

HEC MONTRÉAL

**Création d'actifs synthétiques par treillis stochastiques**

Par

Hugues Langlois

Sous la direction de  
Pr. Bruno Rémillard  
et  
Pr. Nicolas Papageorgiou

*Mémoire présenté en vue de l'obtention du grade de maîtrise ès  
sciences (M.Sc.)  
(Ingénierie financière)*

Mai 2008

© Hugues Langlois, 2008

# Remerciements

Je tiens tout d'abord à remercier l'Institut de finance mathématique de Montréal (IFM2), HEC Montréal ainsi que le Groupe de recherche en placements alternatifs Desjardins Gestion Internationale d'Actifs - HEC Montréal pour un support financier grandement apprécié. J'aimerais aussi souligner le privilège que j'ai eu à développer ce mémoire sous la direction de Pr. Bruno Rémillard et Pr. Nicolas Papageorgiou. Vous m'avez permis d'en apprendre toujours un peu plus à chaque jour et m'avez donné de judicieux conseils, et pour cela, je vous dois toute ma reconnaissance. De plus, j'ai beaucoup appris au cours de ma maîtrise à travers mes interactions avec les autres étudiants ainsi que mes collègues du groupe de recherche. Cependant, je dois reconnaître en particulier l'apport important d'Alexandre Hocquard sans qui j'aurais littéralement progressé deux fois moins vite dans l'élaboration de ce mémoire.

Finalement, je remercie mes parents Mario et Cécile ainsi que mon frère Simon pour leur support inconditionnel. Je vous dois absolument tout.

# Table des matières

<b>Remerciements</b>	<b>i</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Revue de littérature</b>	<b>3</b>
2.1 Actifs synthétiques . . . . .	3
2.2 Treillis stochastiques . . . . .	7
<b>3 Création d'actifs synthétiques</b>	<b>10</b>
3.1 Modélisation des rendements . . . . .	10
3.2 Fonction de flux monétaires . . . . .	12
3.3 Réplication de l'option . . . . .	14
3.3.1 Algorithme de réplication optimale . . . . .	14
<b>4 Treillis stochastiques</b>	<b>21</b>
4.1 Fondements théoriques . . . . .	21
4.2 Choix de la densité de transition . . . . .	24
4.3 Approche combinée . . . . .	27
<b>5 Méthode multipôle rapide</b>	<b>29</b>
5.1 Transformée gaussienne rapide . . . . .	31
5.1.1 Fondements théoriques . . . . .	32

---

5.1.2	Implémentation de la transformation . . . . .	36
5.1.3	Exemple . . . . .	37
5.2	Extension pour les mixtures de densités gaussiennes . . . . .	39
5.3	Application à l'algorithme de répliation optimale . . . . .	40
5.4	Transformée gaussienne rapide améliorée . . . . .	41
5.4.1	Algorithme de regroupement . . . . .	42
5.4.2	Expansion de Taylor . . . . .	43
5.5	Treillis stochastiques améliorés . . . . .	46
<b>6</b>	<b>Performances</b>	<b>49</b>
6.1	Sensibilité aux paramètres du treillis stochastique amélioré . . . . .	50
6.2	Variance des estimateurs et temps de calcul . . . . .	52
6.3	Évaluation de la stratégie de répliation . . . . .	52
6.4	Cas non linéaire . . . . .	55
<b>7</b>	<b>Conclusion</b>	<b>58</b>
	<b>Annexe</b>	<b>60</b>
	<b>Bibliographie</b>	<b>62</b>

# Table des figures

4.1	(a) Génération de vecteur $X_t$ avec $b = 3$ et $T = 2$ (b) treillis stochastique. . . . .	22
5.1	Séparation uniforme des grilles. . . . .	42
5.2	Séparation par l'algorithme du point le plus lointain et séparation uniforme de 50000 points provenant d'une mixture de gaussienne avec deux régimes ( $\mu_1 = (-3, -3)^\top$ , $\mu_2 = (3, 3)^\top$ , $\rho_1 = 0$ , $\rho_2 = 0.5$ ). . . . .	44
5.3	Calcul du vecteur de coefficients pour l'expansion de Taylor multivariée avec $d = 2$ et $n = 3$ . . . . .	45
6.1	Fonction de flux monétaires à échéance donnée par $\tilde{g}$ . . . . .	51
6.2	Temps de calcul selon le nombre de trajectoires. . . . .	54
6.3	Fonction de flux monétaires à échéances non-linéaire donnée par $\tilde{g}$ . . . . .	57

# Liste des tableaux

6.1	Paramètre des mixtures de gaussiennes avec 4 régimes pour modéliser les rendements journaliers. . . . .	50
6.2	Paramètre de la loi gaussienne pour modéliser les rendements mensuels.	50
6.3	Sensibilité au paramètre $p$ obtenue à partir d'un échantillon de 5000 trajectoires avec $v = 8$ et $r = 0.5$ . . . . .	51
6.4	Sensibilité au paramètre $v$ obtenue à partir d'un échantillon de 5000 trajectoires avec $p = 8$ et $r = 0.5$ . . . . .	51
6.5	Sensibilité au paramètre $r$ obtenue à partir d'un échantillon de 5000 trajectoires avec $p = 8$ et $v = 8$ . . . . .	52
6.6	Moyennes et écart-types des estimateurs obtenus à partir de 1000 resimulations. . . . .	53
6.7	Résultats de réplication avec 250000 resimulations pour $g(R_T^{(1)}, R_T^{(2)}) = \log(\frac{C_T}{100})$ et $\log(\frac{V_T}{100})$ . . . . .	55
6.8	Résultats de réplication avec 250000 resimulations pour $g(R_T^{(1)}, R_T^{(2)}) = \log(\frac{C_T}{100})$ non linéaire et $\log(\frac{V_T}{100})$ . . . . .	57

# Chapitre 1

## Introduction

Les actifs ou fonds synthétiques sont un sujet de recherche récent qui a été enfanté par la vaste littérature sur l'évaluation de la performance. Tout d'abord, des modèles de réplication de fonctions de répartition et de mesures de dépendance développés dans Amin et Kat (2003, [1]) et Kat et Palaro (2005, [21]) avaient pour but d'évaluer la performance de fonds de couvertures. En effet, en tentant de répliquer la distribution historique des rendements d'un fond en particulier grâce à une règle d'investissements dynamiques, le talent et la performance du gestionnaire pouvaient être comparés au coût de la réplication de leur performance, permettant ainsi évaluer si les frais demandés étaient justifiés.

Cette méthodologie a par la suite soulevé la possibilité de créer des actifs possédant des caractéristiques préétablies. Précisément, il s'agit de créer un actif synthétique  $A^{(3)}$  dont les rendements suivent une distribution de probabilité et démontrent une dépendance avec un actif de référence  $A^{(1)}$ . Cette méthode présente l'avantage majeur de permettre de choisir à l'avance cette distribution de rendements et cette dépendance. Dans la majorité des cas, cet actif de référence sera le portefeuille initialement détenu par l'investisseur. La création de l'actif synthétique impliquera aussi un actif de réserve  $A^{(2)}$  qui servira de source de risque additionnelle dans la création de  $A^{(3)}$ . Bref, l'actif de référence et l'actif de réserve seront transigés dynamiquement afin de répliquer un actif fictif ayant les caractéristiques statistiques voulues.

Le potentiel de cette technologie dépasse largement l'évaluation de la performance.

Non seulement est-il possible de créer un actif qui possédera les caractéristiques optimales afin de diversifier le risque intrinsèque de l'actif de référence dans le contexte où ce dernier est un portefeuille détenu par l'investisseur, mais les applications en gestion des risques sont aussi nombreuses. Par exemple, la possibilité de créer un actif synthétique sera de grande utilité pour un fond de pension aux prises avec un débalancement entre actifs et passifs.

Alors que le prochain chapitre présente une revue de la littérature pertinente, le troisième chapitre détaille les étapes requises dans la création d'actifs synthétiques, le chapitre 4 survole la méthode de treillis stochastique alors que le chapitre 5 apporte des améliorations techniques. Finalement, le sixième chapitre évalue la performance des algorithmes utilisés.

## Chapitre 2

# Revue de littérature

Cette revue se présente en deux grandes étapes. Tout d'abord, nous reviserons l'évolution de la littérature concernant le modèle de réplication de distribution de rendements et de mesures de dépendance. Ensuite, les raffinements de cette méthode apportés par Papageorgiou, Rémillard et Hocquard (2007, [24]) requièrent une technologie de réplication plus avancée qu'une tarification dans un monde Black-Scholes-Merton traditionnel. Précisément, l'algorithme récursif proposé s'apparente à un problème de tarification d'options américaines bien qu'il ne soit nullement question d'exercice avant échéance. Alors que l'estimation d'une valeur d'option implique traditionnellement une intégration numérique à partir d'une génération de trajectoires aléatoires, le caractère américain d'un produit dérivé demande une approche récursive de programmation dynamique afin de déterminer une zone optimale d'exercice. Plusieurs hybrides, c'est-à-dire un modèle combinant une génération de trajectoires aléatoires et des aspects de programmation dynamique, ont été proposés afin de remédier à ce problème et un traitement général se retrouve dans Glasserman (2004, [15]). Par contre, nous nous concentrerons dans la deuxième section de cette revue de la littérature sur une technique de treillis stochastiques détaillée dans Broadie et Glasserman (2004, [6]).

### 2.1 Actifs synthétiques

À l'aide d'arbres binomiaux, Dybvig (1988, [12]) évalue le coût de l'inefficience de certains fonds d'investissement. Spécifiquement, en imposant les hypothèses ci-dessous

sous-jacentes à son modèle de tarification de distribution de flux monétaires (*Payoff Distribution Pricing Model*), l'auteur arrive à déterminer le coût de la stratégie permettant de générer une fonction de répartition.

1. les préférences des agents économiques dépendent seulement de la fonction de répartition de la richesse finale,
2. les agents économiques préfèrent plus de richesse à moins,
3. le marché est complet et parfait (pas de taxes, de coûts de transactions ou d'asymétrie d'information),

De plus, Dybvig (1988, [11]) soutient que la façon la moins onéreuse pour obtenir une distribution de flux monétaires est d'allouer la richesse terminale par une fonction décroissante de la densité état-prix. En se basant sur les travaux de Dybvig (1988, [12]), Amin et Kat (2003, [1]) introduisent un système d'investissement dynamique permettant de reproduire une fonction de densité de probabilité. Le point majeur est que, malgré leurs calculs effectués dans un monde Black-Scholes-Merton, les stratégies de réplication dynamique permettent de recréer les moments d'ordre supérieurs de la distribution marginale des rendements du fond cible. Développant la méthodologie de Amin et Kat (2003, [1]), Kat et Palaro (2005, [21]) proposent de répliquer le coefficient de corrélation de Pearson en relation avec un portefeuille initial de l'investisseur en plus de la loi marginale du fond de couverture. Cette amélioration est fondée sur le fait que la surperformance de ces fonds a fondu au cours des années et leur argument majeur de vente est maintenant leur faible corrélation avec des portefeuilles traditionnels. Afin d'intégrer cette nouvelle notion dans le cadre théorique de Dybvig (1988, [12]), ils modifient la première hypothèse ci-dessus afin que les préférences des agents économiques dépendent de la distribution de probabilité conjointe de leur richesse finale et d'un actif de référence.

La méthode Kat-Palaro se divise en trois étapes. Il est tout d'abord approprié de modéliser les rendements du portefeuille de l'investisseur et du portefeuille de réserve. Ce dernier servira de source de risque additionnelle afin de pouvoir générer la fonction de répartition voulue. Cette étape comprend aussi le choix de la distribution cible de l'actif synthétique que l'on cherche à créer. Il est à noter que les dépendances entre le portefeuille et la réserve et entre le portefeuille et l'actif synthétique désiré sont modélisées à l'aide de copules. Donc, le choix des caractéristiques synthétiques comprend aussi la forme de la dépendance avec le portefeuille détenu par l'investisseur. Ensuite, il s'agit de calculer une fonction  $\tilde{g}$  de flux monétaires à échéance dont

une dérivation complète se retrouve dans Kat et Palaro (2005, [21]). Finalement, la dernière étape se réduit au problème fondamental de l'ingénierie financière, c'est-à-dire tarifier l'option payant à échéance la fonction  $g$  et déterminer la meilleure stratégie de réplication.

En guise d'exemple, Kat et Palaro (2005, [20]) tentent de répliquer la loi des fonds *Quantum Emerging Growth Fund* de Soros Fund Management et *Vega Global Fund* de Vega Asset Management en utilisant une réserve constituée de contrats à terme sur l'Eurodollar à 3 mois et un portefeuille formé de l'indice S&P 500 et de T-Bond. Dans les deux situations, les auteurs attribuent la plus grande volatilité de leur réplique aux erreurs de réplication tandis que le rendement inférieur dans le premier cas ne serait dû qu'au talent d'investisseur de George Soros.

Bien que le but initial de cette littérature soit l'évaluation de la performance de fonds d'investissement, le potentiel des méthodes proposées déborde largement ce cadre. En effet, tel qu'avancé par Kat et Palaro (2005, [20], [21]), les investisseurs ne sont plus bornés à chercher dans le marché les actifs possédant les caractéristiques qu'ils désirent. La création et la combinaison optimale d'actifs démontrant un profil de risque fait sur mesure ouvrent de nouvelles possibilités, par exemple, une meilleure gestion des risques par les compagnies d'assurance et fonds de pension. Kat et Palaro (2006, [22]) testent cette approche en répliquant 4 fonds démontrant des propriétés statistiques différentes. Il est primordial de mentionner que les caractéristiques statistiques recherchées sont les quatre premiers moments ainsi que la dépendance avec un portefeuille de référence. Par contre, bien que le premier moment soit répliqué, le rendement observé du portefeuille dépendra de plusieurs facteurs tels que la capacité à répliquer l'option, la composition du portefeuille de réserve et les conditions économiques. Bref, la prime de rendement, c'est-à-dire la différence entre le rendement observé et l'espérance de rendement ciblée sera déterminée en partie par les marchés des capitaux. Ainsi, les fonds produits ne constituent pas à prime abord des investissements véritablement supérieurs lorsqu'ils sont considérés séparément, par contre, leurs propriétés statistiques prédéterminées font d'eux d'excellents ajouts dans un contexte d'optimisation de portefeuille. Ces conclusions, en plus de la vitesse de la convergence des moments, de la corrélation ainsi que des coûts de transactions engendrés par les rebalancements quotidiens sont aussi discutés dans Kat (2006, [19]). Également, l'efficacité de la réplication sera facilitée si le por-

tefeuille de réserve possède un coefficient d'asymétrie et d'aplatissement similaires à ceux recherchés. Le même argument ne s'applique cependant pas pour le coefficient de corrélation. Dans une même optique, les résultats de Papageorgiou, Rémillard et Hocquard (2007, [24]) démontrent l'importance du choix du portefeuille de réserve sur l'efficacité de la réplication dans un contexte de réplication de rendements de fonds de couverture.

La méthodologie de Kat et Palaro (2005, [21]), bien qu'innovatrice, présente cependant certaines failles. Dans la première étape de la réplication, ils testent 3 différentes fonctions marginales, soient la gaussienne, la student et la Johnson SU. De plus, ils examinent les copules gaussienne et student de la famille méta-elliptique ainsi que les copules Gumbel, Cook-Johnson et Frank de la famille archimédienne afin de modéliser la dépendance entre les rendements. Une dernière copule ajoutée à la liste, soit la symétrique Joe-Clayton introduite par Patton(2005, [25]) dans un contexte de dépendance entre taux de change, permet de contrôler séparément la dépendance dans les ailes supérieure et inférieure de la distribution. La première faille se présente au niveau de cette étape. En effet, la compatibilité entre les distributions des rendements journaliers nécessaires à la réplication dynamique et les lois mensuelles requises dans le calcul de la fonction  $g$  n'est pas discutée. Afin de palier à ce problème, Papageorgiou, Rémillard et Hocquard (2007, [24]) proposent de modéliser les rendements journaliers à l'aide de mixtures de gaussiennes et de résoudre par la suite pour la distribution mensuelle afin que la somme des densités quotidiennes sur le mois qui sera aussi une mixture de gaussiennes soit compatible avec la loi des rendements mensuels. De plus, ils proposent un test d'adéquation permettant de déterminer le nombre approprié de régimes.

Ensuite, alors que les rendements des actifs sont non-gaussiens, la tarification de l'option implicite à la réplication de Kat et Palaro est opérée dans un cadre Black-Scholes-Merton. Papageorgiou, Rémillard et Hocquard (2007, [24]) adaptent une méthodologie de tarification d'options américaines dans laquelle l'espérance du carré de l'erreur de réplication est minimisée. Spécifiquement, ils proposent un algorithme récursif qui permet de trouver la valeur de la stratégie de réplication. De plus, cette technique de tarification, qui s'assimile à la programmation dynamique retrouvée dans la tarification de produits dérivés américains, utilise les probabilités physiques.

## 2.2 Treillis stochastiques

Le type d'algorithme récursif discuté dans la dernière section demande de façon répétitive l'évaluation d'espérances conditionnelles. Ce problème est souvent rencontré lors de la tarification d'options américaines et plusieurs modèles ont vu le jour. Broadie et Glasserman (1997, [5]), ayant développé une méthodologie d'arbres simulés ne cachaient pas le coût computationnel exorbitant variant exponentiellement avec le nombre de dates d'exercice. Broadie et Glasserman (2004, [6]) (en circulation depuis 1997) introduisent une technique de treillis stochastiques afin de tarifer une option américaine dans laquelle le temps de calcul est linéairement relié au nombre de pas de temps. De plus, ils développent leur méthode dans une optique de produits dérivés ayant plusieurs sources de risques sous-jacentes.

Le principe de base des treillis stochastiques est de simuler un nombre fixe de points à chaque pas de temps et de permettre la transition entre un point et tous les points du pas de temps suivant. Cette méthode procure beaucoup de flexibilité dans son application et le choix de la densité de transition, c'est-à-dire la probabilité associée à la transition d'un nœud aux autres au pas de temps suivant, doit être judicieusement considéré. Les auteurs proposent un choix de densité de probabilité faisant en sorte que la valeur de la variante européenne est la moyenne des flux monétaires à échéance. En fait, les méthodes proposées par Longstaff et Schwartz (2001, [23]) et Tsitsiklis et Van Roy (1999, [28]) impliquent aussi un choix de poids affectés aux différentes branches, mais ceux-ci sont déterminés par régression. Une présentation plus générale de la méthode de treillis stochastiques se retrouve aussi dans Glasserman (2004, [15]). Broadie, Glasserman et Ha (2000, [7]) investiguent plus profondément le choix de poids, particulièrement dans les cas où la densité de transition du processus sous-jacent est inconnue ou n'existe pas. Les poids sont choisis par optimisation sous certaines contraintes d'appariement des moments. En général, les contraintes seront moins nombreuses que les poids et le problème sera par conséquent sous-déterminé. Afin de choisir parmi l'ensemble de poids satisfaisant les contraintes, ils implémentent deux critères d'optimisation, soit par maximum d'entropie et par moindres carrés. Bien que l'utilisation des moindres carrés soit moins onéreuse en temps de calcul, le critère du maximum d'entropie nous procure une solution non négative si elle existe. Ce choix de poids est étroitement lié aux méthodes de simulation Monte Carlo pondérée. Dans ce cas, Glasserman et Yu (2005, [16]) démontrent que pour les estimateurs biaisés obtenus lors de telles simulations,

le choix du critère d'optimisation affectera les propriétés asymptotiques mais les résultats obtenus ne leur permettent pas de favoriser un critère en particulier.

