

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ESTIMATION DE MODÈLES DE RISQUE DE CRÉDIT AVEC DES
FILTRES À PARTICULES

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN MATHÉMATIQUES

PAR
FABIO ORLANDO GUACANEME CASTIBLANCO

JUIN 2016

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.07-2011). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je voudrais remercier toutes les personnes qui m'ont soutenu au cours de mes études et qui ont contribué de près ou de loin à ce projet.

D'abord, un gros merci au professeur Diego Amaya. Ce document aurait été impossible sans sa disponibilité et ses précieux conseils.

Je remercie le professeur Mathieu Boudreault pour son accompagnement pendant tout mon parcours.

Ma gratitude et mes pensées vont tout naturellement vers mes parents Carlos et Mercedes et mes frères Andrés, Diego et Miguel qui m'ont supporté depuis toujours.

À Milena, qui m'a accompagné dans cette traversée durant toutes ces années, je ne te remercierais jamais assez, car tu as été toujours présent dans les moments de joie et de peine.

À mon frère Luis Alejandro qui me regarde depuis le ciel.

TABLE DES MATIÈRES

LISTE DES TABLEAUX	ix
LISTE DES FIGURES	xi
RÉSUMÉ	xiii
INTRODUCTION	1
CHAPITRE I	
REVUE DE LA LITTÉRATURE	5
1.1 Modèles structurels	5
1.2 Modèles à forme réduite	8
1.3 Modèles hybrides	10
1.4 Estimation	12
1.5 Filtrage	13
CHAPITRE II	
MODÈLES DE RISQUE DE CRÉDIT	17
2.1 Modèles structurels	17
2.1.1 Modèle de Merton	18
2.1.2 Modèle de risque de crédit à taux d'intérêt stochastique	20
2.2 Modèles à forme réduite	22
2.2.1 Changement de la filtration	22
2.2.2 Processus de Cox	23
2.2.3 Obligations corporatives risquées zéro-coupon	25
2.2.4 Obligations corporatives risquées à coupons	26
2.2.5 Le swap sur défaillance de crédit (CDS)	26
2.2.6 Modèles d'intensité stochastique - CIR	28
2.2.7 Modèles d'intensité stochastique avec sauts - JCIR	29

CHAPITRE III	
REPRÉSENTATION DES MODÈLES D'ÉTAT-ESPACE	31
3.1 Représentation d'état-espace du modèle de Merton	32
3.2 Représentation d'état-espace du modèle à taux d'intérêt stochastique	33
3.3 Représentation d'état-espace pour les modèles à forme réduite	35
3.4 Représentation d'état-espace à multiples sources d'information	37
CHAPITRE IV	
FILTRAGE	39
4.1 Filtre à particules	42
4.1.1 Échantillonnage préférentiel séquentiel	43
4.1.2 Le problème de dégénérescence	48
4.1.3 Échantillonnage préférentiel séquentiel avec rééchantillonnage	48
4.2 Lissage	50
4.3 Filtrage proposée pour multiples sources d'information	53
CHAPITRE V	
ESTIMATION DE PARAMÈTRES DE MODÈLES DE RISQUE DE CRÉ- DIT	57
5.1 Estimation de paramètres par maximum de vraisemblance	57
5.2 Estimation de paramètres par espérance-maximisation	58
CHAPITRE VI	
ANALYSE PAR SIMULATION	63
6.1 Analyse du filtrage	64
6.1.1 Paramètres du modèle structurel	64
6.1.2 Paramètres du modèle à forme réduite	65
6.1.3 Paramètres de la méthode de filtrage proposée	66
6.1.4 Mesures de la performance du filtrage	68
6.1.5 Performance du filtrage des modèles structurels	69
6.1.6 Performance du filtrage de modèle à forme réduite	73
6.2 Analyse d'estimation de paramètres par simulation	81

6.2.1	Surface de la fonction de log-vraisemblance et temps de calcul	81
CHAPITRE VII		
ANALYSE EMPIRIQUE 89		
7.1	Dow Jones 2003 89	89
7.1.1	Estimation des paramètres du modèle de Merton 90	90
7.1.2	Estimation des paramètres du modèle à taux d'intérêt stochastique 92	92
CONCLUSION 97		
APPENDICE A		
ÉCHANTILLONNAGE PRÉFÉRENTIEL 99		
APPENDICE B		
DÉVELOPPEMENT DE L'ESPÉRANCE CONDITIONNELLE : ESPÉRANCE-MAXIMISATION 101		
APPENDICE C		
PERFORMANCE DU FILTRAGE : REQM 103		
RÉFÉRENCES 109		

LISTE DES TABLEAUX

Tableau	Page
6.1 Scénarios des paramètres des intensités CIR	66
6.2 Scénarios des paramètres des intensités avec sauts - JCIR	66
6.3 Pondération assignée aux obligations risquées à coupons pour la méthode de filtrage proposée	67
6.4 Pondération assignée aux CDS pour la méthode de filtrage proposée	67
6.5 Moyennes des REQM par maturité, bruit transactionnel et volatilité de la valeur des actifs du modèle de Merton	70
6.6 Moyennes des REQM par maturité, bruit transactionnel et volatilité du taux d'intérêt du modèle à taux d'intérêt stochastique	72
6.7 Gain en pourcentage des 4 obligations conjointement par scénario et bruit transactionnel : Intensités sans sauts.	75
6.8 Gain en pourcentage des 4 obligations conjointement par scénario et bruit transactionnel : Intensités avec sauts.	77
6.9 Gain en pourcentage d'utiliser les six CDS conjointement par scénario et bruit transactionnel : Intensités sans sauts.	79
6.10 Gain en pourcentage d'utiliser les six CDS conjointement par scénario et bruit transactionnel : Intensités avec des sauts.	80
6.11 Rugosité de la fonction de log-vraisemblance en fonction du nombre de particules	84
6.12 Temps de calcul pour estimer les paramètres par les méthodes MV et EM en fonction du nombre de particules	87
7.1 Estimations des paramètres du modèle de Merton	94
7.2 Estimations des paramètres du modèle à taux d'intérêt stochastique	95
7.3 Rapport de vraisemblance (RV) et EAMP	96

C.1	Comparaison des obligations risquées à coupons sans sauts dans les intensités	103
C.2	Comparaison des obligations risquées à coupons avec sauts dans les intensités	104
C.3	Comparaison des CDS sans sauts dans les intensités	105
C.4	Comparaison des CDS avec sauts dans les intensités	106
C.5	Comparaison des filtrages standard (Sd) et proposée (Pr) prenant de multiples sources d'information sans sauts dans les intensités .	107
C.6	Comparaison des filtrages standard (Sd) et proposée (Pr) prenant de multiples sources d'information avec sauts dans les intensités .	108

LISTE DES FIGURES

Figure		Page
4.1	Estimation de la densité des particules	53
4.2	Maximun de vraisemblance et poids d'importance de particules . .	55
4.3	Nouvelle estimation de la densité à particule	56
6.1	Boxplot pour les REQM pour le scénario 1 CIR Obligations . . .	76
6.2	Boxplot pour les REQM pour le scénario 1 CIR Obligations avec sauts	77
6.3	Boxplot pour les REQM pour le scénario 1 CIR CDS	79
6.4	Boxplot pour les REQM pour le scénario 1 CIR CDS avec sauts .	80
6.5	Surfaces de log-vraisemblance	83
6.6	Contour de la vraisemblance du modèle Vasicek	85
6.7	Itérations de la méthode (EM) et surface de la fonction $Q(\Theta_k, \Theta_k)$	87
7.1	Points de départ et des estimations du modèle de Merton	91

RÉSUMÉ

Ce document porte sur l'estimation de paramètres et de variables latentes qui sont implicites dans les modèles de risque de crédit. Deux catégories des modèles sont considérés : Les modèles structurels et les modèles à forme réduite.

Sous la première catégorie, la valeur marchande des actifs est la principale mesure de solvabilité. Bien qu'elle n'est pas directement observée, elle peut être estimée grâce à sa relation avec la valeur de fonds propres. De même, sous la deuxième catégorie, les intensités de défaut qui ne sont pas observées peuvent être estimées en fonction des prix des instruments financiers susceptibles de défaut.

Cependant, tant la valeur de fonds propres que le prix des instruments financiers observés dans le marché ne reflètent pas ni la variation de la solvabilité de la firme ni l'intensité de défaut en raison de la présence de bruits transactionnels. Il est donc nécessaire de représenter ce bruit transactionnel et de filtrer cette quantité.

C'est à ce moment que les techniques de filtrage deviennent un outil important pour estimer ces variables latentes et les paramètres des modèles de risque de crédit en présence de bruit transactionnel.

Mots clés : Risque de défaut, techniques de filtrage, filtre à particules, évaluation d'options, processus de diffusion, méthodes de simulation.

INTRODUCTION

Dans les dernières années, les modèles de risque de crédit ont été sujets à une importante évolution dans le domaine de la recherche. En plus, la mesure du risque de crédit est cruciale pour beaucoup de participants qui opèrent sur le marché financier, par exemple les investisseurs d'obligations corporatives, les opérateurs de dérivés de crédit, les banques, les compagnies d'assurance, etc.

