h^2+\frac{h}{\gamma})\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}|y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) - [\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y|^2 \\
&+ C(h^2+\frac{h}{\gamma})\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{l=1}^{q}\sum_{m=1}^{M}|z_{l,k}^{N,R}(P_{t_k}^{N,m}) - [\alpha_{l,k}^M.p_{l,k}^m]_z|^2.
\end{aligned}
$$

Reste alors à choisir $\gamma = 2C$, à supposer h suffisamment petit et transformer le terme $\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}) - [\alpha_{0,k+1}^M.\tilde{p}_{0,k+1}^m]_y|^2$, qui intervient encore dans le membre de droite de l'inégalité ci-dessus, en $\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_{k+1}^{N,R}(P_{t_{k+1}}^{N,m}) - [\alpha_{0,k+1}^M.p_{0,k+1}^m]_y|^2$ à l'aide de l'évènement A_k^M de la même façon que pour (II.3.3) pour finalement parvenir à la conclusion du théorème. □

De la même façon que pour le théorème II.3, il est alors facile d'établir

Théorème II.5

$$
\begin{aligned}
&h\mathbb{E}\sum_{k=0}^{N-1}\sum_{l=1}^{q}\int |z_{l,k}^{N,R}(\cdot)-z_{l,k}^{N,R,M}(\cdot)|^2\mu_k(dx)\\
\leq& Ch^{\beta-1}+C\frac{C_y(R)^2+\|f^R\|_\infty^2}{M}\sum_{k=0}^{N-1}\max_{0\leq l\leq q}\mathbb{E}(K_{l,k}(\omega))+C\sum_{k=0}^{N-1}\inf_{\alpha}\mathbb{E}\big(|y_k^{N,R}(P_{t_k}^N)-\alpha.p_{0,k}(P_{t_k}^N)|^2\big)\\
&+C\sum_{k=0}^{N-1}\sum_{l=1}^{q}\{\inf_{\alpha}\mathbb{E}\big(|\alpha.p_{l,k}(P_{t_k}^N)-\sqrt{h}z_{l,k}^{N,R}(P_{t_k}^N)|^2\big)+\frac{C_y(R)^2K_{l,k}}{M}\log(M)\}\\
&+\frac{C}{h}C_y(R)^2e^{-\frac{Mh^{\beta+1}}{144C_y(R)^2}}\sum_{k=0}^{N-1}\mathbb{E}\bigg(\mathcal{N}_2\big(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}},[\mathcal{P}_{0,k+1}]_y-y_{k+1}^{N,R},(P_{t_{k+1}}^{N,m})_{1\leq m\leq M},(\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M}\big)\bigg).
\end{aligned}
$$

3.3 Utilisation des majorations obtenues

3.3.1 Estimation des nombres de couverture

Dans les majorations d'erreur que l'on a obtenues au chapitre précédent interviennent les nombres de couverture $\mathcal{N}_2\big(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}},[\mathcal{P}_{0,k+1}]_y-y_{k+1}^{N,R}(\cdot),(P_{t_{k+1}}^{N,m})_{1\leq m\leq M},(\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M}\big)$. Pour rendre complètement explicites ces majorations, il nous faut à présent borner ces nombres. Dans le cas général, on peut reprendre une partie de la démonstration du théorème II.3 pour déduire, par une application du théorème A.6, que

$$
\begin{aligned}
&\mathcal{N}_2\big(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}},[\mathcal{P}_{0,k+1}]_y-y_{k+1}^{N,R}(\cdot),(P_{t_{k+1}}^{N,m})_{1\leq m\leq M},(\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M}\big)\\
\leq&3\bigg(\frac{216C_y(R)^2}{h^{\beta+1}}\bigg)^{2(K_{0,k+1}+1)}\\
\leq&C\exp(CK_{0,k+1}\log\frac{C_y(R)}{h^{\beta+1}}).
\end{aligned}
\tag{II.3.4}
$$

On obtient donc un contrôle du nombre de couverture exponentiel dans le nombre de fonctions de base. On déduit de (II.3.4) la majoration d'erreur suivante :

$$
\begin{aligned}
&\max_{0\le k\le N}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m})-y_k^{N,R,M}(P_{t_k}^{N,m})|^2\\
&+h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}\sum_{l=1}^{q}|z_{l,k}^{N,R}(P_{t_k}^{N,m})-z_{l,k}^{N,R,M}(P_{t_k}^{N,m})|^2\\
\le& C\frac{C_y(R)^2+\|f^R\|_\infty^2}{M}\sum_{k=0}^{N-1}\max_{0\le l\le q}\mathbb{E}(K_{l,k}(\omega))+C\sum_{k=0}^{N-1}\inf_{\alpha}\mathbb{E}|\alpha.p_{0,k}(P_{t_k}^N)-y_k^{N,R}(P_{t_k}^N)|^2\\
&+C\sum_{k=0}^{N-1}\sum_{l=1}^{q}\inf_{\alpha}\mathbb{E}|\alpha.p_{l,k}(P_{t_k}^N)-\sqrt{h}z_{l,k}^{N,R}(P_{t_k}^N)|^2+Ch^{\beta-1}\\
&+\frac{C}{h^2}C_y(R)^2\exp\left(C\max_k K_{0,k}\log(\frac{C_y(R)}{h^{\beta+1}})-\frac{Mh^{\beta+1}}{144C_y(R)^2}\right).
\end{aligned}
$$

Nous pouvons utiliser ce résultat de deux manières. Nous pouvons soit faire varier en même temps les paramètres $h, K_{l,k}$ et M de manière à obtenir des convergences asymptotiques vers la solution de l'EDSR, soit à complexité fixée (et une fois identifiée la dépendance de la complexité en les différents paramètres), essayer de minimiser en $h, K_{l,k}$ et M le majorant de l'erreur.

3.3.2 Variation simultanée des paramètres

Supposons que nous sachions, à chaque instant de discrétisation t_k, estimer les erreurs d'approximation de $y_k^{N,R}(\cdot)$ et de $z_{l,k}^{N,R}(\cdot)$ en fonction du choix des bases $p_{l,k}$. Ceci est bien envisageable car nous avons des informations sur ces fonctions inconnues et en particulier, vu la remarque de la sous-section 1.3.2, nous savons qu'elles sont Lipschitz. Nous renvoyons à la partie numérique pour des calculs plus précis et supposons que la base $p_{l,k}$ est la même quel que soit $l \ge 0$ et quel que soit k. Ainsi pour $0 \le k \le N-1$ et $0 \le l \le q$ $K_{l,k} = K$. Nous notons sans plus détailler $\epsilon_k(K)$ la somme des erreurs d'approximation

$$
\inf_{\alpha}\mathbb{E}\big(|y_k^{N,R}(P_{t_k}^N)-\alpha.p_{0,k}(P_{t_k}^N)|^2\big)+\sum_{l=1}^{q}\inf_{\alpha}\mathbb{E}\big(|\alpha.p_{l,k}(P_{t_k}^N)-\sqrt{h}z_{l,k}^{N,R}(P_{t_k}^N)|^2\big).
$$

La majoration de l'erreur commise dans l'algorithme se réécrit

$$\begin{aligned}
&\max_{0\leq k\leq N}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R,M}(P_{t_k}^{N,m})-y_k^{N,R}(P_{t_k}^{N,m})|^2\\
&+h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}\sum_{l=1}^{q}|z_{l,k}^{N,R}(P_{t_k}^{N,m})-z_{l,k}^{N,R,M}(P_{t_k}^{N,m})|^2\\
\leq& C(\frac{1}{N})^{\beta-1}+CNK\frac{C_y(R)^2+\|f^R\|_\infty^2}{M}\\
&+C\sum_{k=0}^{N-1}\epsilon_k(K)+\frac{CC_y(R)^2}{h^2}\exp\left(CK\log(\frac{NC_y(R)}{T})-\frac{MT^{\beta+1}}{144N^{\beta+1}C_y(R)^2}\}\right). \qquad \text{(II.3.5)}
\end{aligned}$$

On voit alors que pour que l'algorithme converge lorsque K, N et M tendent vers l'infini, il faut tout d'abord que $\beta > \mathbf{1}$, puis que $\max_k \epsilon_k(K)$ tende vers 0 lorsque K tend vers l'infini mais également que $\max_k \epsilon_k(K)$ tende suffisamment vite vers 0 pour que $\sum_{k=0}^{N-1}\epsilon_k(K)$ tende vers 0. Ceci n'est pas nouveau et apparaissait déjà dans le théorème I.2.
Ce qui est nouveau, c'est que l'on sait maintenant comment fixer M en fonction des autres paramètres pour faire converger l'algorithme. En effet, la condition $\sum_{k=0}^{N-1}\epsilon_k(K)\to 0$ impose que K croisse vers l'infini comme une certaine fonction $K(N)$. Par exemple, pour le cas de la base d'indicatrices d'hypercubes, on verra dans la partie numérique que l'on aura

$$\epsilon_k(K)\approx K^{-\frac{2}{d}}.$$

Ici, on doit donc avoir $\mathbf{NK^{-\frac{2}{d}}\to 0}$ pour que $\sum_{k=0}^{N-1}\epsilon_k(K)$ tende vers 0. On doit alors s'assurer que l'on applique l'algorithme avec un nombre de trajectoires M tel que $\epsilon(M)=CK\log(\frac{NC_y(R)}{T})-\frac{MT^{\beta+1}}{144N^{\beta+1}C_y(R)^2}\to-\infty$ et suffisamment vite pour que $C_y(R)^2N^2\exp(\epsilon(M))\to 0$. Dans notre cas, cela revient à prendre $\mathbf{M=CC_y(R)^2N^{\beta+1}K\log(\frac{NC_y(R)}{T})}$ pour une constante C assez grande.
Nous devrions alors théoriquement prendre en compte que R doit tendre vers l'infini pour faire tendre vers 0 l'erreur de localisation. Nous négligeons cet aspect car si la localisation est suffisamment bien choisie dès le départ, l'erreur induite est négligeable. Ceci revient à imposer finalement que $\mathbf{M=CN^{\beta+1}K\log(N)}$ pour une constante C assez grande.

3.3.3 Erreur et complexité

Calcul de complexité

On raisonne avec des dimensions d', q fixées et on suppose $K_{l,k}=K$ pour tout (l,k). L'algorithme utilise M trajectoires de P^N. Ces trajectoires nécessitent $O(MN)$ calculs. Puis, pour la mise en place de l'algorithme, on doit à chaque pas de temps former les

$q+1$ matrices de régression soit $O(MK)$ opérations. Puis, on doit résoudre les problèmes de moindres carrés associés à ces matrices. On suppose que ces problèmes sont résolus en effectuant des décompositions en valeurs singulières (SVD) des matrices de régression. La complexité d'une SVD dépend des algorithmes employés mais dans notre cas, en se référant à [GVL96], on peut compter sur une complexité $O(MK^2)$ en effectuant l'hypothèse que $M \gg K$. Ainsi, la complexité totale est de l'ordre de $\mathbf{NMK^2}$. Bien sûr, pour des cas particuliers de fonctions de base, la complexité peut être moindre, comme par exemple avec la base d'hypercubes d'arête δ où, dans chaque hypercube, on doit juste effectuer des sommes de quantités dépendant des simulations qui sont tombées dans cet hypercube. A l'instant t_k le terme dominant de la complexité provient alors de savoir quelles sont les trajectoires qui sont tombées dans un hypercube donné, ce qui est bien moins coûteux que $O(MK^2)$. Nous reviendrons sur ce point un peu plus bas. Notons enfin que la complexité de l'algorithme avec simulation d'incréments browniens supplémentaires est au plus égale à un facteur constant multiplié par la complexité de l'algorithme sans simulation d'incréments browniens supplémentaires. Cette simulation supplémentaire ne coûte donc pas très cher et en particulier beaucoup moins cher que d'utiliser des trajectoires indépendantes à chaque pas de temps, comme dans [BT04] par exemple.

Erreur en fonction de la complexité

Nous allons illustrer par un exemple comment on peut relier l'erreur commise dans l'algorithme à la complexité. Nous désignons ici par erreur le membre de gauche de (II.3.5) (ce qui correspondrait plutôt au carré d'une erreur). On va faire un raisonnement sans tenir compte des constantes. Vu que la complexité est croissante en M et que le choix $M \approx N^{\beta+1+\epsilon}K$ où $\epsilon > 0$ assure que le terme exponentiel sera négligeable devant les autres termes d'erreurs (pour N et K assez grands) et que le terme $CNK\frac{C_y(R)^2+\|f^R\|_\infty^2}{M}$ est négligeable devant $h^{\beta-1}$, on a intérêt à prendre juste $M \approx N^{\beta+1+\epsilon}K$ et pas plus de simulations. Ainsi (II.3.5) nous donne que l'erreur commise dans l'algorithme est majorée par :

$$h^{\beta-1} + \frac{1}{hK^{\frac{2}{d}}},$$

alors que la complexité $\mathcal{C}$ vaut $\mathcal{C} = N^{\beta+2+\epsilon}K^3$ soit $K = \mathcal{C}^{\frac{1}{3}}N^{-\frac{\beta+2+\epsilon}{3}}$. En réinjectant, l'erreur vaut en fonction de N et $\mathcal{C}$:

$$(\frac{1}{N})^{\beta-1} + N^{1+\frac{2(\beta+2+\epsilon)}{3d}}\mathcal{C}^{-\frac{2}{3d}}.$$

A l'optimum, les deux termes (l'un décroissant en N l'autre croissant en N) sont du même ordre ce qui donne :

$$N = \mathcal{C}^{\frac{2}{(3d+2)\beta+2(2+\epsilon)}}.$$

L'erreur associée est alors $(\frac{1}{N})^{\beta-1}$ soit

$$\mathcal{C}^{-\frac{2(\beta-1)}{(3d+2)\beta+2(2+\epsilon)}}.$$

Par exemple pour $\beta = 2$, on obtient la même vitesse de convergence que la simple erreur de discrétisation en temps et pour $d = 1$ (et en prenant le cas limite $\epsilon = 0$) cela donne une erreur en $\mathcal{C}^{-\frac{1}{7}}$.

Cas de la base d'hypercubes et comparaison avec [BT04]

Nous avons fait un calcul ci-dessus dans un cas très général où on ne spécifie pas les fonctions de base. Nous utilisons dans la partie numérique essentiellement des bases locales. L'exemple le plus simple est celui de la base d'hypercubes. A un instant donné, le terme le plus coûteux dans la complexité est de savoir, pour chaque hypercube, quelles sont les simulations qui sont tombées dedans. Pour chacune des M simulations, on doit savoir auquel des K hypercubes elle appartient. Ceci a un coût $Md\log(K)$. Les autres termes de la complexité sont négligeables car dans chaque hypercube, on fait un nombre d'opérations proportionnel au nombre de simulations qui sont tombées dedans donc au total M opérations pour l'ensemble des K hypercubes. Ainsi dans le cas de la base d'hypercubes la complexité totale (lorsqu'on somme le nombre d'opérations sur l'ensemble des dates de discrétisation) est de l'ordre de $\mathbf{NMd}\log(\mathbf{K})$. On voit que l'on gagne beaucoup par rapport au cas général où le terme dominant de la complexité valait $\mathbf{NMdK^2}$.
En négligeant le terme en $\log(K)$ dans la complexité, le même calcul que celui effectué dans le cas général donne une erreur (pour $\beta = 2$) de l'ordre de $\mathcal{C}^{-\frac{1}{4+\epsilon+\mathbf{d}}}$, où $\epsilon > 0$ est tel que $M \approx N^{\mathbf{1}+\beta+\epsilon}K$. Nous allons comparer avec l'algorithme décrit dans [BT04].
Pour cet algorithme, à chaque instant de discrétisation t_k et pour chacune des M trajectoires simulées pour l'instant t_k (rappelons que N paquets de M trajectoires sont simulés) on doit effectuer un calcul impliquant les M trajectoires simulées pour l'instant t_{k+1}. Ces termes incluent des intégrales de Skorohod qui sont difficiles à calculer dans le cas d'une diffusion générale mais qui ont une forme explicite dans le cas du modèle Black-Scholes. Plaçons nous dans ce cas, qui est le plus favorable pour cette méthode (la complexité de notre algorithme est quasiment indépendante du modèle utilisé). La complexité est alors de l'ordre de $\mathbf{M^2N}$. Or, l'erreur associée à l'algorithme est donné dans le théorème 6.2 dans [BT04] et est de l'ordre de $\frac{\mathbf{N^{2+\frac{d}{4}}}}{\sqrt{\mathbf{M}}}$. Si on exprime l'erreur en fonction de la complexité, on trouve $\mathcal{C}^{-\frac{1}{13+\mathbf{d}}}$. La différence est donc très importante avec l'algorithme décrit dans cette partie. Si on veut diminuer l'erreur par un facteur 2 on doit augmenter le temps de calcul de 2^{13+d} dans le cas [BT04] contre $2^{4+d+\epsilon}$ ici.
Remarquons enfin, que si sur chaque hypercube on définit des bases polynomiales locales, la complexité ne change pas alors que l'erreur d'approximation est améliorée. Ceci améliore

encore le rapport erreur/complexité de notre algorithme. Nous vérifierons ce phénomène dans la partie numérique.

Chapitre 4

Extensions

Nous avons présenté ici un algorithme qui utilise à chaque pas de temps M incréments browniens auxiliaires $(\Delta\tilde{W}^m_{k,k+1})_{1\leq m\leq M}$. La raison est de pouvoir annuler les termes non-diagonaux de la matrice $A=\tilde{\mathbb{E}}_{k+1}\big(\{V-\tilde{\mathbb{E}}_{k+1}(V)\}\{V-\tilde{\mathbb{E}}_{k+1}(V)\}^*\big)$ où V est le vecteur de $\mathbb{R}^M$ de coordonnées $([\alpha^M_{0,k+1}.\tilde{p}^m_{0,k+1}]_y\frac{\Delta\tilde{W}^m_{l,k}}{\sqrt{h}})_{1\leq m\leq M}$ ou $([\alpha^M_{0,k+1}.\tilde{p}^m_{0,k+1}]_y)_{1\leq m\leq M}$. L'annulation des termes non-diagonaux de A nous permet de contrôler $\|A\|_2$ et ainsi de contrôler les termes $\tilde{\mathbb{E}}_{k+1}\frac{1}{M}\sum_{m=1}^M|\{\alpha^M_{l,k}-\tilde{\mathbb{E}}_{k+1}(\alpha^M_{l,k})\}.p^m_{l,k}|^2$ et $\tilde{\mathbb{E}}_{k+1}\frac{1}{M}\sum_{m=1}^M|\{\alpha^M_{0,k}-\tilde{\mathbb{E}}_{k+1}(\alpha^M_{0,k})\}.p^m_{0,k}|^2$ (voir la majoration des termes $T_{2,k,M}$ et $T_{5,k,l,M}$). En effet, **pour contrôler la présence de Z dans le driver** , nous avons besoin d'appliquer une inégalité de Cauchy-Schwarz pour établir (II.2.20) et pour appliquer cette inégalité, nous avons besoin de travailler avec $\tilde{\mathbb{E}}_{k+1}(\alpha^M_{l,k})$ et non $\alpha^M_{l,k}$. Remarquons que l'annulation des termes non-diagonaux de A n'est pas du tout évidente si on ne simule pas des incréments browniens supplémentaires : si $V_m=[\alpha^M_{0,k+1}.p^m_{0,k+1}]_y$, on ne peut pas écrire

$$\begin{aligned}&\mathbb{E}\big([\alpha^M_{0,k+1}.p^m_{0,k+1}]_y[\alpha^M_{0,k+1}.p^{m'}_{0,k+1}]_y|(P^{N,m}_{t_k})_{1\leq m\leq M}\big)\\ =&\mathbb{E}\big([\alpha^M_{0,k+1}.p^m_{0,k+1}]_y|(P^{N,m}_{t_k})_{1\leq m\leq M}\big)\mathbb{E}\big([\alpha^M_{0,k+1}.p^{m'}_{0,k+1}]_y|(P^{N,m}_{t_k})_{1\leq m\leq M}\big)\end{aligned}$$

pour $m\neq m'$ à cause de la présence de $\alpha^M_{0,k+1}$ qui dépend de toutes les simulations $(P^{N,m}_{t_k})_{1\leq m\leq M,0\leq k\leq N}$.
Lorsque le driver f ne dépend pas de Z, on devrait pouvoir se passer de la simulation des incréments browniens auxiliaires. C'est ce que l'on va constater tout de suite en traitant un exemple simplifié mais représentatif du cas f indépendant de Z.

4.1 Driver nul

Supposons que l'on veuille approcher $Y_t=y(t,W_t)=\mathbb{E}\big(\phi(W_T)\big|\mathcal{F}_t\big)$ où $T>0$ est fixé, où W est un mouvement brownien standard de dimension 1 et où ϕ est une fonction

mesurable bornée. Remarquons que (Y, Z) est la solution de

$$Y_t = \phi(W_T) - \int_t^T Z_s dW_s.$$

Pour simplifier les notations, on note $\mathbb{E}_k^M = \mathbb{E}\big(\cdot\,|(W_{t_k}^m)_{1\leq m\leq M}\big)$.
L'erreur ϵ_k dont nous voulons étudier la propagation est définie comme :

$$\epsilon_k = \mathbb{E}\frac{1}{M}\sum_{m=1}^{M} |y^{N,M}(t_k, W_{t_k}^m) - y(t_k, W_{t_k}^m)|^2$$

où $y^{N,M}(t_k, \cdot)$ est défini comme $[\alpha_k^M.p_k]_y(\cdot)$ où $p_k(\cdot)$ est une base de fonctions de cardinal K_k (on note cette fois $p_k^m = p_k(W_{t_k}^m)$) et α_k^M est solution du problème de moindres carrés :

$$\inf_\alpha \frac{1}{M}\sum_{m=1}^{M} |y^{N,M}(t_{k+1}, W_{t_{k+1}}^m) - \alpha.p_k(W_{t_k}^m)|^2$$

et où

$$[\psi]_y(\cdot) = -\|\phi\|_\infty \vee (\psi(\cdot) \wedge \|\phi\|_\infty)$$

pour une fonction ψ donnée.
Nous allons montrer :

Proposition II.4 *On désigne par $K_k(\omega)$ le rang de la matrice de régression à l'instant t_k, de taille $M \times K_k$ et dont les lignes sont données par $(p_k(W_{t_k}^m)^*)_{1\leq m\leq M}$. Pour $k = N-1$, on a*

$$\epsilon_{N-1} \leq \inf_\alpha \mathbb{E}|y(t_{N-1}, W_{t_{N-1}}) - \alpha.p_{N-1}(W_{t_{N-1}})|^2 + \frac{\|\phi\|_\infty^2 \mathbb{E}(K_{N-1}(\omega))}{M}$$

et pour $k \leq N-2$:

$$\epsilon_k \leq (1+h)\epsilon_{k+1} + \inf_\alpha \mathbb{E}|y(t_k, W_{t_k}) - \alpha.p_k(W_{t_k})|^2 + C\frac{\|\phi\|_\infty^2 \mathbb{E}(K_k(\omega)))}{Mh}.$$

Remarque II.2 *Si on compare avec la majoration du théorème II.2, on paie un facteur h^{-1} dans le terme $C\frac{\|\phi\|_\infty^2 \mathbb{E}(K_k(\omega)))}{Mh}$ qui aurait été $C\frac{\|\phi\|_\infty^2 \mathbb{E}(K_k(\omega)))}{M}$ si on avait utilisé des browniens auxiliaires. Par contre, dans le cas des browniens auxiliaires, les termes $h^\beta + C_y(R)^2 \exp(-\frac{Mh^{\beta+1}}{144C_y(R)^2})\mathbb{E}\Big(\mathcal{N}_2\big(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}}, [\mathcal{P}_{0,k+1}]_y - y_{k+1}^{N,R}(\cdot), (P_{t_{k+1}}^{N,m})_{1\leq m\leq M}, (\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M}\big)\Big)$ apparaissaient pour transformer la norme $\|\cdot\|_{k+1,\tilde{M}}$ en $\|\cdot\|_{k+1,M}$. Ces termes sont ici absents.*

PREUVE: Pour $k = N-1$, on se trouve directement dans le cadre du théorème A.8. Pour $k \leq N-2$, introduisons β_k^M solution de :

$$\beta_k^M = \arg\inf_{\beta} \frac{1}{M} \sum_{m=1}^{M} |y(t_{k+1}, W_{t_{k+1}}^m) - \beta.p_k(W_{t_k}^m)|^2.$$

Alors à l'aide de β_k^M on a :

$$\begin{aligned}
&\mathbb{E}_k^M \frac{1}{M} \sum_{m=1}^{M} |y^{N,M}(t_k, W_{t_k}^m) - y(t_k, W_{t_k}^m)|^2 \\
\leq &\mathbb{E}_k^M \frac{1}{M} \sum_{m=1}^{M} |\alpha_k^M.p_k^m - y(t_k, W_{t_k}^m)|^2 \\
= &\mathbb{E}_k^M \frac{1}{M} \sum_{m=1}^{M} |\alpha_k^M.p_k^m - \mathbb{E}_k^M(\beta_k^M).p_k^m|^2 + \mathbb{E}_k^M \frac{1}{M} \sum_{m=1}^{M} |\mathbb{E}_k^M(\beta_k^M).p_k^m - y(t_k, W_{t_k}^m)|^2
\end{aligned}$$

en utilisant pour la dernière égalité que $\mathbb{E}_k^M(\beta_k^M)$ est la solution de

$$\inf_{\alpha} \frac{1}{M} \sum_{m=1}^{M} |y(t_k, W_{t_k}^m) - \alpha.p_k^m|^2.$$

Le deuxième terme du membre de droite se traite de la manière que $T_{1,k,M}$. Nous avons à présent deux alternatives pour relier $\mathbb{E}_k^M \frac{1}{M} \sum_{m=1}^{M} |\alpha_k^M.p_k^m - \mathbb{E}_k^M(\beta_k^M).p_k^M|^2$ à ϵ_{k+1}. Soit on utilise que

$$\begin{aligned}
\mathbb{E}_k^M \frac{1}{M} \sum_{m=1}^{M} |\alpha_k^M.p_k^m - \mathbb{E}_k^M(\beta_k^M).p_k^m|^2 = &\mathbb{E}_k^M \frac{1}{M} \sum_{m=1}^{M} |\mathbb{E}_k^M(\alpha_k^M).p_k^m - \mathbb{E}_k^M(\beta_k^M).p_k^m|^2 \\
&+ \mathbb{E}_k^M \frac{1}{M} \sum_{m=1}^{M} |\{\alpha_k^M - \mathbb{E}_k^M(\alpha_k^M)\}.p_k^m|^2
\end{aligned}$$

et on applique la propriété de contraction de la projection sur $\left((p_k^m)^*\right)_{1 \leq m \leq M}$ pour le premier terme du membre de droite. Par contre, comme on l'a déjà mentionné, pour contrôler le deuxième terme du membre de droite, on ne peut tirer parti de l'annulation des termes non-diagonaux de la matrice $\mathbb{E}_k^M\left(\{V - \mathbb{E}_k^M(V)\}\{V - \mathbb{E}_k^M(V)\}^*\right)$ où V est le vecteur de $\mathbb{R}^M$ de coordonnées $([\alpha_{k+1}^M.p_{k+1}^m]_y)_{1 \leq m \leq M}$. Or, dans le cas où le driver f ne dépend pas de Z, on peut simplement écrire :

$$\begin{aligned}
\mathbb{E} \frac{1}{M} \sum_{m=1}^{M} |\alpha_k^M.p_k^m - \mathbb{E}_k^M(\beta_k^M).p_k^m|^2 \leq &(1+h)\mathbb{E} \frac{1}{M} \sum_{m=1}^{M} |\alpha_k^M.p_k^m - \beta_k^M.p_k^m|^2 \\
&+ (1+h^{-1})\mathbb{E} \frac{1}{M} \sum_{m=1}^{M} |\{\beta_k^M - \mathbb{E}_k^M(\beta_k^M)\}.p_k^m|^2.
\end{aligned}$$

La propriété de contraction de la projection sur $\left((p_k^m)^*\right)_{1\leq m\leq M}$ permet de majorer le premier terme du membre de droite par

$$(1+h)\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|[\alpha_{k+1}^M.p_{k+1}^m]_y - y(t_{k+1},W_{t_{k+1}}^m)|^2 = (1+h)\epsilon_{k+1}.$$

Pour le second terme, on utilise les mêmes arguments que pour $T_{3,k,M}$. Ceci permet d'établir l'inégalité de la proposition. □

Ainsi, on voit qu'il est possible dans le cas où le driver ne dépend pas de Z d'établir des résultats similaires à ceux du théorème II.2 sans utiliser des incréments browniens supplémentaires. De là à penser que l'on peut appliquer l'algorithme décrit au début de cette partie dans le cas général où f peut dépendre de Z et sans utiliser de browniens supplémentaires, il n'y a qu'un pas que nous franchissons maintenant. Par manque de temps, nous donnons simplement une partie de la démonstration et l'intuition de pourquoi cela va marcher.

4.2 Analyse de l'algorithme sans incréments browniens supplémentaires

On va supposer ici que l'on a localisé également les incréments browniens. Ainsi, on suppose que $|\Delta W_{l,k}^m| + |\Delta\tilde{W}_{l,k}^m| \leq C_w(R)\sqrt{h}$. Remarquons que l'erreur de localisation en fonction de R se traite de la même façon que pour la proposition II.2.
Pour les raisons exposées plus haut, il est important de pouvoir établir une inégalité du type (II.2.20) ce qui est possible en introduisant des incréments browniens supplémentaires. Dans le cas où on appliquerait un algorithme sans browniens supplémentaires, on aurait $y_k^{N,R,M}(\cdot) = [\alpha_{0,k}^M.p_{0,k}]_y(\cdot)$ et $z_{l,k}^{N,R,M}(\cdot) = [\alpha_{l,k}^M.p_{l,k}]_z(\cdot)$ où $(\alpha_{l,k}^M)_{0\leq l\leq q}$ est défini par :

$$\begin{cases} \inf\limits_{\alpha_l} \dfrac{1}{M}\displaystyle\sum_{m=1}^{M}|y_{k+1}^{N,R,M}(P_{t_{k+1}}^{N,m})\frac{\Delta W_{l,k}^m}{h} - \alpha_l.p_{l,k}^m|^2, \\ \inf\limits_{\alpha_0} \dfrac{1}{M}\displaystyle\sum_{m=1}^{M}|y_{k+1}^{N,R,M}(P_{t_{k+1}}^{N,m}) + hf(t_k,X_{t_k}^{N,m},y_{k+1}^{N,R,M}(P_{t_{k+1}}^{N,m}),z_{l,k}^{N,R,M}(P_{t_k}^{N,m})) - \alpha_0.p_{0,k}^m|^2. \end{cases} \tag{II.4.1}$$

La différence avec (II.1.6-II.1.7) est donc que l'on a remplacé $(\tilde{P}_{t_{k+1}}^{N,m},\Delta\tilde{W}_k^m)_{1\leq m\leq M}$ par $(P_{t_{k+1}}^{N,m},\Delta W_k^m)_{1\leq m\leq M}$.
On va voir ci-dessous que les deux problèmes ne sont pas très éloignés et on va en particulier établir la proposition qui joue le rôle des inégalités (II.2.19-II.2.20) dans le cas de l'algorithme avec browniens auxiliaires :

Proposition II.5 *Pour $k \leq N-2$:*

$$
\begin{aligned}
&\mathbb{E}\frac{h}{M}\sum_{m=1}^{M}|\alpha_{l,k}^{M}.p_{l,k}^{m}-z_{l,k}^{N,R}(P_{t_k}^{N,m})|^2\\
\leq& C\frac{C_y(R)^2}{M}\mathbb{E}(K_{l,k}(\omega))+\inf_{\alpha}\mathbb{E}|\sqrt{h}z_{l,k}^{N,R}(P_{t_k}^{N})-\alpha.p_{l,k}(P_{t_k}^{N})|^2+Ch^{\beta}\\
&+C\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|[\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y-y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})|^2\\
&-C\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\tilde{\mathbb{E}}\{[\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y-y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})\}|^2\\
&+C\mathbb{E}\left(\mathcal{N}_2\big(\frac{h^{\frac{\beta}{2}}}{\sqrt{18C_w(R)^2K_{l,k}(\omega)}},[\mathcal{P}_{0,k+1}]_y,(P_{t_{k+1}}^{N,m})_{1\leq m\leq M},(\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M}\big)e^{-\frac{Mh^{\beta}}{36K_{l,k}(\omega)C_y(R)^2C_w(R)^2}}\right)
\end{aligned}
$$

où $\tilde{P}_{t_{k+1}}^{N,m}=T_k(P_{t_k}^{N,m},\Delta\tilde{W}_{k,k+1}^{m})$ avec $(\Delta\tilde{W}_{k,k+1}^{m})_{1\leq m\leq M}$ une copie i.i.d. de $(\Delta W_{k,k+1}^{m})_{1\leq m\leq M}$ et avec la notation $\tilde{\mathbb{E}}(\cdot)=\mathbb{E}(\cdot|(\Delta W_{k,k+1}^{m})_{1\leq m\leq M,0\leq k\leq N-1})$.

Remarque II.3 *On constate qu'on arrive à relier l'algorithme sans simulation d'incréments browniens auxiliaires à un algorithme où on simulerait pour la régression à l'instant t_k $(\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M}$ à l'aide de $(\Delta\tilde{W}_{k,k+1}^{m})_{1\leq m\leq M}$. Ceci nous permet d'appliquer une inégalité de Cauchy-Scwharz comme dans* (II.2.20)*, ce qui nous est utile pour assurer la bonne propagation de l'erreur dans l'algorithme. Le prix à payer est l'apparition du terme h^{β} et le dernier terme dans l'inégalité ci-dessus, qui étaient absents dans les majorations (II.2.19-II.2.20) que l'on obtenait dans l'algorithme avec simulation d'incréments browniens supplémentaires.*

PREUVE:
Remarquons que

$$z_{l,k}^{N,R}(P_{t_k}^{N,m})=\tilde{\mathbb{E}}(y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})\frac{\Delta\tilde{W}_{l,k}^{m}}{h})$$

où $(\Delta\tilde{W}_{l,k}^{m})_{1\leq m\leq M}$ est une copie indépendante de $(\Delta W_{l,k}^{m})_{1\leq m\leq M}$ et où $\tilde{\mathbb{E}}(\cdot)=\mathbb{E}\big(\cdot\,|(\Delta W_{k,k+1}^{m})_{1\leq m\leq M,0\leq k\leq N-1}\big)$. Ainsi, en introduisant le vecteur $\beta_{l,k}^{M}$ solution de

$$\inf_{\alpha}\frac{1}{M}\sum_{m=1}^{M}|y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})\frac{\Delta\tilde{W}_{l,k}^{m}}{h}-\alpha.p_{l,k}^{m}|^2$$

il vient :

$$\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\alpha_{l,k}^{M}.p_{l,k}^{m}-z_{l,k}^{N,R}(P_{t_k}^{N,m})|^2 =\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\alpha_{l,k}^{M}.p_{l,k}^{m}-\tilde{\mathbb{E}}(\beta_{l,k}^{M}).p_{l,k}^{m}|^2 + \mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\tilde{\mathbb{E}}(\beta_{l,k}^{M}).p_{l,k}^{m}-z_{l,k}^{N,R}(P_{t_k}^{N,m})|^2.$$

Le deuxième terme du membre de droite se traite avec les mêmes arguments que pour $T_{4,k,l,M}$. Pour l'autre terme, on a :

$$\frac{1}{M}\sum_{m=1}^{M}|\alpha_{l,k}^{M}.p_{l,k}^{m}-\tilde{\mathbb{E}}(\beta_{l,k}^{M}).p_{l,k}^{m}|^2 \leq\frac{2}{M}\sum_{m=1}^{M}|\tilde{\alpha}_{l,k}^{M}.p_{l,k}^{m}-\alpha_{l,k}^{M}.p_{l,k}^{m}|^2 + \frac{2}{M}\sum_{m=1}^{M}|\tilde{\alpha}_{l,k}^{M}.p_{l,k}^{m}-\tilde{\mathbb{E}}(\beta_{l,k}^{M}).p_{l,k}^{m}|^2$$

où $\tilde{\alpha}_{l,k}^{M}$ est la solution de

$$\inf_{\alpha}\frac{1}{M}\sum_{m=1}^{M}|[\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y\frac{\Delta\tilde{W}_{l,k}^{m}}{h}-\alpha.p_{l,k}^{m}|^2. \tag{II.4.2}$$

Le dernier terme de l'inégalité ci-dessus se traite comme suit :

$$\begin{aligned}
&\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\tilde{\alpha}_{l,k}^{M}.p_{l,k}^{m}-\tilde{\mathbb{E}}(\beta_{l,k}^{M}).p_{l,k}^{m}|^2\\
=&\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\tilde{\mathbb{E}}(\tilde{\alpha}_{l,k}^{M}).p_{l,k}^{m}-\tilde{\mathbb{E}}(\beta_{l,k}^{M}).p_{l,k}^{m}|^2+\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\tilde{\alpha}_{l,k}^{M}.p_{l,k}^{m}-\tilde{\mathbb{E}}(\tilde{\alpha}_{l,k}^{M}).p_{l,k}^{m}|^2\\
\leq&\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\tilde{\mathbb{E}}\{([\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y-y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m}))\frac{\Delta\tilde{W}_{l,k}^{m}}{h}\}|^2\\
&+\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\tilde{\alpha}_{l,k}^{M}.p_{l,k}^{m}-\tilde{\mathbb{E}}(\tilde{\alpha}_{l,k}^{M}).p_{l,k}^{m}|^2\\
\leq&h^{-1}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|[\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y-y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})|^2\\
&-h^{-1}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\tilde{\mathbb{E}}\{[\alpha_{0,k+1}^{M}.\tilde{p}_{0,k+1}^{m}]_y-y_{k+1}^{N,R}(\tilde{P}_{t_{k+1}}^{N,m})\}|^2\\
&+\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|\tilde{\alpha}_{l,k}^{M}.p_{l,k}^{m}-\tilde{\mathbb{E}}(\tilde{\alpha}_{l,k}^{M}).p_{l,k}^{m}|^2.
\end{aligned}$$

Le dernier terme se traite comme $T_{5,k,l,M}$ car $\alpha^M_{0,k+1}$ est indépendant de $(\Delta\tilde{W}^m_{k,k+1})_m$ et on a bien pu établir une inégalité de Cauchy-Schwarz de la même façon que pour (II.2.20) comme en témoignent les deux premiers termes du membre de droite de la dernière inégalité. Reste le terme $\frac{1}{M}\sum_{m=1}^M |\tilde{\alpha}^M_{l,k}.p^m_{l,k} - \alpha^M_{l,k}.p^m_{l,k}|^2$. En comparant les équations (II.4.1) et (II.4.2), on imagine sans peine que les coefficients $\tilde{\alpha}^M_{l,k}$ et $\alpha^M_{l,k}$ vont être proches. Ce qui est moins évident, c'est que l'on peut obtenir des bornes d'erreurs sans faire intervenir les moments d'ordre supérieur à 2 de $p_{l,k}$. Supposons sans perte de généralité (voir annexe A.5) que $P^M_{l,k} = \mathrm{Id}$. Alors

$$\begin{aligned}\frac{1}{M}\sum_{m=1}^M |\tilde{\alpha}^M_{l,k}.p^m_{l,k} - \alpha^M_{l,k}.p^m_{l,k}|^2 =&(\tilde{\alpha}^M_{l,k} - \alpha^M_{l,k}).P^M_{l,k}(\tilde{\alpha}^M_{l,k} - \alpha^M_{l,k})\\ =&\|\tilde{\alpha}^M_{l,k} - \alpha^M_{l,k}\|^2.\end{aligned}$$