Il est à noter que la méthode de treillis stochastique énoncée dans Broadie et Glasserman (2004, [6]) produit un estimateur pour la valeur de l'option américaine biaisé vers le haut et un biaisé vers le bas. Ces estimateurs sont ensuite combinés pour engendrer un intervalle de confiance d'un niveau au moins  $(1-\alpha)$  % pour la valeur de l'option américaine. En d'autres mots, la borne inférieure de l'estimateur biaisé vers le bas et la borne supérieure de celui biaisé vers le haut créent un intervalle conservateur. Dans un même ordre d'idée, Boyle, Kolkievicz et Tan (2000, [4]) développent des techniques pour réduire les biais des estimateurs en utilisant des suites à discrétion faible. Similairement, Broadie et Glasserman (2004, [6]) soulignent que le succès de la méthode repose sur une utilisation extensive de techniques de réduction de la variance. Précisément, ils procèdent à des simulations Quasi-Monte Carlo accompagnées de plusieurs variables de contrôle. Toujours afin de quantifier l'erreur produite par cette méthode, Avramidis et Matzinger (2002, [2]) développent une borne supérieure sur la probabilité d'erreur de l'estimateur en fonction du nombre de nœuds  $b$  du treillis. De plus, ils soulignent que contrairement aux méthodes de régression de Longstaff et Schwartz (2001, [23]) et Tsitsiklis et Van Roy (1999, [28]) qui impliquent une erreur lorsque la valeur de continuation est estimée en projetant les valeurs futures de l'option sur l'espace euclidien engendré par les fonctions de bases utilisées, la convergence de l'estimateur du treillis stochastique est seulement fonction de la taille de l'échantillon Monte Carlo.

Le principal avantage de la méthode Broadie-Glasserman est que le coût computationnel est linéairement proportionnel au nombre de dates d'exercice. Par contre, cette technique est toujours affligée d'un temps de calcul relié quadratiquement au nombre de nœuds  $b$  utilisé dans le treillis puisque chaque embranchement requiert le calcul de  $b$  poids de transition afin d'obtenir la moyenne pondérée. Broadie et Yamamoto (2002, [8]) utilisent la méthode de transformée gaussienne rapide (*Fast Gauss Transform*) pour réduire le coût computationnel à chaque pas de temps dans l'algorithme récursif d'une opération  $O(b^2)$  à  $O(b)$ . En procédant à une approximation de l'évaluation des poids grâce à la méthode de transformée gaussienne rapide, une accélération considérable du temps de calcul est notée lors de la tarification en une, deux ou trois dimensions. Introduite par Greengard et Strain (1991,[18]), la

transformée gaussienne rapide permet de réduire significativement le temps de calcul lorsqu'un problème implique l'évaluation d'une certaine interaction entre chacune des paires possibles dans un ensemble de particules. Alors que cette technique engendre une erreur dépendant de la précision recherchée, les borne pour cette erreur initialement dérivée par Greengard et Strain (1991,[18]) ont été corrigée successivement par Baxter et Roussos (2002, [26]) et par Wan et Karniadakis (2006, [29]). Finalement, plusieurs critiques de la méthode sont avancées par Yang, Duraiswami, Gumerov et Davis (2003, [30]) qui proposent une autre transformée gaussienne rapide dans un contexte d'estimateur à noyau de densité utilisé en reconnaissance d'images. Parmi les changements apportés, il utilisent des algorithmes de regroupement de points tels que ceux fournis par Gonzalez (1985, [17]) ou par Feder et Greene (1988, [13]) afin de procéder à la partition de l'espace nécessaire à l'application de la méthode.

Au cours du prochain chapitre, nous débuterons par l'énoncé de toutes les étapes requises à la création d'actifs synthétiques.

## Chapitre 3

# Création d'actifs synthétiques

Nous débuterons par l'énumération des grandes étapes requises dans le processus de création d'actifs synthétiques. Globalement, la technologie de réplication de distributions de rendements et de mesures de dépendance peut être séparée en trois parties qui seront décrites plus en détail dans les sections suivantes.

1. Modélisation des lois marginales de  $A^{(1)}$  et  $A^{(2)}$  ainsi que de la copule entre ces deux actifs.
2. Choix de la fonction de répartition désirée de  $A^{(3)}$ , de la copule entre  $A^{(1)}$  et  $A^{(3)}$  ainsi que le calcul de la fonction  $g$ .
3. Réplication de l'option avec fonction de flux monétaires  $\tilde{g}$  à échéance.

### 3.1 Modélisation des rendements

Soit l'actif  $A^{(1)}$  représentant dans la majorité des cas le portefeuille déjà détenu de l'investisseur et l'actif  $A^{(2)}$  un autre produit investissable. Tel que nous le verrons dans la prochaine section, ces deux actifs seront transigés de façon dynamique afin de reproduire les propriétés statistiques de l'actif désiré. Bien que la modélisation de la loi des rendements d'actifs financiers constitue un des problèmes fondamentaux de la finance empirique, une difficulté supplémentaire se présente dans ce contexte. En effet, alors que la reproduction des propriétés statistiques voulues sera calculée sur la base des distributions mensuelles de  $A^{(1)}$ ,  $A^{(2)}$  et  $A^{(3)}$ , le portefeuille de réplication sera transigé sur une base quotidienne. Donc, il sera impératif que les lois quotidiennes utilisées soient compatibles avec les lois mensuelles. L'approche adop-

tée par Papageorgiou, Rémillard et Hocquard (2007, [24]) consiste à modéliser la loi conjointe de  $A^{(1)}$  et  $A^{(2)}$  avec des mixtures de gaussiennes bivariées. Brièvement, un vecteur aléatoire multivarié  $X$  de dimension  $d$  est une mixture de gaussiennes de  $M$  régimes avec paramètres de poids  $\{\nu_m\}_{m=1}^M$ , de moyennes  $\{\mu_m\}_{m=1}^M$  et covariances  $\{\Sigma_m\}_{m=1}^M$  si sa densité est donnée par

$$f(x) = \sum_{m=1}^M \nu_m \phi_d(x; \mu_m, \Sigma_m) \quad (3.1)$$

où

$$\phi_d(x; \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)} \quad (3.2)$$

est la densité gaussienne multivariée. La propriété importante dans ce cas est que la somme de mixtures de densités gaussiennes indépendantes est aussi une mixture de densités gaussiennes. Donc, la somme de ces lois journalières arrivera aussi à une mixture de gaussiennes pour les lois mensuelles. Formellement, soit

$$X = \sum_{i=1}^n X_i \quad (3.3)$$

où  $X_i, i \in \{1, \dots, n\}$  sont des mixtures de gaussiennes indépendamment et identiquement distribuées. Alors,  $X$  est aussi une mixture de gaussiennes. Afin de décrire les paramètres de cette dernière mixture, nous utiliserons l'ensemble  $\mathcal{A}$  composé des éléments  $\alpha_i$  représentant le nombre de fois où le régime  $i$  se retrouve parmi les  $n$  tirages

$$\mathcal{A} = \left\{ \alpha = (\alpha_1, \dots, \alpha_m); \alpha_j \geq 0 \text{ et } \sum_{i=1}^m \alpha_i = n \right\}. \quad (3.4)$$

Le nombre de régimes de la mixture  $X$  est donnée par la cardinalité de l'ensemble  $\mathcal{A}$ , c'est-à-dire  $C_{m-1}^{n+m-1}$ . Les paramètres de la distribution de  $X$  sont  $(\nu_\alpha)_{\alpha \in \mathcal{A}}$ ,  $(\mu_\alpha)_{\alpha \in \mathcal{A}}$ ,  $(\Sigma_\alpha)_{\alpha \in \mathcal{A}}$ , où pour chacun des  $\alpha \in \mathcal{A}$

$$\begin{aligned}
\nu_\alpha &= \frac{n!}{\alpha_1! \dots \alpha_m!} \prod_{m=1}^M \nu_m^{\alpha_m}, \\
\mu_\alpha &= \sum_{m=1}^M \alpha_m \mu_m, \\
\Sigma_\alpha &= \sum_{m=1}^M \alpha_m \Sigma_m.
\end{aligned} \tag{3.5}$$

Papageorgiou, Rémillard et Hocquard (2007, [24]) remarquent que si  $n$  est moindrement grand, alors  $m^n$  est très grand et il est ardu d'estimer chacun des paramètres. Par contre, la plupart des poids  $\nu_m$  pourraient être minimes et la mixture de gaussiennes  $X$  serait en toute fin pratique une mixture de quelque régime. Par conséquent, l'approche qu'ils adoptent consiste à estimer les paramètres mensuels directement à partir des rendements mensuels.

Nous ne nous préoccupons pas d'estimation de distributions dans ce mémoire. Par contre, il est important de mentionner que ces modèles sont estimés en utilisant l'algorithme EM (*Expectation-Maximization*) de Dempster, Laird et Rubin (1977, [10]) et Papageorgiou, Rémillard et Hocquard (2007, [24]) proposent un test d'adéquation basé sur les travaux de Genest, Rémillard et Beaudoin (2007, [14]) utilisant la transformée de Rosenblatt.

### 3.2 Fonction de flux monétaires

Cette section présente les étapes de la dérivation de la fonction de flux monétaires à répliquer afin de créer synthétiquement les propriétés statistiques voulues. Une dérivation détaillée se retrouve aussi dans Kat et Palaro (2005, [21]).

À partir de deux actifs  $A^{(1)}$  et  $A^{(2)}$ , il s'agit de trouver une fonction  $g$  telle que pour un actif fictif  $A^{(3)}$  possédant les caractéristiques prédéterminées, nous avons

$$P\{A^{(1)} \leq x, g(A^{(1)}, A^{(2)}) \leq y\} = P\{A^{(1)} \leq x, A^{(3)} \leq y\} \quad x, y \in \mathbb{R}. \tag{3.6}$$

En d'autres termes, il s'agit de répliquer la loi conjointe de  $A^{(1)}$  et  $A^{(3)}$  en investissant

seulement dans  $A^{(1)}$  et  $A^{(2)}$ . Bien entendu, les propriétés statistiques recherchées sont habituellement exprimées en terme de rendements. Donc en normalisant tous les actifs à un prix initial de 100\$ et en passant en log-rendements, c'est-à-dire en définissant  $R_i = \ln\{\frac{A^{(i)}}{100}\}, \forall i \in \{1, 2, 3\}$ , l'équation (3.6) peut être exprimée en terme de probabilités conditionnelles

$$P\{g(R_1, R_2) \leq y \mid R_1 \leq x\} = P\{R_3 \leq y \mid R_1 \leq x\} \quad x, y \in \mathbb{R} \quad (3.7)$$

ce qui nous permet de déduire

$$g(x, y) = F_{3|1}^{-1}[F_{2|1}(y \mid x) \mid x] \quad x, y \in \mathbb{R} \quad (3.8)$$

où  $F_{\cdot|1}$  est une fonction de répartition conditionnelle. De plus, il est aussi possible de travailler séparément avec les fonctions de répartition marginales et les copules. Dans ce cas, (3.8) devient

$$g(x, y) = \kappa_{3|1}^{-1}[\kappa_{2|1}(y \mid x) \mid x] \quad x, y \in \mathbb{R} \quad (3.9)$$

où

$$\kappa_{j|i}(y \mid x) = \frac{\partial C_{i,j}(u, v)}{\partial u} \Big|_{u=F_i(x), v=F_j(y)} \quad x, y \in \mathbb{R}$$

où  $C$  est une copule. Tel que mentionné dans la section précédente, la loi conjointe de  $A^{(1)}$  et  $A^{(2)}$  est modélisée par une mixture de densités gaussiennes. Donc, nous terminons par énoncer la distribution conditionnelle d'une mixture de densité gaussienne bivariée qui sera nécessaire dans le calcul de la fonction  $g$ . Soit  $\beta_m = \rho_m \frac{\sigma_m^{(2)}}{\sigma_m^{(1)}}$  et  $\alpha_m = \mu_m^{(2)} - \beta_m \mu_m^{(1)}$ ,  $m \in \{1, \dots, M\}$ . La distribution conditionnelle de  $A^{(2)}$  sachant  $A^{(1)} = x$  est aussi une mixture de gaussiennes avec paramètres  $\{\tilde{\nu}_m(x)\}_{m=1}^M$ ,  $\{\tilde{\mu}_m(x)\}_{m=1}^M$ ,  $\{\tilde{\sigma}_m^2(x)\}_{m=1}^M$ , où

$$\begin{aligned}
\tilde{\nu}_m(x) &= \frac{\nu_m \phi(x; \mu_m^{(1)}, (\sigma_m^{(2)})^2)}{\sum_{j=1}^M \nu_j \phi(x; \mu_j^{(1)}, (\sigma_j^{(1)})^2)}, \\
\tilde{\mu}_m(x) &= \alpha_m + \beta_m x, \\
\tilde{\sigma}_m^2 &= \sigma_m^2 (1 - \rho_m^2), \\
\phi(x; \mu, \sigma^2) &= \frac{e^{-\frac{(x-\mu)^2}{\sigma^2}}}{\sqrt{2\pi\sigma^2}}.
\end{aligned} \tag{3.10}$$

### 3.3 Réplication de l'option

Finalement, le succès de la méthode repose sur notre capacité à répliquer un produit dérivé dont la fonction de flux monétaires à échéance est

$$\tilde{g}(A^{(1)}, A^{(2)}) = 100e^{g[\ln(\frac{A^{(1)}}{100}), \ln(\frac{A^{(2)}}{100})]}. \tag{3.11}$$

En d'autres mots, un actif synthétique est une option donnant à son détenteur un rendement sur un notional de 100 \$ provenant d'une distribution voulue. Alternativement, le rendement de détention de cette option est une réalisation de la distribution choisie et ainsi, un achat successif de telles options produira une séquence de rendements démontrant les propriétés statistiques désirées.

#### 3.3.1 Algorithme de réplication optimale

Une réplication efficace de la fonction de flux monétaires à échéance  $\tilde{g}$  permettra de déterminer la valeur initiale  $v_0$ . Par contre, il s'avère beaucoup plus important dans la création d'actifs synthétiques de déterminer la stratégie de réplication dynamique, c'est-à-dire les poids  $\varphi_t^1$  et  $\varphi_t^2$  à investir dans  $A_t^{(1)}$  et  $A_t^{(2)}$  respectivement. Nous adopterons l'approche de Papageorgiou, Rémillard et Hocquard (2007, [24]), appelée *réplication optimale*, visant à minimiser l'espérance du carré des erreurs de réplication

$$E[G_T^2(v_0, \varphi)] = E[\beta_T^2 \{V_T(v_0, \varphi) - \tilde{g}(A_T^{(1)}, A_T^{(2)})\}^2]. \tag{3.12}$$

Nous présentons ici une description détaillée de la méthodologie utilisée dans Papa-

georgiou, Rémillard et Hocquard (2007, [24]). Il est à noter que ce qui suit est une extension de Schweizer (1995, [27]) et qu'il ne s'agit pas d'une évaluation sous une mesure neutre au risque. En effet, tous les calculs sont effectués sous une mesure de probabilité physique.

Soit  $(\Omega, \mathcal{P}, \mathcal{F})$  un espace probabilisé avec filtration  $\mathbb{F} = \{\mathcal{F}_0, \dots, \mathcal{F}_T\}$  sur lequel les processus stochastiques sont définis. Nous avons le vecteur de prix  $A_t = (A_t^{(1)}, A_t^{(2)})^\top$ , le facteur d'actualisation  $\beta_t$  ainsi que le vecteur de poids  $\varphi_t = (\varphi_t^{(1)}, \varphi_t^{(2)})^\top$ . Il est à noter que  $\beta_t$  et  $\varphi_t$  sont  $\mathcal{F}_{t-1}$  mesurable pour tout  $t = 1, \dots, T$ . En effet, nous connaissons dès l'instant  $t-1$  la valeur au temps  $t$  d'un montant investi dans l'actif sans risque au temps  $t-1$ . Aussi,  $\varphi_t$  est un processus stochastique prévisible puisqu'il représente les investissements dans les actifs pour la période  $(t-1, t]$  (avec  $\varphi_0 = \varphi_1$ ) et ne dépend que de  $A_0, \dots, A_{t-1}$ . Puisque la stratégie de réplication doit être auto-financée, nous avons

$$\beta_t V_t(V_0, \varphi) - \beta_{t-1} V_{t-1}(V_0, \varphi) = \varphi_t^\top (\beta_t A_t - \beta_{t-1} A_{t-1}), \quad \forall t \in \{1, \dots, T\} \quad (3.13)$$

et par conséquent,

$$\beta_T V_T(V_0, \varphi) = v_0 + \sum_{t=1}^T \varphi_t^\top (\beta_t A_t - \beta_{t-1} A_{t-1}). \quad (3.14)$$

Définissons maintenant

$$\Delta_t = A_t - E[A_t \mid \mathcal{F}_{t-1}], \quad t = 1, \dots, T. \quad (3.15)$$

En supposant que les processus  $A_t$  et  $\varphi_t^\top (\beta_t A_t - \beta_{t-1} A_{t-1})$  sont de carré intégrables, la matrice de covariance conditionnelle existe et est donnée par

$$\Sigma_t = E[\Delta_t \Delta_t^\top \mid \mathcal{F}_{t-1}], \quad 1 \leq t \leq T. \quad (3.16)$$

**Theorème 1** (Papageorgiou, Rémillard et Hocquard (2007, [24])). *Si  $\Sigma_t$  est inversible pour tout  $t = 1, \dots, T$ , alors le risque  $E[G_T^2(v_0, \varphi)]$  est minimisé en choisissant récursivement  $\varphi_T, \dots, \varphi_1$  satisfaisant*

$$\varphi_t = (\Sigma_t)^{-1} E[\{A_t - E[A_t | \mathcal{F}_{t-1}]\} C_t | \mathcal{F}_{t-1}], \quad t = T, \dots, 1 \quad (3.17)$$

où  $C_T, \dots, C_0$  sont déterminés récursivement en initialisant  $C_T = C = \tilde{g}(A_T^{(1)}, A_T^{(2)})$  et

$$C_{t-1} = \frac{\beta_t}{\beta_{t-1}} E[C_t | \mathcal{F}_{t-1}] - \varphi_t^\top E\left[\frac{\beta_t}{\beta_{t-1}} A_t - A_{t-1} | \mathcal{F}_{t-1}\right] \quad t = T, \dots, 1. \quad (3.18)$$

Dans ce cas, la valeur optimale de  $V_0$  est  $C_0$ , et

$$E[G^2] = \sum_{t=1}^T E[\beta_t^2 G_t^2], \quad (3.19)$$

où

$$G_t = \varphi_t^\top \{A_t - E[A_t | \mathcal{F}_{t-1}]\} - \{C_t - E[C_t | \mathcal{F}_{t-1}]\}, \quad 1 \leq t \leq T. \quad (3.20)$$