Il y a une ample gamme de modèles et de méthodes qui permettent de mesurer le risque de crédit, parmi eux figurent les méthodes de filtrage. En fait, une des principales motivations de ce document est de connaître en profondeur ses propriétés, de trouver ses possibles faiblesses et de proposer une extension. En plus, on s'intéresse aux méthodes d'estimation de paramètres par maximum de vraisemblance et par espérance-maximisation.

De façon générale, le risque de crédit concerne la possibilité de pertes financières en raison de changements dans la qualité de crédit des participants du marché. La plupart des changements radicaux sur la qualité de crédit sont liés à des événements de défaut. Selon *l'International Swaps and Derivatives Association (ISDA)*, trois types d'événements de crédit sont possibles, à savoir la faillite de l'entité de référence, le défaut de paiement et la restructuration.

Ainsi, le défaut est considéré comme un événement rare après lequel une société peut cesser d'être viable en générant de grandes pertes financières aux porteurs de quelques titres. Cette idée peut s'étendre aussi au risque de crédit des pays (dette souveraine) avec un risque non négligeable de défaut, comme celui de pays en voie de développement.

En pratique, beaucoup d'outils peuvent être utilisés pour analyser et évaluer le risque de crédit, mais deux métriques sont souvent utilisées : des ratios de couverture d'intérêt et des ratios de capitalisation.

D'autres indicateurs répondent à la question de savoir comment évaluer le niveau d'endettement d'une entreprise. Trois ratios financiers comptent parmi les principaux : le ratio fonds de roulement, le ratio dette long terme, et le ratio d'endettement total ou levier financier.

Évidemment, les mesures de risque de crédit impliquent non seulement la connaissance de la valeur des actifs et des fonds propres, mais aussi l'évaluation d'instruments financiers tels que les obligations corporatives et les swaps sur défaillance de crédit (CDS)¹.

L'émission d'obligations est une alternative que les compagnies utilisent pour se financer et attirer des capitaux. Ces obligations offrent des rendements plus élevés que celles émises par les États, car elles présentent un risque plus élevé. Elles se distinguent par leur échéance et par le risque de crédit. Des échéances à court terme (5 ans ou moins), à moyen terme (entre 5 et 12 ans) et à long terme (12 ans ou plus) sont disponibles sur ce marché.

En ce qui concerne le risque de crédit, il y a quatre principales sociétés de notation financière, *Standard & Poor's*, *Moody's*, *Fitch Ratings* et *Dagong* qui s'occupent d'évaluer la qualité des obligations en fonction des perspectives de remboursement de ses engagements envers ses créanciers.

En se basant sur les instruments financiers, les premiers modèles de risque de crédit sont été concentrés sur une structure basique du capital d'une société contenant seulement la dette et des fonds propres. D'après (Merton, 1974), ces modèles

1. Credit Default Swaps.

essaient de décrire les possibles gains ou pertes des détenteurs d'obligations comme résultat de changements dans la probabilité de défaut et non pas sur les gains ou pertes à cause du changement du taux d'intérêt.

Ces modèles ont évolué particulièrement après le boom des dérivés de crédit qui ont été commercialisés dans les décennies récentes. En fait, ils peuvent permettre d'avoir des composants stochastiques exogènes qui gouvernent la probabilité de défaut.

Au cours des dernières années, les chercheurs ont aussi proposé des modèles impliquant l'estimation de variables qui ne sont pas directement observables ou variables latentes ainsi que différentes façons d'estimer les paramètres qui sont reliés à la probabilité de défaut.

Ils se servent des modèles d'état-espace non linéaires, gaussiens ou non gaussiens, qui n'admettent pas généralement des solutions analytiques. C'est à ce niveau que les méthodes de filtrage deviennent une catégorie très populaire en finance pour estimer ces variables latentes.

Le chapitre 1 fournit une revue de la littérature dans le cadre des trois principales approches de modèles de risque de crédit : les modèles structurels, les modèles à forme réduite et les modèles hybrides. Dans le chapitre 2, on présente les principales propriétés des modèles de risque de crédit pour les deux premières approches. Dans le chapitre 3, on représente ces modèles comme des modèles d'état-espace. Dans le chapitre 4, on fait une description du principal outil que l'on utilise pour estimer les variables latentes : la méthode de filtre à particules. Le chapitre 5 présente deux méthodes d'estimation de paramètres via maximum de vraisemblance et espérance-maximisation. Des analyses par simulation sont illustrées dans le chapitre 6. Finalement, on fait une analyse empirique des 30 compagnies qui constituent l'indice *Dow Jones* dans le chapitre 7. On utilise les

paramètres estimés pour calculer la valeur des actifs de chacune des compagnies.

CHAPITRE I

REVUE DE LA LITTÉRATURE

Ce chapitre présente d'une manière brève une revue de la littérature sur l'évolution des modèles de risque de crédit et les méthodes d'estimation dans le cadre des trois principales approches : les modèles structurels, les modèles à forme réduite et les modèles hybrides qui sont une combinaison des deux premières approches.

Deux méthodes d'estimation de paramètres et des techniques de filtrage figurent également dans ce chapitre.

1.1 Modèles structurels

Le modèle de risque de crédit proposé par (Merton, 1974) est le point de départ de plusieurs modèles dans le cadre de l'approche structurelle (ou approche de la valeur de la firme). L'une des principales hypothèses est qu'un événement de crédit ne survient qu'à l'échéance d'une obligation ; en d'autres termes, si la compagnie n'est pas capable de faire face à ses engagements à l'échéance, on considérera cet événement comme défaut.

L'approche basée sur les modèles structurels est liée directement à la valeur des actifs d'une compagnie. La compagnie est considérée en défaut lorsque la valeur des actifs touche ou croise une barrière. Ici, le temps de défaut τ est un temps d'arrêt prévisible par rapport à un flux d'information disponible. Le temps aléatoire de

défaut est annoncé par une séquence croissante de temps d'arrêt.

Cette approche est intuitive, car l'évaluation et la théorie sont reliées à la technique classique de Black-Scholes (1973). En particulier, il est d'ailleurs très important d'effectuer l'évaluation des obligations d'entreprises et des fonds propres en fonction de ses actifs. Cette approche a l'avantage d'utiliser les données et les informations qui reflètent le bilan du débiteur. Cependant, dans la plupart des cas, la valeur des actifs ne peut pas être observable parce qu'elle n'est pas un titre transigé¹ et il se peut qu'elle soit complexe à identifier, ce qui représente un important inconvénient.

D'autres hypothèses importantes sont que le taux d'intérêt sans risque et la variance de la valeur des actifs sous-jacents de la compagnie s'avèrent des fonctions déterministes. Néanmoins, il existe des modèles qui assouplissent ces hypothèses. Des formules fermées ou semi-fermées de prix d'options européennes sont disponibles pour évaluer des obligations d'entreprises.

Les premières extensions du modèle de (Merton, 1974) admettent la possibilité que le défaut survienne dans le premier instant où la valeur des actifs sous-jacents de la compagnie croise une certaine barrière.

D'abord, la barrière est établie d'une manière exogène. Elle peut être une fonction déterministe dépendant du temps, comme dans le cas du modèle de (Black et Cox, 1976)² ou bien elle peut être constante comme dans le modèle de (Longstaff et Schwartz, 1995). Dans le modèle de (Black et Cox, 1976), la valeur des fonds

1. Il convient de remarquer qu'il s'agit ici non pas des valeurs comptables des actifs et des fonds propres, mais de leur valeur de marché.

2. Une barrière naturelle dépendant du temps, mais certainement pas la seule, peut prendre la forme $K(t) = K_0 e^{kt}$ avec $K_0 \leq K e^{-kT}$.

propres peut être calculée comme étant une option d'achat «*down-and-out*» du modèle Black-Scholes (1973). On trouve une expression fermée dans (Merton, 1974).

Les modèles de (Leland, 1994) et (Leland et Toft, 1996) établissent la barrière de défaut d'une manière endogène, en faisant un compromis entre les coûts de défaut (avocats, comptables, tribunaux, etc) et les avantages fiscaux qui favorisent les niveaux d'endettement plus élevés. Afin d'éviter le défaut, la compagnie doit émettre des fonds propres au service de sa dette. Une barrière optimale peut être déduite par les détenteurs d'actions par maximisation de la valeur de fonds propres en trouvant la valeur la plus basse possible de sorte que les fonds propres ne soient pas négatifs.

Comme nous l'avons déjà mentionné, les hypothèses reliées aux fonctions déterministes qui décrivent le taux d'intérêt sans risque et la variance de la valeur des actifs sous-jacents de la compagnie peuvent être assouplies, (Bu et Liao, 2013) généralisent le modèle de risque de crédit sous l'approche structurelle en imposant le fait que la volatilité de la firme suit un processus stochastique, cette volatilité suit particulièrement le modèle de (Heston, 1993), mais on peut aussi considérer le modèle de (Heston et Nandi, 1997). Dans les deux cas, une forme semi-fermée d'une option d'achat européenne est disponible pour calculer la valeur des fonds propres.

D'un autre côté, (Longstaff et Schwartz, 1995) introduisent aussi le taux d'intérêt stochastique corrélé avec le processus stochastique que suit la valeur des actifs de la compagnie. Une référence est l'article d'(Amin et Jarrow, 1992), dans lequel il est proposé une forme fermée d'une option d'achat européenne avec taux d'intérêt stochastique dont on se sert pour calculer la valeur des fonds propres.