On fixe β tel que $1 < \beta \leq 2$ et on adopte la notation $\mathbb{P}^M_k = \mathbb{P}(\cdot|(P^{N,m}_{t_k})_{1\leq m\leq M})$. Sous l'hypothèse $P^M_{l,k} = \mathrm{Id}$ on a

$$\begin{aligned}&\mathbb{P}^M_k\Big(\|\tilde{\alpha}^M_{l,k} - \alpha^M_{l,k}\|^2 \geq h^{\beta-1}\Big)\\
=&\mathbb{P}^M_k\Big(\sum_{i=1}^{K_{l,k}(\omega)} \big|\frac{1}{M}\sum_{m=1}^M p^m_{l,k,i}\{[\alpha^M_{0,k+1}.\tilde{p}^m_{0,k+1}]_y\frac{\Delta\tilde{W}^m_{l,k}}{h} - [\alpha^M_{0,k+1}.p^m_{0,k+1}]_y\frac{\Delta W^m_{l,k}}{h}\}\big|^2 \geq h^{\beta-1}\Big)\\
\leq&\sum_{i=1}^{K_{l,k}(\omega)} \mathbb{P}^M_k\Big(\big|\frac{1}{M}\sum_{m=1}^M p^m_{l,k,i}\{[\alpha^M_{0,k+1}.\tilde{p}^m_{0,k+1}]_y\frac{\Delta\tilde{W}^m_{l,k}}{h} - [\alpha^M_{0,k+1}.p^m_{0,k+1}]_y\frac{\Delta W^m_{l,k}}{h}\}\big|^2 \geq \frac{h^{\beta-1}}{K_{l,k}(\omega)}\Big)\\
\leq&\sum_{i=1}^{K_{l,k}(\omega)} \mathbb{P}^M_k\Big(\exists\psi\in[\mathcal{P}_{0,k+1}]_y : \big|\frac{1}{M}\sum_{m=1}^M p^m_{l,k,i}\{\psi(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - \psi(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}\big|^2 \geq \frac{h^{\beta-1}}{K_{l,k}(\omega)}\Big)\\
=&\sum_{i=1}^{K_{l,k}(\omega)} \mathbb{P}^M_k\Big(\exists\psi\in[\mathcal{P}_{0,k+1}]_y : \big|\frac{1}{M}\sum_{m=1}^M p^m_{l,k,i}U_m\{\psi(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - \psi(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}\big| \geq \sqrt{\frac{h^{\beta-1}}{K_{l,k}(\omega)}}\Big)\end{aligned}$$

où (U_m) est une suite de variables aléatoires i.i.d. (et indépendantes de toutes les autres variables aléatoires simulées) valant 1 ou -1 avec proba $\frac{1}{2}$. Introduisons alors une couverture $\mathcal{G}$ de $[\mathcal{P}_{0,k+1}]_y$ au sens où pour toute $\psi \in [\mathcal{P}_{0,k+1}]_y$, il existe $g_j \in \mathcal{G}$, bornée par $C_y(R)$ telle que

$$\frac{1}{2M}\sum_{m=1}^M \{|\psi(P^{N,m}_{t_{k+1}}) - g_j(P^{N,m}_{t_{k+1}})|^2 + |\psi(\tilde{P}^{N,m}_{t_{k+1}}) - g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2\} \leq \frac{h^\beta}{18C_w(R)^2K_{l,k}(\omega)}$$

où on rappelle que $C_w(R)$ est tel que $|\Delta W^m_{l,k}| + |\Delta\tilde{W}^m_{l,k}| \leq C_w(R)\sqrt{h}$. Notons que le cardinal de $\mathcal{G}$ est $\mathcal{N}_2\big(\frac{h^{\frac{\beta}{2}}}{\sqrt{18C_w(R)^2K_{l,k}(\omega)}}, [\mathcal{P}_{0,k+1}]_y, (P^{N,m}_{t_{k+1}})_{1\leq m\leq M}, (\tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}\big)$. Alors soit

$\psi \in [\mathcal{P}_{0,k+1}]_y$ et g_j vérifiant l'inégalité ci-dessus. On a

$$
\begin{aligned}
&\big|\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{\psi(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - \psi(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}\big| \\
\leq &\big|\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{\psi(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - g_j(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h}\}\big| \\
&+ \big|\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{g_j(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - g_j(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}\big| \\
&+ \big|\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{g_j(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h} - \psi(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}\big|
\end{aligned}
$$

Or on a

$$
\begin{aligned}
&\big|\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{\psi(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - g_j(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h}\}\big|^2 \\
=&\frac{1}{M}\sum_{m=1}^{M} |\psi(\tilde{P}^{N,m}_{t_{k+1}}) - g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 |\frac{\Delta\tilde{W}^m_{l,k}}{h}|^2 \\
\leq&\frac{2C_w(R)^2}{2hM}\sum_{m=1}^{M}\{\psi(\tilde{P}^{N,m}_{t_{k+1}}) - g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 + |\psi(P^{N,m}_{t_{k+1}}) - g_j(P^{N,m}_{t_{k+1}})|^2\} \\
\leq&\frac{h^{\beta-1}}{9K_{l,k}(\omega)}
\end{aligned}
$$

en utilisant que $\frac{1}{M}\sum_{m=1}^{M}|p^m_{l,k,i}|^2 = 1$ par hypothèse. Comme la même inégalité est valable pour le terme $\big|\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{g_j(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h} - \psi(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}\big|$, on obtient :

$$
\begin{aligned}
&\mathbb{P}^M_k\Big(\exists\psi\in[\mathcal{P}_{0,k+1}]_y : \big|\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{\psi(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - \psi(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}\big| \geq \sqrt{\frac{h^{\beta-1}}{K_{l,k}(\omega)}}\Big) \\
\leq&\mathbb{P}^M_k\Big(\exists j : \big|\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{g_j(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - g_j(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}\big| \geq \frac{1}{3}\sqrt{\frac{h^{\beta-1}}{K_{l,k}(\omega)}}\Big).
\end{aligned}
$$

Pour borner ce dernier terme, on conditionne en plus par rapport à $(\Delta W^m_{k,k+1}, \Delta \tilde{W}^m_{k,k+1})_{1\leq m\leq M}$. On note $\tilde{\mathbb{P}}^M_{k+1}$ cette nouvelle probabilité. Il vient :

$$\begin{aligned}
&\tilde{\mathbb{P}}^M_{k+1}\Big(\exists j : |\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{g_j(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - g_j(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}| \geq \frac{1}{3}\sqrt{\frac{h^{\beta-1}}{K_{l,k}(\omega)}}\Big)\\
\leq &\mathcal{N}_2\Big(\frac{h^{\frac{\beta}{2}}}{\sqrt{18C_w(R)^2K_{l,k}(\omega)}}, [\mathcal{P}_{0,k+1}]_y, (P^{N,m}_{t_{k+1}})_{1\leq m\leq M}, (\tilde{P}^{N,m}_{t_{k+1}})_{1\leq m\leq M}\Big)\\
&\max_j \tilde{\mathbb{P}}^M_{k+1}\Big(|\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{g_j(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - g_j(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}| \geq \frac{1}{3}\sqrt{\frac{h^{\beta-1}}{K_{l,k}(\omega)}}\Big)
\end{aligned}$$

Notons $X_m = p^m_{l,k,i} U_m\{g_j(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - g_j(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}$ alors $\tilde{\mathbb{E}}^M_{k+1}(X_m) = 0$ et $|X_m| \leq \frac{2C_y(R)C_w(R)}{\sqrt{h}}|p^m_{l,k,i}|$. En appliquant l'inégalité de Hoeffding, on obtient :

$$\begin{aligned}
&\tilde{\mathbb{P}}^M_{k+1}\Big(|\frac{1}{M}\sum_{m=1}^{M} p^m_{l,k,i} U_m\{g_j(\tilde{P}^{N,m}_{t_{k+1}})\frac{\Delta\tilde{W}^m_{l,k}}{h} - g_j(P^{N,m}_{t_{k+1}})\frac{\Delta W^m_{l,k}}{h}\}| \geq \frac{1}{3}\sqrt{\frac{h^{\beta-1}}{K_{l,k}(\omega)}}\Big)\\
\leq &2\exp\Big(-\frac{Mh^\beta}{36K_{l,k}(\omega)C_y(R)^2C_w(R)^2\frac{1}{M}\sum_{m=1}^{M}|p^m_{l,k,i}|^2}\Big)\\
= &2\exp\Big(-\frac{Mh^\beta}{36K_{l,k}(\omega)C_y(R)^2C_w(R)^2}\Big)
\end{aligned}$$

en utilisant l'hypothèse $\frac{1}{M}\sum_{m=1}^{M}|p^m_{l,k,i}|^2 = 1$ ($P^M_{l,k} = \mathrm{Id}$). On a donc réussi à traduire la proximité entre $\alpha_{l,k}$ et $\tilde{\alpha}_{l,k}$ sans utiliser les moments d'ordre supérieur à 2 de $p_{l,k}$.

□

Troisième partie

Des extensions

Chapitre 1

Vitesse de convergence dans le cas driver linéaire

Nous avons obtenu dans la partie I la vitesse

$$\max_{0\leq k\leq N} \mathbb{E}|Y_{t_k} - Y^N_{t_k}|^2 \leq C\sqrt{h}$$

pour l'approximation de Y par la solution de l'EDSR discrète définie par (I.2.2-I.2.3). Dans le cas où le driver f est linéaire en (y, z), il est possible d'obtenir une vitesse plus rapide pour l'approximation de Y par Y^N. Nous allons démontrer, en utilisant les idées développées dans [KH01], [GM05] et [KHP02], que l'on obtient la vitesse suivante :

$$|Y_{t_0} - Y^N_{t_0}| \leq Ch.$$

Ce résultat est dans l'esprit des travaux de [RM93] et [Pri99] où les auteurs considèrent un cadre plus général mais n'obtiennent pas de vitesse de convergence.

1.1 Cadre et hypothèses

Précisons tout d'abord les hypothèses sous lesquelles nous allons travailler :

Hypothèse III.1 *On suppose que $d = q = 1$, et que les fonctions $b(t, x)$ et $\sigma(t, x)$ définies dans l'hypothèse 0.1 sont bornées à dérivées bornées (que l'on note C^∞_b dans la suite). En outre, on suppose que la fonction $\phi(x)$, qui définit la condition terminale de l'EDSR, satisfait le contrôle*

$$|\phi(x)| \leq C \exp(C|x|).$$

Enfin, on suppose que σ est uniformément elliptique au sens où $|\sigma(t, x)| \geq \sigma_0 > 0$ $\forall(t, x)$.

Nous allons également supposer que l'EDSR (I.1.1) a un driver linéaire en (y, z) et une condition terminale dépendant uniquement de la valeur terminale X_T. En particulier, nous allons considérer l'EDSR suivante :

$$\begin{cases} -dY_t = -(\theta(t, X_t)Z_t + rY_t)dt - Z_t dW_t, \\ Y_T = \phi(X_T), \end{cases}$$

où $\theta(t,x) = \frac{b(t,x)-r}{\sigma(t,x)}$ est de classe C_b^∞ sous l'hypothèse III.1. De même que pour les parties précédentes, on simule X^N le schéma d'Euler de X et on approche l'EDSR à temps continu par

$$Y_{t_N}^N = \phi(X_T^N), \tag{III.1.1}$$
$$Y_{t_k}^N = \mathbb{E}_{t_k}(Y_{t_{k+1}}^N) - h\theta(t_k, X_{t_k}^N)Z_{t_k}^N - rhY_{t_k}^N, \tag{III.1.2}$$
$$hZ_{t_k}^N = \mathbb{E}_{t_k}(Y_{t_{k+1}}^N \Delta W_k). \tag{III.1.3}$$

1.2 Vitesse de convergence

1.2.1 Résultat principal

Nous allons donc nous efforcer de démontrer le théorème suivant :

Théorème III.1 *Sous l'hypothèse III.1, on a*

$$|Y_{t_0} - Y_{t_0}^N| \leq Ch.$$

Dans ce chapitre, on utilisera la fonction $\eta : [0,T] \to \{t_0, \cdots, t_N\}$ définie par

$$\eta(t) = \sum_{k=0}^{N-1} t_k 1_{[t_k, t_{k+1}[}(t) + T1_T(t).$$

PREUVE: Vu (III.1.2-III.1.3), on obtient :

$$Y_{t_k}^N = \mathbb{E}_{t_k}\big(Y_{t_{k+1}}^N \frac{1 - \theta(t_k, X_{t_k}^N)\Delta W_k}{1 + rh}\big)$$

ce qui en itérant sur k donne :

$$Y_{t_0}^N = \mathbb{E}\big(\frac{1}{(1+rh)^N}\phi(X_T^N)\xi_T^N\big)$$

où

$$\xi_T^N = \prod_{k=0}^{N-1}(1 - \theta(t_k, X_{t_k}^N)\Delta W_k).$$

L'intérêt d'écrire $Y_{t_0}^N$ sous cette forme est qu'en appliquant la proposition A.1, on obtient :

$$Y_t = \mathbb{E}\big(\phi(X_T)\Gamma_T^t|\mathcal{F}_t\big) \text{ p.s.}$$

où

$$d\Gamma_s^t = -\Gamma_s^t(rds + \theta(s, X_s)dW_s)\,, \Gamma_t^t = 1.$$

En résolvant cette dernière EDS, on obtient

$$Y_0 = \mathbb{E}\big(\exp(-rT)\phi(X_T)\xi_T\big)$$

où

$$\xi_T = \exp(-\int_0^T \theta(t, X_t)dW_t - \frac{1}{2}\int_0^T |\theta(t, X_t)|^2 dt\big).$$

On reconnait bien sûr la règle de valorisation risque-neutre à savoir que Y_0 s'exprime comme l'espérance du payoff actualisé sous la probabilité risque-neutre $\mathbb{Q}$ définie par $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_T} = \xi_T$. Ainsi, l'erreur $|Y_{t_0} - Y_{t_0}^N|$ se majore par la somme de trois termes :

$$\begin{aligned}
&\big|\mathbb{E}\big(\exp(-rT)\phi(X_T)\xi_T\big) - \mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T\big)\big| \\
+&\big|\mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T\big) - \mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T^N\big)\big| \\
+&\big|\mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T^N\big) - \mathbb{E}\big(\frac{1}{(1+rh)^N}\phi(X_T^N)\xi_T^N\big)\big|.
\end{aligned}$$

Nous allons borner chacun de ces trois termes.

Troisième terme. Ce terme se borne facilement en remarquant que

$$\begin{aligned}
\exp(-rT) - \frac{1}{(1+rh)^N} =& \exp(-rT) - \exp(-N\log(1+rh)) \\
=& \exp(-rT)\big(1 - \exp[rT - \frac{T}{h}(rh - \frac{r^2h^2}{2} + o(h^2))]\big) \\
=& Ch + o(h).
\end{aligned}$$

Les deux autres termes nécessitent des résultats intermédiaires.

1.2.2 Des résultats intermédiaires

Nous adoptons les notations de [Nua95] pour les espaces $\mathbb{D}^{k,p}$ du calcul de Malliavin.

Lemme III.1 *([GM05]) On considère une EDS de dimension d $(X_t)_{t\geq 0}$ définie par des coefficients C^∞ à dérivées bornées et $(X_t^N)_{t\geq 0}$ son schéma d'Euler :*

$$X_t = x + \int_0^t b(s, X_s)ds + \sum_{j=1}^q \int_0^t \sigma_j(s, X_s)dW_s^j,$$
$$X_t^N = x + \int_0^t b(\eta(s), X_{\eta(s)}^N)ds + \sum_{j=1}^q \int_0^t \sigma_j(\eta(s), X_{\eta(s)}^N)dW_s^j.$$

Alors, pour $1 \leq k \leq d$:

$$X_{k,t} - X_{k,t}^N = \sum_{i,j=0}^q c_{i,j,k}^{X,0}(t) \int_0^t c_{i,j,k}^{X,1}(s)\{\int_{\eta(s)}^s c_{i,j,k}^{X,2}(u)dW_u^i\}dW_s^j$$

avec la convention $dW_t^0 = dt$ et pour des processus adaptés $\{(c_{i,j,k}^{X,i_1}(t))_{t\geq 0} : 0 \leq i, j \leq q, 1 \leq k \leq d, 0 \leq i_1 \leq 2\}$ satisfaisant $\sup_{N,t} \|c_{i,j,k}^{X,i_1}(t)\|_{k',p} < \infty$ pour tout $k', p \geq 1$.

Nous aurons également besoin de la formule d'intégration par partie suivante ([BN95]) :

Proposition III.1 *Soit $F \in (\mathbb{D}^\infty)^m$ tel que sa matrice de Malliavin γ_F soit inversible p.s. et satisfasse $\gamma_F^{-1} \in \cap_{p\geq 1}\mathbf{L}^p$. Soit $\phi \in C_b^\infty(\mathbb{R}^m)$ et $G \in \mathbb{D}^\infty$. Alors $(\det \gamma_F)^{-1} \in \mathbb{D}^\infty$ et pour tout multi-indice α, il existe une variable aléatoire $H_\alpha(F, G) \in \mathbb{D}^\infty$ telle que :*

$$\mathbb{E}\big(G(\partial_x^\alpha \phi)(F)\big) = \mathbb{E}\big(\phi(F)H_\alpha(F, G)\big).$$

De plus, pour tout $p > 1$ et tout multi-indice α, il existe une constante $C(p, \alpha)$, des entiers naturels n_1, n_2 et des indices k, d, d', b, b' dépendant également de p et α tels que

$$\|H_\alpha(F, G)\|_p \leq C(p, \alpha)\big(\|\gamma_F^{-1}\|_k^{n_1}\|F\|_{d,b}^{n_2}\|G\|_{d',b'}\big).$$

Le théorème suivant (théorème 52 page 266 dans [Pro90]), qui exprime sous forme explicite les solutions d'EDS linéaires, nous servira également :

Théorème III.2 *Soit H une semimartingale et Z une semimartingale continue telle que $Z_0 = 0$. Alors la solution $\mathcal{E}_H(Z)$ de l'équation*

$$X_t = H_t + \int_0^t X_{s-}dZ_s$$

est donnée par

$$\mathcal{E}_H(Z)_t = \mathcal{E}(Z)_t\{H_0 + \int_{0+}^t \mathcal{E}(Z)_s^{-1}d(H_s - [H, Z]_s)\}$$

où $\mathcal{E}(Z)_t = \exp(Z_t - \frac{1}{2}[Z, Z]_t)$.

Enfin, nous avons besoin du lemme suivant :

Lemme III.2 *Pour $q, p > 1$, on a*

$$\sup_t \|X_t\|_{q,p} + \sup_t \|\xi_t\|_{q,p} + \sup_{t,N} \|X_t^N\|_{q,p} + \sup_{k,N} \|\xi_{t_k}^N\|_{q,p} < \infty$$

avec $(\xi_t^N)_t$ défini par

$$\xi_t^N = 1 - \int_0^t \xi_{\eta(s)}^N \theta(\eta(s), X_{\eta(s)}^N) dW_s.$$

PREUVE: Le résultat pour les trois premiers termes découle facilement du fait que b, σ et θ sont bornées à dérivées bornées. Pour le dernier terme, on pourrait être tenté de voir le vecteur $\begin{pmatrix} X^N \\ \xi^N \end{pmatrix}$ comme le schéma d'Euler de l'EDS $\begin{pmatrix} X \\ \xi \end{pmatrix}$ et appliquer alors directement des résultats classiques pour en déduire des contrôles dans $\mathbb{D}^{q,p}$. Mais les coefficients de l'EDS $\begin{pmatrix} X \\ \xi \end{pmatrix}$ ne sont pas à dérivées bornées (à cause de ξ) et une preuve particulière doit être effectuée en utilisant la bornitude de θ.
Tout d'abord, montrons que $\sup_{k \le N, N} \|\xi_{t_k}^N\|_p < \infty$ pour $p \ge 1$. Vu la définition de ξ_t^N, on a pour $1 \le i \le N$:

$$\xi_{t_i}^N = \xi_{t_{i-1}}^N (1 - \theta(t_{i-1}, X_{t_{i-1}}^N) \Delta W_{i-1}). \tag{III.1.4}$$

Ainsi, on obtient pour $p \ge 1$:

$$\begin{aligned}
\mathbb{E}|\xi_{t_i}^N|^{2p} =& \mathbb{E}\left(|\xi_{t_{i-1}}^N|^{2p} \mathbb{E}_{t_{i-1}} \{ \sum_{k=0}^{2p} C_{2p}^k (-1)^{2p-k} \theta(t_{i-1}, X_{t_{i-1}}^N)^k \Delta W_{i-1}^k \} \right) \\
=& \mathbb{E}\left(|\xi_{t_{i-1}}^N|^{2p} \mathbb{E}_{t_{i-1}} \{ \sum_{k=0}^{p} C_{2p}^{2k} (-1)^{2(p-k)} \theta(t_{i-1}, X_{t_{i-1}}^N)^{2k} \Delta W_{i-1}^{2k} \} \right) \\
\le& (1 + C_p h) \mathbb{E}(|\xi_{t_{i-1}}^N|^{2p})
\end{aligned}$$

pour h suffisamment petit et en utilisant pour la dernière inégalité la bornitude de θ. On en déduit alors facilement que $\sup_{k \le N, N} \|\xi_{t_k}^N\|_p < \infty$.
Ensuite, considérons u tel que $t_{k-1} < u \le t_k$. En tant que fonctionnelle régulière des accroissements browniens, chaque $\xi_{t_i}^N \in \mathbb{D}^\infty$. En différentiant (III.1.4), on obtient pour $i \ge k+1$:

$$D_u \xi_{t_i}^N = D_u \xi_{t_{i-1}}^N - D_u[\xi_{t_{i-1}}^N \theta(t_{i-1}, X_{t_{i-1}}^N)] \Delta W_{i-1}.$$

En itérant on obtient pour $i \geq k+1$:

$$D_u\xi^N_{t_i} = D_u\xi^N_{t_k} - \sum_{j=k}^{i-1} D_u[\xi^N_{t_j}\theta(t_j, X^N_{t_j})]\Delta W_j.$$

En appliquant l'inégalité de Burkholder-Davis-Gundy à la martingale $\sum_j D_u[\xi^N_{t_j}\theta(t_j, \xi^N_{t_j})]\Delta W_j$, il vient :

$$\begin{aligned}
\mathbb{E}|D_u\xi^N_{t_i}|^p \leq & C_p\mathbb{E}|D_u\xi^N_{t_k}|^p + C_p\mathbb{E}\Big|\sum_{j=k}^{i-1} D_u[\xi^N_{t_j}\theta(t_j, X^N_{t_j})]\Delta W_j\Big|^p \\
\leq & C_p\mathbb{E}|D_u\xi^N_{t_k}|^p + C_p\mathbb{E}\Big|h\sum_{j=k}^{i-1} |D_u[\xi^N_{t_j}\theta(t_j, X^N_{t_j})]|^2\Big|^{\frac{p}{2}} \\
\leq & C_p\mathbb{E}|D_u\xi^N_{t_k}|^p + C_p h\mathbb{E}\sum_{j=k}^{i-1} |D_u[\xi^N_{t_j}\theta(t_j, X^N_{t_j})]|^p \\
\leq & C_p(1+\mathbb{E}|D_u\xi^N_{t_k}|^p) + C_p h\sum_{j=k+1}^{i-1} \mathbb{E}|D_u\xi^N_{t_j}|^p
\end{aligned}$$

car θ est bornée, $\|\xi^N_{t_j}\|_q$ et $\|X^N_{t_j}\|_{1,q}$ sont, pour tout $q \geq 1$, finis uniformément en N et j. En appliquant le lemme de Gronwall discret, on obtient :

$$\max_{k\leq i\leq N} \mathbb{E}|D_u\xi^N_{t_i}|^p \leq C_p(1+\mathbb{E}|D_u\xi^N_{t_k}|^p).$$

Par ailleurs :

$$D_u\xi^N_{t_k} = -\xi^N_{t_{k-1}}\theta(t_{k-1}, X^N_{t_{k-1}})$$

d'où on déduit

$$\max_{k\leq i\leq N} \mathbb{E}|D_u\xi^N_{t_i}|^p \leq C_p(1+\mathbb{E}|\xi^N_{t_{k-1}}|^p).$$

Les dérivées d'ordre supérieur se traitent de la même façon.

□

1.2.3 Preuve du théorème III.1

Nous pouvons à présent continuer la démonstration du théorème III.1 et borner les deux autres termes de la somme.

Premier terme

Définissons la norme $|||\cdot|||_2$ par :

$$|||\phi|||_2 = \sqrt{\mathbb{E}|\phi(X_T)|^2} + \sqrt{\mathbb{E}|\phi(X_T^N)|^2} + \big(\mathbb{E}\int_0^1 |\phi(X_T^{N,\lambda})|^2 d\lambda\big)^{\frac{1}{2}}$$

où $X_t^{N,\lambda} = \lambda X_t + (1-\lambda)X_t^N$. Nous allons tout d'abord supposer sans restriction que ϕ est de classe C^∞ à support compact (C_0^∞), à condition de pouvoir montrer que

$$\big|\mathbb{E}\big(\exp(-rT)\phi(X_T)\xi_T\big) - \mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T\big)\big| \leq C|||\phi|||_2 h$$

avec ϕ de classe C_0^∞. En effet, le cas de ϕ mesurable s'en déduit en considérant une suite de fonctions ϕ_m dans C_0^∞ qui tend vers ϕ pour la norme $|||\cdot|||_2$. Cette hypothèse effectuée, on peut réécrire l'erreur comme

$$\begin{aligned}&\mathbb{E}\big(\exp(-rT)\phi(X_T)\xi_T\big) - \mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T\big)\\ &= \int_0^1 \mathbb{E}\big(\exp(-rT)\xi_T\partial_x\phi(X_T^{N,\lambda})(X_T - X_T^N)\big)d\lambda.\end{aligned}$$

En utilisant le lemme III.1, on obtient

$$\begin{aligned}&\mathbb{E}\big(\exp(-rT)\xi_T\partial_x\phi(X_T^{N,\lambda})(X_T - X_T^N)\big)\\ &= \sum_{i,j=0}^{1} \mathbb{E}\big(\exp(-rT)\xi_T\partial_x\phi(X_T^{N,\lambda})c_{i,j}^{X,0}(T)\int_0^T c_{i,j}^{X,1}(s)\big(\int_{\eta(s)}^s c_{i,j}^{X,2}(u)dW_u^i\big)dW_s^j\big).\end{aligned}$$

Nous allons traiter le terme correspondant à $i=j=1$. En utilisant la formule de dualité ([Nua95]) et le théorème de Fubini, on obtient :

$$\begin{aligned}&\mathbb{E}\big(\exp(-rT)\xi_T\partial_x\phi(X_T^{N,\lambda})c_{1,1}^{X,0}(T)\int_0^T c_{1,1}^{X,1}(s)\big(\int_{\eta(s)}^s c_{1,1}^{X,2}(u)dW_u\big)dW_s\big)\\ &= \int_0^T \mathbb{E}\big(D_s\{\exp(-rT)\xi_T\partial_x\phi(X_T^{N,\lambda})c_{1,1}^{X,0}(T)\}c_{1,1}^{X,1}(s)\big(\int_{\eta(s)}^s c_{1,1}^{X,2}(u)dW_u\big)\big)ds\\ &= \int_0^T\int_{\eta(s)}^s \mathbb{E}\Big(D_u\big[D_s\{\exp(-rT)\xi_T\partial_x\phi(X_T^{N,\lambda})c_{1,1}^{X,0}(T)\}c_{1,1}^{X,1}(s)\big]c_{1,1}^{X,2}(u)\Big)duds.\end{aligned}$$

Le terme $\mathbb{E}\big(D_u[D_s\{\exp(-rT)\xi_T\partial_x\phi(X_T^{N,\lambda})c_{1,1}^{X,0}(T)\}c_{1,1}^{X,1}(s)]c_{1,1}^{X,2}(u)\big)$ peut se réécrire sous la forme d'une somme de termes $\mathbb{E}\big(\partial_x^\alpha\phi(X_T^{N,\lambda})G\big)$ que l'on peut majorer en utilisant la proposition III.1 et l'inégalité de Cauchy-Schwarz par $C\|\phi(X_T^{N,\lambda})\|_2$. Remarquons que l'on peut utiliser la proposition III.1 car $(X_t^{N,\lambda})_t$ est un processus d'Itô uniformément elliptique ([KHP02]). Ici, la dimension $d=1$ est cruciale et évite des détours comme dans [GM05]. Ainsi, on obtient finalement que

$$\mathbb{E}\big(\exp(-rT)\xi_T\partial_x\phi(X_T^{N,\lambda})c_{1,1}^{X,0}(T)\int_0^T c_{1,1}^{X,1}(s)\big(\int_{\eta(s)}^s c_{1,1}^{X,2}(u)dW_u\big)dW_s\big) \leq C\|\phi(X_T^{N,\lambda})\|_2 h.$$

Les autres termes se traitent de même pour finalement conclure à

$$\left|\mathbb{E}\big(\exp(-rT)\phi(X_T)\xi_T\big) - \mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T\big)\right| \leq C|||\phi|||_2 h.$$

Deuxième terme. Le deuxième terme

$$\left|\mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T\big) - \mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T^N\big)\right|$$

exprime l'approximation d'une dérivée de Radon ξ_T par une variable aléatoire d'espérance 1 mais pouvant prendre des valeurs négatives ξ_T^N. Comme pour le premier terme, nous pouvons supposer que ϕ est dans C_0^∞ à condition de montrer que

$$\left|\mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T\big) - \mathbb{E}\big(\exp(-rT)\phi(X_T^N)\xi_T^N\big)\right| \leq Ch|||\phi|||_2.$$

Nous allons exprimer la différence $\xi_T - \xi_T^N$ à l'aide du théorème III.2. Remarquons tout d'abord que

$$d\xi_t = -\xi_t\theta(t, X_t)dW_t$$

et que ξ_T^N est la valeur terminale de l'équation d'Itô

$$\xi_t^N = 1 - \int_0^t \xi_{\eta(s)}^N \theta(\eta(s), X_{\eta(s)}^N)dW_s.$$

On peut donc réexprimer la différence comme

$$\begin{aligned}
\xi_t - \xi_t^N =& \int_0^t \xi_{\eta(s)}^N \theta(\eta(s), X_{\eta(s)}^N)dW_s - \int_0^t \xi_s\theta(s, X_s)dW_s \\
=& -\int_0^t (\xi_s - \xi_s^N)\theta(s, X_s)dW_s + \int_0^t [\xi_{\eta(s)}^N\theta(\eta(s), X_{\eta(s)}^N) - \xi_s^N\theta(s, X_s)]dW_s \\
=& \mathcal{E}(\alpha)_t \int_0^t \mathcal{E}(\alpha)_s^{-1}\epsilon_s^N[dW_s + \theta(s, X_s)ds]
\end{aligned}$$

où $\alpha_t = -\int_0^t \theta(s, X_s)dW_s$ (voir le théorème III.2 pour la notation $\mathcal{E}(\cdot)$) et $\epsilon_s^N = \xi_{\eta(s)}^N\theta(\eta(s), X_{\eta(s)}^N) - \xi_s^N\theta(s, X_s)$. On peut alors réécrire le terme d'erreur ϵ_s^N sous la forme :

$$\begin{aligned}
\epsilon_s^N =& \xi_{\eta(s)}^N\theta(\eta(s), X_{\eta(s)}^N) - \xi_s^N\theta(s, X_s) \\
=& \xi_{\eta(s)}^N\theta(\eta(s), X_{\eta(s)}^N) - \xi_{\eta(s)}^N\big[1 - \theta(\eta(s), X_{\eta(s)}^N)(W_s - W_{\eta(s)})\big]\theta(s, X_s) \\
=& \xi_{\eta(s)}^N[\theta(\eta(s), X_{\eta(s)}^N) - \theta(s, X_s)] + \xi_{\eta(s)}^N\theta(\eta(s), X_{\eta(s)}^N)\theta(s, X_s)(W_s - W_{\eta(s)}) \\
=& \xi_{\eta(s)}^N[\theta(\eta(s), X_{\eta(s)}^N) - \theta(\eta(s), X_{\eta(s)})] + \xi_{\eta(s)}^N[\theta(\eta(s), X_{\eta(s)}) - \theta(s, X_s)] + \\
& \xi_{\eta(s)}^N\theta(\eta(s), X_{\eta(s)}^N)\theta(s, X_s)(W_s - W_{\eta(s)}).
\end{aligned}$$

On va pouvoir alors appliquer la même technique que pour le premier terme, combinant théorème de Fubini, formule de dualité et d'intégration par partie, à condition que $\xi_{t_k}^N \in$

$\mathbb{D}^\infty$ et que $\sup_{N,k} \|\xi^N_{t_k}\|_{k',p} < \infty$ pour $k', p > 1$. C'est exactement le résultat du Lemme III.2. Il découle finalement

$$\begin{aligned}\left|\mathbb{E}\big(\exp(-rT)\phi(X^N_T)\xi_T\big) - \mathbb{E}\big(\exp(-rT)\phi(X^N_T)\xi^N_T\big)\right| \leq & Ch\|\phi(X^N_T)\|_2 \\ \leq & Ch|||\phi|||_2.\end{aligned}$$

□

Chapitre 2

Algorithme pour les équations différentielles stochastiques rétrogrades réfléchies

2.1 Le cadre et les hypothèses

On considère toujours la même dynamique pour X

$$X_{i,t} = X_{i,0} + \int_0^t b_i(s, X_s)ds + \int_0^t \sigma_i(s, X_s)dW_s,\ 1 \leq i \leq d$$

et on renforce l'hypothèse 0.1 :

Hypothèse III.2 *L'hypothèse 0.1 est vérifiée. En outre, les fonctions b et σ sont continues, différentiables avec des dérivées spatiales bornées uniformément. Enfin, on suppose une condition d'uniforme ellipticité sur $\sigma\sigma^*$, i.e. que $\sigma\sigma^* \geq \epsilon_0 \mathrm{Id}$ avec $\epsilon_0 > 0$.*

En effet, nous aurons besoin de travailler avec un processus d'Itô uniformément elliptique pour appliquer les résultats de [MZ05]. Nous supposons également que l'hypothèse I.1 est vérifiée.
Dans cette partie, nous changeons l'EDSR (I.1.1) et allons désormais chercher un triplet (Y, Z, K) qui vérifie :

$$\begin{aligned} &Y_t = g(T, X_T) + \int_t^T f(s, S_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s + K_T - K_t, \\ &Y_t \geq g(t, X_t),\ t \geq 0,\ \int_0^T (Y_t - g(t, X_t))dK_t = 0. \end{aligned} \tag{III.2.1}$$

Nous nous référons à [EKKP$^+$97] pour les bons espaces dans lesquels chercher un tel triplet (Y, Z, K) solution, ainsi que pour des résultats d'existence et d'unicité (voir également

l'annexe A.2). L'interprétation financière est la suivante. On désire toujours construire un portefeuille de valeur $(Y_t)_t$ et où $(Z_t)_t$ est relié à la position en actif S dans le portefeuille. Cette fois, on impose en plus que la valeur du portefeuille soit toujours supérieure à un plancher $g(t, X_t)$. C'est exactement la description d'une option américaine où par exemple, dans le cas d'un put américain, on aurait $g(t,x) = (K - e^x)_+$.
Nous allons voir dans le prochain chapitre qu'il est très simple de modifier les algorithmes proposés dans les parties I et II pour tenir compte de la contrainte de plancher.
Etant donné que le second algorithme et ses techniques de preuve conduisent à des vitesses de convergence plus explicites que dans la partie I, on va plutôt exposer la modification de cet algorithme pour tenir compte du plancher. L'adaptation du premier algorithme présenté, avec itérations de Picard, est similaire. Les hypothèses sur l'obstacle g sont les suivantes

Hypothèse III.3 *$g(\cdot,\cdot)$ est Höldérienne d'indice $\frac{1}{2}$ en temps et uniformément Lipschitz en la variable spatiale.*

Remarquons que pour notre exemple du put américain en dimension 1, cette hypothèse est bien vérifiée étant donné que l'obstacle $(K - e^x)^+$ est nul pour $x \geq \log(K)$.

2.2 L'algorithme

On définit g^R de manière analogue à $\phi^{N,R}$ et f^R dans la partie II. Les bornes $C_y(R)$ et $C_z(R)$ sont définies comme précédemment. Un calcul similaire à celui de la proposition II.1 permet de fixer leurs valeurs explicitement.
L'algorithme se présente toujours sous une forme rétrograde :

→ On initialise pour $k = N$ par $y_N^{N,R,M}(\cdot) = g^R(t_N, \cdot)$.

→ A un instant de discrétisation t_k, $(\alpha_{l,k}^M)_{1\leq l\leq q}$ est défini comme un minimiseur du problème de moindres carrés :

$$\inf_{\alpha_l} \frac{1}{M}\sum_{m=1}^{M} |y_{k+1}^{N,R,M}(\tilde{X}_{t_{k+1}}^{N,m})\frac{\Delta\tilde{W}_{l,k}^m}{h} - \alpha_l.p_{l,k}^m|^2$$

avec $\tilde{X}_{t_{k+1}}^{N,m}$ simulé, de la même façon que $\tilde{P}_{t_{k+1}}^{N,m}$ dans la partie II, à partir de $X_{t_k}^{N,m}$ et l'incrément brownien auxiliaire $\Delta\tilde{W}_k^m$ par $\tilde{X}_{t_{k+1}}^{N,m} = X_{t_k}^{N,m} + b(t_k, X_{t_k}^{N,m})h + \sigma(t_k, X_{t_k}^{N,m})\Delta\tilde{W}_k^m$. On définit alors $z_{l,k}^{N,R,M}(\cdot)$ par $z_{l,k}^{N,R,M}(\cdot) = [\alpha_{l,k}^M.p_{l,k}]_z(\cdot)$. Puis $\alpha_{0,k}^M$ est défini comme un minimiseur du problème de moindres carrés :

$$\inf_{\alpha_0} \frac{1}{M}\sum_{m=1}^{M} |y_{k+1}^{N,R,M}(\tilde{X}_{t_{k+1}}^{N,m}) + hf(t_k, X_{t_k}^{N,m}, y_{k+1}^{N,R,M}(\tilde{X}_{t_{k+1}}^{N,m}), z_{l,k}^{N,R,M}(X_{t_k}^{N,m})) - \alpha_0.p_{0,k}^m|^2.$$

On définit alors $y_k^{N,R,M}(\cdot)$ par $y_k^{N,R,M}(\cdot) = \max\left([\alpha_{0,k}^M.p_{0,k}]_y(\cdot), g^R(t_k,\cdot)\right)$.

$\rightarrow$ On itère jusqu'à l'instant t_0.

On peut bien sûr imaginer le même algorithme sans browniens auxiliaires (même remarque que pour la section 4.2, partie II). Ainsi, l'algorithme est quasiment semblable à celui décrit dans la partie II, à ceci près qu'à chaque instant, on force notre approximation $y_k^{N,R,M}$ à être supérieure au plancher $g^R(t_k,\cdot)$. Le même type d'algorithme a déjà été proposé et analysé dans [BPP05] et [BT04] mais en utilisant des techniques différentes de quantification pour le premier article et de calcul de Malliavin pour le second. Dans ces deux articles, le cas d'un driver dépendant de Z n'est pas envisagé alors que nous autorisons cette possibilité ici. Nous proposerons également d'autres types d'algorithmes pour les EDSR réfléchies dans la partie numérique mais avec une analyse plus empirique.
Naturellement, on peut se poser la question de l'erreur commise entre $(y^{N,R,M}(X^N), z^{N,R,M}(X^N))$ et (Y^N, Z^N), une discrétisation en temps de (III.2.1) que nous décrivons maintenant.

2.3 Erreur de localisation

De manière analogue à (II.1.1-II.1.2), on définit une discrétisation (Y^N, Z^N) de (III.2.1) de manière rétrograde :

$\rightarrow$ On pose $Y_{t_N}^N = g(t_N, X_{t_N}^N)$

$\rightarrow$ A un instant t_k, on définit $Z_{t_k}^N$ par

$$hZ_{t_k}^N = \mathbb{E}_{t_k}(Y_{t_{k+1}}^N \Delta W_k), \tag{III.2.2}$$

puis on définit $\tilde{Y}_{t_k}^N$ tel que

$$\tilde{Y}_{t_k}^N = \mathbb{E}_{t_k}(Y_{t_{k+1}}^N) + h\mathbb{E}_{t_k} f(t_k, X_{t_k}^N, Y_{t_{k+1}}^N, Z_{t_k}^N), \tag{III.2.3}$$

et enfin on pose $Y_{t_k}^N = \max\left(g(t_k, X_{t_k}^N), \tilde{Y}_{t_k}^N\right)$.

On définit également une version localisée $(Y^{N,R}, Z^{N,R})$ en remplaçant g et f par g^R et f^R .

A cause de la propriété de Markov de X^N on a bien sûr $Y_{t_k}^{N,R} = y_k^{N,R}(X_{t_k}^N)$, $\tilde{Y}_{t_k}^{N,R} = \tilde{y}_k^{N,R}(X_{t_k}^N)$ et $Z_{t_k}^{N,R} = z_k^{N,R}(X_{t_k}^N)$.

De même que pour la proposition II.2, on peut établir une borne pour l'erreur commise en localisant l'EDSR réfléchie.