*Démonstration.* Soit

$$G_t = C_t - E[C_t | \mathcal{F}_{t-1}] - \varphi_t^\top (A_t - E[A_t | \mathcal{F}_{t-1}]), \quad (3.21)$$

où  $C_T = C$  et

$$\beta_{t-1} C_{t-1} = E[\beta_t C_t | \mathcal{F}_{t-1}] - \varphi_t^\top E[\beta_t A_t - \beta_{t-1} A_{t-1} | \mathcal{F}_{t-1}], \quad (3.22)$$

d'où nous obtenons

$$\beta_t G_t = \beta_t C_t - \beta_{t-1} C_{t-1} - \varphi_t^\top (\beta_t A_t - \beta_{t-1} A_{t-1}), \quad t \in 1, \dots, T. \quad (3.23)$$

Remarquons que  $G_t$  est  $\mathcal{F}_t$  mesurable et que  $E[G_t | \mathcal{F}_{t-1}] = 0$ , pour tout  $1 \leq t \leq T$ . Rappelons que l'objectif de l'algorithme est de minimiser l'espérance du carré des erreurs de réplication  $G_T(V_0, \varphi) = \beta_T(C - V_T(V_0, \varphi))$ . En utilisant (3.14) et (3.23),

nous obtenons

$$\begin{aligned}
\sum_{t=1}^T \beta_t G_t &= \sum_{t=1}^T [\beta_t C_t - \beta_{t-1} C_{t-1} - \varphi_t^\top (\beta_t A_t - \beta_{t-1} A_{t-1})] \\
&= \beta_T C - C_0 - \sum_{t=1}^T \varphi_t^\top (\beta_t A_t - \beta_{t-1} A_{t-1}) \\
&= \beta_T C - C_0 + V_0 - \beta_T V_T \\
&= G_T - C_0 + V_0
\end{aligned} \tag{3.24}$$

et  $E[G_T] = E[G_T | \mathcal{F}_0] = C_0 - V_0$  puisque  $E[G_t | \mathcal{F}_{t-1}] = 0$  pour tout  $t \in \{1, \dots, T\}$ . En utilisant les propriétés des espérances conditionnelles, nous avons aussi

$$\begin{aligned}
E[G_T^2] &= E[G_T^2 | \mathcal{F}_0] \\
&= (C_0 - V_0)^2 + \sum_{t=1}^T E[\beta_t^2 G_t^2 | \mathcal{F}_0] \\
&= (C_0 - V_0)^2 + \sum_{t=1}^T E[\beta_t^2 E\{G_t^2 | \mathcal{F}_{t-1}\} | \mathcal{F}_0].
\end{aligned} \tag{3.25}$$

Puisque que  $G_t$  ne dépend de  $\varphi_t, \dots, \varphi_T$  que par  $C_t$ , il suffit de trouver  $\varphi_T$  minimisant  $E[G_T^2 | \mathcal{F}_0]$ , puis de trouver  $\varphi_{T-1}$  minimisant  $E[G_{T-1}^2 | \mathcal{F}_0]$  et ainsi de suite afin de minimiser  $E[G^2]$ . En procédant de la sorte, nous obtenons le minimum puisque chacun des terme est non-négatif. Ayant trouvé la stratégie optimale  $\varphi$ , il s'ensuit que le choix optimal pour  $V_0$  est  $C_0$ .

Bref, remarquons en premier que  $G_T = \eta_T - \xi_T^\top \varphi_T$ , où  $\eta_T = C - E[C | \mathcal{F}_{T-1}] = C_T - E[C_T | \mathcal{F}_{T-1}]$  et  $\xi_T = \Delta_T = A_T - E[A_T | \mathcal{F}_{T-1}]$ . En utilisant la proposition 2 donnée en annexe, il est possible de conclure que

$$\varphi_T = (\Sigma_T)^{-1} E[\xi_T \eta_T | \mathcal{F}_{T-1}] = (\Sigma_T)^{-1} E[\{A_T - E[A_T | \mathcal{F}_{T-1}]\} C_T | \mathcal{F}_{T-1}] \tag{3.26}$$

minimise  $E[G_T^2 | \mathcal{F}_0]$ . Ayant trouvé le  $\varphi_T$  optimal,  $C_{T-1}$  est trouvé par (3.22). Supposons maintenant que  $\varphi_T, \dots, \varphi_t$  ont été définis et  $G_{t-1}$  ainsi que  $C_{t-1}$  sont définis selon (3.21) et (3.22). Alors la proposition 2 est utilisée afin de démontrer que

$\varphi_{t-1}$  donné par (3.17) minimise  $E[G_{t-1}^2 | \mathcal{F}_0]$ . Bref, la mesure de risque  $E[G^2 | \mathcal{F}_0]$  est minimisée en choisissant  $\varphi_t$  fournis par (3.17) et finalement, en utilisant (3.25), la valeur optimale de  $V_0$  est  $C_0$ .  $\square$

La relation entre le prix  $C_0$  donné par le théorème 1 et le prix suggéré par la méthode de la mesure martingale est donnée par le corollaire suivant.

**Corollaire 1** (Papageorgiou, Rémillard et Hocquard (2007, [24])). *Soit*

$$\begin{aligned} U_t &= 1 - \Delta_t^\top (\Sigma_t)^{-1} E \left[ A_t - \frac{\beta_{t-1}}{\beta_t} A_{t-1} \mid \mathcal{F}_{t-1} \right], & t \in \{1, \dots, T\}, \\ M_0 &= 1, \\ M_t &= U_t M_{t-1}. \end{aligned} \tag{3.27}$$

Alors,  $(M_t, \mathcal{F}_t)_{t=0}^T$  est une martingale (pas nécessairement positive) et

$$\beta_{t-1} C_{t-1} = E[\beta_t C_t U_t \mid \mathcal{F}_{t-1}]. \tag{3.28}$$

Aussi,  $\beta_t C_t M_t$  est une martingale et  $C_0 = E[\beta_T C_T M_T \mid \mathcal{F}_0]$ . De plus,  $E[\beta_t A_t U_t \mid \mathcal{F}_{t-1}] = \beta_{t-1} A_{t-1}$  et par conséquent,  $\beta_t A_t M_t$  est une martingale.

*Démonstration.* En utilisant le théorème 1, nous avons

$$\begin{aligned} \beta_{t-1} C_{t-1} &= E[\beta_t C_t \mid \mathcal{F}_{t-1}] - \varphi^\top E[\beta_t A_t - \beta_{t-1} A_{t-1} \mid \mathcal{F}_{t-1}] \\ &= E[\beta_t C_t \mid \mathcal{F}_{t-1}] - E \left[ C_t \Delta_t^\top (\Sigma_t)^{-1} E[\beta_t A_t - \beta_{t-1} A_{t-1} \mid \mathcal{F}_{t-1}] \mid \mathcal{F}_{t-1} \right] \\ &= E[C_t U_t \mid \mathcal{F}_{t-1}], \end{aligned} \tag{3.29}$$

où  $U_t$  est défini par (3.27). En particulier,  $E[U_t \mid \mathcal{F}_{t-1}] = 1$  et par conséquent,  $(M_t)_{t=0}^T$  est une martingale.

Finalement, afin de prouver que  $\beta_t A_t M_t$  est une martingale, notons d'abord que

$$\begin{aligned} \beta_t A_t U_t &= \beta_t A_t - \\ &\quad \{\Delta_t + E[A_t | \mathcal{F}_{t-1}]\} \Delta_t^\top (\Sigma_t)^{-1} E[\beta_t A_t - \beta_{t-1} A_{t-1} | \mathcal{F}_{t-1}]. \end{aligned} \quad (3.30)$$

où  $t \in \{1, \dots, T\}$ . De plus, puisque  $E[\Delta_t | \mathcal{F}_{t-1}] = 0$ ,

$$\begin{aligned} E[\beta_t A_t U_t | \mathcal{F}_{t-1}] &= E[\beta_t A_t | \mathcal{F}_{t-1}] - \\ &\quad E[\Delta_t \Delta_t^\top | \mathcal{F}_{t-1}] (\Sigma_t)^{-1} E[\beta_t A_t - \beta_{t-1} A_{t-1} | \mathcal{F}_{t-1}] \\ &\quad - E[A_t | \mathcal{F}_{t-1}] E[\Delta_t^\top | \mathcal{F}_{t-1}] (\Sigma_t)^{-1} E[\beta_t A_t - \beta_{t-1} A_{t-1} | \mathcal{F}_{t-1}] \\ &= E[\beta_t A_t | \mathcal{F}_{t-1}] - \Sigma_t (\Sigma_t)^{-1} E[\beta_t A_t - \beta_{t-1} A_{t-1} | \mathcal{F}_{t-1}] \\ &= E[\beta_t A_t | \mathcal{F}_{t-1}] - E[\beta_t A_t - \beta_{t-1} A_{t-1} | \mathcal{F}_{t-1}] \\ &= \beta_{t-1} A_{t-1}. \end{aligned} \quad (3.31)$$

Le fait que  $E[\beta_t A_t U_t | \mathcal{F}_{t-1}] = \beta_{t-1} A_{t-1}$  est suffisant pour prouver que  $\beta_t A_t M_t$  est une martingale.  $\square$

Examinons maintenant le cas où le processus  $A_t$  est Markovien et les flux monétaires à échéance ne dépendent que des prix à échéances. Si la loi de  $A_t$  sachant  $\mathcal{F}_{t-1}$  est  $f_{t-1}(A_t | A_{t-1})$ , la combinaison de la propriété Markovienne et du théorème 1 fait en sorte que  $C_t = \gamma_t(A_t)$  et  $\varphi_t = \psi_t(A_{t-1})$ , où

$$\begin{aligned}
L_{1,t}(x) &= E[A_t \mid A_{t-1} = x] = \int u f_{t-1}(u \mid A_{t-1} = x) du, \\
L_{2,t}(x) &= E[A_t A_t^\top \mid A_{t-1} = x] = \int uu^\top f_{t-1}(u \mid A_{t-1} = x) du, \\
\Sigma_t(x) &= L_{2,t}(x) - L_{1,t}(x)L_{1,t}(x)^\top, \\
\psi_t(x) &= \Sigma_t(x)^{-1} E[\{A_t - L_{1,t}(x)\} \gamma_t(A_t) \mid A_{t-1} = x] \\
&= \Sigma_t(x)^{-1} \int (u - L_{1,t}(x)) \gamma_t(u) f_{t-1}(u \mid A_{t-1} = x) du, \\
U_t(x, u) &= 1 - (L_{1,t}(x) - \frac{\beta_{t-1}}{\beta_t} x)^\top \Sigma_t(x)^{-1} (u - L_{1,t}(x)), \\
\gamma_{t-1}(x) &= \frac{\beta_t}{\beta_{t-1}} E[U_t(x, A_t) \gamma_t(A_t) \mid A_{t-1} = x] \\
&= \frac{\beta_t}{\beta_{t-1}} \int U_t(x, u) \gamma_t(u) f_{t-1}(u \mid A_{t-1} = x) du. \tag{3.32}
\end{aligned}$$

Papageorgiou, Rémillard et Hocquard (2007, [24]) utilisent dans le cas Markovien la méthodologie de Del Moral et al. (2006, [9]) basée sur des simulations Monte Carlo ainsi que de l'interpolation linéaire. Le but premier de ce mémoire est d'explorer d'autres algorithmes de réplifications possibles qui seront tous basés sur la méthode de treillis stochastiques.

## Chapitre 4

# Treillis stochastiques

L'algorithme de réplication optimale implique l'estimation de plusieurs espérances conditionnelles. Un treillis stochastique est une approche intéressante ne nécessitant pas de simulations additionnelles à différents points des trajectoires. Nous présenterons éventuellement cette technique dans le contexte qui nous intéresse, c'est-à-dire la réplication d'une option avec deux sous-jacents  $A^{(1)}$  et  $A^{(2)}$ . Par contre, nous développerons en exposant la méthode dans un contexte d'options américaines telle que développée dans Broadie et Glasserman (2004, [6]).

### 4.1 Fondements théoriques

La méthode consiste à générer des vecteurs aléatoires  $X_t = (X_t(1), \dots, X_t(b))^\top$  pour  $t = 1, \dots, T$  et de permettre le passage d'un point à n'importe quel point du pas de temps suivant (figure 4.1). Bref, nous nous servirons de tous les points déjà générés afin d'estimer les espérances conditionnelles au lieu de procéder à d'autres simulations. Premièrement, définissons  $f_t(a_{t+1}|A_t = y)$ ,  $f_t(a_t)$  et  $h_t(x_t)$  comme étant respectivement la densité conditionnelle de  $A_{t+1}$  sachant  $A_t = y$ , la densité marginale de  $A_t$  avec  $A_0$  fixé et la densité à partir de laquelle les vecteurs  $X_t$  sont générés. De plus, il est requis que  $h_t(u) > 0$  si  $f_{t-1}(u|A_{t-1} = y) > 0$  pour un  $y$  donné. Il est important de mentionner que nous faisons la distinction entre  $A_t$  et  $X_t$  puisque  $h_t(x_t)$  n'est pas nécessairement égale à  $f_t(a_t | y)$ . Nous verrons dans la prochaine section les implications du choix de  $h_t(x_t)$ .

Soit  $A_t = (A_t^{(1)}, \dots, A_t^{(d)})^\top$  un processus Markovien défini sur  $\mathbb{R}^d$  avec une valeur de

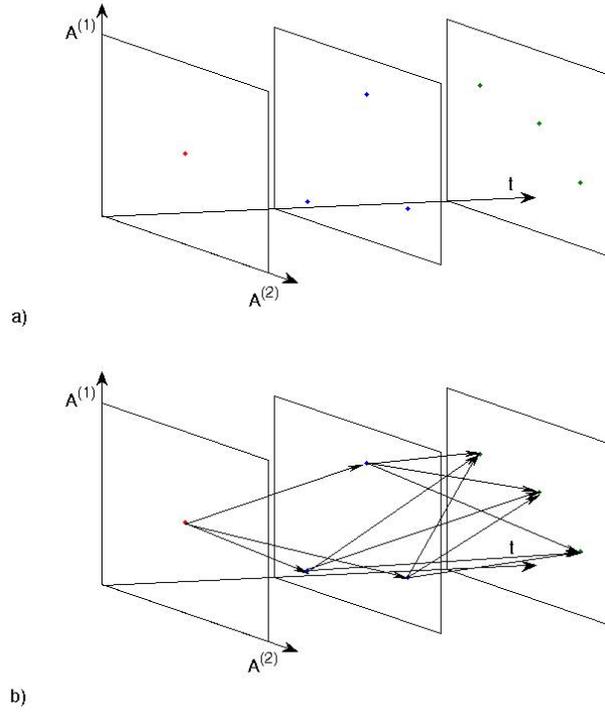


FIG. 4.1 – (a) Génération de vecteur  $X_t$  avec  $b = 3$  et  $T = 2$  (b) treillis stochastique.

départ fixée à  $A_0$  et une discrétisation du temps  $t = 0, 1, \dots, T$ . La tarification d'une option américaine revient à calculer

$$V_0 = \max_{\tau} E[C_{\tau}(A_{\tau})] \quad (4.1)$$

où  $\tau \in \{0, 1, \dots, T\}$  est un temps d'arrêt. De plus,  $A_t$  évolue sous une mesure de probabilité risque neutre et  $C_t$  représente la valeur actualisée au temps 0 du flux monétaire de l'option au temps  $t$ . Après avoir initialisé le vecteur  $\hat{V}_T = C_T(X_T)$ , l'estimateur de la valeur de l'option américaine est calculé de façon récursive par

$$\hat{V}_t(X_t(i)) = \max \left[ C_t(X_t(i)), \frac{1}{b} \sum_{j=1}^b \hat{V}_{t+1}(X_{t+1}(j)) w_t(X_{t+1}(j) | X_t(i)) \right] \quad (4.2)$$

où  $w_t(X_{t+1}(j) | X_t(i))$  est le poids attribué à la transition du point  $X_t(i)$  au point  $X_{t+1}(j)$ . Le biais à la hausse de cet estimateur est démontré par Broadie et Glasser-

man (2004, [6]) en utilisant l'inégalité de Jensen et un argument par induction. De plus, ils énoncent les conditions requises pour prouver la convergence de l'estimateur.

Considérons maintenant l'évaluation de l'espérance conditionnelle d'une fonction  $\psi_t$  de  $A_t$  à un point  $A_{t-1} = y$  donné

$$E[\psi_t(A_t) | A_{t-1} = y] = \int \psi_t(u) f_{t-1}(u | y) du. \quad (4.3)$$

Par contre, nous avons seulement à notre disposition les vecteurs  $X_t$  générés identiquement et indépendamment à partir de  $h_t(x_t)$ . Observons cependant que

$$\begin{aligned} E[\psi_t(A_t) | A_{t-1} = y] &= \int \psi_t(u) f_{t-1}(u | y) du \\ &= \int \psi_t(u) \frac{f_{t-1}(u | y)}{h_t(u)} h_t(u) du \\ &= E \left[ \psi_t(X_t) \frac{f_{t-1}(X_t | y)}{h_t(X_t)} \right] \end{aligned} \quad (4.4)$$

ce qui nous permet de procéder à l'évaluation de  $E[\psi_t(A_t) | A_{t-1} = y]$  pour  $y = X_{t-1}(1), \dots, X_{t-1}(b)$  même si  $X_t$  est généré à partir de  $h_t(x_t)$  et non  $f_{t-1}(x_t | y)$ . L'équation (4.3) est donc estimée par

$$E[\psi_t(\widehat{A_t}) | \widehat{A_{t-1}} = y] = \frac{1}{b} \sum_{j=1}^b \psi_t(X_t(j)) w_{t-1}(X_t(j) | y) \quad (4.5)$$

où  $w_{t-1}(X_t(j) | y) = f_{t-1}(X_t(j) | y) / h_t(X_t(j))$ . La méthode est très flexible quant à la détermination de la densité  $h_t(x_t)$ . Ce choix, qui se révèle crucial, sera traité dans la prochaine section.

Finalement, l'attrait de la méthode de treillis stochastiques repose principalement sur sa capacité à tarifier des options américaines ayant plusieurs sources de risques sous-jacentes. En effet, contrairement à d'autres méthodes telles que l'utilisation d'équations différentielles partielles, le travail engendré par le treillis stochastique ne varie pas exponentiellement avec la dimension du problème confronté. Le coût computationnel requis pour générer les trajectoires est proportionnel à  $b \times T$  alors que l'effort requis pour évaluer les espérances conditionnelles de façon récursive est relatif à  $b^2 \times T$ . Par conséquent, le travail requis par cette méthode est  $O(b^2 T)$ .