1.2 Modèles à forme réduite

Contrairement à l'approche structurelle, l'approche à forme réduite suppose que l'événement de défaut survient par surprise dans à une date totalement inaccessible. Pour cette raison, le temps de défaut est considéré comme non prévisible.

Cette approche est focalisée sur la probabilité conditionnelle de défaut, étant donné que cet événement n'est pas encore arrivé et par l'existence d'un processus d'intensité. C'est pourquoi cette approche est considérée comme étant de nature purement probabiliste.

D'après (Brigo et Mercurio, 2007), l'idée s'articule autour de la description du temps de défaut τ comme le premier moment où un saut d'un processus de Poisson survient. Le défaut n'est pas inclus dans l'information du marché, c'est pour ceci que le défaut est reconnu comme un processus exogène indépendant de l'information marché. La surveillance du marché sans défaut, c.-à-d., les taux d'intérêt, les taux de change, etc., ne nous donnent pas beaucoup d'information sur les processus de défaut et il n'y a aucun raisonnement économique derrière le défaut.

Le modèle à forme réduite est discuté dans l'articles de (Jarrow et Turnbull, 1995) dans lequel l'événement de crédit n'est pas relié à l'information de la compagnie. Ce modèle prend comme donnée une structure à terme stochastique des taux d'intérêt sans risque de défaut et des écarts du risque de crédit stochastique avec une maturité spécifique. De leur côté, (Duffie et Singleton, 1999) modélisent le temps de risque de défaut comme un processus positif en se basant sur un processus de Markov caché. Le modèle proposé par (Madan et Unal, 1998) utilise le prix de dette junior et senior pour identifier les paramètres du processus décrivant le temps de défaut et le risque de recouvrement inhérents dans le défaut.

Le modèle de (Lando, 1998) généralise le modèle markovien proposé par (Jarrow

et al., 1997) en permettant que les intensités dépendent des variables d'état qui régissent l'évolution de la structure à termes. Le processus de Cox est le principal outil pour calculer les prix des instruments financiers dont le risque de crédit est un facteur important.

L'évolution des intensités du processus de Cox est habituellement représenté par le modèle de Cox–Ingersoll–Ross (CIR). En effet, (Brigo et Alfonsi, 2005) représentent non seulement des intensités mais aussi des taux d'intérêts avec ce modèle. Dans leur article, ils discutent la manière de faire la calibration de la structure à terme des taux d'intérêt et celle reliée aux swaps sur défaillance de crédit (CDS) avec des données du marché.

Comme il est bien connu, le modèle CIR a une restriction dans ses paramètres lesquels limitent la volatilité du processus rendent les paramètres, en certains cas, non réalistes. C'est pourquoi, le processus de diffusion CIR avec sauts (JCIR) est souvent utilisé. En fait, il permet, à la fois, de représenter les intensités stochastiques et d'ajouter une composante de variance. De plus, il appartient à la classe des modèles de diffusion affines avec sauts, ce qui le rend mathématiquement attirant.

L'article de (Brigo et El-Bachir, 2006) généralise le modèle proposé par (Brigo et Alfonsi, 2005). En faisant une analyse numérique, ils montrent que le modèle avec sauts peut générer un sourire de volatilité raisonnable.

La maniabilité mathématique de ces modèles est considérée comme un des principaux avantages des modèles à forme réduite.

Les articles publiés par (Jarrow et Protter, 2004) et par (Elizalde, 2005) montrent les différences entre les modèles structurels et à forme réduite. Les modèles structurels assument une connaissance complète de la valeur des actifs. Les modèles

à forme réduite assument par contre une connaissance incomplète de l'état de l'entreprise.

1.3 Modèles hybrides

Les modèles hybrides combinent les meilleures caractéristiques des deux approches discutées précédemment tout en évitant leurs difficultés. Ils reprennent l'attractivité économique et intuitive de l'approche structurelle et la maniabilité et la performance empirique de l'approche à forme réduite.

Une liste de nombreux modèles de risque de crédit hybrides a été proposée. D'après (Boudreault *et al.*, 2013), on peut la diviser en deux sous-classes. On présente les modèles qu'on considère comme les plus importants.

La première sous-classe est basée sur la façon dont l'information disponible aux investisseurs est modélisée. En contraste aux modèles structurels, l'hypothèse d'une information complète est détendue et le défaut perd sa prévisibilité.

Le premier exemple de cette sous-classe est proposé par (Duffie et Lando, 2001); ici, l'intensité est calculée en termes de la densité conditionnelle de la valeur des actifs. En plus, ils supposent que l'on n'observe pas le processus continu de la valeur des actifs; au lieu de ceci, on observe un processus à temps discret avec un bruit. En conséquence, un nouveau filtrage est défini comme sous-ensemble du filtrage du modèle à temps continu de l'approche structurelle, qui est reliée à la valeur des actifs. Finalement, le temps de défaut est fonction d'une barrière constante.

Un autre exemple est le modèle de (Goldberg et Giesecke, 2004). Dans ce cas, le temps de défaut τ dépend d'une barrière aléatoire qui ne peut pas être observée, donc τ est inaccessible. De leur côté, le modèle de (Cetin *et al.*, 2004) regarde la valeur des actifs comme étant un flux financier. Ils définissent une barrière qui

prend la valeur en tout temps. Le temps de défaut τ est donc le premier instant où le flux touche ou croise la barrière. Ces modèles supposent que l'on observe la valeur des actifs comme un modèle stochastique à temps continu.

La deuxième sous-classe de modèles hybrides est bâtie en considérant le processus d'intensité qui dépend des variables structurelles permettant une interprétation économique et financière du défaut.

Des contributions sont trouvées dans des articles comme (Madan et Unal, 2000). Le modèle inclut un taux d'intérêt stochastique qui suit un processus Vasicek ; en plus, les intensités sont représentées à partir d'une fonction déterministe qui dépend à la fois de la valeur des actifs et de ce taux d'intérêt. Ce modèle offre certains avantages, à savoir une courbe complexe des écarts de crédit, y compris les courbes à deux bosses, et des solutions exactes (forme fermée) qui peuvent être obtenues pour calculer la valeur de la dette. Cependant, un inconvénient de ce modèle est que des écarts de crédit peuvent devenir négatifs.

Deux processus sont définis dans l'article de (Chen *et al.*, 2004). Le premier est relié à la dégradation de l'état de l'entreprise qui est une transformation monotone de la valeur de l'actif et peut être compris comme une distance au défaut. Le deuxième est un processus avec une magnitude de saut égal à un et des intensités stochastiques en fonction de l'état de l'entreprise. Le défaut peut être déclenché soit par dégradation successive de l'entreprise (premier processus) ou bien par un saut imprévisible (deuxième processus).

(Boudreault *et al.*, 2013) présentent un modèle hybride dont le défaut est relié à une source externe qui est fortement corrélée avec effet de levier. Une transformation paramétrique du ratio de la dette sert comme un processus d'intensité, ce qui signifie que le défaut peut se produire même si la compagnie a une perspective financière positive ou bien elle peut survivre, même si l'entreprise est fortement

endettée. Avec leur modèle, ils soulignent que le défaut est fortement corrélé avec la dette de l'entreprise, mais ne se produit pas nécessairement dès que la valeur des actifs croise la barrière.

Un modèle hybride plus récent est proposé par (Ballestra et Pacelli, 2014) lequel généralise le modèle de (Madan et Unal, 2000). Dans ce modèle, l'intensité est décrite par un processus stochastique. Il est possible d'avoir des courbes d'écart de crédit et de reproduire des structures par terme de plusieurs formes qui reflètent l'interaction entre la structure des capitaux propres du débiteur et la dynamique de défaut.

1.4 Estimation

Une méthodologie générale, qui a pour but d'estimer les paramètres de la valeur des actifs sous-jacents non observables, a été développée par (Duan, 1994). Cette méthodologie est basée sur l'estimation par la méthode de maximum de vraisemblance. Un fait important à souligner dans cette méthodologie est l'existence d'une bijection entre la valeur des fonds propres et la valeur des actifs sous-jacents.

Plusieurs recherches ont été motivées à partir de cet article; en effet, (Ericsson et Reneby, 2002) utilisent cette méthodologie pour estimer le prix d'obligations en considérant quatre modèles dont la valeur des actifs est présentée en forme fermée, à savoir : Black-Scholes (1973) / Merton (1974), Leland & Torf (1996) et Briys & de Varenne (1997); ainsi que le modèle d'Ericsson & Reneby (2001). De plus, ils comparent les performances des estimations de paramètres par rapport aux méthodologies proposées par la compagnie KMV³, qui était une de références du marché, et le modèle de Ronn & Verma (1986). Les résultats sont fortement

3. La compagnie KMV (Kealhofer, McQuown and Vasicek) fait partie de *Moody's Analytics Enterprise Risk Solutions*

en faveur de l'approche proposée par (Duan, 1994).