Proposition III.2

$$\begin{aligned}
&\max_{0\le k\le N}\mathbb{E}|Y^N_{t_k}-Y^{N,R}_{t_k}|^2+h\mathbb{E}\sum_{k=0}^{N-1}|Z^N_{t_k}-Z^{N,R}_{t_k}|^2\\
\le& C\sum_{k=0}^{N-1}\mathbb{E}\big[\mathbf{1}_{g(t_k,X^N_{t_k})\neq g^R(t_k,X^N_{t_k})}\{|Y^N_{t_k}|^2+C_y(R)^2\}\big]+C\mathbb{E}(|g(t_N,X^N_{t_N})-g^R(t_N,X^N_{t_N})|^2)\\
&+Ch\mathbb{E}\sum_{k=0}^{N-1}|f(t_k,S^N_{t_k},Y^N_{t_{k+1}},Z^N_{t_k})-f^R(t_k,S^N_{t_k},Y^N_{t_{k+1}},Z^N_{t_k})|^2.
\end{aligned}$$

PREUVE:

$$\begin{aligned}
&\mathbb{E}|Y^N_{t_k}-Y^{N,R}_{t_k}|^2\\
=&\mathbb{E}\big[|g(t_k,X^N_{t_k})\vee\tilde{Y}^N_{t_k}-g^R(t_k,X^N_{t_k})\vee\tilde{Y}^{N,R}_{t_k}|^2\mathbf{1}_{g(t_k,X^N_{t_k})=g^R(t_k,X^N_{t_k})}\big]\\
&+\mathbb{E}\big[|Y^N_{t_k}-Y^{N,R}_{t_k}|^2\mathbf{1}_{g(t_k,X^N_{t_k})\neq g^R(t_k,X^N_{t_k})}\big]\\
\le&\mathbb{E}\big[|\tilde{Y}^N_{t_k}-\tilde{Y}^{N,R}_{t_k}|^2\mathbf{1}_{g(t_k,X^N_{t_k})=g^R(t_k,X^N_{t_k})}\big]+2\mathbb{E}\big[(|Y^N_{t_k}|^2+C_y(R)^2)\mathbf{1}_{g(t_k,X^N_{t_k})\neq g^R(t_k,X^N_{t_k})}\big]\\
\le&\mathbb{E}\big[|\tilde{Y}^N_{t_k}-\tilde{Y}^{N,R}_{t_k}|^2\big]+2\mathbb{E}\big[(|Y^N_{t_k}|^2+C_y(R)^2)\mathbf{1}_{g(t_k,X^N_{t_k})\neq g^R(t_k,X^N_{t_k})}\big].
\end{aligned}$$

Le terme $\mathbb{E}\big[|\tilde{Y}^N_{t_k}-\tilde{Y}^{N,R}_{t_k}|^2\big]$ ainsi que le reste de la démonstration se traitent alors de la même manière que pour la proposition II.2. □

2.4 Erreur dûe aux simulations et projections

Il est facile de voir que les théorèmes II.2, II.3, II.4 et II.5 sont toujours valables. En effet, il suffit de remarquer que $|a\vee b-c\vee b|\le|a-c|$ et donc :

$$\begin{aligned}
&\frac{1}{M}\sum_{m=1}^{M}|y^{N,R}_k(X^{N,m}_{t_k})-y^{N,R,M}_k(X^{N,m}_{t_k})|^2\\
=&\frac{1}{M}\sum_{m=1}^{M}|g^R(t_k,X^{N,m}_{t_k})\vee\tilde{y}^{N,R}_k(X^{N,m}_{t_k})-g^R(t_k,X^{N,m}_{t_k})\vee[\alpha^M_{0,k}.p_{0,k}]_y(X^{N,m}_{t_k})|^2\\
\le&\frac{1}{M}\sum_{m=1}^{M}|\tilde{y}^{N,R}_k(X^{N,m}_{t_k}))-[\alpha^M_{0,k}.p_{0,k}]_y(X^{N,m}_{t_k})|^2
\end{aligned}$$

et l'analyse réalisée dans la partie II est valable pour ce dernier terme.

2.5 Erreur de discrétisation

Il reste à établir la convergence de (Y^N, Z^N) vers (Y, Z). Pour cela, on va comparer (Y^N, Z^N) à l'équation rétrograde de solution $(Y^{N,1}, Z^{N,1})$, introduite dans [MZ05] :

→ On initialise en t_N par $Y^{N,1}_{t_N} = g(t_N, X^N_{t_N})$.

→ Puis pour $t \in [t_k, t_{k+1}[$, on résout l'équation rétrograde non réfléchie de solution $(\tilde{Y}^{N,1}, Z^{N,1})$:

$$\tilde{Y}^{N,1}_t = Y^{N,1}_{t_{k+1}} + \int_t^{t_{k+1}} f(s, X^N_s, \tilde{Y}^{N,1}_s, Z^{N,1}_s)ds - \int_t^{t_{k+1}} Z^{N,1}_s dW_s. \tag{III.2.4}$$

→ On définit ensuite pour tout t dans $[t_k, t_{k+1}[$, $Y^{N,1}_t = \tilde{Y}^{N,1}_t \vee g(t, X^N_t)$.

→ On itère jusqu'à la date $t_0 = 0$.

Proposition III.3 *Pour h suffisamment petit, on a*

$$\max_{0\le k\le N} \mathbb{E}(|Y^N_{t_k} - Y^{N,1}_{t_k}|^2) + \sum_{k=0}^{N-1} \mathbb{E}\int_{t_k}^{t_{k+1}} |Z^{N,1}_t - Z^N_{t_k}|^2 dt \le Ch + C\mathbb{E}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} |Z^{N,1}_s - \overline{Z}^N_{t_k}|^2 ds$$

où $h\overline{Z}^N_{t_k} = \mathbb{E}_{t_k}\int_{t_k}^{t_{k+1}} Z^{N,1}_s ds$.

PREUVE: En utilisant $|a \vee b - c \vee b| \le |a - c|$ et en appliquant comme toujours une inégalité de Young à partir des équations (III.2.3-III.2.4), on obtient aisément :

$$\begin{aligned}
&\mathbb{E}|Y^N_{t_k} - Y^{N,1}_{t_k}|^2 \\
\le&\mathbb{E}|\tilde{Y}^N_{t_k} - \tilde{Y}^{N,1}_{t_k}|^2 \\
\le&(1+\gamma h)\mathbb{E}|\mathbb{E}_{t_k}(Y^N_{t_{k+1}} - \tilde{Y}^{N,1}_{t_{k+1}})|^2 \\
&+ h(1+\frac{1}{\gamma h})\mathbb{E}\int_{t_k}^{t_{k+1}} |f(t_k, X^N_{t_k}, Y^N_{t_{k+1}}, Z^N_{t_k}) - f(s, X^N_s, \tilde{Y}^{N,1}_s, Z^{N,1}_s)|^2 ds \\
\le&(1+\gamma h)\mathbb{E}|\mathbb{E}_{t_k}(Y^N_{t_{k+1}} - \tilde{Y}^{N,1}_{t_{k+1}})|^2 + C(1+\frac{1}{\gamma h})h^2 \sup_{t_k\le s\le t_{k+1}} (\mathbb{E}|X^N_{t_k} - X_s|^2 + |t_k - s|) \\
&+ C(1+\frac{1}{\gamma h})h^2 \sup_{t_k\le s\le t_{k+1}} \mathbb{E}|Y^N_{t_{k+1}} - \tilde{Y}^{N,1}_s|^2 + Ch(1+\frac{1}{\gamma h})\mathbb{E}\int_{t_k}^{t_{k+1}} |Z^N_{t_k} - Z^{N,1}_s|^2 ds \\
\le&(1+Ch)\mathbb{E}(|Y^N_{t_{k+1}} - Y^{N,1}_{t_{k+1}}|^2) + Ch^2 + Ch\mathbb{E}\int_{t_k}^{t_{k+1}} |Z^{N,1}_s|^2 ds + C\mathbb{E}\int_{t_k}^{t_{k+1}} |Z^{N,1}_s - \overline{Z}^N_{t_k}|^2 ds
\end{aligned}$$

en utilisant le lemme A.2 pour borner par Ch le terme $\sup_{t_k\le s\le t_{k+1}} \mathbb{E}|Y^{N,1}_{t_{k+1}} - \tilde{Y}^{N,1}_s|^2$ provenant de $\mathbb{E}|Y^N_{t_{k+1}} - \tilde{Y}^{N,1}_s|^2 \le 2\mathbb{E}|Y^N_{t_{k+1}} - Y^{N,1}_{t_{k+1}}|^2 + 2\mathbb{E}|Y^{N,1}_{t_{k+1}} - \tilde{Y}^{N,1}_s|^2$.
La différence avec la partie I est qu'on a besoin d'avoir une estimation de $\mathbb{E}\int_{t_k}^{t_{k+1}} |Z^{N,1}_s|^2 ds$ pour pouvoir conclure par le lemme de Gronwall discret. Le contrôle usuel par la norme

$\mathbf{L}^2$ de la condition terminale $\mathbb{E}|Y^{N,1}_{t_{k+1}}|^2$ n'est pas suffisant car cela donne un terme d'ordre 1 lorsqu'on applique le lemme de Gronwall discret. Par contre, on peut utiliser le Lemme A.2 qui nous donne $\sup_{t_k\leq t\leq t_{k+1}}\mathbb{E}|Z^{N,1}_t|^2\leq C(1+|X_0|^2)$. Ainsi, on obtient finalement

$$\mathbb{E}|Y^N_{t_k}-Y^{N,1}_{t_k}|^2\leq(1+Ch)\mathbb{E}(|Y^N_{t_{k+1}}-Y^{N,1}_{t_{k+1}}|^2)+Ch^2+C\mathbb{E}\int_{t_k}^{t_{k+1}}|Z^{N,1}_s-\overline{Z}^N_{t_k}|^2ds.$$

Le même type de calculs que dans la partie I permet d'établir la borne pour

$$\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z^{N,1}_t-Z^N_{t_k}|^2dt.$$

□

Ainsi, la vitesse de convergence est une fois de plus régie par le terme $\mathbb{E}\int_{t_k}^{t_{k+1}}|Z^{N,1}_s-\overline{Z}^N_{t_k}|^2ds$ puis par la vitesse de convergence de $(Y^{N,1},Z^{N,1})$ vers (Y,Z). Ces deux vitesses dépendent de la régularité de l'obstacle $g(\cdot,\cdot)$.

2.5.1 Régularité $C^{1,2}$

On fait ici l'hypothèse

Hypothèse III.4 *g est de classe $C^{1,2}$ avec ses dérivées bornées uniformément.*

Cette hypothèse reprend en fait les hypothèses de [MZ05] et permet d'appliquer leurs résultats. En combinant le Lemme A.1 et le Théorème A.4 , on obtient que

$$\mathbb{E}\{\sup_{0\leq t\leq T}|Y^{N,1}_t-Y_t|^2+\int_0^T|Z_s-Z^{N,1}_s|^2ds\}\leq C(1+|X_0|^4)\sqrt{h}.$$

De plus, on peut majorer $\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z^{N,1}_s-\overline{Z}^N_{t_k}|^2ds$ par

$$\begin{aligned}&C\mathbb{E}\int_0^T|Z^{N,1}_t-Z_t|^2dt+C\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_t-\frac{1}{h}\mathbb{E}_{t_k}\int_{t_k}^{t_{k+1}}Z_sds|^2dt\\ \leq&C(1+|X_0|^4)\sqrt{h}\end{aligned}$$

en combinant le lemme A.1 et le théorème A.4 pour le premier terme et en utilisant le théorème A.5 pour le second terme.
Ainsi, dans le cas où g est de classe $C^{1,2}$ avec des dérivées bornées, on obtient

$$\max_{0\leq k\leq N}\mathbb{E}(|Y^N_{t_k}-Y_{t_k}|^2)+\sum_{k=0}^{N-1}\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_t-Z^N_{t_k}|^2dt\leq C(1+|X_0|^4)\sqrt{h}.$$

Remarque III.1 *En comparant avec le théorème I.1, on constate que la vitesse de convergence s'est dégradée de $\sqrt{h}$ à $h^{\frac{1}{4}}$. Cette perte de vitesse est dûe à la présence de l'obstacle g et semble inévitable en toute généralité.*

Notre but est d'appliquer cet algorithme au cas d'options américaines où l'obstacle n'est pas de classe $C^{1,2}$ comme par exemple dans le cas du payoff $(K - e^x)_+$. Or, nous montrons que nous pouvons relaxer les hypothèses sur l'obstacle g et néanmoins conserver les conclusions données dans [MZ05].

2.5.2 Relaxation des hypothèses sur l'obstacle

On fait dans cette partie l'hypothèse suivante

Hypothèse III.5 *On suppose que g est de classe C^1 en temps, toujours uniformément Lipschitz en x et que $g(t, X_t)$ et $g(t, X_t^N)$ satisfont les développements :*

$$\begin{aligned} g(t, X_t) =& g(0, X_0) + \int_0^t U_s ds + \int_0^t V_s dW_s + A_t, \\ g(t, X_t^N) =& g(0, X_0^N) + \int_0^t U_s^N ds + \int_0^t V_s^N dW_s + A_t^N, \end{aligned}$$

où A, A^N sont continus, croissants, intégrables et tels que les mesures dA_t et dA_t^N sont singulières par rapport à dt et avec les conditions d'intégrabilité suivantes $\sup_t \mathbb{E}(|U_t|^p) + \sup_t \mathbb{E}(|U_t^N|^p) + \sup_t \mathbb{E}(|V_t|^p) + \sup_t \mathbb{E}(|V_t^N|^p) < \infty$ *pour $p \geq 2$.*

Remarquons que, par exemple, pour le cas du put américain en dimension 1, l'hypothèse III.5 est vérifiée car $g(t, x) = \psi(t, \exp(x))$ avec $\psi(t, \cdot)$ convexe. Ainsi, on obtient les formules d'Itô suivantes (voir théorème 7.1 page 218 dans [KS88]) :

$$\begin{aligned} g(t, X_t) = \psi(t, \exp(X_t)) =& g(0, X_0) + \int_0^t U_s ds + \int_0^t V_s dW_s + A_t \\ g(t, X_t^N) = \psi(t, \exp(X_t^N)) =& g(0, X_0^N) + \int_0^t U_s^N ds + \int_0^t V_s^N dW_s + A_t^N \end{aligned}$$

avec pour $t_k \leq t < t_{k+1}$

$$\begin{aligned} U_t =& \partial_t \psi(t, \exp(X_t)) + \partial_x \psi^-(t, \exp(X_t)) \exp(X_t) \{b(t, X_t) + \frac{1}{2}\sigma^2(t, X_t)\}, \\ U_t^N =& \partial_t \psi(t, \exp(X_t^N)) + \partial_x \psi^-(t, \exp(X_t^N)) \exp(X_t^N) \{b(t_k, X_{t_k}^N) + \frac{1}{2}\sigma^2(t_k, X_{t_k}^N)\}, \\ V_t =& \partial_x \psi^-(t, \exp(X_t)) \exp(X_t) \sigma(t, X_t), \\ V_t^N =& \partial_x \psi^-(t, \exp(X_t^N)) \exp(X_t^N) \sigma(t_k, X_{t_k}^N). \end{aligned}$$

Comme $\partial_x \psi^-$ est bornée dans le cas du put américain en dimension 1, les conditions d'intégrabilité de l'hypothèse III.5 sont vérifiées.
Lorsqu'on examine les preuves données dans [MZ05], on s'aperçoit que la condition g de classe $C^{1,2}$ intervient essentiellement à 2 endroits :

– Dans le lemme A.1. La régularité de g est utilisée pour montrer les contrôles sur les processus croissants K et $K^{N,0}$. Or, c'est justement ce que l'on impose dans l'hypothèse III.5. L'intérêt est que l'on peut à présent appliquer la remarque A.3 qui nous donne

$$\begin{aligned} 0 \leq dK_t &\leq \{f(t, X_t, g(t, X_t), V_t) + U_t\}^- dt = k_t dt, \\ 0 \leq dK_t^{N,0} &\leq \{f(t, X_t^N, g(t, X_t^N), V_t^N) + U_t^N\}^- dt = k_t^{N,0} dt, \end{aligned} \tag{III.2.5}$$

et on obtient les mêmes contrôles sur les processus croissants que dans le cas où g est de classe $C^{1,2}$. En particulier, $\sup_{t,N} \mathbb{E}|k_t^{N,0}|^p < \infty$ et $\sup_t \mathbb{E}|k_t|^p < \infty$ pour $p \geq 2$.

– Dans le théorème A.4. Dans notre cas, on aura seulement le résultat

$$\mathbb{E}\{\sup_{0 \leq t \leq T} |\tilde{Y}_t^{N,1} - Y_t^{N,0}|^2 + \int_0^T |Z_t^{N,1} - Z_t^{N,0}|^2 dt\} \leq C\sqrt{h}$$

soit une perte de vitesse par rapport au théorème A.4 (voir (A.3) pour la définition de $(Y^{N,0}, Z^{N,0})$). Ceci n'est en fait pas très grave vu la vitesse donnée par le lemme A.1. Voyons comment se montre ce résultat :

PREUVE: La preuve est presque similaire à celle de [MZ05]. Nous la redonnons tout de même dans son intégralité pour voir l'endroit où cela change.
On note $\Delta Y_t = Y_t^{N,0} - \tilde{Y}_t^{N,1}$, $\Delta Z_t = Z_t^{N,0} - Z_t^{N,1}$ et on remarque que ΔY_t a la dynamique suivante sur $]t_k, t_{k+1}[$

$$-d\Delta Y_t = [f(t, X_t^N, Y_t^{N,0}, Z_t^{N,0}) - f(t, X_t^N, \tilde{Y}_t^{N,1}, Z_t^{N,1})]dt - (Z_t^{N,0} - Z_t^{N,1})dW_t + dK_t^{N,0}.$$

Ainsi, en appliquant la formule d'Itô on obtient

$$\begin{aligned} d|\Delta Y_t|^2 = &- 2\Delta Y_t [f(t, X_t^N, Y_t^{N,0}, Z_t^{N,0}) - f(t, X_t^N, \tilde{Y}_t^{N,1}, Z_t^{N,1})]dt + 2\Delta Y_t \Delta Z_t dW_t \\ &- 2\Delta Y_t dK_t^{N,0} + |\Delta Z_t|^2 dt. \end{aligned}$$

Choisissons à présent $C > 0$ tel que

$$2|\Delta Y_s||f(s, X_s^N, Y_s^{N,0}, Z_s^{N,0}) - f(s, X_s^N, \tilde{Y}_s^{N,1}, Z_s^{N,1})| \leq C|\Delta Y_s|^2 + \frac{1}{2}|\Delta Z_s|^2 \tag{III.2.6}$$

ce qui est possible en utilisant que f est Lipschitz et une inégalité de Young (A.5). Ainsi, on obtient en appliquant une formule d'Itô à $\Lambda_s |\Delta Y_s|^2$ sur $[t, t_{k+1}]$ avec $\Lambda_s = \exp(Cs)$

(où C est la constante de l'équation précédente) :

$$\begin{aligned}
&\Lambda_t|\Delta Y_t|^2 + \int_t^{t_{k+1}} \Lambda_s|\Delta Z_s|^2 ds \\
=&\Lambda_{t_{k+1}}|\Delta Y_{t_{k+1}}|^2 + 2\int_t^{t_{k+1}} \Lambda_s \Delta Y_s [f(s, X_s^N, Y_s^{N,0}, Z_s^{N,0}) - f(s, X_s^N, \tilde{Y}_s^{N,1}, Z_s^{N,1})] ds \\
&- 2\int_t^{t_{k+1}} \Lambda_s \Delta Y_s \Delta Z_s dW_s + 2\int_t^{t_{k+1}} \Lambda_s \Delta Y_s dK_s^{N,0} - C\int_t^{t_{k+1}} \Lambda_s |\Delta Y_s|^2 ds
\end{aligned}$$

soit en utilisant (III.2.6) :

$$\begin{aligned}
&\Lambda_t|\Delta Y_t|^2 + \frac{1}{2}\int_t^{t_{k+1}} \Lambda_s|\Delta Z_s|^2 ds \\
\leq&\Lambda_{t_{k+1}}|\Delta Y_{t_{k+1}}|^2 - 2\int_t^{t_{k+1}} \Lambda_s \Delta Y_s \Delta Z_s dW_s + 2\int_t^{t_{k+1}} \Lambda_s \Delta Y_s dK_s^{N,0}.
\end{aligned}$$

Il nous reste à contrôler le dernier terme du membre de droite en vue de conclure par un lemme de Gronwall discret. Pour cela, on a :

$$\begin{aligned}
\Delta Y_s dK_s^{N,0} =&[g(s, X_s^N) - \tilde{Y}_s^{N,1}] dK_s^{N,0} \\
=&[g(s, X_s^N) - \mathbb{E}_s\{Y_{t_{k+1}}^{N,1} + \int_s^{t_{k+1}} f(r, X_r^N, \tilde{Y}_r^{N,1}, Z_r^{N,1}) dr\}] dK_s^{N,0} \\
\leq&\mathbb{E}_s[g(s, X_s^N) - g(t_{k+1}, X_{t_{k+1}}^N) - \int_s^{t_{k+1}} f(r, X_r^N, \tilde{Y}_r^{N,1}, Z_r^{N,1}) dr] dK_s^{N,0}.
\end{aligned}$$

C'est le moment où la démonstration change. En effet pour contrôler la différence $g(s, X_s^N) - g(t_{k+1}, X_{t_{k+1}}^N)$, les auteurs appliquent une formule d'Itô ce qui les mène ensuite au contrôle Ch dans l'équation ci-dessous. Dans notre cas, on peut tout de même effectuer une majoration intéressante. Ainsi, en utilisant (III.2.5) et l'hypothèse III.5, ainsi que le lemme A.2 dont on peut vérifier qu'il est toujours valable sous l'hypothèse III.5, on obtient :

$$\begin{aligned}
&\mathbb{E}\int_t^{t_k} \Lambda_r \Delta Y_r dK_r^{N,0} \\
\leq& C\mathbb{E}\int_t^{t_k} \Lambda_r \{|r - t_k| + |X_r^N - X_{t_k}^N| + \int_r^{t_k} |f(s, X_s^N, \tilde{Y}_s^{N,1}, Z_s^{N,1})| ds\} k_r^{N,0} dr \\
\leq& Ch^{\frac{3}{2}}.
\end{aligned}$$

On obtient finalement :

$$\begin{aligned}
\mathbb{E}[\Lambda_t|\Delta Y_t|^2 + \frac{1}{2}\int_t^{t_{k+1}} \Lambda_s|\Delta Z_s|^2 ds] \leq&\Lambda_{t_{k+1}}[\mathbb{E}|\Delta Y_{t_{k+1}}|^2 + Ch^{\frac{3}{2}}] \\
\leq&\Lambda_t(1 + Ch)[\mathbb{E}|\Delta Y_{t_{k+1}}|^2 + Ch^{\frac{3}{2}}].
\end{aligned}$$

Comme Λ_t est croissant, il vient :

$$\mathbb{E}[|\Delta Y_t|^2 + \frac{1}{2}\int_t^{t_{k+1}} |\Delta Z_s|^2 ds] \leq (1+Ch)\mathbb{E}|\Delta Y_{t_{k+1}}|^2 + Ch^{\frac{3}{2}}$$

et une application du Lemme de Gronwall discret mène au résultat. □

On peut alors vérifier que tous les autres résultats dans [MZ05] n'utilisent pas le caractère $C^{1,2}$ de g. Ainsi, on aura le même contrôle

$$\max_{0\leq k\leq N} \mathbb{E}(|Y_{t_k}^N - Y_{t_k}|^2) + \sum_{k=0}^{N-1} \mathbb{E}\int_{t_k}^{t_{k+1}} |Z_t - Z_{t_k}^N|^2 dt \leq C\sqrt{h}.$$

2.6 Conclusion

Si on résume les résultats établis dans ce chapitre, on obtient donc sous les hypothèses III.2, III.3 et soit III.4 soit III.5

$$\begin{aligned}
&\max_{0\leq k\leq N} \mathbb{E}|y_k^{N,R,M}(X_{t_k}^N) - Y_{t_k}|^2 + \sum_{k=0}^{N-1} \mathbb{E}\int_{t_k}^{t_{k+1}} |Z_t - z_k^{N,R,M}(X_{t_k}^N)|^2 dt \\
\leq & C\sqrt{h} + C\sum_{k=0}^{N-1} \mathbb{E}\big[\mathbf{1}_{g(t_k,X_{t_k}^N)\neq g^R(t_k,X_{t_k}^N)}\{|Y_{t_k}^N|^2 + C_y(R)^2\}\big] + C\mathbb{E}(|g(t_N,X_{t_N}^N) - g^R(t_N,X_{t_N}^N)|^2) \\
&+ Ch\mathbb{E}\sum_{k=0}^{N-1} |f(t_k,X_{t_k}^N,Y_{t_{k+1}}^N,Z_{t_k}^N) - f^R(t_k,X_{t_k}^N,Y_{t_{k+1}}^N,Z_{t_k}^N)|^2 \\
&+ C\frac{C_y(R)^2 + \|f^R\|_\infty^2}{M}\sum_{k=0}^{N-1} \max_{0\leq l\leq q} \mathbb{E}(K_{l,k}(\omega)) + Ch^{\beta-1} + C\sum_{k=0}^{N-1} \inf_\alpha \mathbb{E}\big(|y_k^{N,R}(P_{t_k}^N) - \alpha.p_{0,k}(P_{t_k}^N)|^2\big) \\
&+ C\sum_{k=0}^{N-1}\sum_{l=1}^{q} \inf_\alpha \mathbb{E}\big(|\alpha.p_{l,k}(P_{t_k}^N) - \sqrt{h}z_{l,k}^{N,R}(P_{t_k}^N)|^2\big) + C\frac{(C_y(R)^2 + \|f^R\|_\infty^2)\max_l K_{l,k}}{M}\log(M) \\
&+ \frac{CC_y(R)^2}{h}\exp(-\frac{Mh^{\beta+1}}{144C_y(R)^2})\sum_{k=0}^{N-1} \mathbb{E}\bigg(\mathcal{N}_2(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}}, [\mathcal{P}_{0,k+1}]_y - y_{k+1}^{N,R}(\cdot), (P_{t_{k+1}}^{N,m})_{1\leq m\leq M}, (\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M})\bigg).
\end{aligned}$$

Partie numérique

Quatrième partie

Résultats numériques

Chapitre 1

Rappel des chapitres précédents

Dans cette thèse, nous désirons approcher la solution de l'EDSR :

$$Y_t = \Phi(\mathbf{X}) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s \tag{IV.1.1}$$

où W est un mouvement brownien de dimension q et X est la solution de l'équation différentielle stochastique en dimension d :

$$X_{i,t} = X_{i,0} + \int_0^t b_i(s, X_s)ds + \int_0^t \sigma_i(s, X_s)dW_s,\ 1 \leq i \leq d.$$

Φ est une fonctionnelle de toute la trajectoire de la diffusion S, ceci pour tenir compte d'options path-dependent (voir chapitre 2, partie I). Pour se débarasser du caractère path-dependent, on approche la condition terminale $\Phi(\mathbf{X})$ par $\Phi^N(P^N_{t_N})$, où Φ^N est une fonction déterministe et $(P^N_{t_k})_{0\leq k\leq N}$ est une chaine de Markov de dimension $\mathbf{d'} \geq \mathbf{d}$ dont les d premières composantes sont données par celles de $(X^N_{t_k})_{0\leq k\leq N}$. En d'autres termes, on ajoute éventuellement d'autres variables d'état pour rendre markovienne la dynamique implicite de la condition terminale.

On a étudié dans les parties I et II deux algorithmes rétrogrades pour approcher la solution de (IV.1.1). Ces deux algorithmes sont basés sur des problèmes de moindres carrés qu'on résout itérativement le long d'une discrétisation en temps de l'intervalle $[0, T]$ (les $N+1$ dates de discrétisation sont notées $(t_k = kh)_{0\leq k\leq N}$ pour $h = \frac{T}{N}$). En effet, on désire approcher la solution (Y_{t_k}, Z_{t_k}) à chaque instant de discrétisation t_k par des projections sur des bases de fonctions. La base utilisée pour approcher Y_{t_k} est notée $p_{0,k}(P^N_{t_k}) = p_{0,k}$ et les q bases utilisées pour approcher chacunes des q composantes de Z_{t_k} sont notées $(p_{l,k}(P^N_{t_k}) = p_{l,k})_{1\leq l\leq q}$. Le cardinal de la base $p_{l,k}$ est noté $K_{l,k}$.

Rappelons enfin que les problèmes de moindres carrés sont basés sur des simulations de Monte-Carlo de la chaine de Markov $(P_{t_k}^N)_{0\leq k\leq N}$. Le nombre de simulations de Monte-Carlo est toujours noté $\mathbf{M}$ et les simulations sont notées $(P_{t_k}^{N,m})_{0\leq k\leq N, 1\leq m\leq M}$. Les rangs des matrices de régression de ces problèmes de moindres carrés associées à $p_{l,k}$ et aux M simulations sont notés $K_{l,k}(\omega)$.
Avant de tester les deux algorithmes présentés dans les parties I et II, nous devons spécifier le choix des fonctions de base.

Chapitre 2

Choix des fonctions de base

Dans toute la partie numérique, nous testons différents types de fonctions de base. Par souci de clarté, nous avons regroupé la description de ces différents types de fonction dans la même partie.

2.1 Hypercubes

Pour simplifier, nous ne faisons dépendre $p_{l,k}$ ni de l ni de k.
Pour définir la base de fonctions, on choisit un domaine $D \subset \mathbb{R}^{d'}$, centré en P_0, i.e. $D = \prod_{i=1}^{d'}]P_{0,i} - R, P_{0,i} + R]$, que nous partitionnons en petits hypercubes d'arête δ. Ainsi, $D = \cup_{i_1,\cdots,i_{d'}} D_{i_1,\cdots,i_{d'}}$ où $D_{i_1,\cdots,i_{d'}} =]P_{0,1} - R + i_1\delta, P_{0,1} - R + (i_1+1)\delta] \times \cdots \times]P_{0,d'} - R + i_{d'}\delta, P_{0,d'} - R + (i_{d'}+1)\delta]$.
Nous définissons alors $p_{l,k}(\cdot)$ comme la famille de fonctions indicatrices associée à cet ensemble d'hypercubes : $p_{l,k}(\cdot) = \left(\mathbf{1}_{D_{i_1,\cdots,i_{d'}}}(\cdot)\right)_{i_1,\cdots,i_{d'}}$.
Cette base de fonctions est notée **HC**.

De plus, nous considérons éventuellement sur chaque hypercube des polynômes locaux de petit degré. Par exemple, nous pouvons prendre une base de polynômes locaux constituée par $1, x_1, \cdots, x_{d'}$ pour $p_{0,k}$ et 1 pour $p_{l,k}$ ($l \geq 1$). Ainsi,

$$p_{0,k}(x) = \left(\mathbf{1}_{D_{i_1,\cdots,i_{d'}}}(x), x_1\mathbf{1}_{D_{i_1,\cdots,i_{d'}}}(x), \cdots, x_{d'}\mathbf{1}_{D_{i_1,\cdots,i_{d'}}}(x)\right)_{i_1,\cdots,i_{d'}},$$
$$p_{l,k}(x) = \left(\mathbf{1}_{D_{i_1,\cdots,i_{d'}}}(x)\right)_{i_1,\cdots,i_{d'}}, 1 \leq l \leq q.$$

Ce choix particulier de fonctions de base est noté **HC(1,0)** où 1 (resp. 0) représente le degré maximal de la base de polynômes locaux pour $p_{0,k}$ (resp. $p_{l,k}$, $1 \leq l \leq q$).

2.2 Cellules de Voronoï

Ici, nous considérons encore une base de fonctions indicatrices et la même base pour $0 \leq l \leq q$. Cette fois, les ensembles des fonctions indicatrices sont une partition ouverte de Voronoï ([GL00]) dont les centres sont des simulations indépendantes de P^N. Plus précisément, si nous désirons par exemple générer une base de 20 fonctions indicatrices, nous simulons 20 trajectoires supplémentaires de P^N, notée $(P^{N,M+i})_{1\leq i\leq 20}$, indépendamment de $(P^{N,m})_{1\leq m\leq M}$. Puis, nous utilisons à l'instant t_k $(P^{N,M+i}_{t_k})_{1\leq i\leq 20}$ pour définir notre partition de Voronoï $(C_{k,i})_{1\leq i\leq 20}$ où $C_{k,i} = \{x : |x - P^{N,M+i}_{t_k}| < \inf_{j\neq i} |x - P^{N,M+j}_{t_k}|\}$. Enfin, nous posons $p_{l,k}(\cdot) = \left(\mathbf{1}_{C_{k,i}}(\cdot)\right)_i$. Cette base de fonctions est notée **VP** (Voronoi Partition).

De même que pour la base **HC**, nous considérons éventuellement sur chaque cellule de Voronoï des polynômes locaux de petit degré. Pour reprendre le même exemple que pour **HC**, on peut considérer sur chaque cellule, une base polynômiale constituée de $1, x_1, \cdots, x_{d'}$ pour $p_{0,k}$ et 1 pour $p_{l,k}$ $(l \geq 1)$ sur chaque $C_{k,i}$. Ainsi

$$p_{0,k}(x) = \left(\mathbf{1}_{C_{k,i}}(x), x_1\mathbf{1}_{C_{k,i}}(x), \cdots, x_{d'}\mathbf{1}_{C_{k,i}}(x)\right)_i,$$
$$p_{l,k}(x) = \left(\mathbf{1}_{C_{k,i}}(x)\right)_i, 1 \leq l \leq q.$$

Nous notons cette base particulière **VP(1,0)** où $(1, 0)$ a la même signification que pour **HC**.

2.3 Polynômes globaux

Nous définissons ici $p_{0,k}$ comme la base de polynômes de d' variables et de degré inférieur à d_y et $p_{l,k}$ $(l \geq 1)$ comme la base de polynômes de degré inférieur à d_z.

Chapitre 3

Comparaison des approches avec ou sans incréments browniens auxiliaires

Nous avons dans la partie II étudié un algorithme qui utilise à l'instant t_k la simulation d'incréments browniens auxiliaires. Nous avons mentionné dans le dernier chapitre de cette partie des pistes qui montrent qu'il est possible d'appliquer le même algorithme mais sans simulation d'incréments browniens auxiliaires, et probablement de déduire des bornes d'erreurs qui dépendent uniquement du nombre de fonctions de base, de M et de h, du même type que celles établies dans les théorèmes II.2, II.3, II.4 et II.5.
Nous allons ici comparer l'algorithme présenté dans la partie II avec ou sans simulation d'incréments browniens auxiliaires.
On prend ici $d = q = 1$ et on considère pour S la dynamique du modèle Black-Scholes ([BS73]) :

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t.$$

La condition terminale de l'EDSR est $(S_T - K)_+$, celle d'un call européen. Comme driver f, on prend :

$$f(t, x, y, z) = -\theta z - ry + (y - \frac{z}{\sigma})^-(R - r)$$

avec $\theta = \frac{\mu - r}{\sigma}$, la prime de risque.
Les paramètres de la dynamique de S et de l'option sont :

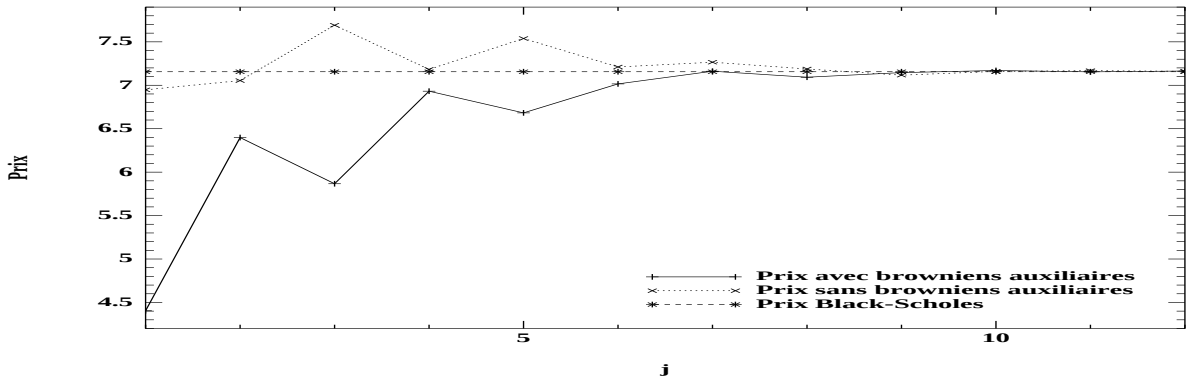
μ	σ	r	R	T	S_0	K
0.05	0.2	0.04	0.06	0.5	100	100

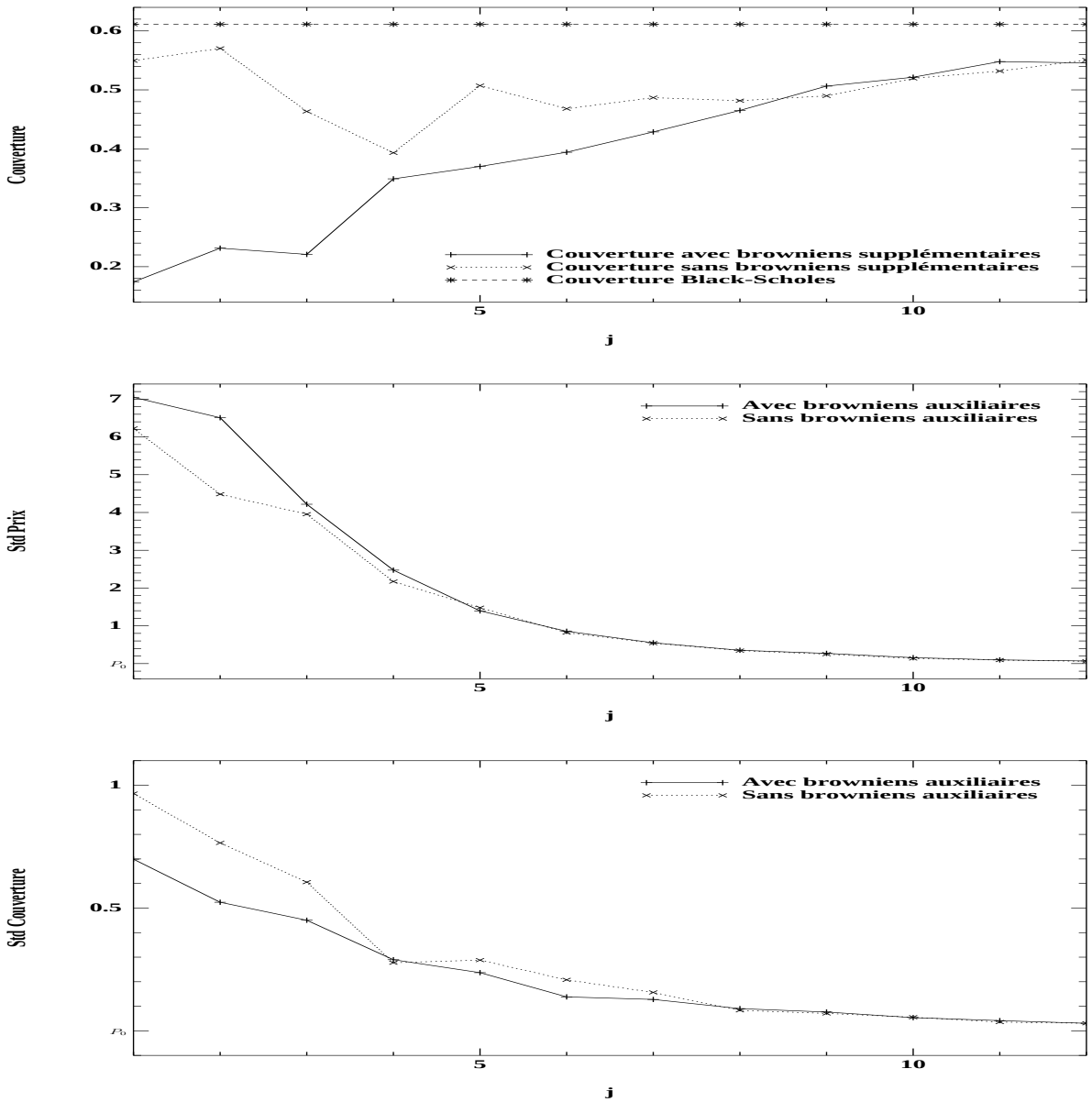
Cette situation correspond au cas où le vendeur de l'option emprunte de l'argent au taux d'intérêt R et prête au taux r avec $r < R$ ([Ber95]). Cette imperfection de marché par

rapport aux hypothèses idéales de [BS73] crée une non-linéarité dans le driver f, qui sinon vaudrait simplement $f(t,x,y,z) = -\theta z - ry$. Nous avons choisi cet exemple car le driver f est non-linéaire mais nous pouvons néanmoins connaitre la valeur Y_0. En effet, le driver f se réécrit :

$$\begin{aligned} f(t,x,y,z) &= -\frac{\mu - R}{\sigma} z - Ry + (y - \frac{z}{\sigma})^+ (R - r) \\ &\geq -\frac{\mu - R}{\sigma} z - Ry. \end{aligned}$$

Ce dernier terme étant le driver correspondant à une dynamique de portefeuille classique avec un unique taux d'intérêt R, on sait en utilisant le théorème de comparaison pour les EDSR ([EKPQ97]) que Y_0 sera supérieur au prix calculé par la formule de Black-Scholes avec taux d'intérêt R. Mais on sait également dans le cas de la réplication dans le modèle Black-Scholes que pour couvrir l'option, le vendeur devra toujours emprunter à la banque, c'est-à-dire que, si on note (Y_t^{BS}, Z_t^{BS}) les solutions de l'EDSR associée au driver $-\frac{\mu-R}{\sigma} z - Ry$, alors $Y_t^{BS} \leq \frac{Z_t^{BS}}{\sigma}$, $\frac{Z_t^{BS}}{\sigma}$ représentant le montant investi en action. Ainsi, on obtient bien l'égalité des drivers f et $-\frac{\mu-R}{\sigma} z - Ry$ le long de la solution (Y_t^{BS}, Z_t^{BS}) et donc l'égalité $Y_0 = Y_0^{BS}$.
Y_0^{BS} se calcule explicitement et vaut **7.156**.
On teste ici la base **VC**. On note K le nombre de cellules de Voronoi utilisées. Pour comparer les algorithmes avec ou sans simulation d'incréments browniens auxiliaires, nous décidons de faire varier simultanément N, M et K. En particulier, nous faisons varier N de telle sorte que $N = [N_0(\sqrt{2})^{j-1}]$ où $j = 1\ldots$ désigne le nombre de valeurs de N que l'on teste ($[\cdot]$ désigne ici la partie entière). Ensuite, nous faisons varier K et M de telle sorte que $K = [K_0(\sqrt{2})^{(j-1)\alpha_K}]$ et $M = [M_0(\sqrt{2})^{(j-1)\alpha_M}]$ avec $N_0 = 2 = M_0 = K_0$.
Pour chaque valeur de (α_K, α_M) nous reportons le prix moyen donné par l'algorithme (calculé en lançant 50 fois l'algorithme), l'écart-type de ce prix, la couverture moyenne en t_0 et l'écart-type de la couverture. Voici les résultats obtenus :





Nous voyons donc que les algorithmes avec ou sans simulation d'incréments browniens auxiliaires semblent se comporter de la même manière : c'est vrai pour cet exemple et également vrai pour les autres exemples qui suivent. Ainsi par concision, dans toute la suite, nous présentons les tests avec l'algorithme sans simulation d'incréments browniens auxiliaires.