## 4.2 Choix de la densité de transition

Un choix intuitif pour  $h_t(x_t)$  serait la densité marginale, c'est-à-dire  $h_t(x_t) = f_t(x_t)$ . Par contre, Broadie et Glasserman (2004, [6]) démontrent que ce choix mène à un estimateur dont la variance croît exponentiellement avec le nombre de pas de temps. Bien entendu, tel qu'il sera discuté dans les prochains chapitres, il n'est pas de notre intérêt de répliquer une distribution de rendement d'une fréquence plus longue que mensuelle. Ainsi, le nombre de pas de temps sera dans notre cas limité (par exemple, 22 jours pour une distribution mensuelle). Cependant, le résultat de Broadie et Glasserman (2004, [6]) pourra nous influencer dans notre choix et une présentation des grandes lignes de la démonstration est pertinente.

Considérons tout d'abord le cas d'une option européenne. Bien entendu, la simplicité d'une simulation Monte Carlo sous une mesure de probabilité neutre au risque est inégalée et une méthode récursive utilisant (4.2) s'avère inutile. Cependant, ce cas nous illuminera dans notre choix de densité de transition. L'estimateur  $\hat{V}_t$  de la valeur d'une option européenne au temps  $t$  est

$$\hat{V}_t(X_t(i)) = \frac{1}{b} \sum_{j=1}^b \hat{V}_{t+1}(X_{t+1}(j)) w_t(X_{t+1}(j) | X_t(i)) \quad (4.6)$$

avec  $\hat{V}_T(X_T(i)) = C_T(X_T(i))$ ,  $i = \{1, \dots, b\}$  les flux monétaires à échéance de l'option. Un développement récursif jusqu'au temps 0 de la valeur de l'option nous permet de l'exprimer comme étant une combinaison linéaire des flux monétaires à échéance, c'est-à-dire

$$\begin{aligned} \hat{V}_0(X_0) &= \frac{1}{b} \sum_{j_T=1}^b \hat{V}_T(X_T(j_T)) \left[ \frac{1}{b^{T-1}} \sum_{j_1, \dots, j_{T-1}}^b \left( \prod_{i=1}^T \frac{f_{i-1}(X_i(j_i) | X_{i-1}(j_{i-1}))}{h_i(X_i(j_i))} \right) \right] \\ &= \frac{1}{b} \sum_{j_T=1}^b \hat{V}_T(X_T(j_T)) L_T(j_T). \end{aligned} \quad (4.7)$$

Selon le choix de la densité  $h_t(X_t)$ , nous nous attendons intuitivement à ce qu'une grande partie de la variance provienne du ratio de vraisemblance  $L_T(j_T)$  multipliant les flux monétaires à échéance plutôt que les flux eux-mêmes. Observons tout d'abord que pour un  $j$  donné

$$E \left[ \frac{f_t(X_{t+1}(j) | y)}{h_{t+1}(X_{t+1}(j))} \right] = \int \frac{f_t(u | y)}{h_{t+1}(u)} h_{t+1}(u) du = \int f_t(u | y) dy = 1 \quad (4.8)$$

et que l'inégalité de Jensen implique que

$$E \left[ \left( \frac{f_t(X_{t+1}(j) | y)}{h_{t+1}(X_{t+1}(j))} \right)^2 \right] > \left( E \left[ \frac{f_t(X_{t+1}(j) | y)}{h_{t+1}(X_{t+1}(j))} \right] \right)^2 = 1 \quad (4.9)$$

sauf si  $f_t(X_{t+1}(j) | y) = h_{t+1}(X_{t+1}(j))$  avec probabilité 1.

Il nous faut donc sélectionner  $h_t(\cdot)$  judicieusement afin de minimiser le deuxième moment du ratio de vraisemblance. Après avoir démontré que la variance de  $L_T(j_T)$  croit exponentiellement avec  $T$ , Broadie et Glasserman (2004, [6]) soulignent qu'il est possible de réduire la variance de  $L_T(j_T)$  à zéro en sélectionnant la *fonction de densité moyenne*, c'est-à-dire en choisissant

$$h_t(u) = f_0(u | A_0) \quad \text{pour } t = 1 \quad (4.10)$$

et

$$h_t(u) = \frac{1}{b} \sum_{j=1}^b f_{t-1}(u | X_{t-1}(j)) \quad \text{pour } t = 2, \dots, T. \quad (4.11)$$

En fait, les ratios de vraisemblance seront tous égaux à 1 dans ce cas. En plus de réduire la variance de  $L_T(j_T)$  à zéro, ce choix a comme particularité de réduire le cas européen considéré à la moyenne des flux monétaires à échéance. En effet,

$$\begin{aligned} \hat{V}_0(X_0) &= \frac{1}{b} \sum_{j_T=1}^b \hat{V}_T(X_T(j_T)) \\ &\times \left[ \frac{1}{b^{T-1}} \sum_{j_1, \dots, j_{T-1}}^b \left( \frac{f_0(X_1(j_1) | X_0(j_0))}{f_0(X_1(j_1) | A_0)} \prod_{i=2}^T \frac{f_{i-1}(X_i(j_i) | X_{i-1}(j_{i-1}))}{\frac{1}{b} \sum_{j=1}^b f_{i-1}(X_i(j_i) | X_{i-1}(j))} \right) \right] \\ &= \frac{1}{b} \sum_{j_T=1}^b \hat{V}_T(X_T(j_T)) \left[ \sum_{j_1, \dots, j_{T-1}}^b \left( \prod_{i=2}^T \frac{f_{i-1}(X_i(j_i) | X_{i-1}(j_{i-1}))}{\frac{1}{b} \sum_{j=1}^b f_{i-1}(X_i(j_i) | X_{i-1}(j))} \right) \right] \\ &= \frac{1}{b} \sum_{j_T=1}^b \hat{V}_T(X_T(j_T)) \left[ \prod_{i=2}^T \left( \frac{\sum_{j_i=1}^b f_{i-1}(X_i(j_i) | X_{i-1}(j_{i-1}))}{\sum_{j=1}^b f_{i-1}(X_i(j_i) | X_{i-1}(j))} \right) \right] \\ &= \frac{1}{b} \sum_{j_T=1}^b \hat{V}_T(X_T(j_T)). \end{aligned} \quad (4.12)$$

De plus, si nous simulons une valeur de  $X_t$  à partir de chacun des  $X_{t-1}(j)$  en utilisant leur densité  $f_{t-1}(X_t | X_{t-1}(j))$  respective et que nous tirons une valeur aléatoirement de  $X_t$ , la valeur obtenue est distribuée selon la *fonction de densité moyenne* conditionnellement à  $X_{t-1}(1), \dots, X_{t-1}(b)$ . En d'autres mots, choisir la *fonction de densité moyenne* revient à simuler des trajectoires comme pour une simulation Monte Carlo traditionnelle et ensuite oublier sur quelle trajectoire chacun des points se trouvait.

Un choix alternatif et beaucoup plus intuitif pour la densité de transition serait de normaliser les densités partant d'une source, c'est-à-dire de sélectionner

$$h_t(u) = f_0(u | A_0) \quad \text{pour } t = 1 \quad (4.13)$$

et

$$\begin{aligned} h_t^{(i)}(u) &= \frac{1}{b} \sum_{j=1}^b f_{t-1}(u | X_{t-1}(j)) \\ &= f_{t-1}(u | X_{t-1}(i)) \quad \text{pour } t = 2, \dots, T \text{ et } i = 1, \dots, b. \end{aligned} \quad (4.14)$$

ce qui produit

$$w_{t-1}(X_t(j) | X_t(i)) = \frac{f_{t-1}(X_t(j) | X_t(i))}{\frac{1}{b} \sum_{k=1}^b f_{t-1}(X_t(k) | X_t(i))} \quad (4.15)$$

et par conséquent,

$$\frac{1}{b} \sum_{j=1}^b w_{t-1}(X_t(j) | X_t(i)) = 1. \quad (4.16)$$

Il est de toute première importance de bien définir les implications de ce choix de densité en particulier. Bien que choisir cette *fonction de densité normalisée* n'offre pas la particularité démontrée par (4.12), cette décision offre néanmoins un avantage computationnel intéressant. En effet, la *fonction de densité moyenne* nous oblige à calculer  $f_{t-1}(X_t(j) | X_{t-1}(i))$  pour  $i = 1, \dots, b$  et  $j = 1, \dots, b$  avant de pouvoir calculer les  $h_t(X_t(j))$  requis pour les calculs d'espérances conditionnelles telles que  $E[A_t | A_{t-1}(i)]$ . Ainsi, cela nous force à conserver en mémoire une matrice  $b \times b$ . En

choisissant la *fonction de densité normalisée*, les calculs requis à chaque  $X_t(i)$  sont indépendants de ceux effectués pour les  $X_t(j)$  où  $j \neq i$ . Bien entendu, la méthode demande toujours un effort computationnel  $O(b^2)$ , mais elle n'impose pas le fardeau de la matrice carrée de dimension  $b$  sur la gestion de la mémoire. De plus, l'avantage de la *fonction de densité moyenne* ne s'applique pas du tout dans notre contexte. En effet, le but étant d'appliquer la méthode de treillis stochastiques en vue d'utiliser l'algorithme de répliation optimale énoncé dans la section 3.3.1, nous ne travaillons pas sous une mesure de probabilité neutre au risque et l'estimateur de la valeur d'une option européenne n'est pas celui donné par (4.12). Bref, la justification donnée par Broadie et Glasserman (2004, [6]) est de moindre importance dans le problème qui nous préoccupe.

Finalement, nous avons vu que la *fonction de densité moyenne* correspond à simuler des trajectoires pour ensuite oublier sur quelle trajectoire les points se trouvent. Dans le cas de la *fonction de densité normalisée*, cette analogie reste valide. En effet, la transformation de l'espérance conditionnelle énoncée en (4.4) servant de justification à l'approche générale de la méthode des treillis stochastique s'applique à une valeur  $y$  donnée, c'est-à-dire à un point  $X_{t-1}(i)$ ,  $i \in \{1, \dots, b\}$  en particulier. Donc, la transformation restera valide tant et aussi longtemps que nous considérerons le vecteur  $X_t$  comme ayant été généré par  $h_t^{(i)}(x_t) = \frac{1}{b} \sum_{j=1}^b f_{t-1}(x_t | X_{t-1}(i)) = f_{t-1}(x_t | X_{t-1}(i))$  lorsque nous évaluons une espérance au point  $X_{t-1}(i)$ .

### 4.3 Approche combinée

Considérons maintenant la combinaison de l'algorithme de *répliation optimale* avec la méthode des treillis stochastiques. Une fois les  $b$  trajectoires simulées et  $C_T$  initialisé à  $\tilde{g}(A_T^{(1)}, A_T^{(2)})$ , il s'agit de calculer de façon récursive pour chaque pas de temps

$$\begin{aligned}
\Delta_t(i) &= A_t - E[A_t \mid A_{t-1}(i)] \\
&= A_t - \sum_{j=1}^b A_t(j) \frac{f_{t-1}(A_t(j) \mid A_{t-1}(i))}{\sum_{k=1}^b f_{t-1}(A_t(k) \mid A_{t-1}(i))}, \\
\Sigma_t(i) &= E[\Delta_t \Delta_t^\top \mid A_{t-1}(i)] \\
&= \sum_{j=1}^b \Delta_t(i) \Delta_t(i)^\top \frac{f_{t-1}(A_t(j) \mid A_{t-1}(i))}{\sum_{k=1}^b f_{t-1}(A_t(k) \mid A_{t-1}(i))}, \\
\varphi_t(i) &= (\Sigma_t(i))^{-1} E[\Delta_t(i) C_t \mid A_{t-1}(i)] \\
&= (\Sigma_t(i))^{-1} \sum_{j=1}^b \Delta_t(i) C_t(j) \frac{f_{t-1}(A_t(j) \mid A_{t-1}(i))}{\sum_{k=1}^b f_{t-1}(A_t(j) \mid A_{t-1}(k))}, \\
C_t(i) &= \frac{\beta_t}{\beta_{t-1}} E[C_t \mid A_{t-1}(i)] - \varphi_t(i)^\top E\left[\frac{\beta_t}{\beta_{t-1}} A_t - A_{t-1}(i) \mid A_{t-1}(i)\right] \\
&= \frac{\beta_t}{\beta_{t-1}} \sum_{j=1}^b C_t(j) \frac{f_{t-1}(A_t(j) \mid A_{t-1}(i))}{\sum_{k=1}^b f_{t-1}(A_t(k) \mid A_{t-1}(i))} \\
&\quad - \varphi_t^\top \sum_{j=1}^b \left\{ \frac{\beta_t}{\beta_{t-1}} A_t(j) - A_{t-1}(i) \right\} \frac{f_{t-1}(A_t(j) \mid A_{t-1}(i))}{\sum_{k=1}^b f_{t-1}(A_t(k) \mid A_{t-1}(i))}, \\
&\quad \forall i \in \{1, \dots, b\}.
\end{aligned} \tag{4.17}$$

La combinaison entre l'algorithme de *réplication optimale* et la méthode de treillis étant établie, nous parlerons dans le prochain chapitre de techniques pour accélérer les calculs requis.

## Chapitre 5

# Méthode multipôle rapide

Tel que mentionné auparavant, le treillis stochastique a amélioré la technique des arbres aléatoires en rendant le temps de calcul linéaire dans le nombre de pas de temps, c'est-à-dire en réduisant la méthode d'une opération  $O(T^2b^2)$  à  $O(Tb^2)$ . Par contre, le temps de calcul demeure toujours un obstacle important lors de l'application des treillis stochastiques. En particulier, le calcul des densités de transition évaluées à chaque point cible et ce pour chaque point source met rapidement en péril la viabilité de la méthode lorsque  $b$  augmente. Broadie et Glasserman (2004, [6]) mentionnent l'importance d'utiliser des variables de contrôle afin de diminuer la grande variance de l'estimation du prix de l'option américaine due en partie au nombre limité de nœuds utilisés. Une approche différente est adoptée dans ce mémoire afin de pallier à ce problème. Broadie et Yamamoto (2002, [8]) utilisent la méthode de transformée gaussienne rapide permettant de réduire le coût computationnel en transformant légèrement le calcul des espérances conditionnelles. Ils démontrent la réduction significative du temps de calcul pour les treillis stochastiques en utilisant un mouvement brownien multidimensionnel pour modéliser les sous-jacents de produits dérivés. Nous développerons donc l'extension nécessaire pour utiliser des mixtures de densités gaussiennes. Par contre, il est impératif d'introduire en premier les méthodes multipôles rapides et d'établir certains fondements théoriques.

Les méthodes multipôles rapides se retrouvent souvent dans un contexte de simulations de particules. Leur avantage premier est de réduire le coût computationnel engendré par le calcul des interactions entre toutes les paires de particules tel que

pour des problèmes de potentiel électrostatique ou gravitationnel. Elles ont aussi une utilité dans la résolution d'équations différentielles partielles. Brièvement, l'évaluation directe d'une certaine interaction entre chacune des paires possibles à l'intérieur d'un ensemble de  $N$  particules requiert un travail  $O(N^2)$ . Certains algorithmes réduisent ce temps à une opération  $O(N^\alpha)$  avec  $1 \leq \alpha < 2$ . Le plus connu de ces algorithmes de sommation rapide est la transformée de Fourier rapide. Cette transformée s'applique lorsque les sources sont disposées sur une grille uniforme et diminue le problème à une opération  $O(N \log(N))$ . Bien que certaines méthodes de transformée de Fourier rapide non-uniforme aient été proposées, les méthodes multipôles rapides ont l'avantage d'être insensibles à la distribution des sources. Cette famille inclut par exemple la méthode multipôle rapide pour l'équation de Laplace, pour l'équation d'Helmholtz ainsi que la transformée gaussienne rapide dont nous traiterons en détails dans les prochaines sections. Beatson et Greengard ([3]) constitue une excellente introduction sur ce sujet. Afin d'en présenter les grandes lignes, considérons la somme suivante

$$G(x_i) = \sum_{j=1}^N q_j f_j(x_i), \quad i \in \{1, \dots, M\}, \quad (5.1)$$

et supposons que les fonction  $f_j$  peuvent être exprimée par

$$f_j(x) = \sum_{\alpha=0}^p \varpi_{\alpha,j}(\bar{x}) \varrho_\alpha(x - \bar{x}) + \epsilon_j(p) \quad (5.2)$$

où  $\varpi_{\alpha,j}$  sont les coefficients de la série,  $\varrho_\alpha$  sont des fonctions de base,  $\bar{x}$  est le centre de l'expansion et  $\epsilon(p)$  est l'erreur causée par la troncature de la série. Ainsi, (5.1) devient

$$\begin{aligned}
G(x_i) &= \sum_{j=1}^N q_j f_j(x_i) \\
&= \sum_{j=1}^N q_j \sum_{\alpha=0}^p \varpi_{\alpha,j}(\bar{x}) \varrho_{\alpha}(x - \bar{x}) + \sum_{j=1}^N q_j \epsilon_j(p) \\
&\cong \sum_{\alpha=0}^p \left[ \sum_{j=1}^N q_j \varpi_{\alpha,j}(\bar{x}) \right] \varrho_{\alpha}(x - \bar{x}) \\
&= \sum_{\alpha=0}^p \Theta_{\alpha} \varrho_{\alpha}(x - \bar{x}), \quad i \in \{1, \dots, M\}.
\end{aligned} \tag{5.3}$$

Ainsi la séparation du travail fait en sorte que le nombre d'opérations requises est  $O((p+1)N) + O((p+1)M)$ . Plusieurs autres expansions ou caractéristiques des méthodes multipôles rapides existent. Cependant, nous inspecterons un cas particulier, c'est-à-dire la transformée gaussienne rapide.

## 5.1 Transformée gaussienne rapide

Plusieurs problèmes en mathématiques appliquées demandent l'évaluation de

$$G(x_i) = \sum_{j=1}^N q_j e^{-\frac{\|x_i - y_j\|^2}{\delta}} \tag{5.4}$$

où  $\|\cdot\|$  désigne la norme  $L^2$ . Il s'agit de la transformation gaussienne discrétisée des coefficients  $\{q_j\}_{j=1}^N$  en fonction des points sources  $\{y_j\}_{j=1}^N$  et des points cibles  $\{x_j\}_{j=1}^M$ . Le calcul de chacun des  $G(x_i)$  demande la somme sur les  $N$  points sources engendrant ainsi un travail  $O(NM)$ . Suivant les principes des méthodes multipôles rapides, Greengard et Strain (1991,[18]) présentent un algorithme où le travail et la mémoire requis est d'ordre  $O(c(N+M))$ . La constante  $c$  dépend de la dimension  $d$  du problème ainsi que de la précision recherchée. Le point de départ de cet algorithme provient de certaines propriétés du noyau gaussien et des expansions d'Hermite que nous présenterons dans la prochaine section. Nous parlerons aussi de l'équation (5.4) comme étant le champ gaussien provenant de la source  $y_j$  de force  $q_j$  évalué à la cible  $x_i$ .