Par la suite, (Duan et Fulop, 2009) présentent une généralisation de l'approche proposé par (Duan, 1994). Ils justifient que les prix des fonds propres observés sont contaminés par un bruit transactionnel provenant du manque de liquidité, de l'asymétrie d'information et de la discrétisation du prix. Ainsi, le bruit transactionnel a une incidence sur la bijection entre la valeur des fonds propres et la valeur des actifs sous-jacents et la méthode d'estimation de (Duan, 1994) ne peut pas appliquée. Cette estimation est faite à travers la méthode de filtre à particules, dont le principal objectif est d'estimer une densité a posteriori des variables d'état (la valeur des actifs) sachant des observations (la valeur des fonds propres).

Dans le cadre de la méthode de filtre à particules, des articles comme (Pitt, 2002) et (Creal, 2012) décrivent la manière d'estimer les paramètres via maximum de vraisemblance. De plus, l'algorithme d'espérance-maximisation, proposé par (Dempster *et al.*, 1977), constitue une alternative qui peut être considérée. Cette alternative est décrite par (Schön *et al.*, 2011) et (Kokkala *et al.*, 2014) en prenant en considération des modèles d'état-espace.

1.5 Filtrage

La méthode de filtrage constitue l'un des piliers sur lesquels repose les méthodes d'estimation de paramètres qui seront implémentés dans ce document.

Une de premières méthodes de filtrage est le filtre de Kalman proposé par (Kalman, 1960) qui est décrite par (Hamilton, 1995). Elle suppose une représentation d'état-espace linéaire et des bruits gaussiens.

Dans le cas où la représentation d'état-espace ne serait pas un système linéaire, le filtre de Kalman est inapplicable, donc il faut se servir du filtre de Kalman

étendu (EKF) proposé par (Anderson et Moore, 1979) qui permet de linéariser localement le problème et d'appliquer les équations du filtre de Kalman classique. Une autre extension est proposée par (Julier et Uhlmann, 1997), est appelée filtre de Kalman « *Unscented* » (UKF).

L'UKF génère une densité d'importance qui correspond de plus près à la véritable densité a posteriori et a la capacité de générer des distributions à queue plus lourde que l'EKF. L'UKF utilise un ensemble de points pour faire correspondre non seulement la moyenne et la variance, mais aussi les moments d'ordre supérieur de la densité de variables d'état. De plus, l'estimation de paramètres avec cette technique de filtrage correspond à la méthode de quasi-vraisemblance.

L'estimation par filtrage est une technique bien établie dans plusieurs courants de la littérature sur la modélisation financière, par exemple (Carr et Wu, 2007) utilisent l'UKF afin d'estimer la volatilité implicite dans le cadre d'options sur devises.

Par ailleurs, le filtre à particules, aussi connu comme filtre bootstrap, est basé sur la méthode de Monte-Carlo et sur une approche bayésienne qui a été introduite par (Gordon *et al.*, 1993). On utilise cette technique dans des cas plus généraux : soit que la représentation d'état-espace est caractérisée par la linéarité et la normalité des variables latentes ou bien par la non-linéarité et la non-normalité des variables latentes. En plus, le filtre à particules a de bonnes propriétés quand l'on ne trouve pas une forme fermée pour l'estimation de la densité des variables d'état conditionnée à toutes les observations disponibles. C'est pourquoi le filtre à particules a plusieurs avantages sur d'autres techniques de filtrage.

L'efficacité de l'estimation par filtre à particules est déterminée par une méthode de pondération. Dans un ensemble de données, on accorde un poids plus élevé aux données qui donnent plus d'information et vice versa. La méthode de pondé-

ration peut être considérée comme la procédure de filtrage. Par conséquent, les données avec plus de contenu d'information seront conservées et seront cruciales dans l'estimation.

Un exemple d'application de la méthode de filtre à particules est présenté dans l'article (Johannes *et al.*, 2009). Les auteurs étudient la structure de sauts et la façon de faire des prévisions de la volatilité stochastique avec trois modèles, à savoir : volatilité stochastique pure, volatilité stochastique avec sauts dans le prix et finalement un modèle à volatilité stochastique avec sauts dans le prix et dans la volatilité. Le prix suit un processus Mouvement Brownien Géométrique et la volatilité suit un processus CIR. En plus, les sauts sont représentés par un processus de Poisson composé, mais on pourrait aussi les représenter par une distribution variance-gamma tel que considéré par (Huang et Wu, 2004). Soulignons que les sauts dans le prix et la volatilité peuvent être corrélés.

En reprenant l'article de (Johannes *et al.*, 2009), les densités a posteriori que l'on obtient avec le filtrage sont utilisées pour estimer les variables latentes (volatilité non observable) et les paramètres du modèle. Dans la première partie de cet article, des estimations sont obtenues sans considérer la présence du bruit transactionnel, cependant dans un exercice postérieur, il est proposé que les prix d'options sont observés avec bruits et on les capture avec un modèle d'erreur multiplicative. Ces erreurs dans les options sont estimées d'après les prix des options d'index observés qui sont dans la monnaie.

Dans le cadre de modèles structurels avec bruit transactionnel, (Huang et Yu, 2010) proposent une technique de filtrage bayésien qui estime la densité a posteriori par l'échantillonnage de Gibbs.

On fait référence à des tutoriels comme (Doucet et Johansen, 2009), (Cappé *et al.*, 2007) ou bien (Arulampalam *et al.*, 2002) qui contiennent plus de détails sur

les méthodes de filtrage.

CHAPITRE II

MODÈLES DE RISQUE DE CRÉDIT

Deux approches pour estimer les paramètres des modèles de risque de crédit sont examinées dans le présent chapitre. On commence par les modèles structurels, ensuite on définit les modèles à forme réduite.

2.1 Modèles structurels

Premièrement, les modèles structurels classiques considèrent un marché parfait c'est-à-dire l'absence de coûts de transaction, d'impôts et de taxes. Deuxièmement, le marché est supposé complet, autrement dit, le nombre d'actifs risqués est égal au nombre de sources de risque. Finalement, on suppose qu'il y a absence d'opportunité d'arbitrage.

On fixe un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ avec \mathcal{F} la sigma algèbre sur l'ensemble Ω et \mathbb{P} une mesure de probabilité. On suppose que sur cet espace existe un mouvement brownien standard W . Soit $\mathcal{F}_t = \sigma\{V_s : 0 \leq s \leq t\}$ une sigma algèbre représentant l'information générée par ce mouvement brownien jusqu'au temps t .

En plus, on suppose que la dynamique d'un actif sans risque ou compte bancaire $B(t)$ est définie comme :

$$dB(t) = r_t B(t) dt, \quad (2.1)$$

avec r_t le taux d'intérêt représenté par un processus stochastique positif. En consé-

quence,

$$B(t) = \exp \left(\int_0^t r_s ds \right). \quad (2.2)$$

On définit aussi le facteur d'actualisation stochastique comme :

$$D(t, T) = \frac{B(t)}{B(T)} = \exp \left(- \int_t^T r_s ds \right), \quad (2.3)$$

Ainsi le prix à t d'une obligation zéro-coupon de maturité T est donné par :

$$P(t, T) = \mathbb{E}^{\mathbb{Q}} [D(t, T) | \mathcal{F}_t], \quad (2.4)$$

où \mathbb{Q} est une mesure de probabilité risque-neutre équivalente à \mathbb{P} .

2.1.1 Modèle de Merton

Le modèle de (Merton, 1974) suppose que la valeur des actifs V_t suit un processus stochastique, plus précisément le mouvement brownien géométrique (MBG) avec $r_t = r$ pour tout t :

$$dV_t = \mu V_t dt + \sigma V_t dW_t^{\mathbb{P}}, \quad (2.5)$$

avec μV_t et σV_t représentant la dérive et la volatilité de V_t respectivement. De plus, $W_t^{\mathbb{P}}$ représente le mouvement brownien standard sous la mesure de probabilité réelle \mathbb{P} .

Si l'on suppose une valeur initiale V_0 , la solution de l'équation différentielle stochastique (2.5) est donnée par :

$$V_t = V_0 \exp \left\{ \left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t^{\mathbb{P}} \right\}. \quad (2.6)$$

Notons que la dynamique des actifs V_t sous une mesure de probabilité risque-neutre \mathbb{Q} équivalente à \mathbb{P} est donné par :

$$dV_t = r V_t dt + \sigma V_t dW_t^{\mathbb{Q}}, \quad (2.7)$$

avec W_t^Q un mouvement brownien standard dont la solution est donnée par :

$$V_t = V_0 \exp \left\{ \left(r - \frac{1}{2} \sigma^2 \right) t + \sigma W_t^Q \right\}. \quad (2.8)$$

Le modèle suppose qu'une compagnie émet deux types de réclamation au temps $t = 0$: une dette et des fonds propres, la première étant représentée par une obligation zéro-coupon avec une valeur faciale L et une maturité T .

On remarque qu'une hypothèse de ce modèle est que l'événement de crédit ne survient qu'à l'échéance de l'obligation. À la date de maturité T , on peut donc envisager deux évolutions possibles de la valeur des actifs V_t .

La première, si $V_T \geq L$, la valeur des actifs est suffisante pour rembourser les créanciers, l'entreprise rembourse la valeur faciale L aux créanciers et donne $V_T - L$ aux actionnaires. La deuxième, si $V_t < L$ l'entreprise est considérée en défaut parce que la valeur des actifs est insuffisante pour rembourser les créanciers. Ainsi, les créanciers reçoivent V_T et deviennent propriétaires de la firme et les actionnaires ne reçoivent rien en retour. De plus, on suppose une entreprise à responsabilité limitée ce qui veut dire que la perte potentielle des associés se limite au montant de leurs apports respectifs.