Chapitre 4

Variation simultanée des paramètres

4.1 Estimation de l'erreur d'approximation pour la base HC

Nous avons obtenu dans la partie II le résultat suivant.

$$
\begin{aligned}
&\max_{0\le k\le N}\mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R}(P_{t_k}^{N,m})-y_k^{N,R,M}(P_{t_k}^{N,m})|^2\\
&+h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}|z_k^{N,R,M}(P_{t_k}^{N,m})-z_k^{N,R}(P_{t_k}^{N,m})|^2\\
\le& C\frac{C_y(R)^2+\|f^R\|_\infty^2}{M}\sum_{k=0}^{N-1}\max_{0\le l\le q}\mathbb{E}(K_{l,k}(\omega))+C\sum_{k=0}^{N-1}\inf_{\alpha}\mathbb{E}|\alpha.p_{0,k}(P_{t_k}^N)-y_k^{N,R}(P_{t_k}^N)|^2\\
&+C\sum_{k=0}^{N-1}\sum_{l=1}^{q}\inf_{\alpha}\mathbb{E}|\alpha.p_{l,k}(P_{t_k}^N)-\sqrt{h}z_{l,k}^{N,R}(P_{t_k}^N)|^2+Ch^{\beta-1}\\
&+\frac{CC_y(R)^2}{h}\exp(-\frac{Mh^{\beta+1}}{144C_y(R)^2})\sum_{k=0}^{N-1}\mathbb{E}\bigg(\mathcal{N}_2(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}},[\mathcal{P}_{0,k+1}]_y-y_{k+1}^{N,R}(\cdot),(P_{t_{k+1}}^{N,m})_{1\le m\le M},(\tilde{P}_{t_{k+1}}^{N,m})_{1\le m\le M})\bigg).
\end{aligned}
\tag{IV.4.1}
$$

Rappelons (voir partie II) que $\big(y^{N,R}(P^N),z^{N,R}(P^N)\big)$ est la solution d'une discrétisation en temps de (IV.1.1) et d'une localisation déterminée par un vecteur R. Cette localisation nous permet de montrer que $y^{N,R}(P^N)$ est borné par une constante $C_y(R)$ qui dépend de la localisation. Rappelons enfin que l'approximation résultant de l'algorithme de la partie II est notée $\big(y^{N,R,M}(P^N),z^{N,R,M}(P^N)\big)$.
Nous allons déduire du résultat ci-dessus la proposition suivante

Proposition IV.1 *En utilisant la base* **HC** *avec les paramètres* δ *et* M *tels que* $\delta = h^\gamma$ *pour* $\gamma > \frac{1}{2}$ *et* $M = h^{-(\beta+1+\gamma d')}$ *avec* $\beta > 1$*, on obtient :*

$$\begin{aligned}
&\max_k \mathbb{E}\frac{1}{M}\sum_{m=1}^{M} |y_k^{N,R,M}(P_{t_k}^{N,m}) - y_k^{N,R}(P_{t_k}^{N,m})|^2 + h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M} |z_k^{N,R,M}(P_{t_k}^{N,m}) - z_k^{N,R}(P_{t_k}^{N,m})|^2 \\
\leq & o(1) + C(R)\sum_{k=0}^{N-1}\mathbb{P}(P_{t_k}^N \notin D)
\end{aligned}$$

où $o(1)$ *tend vers* 0 *quand* h *tend vers* 0*.*

PREUVE: En appliquant (IV.4.1) avec un paramètre β' tel que $1 < \beta' < \beta$ on obtient

$$\begin{aligned}
&\max_{0\leq k\leq N} \mathbb{E}\frac{1}{M}\sum_{m=1}^{M} |y_k^{N,R}(P_{t_k}^{N,m}) - y_k^{N,R,M}(P_{t_k}^{N,m})|^2 \\
&+ h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M} |z_k^{N,R,M}(P_{t_k}^{N,m}) - z_k^{N,R}(P_{t_k}^{N,m})|^2 \\
\leq & C\frac{C_y(R)^2 + \|f^R\|_\infty^2}{M}\sum_{k=0}^{N-1}\max_{0\leq l\leq q}\mathbb{E}(K_{l,k}(\omega)) + C\sum_{k=0}^{N-1}\inf_\alpha \mathbb{E}|\alpha.p_{0,k}(P_{t_k}^N) - y_k^{N,R}(P_{t_k}^N)|^2 \\
&+ C\sum_{k=0}^{N-1}\sum_{l=1}^{q}\inf_\alpha \mathbb{E}|\alpha.p_{l,k}(P_{t_k}^N) - \sqrt{h}z_{l,k}^{N,R}(P_{t_k}^N)|^2 + Ch^{\beta'-1} \\
&+ \frac{CC_y(R)^2}{h}\exp(-\frac{Mh^{\beta'+1}}{144C_y(R)^2})\sum_{k=0}^{N-1}\mathbb{E}\bigg(\mathcal{N}_2(\frac{h^{\frac{\beta'+1}{2}}}{3\sqrt{2}}, [\mathcal{P}_{0,k+1}]_y - y_{k+1}^{N,R}(\cdot), (P_{t_{k+1}}^{N,m})_{1\leq m\leq M}, (\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M})\bigg).
\end{aligned}$$

Tout d'abord, nous savons que $y_k^{N,R}(\cdot)$ et $(\sqrt{h}z_{l,k}^{N,R}(\cdot))_{1\leq l\leq q}$ sont des fonctions Lipschitz uniformément en k et N. Ainsi, pour la base **HC**, nous obtenons :

$$\begin{aligned}
&\mathbb{E}\big(\mathcal{R}_{p_{0,k}}(Y_{t_k}^{N,R})^2\big) \\
\leq & \mathbb{E}\big(|Y_{t_k}^{N,R}|^2\mathbf{1}_{D^c}(P_{t_k}^N)\big) + \sum_{i_1,\cdots,i_{d'}}\mathbb{E}\big(\mathbf{1}_{D_{i_1,\cdots,i_{d'}}}(P_{t_k}^N)|y_k^{N,R}(P_{t_k}^N) - y_k^{N,R}(x_{i_1,\cdots,i_{d'}})|^2\big) \\
\leq & C\delta^2 + C_y(R)^2\mathbb{P}(P_{t_k}^N \notin D),
\end{aligned}$$

où $x_{i_1,\cdots,i_{d'}}$ est un point arbitraire de $D_{i_1,\cdots,i_{d'}}$ et où nous avons utilisé la propriété de Lipschitz de $y_k^{N,R}$ sur D. En posant à présent $\eta = \delta h^{-\frac{1}{2}} \to 0$ par hypothèse, on obtient

$$\inf_\alpha \mathbb{E}|\alpha.p_{0,k}(P_{t_k}^N) - y_k^{N,R}(P_{t_k}^N)|^2 \leq Ch\eta^2 + C_y(R)^2\mathbb{P}(P_{t_k}^N \notin D).$$

Des calculs similaires sont valables pour $\mathbb{E}\big(\mathcal{R}_{p_{l,k}}(\sqrt{h}Z^{N,R}_{l,t_k})^2\big)$ et on en déduit

$$\begin{aligned}
&\max_k \mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R,M}(P_{t_k}^{N,m}) - y_k^{N,R}(P_{t_k}^{N,m})|^2 + h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}|z_k^{N,R,M}(P_{t_k}^{N,m}) - z_k^{N,R}(P_{t_k}^{N,m})|^2 \\
\leq& C(h^{\beta'-1}+\eta^2) + CC_y(R)^2\sum_{k=0}^{N-1}\mathbb{P}(P_{t_k}^N \notin D) + C\frac{C_y(R)^2+\|f^R\|_\infty^2}{M}\sum_{k=0}^{N-1}\max_{0\leq l\leq q}\mathbb{E}(K_{l,k}(\omega)) \\
&+ \frac{CC_y(R)^2}{h}\exp(-\frac{Mh^{\beta'+1}}{144C_y(R)^2})\sum_{k=0}^{N-1}\mathbb{E}\bigg(\mathcal{N}_2(\frac{h^{\frac{\beta'+1}{2}}}{3\sqrt{2}}, [\mathcal{P}_{0,k+1}]_y - y_{k+1}^{N,R}(\cdot), (P_{t_{k+1}}^{N,m})_{1\leq m\leq M}, (\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M})\bigg).
\end{aligned}$$

$K_{l,k}(\omega)$ est alors majoré par le nombre d'hypercubes d'arête δ contenu dans le domaine D. **Majoré et non pas égal** car seuls comptent les hypercubes dans lesquels sont présentes des simulations $(P_{t_k}^{N,m})_{1\leq m\leq M}$, comme on l'a déjà indiqué dans la partie II. Ainsi, on a $K_{l,k}(\omega) \leq (\frac{2R}{\delta})^{d'}$ soit $K_{l,k}(\omega) \leq C(R)\delta^{-d'}$ vu le choix de R et δ effectués précédemment. Il vient :

$$\begin{aligned}
&\max_k \mathbb{E}\frac{1}{M}\sum_{m=1}^{M}|y_k^{N,R,M}(P_{t_k}^{N,m}) - y_k^{N,R}(P_{t_k}^{N,m})|^2 + h\mathbb{E}\sum_{k=0}^{N-1}\frac{1}{M}\sum_{m=1}^{M}|z_k^{N,R,M}(P_{t_k}^{N,m}) - z_k^{N,R}(P_{t_k}^{N,m})|^2 \\
\leq& C(h^{\beta'-1}+\eta^2) + CC_y(R)^2\sum_{k=0}^{N-1}\mathbb{P}(P_{t_k}^N \notin D) + \frac{C(R)\delta^{-d'}}{Mh} \\
&+ \frac{CC_y(R)^2}{h}\exp(-\frac{Mh^{\beta'+1}}{144C_y(R)^2})\sum_{k=0}^{N-1}\mathbb{E}\bigg(\mathcal{N}_2(\frac{h^{\frac{\beta'+1}{2}}}{3\sqrt{2}}, [\mathcal{P}_{0,k+1}]_y - y_{k+1}^{N,R}(\cdot), (P_{t_{k+1}}^{N,m})_{1\leq m\leq M}, (\tilde{P}_{t_{k+1}}^{N,m})_{1\leq m\leq M})\bigg).
\end{aligned}$$

On utilise à présent le contrôle du nombre de couverture $\mathcal{N}_2$ en fonction de δ et h donné par (II.3.4) (voir partie II). Comme on a choisi $M \approx h^{-(\beta+1+\gamma d')}$ avec $\beta > \beta'$ on obtient finalement que les deux termes impliquant M dans l'inéquation ci-dessus tendent vers 0. □

Nous désirons vérifier la validité de cette proposition, c'est-à-dire voir si en faisant varier N, δ et M de telle sorte que le majorant de la proposition tende vers 0, l'algorithme converge bien.
Normalement, nous devrions également faire tendre le paramètre R vers l'infini pour assurer que $\sum_k \mathbb{P}(P_{t_k}^N \notin D)$ tend vers 0 et que $y_0^{N,R}(P_{t_0}^N)$ tend vers Y_0. Nos premiers tests numériques nous montrent que le paramètre R, s'il est choisi suffisamment grand dès le début, n'influe pas de manière significative sur la convergence. Ainsi, nous décidons d'utiliser un domaine de taille fixe.

4.2 Tests sur une option européenne avec différentiel de taux

Nous considérons la même EDSR et les mêmes valeurs de paramètres que pour le chapitre 3 et rappelons que le prix de référence est **7.156**.

4.2.1 Tests avec la base HC

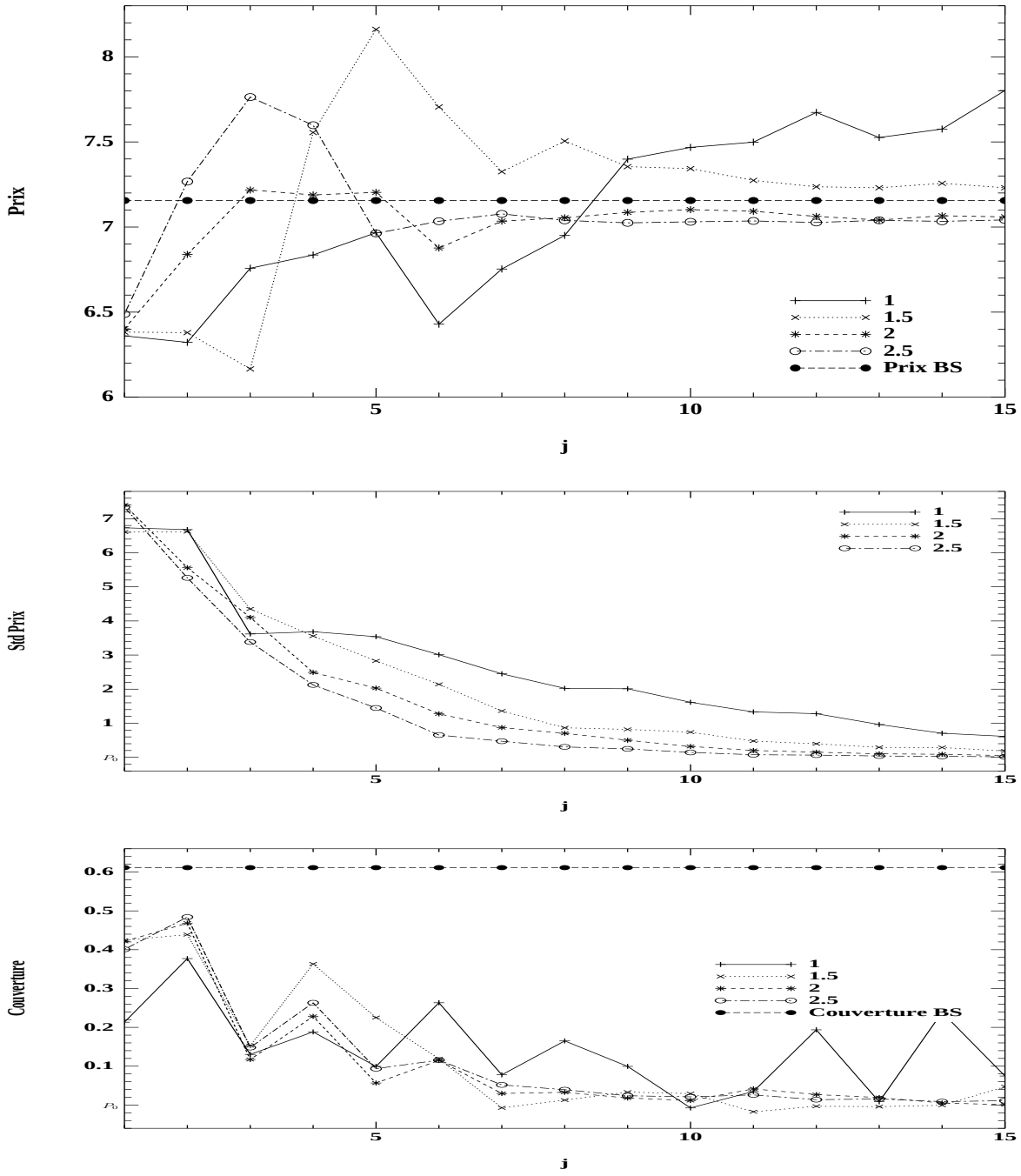
Nous effectuons tout d'abord des tests avec la base **HC**. Comme on l'a déjà dit, on désire tester la variation simultanée de M, δ et N. En effet, le domaine D est fixé une fois pour toute à $[40, 180]$ et la constante $C_y(R)$ est fixé à 100. Ces valeurs paraissent suffisamment larges pour ne pas biaiser les résultats. En effet, un calcul de probabilité de sortie de domaine nous indique que la sortie de $[40, 180]$ (sur les maturités d'options que nous considérons) est extrêmement rare et fixer $C_y(R)$ à $100 = S_0 = K$ parait très raisonnable. Nous faisons varier N de telle sorte que $N = [N_0(\sqrt{2})^{j-1}]$ où $j = 1 \ldots$ désigne le nombre de valeurs de N que l'on teste ($[\cdot]$ désigne ici la partie entière). Ensuite, nous faisons varier δ et M de telle sorte que $\delta = \frac{\delta_0}{\sqrt{2}^{(j-1)\alpha_\delta}}$ et $M = [M_0\sqrt{2}^{(j-1)\alpha_M}]$.
Nous prenons $N_0 = 2 = M_0$ et $\delta_0 = 50$. Nous fixons une valeur de α_δ pour laquelle nous testons différentes valeurs de α_M. Pour chaque valeur de α_δ nous reportons le prix moyen donné par l'algorithme (calculé en lançant 50 fois l'algorithme), l'écart-type de ce prix, la couverture moyenne en t_0 et l'écart-type de la couverture. Pour donner une idée de l'ordre de grandeur des paramètres, disons par exemple que pour $j = 10$, cela signifie que $N = 64$ et suivant les valeurs de α_δ et α_M on donne dans le tableau ci-dessous les valeurs du couple (δ, M) :

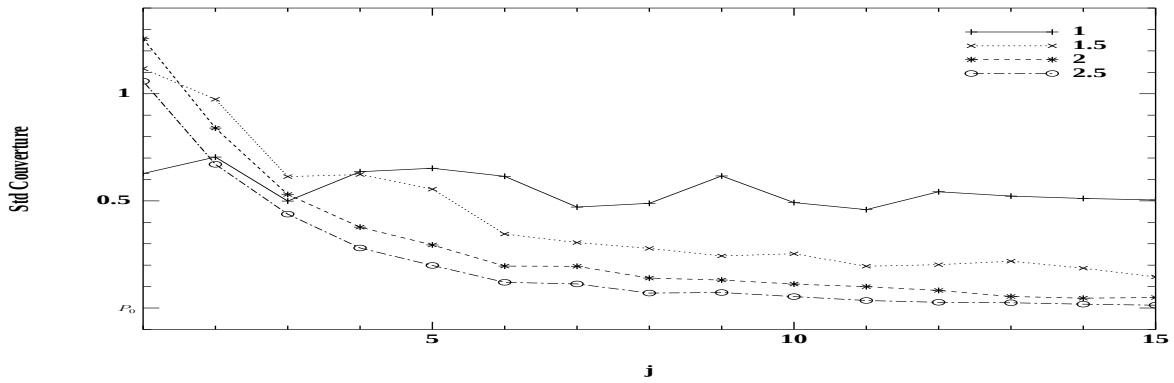
$\alpha_\delta \setminus \alpha_M$	1	1.5	2	2.5	3
0.6	(6.25,64)	**(6.25,362)**	(6.25,2048)	(6.25,11585)	(6.25,65536)
1	(1.56,64)	(1.56,362)	**(1.56,2048)**	(1.56,11585)	(1.56,65536)
1.5	(0.28,64)	(0.28,362)	(0.28,2048)	**(0.28,11585)**	(0.28,65536)
2	(0.05,64)	(0.05,362)	(0.05,2048)	(0.05,11585)	**(0.05,65536)**

En gras, on a indiqué les valeurs correspondants aux seuils de convergence. En effet, pour chaque valeur de α_δ on s'apercevra sur les graphiques ci-après, qu'il existe une valeur seuil pour α_M en dessous de laquelle l'algorithme semble ne pas converger et au dessus de laquelle il converge. Par exemple pour $\alpha_\delta = 0.6$ le seuil que l'on observe numériquement est $\alpha_M = 1.5$ donc on fonce le couple $(6.25, 362)$.
Voici les résultats obtenus pour la variation simultanée des paramètres N, δ et M :

– $\alpha_\delta = \mathbf{0.6}$ et α_M varie

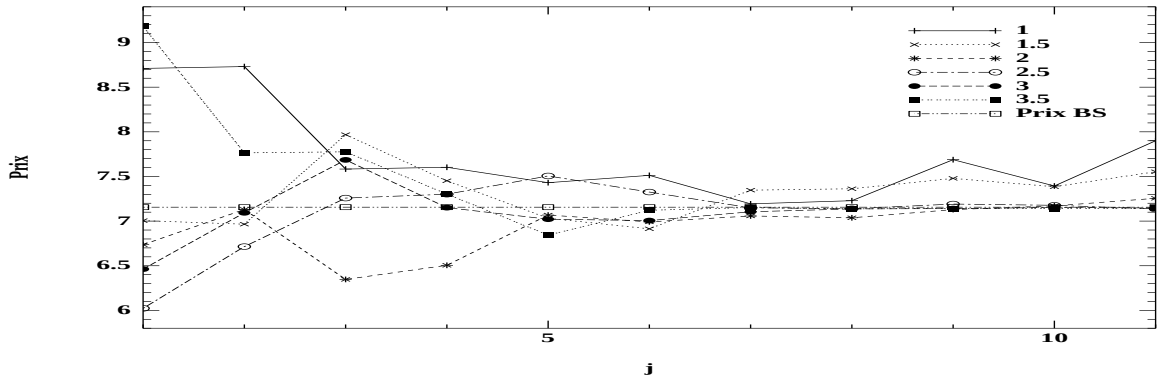
La proposition IV.1 nous impose que δ tende vers 0 plus vite que $h^{\frac{1}{2}}$. Ceci impose $\alpha_\delta > \frac{1}{2}$. C'est juste le cas avec $\alpha_\delta = 0.6$. Le seuil pour M de la proposition IV.1, ici avec $d' = 1$ (et en prenant la valeur minimale pour β), donne $\alpha_{\mathbf{M}} > \mathbf{2.6}$.

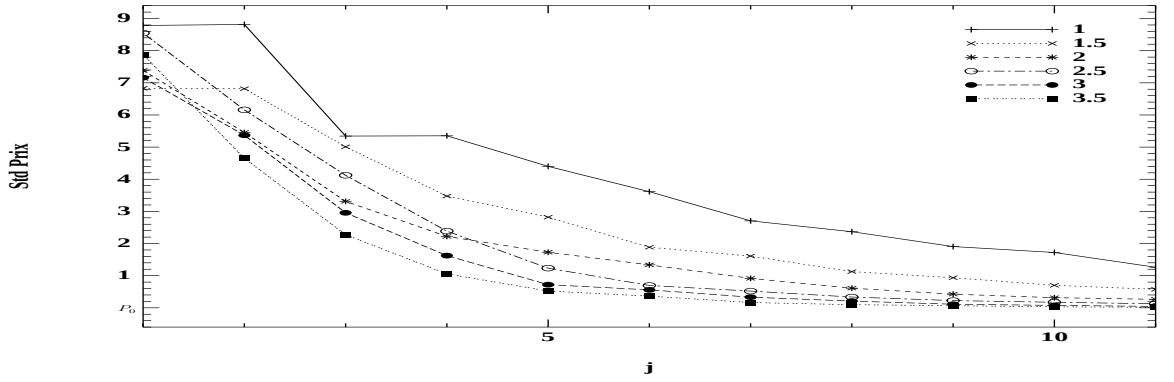


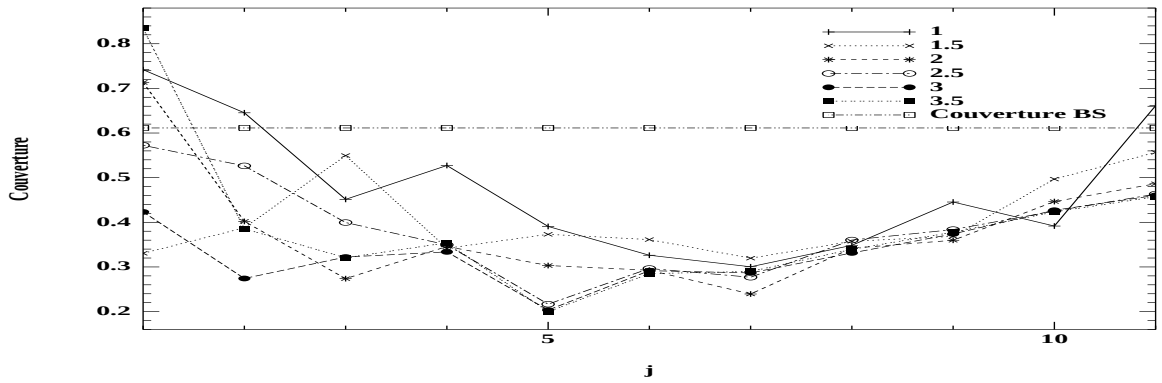


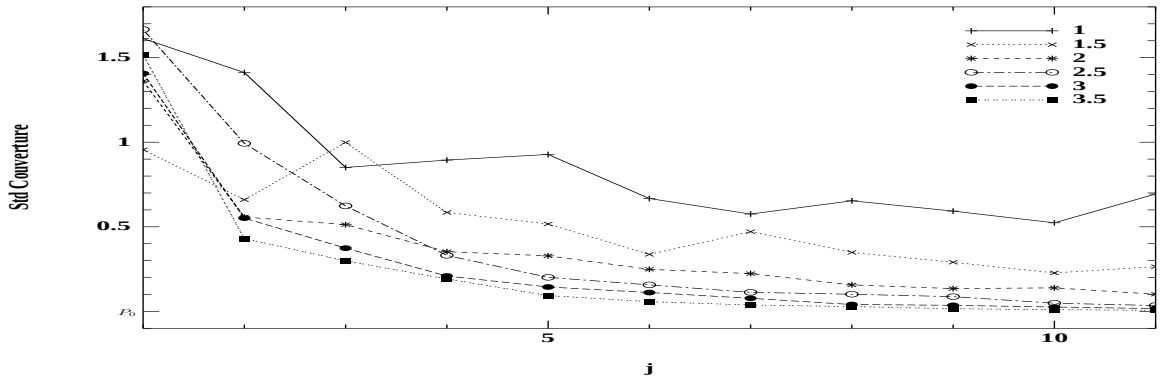
– $\alpha_\delta = \mathbf{1}$

Le seuil pour M donne $\alpha_{\mathbf{M}} > \mathbf{3}$ pour que l'algorithme converge.



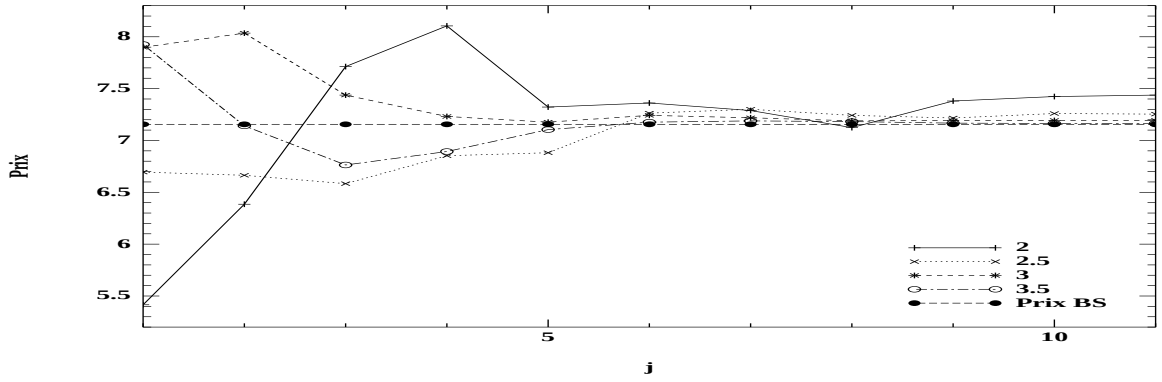


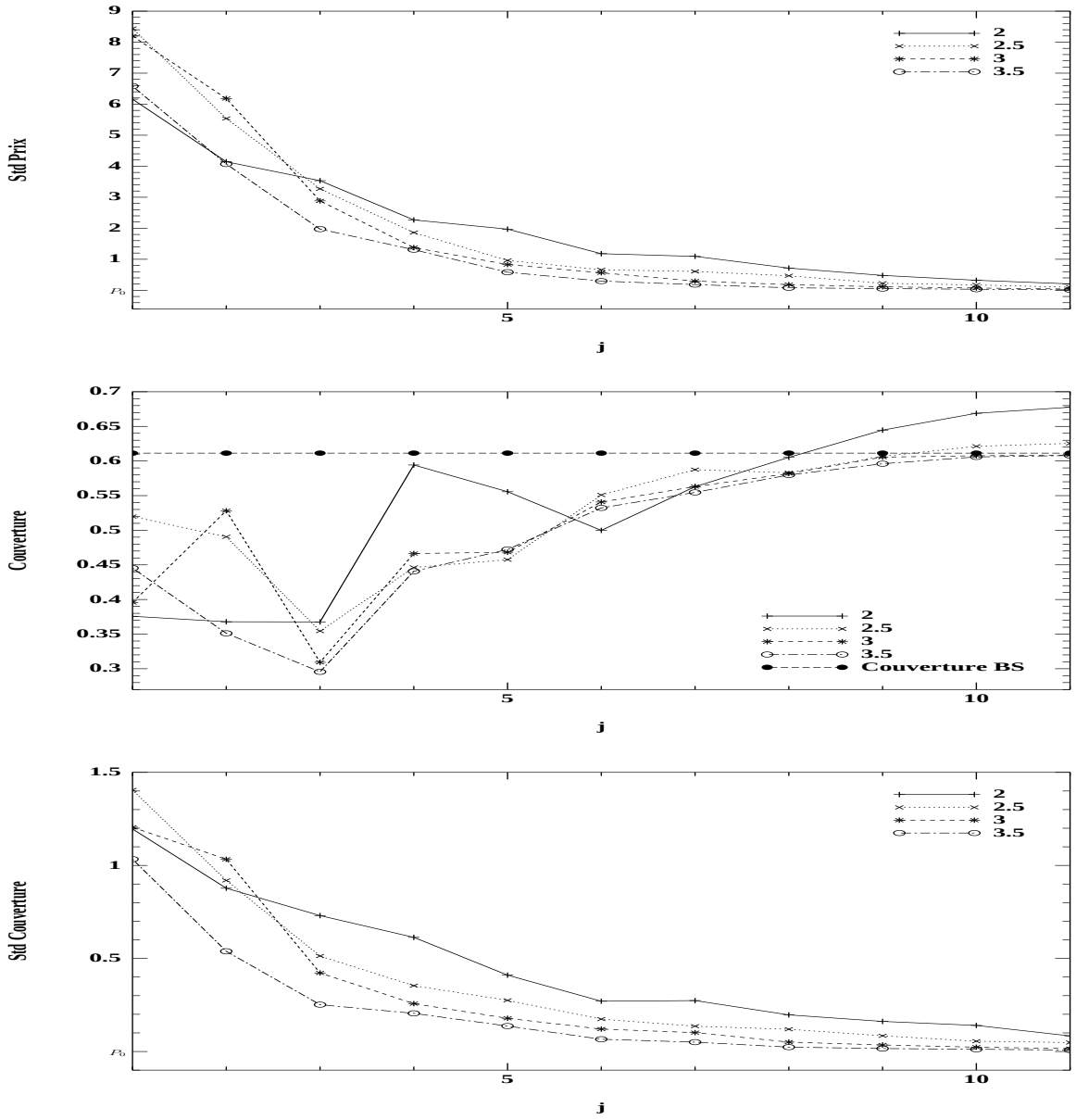




– $\alpha_\delta = \mathbf{1.5}$

Le seuil pour M est de $\alpha_{\mathbf{M}} > \mathbf{3.5}$.





On peut alors tirer plusieurs conclusions de l'examen de ces courbes :

- Pour n'importe quelle valeur de α_δ, dès que $\alpha_M > 1$, l'écart-type de l'approximation du prix ou de la couverture décroit et d'autant plus vite que α_M est grand.
- Le biais ne tend pas vers 0 dans tous les cas. Ainsi pour $\alpha_\delta = 1$, le biais diverge clairement pour $\alpha_M = 1, 1.5$, semble diverger pour $\alpha_M = 2$ et converge pour $\alpha_M \geq 2.5$. Ceci est plus optimiste que les seuils donnés dans la proposition IV.1 et que l'on a rappelés avant chaque graphique. On peut consulter la remarque plus bas pour une explication possible.
- L'approximation ponctuelle de la couverture tend vers la couverture Black-Scholes si l'approximation du prix est suffisamment bonne. En effet pour $\alpha_\delta = 1$, la couverture

ne semble pas converger clairement vers la couverture BS, alors que pour $\alpha_\delta = 1.5$, c'est le cas. Remarquons que les théorèmes établis dans les parties I et II ne concernent jamais des résultats de convergence sur des couvertures ponctuelles mais toujours sur des intégrales de la couverture sur l'intervalle $[0,T]$. L'observation de convergence pour des couvertures ponctuelles s'explique certainement par $hZ^{N,R}_{t_k} = \mathbb{E}_{t_k}(Y^{N,R}_{t_{k+1}}\Delta W_k)$ où on comprend qu'une bonne approximation de $Y^{N,R}$ induit une bonne approximation de $Z^{N,R}$.

– Théoriquement avec $\alpha_\delta = 0.6$, on devrait commencer à converger vers le vrai prix pour M suffisamment grand. Néanmoins, la vitesse de convergence est tellement lente que l'on ne peut l'observer avec un nombre de tests raisonnable.