## 5.1.1 Fondements théoriques

Nous débuterons par la fonction génératrice pour les polynômes d'Hermite

$$e^{2xy-y^2} = \sum_{\alpha=0}^{\infty} \frac{y^\alpha}{\alpha!} H_\alpha(x), \quad (5.5)$$

où les polynômes d'Hermite sont définis par  $H_\alpha(x) = (-1)^\alpha e^{x^2} \frac{d^\alpha}{dx^\alpha} e^{-x^2}$ ,  $x \in \mathbb{R}$ . Nous multiplions (5.5) de chaque côté par  $e^{-x^2}$  pour obtenir

$$e^{-(x-y)^2} = \sum_{\alpha=0}^{\infty} \frac{y^\alpha}{\alpha!} h_\alpha(x), \quad (5.6)$$

où les fonctions d'Hermite sont définies par  $h_\alpha(x) = e^{-x^2} H_\alpha(x)$ . Il est possible de réajuster par  $\delta$  et évaluer  $e^{-(x-y)^2}$  en utilisant une expansion d'Hermite centrée à  $\bar{y}$ . Pour  $\delta > 0$  et  $\bar{y} \in \mathbb{R}$ , nous avons

$$\begin{aligned} e^{-\frac{(x-y)^2}{\delta}} &= e^{-\frac{(x-\bar{y}-(y-\bar{y}))^2}{\delta}} \\ &= \sum_{\alpha=0}^{\infty} \frac{1}{\alpha!} \left( \frac{y-\bar{y}}{\sqrt{\delta}} \right)^\alpha h_\alpha \left( \frac{x-\bar{y}}{\sqrt{\delta}} \right). \end{aligned} \quad (5.7)$$

Il est aussi possible d'exprimer  $e^{-\frac{(x-y)^2}{\delta}}$  comme une série de Taylor autour d'un point cible  $\bar{x}$  dans laquelle les coefficients sont les fonctions d'Hermite. Dans ce cas, nous avons

$$e^{-\frac{(x-y)^2}{\delta}} = \sum_{\alpha=0}^{\infty} \frac{1}{\alpha!} \left( \frac{x-\bar{x}}{\sqrt{\delta}} \right)^\alpha h_\alpha \left( \frac{y-\bar{x}}{\sqrt{\delta}} \right). \quad (5.8)$$

Dans le cas multidimensionnel, nous utilisons un indice  $\alpha_i$  pour chacune des  $d$  dimensions et définissons  $\boldsymbol{\alpha}$  somme étant l'ensemble de ces  $d$  indices. Définissons ensuite

$$\begin{aligned}
|\boldsymbol{\alpha}| &= \sum_{i=1}^d \alpha_i, \\
\boldsymbol{\alpha}! &= \prod_{i=1}^d \alpha_i!, \\
x^\boldsymbol{\alpha} &= \prod_{i=1}^d x_i^{\alpha_i},
\end{aligned}
\tag{5.9}$$

et les polynômes et fonctions d'Hermite multidimensionnels se définissent par

$$\begin{aligned}
H_{\boldsymbol{\alpha}}(x) &= \prod_{i=1}^d H_{\alpha_i}(x_i), \\
h_{\boldsymbol{\alpha}}(x) &= \prod_{i=1}^d h_{\alpha_i}(x_i).
\end{aligned}
\tag{5.10}$$

Nous présenterons maintenant trois lemmes énoncés dans Greengard et Strain (1991, [18]) sur lesquels est basé leur algorithme afin de procéder à l'évaluation de (5.4). Afin de mieux comprendre les opérations effectuées, nous aurons aussi besoin de l'expansion de Taylor de la fonction d'Hermite centré à un point  $\bar{x} \in \mathbb{R}$ . Nous avons tout d'abord

$$h_{\boldsymbol{\alpha}} = (-1)^{|\boldsymbol{\alpha}|} D^{\boldsymbol{\alpha}} e^{-x^2}, \tag{5.11}$$

où  $D^{\boldsymbol{\alpha}} = \prod_{i=1}^d \left( \frac{\partial}{\partial x_i} \right)^{\alpha_i}$  ce qui nous donne

$$D^{\boldsymbol{\beta}} h_{\boldsymbol{\alpha}}(x) = (-1)^{|\boldsymbol{\beta}|} h_{\boldsymbol{\alpha}+\boldsymbol{\beta}}(x). \tag{5.12}$$

La série de Taylor de  $h_{\boldsymbol{\alpha}}(x)$  est donc

$$h_{\alpha}(x) = \sum_{\beta=0}^{\infty} \frac{(x - \bar{x})^{\beta}}{\beta!} (-1)^{|\beta|} h_{\alpha+\beta}(\bar{x}). \quad (5.13)$$

**Lemme 1** (Greengard et Strain (1991,[18])). *Soit  $N$  sources dans un hypercube de dimension  $d$ , de largeur  $r\sqrt{2\delta}$  et centré à  $\bar{\mathbf{y}}$ . L'évaluation du champ gaussien provenant des sources  $y_j$  de force  $q_j$  est égale à une unique expansion d'Hermite autour de  $\bar{\mathbf{y}}$*

$$G(x) = \sum_{\alpha=0}^{\infty} \frac{1}{\alpha!} h_{\alpha} \left( \frac{x - \bar{\mathbf{y}}}{\sqrt{\delta}} \right) \sum_{j=1}^N q_j \left( \frac{y_j - \bar{\mathbf{y}}}{\sqrt{\delta}} \right)^{\alpha} \quad (5.14)$$

avec une erreur due à la troncature de la série après  $p^d$  terme satisfaisant la borne suivante

$$\begin{aligned} |\epsilon(p)| &= \left| \sum_{\alpha \geq p} \frac{1}{\alpha!} h_{\alpha} \left( \frac{x - \bar{\mathbf{y}}}{\sqrt{\delta}} \right) \sum_{j=1}^N q_j \left( \frac{y_j - \bar{\mathbf{y}}}{\sqrt{\delta}} \right)^{\alpha} \right| \\ &\leq Q \sum_{i=1}^d \binom{d}{i} \left( K p^{1/4} \frac{r_p^p}{1 - r_p} \right)^i \end{aligned} \quad (5.15)$$

où  $K = 1.09(2\pi)^{-\frac{1}{4}}$ ,  $r_p = r \sqrt{\frac{\epsilon}{p}} < 1$  et  $Q = \sum_{j=1}^N |q_j|$ .

La borne pour l'erreur de troncature donnée par Greengard et Strain (1991, [18]) a initialement été corrigée par Baxter et Roussos (2002, [26]) puis par Wan et Karniadakis (2006, [29]). Cette dernière correspond à celle que nous retrouvons ci-dessus. Cependant, les bornes pour l'erreur de troncature des deux prochains lemmes sont celles de Greengard et Strain (1991,[18]).

Nous pouvons aussi convertir l'expansion d'Hermite autour de  $\bar{\mathbf{y}}$  en une expansion de Taylor autour de  $\bar{\mathbf{x}}$ .

**Lemme 2** (Greengard et Strain (1991,[18])). *L'expansion d'Hermite (5.14) possède l'expansion de Taylor suivante autour du point  $\bar{\mathbf{x}}$*

$$G(x) = \sum_{\beta=0}^{\infty} \frac{(-1)^{|\beta|}}{\beta!} \left( \frac{x - \bar{\mathbf{x}}}{\sqrt{\delta}} \right)^{\beta} \sum_{\alpha=0}^{\infty} \frac{1}{\alpha!} \sum_{j=1}^N q_j \left( \frac{y_j - \bar{\mathbf{y}}}{\sqrt{\delta}} \right)^{\alpha} h_{\alpha+\beta} \left( \frac{\bar{\mathbf{y}} - \bar{\mathbf{x}}}{\sqrt{\delta}} \right) \quad (5.16)$$

avec une erreur due à la troncature de la série après  $p^d$  terme satisfaisant la borne suivante si  $r < 1$

$$\begin{aligned} |\epsilon(p)| &= \left| \sum_{\beta \geq p} \frac{(-1)^{|\beta|}}{\beta!} \left( \frac{x - \bar{x}}{\sqrt{\delta}} \right)^\beta \sum_{\alpha=0}^{\infty} \frac{1}{\alpha!} \sum_{j=1}^N q_j \left( \frac{y_j - \bar{y}}{\sqrt{\delta}} \right)^\alpha h_{\alpha+\beta} \left( \frac{\bar{y} - \bar{x}}{\sqrt{\delta}} \right) \right| \\ &\leq KQ \left( \frac{1}{p!} \right)^{\frac{d}{2}} \left( \frac{r^{p+1}}{1-r} \right)^d \end{aligned} \quad (5.17)$$

où  $K = 1.09^d$ .

Bien entendu, le succès de cette méthode repose sur la capacité à tronquer les séries et à obtenir néanmoins un résultat précis. Pour leur algorithme, Greengard et Strain (1991,[18]) avancent un dernier lemme dans lequel la série d'Hermite est tronquée avant de la convertir en une série de Taylor et de tronquer cette dernière.

**Lemme 3** (Greengard et Strain (1991,[18])). *Une version de (5.14) dans laquelle la somme infinie est tronquée à  $p$  possède l'expansion de Taylor suivante autour du point  $\bar{x}$*

$$G(x) = \sum_{\beta=0}^p \frac{(-1)^{|\beta|}}{\beta!} \left( \frac{x - \bar{x}}{\sqrt{\delta}} \right)^\beta \sum_{\alpha=0}^p \frac{1}{\alpha!} \sum_{j=1}^N q_j \left( \frac{y_j - \bar{y}}{\sqrt{\delta}} \right)^\alpha h_{\alpha+\beta} \left( \frac{\bar{y} - \bar{x}}{\sqrt{\delta}} \right) \quad (5.18)$$

avec une erreur due à la troncature de la série après  $p^d$  terme satisfaisant la borne suivante

$$\begin{aligned} |\epsilon(p)| &= \left| \sum_{\beta \geq p} \frac{(-1)^{|\beta|}}{\beta!} \left( \frac{x - \bar{x}}{\sqrt{\delta}} \right)^\beta \sum_{\alpha=0}^p \frac{1}{\alpha!} \sum_{j=1}^N q_j \left( \frac{y_j - \bar{y}}{\sqrt{\delta}} \right)^\alpha h_{\alpha+\beta} \left( \frac{\bar{y} - \bar{x}}{\sqrt{\delta}} \right) \right| \\ &\leq \bar{K}Q \left( \frac{1}{p!} \right)^{\frac{d}{2}} \left( \frac{r^{p+1}}{1-r} \right)^d \end{aligned} \quad (5.19)$$

où  $\bar{K} \leq K(a + (p!)^{\frac{d}{2}}) \leq 2K$  si  $r \leq \frac{1}{2}$ .

Puisque nous traitons d'un problème à deux dimensions dans ce mémoire, le développement du lemme 3 dans ce contexte est donné en guise d'exemple.

$$G(x) = \sum_{\substack{0 \leq \beta_1 \leq p \\ 0 \leq \beta_2 \leq p \\ 0 \leq \alpha_1 \leq p \\ 0 \leq \alpha_2 \leq p \\ 1 < j < N}} \prod_{i=1}^2 \frac{(-1)^{\beta_i} q_j}{\beta_i! \alpha_i!} \left( \frac{y_j^{(i)} - \bar{y}^{(i)}}{\sqrt{\delta}} \right)^{\alpha_i} h_{\alpha_i + \beta_i} \left( \frac{\bar{y}^{(i)} - \bar{x}^{(i)}}{\sqrt{\delta}} \right) \left( \frac{x^{(i)} - \bar{x}^{(i)}}{\sqrt{\delta}} \right)^{\beta_i}.$$

### 5.1.2 Implémentation de la transformation

L'avantage principal de la transformée gaussienne rapide est que les calculs dépendants des points sources peuvent être accomplis séparément de ceux dépendants des points cibles. Séparons donc (5.18) afin de mieux comprendre les calculs requis. Nous commençons par calculer pour chacun des  $J$  hypercubes de dimension  $d$  contenant les points sources

$$\Theta_{\alpha, J} = \frac{1}{\alpha!} \sum_{y_j \in J} q_j \left( \frac{y_j - \bar{y}_J}{\sqrt{\delta}} \right)^{\alpha}, \quad \forall \alpha_1, \alpha_2 \in \{0, \dots, p\}. \quad (5.20)$$

Cette première étape a un coût computationnel d'ordre  $O((p+1)^d N)$ . Ensuite, nous calculons les fonctions d'Hermite pour chacune des interactions entre les hypercubes sources et les  $I$  hypercubes cibles, c'est-à-dire que nous calculons

$$\Lambda_{\beta, I} = \sum_J \sum_{\alpha} \Theta_{\alpha, J} h_{\alpha + \beta} \left( \frac{\bar{y}_J - \bar{x}_I}{\sqrt{\delta}} \right), \quad \forall \alpha_1, \alpha_2, \beta_1, \beta_2 \in \{0, \dots, p\}. \quad (5.21)$$

Cette étape dépend du nombre de division de l'espace et elle engendre un travail d'ordre  $O((p+1)^{2d} IJ)$ . Finalement, nous obtenons  $G(x_i)$  pour chacun des  $M$  points cibles  $x_i$  appartenant à l'hypercube  $I$  par

$$G(x_i) = \sum_{\beta} \Lambda_{\beta, I} \frac{(-1)^{|\beta|}}{\beta!} \left( \frac{x_i - \bar{x}}{\sqrt{\delta}} \right)^{\beta}, \quad \forall i \in \{1, \dots, M\}, \beta_1, \beta_2 \in \{0, \dots, p\}. \quad (5.22)$$

ce qui requiert un travail d'ordre  $O((p+1)^d M)$ . Par conséquent, le temps requis pour cet algorithme est d'ordre

$$O\left((p+1)^d(N+M)\right) + O\left((p+1)^{2d}IJ\right).$$

## 5.1.3 Exemple

Présentons à titre d'exemple le calcul de l'espérance conditionnelle au temps  $t - 1$  de  $A_t = (A_t^1, A_t^2)^\top$ . L'estimateur par treillis stochastique utilisant la *fonction de densité moyenne* est donné par l'équation (4.5) et le dénominateur de la densité de transition est donnée par (4.10) et (4.11). Rappelons aussi que nous avons déterminé qu'avec ce choix de densité en particulier, la méthode de treillis stochastique était l'équivalent d'une génération de trajectoires aléatoires. Par conséquent, nous avons  $b = N = M$  pour tous les pas de temps. Bref, le calcul de l'espérance conditionnelle est

$$\begin{aligned} E[A_t \mid \widehat{A_{t-1}} = A_{t-1}(i)] &= \sum_{j=1}^N A_t(j) w_{t-1}(A_t(j) \mid A_{t-1}(i)) \\ &= \sum_{j=1}^N A_t(j) \frac{f_{t-1}(A_t(j) \mid A_{t-1}(i))}{\sum_{k=1}^N f_{t-1}(A_t(j) \mid A_{t-1}(k))}. \end{aligned} \quad (5.23)$$

Observons tout d'abord que le calcul de la densité de transition est invariant sous un changement de variables, c'est-à-dire que

$$\begin{aligned} w_{t-1}(A_t(j) \mid A_{t-1}(i)) &= \frac{f_{t-1}(A_t(j) \mid A_{t-1}(i))}{\sum_{k=1}^N f_{t-1}(A_t(j) \mid A_{t-1}(k))} \\ &= \frac{f_{t-1}(y_t(j) \mid y_{t-1}(i)) \left| \frac{\partial y_t}{\partial A_t} \right|_{A_t=A_t(j)}}{\sum_{k=1}^N f_{t-1}(y_t(j) \mid y_{t-1}(k)) \left| \frac{\partial y_t}{\partial A_t} \right|_{A_t=A_t(j)}} \\ &= \frac{f_{t-1}(y_t(j) \mid y_{t-1}(i))}{\sum_{k=1}^N f_{t-1}(y_t(j) \mid y_{t-1}(k))}. \end{aligned} \quad (5.24)$$

Adoptons maintenant pour  $A^{(1)}$  et  $A^{(2)}$  le modèle Black-Scholes-Merton bivarié :

$$\begin{aligned} dA^{(1)} &= \mu_1 A^{(1)} dt + \sigma_1 A^{(1)} dW_1, \\ dA^{(2)} &= \mu_2 A^{(2)} dt + \sigma_2 A^{(2)} dW_2, \\ dW_1 dW_2 &= \rho dt. \end{aligned} \quad (5.25)$$

où  $W_i$  est un processus de Wiener. L'astuce présentée par Broadie et Yamamoto (2002, [8]) est d'effectuer le changement de variable suivant

$$y = L^{-1}\sigma^{-1}x, \quad (5.26)$$

où

$$x^{(i)} = \ln(A_t^{(i)}) - (\mu_i - \frac{1}{2}\sigma_i^2)t,$$

$$\sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix},$$

et  $L$  est la décomposition de Cholesky de la matrice de corrélation. Nous obtenons un vecteur  $y$  de 2 processus de Wiener indépendants satisfaisant

$$dydy^\top = \mathbf{I}dt, \quad (5.27)$$

à partir duquel nous pouvons calculer nos espérances conditionnelles. Par exemple, (5.23) devient

$$\begin{aligned} E[\widehat{A_t | A_{t-1}}(i)] &= \sum_{j=1}^N A_t(j) w_{t-1}(A_t(j) | A_t(i)) \\ &= \sum_{j=1}^N A_t(j) \frac{f_{t-1}(y_t(j) | y_{t-1}(i))}{\sum_{k=1}^N f_{t-1}(y_t(j) | y_{t-1}(k))} \\ &= \sum_{j=1}^N A_t(j) \frac{\prod_{d=1}^2 \frac{1}{\sqrt{2\pi\Delta t}} e^{-\frac{1}{2\Delta t}(y_t^{(d)}(j) - y_{t-1}^{(d)}(i))^2}}{\sum_{k=1}^N \prod_{d=1}^2 \frac{1}{\sqrt{2\pi\Delta t}} e^{-\frac{1}{2\Delta t}(y_t^{(d)}(j) - y_{t-1}^{(d)}(k))^2}} \\ &= \sum_{j=1}^b A_t(j) \frac{\prod_{d=1}^2 e^{-\frac{1}{2\Delta t}(y_t^{(d)}(j) - y_{t-1}^{(d)}(i))^2}}{\sum_{k=1}^N \prod_{d=1}^2 e^{-\frac{1}{2\Delta t}(y_t^{(d)}(j) - y_{t-1}^{(d)}(k))^2}} \\ &= \sum_{j=1}^b A_t(j) \frac{e^{-\frac{\|y_t(j) - y_{t-1}(i)\|^2}{2\Delta t}}}{\sum_{k=1}^N e^{-\frac{\|y_t(j) - y_{t-1}(k)\|^2}{2\Delta t}}}. \end{aligned} \quad (5.28)$$

Remarquons en terminant que le numérateur et le dénominateur de (5.28) ont la forme de (5.4) ce qui permet d'utiliser la transformée gaussienne rapide énoncée dans les sections précédentes.