En bref, à la maturité, les actionnaires reçoivent $\max(V_T - L, 0)$ et les créanciers reçoivent $\min(L, V_T)$, donc les *payoffs* sont :

$$B_T = \min(L, V_T) = L - \max(L - V_T, 0) \quad (2.9)$$

$$S_T = \max(V_T - L, 0) \quad (2.10)$$

On remarque aussi que, s'il n'y a pas d'autres investisseurs ou entités qui reçoivent de paiements de V_t , signifiant que la compagnie ne doit pas payer des taxes en cas d'une banqueroute, la valeur des actifs de la compagnie est égale à la somme de

la valeur de la dette et la valeur des fonds propres.

$$V_T = B_T + S_T \quad (2.11)$$

et donc on remarque que le choix de L ne modifie pas la valeur de V_T

On veut évaluer le prix de la dette et des fonds propres au temps antérieur aux maturités T . D'une part, remarquons que le *payoff* aux créanciers est la différence de la valeur faciale de l'obligation zéro-coupon et le *payoff* d'une option de vente au temps T avec prix d'exercice L . D'autre part, le *payoff* des actionnaires est le *payoff* d'une option d'achat au temps T avec prix d'exercice L .

En conséquence, la valeur de la dette et des fonds propres au temps t , étant donné les paramètres μ , σ , r et une maturité de l'obligation T , sont calculés à l'aide des formules de prix des options d'achat et vente sous le modèle de Black-Scholes (1973) comme :

$$\begin{aligned} S_t &= \mathbb{E}^{\mathbb{Q}}[D(t, T)(V_T - L)_+ | \mathcal{F}_t] \\ &= V_t N(d_1) - L \cdot e^{-r(T-t)} N(d_2) \end{aligned} \quad (2.12)$$

$$\begin{aligned} B_t &= L \cdot e^{-r(T-t)} - \mathbb{E}^{\mathbb{Q}}[D(t, T)(L - V_T)_+ | \mathcal{F}_t] \\ &= L \cdot D(t, T) - [L \cdot D(t, T)N(-d_2) - V_t N(-d_1)] \end{aligned} \quad (2.13)$$

où $N(\cdot)$ est la fonction de répartition de la loi normale standard et avec :

$$\begin{aligned} d_1 &= \frac{\log(V_t/L) + (T-t)r + \frac{1}{2}(T-t)\sigma^2}{\sigma\sqrt{T-t}} \\ d_2 &= d_1 - \sqrt{T-t}\sigma \end{aligned}$$

2.1.2 Modèle de risque de crédit à taux d'intérêt stochastique

(Longstaff et Schwartz, 1995) proposent un modèle à taux d'intérêt stochastique. D'abord, on établit que la dynamique de la valeur de la firme V_t suit un mouvement

brownien géométrique (MBG) et le taux court instantané r_t suit un processus de Vasicek. On assume qu'il existe une corrélation ρ entre les deux dynamiques, donc :

$$\begin{aligned} dV_t &= r_t V_t dt + \sigma V_t (\rho dW_{t,1}^{\mathbb{Q}} + \sqrt{1 - \rho^2} dW_{t,2}^{\mathbb{Q}}) \\ dr_t &= \kappa_r (\theta_r^* - r_t) dt + \sigma_r dW_{t,1}^{\mathbb{Q}} \end{aligned}$$

avec $W_{t,1}^{\mathbb{Q}}$ et $W_{t,2}^{\mathbb{Q}}$ des mouvements browniens indépendants et avec $\theta_r^* = \theta - \frac{\sigma_r}{\kappa_r} \lambda$, et λ le paramètre d'une unité additionnelle de risque.

Si l'on suppose que la dynamique de r_t est gaussienne, une expression fermée de fonds propres S_t peut être obtenue. En vue d'obtenir cette expression dont la seule différence par rapport au modèle classique (Merton, 1974) est que le taux d'intérêt est stochastique, on a besoin de calculer :

$$\begin{aligned} S_t &= \mathbb{E}^{\mathbb{Q}} [D(t, T)(V_T - L)_+ | \mathcal{F}_t] \\ &= \mathbb{E}^{\mathbb{Q}} \left[\exp \left(- \int_t^T r_t dt \right) (V_T - L)_+ \middle| \mathcal{F}_t \right]. \end{aligned} \quad (2.14)$$

Une importante contribution est donnée par (Amin et Jarrow, 1992). Ils ont exprimé une forme fermée du prix des options européennes lorsque l'on considère un taux d'intérêt stochastique lequel est bâti dans le cadre du modèle de Heath, Jarrow et Morton, voir (Heath *et al.*, 1990; Heath *et al.*, 1992). La solution de l'expression (2.14) se présente ci-dessous avec la notation de l'article de (Kim, 2002). On a

$$S_t = V_t \Phi(d_1) - L \cdot P(0, T) \Phi(d_2), \quad (2.15)$$

où,

$$P(0, T) = \exp \left(\frac{1}{2} \Sigma_{22}^T - B(T) \right)$$

et,

$$d_1 = \frac{\Sigma_{11} + \Sigma_{12} - C(T)}{\sqrt{D_\Sigma}}$$

$$d_2 = d_1 - \sqrt{D_\Sigma}$$

$$D_\Sigma = \frac{1}{2}\Sigma_{11}\Sigma_{12} + \Sigma_{22}$$

où

$$\Sigma_{11} = \sigma^2 T$$

$$\Sigma_{22} = \left(\frac{\sigma_r}{\kappa_r}\right)^2 \left(T - \frac{3 + e^{-\kappa_r T}(e^{-\kappa_r T} - 4)}{2\kappa_r}\right)$$

$$\Sigma_{12} = \frac{\sigma\sigma_r\rho}{\kappa_r} \left(\frac{e^{-\kappa_r T} - 1}{\kappa_r} + T\right)$$

et

$$C(T) = \frac{\Sigma_{11}}{2} - B(T) + \log(L/S_t)$$

$$B(T) = -\frac{1}{\kappa_r} \left[\left(r_t - \frac{\kappa_r \theta_r - \sigma_r \lambda}{\kappa_r} \right) (e^{-\kappa_r T} - 1) - (\kappa_r \theta_r - \sigma_r \lambda) T \right].$$

2.2 Modèles à forme réduite

La deuxième approche qui est utilisée pour évaluer le risque de crédit, est celui connu comme les modèles à forme réduite. On s'intéresse à décrire le défaut à partir d'un processus exogène.

2.2.1 Changement de la filtration

Commençons par définir \mathcal{G}_t comme la *filtration complète* incluant l'information du marché et l'information du défaut, $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t^1$, avec \mathcal{F}_t la filtration du marché qui comprend toutes les informations autres que le fait de défaut ou de survie de

1. Cette expression représente la plus petite σ -algèbre qui contient \mathcal{F} et \mathcal{H}

la compagnie et $\mathcal{H}_t = \sigma\{1_{\{\tau \leq s\}} : 0 \leq s \leq t\}$ l'information indiquant si le défaut a eu lieu jusqu'au temps t , et si oui quand exactement (cette filtration est aussi appelée «*surveillance de défaut*»)

On remarque que l'on a ajouté à la filtration de marché libre de défauts \mathcal{F}_t la filtration qui contient l'information du défaut \mathcal{H}_t et que dans les modèles structurels $\mathcal{F}_t = \mathcal{G}_t$ parce qu'ils n'admettent pas le processus d'intensités.

L'évaluation des dérivés de crédit implique des valeurs espérées par rapport à la filtration \mathcal{G}_t . Dans le but de faciliter les calculs, on peut montrer qu'un changement de la filtration est possible pour obtenir des prix qui sont calculés à partir d'espérances conditionnelles par rapport à la filtration contenue dans \mathcal{F}_t . En effet, on trouve ces preuves dans (Brigo et Mercurio, 2007) et la proposition 3.1 de l'article de (Lando, 1998).

2.2.2 Processus de Cox

La construction du processus de Cox est basée sur un processus de Poisson non-homogène N_t lorsque les intensités sont représentés par une fonction déterministe $\lambda(t)$, qui satisfait :

$$\mathbb{P}(N_t - N_s = k) = \frac{\left(\int_s^t \lambda(u) du\right)^k}{k!} \exp\left(-\int_s^t \lambda(u) du\right), \quad (2.16)$$

pour $t > s$ et avec $N_t - N_s$ un processus de Poisson de moyenne $\int_s^t \lambda(u) du$.

Le processus de Cox est une généralisation du processus de Poisson non-homogène dont on suppose que les intensités sont stochastiques. On suppose que les intensités, dénotées par λ_t , sont représentées par une fonction continue et non négative. Plus de détails sur la construction sont trouvés dans (Lando, 1998).