On voit donc que l'algorithme se comporte de manière cohérente avec ce que l'on peut déduire des estimations de la partie II et même plutôt mieux que prévu. En effet, les seuils α_M en fonction de α_δ nécessaires pour observer une convergence de l'algorithme sont plus bas que ceux prévus par la proposition IV.1 et on comprend mal comment les seuils apparaissant numériquement peuvent faire converger (par exemple) le terme

$$\frac{CC_y(R)^2}{h}\exp(-\frac{Mh^{\beta+1}}{144C_y(R)^2})\sum_{k=0}^{N-1}\mathbb{E}\bigg(\mathcal{N}_2\big(\frac{h^{\frac{\beta+1}{2}}}{3\sqrt{2}},[\mathcal{P}_{0,k+1}]_y - y^{N,R}_{k+1}(\cdot),(P^{N,m}_{t_{k+1}})_{1\le m\le M},(\tilde{P}^{N,m}_{t_{k+1}})_{1\le m\le M}\big)\bigg).$$

Or, il apparait que l'on pourrait effectuer à certains endroits de meilleures majorations. Par exemple, lorsque dans la partie II, on veut majorer $\mathbb{P}([A^M_k]^c)$ (d'où provient le terme ci-dessus), on obtient

$$\begin{aligned}
&\mathbb{P}\bigg(\frac{\frac{1}{M}\sum_{m=1}^M U_m\{|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\}}{\sqrt{\frac{1}{M}\sum_{m=1}^M\{|g_j(P^{N,m}_{t_{k+1}})|^2 + |g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2\}}} > \frac{1}{3}h^{\frac{\beta+1}{2}}\bigg|(P^{N,m}_{t_{k+1}},\tilde{P}^{N,m}_{t_{k+1}})_{1\le m\le M}\bigg)\\
\le& 2\exp\bigg(-\frac{Mh^{\beta+1}\frac{1}{M}\sum_{m=1}^M\{|g_j(P^{N,m}_{t_{k+1}})|^2 + |g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2\}}{18\frac{1}{M}\sum_{m=1}^M\big(|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\big)^2}\bigg).
\end{aligned}$$

Pour arranger le majorant de l'inéquation ci-dessus, on écrit alors

$$\begin{aligned}
&\frac{1}{M}\sum_{m=1}^M\big(|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 - |g_j(P^{N,m}_{t_{k+1}})|^2\big)^2\\
=&\frac{1}{M}\sum_{m=1}^M\big(|g_j(\tilde{P}^{N,m}_{t_{k+1}})| + |g_j(P^{N,m}_{t_{k+1}})|\big)^2\big(|g_j(\tilde{P}^{N,m}_{t_{k+1}})| - |g_j(P^{N,m}_{t_{k+1}})|\big)^2\\
\le&\frac{8C_y(R)^2}{M}\sum_{m=1}^M\big(|g_j(\tilde{P}^{N,m}_{t_{k+1}})|^2 + |g_j(P^{N,m}_{t_{k+1}})|^2\big). \qquad \text{(IV.4.2)}
\end{aligned}$$

Il apparait trop fort de majorer $\big(|g_j(\tilde{P}^{N,m}_{t_{k+1}})| - |g_j(P^{N,m}_{t_{k+1}})|\big)^2$ par $4C_y(R)^2$. En effet, étant donné que $\tilde{P}^{N,m}_{t_{k+1}} = T_k(P^{N,m}_{t_k},\Delta\tilde{W}^m_{k,k+1})$ et $P^{N,m}_{t_{k+1}} = T_k(P^{N,m}_{t_k},\Delta W^m_{k,k+1})$ où on rappelle

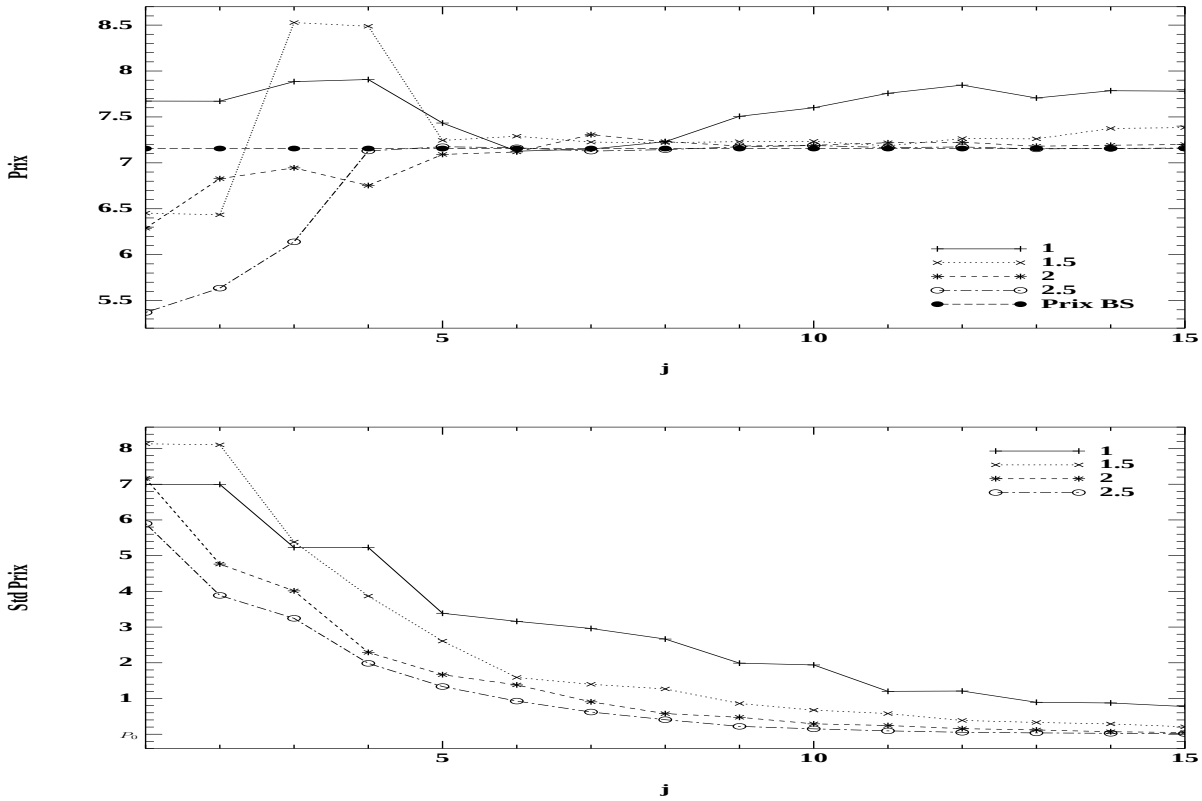
que T_k désigne la transition de la chaine de Markov P^N, il y a de fortes chances que $g_j(\tilde{P}^{N,m}_{t_{k+1}})$ et $g_j(P^{N,m}_{t_{k+1}})$ soient proches. Ainsi, si $P^N = X^N$ et que la classe $\mathcal{G}$ (qui recouvre $[\mathcal{P}_{0,k+1}]_y - y^{N,R}_{k+1}$) est une classe de fonctions Lipschitz on obtient :

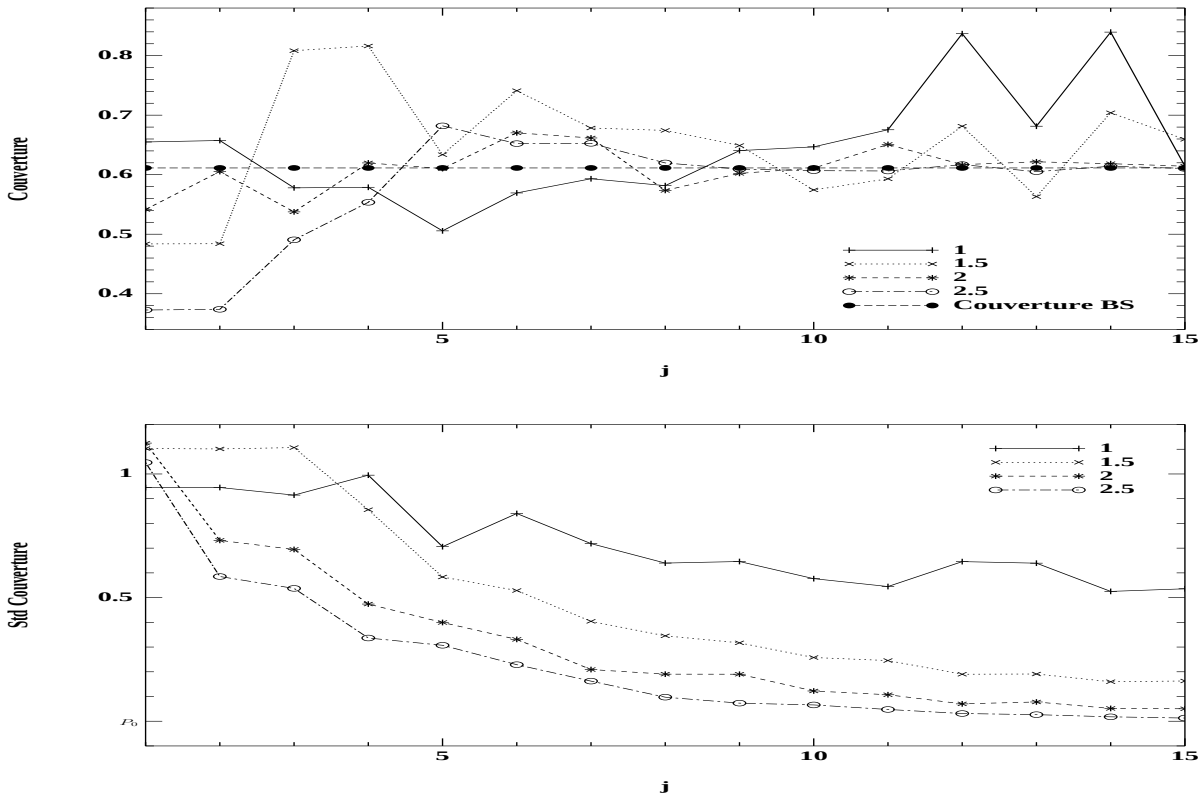
$$|g_j(\tilde{P}^{N,m}_{t_{k+1}}) - g_j(P^{N,m}_{t_{k+1}})| \leq C|\Delta W^m_k - \Delta \tilde{W}^m_k|$$

et on voit que l'on gagne un facteur h. En général, comme la classe $[\mathcal{P}_{0,k+1}]_y$ n'est pas une classe de fonctions Lipschitz $[\mathcal{P}_{0,k+1}]_y - y^{N,R}_{k+1}$ n'a pas de raison d'être une classe de fonctions Lipschitz et donc $\mathcal{G}$ non plus. Mais dans notre cas particulier de la base **HC**, il n'est pas difficile de se convaincre que la majoration (IV.4.2) est tout de même trop forte (voir également la sous-section 4.2.3 où on s'aperçoit que les fonctions estimées ne sont pas loin d'être Lipschitz).

4.2.2 Tests avec la base HC(1,0)

Nous prenons cette fois $\alpha_\delta = 0.6$ et la base **HC(1,0)**. Voici les résultats obtenus pour différentes valeurs de α_M.





On constate donc que si la base **HC** donnait des convergences très lentes pour le prix avec $\alpha_\delta = 0.6$ on gagne ici beaucoup à améliorer l'approximation sur $y^{N,R}$ en prenant des polynômes de degré 1 sur chaque hypercube. On constate également la présence de seuils de convergence pour α_M. Dans cet exemple, il semble nécessaire que $\alpha_M > 2$. En outre avec la base **HC** et $\alpha_\delta = 0.6$ on n'observait pas de convergence sur la couverture en t_0. Ayant amélioré l'approximation de $y^{N,R}$ on constate à présent une convergence de la couverture.

4.2.3 Tracé des fonctions de prix

L'algorithme ne donne pas seulement un prix à l'instant t_0 mais également un prix à tous les instants t_k pour $0 \leq k \leq N$. Il est intéressant de tracer la déformation des fonctions de prix, en partant de $j = N$ où la fonction de prix est égale à la fonction payoff. Dans le cas de la base **VC** avec les paramètres $N = 20$, $K_0 = 20$ cellules de Voronoï et $M = 5000$ (et pour un seul lancer de l'algorithme), nous traçons les fonctions de prix données par l'algorithme et par la formule de Black-Scholes (avec un taux d'intérêt R) aux instants t_j pour $j = 4, 9, 14, 20$:

On remarque qu'assez naturellement, l'erreur sur le prix est grande pour des valeurs de S élevées : peu de simulations sont tombées dans cette zone.

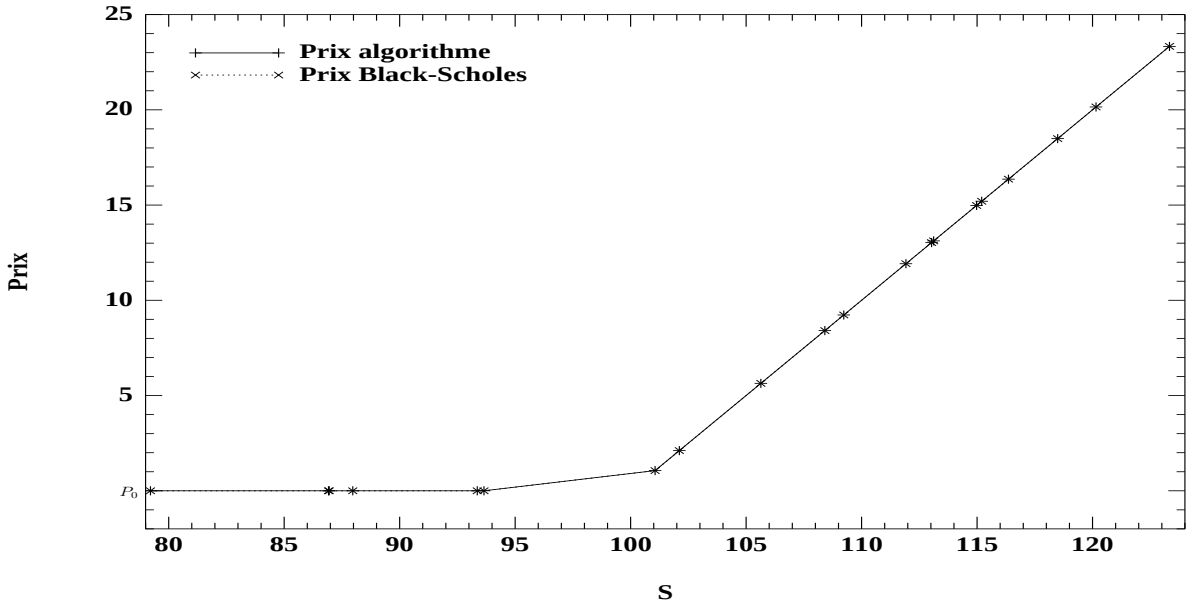


FIG. 4.1 – Fonction de prix en $t_{20} = T = 0.5$

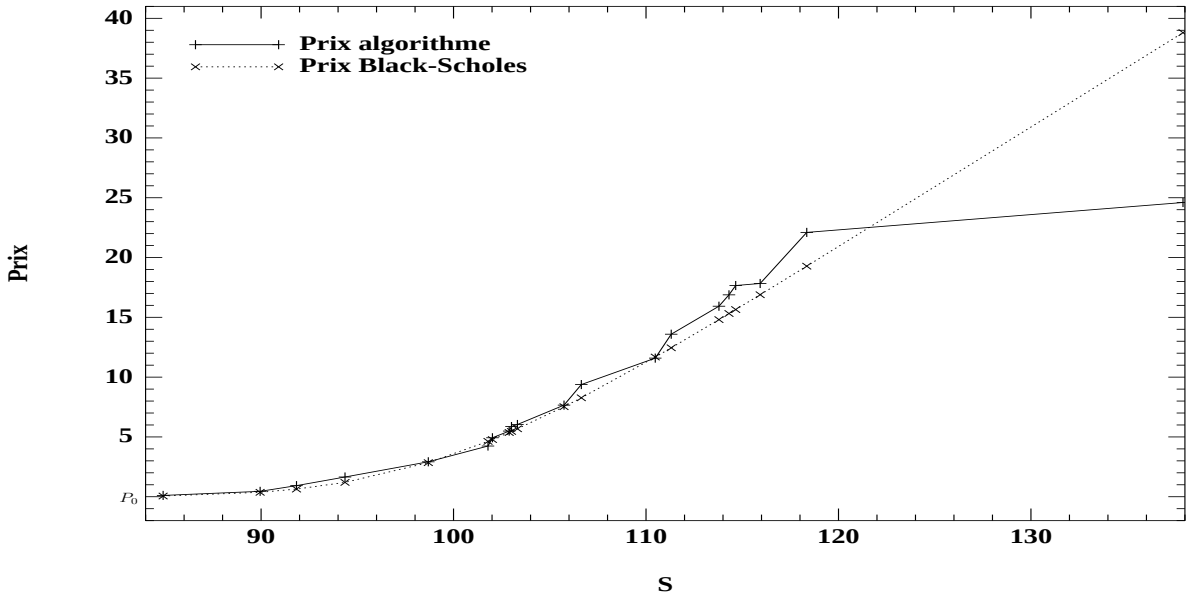


FIG. 4.2 – Fonction de prix en $t_{14} = 0.35$

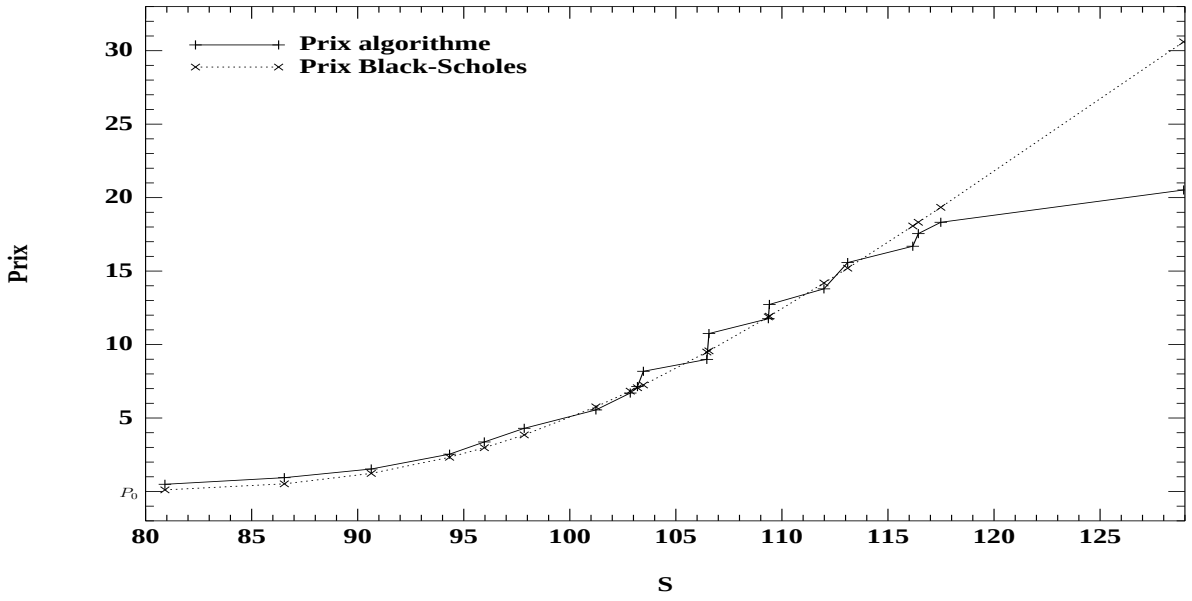


FIG. 4.3 – Fonction de prix en $t_9 = 0.22$

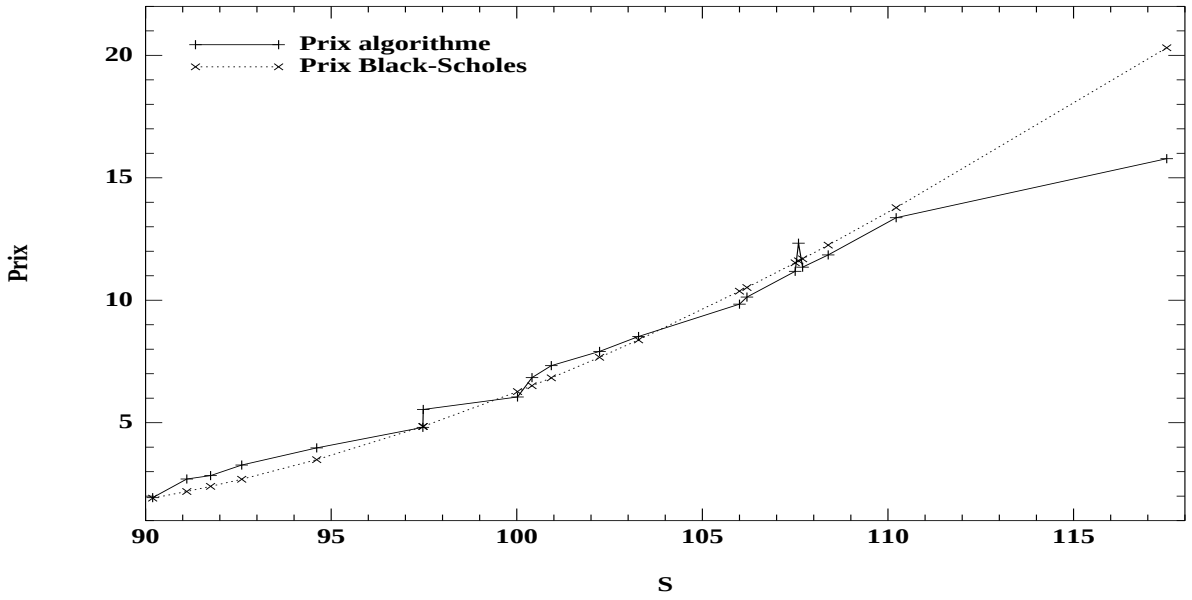


FIG. 4.4 – Fonction de prix en $t_4 = 0.1$

4.3 Tests sur une EDSR réfléchie

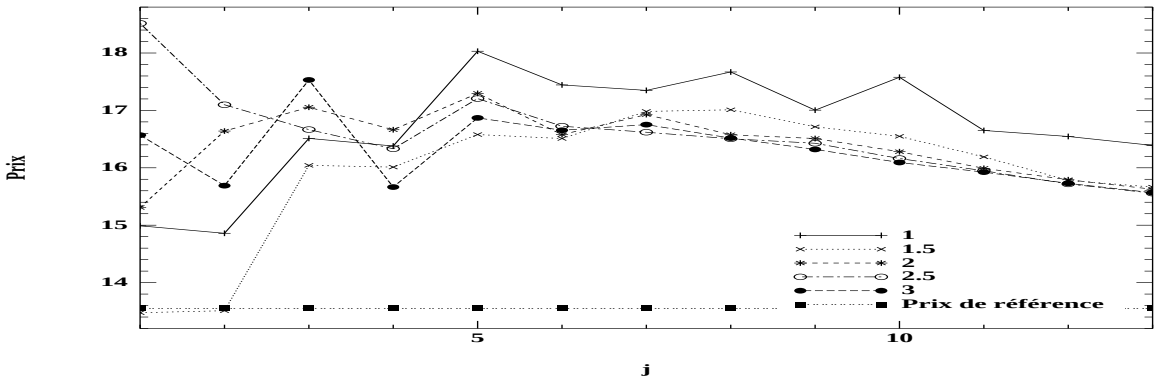
Nous allons tester à présent la variation simultanée des paramètres N, δ et M dans le cas de la base **HC** et pour l'algorithme présenté dans la partie III, chapitre 2, pour les EDSR réfléchies, où on prend le maximum à chaque pas de temps entre le résultat d'un problème de régression et l'obstacle. Dans ce cas, comme indiqué dans la partie III, les théorèmes II.2, II.3, II.4 et II.5 sont encore valables dans ce cas et donc la proposition IV.1 également. Nous nous attendons donc à observer le même seuil de convergence fonction de α_δ pour α_M.
Nous choisissons $d = q = 1$. L'EDSRR que nous voulons approcher est la suivante :

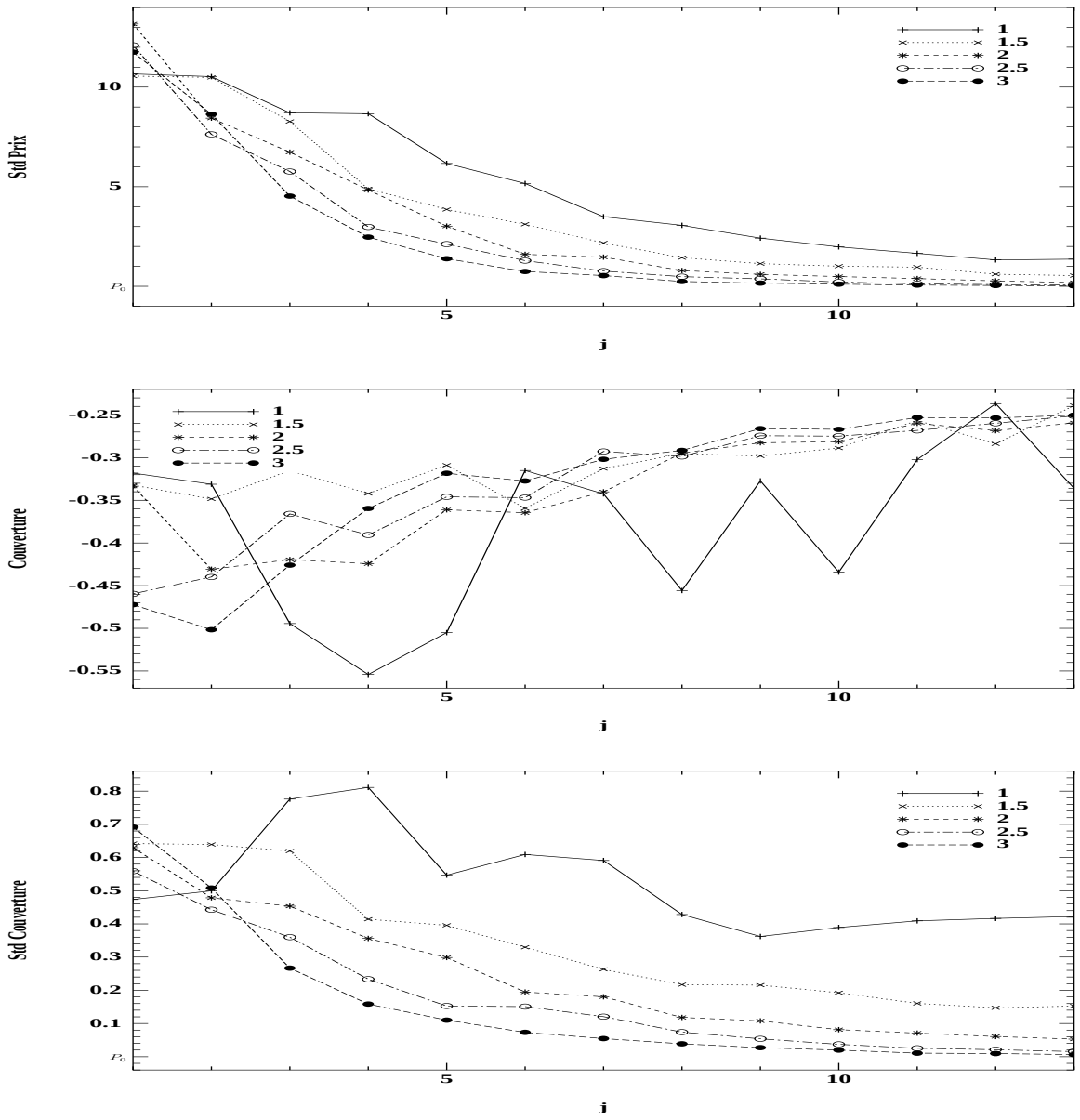
$$\begin{aligned}
&\frac{dS_t}{S_t} = \mu dt + \sigma dW_t, \\
&dY_t = (-rY_t - \theta Z_t)dt - Z_t dW_t + dK_t, \\
&Y_T = (K - S_T)_+, Y_t \geq (K - S_t)_+, 0 \leq t \leq 1, \\
&\int_0^T \{Y_t - (K - S_t)_+\} dK_t = 0
\end{aligned}$$

avec $\theta = \frac{\mu - r}{\sigma}$. Si l'on compare avec les notations du Chapitre 2, cet exemple correspond à $f(t, x, y, z) = -ry - \theta z$ et $g(t, x) = (K - e^x)_+$. Ceci correspond donc au cas d'un put américain de strike K et couvert sans contrainte de portefeuille (d'où un driver f linéaire en (y, z)).
Les paramètres de la dynamique de S et de l'option sont les suivants :

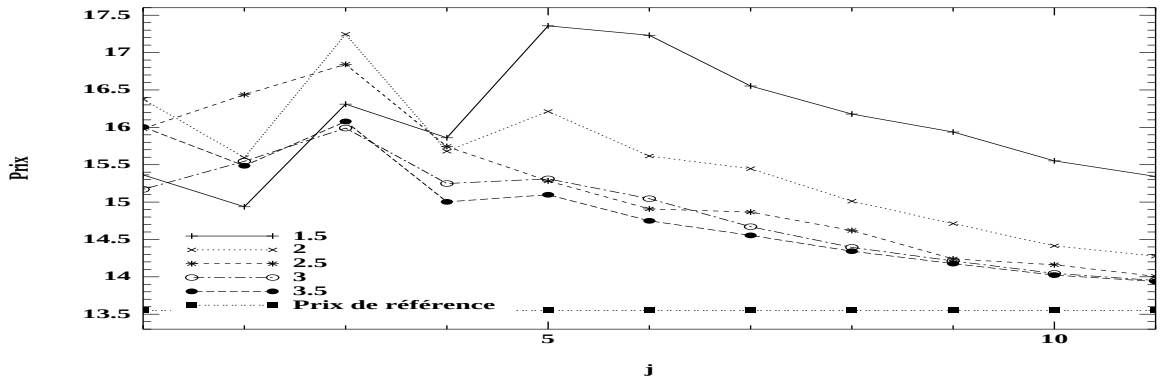
μ	σ	r	T	S_0	K
0.05	0.4	0.05	1	100	100

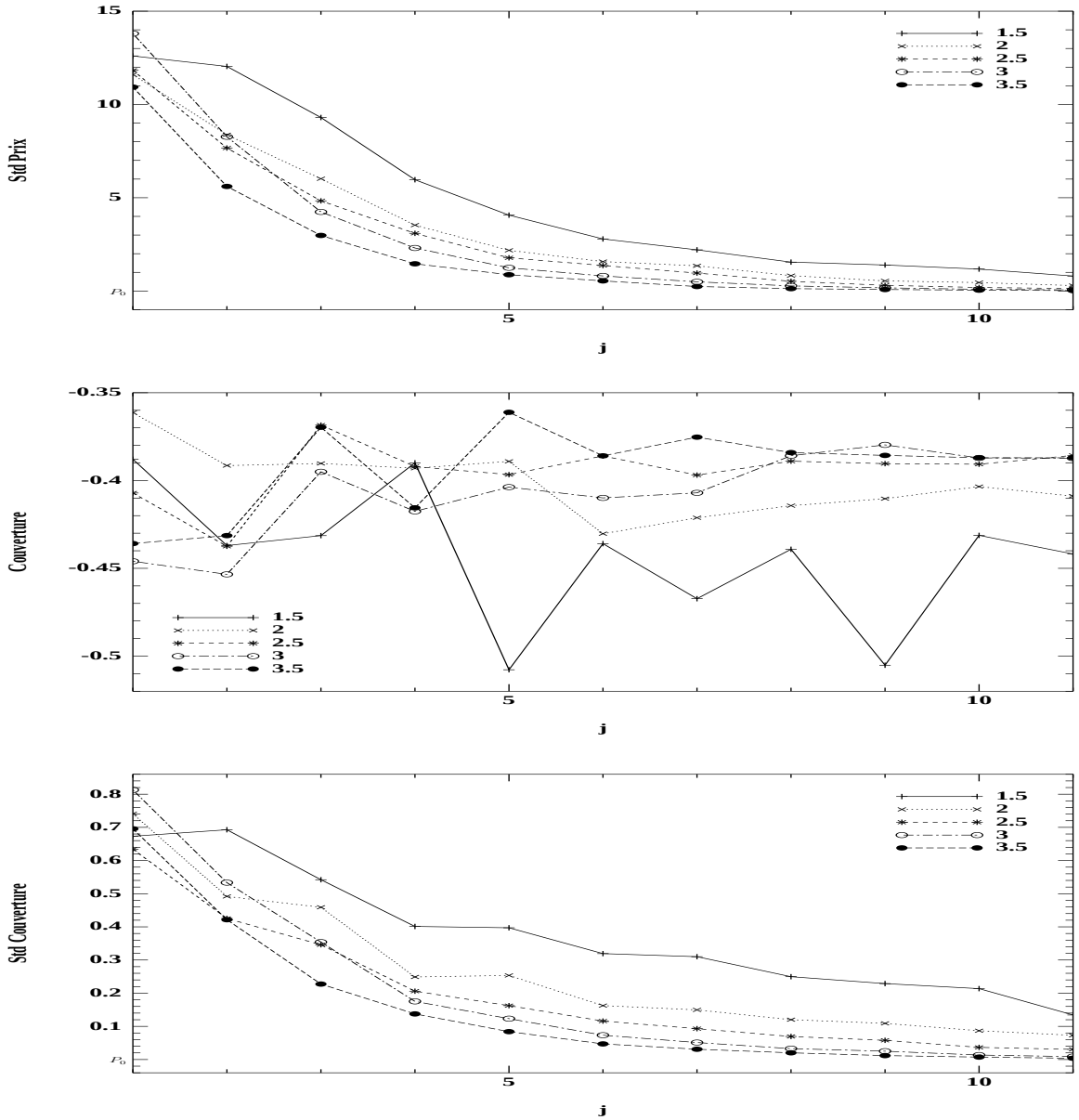
Cet exemple est tiré de [BT04] où les auteurs donnent un prix de référence $\mathbf{Y_0 = 13.55}$. Nous utilisons la même procédure pour tester la variation simultanée des paramètres que pour le cas de l'option avec différentiel de taux d'intérêt.
– $\alpha_\delta = \mathbf{0.6}$



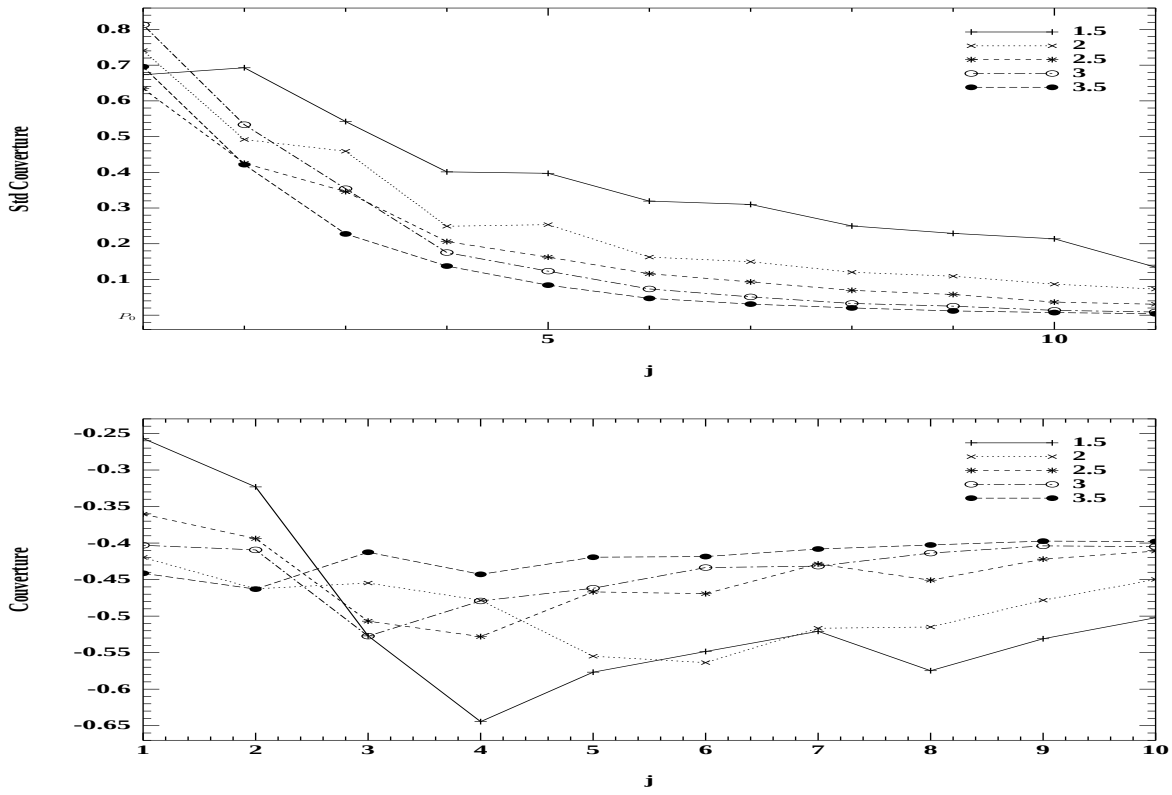


– $\alpha_\delta = 1$.

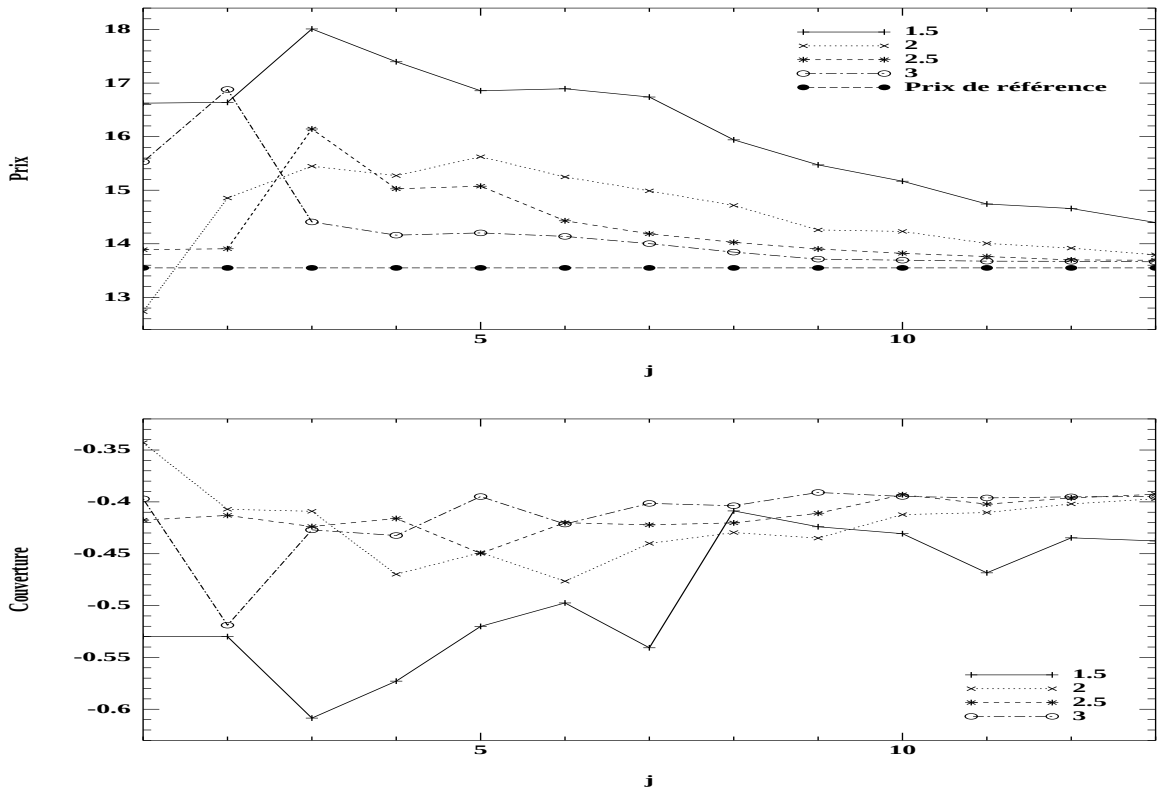




– $\alpha_\delta = \mathbf{1.5}$. On change la valeur de μ qui vaut à présent 0.08. On ne se trouve donc plus sous la probabilité risque-neutre. On veut vérifier que l'algorithme corrige bien le drift dans le cas réfléchi.



– $\alpha_\delta = \mathbf{0.6}$ avec une base $\mathbf{HC(1,0)}$. $\mu = 0.08$.



On observe également l'existence de seuils de convergence pour α_M en fonction de la valeur de α_δ. Ces seuils sont comparables à ceux observés pour le cas EDSR non réfléchie. De plus, en comparant des valeurs de α_δ identiques pour le cas réfléchi et non-réfléchi, on s'aperçoit que la convergence dans le cas réfléchi est plus lente et que l'on a tendance à approcher le prix américain par au-dessus. On constate également en observant les cas $\alpha_\delta = 1.5$ avec la base **HC** et $\alpha_\delta = 0.6$ avec la base **HC(1,0)** que l'algorithme corrige le drift pour retrouver le prix risque-neutre. Ces deux exemples montrent également que lorsque l'approximation de Y est suffisamment bonne, l'approximation sur la couverture est bonne. Enfin, en comparant les cas $\alpha_\delta = 0.6$ pour les bases **HC** et **HC(1,0)** on constate qu'on gagne beaucoup à ajouter des polynômes locaux de degré 1 pour approcher Y.

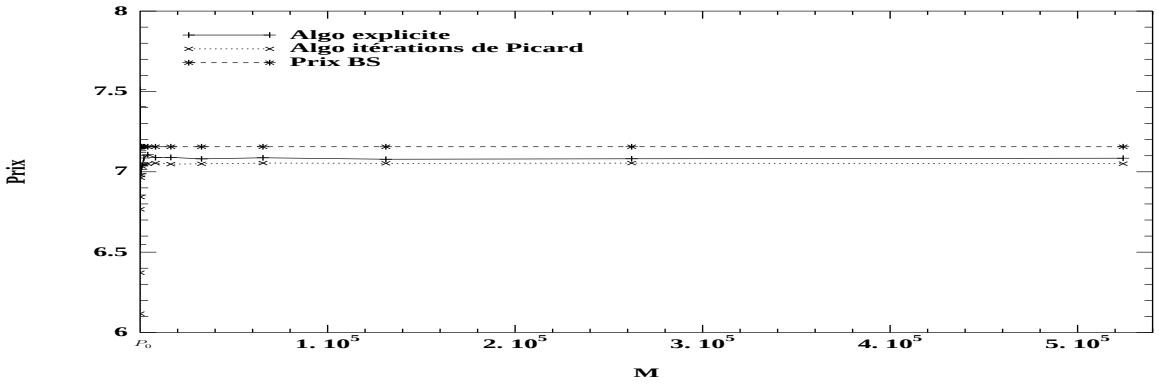
Chapitre 5

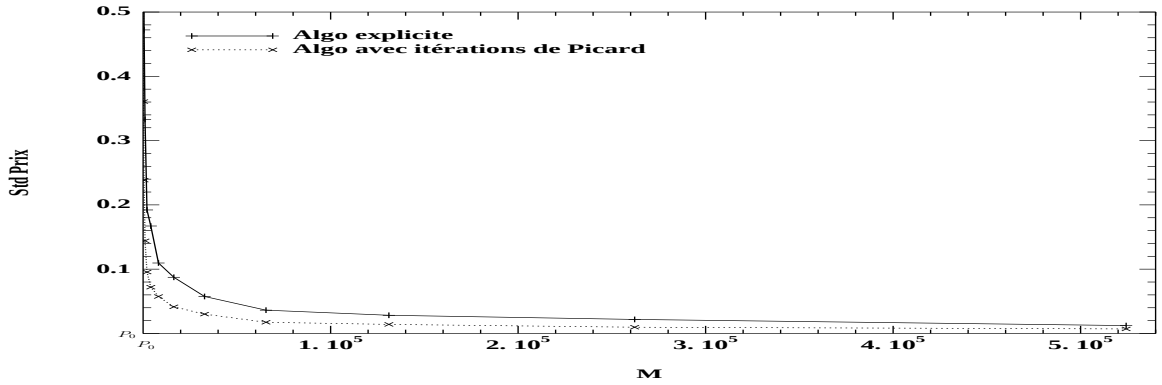
Comparaison des algorithmes des parties I et II

Après avoir vérifié la validité des résultats de convergence donnés dans la partie II, nous pouvons comparer les deux algorithmes : celui de la partie I et celui de la partie II. L'un est basé sur des itérations de Picard à chaque instant de discrétisation t_k et cherche à approcher en même temps Y_{t_k} et Z_{t_k}. Le second n'a pas besoin d'itérations de Picard car le schéma est explicite, et trouve par des régressions séparées Y et Z. Nous allons voir comment se comportent les deux algorithmes sur un exemple type.

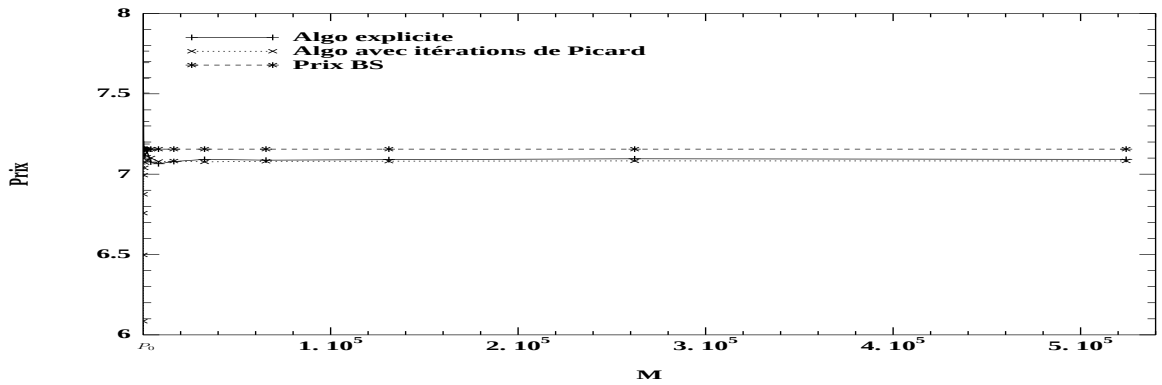
Nous nous plaçons dans le cas de l'option européenne avec différentiel de taux d'intérêt, avec une base **HC**. Le résultat du théorème I.3 ne permet théoriquement pas de faire tendre en même temps N, M vers l'infini et δ vers 0. Ainsi dans un premier temps, nous allons fixer N et δ, un domaine $D = [40, 180]$ et faire tendre M vers l'infini. Nous allons alors comparer les convergences des deux algorithmes.