## 5.2 Extension pour les mixtures de densités gaussiennes

L'extension pour le cas où les rendements de  $A^{(1)}$  et  $A^{(2)}$  sont modélisés par une mixture de gaussiennes bivariées présente à première vue certaines difficultés. En effet, il n'y a pas de changement de variables aussi simple que celui présenté dans l'exemple précédent afin de pouvoir décorréler une mixture de gaussiennes bivariées. Nous arrivons tout de même à reformuler le calcul des espérances conditionnelles afin de pouvoir utiliser la transformée gaussienne rapide. Tout simplement, il s'agit d'appliquer le changement de variables séparément aux prix pour chacun des régimes comme si ceux-ci venaient du régime en question. Plus précisément, le calcul de l'espérance en (5.28) en utilisant cette fois-ci la densité de transition normalisée devient

$$\begin{aligned}
 E[A_t | \widehat{A_{t-1}}(i)] &= \sum_{j=1}^N A_t(j) w_{t-1}(A_t(j) | A_t(i)) \\
 &= \sum_{j=1}^N A_t(j) \frac{\sum_{m=1}^M \nu_m f_{t-1}(y_t^{m(j)} | y_{t-1}^{m(i)}) \left| \frac{\partial y_t^m}{\partial A_t} \right|_{A_t=A_t(j)}}{\sum_{k=1}^N \sum_{m=1}^M \nu_m f_{t-1}(y_t^{m(k)} | y_{t-1}^{m(i)}) \left| \frac{\partial y_t^m}{\partial A_t} \right|_{A_t=A_t(k)}} \\
 &= \sum_{j=1}^N A_t(j) \frac{\sum_{m=1}^M \nu_m \prod_{d=1}^2 \frac{1}{\sqrt{2\pi\Delta t}} e^{-\frac{1}{2\Delta t} (y_t^{(m,d)(j)} - y_{t-1}^{(m,d)(i)})^2} \left| \frac{\partial y_t^m}{\partial A_t} \right|_{A_t=A_t(j)}}{\sum_{k=1}^N \sum_{m=1}^M \nu_m \prod_{d=1}^2 \frac{1}{\sqrt{2\pi\Delta t}} e^{-\frac{1}{2\Delta t} (y_t^{(m,d)(k)} - y_{t-1}^{(m,d)(i)})^2} \left| \frac{\partial y_t^m}{\partial A_t} \right|_{A_t=A_t(k)}} \\
 &= \sum_{j=1}^N A_t(j) \frac{\sum_{m=1}^M \nu_m \prod_{d=1}^2 e^{-\frac{1}{2\Delta t} (y_t^{(m,d)(j)} - y_{t-1}^{(m,d)(i)})^2} \left| \frac{\partial y_t^m}{\partial A_t} \right|_{A_t=A_t(j)}}{\sum_{k=1}^N \sum_{m=1}^M \nu_m \prod_{d=1}^2 e^{-\frac{1}{2\Delta t} (y_t^{(m,d)(k)} - y_{t-1}^{(m,d)(i)})^2} \left| \frac{\partial y_t^m}{\partial A_t} \right|_{A_t=A_t(k)}}
 \end{aligned} \tag{5.29}$$

où  $M$  est le nombre de régime de la mixture et les  $\{\nu_m\}_{m=1}^M$  sont les poids associés. Donc, les  $y_t^m = (y_t^{m,1}, y_t^{m,2})^\top$  sont les variables ayant subies le changement en utilisant les paramètres  $\mu_m$ ,  $\sigma_m$  et  $\rho_m$  respectifs au régime. Bien entendu, aucun de ces changements de variables ne produit un processus de Wiener bidimensionnel indépendant. Intuitivement, puisque la densité d'une mixture de densités gaussiennes est la combinaison linéaire des densités gaussiennes, la procédure revient à calculer chacune des densités requises comme si le processus de prix avait été généré seule-

ment par ce régime et de procéder ensuite à la combinaison linéaire des probabilités obtenues. Finalement, remarquons que les déterminants des Jacobiens que l'on retrouve au numérateur et au dénominateur ne s'annulent pas comme dans l'exemple de la section (5.1.3).

### 5.3 Application à l'algorithme de réplication optimale

Observons maintenant les calculs requis lorsque nous utilisons la transformée gaussienne rapide avec l'algorithme de réplication optimale. Tout d'abord, l'étape définie en (5.20) demande un calcul pour chacune des espérances conditionnelles. Il est donc tout naturel de déterminer en premier lieu le nombre de ces espérances, c'est-à-dire de définir les différents coefficients  $\{q_j\}_{j=1}^N$ . Nous avons

$$\hat{\Sigma}_t(i) = \sum_{j=1}^N (A_t(j) - E[A_t | A_{t-1}(i)]) (A_t(j) - E[A_t | A_{t-1}(i)])^\top w_{t-1}(A_t(j) | A_{t-1}(i)). \quad (5.30)$$

Puisque le but de la transformation est d'éliminer l'enchevêtrement entre les points sources et les points cibles, il est approprié de calculer (5.30) comme

$$\begin{aligned} \hat{\Sigma}_t(i) &= \sum_{j=1}^N A_t(j) A_t(j)^\top w_{t-1}(A_t(j) | A_{t-1}(i)) \\ &\quad - \left\{ \sum_{j=1}^N A_t(j) w_{t-1}(A_t(j) | A_{t-1}(i)) \right\} \left\{ \sum_{j=1}^N A_t(j) w_{t-1}(A_t(j) | A_{t-1}(i)) \right\}^\top \end{aligned} \quad (5.31)$$

ce qui exige l'estimation de  $\frac{d(d+1)}{2} + d$  espérances. Ensuite, nous avons

$$\begin{aligned}
\hat{\varphi}_t(i) &= (\hat{\Sigma}_t)^{-1} \sum_{j=1}^N \{A_t(j) - E[A_t | \widehat{A}_{t-1}(i)]\} C_t(j) w_{t-1}(A_t(j) | A_{t-1}(i)) \\
&= (\hat{\Sigma}_t)^{-1} \sum_{j=1}^N A_t(j) C_t(j) w_{t-1}(A_t(j) | A_{t-1}(i)) \\
&\quad - \left\{ E[A_t | \widehat{A}_{t-1}(i)] \right\} \sum_{j=1}^N C_t(j) w_{t-1}(A_t(j) | A_{t-1}(i))
\end{aligned} \tag{5.32}$$

impliquant l'estimation de  $d + 1$  espérances. Le calcul de  $C_{t-1}$  donné par (3.18) ne requière pas de nouvelles espérances conditionnelles. Cependant, il ne faut pas oublier qu'il faut estimer le dénominateur de  $w_{t-1}(A_t | A_{t-1}(i))$ . Bref, il y a  $\frac{d(d+1)}{2} + d + d + 1 + 1 = \frac{(d+1)(d+4)}{2}$  espérances à calculer et, par conséquent, l'effort computationnel est

$$O\left(\frac{(d+1)(d+4)}{2}(p+1)^d(N+M)\right) + O\left(\frac{(d+1)(d+4)}{2}(p+1)^{2d}IJ\right). \tag{5.33}$$

## 5.4 Transformée gaussienne rapide améliorée

En plus du treillis stochastique et de la combinaison du treillis stochastique avec la transformée gaussienne rapide, un troisième algorithme est présenté dans cette section. Il s'agit aussi d'un treillis stochastique utilisant une transformation mais une approche légèrement différente est adoptée. Ces changements sont inspirés de Yang, Duraiswami, Gumerov et Davis (2003, [30]) qui développent leur technique dans un contexte d'estimateur à noyau de densité utilisé en reconnaissance d'images. Il est à noter que l'épithète *améliorée* utilisée par Yang, Duraiswami, Gumerov et Davis (2003, [30]) pour leur algorithme est conservé dans ce texte mais ne constitue en aucun cas un gage à priori de la supériorité de leur méthode dans notre contexte. Ces auteurs apportent deux principales critiques au niveau de la transformée gaussienne rapide telle que proposée par Greengard et Strain (1991, [18]). Les deux prochaines sections portent sur ces critiques.

### 5.4.1 Algorithme de regroupement

Premièrement, la séparation de l'espace en hypercubes est sous-optimale. À ce sujet, Greengard et Strain (1991, [18]) proposent une heuristique permettant de choisir si l'on procède à une évaluation directe, une transformation partielle ou à la transformation complète telle que stipulée dans le lemme 3 selon le nombre de sources et de cibles dans la paire d'hypercubes considérée. L'idée étant qu'il est laborieux de transformer en expansion lorsqu'il y a peu de points et qu'une évaluation directe dans ce cas est plus rapide. Tel qu'illustré dans la figure 5.2, la deuxième étape de la transformation demande un calcul pour chacune des interactions entre les centres sources et les centres cibles même si peu de points s'y trouvent.

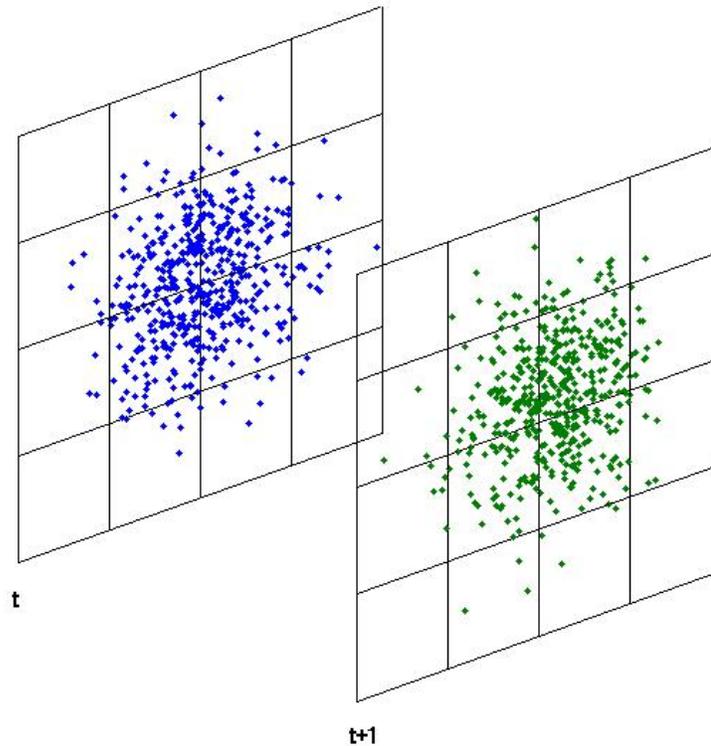


FIG. 5.1 – Séparation uniforme des grilles.

En plus de gaspiller du temps sur des hypercubes vides, cette procédure présente un désavantage énorme : le ratio du volume de l'hypercube sur celui de la sphère de diamètre égal à la largeur de l'hypercube croît exponentiellement avec la dimension

du problème. Donc, lorsque la dimension augmente, la probabilité que les points tombent dans la région hors de la sphère est plus grande. Malheureusement, c'est dans cette région loin du centre de l'hypercube que l'approximation de la transformation est moins bonne. Yang, Duraiswami, Gumerov et Davis (2003, [30]) proposent d'utiliser un algorithme de regroupement du point le plus lointain (*farthest-point clustering*) introduit par Gonzalez (1985, [17]) afin de diviser l'espace de façon plus efficace. Soit un ensemble de points  $\{y_i\}_{i=1}^N$  et un nombre prédéfini  $J$  de régions, l'algorithme commence par choisir un point au hasard parmi les  $N$  points et le rajoute à l'ensemble vide  $C$  de centres. Il s'agit ensuite de calculer la distance pour chacun des  $N - 1$  points restants à l'ensemble  $C$ , c'est-à-dire  $d_i = \min_{c \in C} \|y_i - c\|$ . L'itération se termine en ajoutant à  $C$  le point  $y_i$  satisfaisant  $d_i = \max_{k \in \{1, \dots, N-1\}} d_k$ . Le processus est répété  $J - 2$  fois et chacun des points est assigné à son centre le plus près. Gonzalez (1985, [17]) prouve que cet algorithme produit une partition dont le rayon maximal est au maximum 2 fois celui obtenu lorsque la partition est optimale. Alors que cette technique requiert un temps de calcul  $O(NJ)$ , Feder et Greene (1988, [13]) proposent une méthode engendrant un travail  $O(N \log(J))$ . Dans notre cas, l'algorithme est utilisé non pas avec un nombre prédéfini de régions, mais jusqu'à ce que la distance maximale des points à l'ensemble  $C$  de centres devienne plus petite qu'une valeur voulue. Selon la notation de la section 5.1.1, cette valeur correspond à  $\sqrt{(\frac{1}{2}r\sqrt{2\delta})^2 + (\frac{1}{2}r\sqrt{2\delta})^2} = r\sqrt{\delta}$ .

Illustrons l'amélioration apportée en divisant l'espace par cette méthode. Nous avons généré 50000 points provenant d'une mixture de gaussiennes avec deux régimes ( $\mu_1 = (-3, -3)^\top$ ,  $\mu_2 = (3, 3)^\top$ ,  $\rho_1 = 0$ ,  $\rho_2 = 0.5$ ). La figure (5.2) montre la séparation en 123 zones pour l'algorithme du point le plus lointain et en 210 zones pour la séparation uniforme.

#### 5.4.2 Expansion de Taylor

En second lieu, la troncature à  $p + 1$  termes tel que mentionné dans la section 5.1.2 implique un facteur  $O((p + 1)^d)$  croissant exponentiellement avec la dimension. Développons d'abord l'équation (5.4)

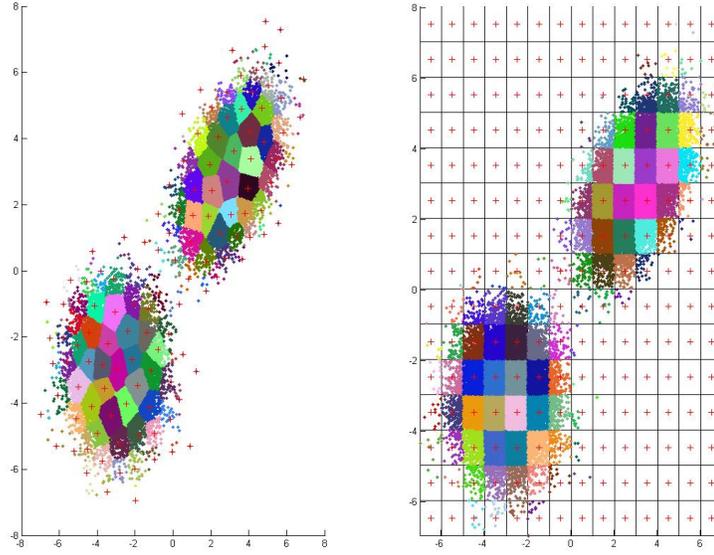


FIG. 5.2 – Séparation par l’algorithme du point le plus lointain et séparation uniforme de 50000 points provenant d’une mixture de gaussienne avec deux régimes ( $\mu_1 = (-3, -3)^\top$ ,  $\mu_2 = (3, 3)^\top$ ,  $\rho_1 = 0$ ,  $\rho_2 = 0.5$ ).

$$\begin{aligned}
 G(x_i) &= \sum_{j=1}^N q_j e^{-\frac{\|x_i - y_j\|^2}{\delta}} \\
 &= \sum_{j=1}^N q_j e^{-\frac{\|(x_i - \bar{y}) - (y_j - \bar{y})\|^2}{\delta}} \\
 &= \sum_{j=1}^N q_j e^{-\frac{\|\Delta_{x_i}\|^2}{\delta}} e^{-\frac{\|\Delta_{y_j}\|^2}{\delta}} e^{\frac{2\Delta_{x_i} \cdot \Delta_{y_j}}{\delta}}
 \end{aligned} \tag{5.34}$$

où  $\Delta_{x_i} = x_i - \bar{y}$  et  $\Delta_{y_j} = y_j - \bar{y}$ . Nous remarquons que les sources et les cibles sont enchevêtrées seulement dans le dernier terme. L’idée avancée par Yang, Duraiswami, Gumerov et Davis (2003, [30]) est de procéder à une expansion de Taylor multivariée pour ce terme en question. Soit  $\Upsilon_n^d$  l’espace des polynômes réels de  $d$  variables avec degré maximal  $n$ . La dimensionnalité de  $\Upsilon_n^d$  est  $C_d^{n+d}$  et il sera efficace de conserver les termes des polynômes dans un vecteur avec un ordre déterminé par le degré total

$|\alpha|$ . Précisément, l'expansion de Taylor multivariée pour le terme en question est

$$e^{\frac{2\Delta x_i \cdot \Delta y_j}{\delta}} = \sum_{\alpha \geq 0} \frac{2^{|\alpha|}}{\alpha!} (\Delta x_i \Delta y_i)^\alpha, \quad (5.35)$$

avec  $\alpha!$ ,  $|\alpha|$  et  $x^\alpha$  définis en (5.9). Les termes du polynôme qui sont conservés dans un vecteur de longueur  $C_d^{n+d}$ , peuvent être calculés de la façon suivante. Supposons que  $d = 2$ , alors à partir du premier élément du vecteur égal à 1, nous obtenons les coefficients avec degré total  $k$  en multipliant ceux avec degré total  $k - 1$  par la première variable pour ensuite ajouter le produit de la deuxième variable par le dernier coefficient avec degré total  $k - 1$  (figure 5.3).

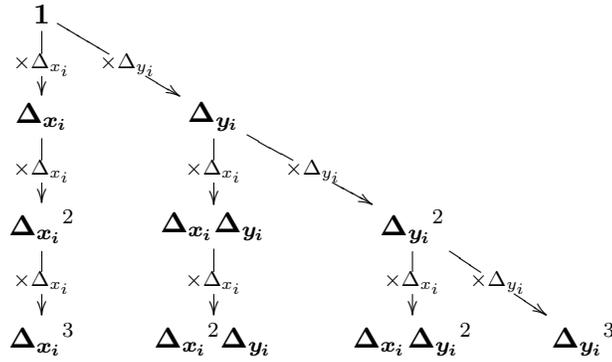


FIG. 5.3 – Calcul du vecteur de coefficients pour l'expansion de Taylor multivariée avec  $d = 2$  et  $n = 3$

Cette opération nécessite  $C_d^{n+d} - 1$  multiplications. Ainsi, l'algorithme pour l'évaluation de  $G(x_i)$  consiste à calculer pour chacune des  $J$  régions

$$\Theta_{\alpha,k} = \frac{2^{|\alpha|}}{\alpha!} \sum_{y_j \in J_k} q_j e^{-\frac{\|y_j - \bar{y}_k\|^2}{\delta}} \left( \frac{y_j - \bar{y}_k}{\sqrt{\delta}} \right)^\alpha, \quad \alpha_1, \alpha_2 \in \{0, 1, \dots, p\}, k \in \{1, \dots, J\}, \quad (5.36)$$

pour ensuite calculer

$$G(x_i) = \sum_{k=1}^J \sum_{\alpha \leq p} \Theta_{\alpha,k} e^{-\frac{\|x_i - \bar{y}_k\|^2}{\delta}} \left( \frac{x_i - \bar{y}_k}{\sqrt{\delta}} \right)^\alpha \quad \forall i \in \{1, \dots, N\}. \quad (5.37)$$

Cet algorithme demande un travail  $O(C_d^{p+d} N)$  pour la première étape (5.36) et

$O(C_d^{p+d}JM)$  pour l'étape (5.37).