Le temps de défaut τ est défini comme le premier moment où le processus de Cox

présente un saut

$$\tau = \inf \left\{ t : \int_0^t \lambda_u \geq E_1 \right\} \quad (2.17)$$

avec E_1 une variable aléatoire exponentielle de paramètre 1 qui est reliée aux temps entre sauts. En plus, l'intensité cumulative est définie comme la variable aléatoire suivante :

$$\Lambda(t) = \int_0^t \lambda_u du \quad \text{avec } \lambda_u > 0 \quad (2.18)$$

Le temps du premier saut, transformé par son intensité cumulée, est une variable aléatoire E_1 qui suit une loi exponentielle de paramètre 1, alors :

$$\Lambda(\tau) =: E_1 \sim \exp(1) \quad (2.19)$$

avec E_1 étant une variable aléatoire indépendante à toute l'information du marché de défaut, par exemple, la filtration \mathcal{F} , des intensités λ , les taux d'intérêt, etc. Notons que non seulement E_1 est une source aléatoire, mais aussi λ . Pour cette raison, le processus de Cox est connu comme processus de Poisson doublement stochastique (*doubly stochastic Poisson processes*).

Dans la définition d'un processus de Cox comme un processus de Poisson d'intensité stochastique, il est implicite que nous avons encore un processus de Poisson conditionnel à la filtration $\mathcal{F}_t^\lambda = \sigma \{ \lambda_s : s \leq t \}$.

La probabilité que le premier saut τ survient après le temps t peut être défini sous la mesure \mathbb{Q} . En fait, il s'agit de la probabilité de survie suivante :

$$\begin{aligned} \mathbb{Q}(\tau > t | \mathcal{F}_0) &= \mathbb{Q}(\Lambda(\tau) > \Lambda(t)) \\ &= \mathbb{Q} \left[E_1 > \int_0^t \lambda_u du \right] \\ &= \mathbb{E} \left[\mathbb{Q} \left[E_1 > \int_0^t \lambda_u du \middle| \mathcal{F}_t^\lambda \right] \right] \\ &= \mathbb{E} \left[e^{-\int_0^t \lambda_u du} \right] \end{aligned} \quad (2.20)$$

2.2.3 Obligations corporatives risquées zéro-coupon

On calcule le prix d'une obligation corporative risquée zéro-coupon au temps t , avec un taux de recouvrement R_{EC} constant qui sera payé à maturité, comme :

$$\mathbf{1}_{\{\tau > t\}} \bar{P}(t, T, R_{EC}) = \mathbb{E}^{\mathbb{Q}}[D(t, T)(R_{EC} \mathbf{1}_{\{\tau \leq T\}} + \mathbf{1}_{\{\tau > T\}}) \mid \mathcal{G}_t] \quad (2.21)$$

$$= R_{EC} P(t, T) + \mathbf{1}_{\{\tau > t\}} (1 - R_{EC}) \mathbb{E}^{\mathbb{Q}} \left(e^{-\int_t^T (r_s + \lambda_s) ds} \mid \mathcal{F}_t \right) \quad (2.22)$$

avec $P(t, T)$ défini comme le prix d'une obligation zéro-coupon standard au temps t .

Notons que la première composante de (2.22) représente le taux de recouvrement R_{EC} , qui sera reçu certainement, escompté au temps t . La deuxième composante représente le résiduel $(1 - R_{EC})$ lorsqu'il n'y a pas un événement de crédit, aussi escompté au temps t . Ce résiduel est escompté avec un taux spot, ajusté par le risque de défaut représenté par le processus λ_t . On a fait le changement de la filtration (de \mathcal{G}_t à \mathcal{F}_t) pour obtenir l'expression 2.22. La démonstration est donnée par (Jarrow et Yu, 2001).

En particulier, si l'on suppose qu'il y a indépendance entre le processus de taux court instantané r_t et celui des intensités λ_t l'expression précédente deviendra :

$$\mathbf{1}_{\{\tau > t\}} \bar{P}(t, T, R_{EC}) = R_{EC} P(t, T) + \mathbf{1}_{\{\tau > t\}} (1 - R_{EC}) P(t, T) \mathbb{Q}(\tau > T \mid \mathcal{F}_t) \quad (2.23)$$

En effet, la deuxième composante de (2.22) est :

$$\begin{aligned} \mathbf{1}_{\{\tau > t\}} \mathbb{E}^{\mathbb{Q}} \left(e^{-\int_t^T (r_s + \lambda_s) ds} \mid \mathcal{F}_t \right) &= \mathbf{1}_{\{\tau > t\}} \mathbb{E}^{\mathbb{Q}} \left(e^{-\int_t^T r_s ds} e^{-\int_t^T \lambda_s ds} \mid \mathcal{F}_t \right) \\ &= \mathbf{1}_{\{\tau > t\}} \mathbb{E}^{\mathbb{Q}} \left(e^{-\int_t^T r_s ds} \mid \mathcal{F}_t \right) \mathbb{E}^{\mathbb{Q}} \left(e^{-\int_t^T \lambda_s ds} \mid \mathcal{F}_t \right) \\ &= \mathbf{1}_{\{\tau > t\}} P(t, T) \mathbb{Q}(\tau > T \mid \mathcal{F}_t) \end{aligned}$$

2.2.4 Obligations corporatives risquées à coupons

On considère une obligation qui paie une série de coupons c_1, c_2, \dots, c_n aux temps T_1, T_2, \dots, T_n ; en plus, on assume un taux de recouvrement constant qui paie R_{EC} au premier moment T_i suivant le défaut. Le *payoff* de ce compromis de paiement, escompté au temps t , est défini comme :

$$\sum_{i=1}^n c_i D(t, T_i) \mathbf{1}_{\{\tau > T_i\}} + D(t, T_n) \mathbf{1}_{\{\tau > T_n\}} + R_{EC} \sum_{i=1}^n D(t, T_i) \mathbf{1}_{\{\tau \in (T_{i-1}, T_i)\}}$$

Pour obtenir le prix, on calcule l'espérance du *payoff* par rapport à la filtration \mathcal{G}_t et l'on utilise les relations de changement de la filtration; en plus, on a supposé l'indépendance entre les processus de taux d'intérêt r_t et des intensités λ_t de la même manière qu'on l'a fait précédemment, donc :

$$\begin{aligned} P^c(t, T, R_{EC}) &= \sum_{i=1}^n c_i \mathbb{E}[D(t, T_i) \mathbf{1}_{\{\tau > T_i\}} | \mathcal{G}_t] + \mathbb{E}[D(t, T_n) \mathbf{1}_{\{\tau > T_n\}} | \mathcal{G}_t] \\ &\quad + R_{EC} \sum_{i=1}^n \mathbb{E}[D(t, T_i) \mathbf{1}_{\{\tau \in (T_{i-1}, T_i)\}} | \mathcal{G}_t] \\ &= \sum_{i=1}^n c_i \mathbf{1}_{\{\tau > t\}} P(t, T_i) \mathbb{Q}(\tau > T_i) + \mathbf{1}_{\{\tau > t\}} P(t, T_n) \mathbb{Q}(\tau > T_n) \\ &\quad + R_{EC} \sum_{i=1}^n \mathbb{E}[D(t, T_i) | \mathcal{G}_t] (\mathbb{Q}(\tau > T_{i-1}) - \mathbb{Q}(\tau > T_i)) \end{aligned} \quad (2.24)$$

2.2.5 Le swap sur défaillance de crédit (CDS)

Rappelons qu'un swap est un produit financier, il est plus précisément un contrat transigé dans le marché gré à gré (*over-the-counter*)² entre deux parties qui s'engagent à faire un échange de flux financiers dans le futur. Un contrat à terme peut être vu comme un contrat swap dans lequel il y a un échange dans une seule

2. Bien que le contrats se transigent dans le marché gré à gré, ils sont encadrés par l'*International Swap and Derivatives Association* (ISDA).

date future. Le swap permet donc d'avoir des échanges de flux sur plusieurs dates futures.

Le swap sur défaillance de crédit (CDS) des produits financiers permet aux compagnies de transférer ou de redistribuer le risque de crédit. Un contrat bilatéral est conçu de manière à ce que la partie qui achète la protection accepte de payer une prime périodique à celle qui vend la protection. Cette dernière s'engage à faire une rétribution dans le cas d'un événement de crédit.

Il existe trois manières d'évaluer un CDS : en considérant le temps exact du défaut, en supposant que le temps de défaut survient dans un temps exact de la partition ou en considérant que la valeur actuelle du vendeur de protection est versée d'avance par la partie qui achète la protection. Dans ce document, on va considérer la première.

Supposons que l'acheteur de protection fait une série de paiements π dans les temps T_1, T_2, \dots, T_n ou jusqu'au temps de défaut $\tau \leq T_n$ en échange d'un paiement unique de protection $1 - R_{EC}$. Soit $T_{\beta(\tau)}$ le premier T_i suivant le défaut, c.-à-d., si $\tau \in (T_{i-1}, T_i]$ donc $T_{\beta(\tau)} = T_i$. On pose $\Delta(T_{i-1}, T_i)$ la fraction de l'année entre T_{i-1} et T_i .

Du point de vue de celle qui vend la protection, le *payoff* escompté au temps exact de défaut est :

$$\begin{aligned} \Psi_{Prot}(t) = & D(t, \tau)(\tau - T_{\beta(\tau)-1}) \cdot \pi \cdot \mathbb{1}_{\{T_1 < \tau < T_n\}} + \sum_{i=1}^n D(t, T_i) \Delta(T_{i-1}, T_i) \cdot \pi \cdot \mathbb{1}_{\{\tau \geq T_i\}} \\ & - \mathbb{1}_{\{T_1 < \tau \leq T_n\}} D(t, \tau)(1 - R_{EC}) \end{aligned} \quad (2.25)$$

Le premier terme représente la fraction de la prime acquise lorsque le défaut survient entre temps T_1 et T_n . Le deuxième terme représente les primes acquises lorsque le défaut n'est pas survenu au temps t et le troisième terme représente le montant que la compagnie qui offre la protection doit payer dans le cas où il y

aurait un événement de défaut.