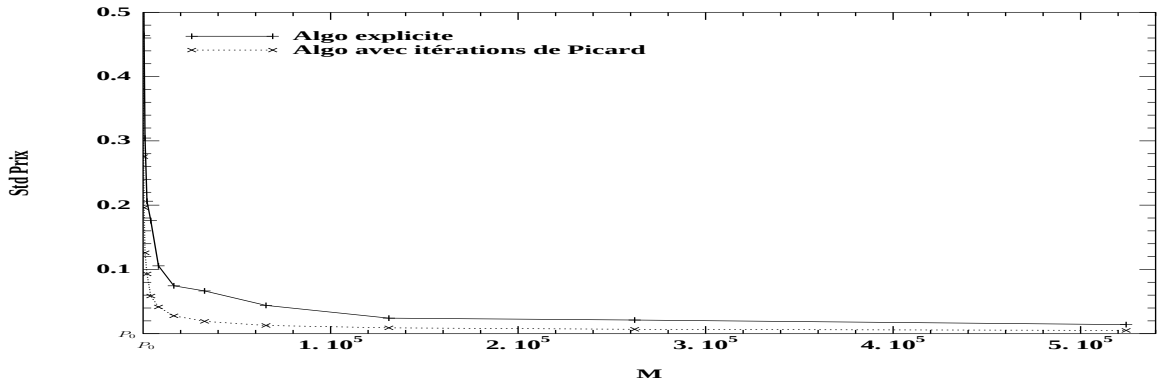
– $N = 5$, $\delta = 20$



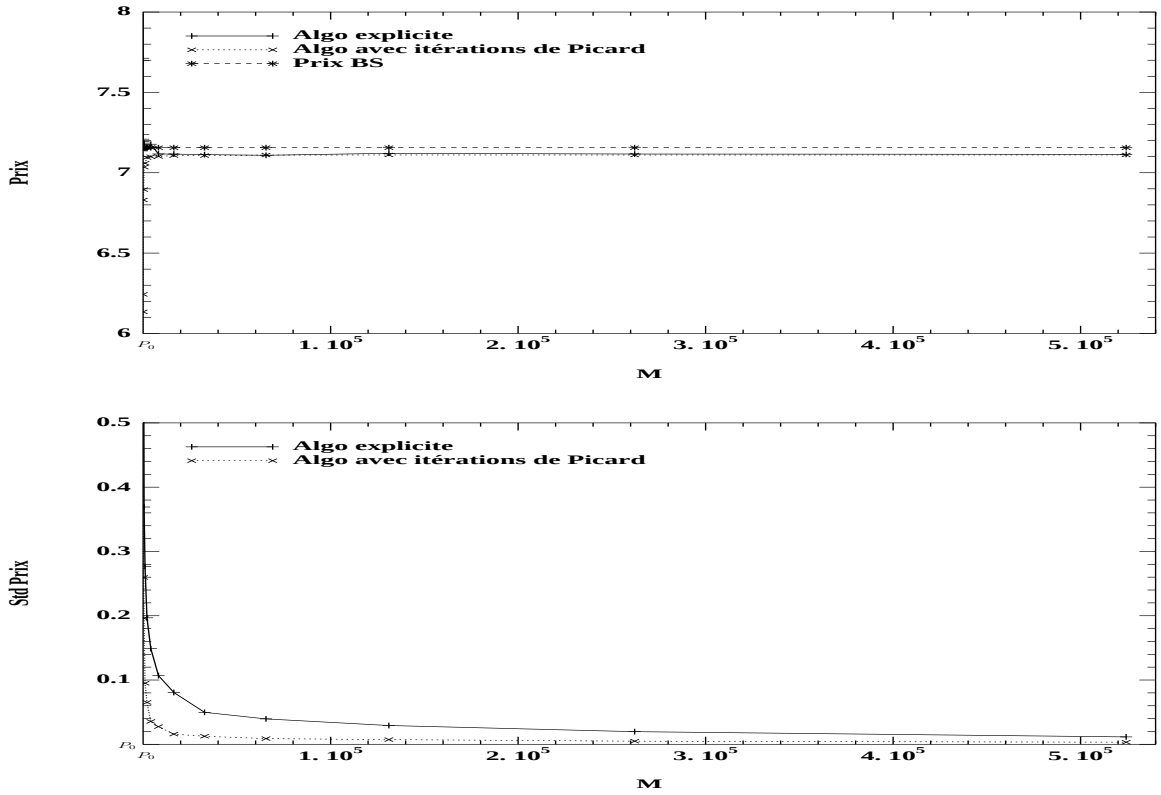


– $N = 10$, $\delta = 10$

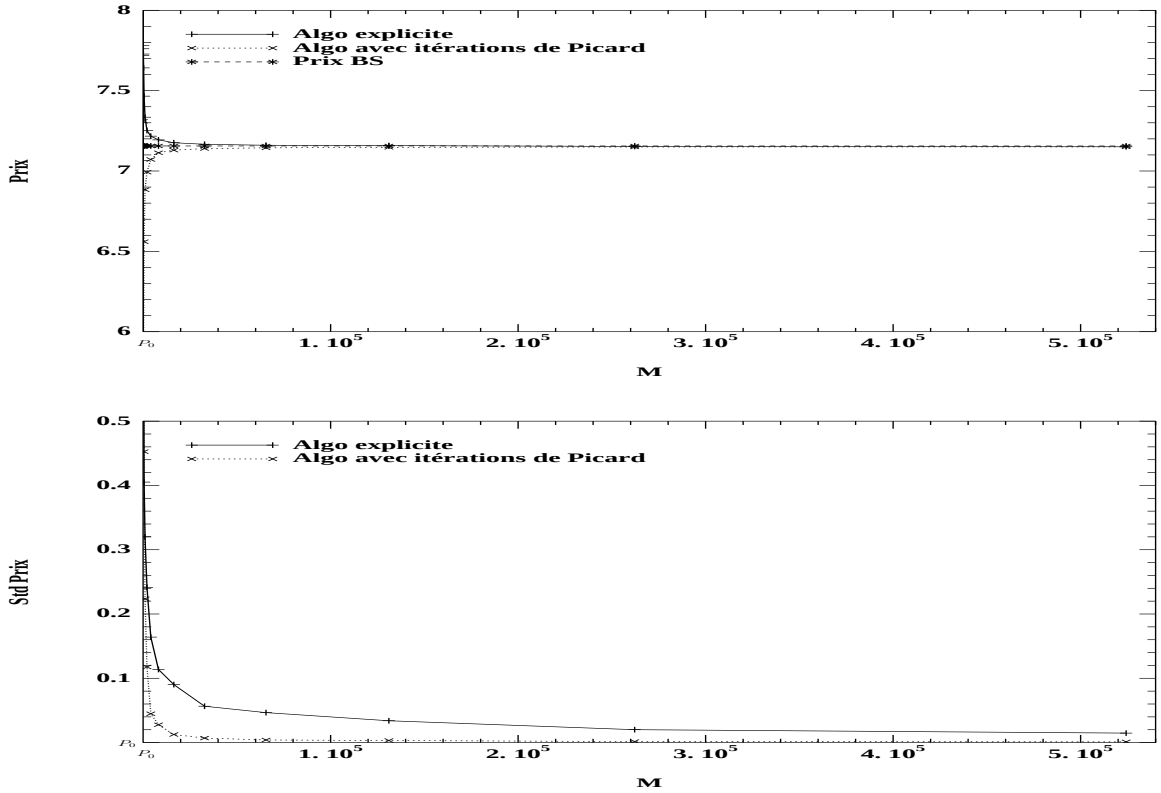




– $N = 20$, $\delta = 5$


Algo explicite
Algo avec itérations de Picard
Prix BS
Prix
M
Algo explicite
Algo avec itérations de Picard
Std Prix
M

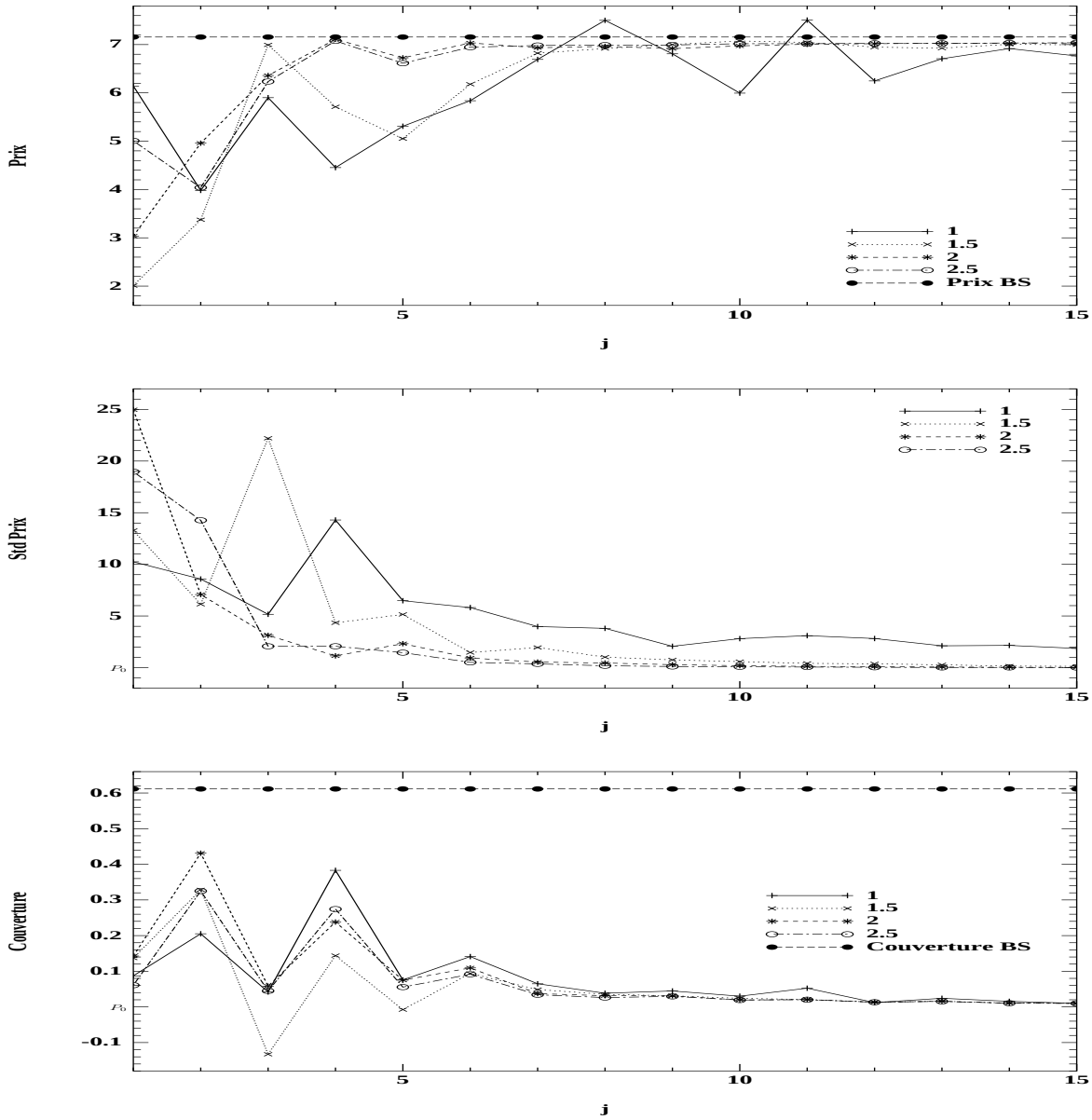
– $N = 50$, $\delta = 1$

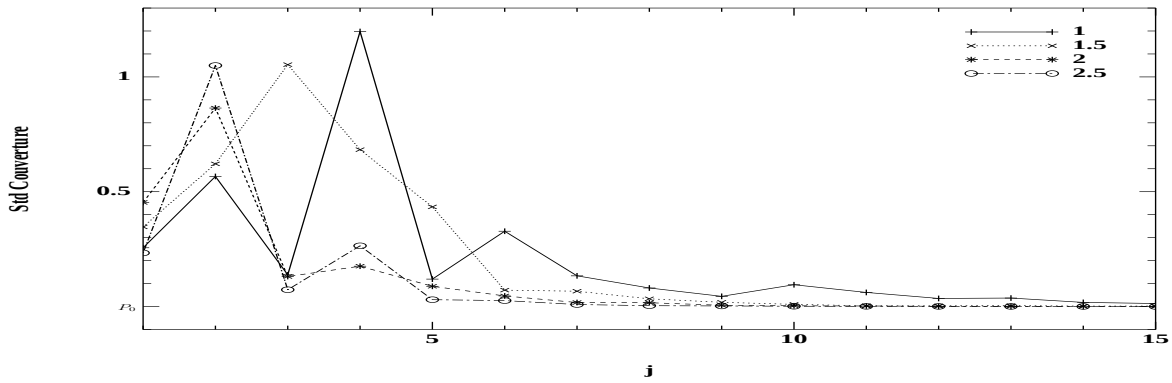

Algo explicite
Algo avec itérations de Picard
Prix BS
Prix
M
Algo explicite
Algo avec itérations de Picard
Std Prix
M

On peut remarquer que les algorithmes se comportent essentiellement de la même façon à ceci près que l'algorithme avec itérations de Picard tend à donner une plus petite variance que l'algorithme explicite. On peut penser que la cause est la minimisation simultanée en (Y, Z) : la partie en $Z\Delta W$ peut jouer le rôle de variable de contrôle. On peut enfin remarquer que le biais entre le prix donné par l'algorithme et le prix de référence est légèrement plus faible dans le cas de l'algorithme explicite.

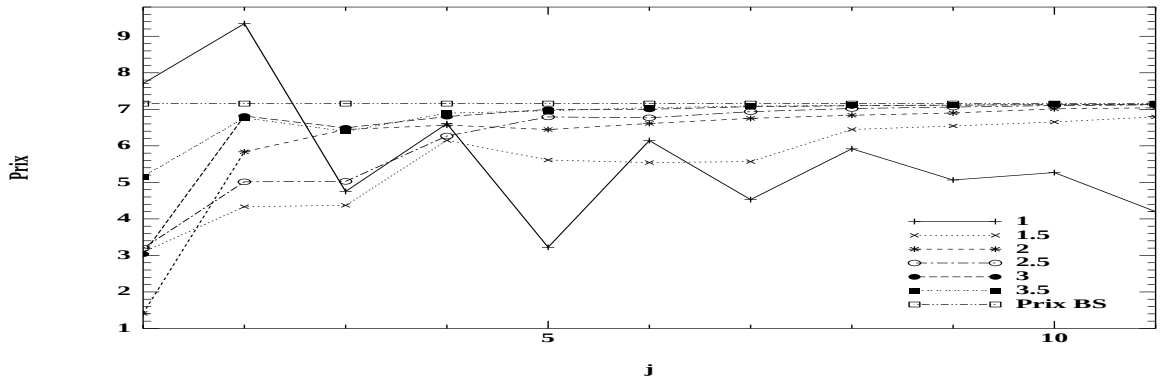
Nous allons à présent essayer de faire varier simultanément les paramètres dans l'algorithme avec itérations de Picard, pour voir si les résultats sont identiques à ceux observés pour l'algorithme explicite.

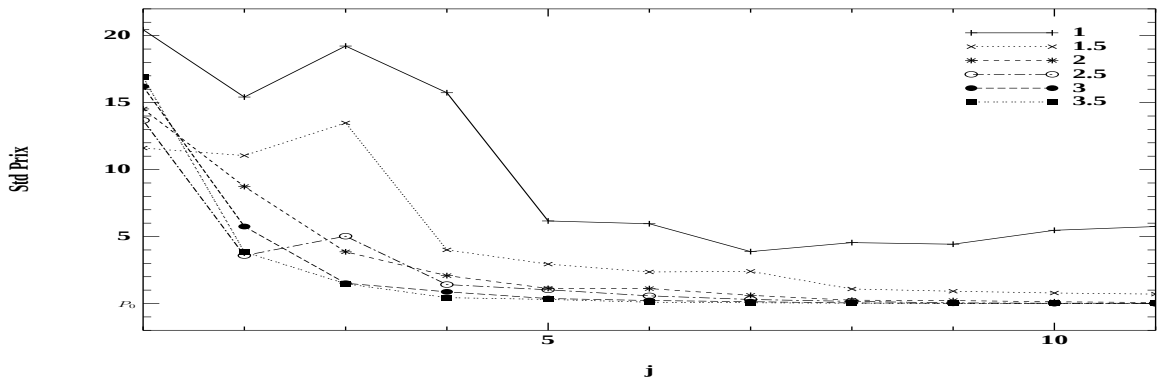
– $\alpha_\delta = 0.6$ et α_M varie.

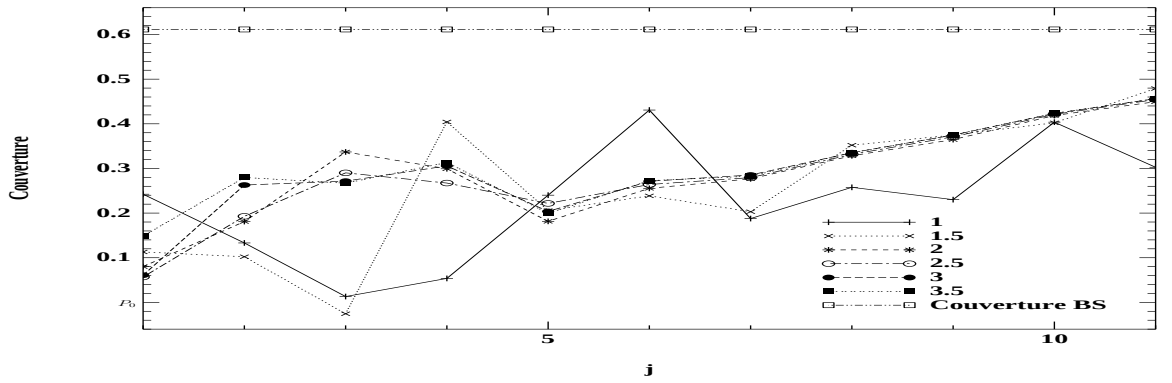


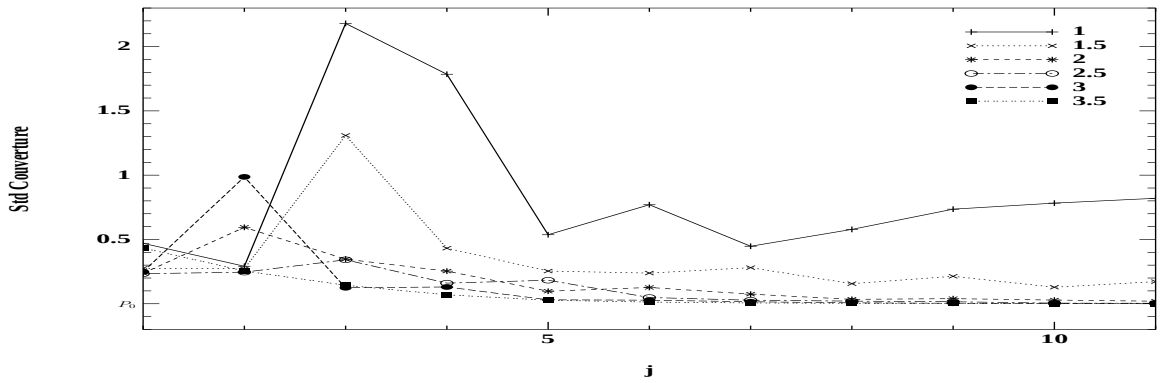


– $\alpha_\delta = 1$

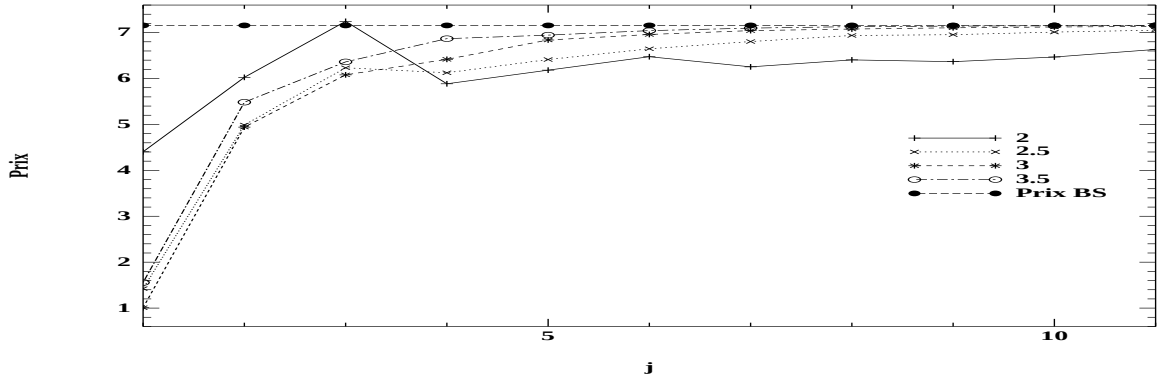


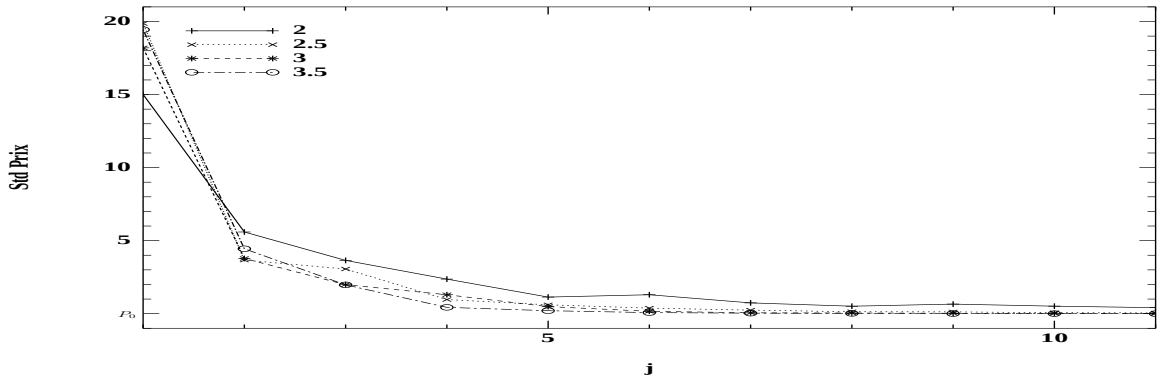


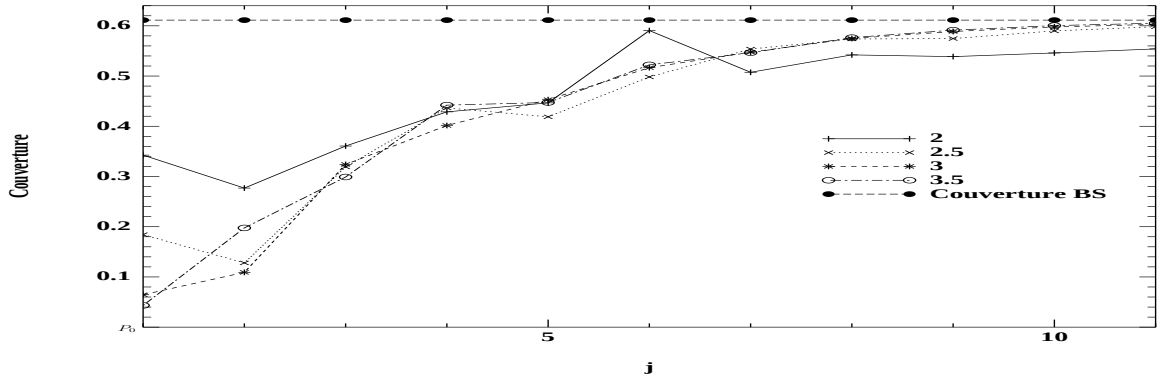


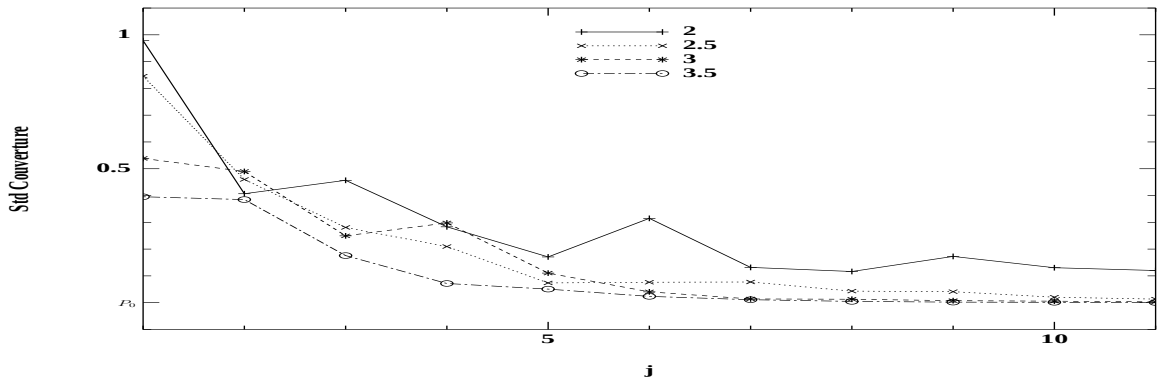


– $\alpha_\delta = 1.5$









Les résultats sont comparables à ceux obtenus avec l'algorithme explicite. On vérifie l'existence d'un seuil pour α_M, fonction de α_δ, en dessous duquel le biais semble ne pas tendre vers 0 et au dessus duquel l'algorithme semble converger normalement. Ainsi, il semble numériquement qu'il soit possible d'analyser l'algorithme avec itérations de Picard de la même façon que l'algorithme explicite.

Chapitre 6

Options américaines

Nous avons déjà testé la variation simultanée des paramètres N, δ, et M dans le cas de la base **HC** pour une EDSR réfléchie. Nous allons dans ce chapitre traiter plus en détail la valorisation et couverture d'options américaines. Outre l'algorithme présenté dans la partie III, chapitre 2, où l'on prend à chaque instant de discrétisation t_k, le maximum entre l'obstacle et la solution du problème de moindres carrés, on peut imaginer également deux autres méthodes pour approcher la solution d'une EDSR. Nous allons à présent détailler ces deux autres méthodes :

Méthode de pénalisation

Une méthode de preuve de l'existence et de l'unicité à l'EDSR réfléchie

$$\begin{aligned} &Y_t = g(T, X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds + K_T - K_t - \int_t^T Z_s dW_s, \\ &Y_t \geq g(t, X_t), \int_0^T (Y_t - g(t, X_t))dK_t = 0 \end{aligned} \tag{IV.6.1}$$

consiste à définir une suite d'EDSR non réfléchies

$$Y_t^n = g(T, X_T) + \int_t^T f(s, X_s, Y_s^n, Z_s^n)ds + n\int_t^T (Y_s^n - g(s, X_s))^- ds - \int_t^T Z_s^n dW_s \tag{IV.6.2}$$

de solution (Y^n, Z^n). On définit alors $K_t^n = n\int_0^t (Y_s^n - g(s, X_s))^- ds$, et on montre (voir [EKKP$^+$97]) que lorsque n tend vers l'infini, (Y^n, Z^n, K^n) tend vers (Y, Z, K) solution de (IV.6.1).
L'idée est alors pour n fixé, d'approcher la solution de l'équation (IV.6.2) à l'aide de l'algorithme exposé dans la partie II. En effet, (IV.6.2) est une EDSR avec un driver Lipschitz. Remarquons que vu les résultats de [EKKP$^+$97], $Y_n \nearrow Y$ et que l'on peut

s'attendre avec cette méthode à trouver des bornes inférieures (de plus en plus précises quand $n \to \infty$) du prix américain.

Méthode de régularisation

Une autre idée consiste à utiliser les résultats de [BCFEK02]. Les auteurs considèrent le cas où l'obstacle $g(t, X_t)$ est une semimartingale càdlàg de la forme

$$g(t, X_t) = g(0, X_0) + \int_0^t U_s ds + \int_0^t V_s dW_s + A_t$$

avec U et V des processus prévisibles vérifiant la condition d'intégrabilité $\mathbb{E}\int_0^T \{|V_t|^2 + |U_t|^2\}dt < \infty$ et A un processus càdlàg, adapté, croissant, et vérifiant la condition d'intégrabilité $\mathbb{E}|A_T| < \infty$. Dans ce cas, on sait que chercher une solution à (IV.6.1) revient à chercher un triplet (Y, Z, α) vérifiant l'EDSR :

$$\begin{aligned} Y_t =& g(T, X_T) + \int_t^T \{f(s, X_s, Y_s, Z_s) + \alpha_s \mathbf{1}_{Y_s = g(s,X_s)}[f(s, X_s, g(s, X_s), V_s) + U_s]^-\}ds \\ & - \int_t^T Z_s dW_s, \\ Y_t \geq& g(t, X_t) \end{aligned} \tag{IV.6.3}$$

avec $\mathbb{E}\int_0^T \{|Y_t|^2 + |Z_t|^2 + |\alpha_t|^2\}dt$.
En comparant avec (III.2.1), on constate que cela revient à exploiter le fait que l'on sait déjà, en utilisant la remarque A.3, que le processus croissant K de la solution de l'EDSR (IV.6.1) sera de la forme :

$$dK_t = \alpha_t \mathbf{1}_{Y_t = g(t,X_t)}[f(t, X_t, g(t, X_t), V_t) + U_t]^- dt,$$

avec $0 \leq \alpha_t \leq 1$. L'inconnue pour déterminer K devient donc le processus α. Pour montrer l'existence et l'unicité de l'équation ci-dessus, les auteurs utilisent une régularisation de (IV.6.3). En particulier, une fois choisi un paramètre de régularisation n, ils introduisent une suite de fonctions $\phi_n \in C^\infty$ telles que $0 \leq \phi_n \leq 1$ et :

$$\begin{aligned} \phi_n(x) =& 1 \qquad \text{si} \qquad |x| \leq \frac{1}{2^n}, \\ \phi_n(x) =& 0 \qquad \text{si} \qquad |x| \geq \frac{2}{2^n}. \end{aligned}$$

A partir de cette suite de fonctions, les auteurs définissent une suite d'EDSR paramétrée par n :

$$Y_t^n = g(T, X_T) + \int_t^T \{f(s, X_s, Y_s^n, Z_s^n) + \phi_n(Y_s^n - g(s, X_s))[f(s, X_s, g(s, X_s), V_s) + U_s]^-\}ds$$
$$- \int_t^T Z_s^n dW_s. \qquad \text{(IV.6.4)}$$

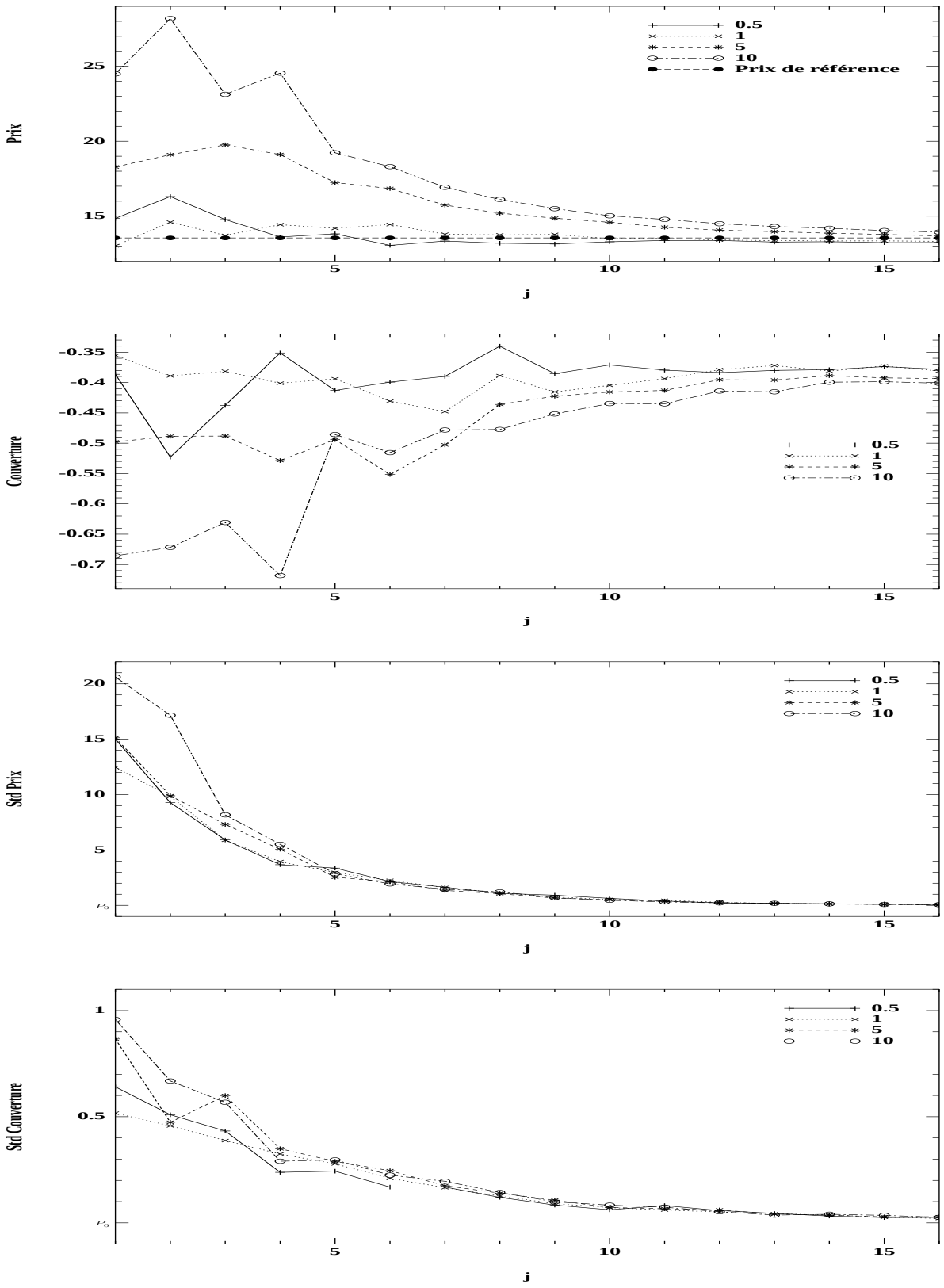
En comparant avec (IV.6.3), on voit que cela revient à régulariser le terme $\alpha_s \mathbf{1}_{Y_s = g(s,X_s)}$. Les auteurs montrent alors que la solution de cette EDSR converge lorsque $n \to \infty$ vers la solution de (IV.6.3). En outre, on a $Y_n \searrow Y$. L'EDSR régularisée (IV.6.4) est une EDSR à driver Lipschitz. On peut donc l'approcher à l'aide de l'algorithme décrit dans la partie II et espérer obtenir des bornes supérieures (de plus en plus précises lorsque $n \to \infty$) du prix américain.

6.1 Tests à coefficient fixés

Ces deux méthodes - pénalisation et régularisation- conduisent à approcher non pas une EDSR mais une suite d'EDSR paramétrée soit par le coefficient de pénalisation soit par le coefficient de régularisation. De plus, la méthode de pénalisation approche Y par en dessous et la méthode de régularisation approche Y par au-dessus.
Nous considérons pour ce test le cas du put américain de la section 4.3, avec les mêmes valeurs de paramètres.
Dans un premier temps, nous allons fixer les paramètres de régularisation et de pénalisation, puis faire varier simultanément N, le nombre de fonctions de base et M. Nous espérons ainsi vérifier un encadrement de Y_0 par les deux méthodes et vérifier que l'encadrement devient plus précis au fur et à mesure que les coefficients de pénalisation et régularisation augmentent.

6.1.1 Méthode de pénalisation

Pour différentes valeurs de n, nous utilisons la base **HC(1,0)** avec $\alpha_\delta = 0.6$ et $\alpha_M = 2$. Nous avons en effet vu dans les chapitres précédents que ce choix de paramètre assurait la convergence du biais et de la variance. Nous montrons les résultats de convergence du prix, de la couverture et de leur écart-type pour n valant $0.5, 1, 5, 10$.



On remarque que quel que soit n, le prix et la couverture convergent, en accord avec la

manière dont on a fait varier N, δ et M. On sait que l'on doit normalement approcher le prix de référence par en dessous. C'est bien ce que l'on observe pour $n = 0.5$ ou 1. Par contre, pour $n = 5$ ou $n = 10$, les dernières valeurs de j testées présentent toujours un prix légèrement supérieur au prix de référence. On peut penser que l'algorithme n'a pas encore convergé pour ces valeurs et que si l'on testait des valeurs de j supérieures, on repasserait en dessous du prix de référence. C'est également cohérent d'observer une convergence moins rapide dans le cas de n grand car le coefficient de Lipschitz du driver intervient dans les estimations d'erreurs que l'on a obtenu. Plus celui-ci est grand, plus les constantes dans les bornes d'erreur sont importantes.

6.1.2 Méthode de régularisation

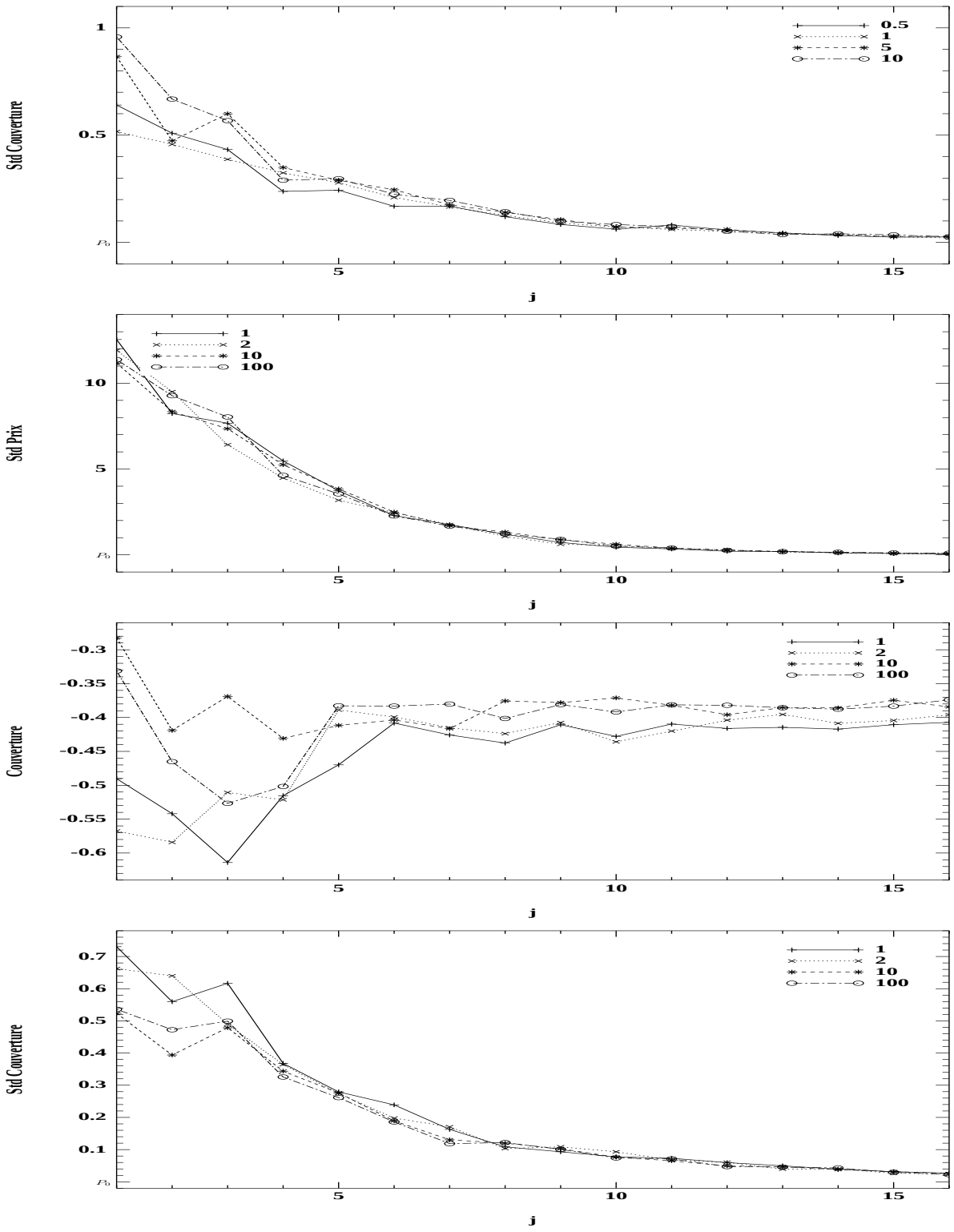
Pour appliquer cette méthode, nous devons spécifier la fonction ϕ_n. Nous modifions légèrement la définition qui en est donnée dans [BCFEK02], étant donné que nous ne prenons pas ϕ_n C^∞ :

$$\begin{aligned}\phi_n(x) =&1 \qquad \text{si} \qquad |x| \leq \frac{1}{2^n},\\ \phi_n(x) =&0 \qquad \text{si} \qquad |x| \geq \frac{2}{2^n}\end{aligned}$$

et ϕ_n reliant linéairement 0 à 1 sur les intervalles $[-\frac{2}{n}, -\frac{1}{n}]$ et $[\frac{1}{n}, \frac{2}{n}]$.
En outre, l'idée de la méthode décrite dans [BCFEK02] est de donner des majorants de Y. Il importe donc de garantir que la discrétisation de (IV.6.4) va bien fournir des approximations par au-dessus. Pour cette raison, on va discrétiser (IV.6.4) de la manière suivante :

$$\begin{aligned}Y^N_{t_k} =&\mathbb{E}_{t_k}\Big(Y^N_{t_{k+1}} + hf(t_k, X^N_{t_k}, Y^N_{t_{k+1}}, Z^N_{t_k})\\ &+ h\phi_n(Y^N_{t_{k+1}} - g(t_{k+1}, X^N_{t_{k+1}}))[f(t_k, X^N_{t_k}, g(t_{k+1}, X^N_{t_{k+1}}), V^N_{t_k})\\ &+ \frac{g(t_{k+1}, X^N_{t_{k+1}}) - g(t_k, X^N_{t_k})}{h}]^-\Big),\\ hZ^N_{t_k} =&\mathbb{E}_{t_k}(Y^N_{t_{k+1}}\Delta W_k),\\ hV^N_{t_k} =&\mathbb{E}_{t_k}(g(t_{k+1}, S^N_{t_{k+1}})\Delta W_k).\end{aligned}$$

L'idée est qu'en utilisant un analogue du théorème de comparaison pour les EDSR discrètes, on peut facilement montrer qu'on a toutes les chances que l'approximation de cette discrétisation donne des prix supérieurs à l'obstacle et donc supérieur à Y.
Ci-dessous, nous présentons les résultats pour différentes valeurs de n.