## 5.5 Treillis stochastiques améliorés

Tel que décrit dans la section précédente, l'algorithme de la transformée gaussienne rapide amélioré engendre un cout computationnel d'ordre

$$O(NJ) + O\left(\frac{(d+1)(d+4)}{2}C_d^{p+d}N\right) + O\left(\frac{(d+1)(d+4)}{2}C_d^{p+d}JM\right) \quad (5.38)$$

alors que le travail associé au premier algorithme est d'ordre

$$O\left(\frac{(d+1)(d+4)}{2}(p+1)^d(N+M)\right) + O\left(\frac{(d+1)(d+4)}{2}(p+1)^{2d}IJ\right). \quad (5.39)$$

Au lieu d'évaluer directement la valeur de chacune de ces avenues, nous inspecterons la possibilité de procéder à une combinaison plus performante. Tout d'abord, l'avantage de toutes les expansions discutées jusqu'à présent est leur précision acceptable même lorsqu'elles sont tronquées après peu de termes. Malheureusement, nous ne traiterons pas en détails de l'impact sur notre algorithme du choix de  $p$  (nous examinerons tout de même la sensibilité à ce paramètre dans une section ultérieure) et nous utiliserons la même valeur que Brodie et Yamamoto (2002, [8]), c'est-à-dire  $p = 8$ .

Donc, avec  $p = 8$  et  $d = 2$ , notons en premier lieu que  $C_d^{p+d} < (p+1)^d$ . Aussi, les calculs requis pour chacune des interactions entre le hypercubes contenant les points sources et les cibles alourdit énormément la charge computationnelle demandée par l'algorithme de la transformée gaussienne rapide. Par contre, l'algorithme de regroupement utilisé dans la méthode de la transformée gaussienne rapide améliorée est onéreuse malgré sa capacité à partitionner l'espace de façon plus efficace. Il faut mentionner que Yang, Duraiswami, Gumerov et Davis (2003, [30]) soutenaient le mérite de cette méthodologie pour des problèmes à haute dimension ce qui n'est pas le cas ici. Donc, nous utiliserons un hybride entre les deux techniques détaillées dans ce chapitre. Précisément, nous conservons la méthode naïve de Greengard et Strain (1991,[18]) en séparant les données sources en carrés uniformes. Bien que cette

procédure soit sous-optimale, sa rapidité est de loin supérieure à l'algorithme de regroupement de la section 5.4.1. Nous adopterons ensuite la méthode avancée par Yang, Duraiswami, Gumerov et Davis (2003, [30]) en procédant à une expansion de Taylor multivariée. Par contre, nous restreindrons la somme sur les régions sources de l'équation (5.37) à celles se trouvant à une certaine distance maximale du point cible étudié. Bref, après avoir séparé les points sources en carré de largeur  $r\sqrt{2\delta}$  et avoir calculé la matrice  $\Theta$  de dimension  $J$  par  $C_d^{p+d}$  donné par (5.36), il s'agit d'évaluer

$$G(x_i) = \sum_{\|x_i - c_j\| \leq \sqrt{\delta}\lambda_x} \sum_{\alpha \leq p} \Theta_{\alpha, J} e^{-\frac{\|x_i - \bar{y}_j\|^2}{\delta}} \left( \frac{x_i - \bar{y}_j}{\sqrt{\delta}} \right)^\alpha, \quad \forall i \in \{1, \dots, M\}. \quad (5.40)$$

L'erreur produite par l'approximation est donnée par Yang, Duraiswami, Gumerov et Davis (2003, [30]) selon l'expression suivante

$$|\epsilon(p)| = Q \left( \frac{2^p}{p!} (\lambda_x)^p (\lambda_y)^p + e^{-\lambda_x^2} \right) \quad (5.41)$$

où  $Q = \sum_{j=1}^N |q_j|$ , et  $\sqrt{\delta}\lambda_y$  est le rayon maximal des régions contenant les points sources. Puisque nous travaillerons avec des carrés de largeur  $r\sqrt{2\delta}$ , le rayon maximal de la région sera  $r\sqrt{\delta}$  et (5.41) devient

$$|\epsilon(p)| = Q \left( \frac{2^p}{p!} (\lambda_x)^p (r)^p + e^{-\lambda_x^2} \right). \quad (5.42)$$

Il est à noter que le fait de restreindre la somme à un nombre prédéfini de régions avoisinantes avait aussi été traité dans Greengard et Strain (1991, [18]). En effet, ceux-ci mentionnaient qu'une erreur bornée par  $Qe^{-2r^2n^2}$  se produit en se limitant au  $(2\nu + 1)^d$  plus proches régions. Suivant leur méthodologie, nous utiliserons  $\nu = 8$ .

Nous apporterons en terminant une dernière amélioration à l'algorithme proposé rendue possible par la nature même du problème étudié. Précisément, nous avons développé (5.34) en centrant les points sources. Un développement similaire se fait en centrant les points cibles. L'avantage vient du fait que nous avons un algorithme récursif et ainsi, nous pouvons séparer les points cibles, c'est-à-dire les points où les espérances conditionnelles sont évaluées, et réutiliser ce partitionnement pour

les points sources de l'itération suivante. Donc, nous utiliserons aussi une version inversée de l'algorithme présenté jusqu'à présent, soit

$$\begin{aligned}\Theta_{\alpha,k} &= \frac{2^{|\alpha|}}{\alpha!} \sum_{\|y_j - \bar{x}_k\| \leq \sqrt{\delta} \lambda_x} q_j e^{-\frac{\|y_j - \bar{x}_k\|^2}{\delta}} \left( \frac{y_j - \bar{x}_k}{\sqrt{\delta}} \right)^{\alpha}, \\ G(x_i) &= \sum_{k=1}^I \sum_{\alpha \leq p} \Theta_{\alpha,k} e^{-\frac{\|x_i - \bar{x}_k\|^2}{\delta}} \left( \frac{x_i - \bar{x}_k}{\sqrt{\delta}} \right)^{\alpha},\end{aligned}\quad (5.43)$$

pour  $\alpha_1, \alpha_2 \in \{0, 1, \dots, p\}$ ,  $k \in \{1, \dots, I\}$  et  $i \in \{1, \dots, M\}$ . Ainsi, nous procédons seulement à  $\left\lceil \frac{T-1}{2} \right\rceil$  séparations de points. Le prochain chapitre se concentrera sur l'évaluation de la performance des méthodes discutées. Il est de toute première importance de mentionner que l'approximation de notre algorithme implique trois paramètres :  $r$  délimitant la largeur des zones contenant les points,  $p$  servant à tronquer les sommes infinies et  $v$  déterminant le nombre de zones avoisinantes à utiliser. Nous débuterons donc le prochain chapitre en étudiant la sensibilité de nos résultats à ces paramètres.

## Chapitre 6

# Performances

Au cours des deux précédents chapitres, nous avons énoncé deux méthodes générales afin de tarifier et de déterminer la stratégie de couverture d'un produit dérivé. Il est de toute première importance de procéder maintenant à une évaluation attentive de leur performance respective. Particulièrement, une étude de la variance des estimateurs du prix, des stratégies de couverture initiales ainsi que du temps de calcul sera effectuée. Toutefois, il est approprié de débiter par évaluer la sensibilité des estimations des treillis stochastiques améliorés aux paramètres d'approximations. Par la suite, la validité des algorithmes sera mise à l'épreuve en testant les erreurs de réplification.

Tel que mentionné auparavant, nous ne nous préoccupons pas du choix d'actifs  $A^{(1)}$  et  $A^{(2)}$  ainsi que de l'estimation de la loi conjointe de leurs rendements. Nous nous servirons du même cas utilisé par Papageorgiou, Rémillard et Hocquard (2007, [24]) pour démontrer la supériorité de leur algorithme de *réplication optimale* sur une couverture Black-Scholes-Merton traditionnelle. Plus précisément, les rendements journaliers de  $A^{(1)}$  et  $A^{(2)}$  sont modélisés par une mixture de 4 régimes gaussiens alors qu'un seul régime suffit à spécifier la loi mensuelle. Les paramètres se trouvent respectivement dans les tableaux 6.1 et 6.2. Quant à  $A^{(3)}$ , la distribution cible est une loi gaussienne de moyenne nulle et de volatilité annuelle égale à 12%. De plus, le coefficient de corrélation de Pearson avec  $A^{(1)}$  est de 0.3. La figure 6.1 illustre la fonction de flux monétaires à échéance pour l'actif synthétique choisi. Tout au long de ce chapitre, nous supposerons que le taux d'intérêt sans risque est égal à 5% et qu'il y a 22 jours dans le mois.

$\nu_m$	$\mu_m^{(1)}$	$\mu_m^{(2)}$	$\sigma_m^{(1)}$	$\sigma_m^{(2)}$	$\rho_m$
0.0956	0.0016	0.0008	0.0039	0.0016	0.9754
0.4673	0.0000	0.0002	0.0069	0.0032	0.7981
0.0763	-0.0003	-0.0005	0.0115	0.0054	0.6964
0.3607	0.0006	0.0005	0.0037	0.0027	0.4613

TAB. 6.1 – Paramètre des mixtures de gaussiennes avec 4 régimes pour modéliser les rendements journaliers.

$\mu^{(1)}$	$\mu^{(2)}$	$\sigma^{(1)}$	$\sigma^{(2)}$	$\rho$
0.007892797	0.0068086	0.029334999	0.034641016	0.700295314

TAB. 6.2 – Paramètre de la loi gaussienne pour modéliser les rendements mensuels.

## 6.1 Sensibilité aux paramètres du treillis stochastique amélioré

Rappelons que le treillis stochastique amélioré implique trois paramètres d'approximation. Nous avons d'abord le paramètre  $p$  représentant le nombre de termes dans les sommes tronquées. Ensuite, le nombre de zones avoisinantes utilisés dans la sommation (5.40) ou (5.43) est déterminé par  $(2v + 1)^d$ . Finalement, la variable  $r$  est utilisée pour définir la largeur des zones de regroupement des points et la borne de l'erreur due à la troncature des sommes infinies. L'algorithme a été testé pour un échantillon fixe de 5000 trajectoires et, tel que mentionné auparavant, les valeurs de base utilisées sont  $p = 8$ ,  $v = 8$  et  $r = 0.5$ .

Les résultats donnés dans les tableaux 6.3, 6.4 et 6.5 sont éloquentes. En effet, la variation de chacun des paramètres ne semble pas faire une différence notable sur les estimateurs. Naturellement, le temps de calcul varie beaucoup.

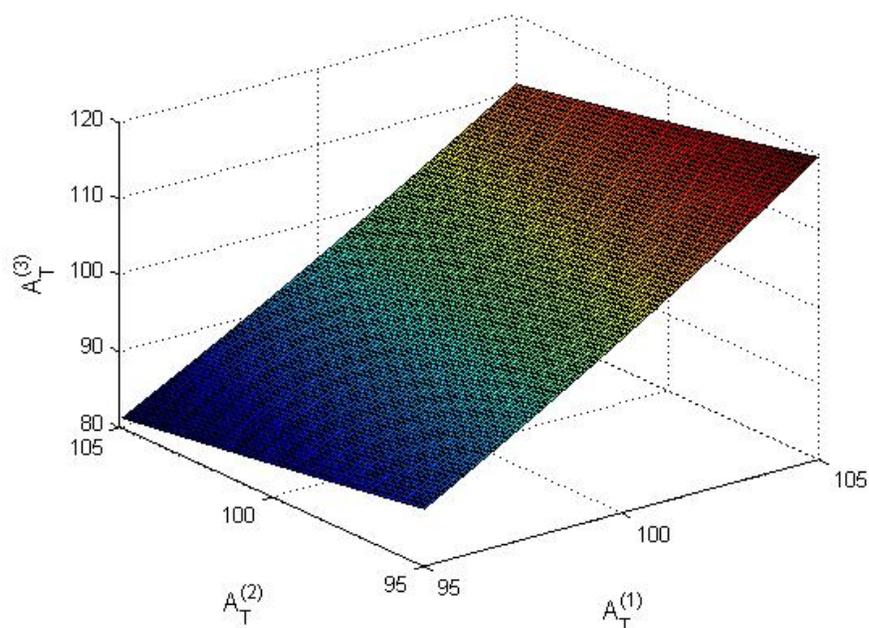


FIG. 6.1 – Fonction de flux monétaires à échéance donnée par  $\tilde{g}$ .

$p$	$V_0$	$\varphi_{A^{(1)}}$	$\varphi_{A^{(2)}}$	temps(secondes)
2	99.0824	-0.7436	3.1315	41.84
4	99.0823	-0.7436	3.1314	58.13
6	99.0823	-0.7436	3.1314	111.53
8	99.0823	-0.7436	3.1314	112.50
10	99.0823	-0.7436	3.1314	148.39
12	99.0823	-0.7436	3.1314	193.42

TAB. 6.3 – Sensibilité au paramètre  $p$  obtenue à partir d'un échantillon de 5000 trajectoires avec  $v = 8$  et  $r = 0.5$ .

$v$	$V_0$	$\varphi_{A^{(1)}}$	$\varphi_{A^{(2)}}$	temps(secondes)
4	99.0823	-0.7436	3.1314	46.68
6	99.0823	-0.7436	3.1314	75.39
8	99.0823	-0.7436	3.1314	111.12
10	99.0823	-0.7436	3.1314	153.58
12	99.0823	-0.7436	3.1314	193.11

TAB. 6.4 – Sensibilité au paramètre  $v$  obtenue à partir d'un échantillon de 5000 trajectoires avec  $p = 8$  et  $r = 0.5$ .

$r$	$V_0$	$\varphi_{A(1)}$	$\varphi_{A(2)}$	temps(secondes)
0.30	99.0823	-0.7436	3.1314	151.22
0.35	99.0823	-0.7436	3.1314	137.75
0.40	99.0823	-0.7436	3.1314	126.19
0.45	99.0823	-0.7436	3.1314	112.24
0.50	99.0823	-0.7436	3.1314	111.23

TAB. 6.5 – Sensibilité au paramètre  $r$  obtenue à partir d’un échantillon de 5000 trajectoires avec  $p = 8$  et  $v = 8$ .

## 6.2 Variance des estimateurs et temps de calcul

Il est maintenant primordial d’inspecter les erreurs commises sur les estimateurs ainsi que le temps de calcul requis selon la taille d’échantillon utilisée. Les résultats du tableau 6.6 concernent 1000 resimulations des deux algorithmes pour 1000, 2000, 3000, 4000 et 5000 trajectoires. Les paramètres discutés dans la section précédente ont été laissés à leur valeur de base. Bien entendu, la rapidité du treillis stochastique pourrait être accrue comparée à celle retrouvée ci-dessous en augmentant notre tolérance à l’erreur de troncature. Remarquons tout d’abord que le temps de calcul du treillis stochastique évolue proportionnellement à  $N^2$ . Aussi, tel que désiré, le temps de calcul du treillis stochastique amélioré croît linéairement avec le nombre de trajectoires. Finalement, le treillis stochastique traditionnel est plus rapide que dans le cas où le nombre de trajectoires est égal à 1000. Bref, il y a un coût fixe minimale à utiliser le treillis stochastique amélioré, mais sa supériorité se démontre rapidement lorsque la taille d’échantillon augmente (figure 6.2).

## 6.3 Évaluation de la stratégie de réplication

Après nous avoir assuré de la précision des estimateurs obtenus par les deux méthodes de treillis stochastiques, il est approprié d’inspecter la capacité de l’algorithme à bien répliquer le produit dérivé en question. Le tableau 6.7 présente les statistiques obtenues en resimulant 250000 trajectoires et en calculant les erreurs de réplication. Les portefeuilles de réplication sont formés par interpolation linéaire à partir des  $V_0$ ,  $\varphi_t^{(1)}$  et  $\varphi_t^{(2)}$  estimés en utilisant un échantillon de 50000 trajectoires.

Premièrement, le tableau 6.7 illustre les statistiques de la fonction  $g$  sur les 250000

	Treillis stochastique		Treillis stochastique amélioré	
	moyenne	écart-type	moyenne	écart-type
<i>N = 1000</i>				
$V_0$	99.1040	0.0004	99.1040	0.0003
$\varphi_{A^{(1)}}$	-0.7443	0.0006	-0.7443	0.0006
$\varphi_{A^{(2)}}$	3.1331	0.0017	3.1331	0.0016
temps (secondes)	11.3967	0.3154	22.7133	0.4948
<i>N = 2000</i>				
$V_0$	99.1040	0.0003	99.1040	0.0003
$\varphi_{A^{(1)}}$	-0.7443	0.0005	-0.7443	0.0005
$\varphi_{A^{(2)}}$	3.1331	0.0013	3.1331	0.0013
temps (secondes)	45.2883	1.0434	46.2410	0.9823
<i>N = 3000</i>				
$V_0$	99.1040	0.0002	99.1040	0.0002
$\varphi_{A^{(1)}}$	-0.7443	0.0004	-0.7443	0.0004
$\varphi_{A^{(2)}}$	3.1331	0.0010	3.1331	0.0010
temps (secondes)	102.4810	1.5856	69.3031	1.2288
<i>N = 4000</i>				
$V_0$	99.1040	0.0002	99.1040	0.0002
$\varphi_{A^{(1)}}$	-0.7443	0.0003	-0.7443	0.0003
$\varphi_{A^{(2)}}$	3.1331	0.0009	3.1331	0.0009
temps (secondes)	186.5285	6.6294	94.0062	1.6622
<i>N = 5000</i>				
$V_0$	99.1040	0.0002	99.1040	0.0002
$\varphi_{A^{(1)}}$	-0.7443	0.0003	-0.7443	0.0003
$\varphi_{A^{(2)}}$	3.1331	0.0008	3.1331	0.0008
temps (secondes)	309.2332	10.9151	118.8672	2.0425

TAB. 6.6 – Moyennes et écart-types des estimateurs obtenus à partir de 1000 résimulations.