D'une manière générale, on peut calculer l'écart de crédit d'un CDS sous la mesure neutre-risque à partir de la relation suivante :

$$\text{CDS}(t, \pi, R_{EC}) = \mathbb{E}^{\mathbb{Q}}[\Psi_{Prot}(t) \mid \mathcal{G}_t] \quad (2.26)$$

L'écart de crédit (ou *spread de crédit*) au temps t est la valeur de π qui fait que l'expression 2.26 est égale à zéro. On résout cette équation en considérant la formule de changement de la filtration.

$$\pi = \frac{(1 - R_{EC}) \mathbb{E}^{\mathbb{Q}}(D(t, \tau) \mathbb{1}_{\{T_a < \tau < T_b\}} \mid \mathcal{F}_t)}{\sum_{i=a+1}^b \Delta(T_{i-1}, T_i) \mathbb{Q}(\tau > t \mid \mathcal{F}_t) \bar{P}(t, T_i, R_{EC}) + Ac_t} \quad (2.27)$$

avec $\bar{P}(t, T_i, R_{EC})$ le prix d'une obligation zéro-coupon (2.22) avec taux de recouvrement R_{EC} . En plus, $Ac_t = \mathbb{E}^{\mathbb{Q}}(D(t, \tau)(\tau - T_{\beta(\tau)-1}) \mathbb{1}_{\{T_a < \tau < T_b\}} \mid \mathcal{F}_t)$ le terme relié à la prime acquise.

2.2.6 Modèles d'intensité stochastique - CIR

Le choix du processus stochastique λ_t définit les propriétés de la forme et les propriétés de la probabilité de survie.

Un des processus stochastiques les plus utilisés pour modéliser les intensités stochastiques est nommé le modèle racine carrée de retour à la moyenne ou simplement le processus de Cox-Ingersoll-Ross (CIR). Il a de bonnes propriétés permettant de garantir que les intensités demeurent positives, étant donné que la solution à son équation différentielle stochastique (EDS) suit une loi khi-carré.

L'EDS du processus de CIR est définie comme :

$$d\lambda_t = \kappa(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t \quad (2.28)$$

avec κ la vitesse de retour à la moyenne des intensités ; θ niveau des intensités à long terme ; $\sigma\sqrt{\lambda_t}$ la volatilité des intensités ; W_t un mouvement brownien standard et $\lambda_t > 0$ pour $t \geq 0$

Si $2\kappa\theta > \sigma^2$ et que $\kappa, \theta, \sigma \in \mathbb{R}^+$, le processus est bien défini. Cette restriction est importante parce que le processus sera strictement positif.

Le processus CIR appartient à l'ensemble des processus qui peuvent s'écrire comme des modèles de structure affines. On utilise cet important résultat pour calculer la probabilité de survie (2.20).

$$\mathbb{Q}^{CIR}(\tau > T) = A(t, T) \exp(-B(t, T)\lambda_t) \quad (2.29)$$

avec

$$B(t, T) = \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma + \kappa)(e^{\gamma(T-t)} - 1) + 2\gamma} \quad (2.30)$$

$$A(t, T) = \left(\frac{2\gamma e^{(\gamma+\kappa)(T-t)/2}}{(\gamma + \kappa)(e^{\gamma(T-t)} - 1) + 2\gamma} \right)^{\frac{2\kappa\theta}{\sigma^2}} \quad (2.31)$$

avec $\gamma = \sqrt{\kappa^2 + 2\sigma^2}$

D'après (Brigo et Mercurio, 2007), un des problèmes du modèle CIR est relié à sa restriction. Notons que si σ^2 prend des valeurs grandes, κ ou θ seront aussi grandes et, en conséquence, on aura des paramètres possiblement non réalistes.

2.2.7 Modèles d'intensité stochastique avec sauts - JCIR

D'une manière semblable, le processus de diffusion CIR avec sauts (JCIR) permet de modéliser les intensités stochastiques, mais aussi d'ajouter une composante de variance, avec

$$d\lambda_t = \kappa(\theta - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t + dJ_t^{\alpha, \gamma} \quad (2.32)$$

où $J_t^{\alpha,\gamma}$ est un processus à sauts purs ; α le taux d'arrivée des sauts ; γ la magnitude ou la taille des sauts. Remarquons que :

$$J_t^{\alpha,\gamma} = \sum_{i=1}^{M_t} Y_i \quad (2.33)$$

où M_t est un processus de Poisson à temps homogène avec intensité α ; Y_i sont des variables aléatoires qui suivent une loi exponentielle de paramètre γ .

D'après (Brigo et Mercurio, 2007), le processus JCIR appartient à la classe de *diffusions affines avec sauts* permettant d'écrire la probabilité de survie comme :

$$\mathbb{Q}^{JCIR}(\tau > T) = \bar{\alpha}(t, T) \exp(-\bar{\beta}(t, T)\lambda_t) \quad (2.34)$$

avec

$$\bar{\alpha}(t, T) = A(t, T) \left(\frac{2h \exp\left(\frac{h+\kappa+2\gamma}{2}(T-t)\right)}{2h + (h + \kappa + 2\gamma)(\exp(h(T-t)) - 1)} \right)^{\frac{2\alpha\gamma}{\sigma^2 - 2\kappa\gamma - 2\gamma^2}} \quad (2.35)$$

$$\bar{\beta}(t, T) = B(t, T) \quad (2.36)$$

avec $h = \sqrt{\kappa^2 + 2\sigma^2}$ et avec $A(t, T)$ et $B(t, T)$ tels que définis dans le processus CIR dans les expressions (2.30) et (2.31).

CHAPITRE III

REPRÉSENTATION DES MODÈLES D'ÉTAT-ESPACE

Sous l'approche structurelle, on a mentionné que la valeur marchande de la compagnie V_t pourrait être non observable. De même, sous l'approche à forme réduite, les intensités λ_t sont aussi des variables non-observables (ou variables latentes). Les modèles d'état-espace sont des outils qui nous aident à représenter des variables latentes d'un processus. Donc, tous les modèles présentés dans le chapitre antérieur peuvent être vus comme ce type de modèles. Dans ce chapitre, on présente la représentation d'état-espace pour tous les modèles.

On peut formuler une représentation d'état-espace lorsque l'on a besoin d'estimer l'état d'un système qui change à travers le temps. On suppose qu'il existe une séquence temporelle des états latents x_t qui peuvent être estimés indirectement à travers des observations données y_t . Le vecteur d'état contient toutes les informations pertinentes qui décrivent le système. Comme on le verra, il s'agit d'un problème d'estimation récursive.

Au début, on définit l'état non observé comme $\{x_t; t \in \mathbb{N}\}$ qui est un processus stochastique markovien d'ordre 1 avec une valeur aléatoire initiale x_0 de densité $p_{\Theta}(x_0)$. De plus, on suppose que l'on connaît la *densité de transition* $p_{\Theta}(x_t | x_{t-1})$ avec laquelle on est capable de générer des valeurs aléatoires.

De plus, soit $\{y_t; t \in \mathbb{N}\}$ le processus stochastique qui décrit les observations disponibles avec une *densité de mesure* $p_{\Theta}(y_t | x_t)$. Remarquons que les trois densités dépendent d'un vecteur de paramètres Θ . La représentation du modèle d'état-espace est alors l'expression suivante :

$$\begin{aligned} x_0 &\sim p_{\Theta}(x_0) \\ x_t | x_{t-1} &\sim p_{\Theta}(x_t | x_{t-1}) \quad \text{pour } t \geq 0 \end{aligned} \quad (3.1)$$

$$y_t | x_t \sim p_{\Theta}(y_t | x_t) \quad \text{pour } t \geq 0. \quad (3.2)$$

On peut interpréter $p_{\Theta}(x_0)$ comme la densité a priori sur l'état initial du système. On dénote l'ensemble de variables d'état au temps t comme $\mathbf{x}_{0:t} = \{x_0, \dots, x_t\}$ et l'ensemble de variables observées comme $\mathbf{y}_{0:t} = \{y_0, \dots, y_t\}$.

3.1 Représentation d'état-espace du modèle de Merton

Sachant que la valeur non observable des actifs de la compagnie suit un mouvement brownien géométrique (MBG)

$$dV_t = rV_t dt + \sigma V_t dW_t^{\mathbb{Q}},$$

on pose $x_t = \log(V_t)$ la variable latente. Par le lemme d'Itô, le processus de x_t suit le mouvement brownien arithmétique (MBA) :

$$dx_t = \left(r - \frac{\sigma^2}{2} \right) dt + \sigma dW_t^{\mathbb{Q}}$$

On s'appuie sur le schéma d'Euler-Maruyama (voir par exemple (Glasserman, 2003)) qui nous permet de discrétiser l'expression précédente sur une partition de l'intervalle $[0, T]$ ($0 = t_0 < t_1 < \dots < t_n = T$) ou bien sur la solution exacte.