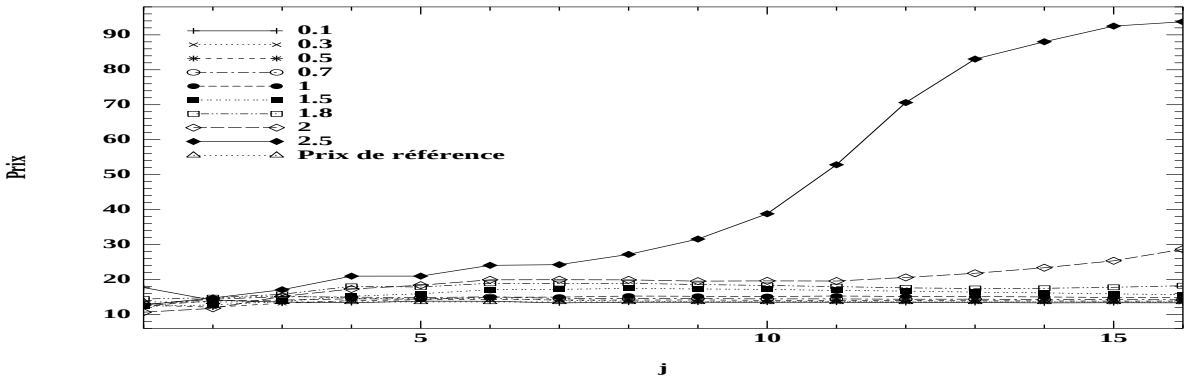
Nous constatons que l'on approche bien le prix américain par au-dessus. L'écart entre le prix vers lequel converge l'algorithme et le prix américain se resserre lorsque n croit. Par contre si n est trop grand ($n = 100$), la régularisation n'intervient pas et le prix vers

lequel l'algorithme converge est inférieur au prix américain.

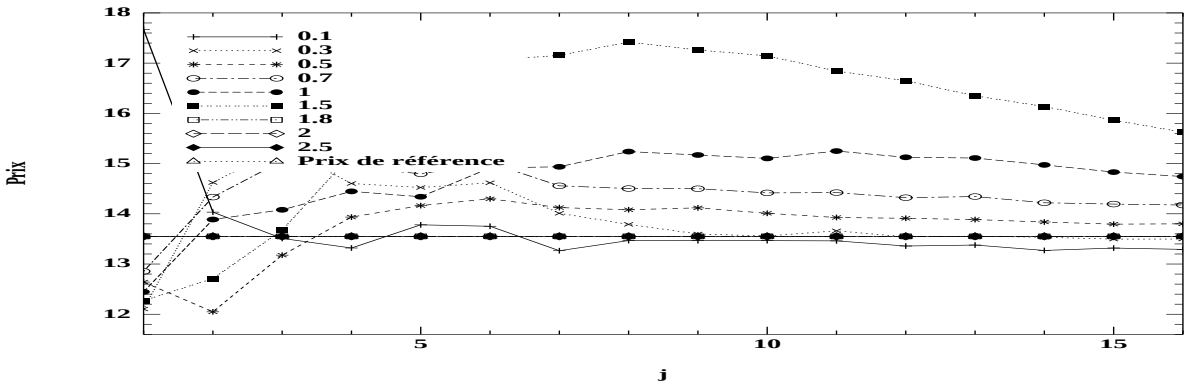
6.2 Tests à coefficients variables

6.2.1 Méthode de pénalisation

Nous avons déjà observé les propriétés de convergence de la méthode de pénalisation à pénalité fixe ($n = 0.5, 1, 5, 10$). Une idée naturelle est d'appliquer l'algorithme avec une pénalité variable, qui varierait en même temps que N, δ et M dans le cas de la base **HC(1,0)**. Pour tester cette variation simultanée, on pose $n = n_0(\sqrt{2})^{(j-1)\alpha_n}$ pour $n_0 = 0.5$ et on teste la convergence de l'algorithme pour différentes valeurs de α_n.



On constate que si α_n est trop important, le prix diverge. C'est clair pour $\alpha_n = 2$ et 2.5. On peut supposer la même chose pour $\alpha_n = 1.8$. Pour les autres valeurs, on a l'impression que le prix converge vers le prix de référence, même si la vitesse de convergence est très différente suivant les valeurs de α_n. Pour comparer ces vitesses, traçons le même graphique en enlevant les valeurs $\alpha_n = 1.8, 2, 2.5$.

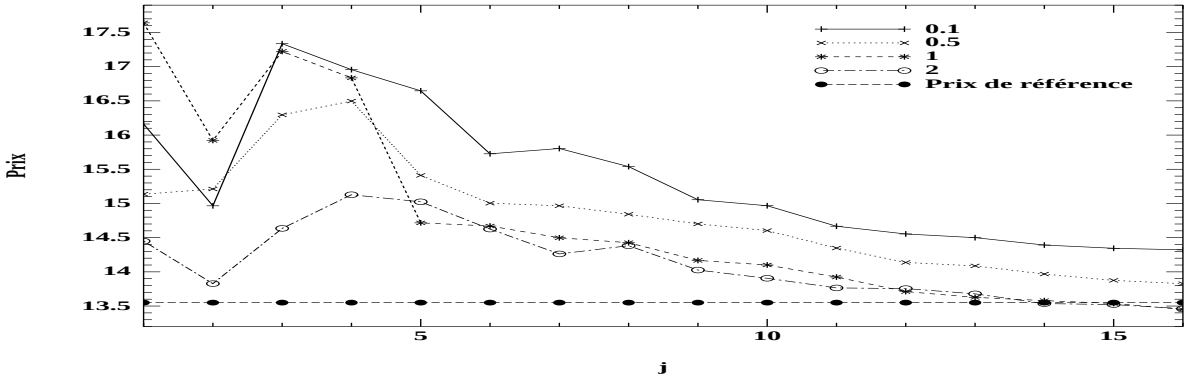


La valeur $\alpha_n = 0.1$ apparait trop faible car on converge rapidement vers une valeur inférieure au prix américain (ce qui est cohérent étant donné que la méthode de pénalisation approche par en-dessous le prix américain) mais cette valeur évolue trop peu rapidement

vers le prix américain. Les valeurs $\alpha_n \geq 0.7$ semblent trop fortes car on se retrouve au-dessus du prix américain ce qui est peu cohérent avec les résultats attendus de la méthode. Restent les valeurs $\alpha_n = 0.3$ ou 0.5 qui donnent les résultats les plus cohérents. Les résultats sur la couverture confirment ces observations. On trouve donc une heuristique de $\mathbf{n \approx h^{-\frac{1}{2}}}$ pour assurer une convergence de la méthode de pénalisation en dimension 1. D'autres tests en dimension plus élevée seraient nécessaires pour voir si cette heuristique dépend de la dimension.

6.2.2 Méthode de régularisation

De la même manière que pour la méthode de pénalisation, on fait varier le coefficient de régularisation n en même temps que N, δ et M dans le cas de la base **HC(1,0)**. On prend toujours $n = n_0(\sqrt{2})^{(j-1)\alpha_n}$ pour $n_0 = 0.5$ et on teste la convergence de l'algorithme pour différentes valeurs de α_n.



On constate que pour α_n petit, le prix donné par l'algorithme se trouve au-dessus du prix américain et pour α_n trop grand, la régularisation ne joue pas bien son rôle et le prix donné par l'algorithme passe en dessous du prix américain. Par contre, la méthode de régularisation apparait plus stable que la méthode de pénalisation. En effet, même lorsque n est très grand, l'algorithme ne diverge pas. C'est normal car ce que l'on approche dans la méthode de régularisation est borné par 1 alors que dans la méthode de pénalisation n est présent dans les constantes apparaissant dans les contrôles d'erreurs. Par contre, on ne tient compte de l'obstacle dans la méthode de régularisation que si Y est dans une bande autour de l'obstacle d'autant plus étroite que n est grand. Le risque, si n est trop grand, est que la méthode ne voit pas qu'il y a un obstacle. Ceci explique pourquoi on observe un prix inférieur au prix américain pour n trop grand. Pour corriger ce problème, il conviendrait de proposer d'autres types de fonctions ϕ_n qui seraient par exemple disymétriques et dont le support serait plus large qu'une bande autour de l'obstacle.

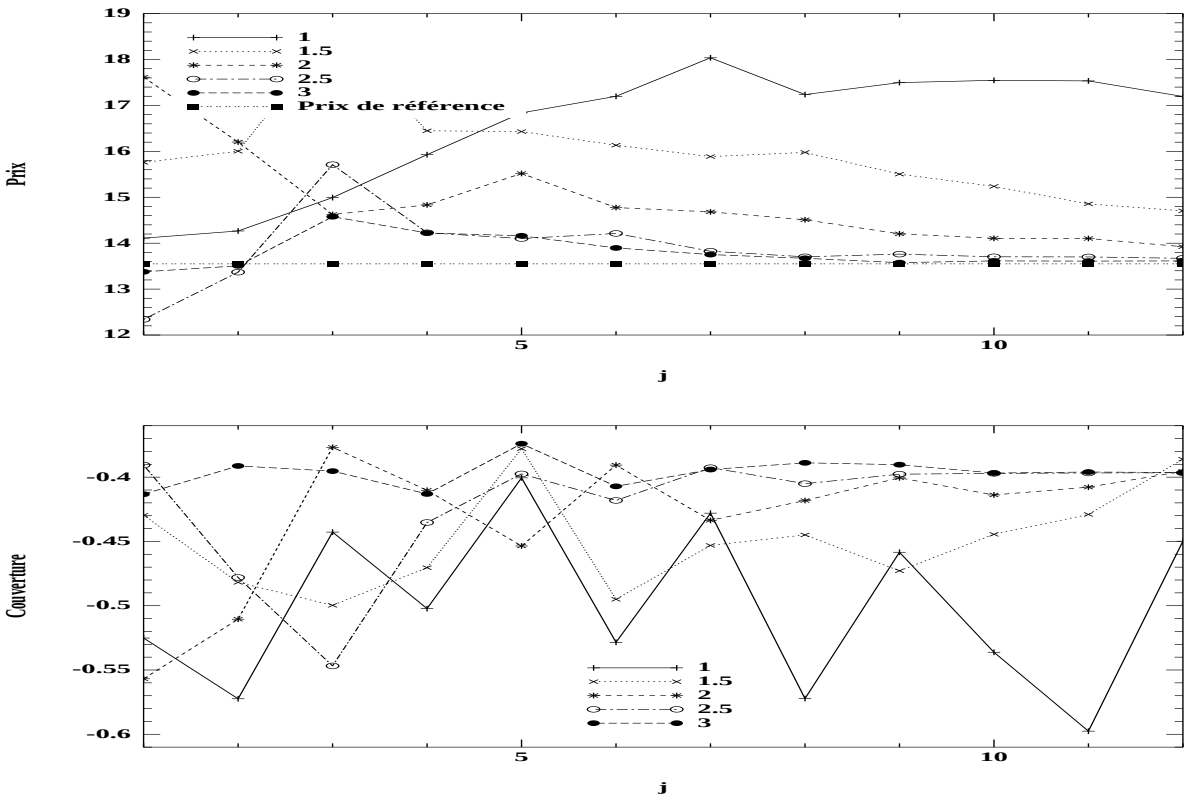
6.3 Liens avec la méthode de Longstaff-Schwartz [LS01]

On peut également utiliser l'algorithme pour déterminer un temps d'arrêt optimal suivant la méthode de Longstaff-Schwartz. Rappelons le principe de cette méthode :

- On initialise avec $y_N^{N,M}(\cdot) = g(t_N, \cdot)$ et pour tout m on pose $\tau^m = t_N$.
- On calcule par une régression linéaire $y_k^{N,M}(\cdot)$ puis pour tout m, si $y_k^{N,M}(X_{t_k}^{N,m}) \leq g(t_k, X_{t_k}^{N,m})$ alors on actualise le temps d'arrêt optimal pour la trajectoire m en posant $\tau^m = t_k$.
- Une fois parvenu à l'instant t_0, on calcule le prix américain comme

$$\frac{1}{M}\sum_{m=1}^{M} \exp(-r\tau^m) g(\tau^m, X_{\tau^m}^{N,m}).$$

Nous pouvons appliquer cet algorithme en déterminant à chaque instant $y_k^{N,M}(\cdot)$ par une régression sur la base **HC**. Voici les résultats :



On constate que l'algorithme de Longstaff-Schwartz donne des résultats qui sont cohérents avec les résultats déjà trouvés par ailleurs :

- Si α_M n'est pas assez grand par rapport à α_δ on observe des problèmes de convergence en particulier pour la couverture (par exemple pour $\alpha_M = 1$ ou 1.5).
- L'algorithme converge bien vers le prix de référence pour α_M assez grand.

– L'algorithme donne une couverture (pour α_M assez grand) du même ordre que celle obtenue par les autres méthodes.

Chapitre 7

Exemples financiers

7.1 Modèle factoriel sur le marché de l'électricité

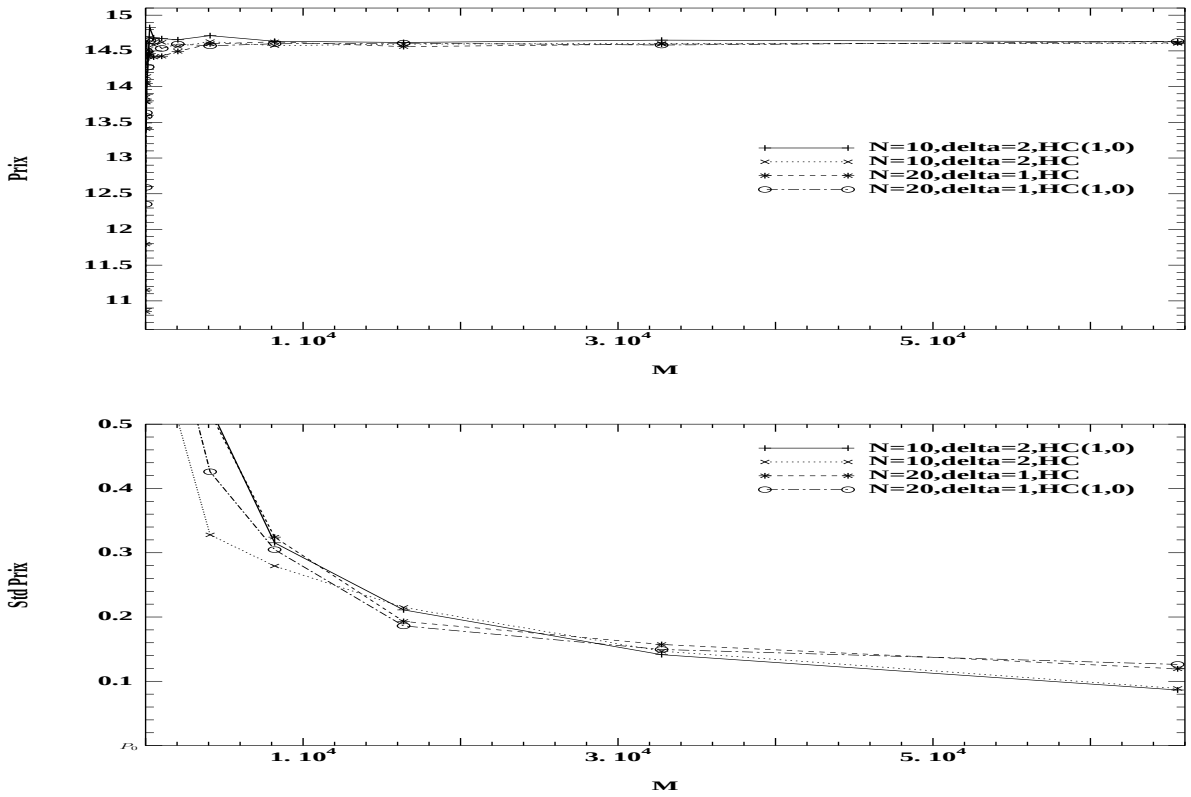
On va traiter le cas du modèle factoriel pour l'évolution des forwards sur électricité pour montrer quels sont les résultats que la méthode donne dans ce cas. En particulier, on désire valoriser une option call européen de maturité $T = 10$ jours. Le sous-jacent est un contrat forward de dynamique :

$$\frac{dF(t,T)}{F(t,T)} = \sigma e^{-a(T-t)} dW_t.$$

Les valeurs des paramètres sont $a = 0.19$, $\sigma = 0.23$ et le taux d'intérêt $r = 0.05$ en valeur annuelle. La valeur du strike est $K = 100$ et la valeur initiale du sous-jacent vaut $F(0,T) = 100$.
Valoriser l'option revient à trouver le couple de processus (Y, Z) satisfaisant l'équation rétrograde :

$$Y_t = (F(T,T) - K)_+ + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s$$

où $f(t,y,z) = -(y - \frac{z}{\sigma}e^{a(T-t)})r$. On applique l'algorithme décrit dans la partie II avec les bases **HC** et **HC(1,0)** et nous reportons les résultats obtenus pour une couverture journalière ($N = 10$) et une taille d'hypercubes $\delta = 2$ puis pour une couverture 2 fois par jour ($N = 20$) et une taille d'hypercubes $\delta = 1$.



Un prix de référence pour cette option, obtenu par EDP, est de **14.72**. On observe sur les courbes que le biais est amélioré lorsqu'on utilise la base $\mathbf{HC(1,0)}$ et que l'écart-type est plus important pour $(N,\delta) = (20,1)$ que pour $(10,2)$ en accord avec les résultats de la partie II.

7.2 Combinaison de calls avec différentiel de taux d'intérêt

On a déjà testé un exemple de différentiel de taux d'intérêt dans la partie précédente et nous allons considérer à nouveau ce cas ici. Cependant, nous changeons la condition terminale $\Phi(\mathbf{S}) = (S_T - K_1)^+ - 2(S_T - K_2)^+$. Le choix de paramètres est

μ	σ	r	R	T	S_0	K_1	K_2
0.05	0.2	0.01	0.06	0.25	100	95	105

On désigne par $BS_i(r)$ le prix Black-Scholes évalué avec le strike K_i et le taux d'intérêt r. Si on tente de prédire Y_0 par une combinaison linéaire de prix Black-Scholes, on obtient

$BS_1(R) - 2BS_2(R)$	2.75
$BS_1(r) - 2BS_2(r)$	2.76
$BS_1(r) - 2BS_2(R)$	1.92
$BS_1(R) - 2BS_2(r)$	3.60

En utilisant un théorème de comparaison ([EKPQ97]), on peut vérifier que les trois premières lignes sont des bornes inférieures pour Y_0 alors que la dernière ligne fournit une borne supérieure.
On reporte ci-dessous la moyenne des prix obtenus par l'algorithme sur 50 lacements et entre parenthèses l'écart-type empirique, pour différents types de fonctions de base et différentes valeurs de M.

M	$N = 5$ $D = [60, 140]$ $\delta = 5$	$N = 20$ $D = [60, 200]$ $\delta = 1$	$N = 50$ $D = [40, 200]$ $\delta = 0.5$
128	3.05(0.27)	3.71(0.95)	3.69(4.15)
512	2.93(0.11)	3.14(0.16)	3.48(0.54)
2048	2.92(0.05)	3.00(0.03)	3.08(0.12)
8192	2.91(0.03)	2.96(0.02)	2.99(0.02)
32768	2.90(0.01)	2.95(0.01)	2.96(0.01)

TAB. 7.1 – Résultats obtenus avec la base **HC**.

M	base **VP** 16 régions de Voronoi $N = 5$	base **VP** 64 reg. Vor. $N = 20$	base **VP** 10 reg. Vor. $N = 20$	base **VP(1,0)** 10 reg. Vor. $N = 20$
128	3.23(0.30)	4.50(1.71)	3.08(0.25)	3.23(0.23)
512	3.05(0.13)	3.36(0.10)	2.91(0.11)	3.03(0.08)
2048	2.94(0.06)	3.05(0.04)	2.90(0.06)	2.97(0.04)
8192	2.92(0.03)	2.96(0.02)	2.86(0.03)	2.95(0.02)
32768	2.90(0.02)	2.94(0.01)	2.86(0.02)	2.95(0.01)

TAB. 7.2 – Résultats obtenus avec les bases **VP** et **VP(1,0)**.

Vu les résultats de **HC** et **VP**, $\overline{Y}_{t_0}^{N,M}$ semble converger vers **2.95**. Cette valeur ne peut pas être prédite par une combinaison linéaire de prix Black-Scholes : dans cet exemple, la non-linéarité de f a un réel impact sur Y_0. L'interprétation financière est que le vendeur de l'option doit alternativement emprunter et prêter de l'argent pour répliquer le payoff de l'option.

M	$N=5$ $d_y=1$, $d_z=0$	$N=20$ $d_y=2$, $d_z=1$	$N=50$ $d_y=4$, $d_z=2$	$N=50$ $d_y=9$, $d_z=9$
128	2.87(0.39)	3.01(0.24)	3.02(0.22)	3.49(1.57)
512	2.82(0.20)	2.94(0.12)	2.97(0.09)	3.02(0.1)
2048	2.78(0.07)	2.92(0.07)	2.92(0.04)	2.97(0.03)
8192	2.78(0.05)	2.92(0.04)	2.92(0.02)	2.96(0.01)
32768	2.79(0.03)	2.91(0.02)	2.91(0.01)	2.95(0.01)

TAB. 7.3 – Résultats obtenus avec la base **GP**.

En comparant les différents choix de fonctions de base, on peut remarquer que la colonne $N=5$ de **VP** (Tableau 7.2) présente des résultats similaires avec un nombre équivalent de fonctions de base que la colonne $N=5$ de **HC** (Tableau 7.1). Dans le Tableau 7.2, les deux dernières colonnes montrent qu'utiliser une base de polynômes locaux peut améliorer significativement la précision. Nous avions déjà observé cela dans la sous-section 4.2 pour les bases **HC** et **HC(1,0)**. Finalement, vu le Tableau 7.3, la base de polynômes globaux parvient également à atteindre la valeur attendue, lorsque nous augmentons le nombre de polynômes dans la base. Ajoutons que pour la dernière colonne du Tableau 7.3, les bases à l'instant t_k ne sont pas (avec $d=q=1$) $p_{0,k}=1,x,\cdots,x^{d_y}$ (resp. $p_{1,k}=1,\cdots,x^{d_z}$) mais $1,\frac{x-\min_m S_{t_k}^{N,m}}{\max_m S_{t_k}^{N,m}-\min_m S_{t_k}^{N,m}},\cdots,(\frac{x-\min_m S_{t_k}^{N,m}}{\max_m S_{t_k}^{N,m}-\min_m S_{t_k}^{N,m}})^{d_y}$ (et la même chose pour $p_{1,k}$). Cette normalisation est nécessaire lorsque le nombre de polynômes devient grand, pour éviter d'avoir des matrices de régression mal conditionnées.

7.3 Options asiatiques

On considère toujours $d=q=1$ et le modèle de Black-Scholes pour S. La condition terminale de l'EDSR vaut à présent $\Phi(\mathbf{S})=(\frac{1}{T}\int_0^T S_t dt-K)_+$, soit le cas d'une option asiatique, et nous prenons les paramètres suivants :

μ	σ	r	T	S_0	K
0.06	0.2	0.1	1	100	100

Pour approcher cette condition terminale path-dependent, comme suggéré dans la partie I, nous prenons $d'=2$ et simulons $P_{t_k}^N=(S_{t_k}^N,\frac{1}{k+1}\sum_{i=0}^k S_{t_i}^N)^*$ (voir [LT01]). Les résultats présentés dans le tableau 7.4 sont cohérents car le prix donné par l'algorithme n'est pas très éloigné du prix de référence **7.04** donné dans [LT01].
Comme mentionné dans [LT01], l'utilisation de $\frac{1}{N+1}\sum_{i=0}^N S_{t_i}^N$ pour approcher $\frac{1}{T}\int_0^T S_t dt$ est loin d'être optimal. On peut vérifier ce qui se passe si on change P^N pour mieux approcher $\frac{1}{T}\int_0^T S_t dt$. Comme proposé dans [LT01], nous approchons $\frac{1}{T}\int_0^T S_t dt$ par

M	$N=5$ $\delta=5$ $D=[60,200]^2$	$N=20$ $\delta=1$ $D=[60,200]^2$	$N=50$ $\delta=0.5$ $D=[60,200]^2$
128	6.33(0.41)	4.47(3.87)	3.48(13.08)
512	6.65(0.21)	6.28(0.76)	5.63(2.37)
2048	6.80(0.09)	6.76(0.24)	6.48(0.49)
8192	6.83(0.04)	6.95(0.06)	6.86(0.12)
32768	6.83(0.02)	6.98(0.03)	6.99(0.04)

TAB. 7.4 – Résultats pour l'option asiatique en utilisant la base **HC**.

$\frac{1}{N}\sum_{i=0}^{N-1} S_{t_i}^N(1+\mu\frac{h}{2}+\frac{\sigma}{2}\Delta W_{t_i})$, ce qui conduit à $P_{t_k}^N = \left(S_{t_k}^N, \frac{1}{k}\sum_{i=0}^{k-1} S_{t_i}^N(1+\mu\frac{h}{2}+\frac{\sigma}{2}\Delta W_{t_i})\right)^*$ pour $k \geq 1$. Les résultats (voir tableau 7.5) sont bien meilleurs avec ce choix pour P^N. De plus, nous observons la cohérence de l'algorithme qui prend en input des simulations de S^N sous la probabilité historique ($\mu \neq r$) et corrige le drift pour donner le prix risque-neutre.

M	2	8	32	128	512	2048	8192	32768
Prix	2.26	0.90	4.49	6.68	6.15	6.88	6.99	7.02
Ecart-type	4.08	7.80	11.27	4.64	1.11	0.21	0.07	0.02

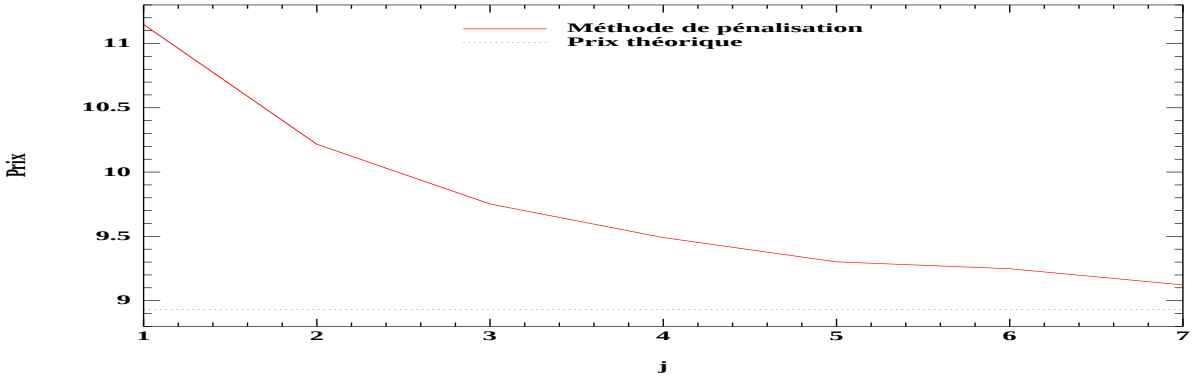
TAB. 7.5 – Résultats pour l'option asiatique en utilisant une meilleure approximation de $\frac{1}{T}\int_0^T S_t dt$ et la base **HC** ($N=20$, $\delta=1$, $D=[60,200]^2$).

7.4 Options américaines et influence de la dimension

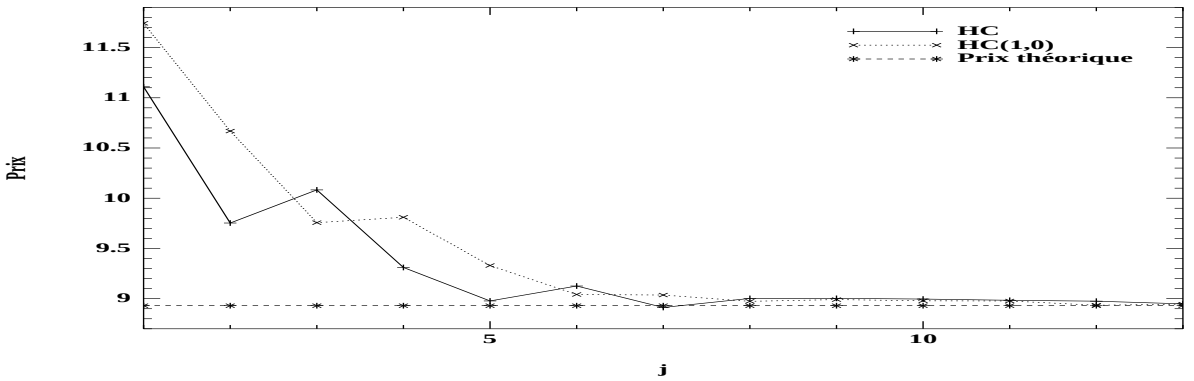
La plupart des exemples que nous avons traités jusqu'à présent sont effectués en dimension 1. Nous allons illustrer sur un exemple les problèmes de la dimension. Comme benchmark, nous choisissons à nouveau un exemple tiré de [BT04] mais cette fois-ci en dimension 3. S suit la dynamique Black-Scholes en dimension 3 et le payoff vaut $(K-(\prod_{i=1}^3 S_t^i)^{\frac{1}{3}})^+$. Les valeurs des paramètres sont $\mu=0.05$, $r=0.05$, $T=1$, les composantes de S_0 valent 100 et $K=100$. La matrice σ vaut $\sigma = \begin{pmatrix} 0.4 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.4 \end{pmatrix}$. Le prix de référence de cette option donné dans [BT04] est **8.93**. Nous désirons établir des bornes inférieures pour ce prix américain et utilisons donc l'algorithme avec pénalisation que nous choisissons égale à **0.5**.

Nous présentons tout d'abord des résultats avec la base **HC** en choisissant $\alpha_\delta = 1$ et $\alpha_M = 4.5$. En effet la proposition IV.1 impose à M de varier comme $h^{-(1+\beta+\alpha_\delta d')}$ avec

$\beta > 1$. Ici cela donne que M doit tendre vers l'infini plus vite que h^{-5}. Nous avons également constaté que ce seuil était trop pessimiste par rapport aux résultats numériques observés. C'est la raison pour laquelle nous prenons $\alpha_M = 4.5$. Nous obtenons les résultats suivants :



Nous remarquons que la convergence est très lente : le prix donné par l'algorithme est encore loin du prix théorique pour $j = 7$ et l'on devrait surtout observer des prix inférieurs au prix théorique étant donné que l'on utilise une méthode de pénalisation. Ainsi, le prix donné par l'algorithme est encore loin d'avoir convergé. C'est l'effet de la dimension car si on voit le problème comme un problème en dimension 1 en prenant comme nouvelle variable d'état $(\prod_{i=1}^{3} S^i)^{\frac{1}{3}}$ on obtient les résultats suivants en utilisant d'une part les bases **HC** et **HC(1,0)** en dimension 1 avec cette nouvelle variable d'état :



La convergence est bien meilleure même si le prix n'a pas encore tout à fait convergé pour les dernières valeurs de j testées (on observe encore des prix légèrement supérieurs au prix théorique) .

Ainsi, le problème de l'augmentation de la dimension est de trouver de bonnes variables d'état markoviennes ou de bonnes fonctions de base pour éviter d'avoir à utiliser des bases de fonctions rudimentaires telles **HC** qui certes convergent mais à un coût de calcul élevé.

Conclusion

Nous avons étudié dans cette thèse une méthode de Monte-Carlo permettant d'approcher la solution d'une équation différentielle stochastique rétrograde markovienne. Dans la partie I, nous avons analysé les erreurs dûes au pas de temps h, à l'utilisation d'un nombre fini de fonctions de base K et à un nombre fini M de simulations de Monte-Carlo. Toutefois, les bornes pour l'erreur dûe à M n'étant pas assez explicites pour une utilisation en pratique, nous avons amélioré ces estimations pour un algorithme légèrement modifié dans la partie II. Ces nouvelles bornes nous ont alors permis de faire varier simultanément h, le nombre de fonctions de base et M simultanément. Nous avons vérifié numériquement la validité de ces estimations dans la partie IV. Cependant, cette dernière partie soulève de nombreuses questions :

- Nous avons établi des résultats de bornes pour la norme $\mathbf{L}^2$ de l'erreur ce qui nous a conduit à lancer, pour chaque jeu de paramètres (N, M, K), 50 fois l'algorithme de manière à vérifier la validité de ces bornes. En pratique, on ne fait pas tourner l'algorithme 50 fois pour obtenir un prix et une couverture. Ainsi, il conviendrait d'établir également des résultats de convergence p.s. pour l'algorithme, qui pourrait éventuellement provenir des bornes pour la norme $\mathbf{L}^2$ lorsque ces bornes tendent suffisamment vite vers 0.
- Nous avons testé certains choix de bases de fonctions qui nous paraissaient cohérents mais de nombreux autres sont à tester. En particulier, un choix naturel serait de définir le support d'une base d'hypercubes, de manière à ce que chaque hypercube de la base contienne le même nombre de simulations à un instant donné. Il faut alors étudier l'erreur d'approximation d'une fonction Lipschitz par une telle base.
- L'algorithme nécessite que la condition terminale de l'EDSR soit une fonction déterministe de la valeur finale d'une variable d'état markovienne. Cependant, cette condition est parfois coûteuse à cause de la dimension de la variable d'état markovienne et il serait intéressant de tester des alternatives comme de prendre pour les régressions des fonctions de base qui sont fonctions d'une autre variable d'état, pas forcément markovienne, et de dimension plus petite.
- Nous avons testé plusieurs méthodes d'approximation pour les options américaines. Pour la méthode de pénalisation, une heuristique apparait pour la manière de faire varier simultanément les paramètres (N, M, K) et le coefficient de pénalisation n : il conviendrait de montrer théoriquement cette règle. Pour la méthode de régularisation,

la nécessité de fonctions de régularisation ϕ_n adaptées au problème discrétisé se fait sentir.

- Toujours pour les options américaines, nous avons fait le lien entre notre algorithme et la méthode de Longstaff-Schwarz mais il serait également intéressant de voir ce que donne l'utilisation de $\sum z_k^{N,R,M}(P_{t_k}^N)\Delta W_k$ comme martingale candidate à la méthode duale décrite par Rogers dans [Rog02].
- Nous avons limité nos tests numériques au cas d'un aléa brownien et une généralisation à d'autres aléas serait intéressante.
- Il parait intéressant de forcer l'approximation à être une fonction Lipschitz pour gagner un facteur h (voir la discussion page 166) et donc de modifier l'algorithme pour tenir compte de cet aspect .
- Il faudrait étudier si l'algorithme se généralise au cas des EDS progressives rétrogrades.
- Il apparait également tentant d'utiliser directement des historiques d'évolution de P^N à la place des simulations de Monte-Carlo (évitant ainsi l'estimation d'un modèle). C'est l'idée exposée dans [PBS01] qu'il conviendrait de vérifier théoriquement.

Annexes

A.1 Equations différentielles stochastiques rétrogrades

A.1.1 Existence et unicité

Soit $T > 0$ l'horizon du problème. On considère un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ qui supporte un mouvement brownien W de dimension q. On rappelle quelques notations classiques des EDSR (voir [EKPQ97]) :

- $(\mathcal{F}_t)_{0\leq t\leq T}$ désigne la filtration engendrée par W et augmentée. On note $\mathcal{P}$ la tribu des sous-ensembles prévisibles de $\Omega \times [0,T]$.
- $\mathbf{L}^2_t(\mathbb{R}^q)$ désigne l'ensemble des variables aléatoires à valeurs $\mathbb{R}^q$, $\mathcal{F}_t$ mesurables et de carré intégrable. On note simplement $\mathbf{L}^2_t$ pour $\mathbf{L}^2_t(\mathbb{R})$.
- $\mathbb{H}^2_T(\mathbb{R}^q)$ désigne l'ensemble des processus prévisibles $\phi : \Omega \times [0,T] \to \mathbb{R}^q$ (considérés comme des vecteurs lignes) tels que $\mathbb{E}\int_0^T |\phi_t|^2 dt$ soit fini. On note simplement $\mathbb{H}^2_T$ pour $\mathbb{H}^2_T(\mathbb{R})$.
- Pour $\beta > 0$ et $\phi \in \mathbb{H}^2_T(\mathbb{R}^q)$, $\|\phi\|^2_\beta$ désigne $\mathbb{E}\int_0^T \exp(\beta t)|\phi_t|^2 dt$. $\mathbb{H}^2_{T,\beta}(\mathbb{R}^q)$ désigne l'espace $\mathbb{H}^2_T(\mathbb{R}^q)$ muni de la norme $\|\cdot\|_\beta$.

Résoudre une équation différentielle stochastique rétrograde consiste à trouver un couple de processus (Y,Z) adapté à $(\mathcal{F}_t)_{0\leq t\leq T}$ qui vérifie :

$$\begin{cases} Y_T = \xi, \\ -dY_t = f(t, Y_t, Z_t)dt - Z_t dW_t \end{cases} \tag{A.1}$$

où ξ est une variable aléatoire $\mathcal{F}_T$-mesurable et $f : \Omega \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^q \to \mathbb{R}$ est $\mathcal{P} \bigotimes \mathcal{B} \bigotimes \mathcal{B}^q$-mesurable. f est appelé le driver de l'EDSR. Lorsque ξ appartient à $\mathbf{L}^2_T$, que $f(\cdot, 0, 0)$ appartient à $\mathbb{H}^2_T$ et que f est uniformément Lipschitz au sens où :

$$\exists C > 0, d\mathbb{P} \times dt\, p.s. |f(\omega, t, y_1, z_1) - f(\omega, t, y_2, z_2)| \leq C(|y_1 - y_2| + |z_1 - z_2|)\, \forall (y_1, z_1, y_2, z_2),$$

on dit que (f, ξ) est un couple de paramètres standard.
On a alors (voir [PP90] ou [EKPQ97]) :

Théorème A.1 *Soit (f, ξ) un couple de paramètres standard, il existe un unique couple de processus $(Y,Z) \in \mathbb{H}^2_T \times \mathbb{H}^2_T(\mathbb{R}^q)$ qui résout* (A.1).