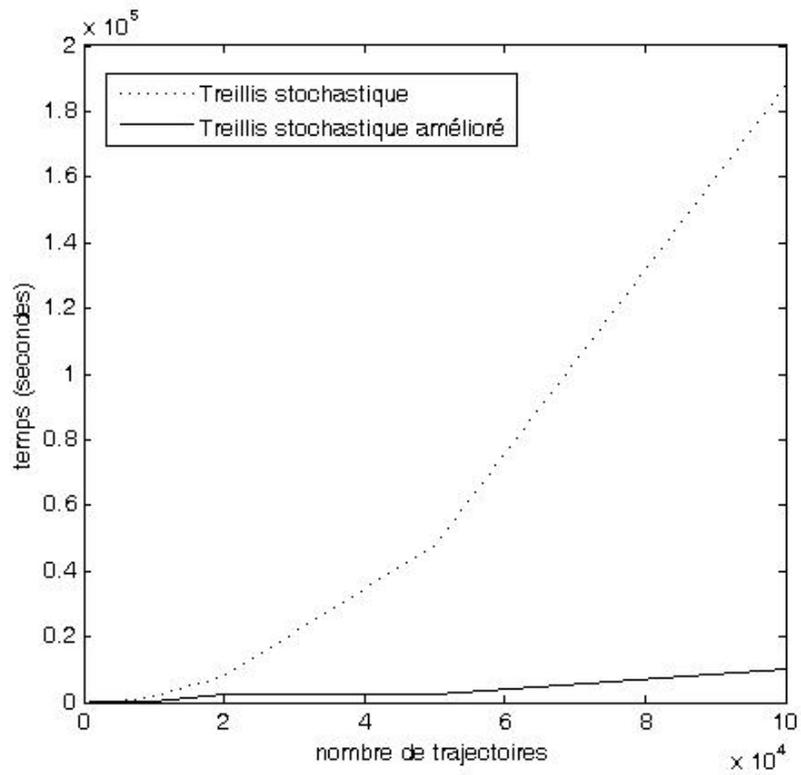


FIG. 6.2 – Temps de calcul selon le nombre de trajectoires.

trajectoires. En d'autres mots, cela représente les statistiques qui seraient obtenues par une répliation parfaite de la fonction  $g$ . En second lieu, remarquons que les statistiques répliquées dans la dernière colonne sont très près de celles ciblées. De plus, la très faible moyenne et l'écart-type des erreurs de répliation confirment la capacité de l'algorithme à bien répliquer l'option.

paramètres	valeurs cibles	$g$	répliation optimale
moyenne	0	7.4446E-05	-6.9886E-05
écart-type	0.03464102	0.03489708	0.03486798
asymétrie	0	-0.01855822	-0.04290073
aplatissement	0	0.02596796	0.02420793
corrélation	0.3	0.30742898	0.30740647
<i>erreurs de répliation</i>			
moyenne			-0.00262611
écart-type			0.01871772

TAB. 6.7 – Résultats de répliation avec 250000 resimulations pour  $g(R_T^{(1)}, R_T^{(2)}) = \log(\frac{C_T}{100})$  et  $\log(\frac{V_T}{100})$ .

## 6.4 Cas non linéaire

Les résultats obtenus jusqu'à présent sont très intéressants. Par contre, le cas étudié fait en sorte que la fonction de flux monétaires à échéance de l'option est linéaire ce qui facilite sa répliation. Par conséquent, nous tenterons maintenant de créer un actif synthétique ayant des propriétés différentes. Nous conserverons une copule normale entre  $A^{(1)}$  et  $A^{(3)}$ , mais nous lui assignerons un coefficient de corrélation égal à -0.5. Plus important, l'actif synthétique suivra maintenant une loi gaussienne tronquée. Brièvement, une variable aléatoire gaussienne de normale  $\mu$  et variance  $\sigma^2$  se trouvant dans l'intervalle  $[a, b]$ ,  $-\infty \leq a < b \leq \infty$ , suivra une loi gaussienne tronquée dont la fonction de densité de probabilité est donnée par

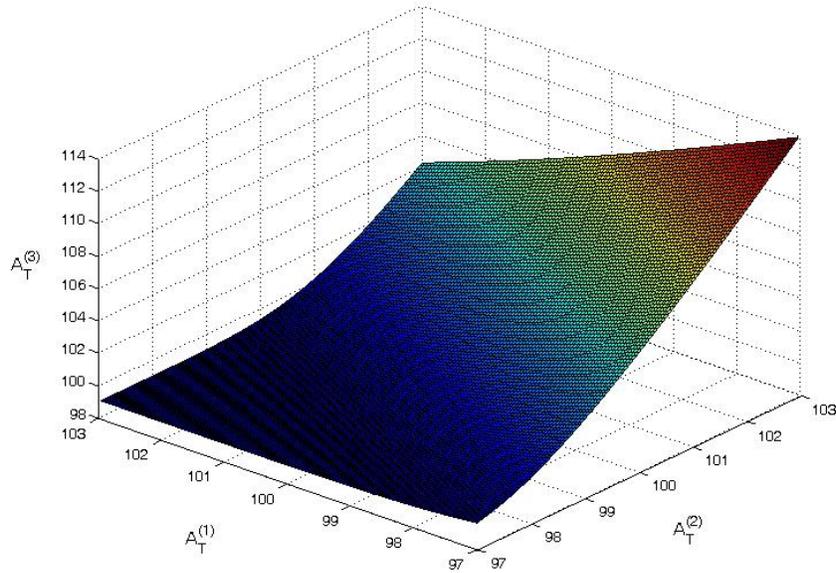
$$f(x; \mu, \sigma^2, a, b) = \frac{\frac{1}{\sigma} \phi(\frac{x-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})} I_{[a,b]}(x), \quad (6.1)$$

où  $\phi(x)$ ,  $\Phi(x)$  et  $I_{[a,b]}(x)$  sont respectivement la fonction de densité de probabilité et la fonction de répartition de la loi gaussienne standardisée ainsi qu'une variable indicatrice égale à 1 si  $x$  est compris dans l'intervalle en question. De plus, la fonction de répartition, l'espérance et la variance sont

$$\begin{aligned}
 F(x; \mu, \sigma^2, a, b) &= \frac{\Phi\left(\frac{x-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} I_{[a,b]}(x), \\
 E[X] &= \mu + \sigma \left[ \frac{\phi\left(\frac{a-\mu}{\sigma}\right) - \phi\left(\frac{b-\mu}{\sigma}\right)}{\left[\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right]} \right], \\
 E[(X - \mu)^2] &= \sigma^2 \left[ 1 + \frac{\left(\frac{a-\mu}{\sigma}\right)\phi\left(\frac{a-\mu}{\sigma}\right) - \left(\frac{b-\mu}{\sigma}\right)\phi\left(\frac{b-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} \right] \\
 &\quad - \sigma^2 \left[ \frac{\phi\left(\frac{a-\mu}{\sigma}\right) - \phi\left(\frac{b-\mu}{\sigma}\right)}{\left[\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right]} \right]^2.
 \end{aligned} \tag{6.2}$$

L'objectif atteint en tronquant à gauche la distribution créée est évident. En effet, nous désirons limiter les rendements à la baisse tout en conservant tout le potentiel de rendements positifs. Par exemple, nous fixons  $a = -0.01$  et  $b = \infty$  ce qui implique que la possibilité d'avoir un rendement inférieur à 1% proviendra seulement d'une erreur de réplcation de l'option. L'exercice de la section précédente est repris et les résultats se trouvent dans le tableau 6.8.

Remarquons d'abord que la moyenne, l'écart-type et la corrélation sont aussi bien réplquées que dans l'exemple précédent. Par contre, les résultats obtenus pour les coefficients d'asymétrie et d'aplatissement sont beaucoup moins précis. Finalement, l'ordre de grandeur de la moyenne et de l'écart-type des erreurs de réplcation est plus grand reflétant la tâche plus ardue de réplcation d'une fonction non-linéaire.

FIG. 6.3 – Fonction de flux monétaires à échéances non-linéaire donnée par  $\tilde{g}$ .

paramètres	valeurs cibles	$g$	réplication optimale
moyenne	0.02160381	0.02163201	0.02056127
écart-type	0.02274285	0.02267803	0.02261321
asymétrie	0.73105921	0.88991404	0.58096182
aplatissement	-0.05870052	0.60986672	0.23901722
corrélation	-0.5	-0.47487841	-0.48205194
<i>erreurs de réplication</i>			
moyenne			-0.02001143
écart-type			0.18054301

TAB. 6.8 – Résultats de réplication avec 250000 resimulations pour  $g(R_T^{(1)}, R_T^{(2)}) = \log(\frac{C_T}{100})$  non linéaire et  $\log(\frac{V_T}{100})$ .

## Chapitre 7

# Conclusion

La méthode introduite par Kat et Palaro (2005, [21]) permet de créer des actifs possédant des caractéristiques prédéterminées, soient une loi marginale ainsi qu'une fonction de dépendance avec un portefeuille référence. Les applications de cette technique en gestion des risques sont vastes. Par contre, plusieurs failles dans l'application nous obligent à y apporter des améliorations. En particulier, une méthode de treillis stochastique peut être implémentée afin de tarifier la stratégie de réplication par l'algorithme récursif proposé par Papageorgiou, Rémillard et Hocquard (2007, [24]).

Après avoir introduit en détail le concept d'actif synthétique, nous avons inspecté les considérations théoriques de cette approche. Par la suite, nous avons tenté d'améliorer la méthode de treillis stochastique utilisée en utilisant la transformée gaussienne rapide. Nous avons démontré la rapidité accrue de la méthode de treillis stochastique amélioré proposée en plus d'examiner la précision des estimateurs ainsi que les statistiques des erreurs de réplication.

La méthode de treillis stochastique amélioré s'applique à la tarification et réplication d'options autres que celles étudiées dans ce mémoire. En particulier, la technique de treillis stochastique amélioré serait adéquate pour la tarification et réplication de produits dérivés où la dimension du problème est supérieure à 2. Il serait donc intéressant d'évaluer les performances de la méthodologie appliquée à d'autres situations. De plus, la modélisation par mixture de gaussiennes utilisée jusqu'à présent, bien qu'elle soutient la cohérence entre les lois journalières et mensuelles, assume que les rendements des actifs ne démontrent pas de dépendance sérielle. Ce problème

soulève le besoin d'étudier une modélisation pouvant incorporer la dépendance sérielle possible des rendements. Finalement, il pourrait être préférable ou requis dans certains contextes de créer un actif synthétique sur une base autre que mensuelle. Dans ce cas, une avenue de recherche intéressante serait d'étudier l'impact des frais de transactions ainsi que de la fréquence de rebalancements sur la validité de la méthode.

# Annexe

Voici deux propositions énoncées dans Papageorgiou, Rémillard et Hocquard (2007, [24]) afin de procéder à la démonstration du théorème 1. Dans ce qui suit,  $L^2 = L^2(\Omega, \mathcal{F}, \mathcal{P})$  est l'ensemble des toutes les variables aléatoires de carré intégrables définies sur  $(\Omega, \mathcal{F})$ .

**Proposition 1.** *Soit  $X$  une variable aléatoire non-négative définie sur  $(\Omega, \mathcal{F}, \mathcal{P})$  satisfaisant  $E[X] < \infty$ . Si  $\mathcal{G}$  est une tribu de  $\mathcal{F}$  et  $E[X | \mathcal{G}] \geq 0$ ,  $\mathcal{P}$ -presque sûrement, alors, pour une variable aléatoire non-négative  $\mathcal{G}$ -mesurable  $\xi$ , nous avons l'égalité suivante*

$$E[\xi X] = E[\xi E[X | \mathcal{G}]]. \quad (1)$$

*Démonstration.* Dans le cas où  $\xi$  est une variable aléatoire bornée, le résultat se déduit de la définition même d'une espérance conditionnelle. En particulier, ceci est vrai pour  $\xi_n = \min(x, \xi) \geq 0$ , pour tout  $n \geq 1$ . Puisque  $\xi_n \uparrow \xi$ , nous obtenons à partir du théorème Beppo-Lévy

$$E[\xi X] = \lim_{n \rightarrow \infty} E[\xi_n X] = \lim_{n \rightarrow \infty} E[\xi_n E[X | \mathcal{G}]] = E[\xi E[X | \mathcal{G}]]. \quad (2)$$

□

**Proposition 2.** *Soit  $\xi \in \mathbb{R}^d$  et  $\eta \in \mathbb{R}$  des variables aléatoires  $L^2$  définies sur  $(\Omega, \mathcal{F})$  et  $A = E[\xi \xi^\top | \mathcal{G}]$  où  $A$  est inversible et  $\mathcal{G}$  est une tribu de  $\mathcal{F}$ . Alors  $\varphi \in \mathbb{R}^d$  minimise  $E\{(\varphi^\top \xi - \eta)^2\}$  parmi tous les  $\varphi \in \mathcal{G}$  tel que  $\varphi^\top \xi \in L^2$  si et seulement si  $\varphi = A^{-1}E[\xi \eta | \mathcal{G}]$ . De plus,  $\varphi^\top \xi$  est de carré intégrable.*

*Démonstration.* Soit  $\varphi = A^{-1}b$  où  $b = E[\xi \eta | \mathcal{G}]$ . Nous utilisons la proposition 1 pour démontrer que  $\varphi^\top \xi \in L^2$ . Précisément,

$$\begin{aligned}
E[(\varphi^\top \xi)^2] &= \sum_{i=1}^d E[\varphi_i^2 \xi_i^2] \\
&= \sum_{i=1}^d E[\varphi_i^2 E[\xi_i^2 \mid \mathcal{G}]] \\
&= \sum_{i=1}^d E[\varphi_i^2 A_{ii}] \\
&= E[b^\top A^{-1} b].
\end{aligned} \tag{3}$$

Puisque  $A$  est symétrique et définie positive, il existe une matrice  $d \times d$   $M \in \mathcal{G}$  telle que  $M^{-1} = M^\top$  et une matrice diagonale  $d \times d$   $\Delta \in \mathcal{G}$  telle que  $A = M\Delta M^\top$ . Soit  $\tilde{\xi} = M^\top \xi$  et  $\tilde{b} = M^\top b$ . Donc,  $\Delta = E[\tilde{\xi} \tilde{\xi}^\top \mid \mathcal{G}]$ ,  $\tilde{b} = E[\tilde{\xi} \eta \mid \mathcal{G}]$ ,  $E[\tilde{\xi}_i^2 \mid \mathcal{G}] = \Delta_{ii} > 0$  par hypothèse, et

$$\begin{aligned}
b^\top A^{-1} b &= \tilde{b}^\top \Delta^{-1} \tilde{b} \\
&= \sum_{i=1}^d \frac{E^2[\tilde{\xi}_i \eta \mid \mathcal{G}]}{E[\tilde{\xi}_i^2 \mid \mathcal{G}]} \\
&\leq dE[\eta^2 \mid \mathcal{G}], p.s.
\end{aligned} \tag{4}$$

en utilisant l'inégalité de Cauchy-Schwarz. Subséquemment,

$$E[(\varphi^\top \xi)^2] \leq dE[\eta^2] < \infty. \tag{5}$$

Finalement, soit  $\psi \in \mathcal{G}$  un vecteur aléatoire tel que  $\psi^\top \xi \in L^2$ . Alors

$$E[(\psi^\top \xi - \eta)^2] = E[E[(\psi^\top \xi - \eta)^2 \mid \mathcal{G}]] \tag{6}$$

et nous pouvons voir que

$$\begin{aligned} E[(\psi^\top \xi - \eta)^2 \mid \mathcal{G}] &= \psi^\top A \psi - 2\psi^\top b + c \\ &= (\psi - \varphi)^\top A (\psi - \varphi) + \varphi^\top A \varphi - 2\varphi^\top b + c \\ &= (\psi - \varphi)^\top A (\psi - \varphi) + E[(\varphi^\top \xi - \eta)^2 \mid \mathcal{G}] \\ &\geq E[(\varphi^\top \xi - \eta)^2 \mid \mathcal{G}] \end{aligned} \tag{7}$$

correspondant à ce qu'il fallait démontrer.

□

# Bibliographie

- [1] G. Amin and H.M. Kat. Hedge fund performance 1990-2000 : Do the “money machines” really add value. *Journal of Financial and Quantitative Analysis*, 38(2) :251–275, 2003.
- [2] A.N. Avramidis and H. Matzinger. Convergence of the stochastic mesh estimator for pricing American options. *Proceedings of the Winter Simulation Conference*, pages 1560–1567, 2002.
- [3] R. Beatson and L. Greengard. A short course on fast multipole methods.
- [4] P. Boyle, A. Kolkiewicz, and K.-S. Tan. Pricing American derivatives using simulation : a biased low approach. *Monte Carlo and Quasi-Monte Carlo Methods 2000*, pages 181–200, 2000.
- [5] M. Broadie and P. Glasserman. Pricing American-style securities using simulation. *Journal of Economic Dynamics and Control*, 21 :1323–1352, 1997.
- [6] M. Broadie and P. Glasserman. A stochastic mesh method for pricing high-dimensional American options. *Journal of Computational Finance*, 7 :35–72, 2004.
- [7] M. Broadie, P. Glasserman, and Z. Ha. Pricing American options by simulation using a stochastic mesh with optimized weights. *Probabilistic Constrained Optimization : Methodology and Applications*, pages 32–50, 2000.
- [8] M. Broadie and Y. Yamamoto. Application of the fast gauss transform to option pricing. *Management Science*, 49(8) :1071–1088, 2003.
- [9] P. Del Moral, B. Rémillard, and S. Rubenthaler. Monte Carlo approximations of American options. Technical report, GERAD, 2006.
- [10] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B*, 39 :1–38, 1977.

- 
- [11] P. Dybvig. Distributional Analysis of Portfolio Choice. *Journal of Business*, 61 :369–393, 1988a.
- [12] P. Dybvig. Inefficient Dynamic Portfolio Strategies or How to Throw Away a Million Dollars in the Stock Market. *Review of Financial Studies*, 1 :67–88, 1988b.
- [13] T. Feder and D. Greene. Optimal algorithms for approximate clustering. In *STOC '88 : Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 434–444, New York, NY, USA, 1988. ACM.
- [14] C. Genest, B. Rémillard, and D. Beaudoin. Omnibus goodness-of-fit tests for copulas : A review and a power study. *Insurance Math. Econom.*, 40 :in press, 2007.
- [15] P. Glasserman. *Monte Carlo methods in financial engineering*, volume 53 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 2004. Stochastic Modelling and Applied Probability.
- [16] P. Glasserman and B. Yu. Large sample properties of weighted Monte Carlo estimators. *Operations Research*, 53 :298–312, 2005.
- [17] T. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 38 :293–306, 1985.
- [18] L. Greengard and J. Strain. The fast gauss transform. *Journal on Scientific and Statistical Computing*, 12 :79–94, January 1991.
- [19] H.M. Kat. Synthetic funds and the Mongolian Barbeque. Technical report, Cass Business School, City University, 2006.
- [20] H.M. Kat and H.P. Palaro. Hedge fund returns : you can make them yourself! Technical report, Cass Business School, City University, 2005.
- [21] H.M. Kat and H.P. Palaro. Who needs hedge funds ? A copula-based approach to hedge fund return replication. Technical report, Cass Business School, City University, 2005.
- [22] H.M. Kat and H.P. Palaro. Tell me what you want, what you really really want ! An exercise in tailor made synthetic fund creation. Technical report, Cass Business School, City University, 2006.
- [23] F.A. Longstaff and E.S. Schwartz. Valuing American options by simulation : A simple least-square approach. *The Review of Financial Studies*, 14 :113–147, 2001.

- 
- [24] N. Papageorgiou, B. Rémillard, and A. Hocquard. Replicating the properties of hedge fund returns. Technical Report 1, DGAM-HEC Alternative Investments Research, 2007.
- [25] A. Patton. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2) :527–556, May 2006.
- [26] B.J.C. Roussos and G. Baxter. A new error estimate for the fast gauss transform. *SIAM Journal of Scientific Computing*, 24(1) :257–259, 2002.
- [27] M. Schweiser. Variance-Optimal Hedging in Discrete Time. *Mathematics of Operations Research*, 20(1) :10–32, 1995.
- [28] J. 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 Transactions on Automatic Control*, 44 :1840–51, 1999.
- [29] X. Wan and G. Em Karniadakis. A sharp error estimate for the fast gauss transform. *J. Comput. Phys.*, 219(1) :7–12, 2006.
- [30] C. Yang, R. Duraiswami, N. Gumerov, and L. Davis. Improved fast gauss transform and efficient kernel density estimation, 2003.