La proposition suivante ([EKPQ97]) permet d'exprimer explicitement Y lorsque le driver f est linéaire en (y, z).

Proposition A.1 *Soit (δ, γ) un processus prévisible, borné, à valeurs $(\mathbb{R}, \mathbb{R}^q)$, ϕ un élément de $\mathbb{H}^2_T(\mathbb{R})$ et $\xi \in \mathbf{L}^2_T$. Alors l'EDSR*

$$\begin{cases} Y_T = \xi, \\ -dY_t = [\phi_t + Y_t\delta_t + Z_t\gamma_t]dt - Z_t dW_t \end{cases}$$

a une unique solution (Y, Z) dans $\mathbb{H}^2_{T,\beta}(\mathbb{R}) \times \mathbb{H}^2_{T,\beta}(\mathbb{R}^q)$ et Y est donné par la formule

$$Y_t = \mathbb{E}\big(\xi\Gamma^t_T + \int_t^T \Gamma^t_s\phi_s ds|\mathcal{F}_t\big)$$

où Γ^t_s est le processus adjoint défini pour $s \geq t$ par l'EDS forward

$$\begin{cases} d\Gamma^t_s = \Gamma^t_s[\delta_s ds + \gamma^*_s dW_s], \\ \Gamma^t_t = 1. \end{cases}$$

En particulier, si ξ et ϕ sont positifs, le processus Y est positif. Si de plus, $Y_0 = 0$, alors pour tout t, $Y_t = 0$ p.s., $\xi = 0$ p.s. et $\phi_t = 0$ $d\mathbb{P} \otimes dt$ p.s.

A.1.2 Propriétés fines des EDSR

Supposons que l'EDSR (A.1) prenne la forme particulière suivante :

$$\begin{aligned} X_t =& X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \\ Y_t =& \Phi(\mathbf{X}) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s \end{aligned} \tag{A.2}$$

où b, σ, f, Φ vérifient l'hypothèse :

Hypothèse A.1 *Les fonctions b, σ et f sont Höldériennes d'indice $\frac{1}{2}$ en temps et uniformément Lipschitz dans les variables d'espace. On suppose que $\Phi : \mathbb{D}^d \to \mathbb{R}$ est une fonctionnelle $\mathbf{L}^\infty$-Lipschitz au sens où, $\forall x_1, x_2 \in \mathbb{D}^d$*

$$|\Phi(x_1) - \Phi(x_2)| \leq C \sup_{0\leq t\leq T} |x_1(t) - x_2(t)|$$

avec $\mathbb{D}$ l'espace des fonctions continues à droite, avec des limites à gauche (càdlàg), définies sur $[0, T]$. On note par $K > 0$ toutes les constantes de Lipschitz et on suppose que

$$\sup_{0\leq t\leq T} \{|b(t, 0)| + |\sigma(t, 0)| + |f(t, 0, 0, 0)|\} + |\Phi(\mathbf{0})| \leq K$$

où $\mathbf{0}$ est la fonction constante valant 0 sur $[0, T]$.

Remarque A.1 *([Zha04]) A cause de la remarque 3.2.3 dans [Zha01], Z possède une version càdlàg si, en plus de l'hypothèse A.1, σ est uniformément Lipschitz par rapport à t.*

Théorème A.2 *([Zha04]) On suppose que l'hypothèse A.1 est vérifiée et que Z est càdlàg. Soit $\pi = (t_i)_{0\leq i\leq n}$ une discrétisation de $[0,T]$, alors*

$$\sum_{i=1}^{n} \mathbb{E}\int_{t_{i-1}}^{t_i} \{|Z_t - Z_{t_{i-1}}|^2 + |Z_t - Z_{t_i}|^2\}dt \leq C(1+|X_0|^2)|\pi|$$

où $C>0$ est une constante dépendant de T et K mais indépendante de la discrétisation π.

A.2 Equations différentielles stochastiques rétrogrades réfléchies

En nous appuyant sur [EKKP$^+$97] (dont on reprend les notations), une équation différentielle stochastique rétrograde réfléchie (EDSRR) est définie par la donnée de

(i) une condition terminale $\xi \in \mathbf{L}^2$,

(ii) une fonction $f : \Omega\times[0,T]\times\mathbb{R}\times\mathbb{R}^q \to \mathbb{R}$, telle que $\forall (y,z)\in\mathbb{R}\times\mathbb{R}^q$, $f(\cdot,y,z)\in\mathbb{H}_T^2$ et telle qu'il existe $C>0$ satisfaisant $|f(t,y,z)-f(t,y',z')| \leq C(|y-y'|+|z-z'|) \quad \forall y,y'\in\mathbb{R}, z,z'\in\mathbb{R}^q, p.s.$,

(iii) un obstacle $(S_t)_{0\leq t\leq T}$, qui est un processus à valeurs réelles, continu et progressivement mesurable satisfaisant $\mathbb{E}(\sup_{0\leq t\leq T}(S_t^+)^2) < \infty$. On suppose que $S_T \leq \xi \quad p.s.$

La solution de l'EDSRR définie par (ξ, f, S) est un triplet $(Y_t, Z_t, K_t)_{0\leq t\leq T}$ de processus $\mathcal{F}_t$ progressivement mesurables à valeurs $\mathbb{R}$, $\mathbb{R}^q$ et $\mathbb{R}_+$ et satisfaisant

(iv) $Z\in\mathbb{H}_T^2(\mathbb{R}^q)$;

(v) Y est un processus prévisible tel que $\mathbb{E}(\sup_{0\leq t\leq T}|Y_t|^2) < \infty$;

(vi) $Y_t = \xi + \int_t^T f(s,Y_s,Z_s)ds + K_T - K_t - \int_t^T Z_s dW_s$, $0\leq t\leq T$;

(vii) $Y_t \geq S_t$, $0\leq t\leq T$;

(viii) $(K_t)_{0\leq t\leq T}$ est un processus continu, croissant, nul en 0 et tel que $\int_0^T (Y_t - S_t)dK_t = 0$.

Le théorème 5.2 dans [EKKP$^+$97] permet alors d'affirmer que

Théorème A.3 *L'EDSRR définie par (i)-(ii)-(iii) admet une unique solution (Y,Z,K) satisfaisant (iv)-(v)-(vi)-(vii)-(viii).*

Nous allons à présent considérer des EDSRR markoviennes. En particulier, on considère la dynamique (A.2) pour X (on note par X^N son schéma d'Euler) et les deux EDSRR :

$$Y_t = \phi(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s + K_T - K_t,$$
$$Y_t \geq g(t, X_t),\ t \geq 0,\ \int_0^T (Y_t - g(t, X_t))dK_t = 0,$$

avec $\phi(\cdot) \geq g(T, \cdot)$ et

$$Y_t^{N,0} = \phi(X_T^N) + \int_t^T f(s, X_s^N, Y_s^{N,0}, Z_s^{N,0})ds - \int_t^T Z_s^{N,0} dW_s + K_T^{N,0} - K_t^{N,0},$$
$$Y_t^{N,0} \geq g(t, X_t^N),\ t \geq 0,\ \int_0^T (Y_t^{N,0} - g(t, X_t^N))dK_t^{N,0} = 0. \tag{A.3}$$

Les hypothèses de [MZ05] sont les suivantes (X et W sont supposés de dimension d, qui est ensuite fixée à 1 dans [MZ05] avec une extension facile au cas $d > 1$) :

Hypothèse A.2 *b et σ sont continues, différentiables et avec des dérivées uniformément bornées par rapport à la variable x, et $\sigma\sigma^* \geq \epsilon_0 Id$ où $\epsilon_0 > 0$.*

Hypothèse A.3 *f et ϕ sont continues et uniformément Lipschitz par rapport aux variables (x, y, z).*

Hypothèse A.4 *$g \in C^{1,2}([0, T] \times \mathbb{R}^d)$ avec toutes ses dérivées uniformément bornées.*

Remarque A.2 *Dans [MZ05], l'obstacle est noté $h(\cdot, \cdot)$. Nous avons préféré modifié les notations car dans notre cas, il pourrait y avoir une confusion avec le pas de temps $h = \frac{T}{N}$.*

Citons à présent le lemme 2.2 de [MZ05], que nous écrivons dans un cadre multidimensionnel :

Lemme A.1 *On suppose que les hypothèses A.2, A.3 et A.4 sont vraies.*

1. *Il existe une constante universelle $C > 0$ telle que*

$$\mathbb{E}\{\sup_{0 \leq t \leq T} |Y_t|^2 + \int_0^T |Z_t|^2 dt + K_T^2\} \leq C(1 + |X_0|^2)$$
$$\mathbb{E}\{\sup_{0 \leq t \leq T} |Y_t^{N,0}|^2 + \int_0^T |Z_t^{N,0}|^2 dt + |K_T^{N,0}|^2\} \leq C(1 + |X_0|^2).$$

2.

$$\mathbb{E}\left\{\sup_{0 \leq t \leq T} [|Y_t^{N,0} - Y_t|^2 + |K_t^{N,0} - K_t|^2] + \int_0^T |Z_t^{N,0} - Z_t|^2 dt\right\} \leq C(1 + |X_0|^2)\sqrt{h}.$$

3. On note pour $t \in [t_k, t_{k+1}[$

$$\begin{aligned} k_t =& |\partial_t g(t, X_t)| + |\sum_{i=1}^{d} \partial_{x_i} g(t, X_t) b_i(t, X_t)| + \frac{1}{2} |\sum_{i,j=1}^{d} (\sigma\sigma^*)_{i,j}(t, X_t) \partial^2_{x_i,x_j} g(t, X_t)| \\ &+ |f(t, X_t, Y_t, Z_t)|, \\ k_t^{N,0} =& |\partial_t g(t, X_t^N)| + |\sum_{i=1}^{d} \partial_{x_i} g(t, X_t^N) b_i(t_k, X_{t_k}^N)| + \frac{1}{2} |\sum_{i,j=1}^{d} (\sigma\sigma^*)_{i,j}(t_k, X_{t_k}^N) \partial^2_{x_i,x_j} g(t, X_t^N)| \\ &+ |f(t, X_t^N, Y_t^{N,0}, Z_t^{N,0})|. \end{aligned}$$

Alors, au sens des mesures aléatoires, on a

$$0 \leq dK_t \leq k_t dt, 0 \leq dK_t^{N,0} \leq k_t^{N,0} dt, \forall t \in [0, T], p.s.$$

Ce lemme s'appuie sur des résultats de [EKKP+97] et en particulier

Proposition A.2 *On suppose que l'obstacle $g(t, X_t)$ est une semimartingale de la forme*

$$g(t, X_t) = g(0, X_0) + \int_0^t U_s ds + \int_0^t V_s dW_s$$

où U et V sont des processus $\mathcal{F}_t$-progressivement mesurables satisfaisant

$$\int_0^T (|U_t| + |V_t|^2) dt < \infty \ p.s.$$

Alors

$$Z_t = V_t, d\mathbb{P} \times dt \, p.p. \, sur \, l'ensemble \, \{Y_t = g(t, X_t)\} \tag{A.4}$$

et

$$0 \leq dK_t \leq \mathbf{1}_{Y_t = g(t, X_t)} [f(t, X_t, g(t, X_t), V_t) + U_t]^- dt.$$

On utilisera également une remarque liée à ce résultat :

Remarque A.3 *La propriété ci-dessus peut être généralisée facilement à un obstacle $g(t, X_t)$ qui est une semimartingale plus générale*

$$g(t, X_t) = g(0, X_0) + \int_0^t U_s ds + \int_0^t V_s dW_s + A_t$$

où A est un processus continu à variation intégrable tel que la mesure dA_t est singulière par rapport à dt et qui admet une décomposition $A_t = A_t^+ - A_t^-$, où A_t^+ et A_t^- sont des processus intégrables. Alors, $(U_t)_t$ et $(V_t)_t$ sont des processus $\mathcal{F}_t$-progressivement mesurables respectivement à valeurs dans $\mathbb{R}$ et $\mathbb{R}^d$ satisfaisant :

$$\int_0^T |U_t| + |V_t|^2 dt + A_T^+ + A_T^- < \infty \, p.s.$$

L'égalité (A.4) *est toujours satisfaite et le reste est modifié par (avec L_t le temps local en 0 de la semi-martingale $Y_t - g(t, X_t)$) :*

$$dK_t + \frac{1}{2} dL_t = \mathbf{1}_{Y_t = g(t,X_t)} [f(t, X_t, g(t, X_t), V_t) + U_t]^- dt + dA_t^-.$$

Il en découle l'existence d'un processus prévisible $(\alpha_t, 0 \leq t \leq T)$ tel que $0 \leq \alpha_t \leq 1$ et

$$dK_t = \alpha_t \mathbf{1}_{Y_t = g(t,X_t)} ([f(t, X_t, g(t, X_t), V_t) + U_t]^- dt + dA_t^-).$$

Les auteurs de [MZ05] définissent alors une approximation de (A.3) par
- On initialise en t_N par $Y_{t_N}^{N,1} = \phi(X_{t_N}^N)$.
- Puis pour $t \in [t_k, t_{k+1}[$, on résout l'équation rétrograde non réfléchie de solution $(\tilde{Y}^{N,1}, Z^{N,1})$:

$$\tilde{Y}_t^{N,1} = Y_{t_{k+1}}^{N,1} + \int_t^{t_{k+1}} f(s, X_s^N, \tilde{Y}_s^{N,1}, Z_s^{N,1}) ds - \int_t^{t_{k+1}} Z_s^{N,1} dW_s.$$

- On définit ensuite pour tout t dans $[t_k, t_{k+1}[$, $Y_t^{N,1} = \tilde{Y}_t^{N,1} \vee g(t, X_t^N)$.
- On itère jusqu'à la date $t_0 = 0$.

Il est facile de voir que $Y_{t_k}^{N,1} = u_k^{N,1}(X_{t_k}^N)$ pour une certaine fonction mesurable $u_k^{N,1}$. Le lemme 3.2 dans [MZ05] donne des contrôles sur $(u_k^{N,1})_k$, $Y^{N,1}$ et $Z^{N,1}$.

Lemme A.2 *On suppose les hypothèses A.2, A.3 et A.4 sont vérifiées.*
- *Les fonctions $u_k^{N,1}$ sont Lipschitz uniformément en N et k.*
- *Il existe une constante C, indépendante de N, telle que $\forall k$ and $\forall t \in [t_k, t_{k+1}[$*

$$\mathbb{E}\Big\{ \sup_{t_k \leq t \leq t_{k+1}} \big[|\tilde{Y}_t^{N,1}|^2 + |Z_t^{N,1}|^2 \big] \Big\} \leq C(1 + |X_0|^2)$$
$$\mathbb{E}\big\{ |\tilde{Y}_t^{N,1} - Y_t^{N,1}|^2 \big\} \leq C(1 + |X_0|^2) |t_{k+1} - t|.$$

Le théorème 3.3 dans [MZ05] borne l'erreur entre $(Y^{N,0}, Z^{N,0}, K^{N,0})$ et $(Y^{N,1}, Z^{N,1}, K^{N,1})$:

Théorème A.4 *On suppose que les hypothèses A.2, A.3 et A.4 sont vraies. Alors :*

$$\mathbb{E}\Big\{ \sup_{0 \leq t \leq T} \big[|Y_t^{N,1} - Y_t^{N,0}|^2 + |K_t^{N,1} - K_t^{N,0}|^2 \big] + \int_0^T |Z_t^{N,1} - Z_t^{N,0}|^2 dt \Big\} \leq C(1 + |X_0|^4) h.$$

Enfin, le théorème 6.3 dans [MZ05] permet d'évaluer la régularité $\mathbf{L}^2$ de Z. Ce théorème suppose que

Hypothèse A.5 *b, σ et f sont uniformément Höldériennes d'indice $\frac{1}{2}$ en temps.*

Théorème A.5 *On suppose que les hypothèses A.2, A.3, A.4 et A.5 sont vraies. Alors il existe une constante $C > 0$ telle que*

$$\sum_{k=0}^{N-1} \sup_{s\in[t_k,t_{k+1}]} \mathbb{E}\int_{t_k}^{t_{k+1}} |Z_t - Z_s|^2 dt \leq C(1+|X_0|^4)\sqrt{h}.$$

A.3 Inégalités classiques

Un certain nombre d'inégalités sont constamment utilisées.
Inégalité de Young

$$\forall\ \gamma > 0, \forall\ (a,b) \in \mathbb{R}^2, \quad (a+b)^2 \leq (1+\gamma h)a^2 + (1+\frac{1}{\gamma h})b^2. \tag{A.5}$$

Le lemme de Gronwall discret

Lemme A.3 *Soient $T > 0$ et $t_k = kh$ pour $0 \leq k \leq N = \frac{T}{h}$ une discrétisation de $[0,T]$ de pas h. Pour toutes suites positives $(a_k)_{0\leq k\leq N}, (b_k)_{0\leq k\leq N}$ et $(c_k)_{0\leq k\leq N}$ telles que $a_{k-1} + c_{k-1} \leq (1+\gamma h)a_k + b_{k-1}$ (with $\gamma > 0$), on a*

$$a_k + \sum_{i=k}^{N-1} c_i \leq e^{\gamma(T-t_k)}\big[a_N + \sum_{i=k}^{N-1} b_i\big]. \tag{A.6}$$

A.4 Régression non-paramétrique

Nous listons ici un certain nombre de définitions et résultats exposés dans [GKKW02]. On peut également consulter sur ce sujet [Pol84], [LT91], [VDVW96] ou [Dud99]. Nous remercions N. Vayatis pour la traduction française des termes liés aux nombres de couverture.

A.4.1 Nombres de couverture

Définition A.1 *Soient $\epsilon > 0$, $\mathcal{G}$ une classe de fonctions $\mathbb{R}^d \to \mathbb{R}$, $1 \leq p < \infty$, et ν une mesure de probabilité sur $\mathbb{R}^d$. Pour une fonction $g : \mathbb{R}^d \to \mathbb{R}$, on définit*

$$\|g\|_{L_p(\nu)} := \{\int |g(z)|^p d\nu\}^{\frac{1}{p}}.$$

1. *Toute famille finie de fonctions $g_1, \cdots, g_N : \mathbb{R}^d \to \mathbb{R}$ telle que pour toute fonction $g \in \mathcal{G}$, il existe $j = j(g) \in \{1, \cdots, N\}$ tel que*

$$\|g - g_j\|_{L_p(\nu)} < \epsilon$$

est appelée une ϵ-couverture de $\mathcal{G}$ par rapport à $\|\cdot\|_{L_p(\nu)}$.

2. *On note $\mathcal{N}(\epsilon, \mathcal{G}, \|\cdot\|_{L_p(\nu)})$ la taille de la plus petite ϵ-couverture de $\mathcal{G}$ par rapport à $\|\cdot\|_{L_p(\nu)}$. On prend $\mathcal{N}(\epsilon, \mathcal{G}, \|\cdot\|_{L_p(\nu)}) = \infty$ s'il n'existe pas de ϵ-couverture. Alors $\mathcal{N}(\epsilon, \mathcal{G}, \|\cdot\|_{L_p(\nu)})$ est appelé nombre de ϵ-couverture de $\mathcal{G}$ par rapport à $\|\cdot\|_{L_p(\nu)}$.*

3. *Soit $z_1^M = (z_1, \cdots, z_M)$ M points fixes de $\mathbb{R}^d$. Soit ν_M la mesure empirique correspondante, i.e.*

$$\nu_n(A) = \frac{1}{M} \sum_{m=1}^{M} 1_A(z_m),\ (A \subset \mathbb{R}^d).$$

Alors

$$\|g\|_{L_p(\nu_M)} = \{\frac{1}{M} \sum_{m=1}^{M} |g(z_m)|^p\}^{\frac{1}{p}}$$

et toute ϵ-couverture de $\mathcal{G}$ par rapport à $\|\cdot\|_{L_p(\nu_M)}$ sera appelé une L_p ϵ-couverture de $\mathcal{G}$ sur z_1^M et le nombre de ϵ-couverture de $\mathcal{G}$ par rapport à $\|\cdot\|_{L_p(\nu_M)}$ sera noté

$$\mathcal{N}_p(\epsilon, \mathcal{G}, z_1^M).$$

En d'autres termes, $\mathcal{N}_p(\epsilon, \mathcal{G}, z_1^M)$ est l'entier N minimal tel qu'il existe des fonctions $g_1, \cdots, g_N : \mathbb{R}^d \to \mathbb{R}$ avec la propriété que pour toute $g \in \mathcal{G}$, il existe $j = j(g) \in \{1, \cdots, N\}$ tel que

$$\{\frac{1}{M} \sum_{m=1}^{M} |g(z_m) - g_j(z_m)|^p\}^{\frac{1}{p}} < \epsilon.$$

A.4.2 Coefficients d'éclatement et dimension VC

Définition A.2 *Soit $\mathcal{A}$ une classe de sous-ensembles de $\mathbb{R}^d$ et soit $M \in \mathbb{N}$.*

1. *Pour $z_1, \cdots, z_M \in \mathbb{R}^d$, on définit*

$$s(\mathcal{A}, \{z_1, \cdots, z_M\}) = \big|\{A \cap \{z_1, \cdots, z_M\} : A \in \mathcal{A}\}\big|,$$

c'est-à-dire que $s(\mathcal{A}, \{z_1, \cdots, z_M\})$ est le nombre de sous-ensembles distincts de $\{z_1, \cdots, z_M\}$ en intersectant avec des éléments de $\mathcal{A}$.

2. *Soit G un sous-ensemble de $\mathbb{R}^d$ de cardinal M. On dit que $\mathcal{A}$ fragmente G si $s(\mathcal{A}, G) = 2^M$, i.e. si tout sous-ensemble de G peut être représenté sous la forme $A \cap G$ pour un élément $A \in \mathcal{A}$.*

3. *Le M-ème coefficient d'éclatement $\mathcal{S}(\mathcal{A}, M)$ de $\mathcal{A}$ est défini comme*

$$\mathcal{S}(\mathcal{A}, M) = \max_{\{z_1, \cdots, z_M\} \subset \mathbb{R}^d} s(\mathcal{A}, \{z_1, \cdots, z_M\}).$$

Ainsi, le M-ème coefficient d'éclatement est le nombre maximal de sous-ensembles différents d'un ensemble de M points qui peuvent être obtenus par intersection avec les éléments de $\mathcal{A}$.

Définition A.3 *Soit $\mathcal{A}$ une classe de sous-ensembles de $\mathbb{R}^d$ avec $\mathcal{A} \neq \varnothing$. La dimension* **VC** *(ou dimension de Vapnik-Chervonenkis) $V_{\mathcal{A}}$ de $\mathcal{A}$ est définie par*

$$V_{\mathcal{A}} = \sup\{M \in \mathbb{N} : \mathcal{S}(\mathcal{A}, M) = 2^M\},$$

i.e. $V_{\mathcal{A}}$ est le plus grand entier M tel qu'il existe un ensemble de M points qui peut être éclaté par $\mathcal{A}$.

Définition A.4 *Soit $\mathcal{G}$ une classe de fonctions de $\mathbb{R}^d$ prenant ses valeurs dans $\mathbb{R}$. On note $\mathcal{G}^+$ le sous-ensemble de $\mathbb{R}^{d+1}$ défini par*

$$\mathcal{G}^+ = \big\{\{(z, t) \in \mathbb{R}^d \times \mathbb{R}; t \leq g(z)\}, g \in \mathcal{G}\big\}$$

soit l'ensemble des sous-graphes de fonctions de $\mathcal{G}$.

Théorème A.6 *Soit $\mathcal{G}$ une classe de fonctions $g : \mathbb{R}^d \to [0, B]$ avec $2 \leq V_{\mathcal{G}^+} < \infty$, soit $p \geq 1$, soit ν une mesure de probabilités sur $\mathbb{R}^d$, et soit $0 < \epsilon < \frac{B}{4}$. Alors*

$$\mathcal{N}(\epsilon, \mathcal{G}, \|\cdot\|_{L_p(\nu)}) \leq 3\left(\frac{2eB^p}{\epsilon^p} \log \frac{3eB^p}{\epsilon^p}\right)^{V_{\mathcal{G}^+}}.$$

Théorème A.7 *Soit $\mathcal{G}$ un espace vectoriel de dimension r de fonctions de $\mathbb{R}^d \to \mathbb{R}$ et soit*

$$\mathcal{A} = \big\{\{z : g(z) \geq 0\} : g \in \mathcal{G}\big\}.$$

Alors

$$V_{\mathcal{A}} \leq r.$$

A.4.3 Vitesse de convergence pour l'estimation moindres carrés de la fonction de régression

Supposons que l'on observe M réalisations indépendantes du couple (X, Y) notées $(X_m, Y_m)_{1\leq m\leq M}$. On note F_M un espace vectoriel de fonctions qui peut dépendre de $(X_1 \cdots, X_M)$. On définit $\tilde{m}_M$, la fonction de régression estimée, comme :

$$\tilde{m}_M(\cdot) = \arg \inf_{\psi \in F_M} \frac{1}{M} \sum_{m=1}^{M} |\psi(X_m) - Y_m|^2.$$

On introduit la notation

$$\|\psi\|_M^2 = \frac{1}{M} \sum_{m=1}^{M} |\psi(X_m)|^2.$$

Alors, on a la théorème suivant (théorème 11.1 page 184 dans [GKKW02]) :

Théorème A.8 *Supposons que*

$$\sigma^2 = \sup_{x \in \mathbb{R}^d} Var(Y|X = x) < \infty.$$

Soit K_M la dimension de F_M, alors :

$$\mathbb{E}\left(\|\tilde{m}_M - m\|_M^2 \middle| X_1, \ldots, X_M \right) \leq \sigma^2 \frac{K_M}{M} + \min_{\psi \in F_M} \|\psi - m\|_M^2.$$

Si de plus, F_M est indépendant des réalisations $(X_m, Y_m)_{1\leq m\leq M}$, on a

$$\mathbb{E}\left(\|\tilde{m}_M - m\|_M^2 \right) \leq \sigma^2 \frac{K_M}{M} + \min_{\psi \in F_M} \mathbb{E}|\psi(X) - m(X)|^2.$$

Le théorème 11.2 page 187 dans [GKKW02] nous donne

Théorème A.9 *Soit $\mathcal{G}$ une classe de fonctions $g : \mathbb{R}^d \to \mathbb{R}$ bornée par B et soit $\epsilon > 0$. Alors*

$$\mathbb{P}\big(\exists g \in \mathcal{G} : \|g\| - 2\|g\|_M > \epsilon\big) \leq 3\mathbb{E}\mathcal{N}_2(\frac{\sqrt{2}}{24}\epsilon, \mathcal{G}, X_1^{2M}) \exp(-\frac{M\epsilon^2}{288B^2})$$

où $X_1^{2M} = (X_1, \cdots, X_M, X_{M+1}, \cdots, X_{2M})$ avec $(X_{M+m})_{1\leq m\leq M}$ une copie i.i.d de $(X_m)_{1\leq m\leq M}$.

A.4.4 Inégalité de Hoeffding

Théorème A.10 *Soient $X_1, \cdots, X_M$ des variables aléatoires réelles indépendantes. Soient $a_1, b_1, \cdots, a_M, b_M$ des réels et on suppose que $X_m \in [a_m, b_m]$ avec probabilité 1 pour $1 \leq m \leq M$. Alors, $\forall \epsilon > 0$:*

$$\mathbb{P}\left(\left| \frac{1}{M} \sum_{m=1}^{M} (X_m - \mathbb{E}(X_m)) \right| > \epsilon \right) \leq 2 \exp\left(- \frac{2M\epsilon^2}{\frac{1}{M}\sum_{m=1}^{M} |b_m - a_m|^2} \right)$$

A.5 Analyse matricielle

A plusieurs reprises, on doit préciser quelle solution on choisit du problème de moindres carrés :

$$\inf_{\alpha} \frac{1}{M} \sum_{m=1}^{M} |Y_m - \alpha.p(X_m)|^2$$

où $p(\cdot)$ est une base de fonctions dépendant éventuellement des données $(X_m)_{1 \leq m \leq M}$ et de cardinal K (on suppose $K \leq M$). En outre, on suppose souvent que sans perte de généralité $\frac{B^*B}{M} = Id$ où B est la matrice de taille $M \times K$ dont les lignes sont $p(X_m)^*$.
Il n'y a pas de problème dans le cas où la matrice B^*B est inversible car il existe une unique solution au problème de moindres carrés et un procédé de Schmidt permettra bien, par une transformation linéaire de la base $p(\cdot)$ d'orthonormaliser les colonnes de la matrice B.
Par contre, dans le cas où il existe des colinéarités entre les colonnes de B, il convient de choisir une manière de résoudre le problème de moindres carrés (i.e. de choisir une solution particulière) qui soit compatible avec le fait de supposer $\frac{B^*B}{M} = Id$ sans perte de généralité. Une possibilité est la suivante.
On effectue une décomposition en valeurs singulières (SVD, voir [GVL96]) de $\frac{B}{\sqrt{M}}$ qui s'écrit $\frac{B}{\sqrt{M}} = U\Sigma V^*$ avec U une matrice orthogonale ($U^*U = Id$) de taille $M \times M$, Σ une matrice de la forme

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & \sigma_K \\ 0 & \dots & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

où $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_K \geq 0$ sont les valeurs singulières de $\frac{B}{\sqrt{M}}$ et V une matrice orthogonale de taille $K \times K$. Si toutes les valeurs singulières sont strictement positives alors on se retrouve dans le cadre d'une solution unique. Si elles sont non nulles jusqu'à un rang $r < K$ puis $\sigma_{r+1} = \cdots = \sigma_K = 0$, alors on choisit de remplacer B par $[u_1, \cdots, u_r]$ les r premières colonnes de U. Par propriété de la SVD (voir [GVL96]), on sait que ces r premières colonnes forment une base orthonormale du sous-espace vectoriel de $\mathbb{R}^M$ engendré par les colonnes de B. En outre, il est alors facile de vérifier que la solution du problème de moindres carrés une fois cette transformation effectuée est donnée par la solution de norme euclidienne minimale.
En conclusion, on a donc spécifié quelle solution au problème de moindres carrés choisir

et que l'on pouvait bien supposer que les colonnes de $\frac{B}{\sqrt{M}}$ sont orthonormales, quitte à remplacer K par $r < K$.

Bibliographie

[ABR96] L. Andersen and R. Brotherton-Ratcliffe. Exact exotics. *Risk*, 9 :85–89, 1996.

[Bal97] V. Bally. Approximation scheme for solutions of BSDE. *In Backward stochastic differential equations (Paris, 1995–1996), Edited by N. El Karoui and L. Mazliak, volume 364 of Pitman Res. Notes Math. Ser., pages 177–191. Longman, Harlow*, 1997.

[Bar02] M.T. Barlow. A diffusion model for electricity prices. *Mathematical Finance*, 12 :287–298, 2002.

[BCFEK02] V. Bally, M.E. Caballero, B. Fernandez, and N. El Karoui. Reflected BSDE's, PDE's and variational inequalities. *INRIA Report 4455*, 2002.

[BDM01] P. Briand, B. Delyon, and J. Mémin. Donsker-type theorem for BSDEs. *Electron. Comm. Probab.*, 6 :1–14, 2001.

[BEHN03] F.E. Benth, L. Ekland, R. Hauge, and B.F. Nielsen. On arbitrage-free pricing of forward contracts in energy markets. *Applied Mathematical Finance*, 10(4) :325–336, 2003.

[Ber95] Y.Z. Bergman. Option pricing with differential interest rates. *Rev. of Financial Studies*, 8(2) :475–500, 1995.

[BET04] B. Bouchard, I. Ekeland, and N. Touzi. On the Malliavin approach to Monte Carlo approximation of conditional expectations. *Finance Stochastics*, 8 :45–71, 2004.

[BN95] M.T. Barlow and D. Nualart. *Lectures on probability theory and statistics, Ecole d'été de probabilités de Saint-Flour*, volume 25 of *Lecture Notes in Mathematics 1690*. Springer, 1995.

[BP03] V. Bally and G. Pagès. Error analysis of the optimal quantization algorithm for obstacle problems. *Stochastic Process. Appl.*, 106(1) :1–40, 2003.

[BPP05] V. Bally, G. Pagès, and J. Printemps. A quantization tree method for pricing and hedging multidimensional American options. *Mathematical Finance*, 15(1), 2005.

[BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81, 1973.

[BT04] B. Bouchard and N. Touzi. Discrete time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic processes and their applications*, 111 :175–206, 2004.

[Che97] D. Chevance. Numerical methods for backward stochastic differential equations. *In numerical methods in finance, Edited by L.C.G. Rogers and D. Talay*, pages 232–244, 1997.

[CLP02] E. Clément, D. Lamberton, and P. Protter. An analysis of a least squares regression method for American option pricing. *Finance Stoch.*, 6(4) :449–471, 2002.

[CS00] L. Clewlow and C. Strickland. *Energy derivatives. Pricing and Risk management.* Lacima Publications, London, 2000.

[DE92a] D. Duffie and L. Epstein. Asset pricing with stochastic differential utility. *Review of Financial Studies*, 5 :411–436, 1992.

[DE92b] D. Duffie and L. Epstein. Stochastic differential utility. *Econometrica*, 60(2) :353–394, 1992.

[Den98] S. Deng. Stochastic models of energy commodity prices and their applications : mean-reversion with jumps and spikes. *Preprint*, 1998.

[DM04] F. Delarue and S. Menozzi. A Forward-Backward stochastic algorithm for Quasi-linear PDEs. *Rapport Interne 551, CMAP, Ecole Polytechnique*, 2004.

[DMP96] J. Douglas, J. Ma, and P. Protter. Numerical methods for forward-backward stochastic differential equations. *Annals of Applied Probability*, 6(3) :940–968, 1996.

[Dud99] R.M. Dudley. *Uniform central limit theorems.* Cambridge studies in advanced mathematics, 1999.

[DW88] J. Dupacova and R.J.-B. Wets. Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. *The Annals of Statistics*, 16 :1517–1549, 1988.

[Egl05] D. Egloff. Monte Carlo algorithms for optimal stopping and statistical learning. *To appear in Annals of Applied Probability*, 2005.

[EKKP+97] N. El Karoui, C. Kapoudjian, E. Pardoux, S. Peng, and M.C. Quenez. Reflected solutions of backward SDE's and related obstacle problems for PDE's. *Ann. Probab.*, 25(2) :702–737, 1997.

[EKPQ97] N. El Karoui, S. Peng, and M.C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1) :1–71, 1997.

[EKQ95] N. El Karoui and M.C. Quenez. Dynamic programming and pricing of contingent claims in incomplete markets. *SIAM J. Control Optim.*, 33(1) :29–66, 1995.

[FL99] H. Föllmer and P. Leukert. Quantile hedging. *Finance and Stochastics*, 3 :251–274, 1999.

[FS91] H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. In *Applied stochastic analysis (London, 1989)*, volume 5 of *Stochastics Monogr.*, pages 389–414. Gordon and Breach, New York, 1991.

[GKKW02] L. Gyorfi, M. Kohler, A. Krzyzak, and H. Walk. *A distribution-free theory of nonparametric regression.* Springer Series in Statistics, 2002.

[GL00] S. Graf and H. Luschgy. *Foundations of quantization for probability distributions.* Lecture Notes in Mathematics, 1730, Springer-Verlag, Berlin, 2000.

[GLW05] E. Gobet, J.P. Lemor, and X. Warin. A regression-based Monte-Carlo method for backward stochastic differential equations. *To appear in Annals of Applied Probability*, 2005.

[GM05] E. Gobet and R. Munos. Sensitivity analysis using Itô-Malliavin calculus and martingales. Application to stochastic optimal control. *To appear in SIAM Journal on Control and Optimization*, 2005.

[GVL96] G. Golub and C.F. Van Loan. *Matrix computations.3rd ed.* Baltimore, MD : The Johns Hopkins Univ. Press. xxvii, 694 p., 1996.

[KH01] A. Kohatsu-Higa. Weak approximations. A Malliavin calculus approach. *Math. Comp.*, 70(233) :135–172, 2001.

[KHP02] A. Kohatsu-Higa and R. Pettersson. Variance reduction methods for simulation of densities on Wiener space. *SIAM Journal on Numerical Analysis*, 40 :431–450, 2002.

[KS88] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus.* Springer-Verlag, 1988.

[KS98] I. Karatzas and S.E. Shreve. *Methods of Mathematical Finance.* Springer, 1998.

[LS01] F.A. Longstaff and E.S. Schwartz. Valuing American options by simulation : a simple least-squares approach. *Review of Financial Studies*, 14 :113–148, 2001.

[LS02] J. Lucia and E.S. Schwartz. Electricity prices and power derivatives : Evidence from the nordic power exchange. *Review of Derivatives Research*, 5(1) :5–50, 2002.

[LT91] M. Ledoux and M. Talagrand. *Probability in Banach Spaces.* Springer, 1991.

[LT01] B. Lapeyre and E. Temam. Competitive Monte Carlo methods for the pricing of asian options. *Journal of Computational Finance*, 5(1) :39–59, 2001.

[MPSMT02] J. Ma, P. Protter, J. San Martín, and S. Torres. Numerical method for backward stochastic differential equations. *Ann. Appl. Probab.*, 12(1) :302–316, 2002.

[MPY94] J. Ma, P. Protter, and J.M. Yong. Solving forward-backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields*, 98(3) :339–359, 1994.

[MZ05] J. Ma and J. Zhang. Representations and regularities for solutions to BSDEs with reflections. *To appear in Stochastic Processes and their applications*, 2005.

[Nua95] D. Nualart. *The Malliavin calculus and related topics.* Probability and its Applications. Springer-Verlag, New York, 1995.

[Par98] E. Pardoux. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. *In Stochastic Analysis and Related Topics : The Geilo Workshop, 1996, L. Decreusefond, J. Gjerde, B. Oksendal, A.S. Ustüunel eds., Birkhäuser*, pages 79–127, 1998.

[PBS01] M. Potters, J.P. Bouchaud, and D. Sestovic. Hedged Monte-Carlo : low variance derivative pricing with objective probabilities. *Phys. A*, 289(3-4) :517–525, 2001.

[Pha00] H. Pham. Dynamic L^p-hedging in discrete time under cone constraints. *SIAM J. Control Optim.*, 38(3) :665–682, 2000.

[Pol84] D. Pollard. *Convergence of stochastic processes.* Springer Series in Statistics. Springer-Verlag, New York, 1984.

[PP90] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control. Lett.*, 14 :55–61, 1990.

[PP92] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Lecture Notes in CIS, Springer-Verlag*, 176 :200–217, 1992.

[Pri99] J.L. Prigent. Incomplete markets : Convergence of options values under the minimal martingale measure. *Adv. Appl. Prob*, 31 :1058–1077, 1999.

[Pro90] P. Protter. *Stochastic integration and differential equations. A new approach*, volume 21 of *Applications of Mathematics.* Springer-Verlag, Berlin, 1990.

[REK00] R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *Mathematical Finance*, 10(2) :259–276, 2000.

[RM93] W.J. Runggaldier and Schweizer M. Convergence of option values under incompleteness. *In Seminar on Stochastic Analysis, eds E. Bolthausen, M. Dozzi and F. Russo*, pages 365–384, 1993.

[Rog02] L.C.G. Rogers. Monte Carlo valuation of American options. *Mathematical Finance*, 12 :271–286, 2002.

[Sch94] M. Schweizer. Approximating random variables by stochastic integrals. *Annals of probability*, 22 :1536–1575, 1994.

[Sch95] M. Schweizer. On the minimal martingale measure and the Föllmer-Schweizer decomposition. *Stochastic Anal. Appl.*, 13(5) :573–599, 1995.

[Sch01] M. Schweizer. A guided tour through quadratic hedging approaches. *Option pricing, interest rates and risk management, E. Jouini, J. Cvitanic, M. Musiela (eds.)*, pages 538–574, 2001.

[TVR99] J.N. Tsitsiklis and B. Van Roy. Optimal stopping of markov processes : Hilbert space theory, approximation algorithms and an application to pricing high-dimensional financial derivatives. *IEEE Trans. Autom. Control*, 44(10) :1840–1851, 1999.

[VDVW96] A.W. Van Der Vaart and J.A. Wellner. *Weak convergence and empirical processes, with applications to statistics.* Springer-Verlag, New-York, 1996.

[Zha01] J. Zhang. Some fine properties of backward stochastic differential equations. *PhD thesis, Purdue University*, 2001.

[Zha04] J. Zhang. A numerical scheme for BSDEs. *The Annals of Applied Probability*, 14(1) :459–488, 2004.