On suppose que la valeur de fonds propres est observée avec un bruit transactionnel multiplicatif de la forme $S_t e^{\delta \nu_t}$. Donc, on pose y_{t_i} représentant le logarithme des

fonds propres ou prix des actions $\log(S_{t_i})$ plus un bruit qui suit une loi normale centrée et de variance δ^2 . Notons que d'après (2.12), la valeur des fonds propres S_t est fonction de la valeur non observable de la compagnie V_t , c'est-à-dire, $S_{t_i} = P^M(V_{t_i}) = P^M(\exp(x_{t_i}))$.

En conséquence, on a le système suivant :

$$\begin{aligned} x_{t_i} &= x_{t_{i-1}} + \left(r - \frac{\sigma^2}{2}\right) \Delta t_i + \sigma \sqrt{\Delta t_i} \epsilon_{t_i} \\ y_{t_i} &= \log(P^M(\exp(x_{t_i}))) + \delta \nu_{t_i} \end{aligned} \quad (3.3)$$

avec $\Delta t_i = t_i - t_{i-1}$, ν_{t_i} et ϵ_{t_i} deux variables aléatoires indépendantes qui suivent une loi normale centrée et réduite.

Donc, si x_0 est la valeur initiale du processus, on obtient la représentation d'état-espace du modèle de Merton dont les densités de transition et de mesure sont :

$$\begin{cases} x_{t_i} | x_{t_{i-1}} & \sim \mathcal{N}\left(x_{t_{i-1}} + \left(r - \frac{\sigma^2}{2}\right) \Delta t_i; \sigma^2 \Delta t_i\right) \\ y_{t_i} | x_{t_i} & \sim \mathcal{N}\left(\ln(P^M(\exp(x_{t_i}))) ; \delta^2\right) \end{cases}$$

3.2 Représentation d'état-espace du modèle à taux d'intérêt stochastique

Dans ce modèle on suppose que la valeur de la compagnie suit un mouvement brownien géométrique et que le taux d'intérêt stochastique suit un processus de Vasicek, de plus on suppose que les processus V_t et r_t sont corrélés.

$$\begin{aligned} dr_t &= \kappa(\theta^* - r_t)dt + \sigma_r dW_{t,1}^{\mathbb{Q}} \\ dV_t &= r_t V_t dt + \sigma V_t dW_{t,2}^{\mathbb{Q}} \end{aligned}$$

On pose $x_t = \log(V_t)$. Appliquant le lemme d'Itô, on obtient que :

$$\begin{aligned} dr_t &= \kappa(\theta^* - r_t)dt + \sigma_r dW_{t,1}^{\mathbb{Q}} \\ dx_t &= \left(r_t - \frac{\sigma^2}{2}\right) dt + \sigma dW_{t,2}^{\mathbb{Q}} \end{aligned}$$

Premièrement, notons que la solution exacte du processus de Vasicek est obtenue en utilisant le lemme d'Itô sur $U_t = e^{\kappa t} r_t$. Deuxièmement, le processus x_t suit mouvement brownien arithmétique (MBA). Encore une fois le schéma d'Euler-Maruyama nous permet de discrétiser les deux processus la même manière qu'on l'a fait avec le modèle de Merton.

On suppose que y_{t_i} est le logarithme des fonds propres $\log(S_{t_i})$ plus un bruit qui suit une loi normale centrée et de variance δ^2 . Notons que sous le modèle à taux d'intérêt stochastique, S_{t_i} est fonction de V_{t_i} et r_{t_i} et prend la forme de l'équation (2.15), donc la $S_{t_i} = P^V(V_{t_i}, r_{t_i}) = P^V(\exp(x_{t_i}), r_{t_i})$, avec $P^V(\cdot)$ le prix d'une option européenne du modèle à taux d'intérêt stochastique dont le taux d'intérêt suit un processus Vasicek.

Finalement, le système suivant est obtenu :

$$\begin{aligned} r_{t_i} &= e^{-\kappa \Delta t_i} r_{t_{i-1}} + \theta^*(1 - e^{-\kappa \Delta t_i}) + \sigma_r \sqrt{\frac{1}{2\kappa}(1 - e^{-2\kappa \Delta t_i})} \eta_{t_i} \\ x_{t_i} &= x_{t_{i-1}} + \left(r_{t_i} - \frac{\sigma^2}{2} \right) \Delta t_i + \sigma \sqrt{\Delta t_i} \epsilon_{t_i} \\ y_{t_i} &= \log(P^V(\exp(x_{t_i}), r_{t_i})) + \delta \nu_{t_i} \end{aligned} \quad (3.4)$$

avec $\Delta t_i = t_i - t_{i-1}$ et avec η_{t_i} , ν_{t_i} et ϵ_{t_i} trois variables aléatoires indépendantes qui suivent une loi normale centrée et réduite.

La représentation d'état-espace du modèle à taux d'intérêt a deux densités de transition $p_{\Theta}(r_{t_i} | r_{t_{i-1}})$ et $p_{\Theta}(x_{t_i} | x_{t_{i-1}}, r_{t_i})$ et une densité de mesure $p_{\Theta}(y_{t_i} | x_{t_i}, r_{t_i})$:

$$\begin{cases} r_{t_i} | r_{t_{i-1}} & \sim \mathcal{N}\left(e^{-\kappa \Delta t_i} r_{t_{i-1}} + \theta^*(1 - e^{-\kappa \Delta t_i}); \frac{\sigma_r^2}{2\kappa}(1 - e^{-2\kappa \Delta t_i})\right) \\ x_{t_i} | x_{t_{i-1}}, r_{t_i} & \sim \mathcal{N}\left(x_{t_{i-1}} + \left(r_{t_i} - \frac{\sigma^2}{2}\right) \Delta t_i; \sigma^2 \Delta t_i\right) \\ y_{t_i} | x_{t_i}, r_{t_i} & \sim \mathcal{N}\left(\ln(P^V(\exp(x_{t_i}), r_{t_i})); \delta^2\right) \end{cases}$$

3.3 Représentation d'état-espace pour les modèles à forme réduite

On peut aussi approximer la solution d'équation différentielle stochastique du processus CIR en utilisant le schéma de discrétisation d'Euler-Maruyama. Cependant, si le processus de diffusion à sauts (JCIR) est considéré pour représenter les intensités, on opte pour le schéma de discrétisation proposé par (Mikulevičius et Platen, 1988) tel que suggéré et utilisé par (Brigo et El-Bachir, 2006).

Ce dernier schéma de discrétisation est basé sur l'hypothèse d'indépendance entre le mouvement brownien W_t de la partie de diffusion pure du processus JCIR et le processus à sauts J_t . Ainsi, entre les sauts du processus de Poisson, les intensités évoluent comme un processus CIR standard.

Afin de simuler le processus à sauts, ils proposent de générer des temps de sauts et des amplitudes de sauts, puis procéder avec le schéma de simulation de diffusion en ajoutant les sauts aux moments où ils se produisent.

Spécifiquement, pour une trajectoire, on génère une variable Poisson de paramètre $\alpha(T-t)$ qui nous donne le nombre de sauts à simuler pendant la période de temps $[t, T]$, ensuite, on simule tous les temps de sauts comme étant une variable aléatoire uniforme sur l'intervalle $[t, T]$, finalement on génère la magnitude de sauts à chaque temps de saut comme une variable aléatoire exponentielle.

Soit y_t le logarithme d'un instrument financier susceptible de défaut $PF(\lambda_t)$ (une obligation corporative risquée ou bien CDS) avec intensités JCIR. On suppose que y_t est observée avec un bruit qui suit une loi normale centrée et de variance δ^2 , on a que :

$$\begin{aligned}\lambda_{t_i} &\approx \lambda_{t_{i-1}} + \kappa(\theta - \lambda_{t_{i-1}})\Delta t_i + \sigma\sqrt{\lambda_{t_{i-1}}\Delta t_i}\epsilon_{t_i} + I_t^{\alpha,\gamma} \\ y_{t_i} &= \log(PF(\lambda_{t_i})) + \delta\nu_{t_i}\end{aligned}$$

avec $\Delta t_i = t_i - t_{i-1}$ et ν_t et ϵ_t deux variables aléatoires indépendantes qui suivent une loi normale centrée et réduite. De plus, $I_t^{\alpha,\gamma}$ représente la magnitude du saut dans le temps t .

Si $I_t^{\alpha,\gamma} = 0$ pour $t = 0, 1, 2, \dots, T$, on aura la représentation d'état-espace du modèle CIR. Notons que les densités de transition et de mesure de la représentation d'état-espace sont :

$$\lambda_{t_i} | \lambda_{t_{i-1}} \sim \mathcal{N}(\lambda_{t_{i-1}} + \kappa(\theta - \lambda_{t_{i-1}})\Delta t_i; \sigma^2 \lambda_{t_{i-1}} \Delta t_i)$$

$$y_{t_i} | \lambda_{t_i} \sim \mathcal{N}(\log(PF(\lambda_{t_i})); \delta^2)$$

Si $I_t^{\alpha,\gamma} > 0$ pour $t = t_s$ et zéro si $t \neq t_s$, le processus de diffusion peut être vu comme un processus à deux état. Le premier état correspond à la partie du processus CIR standard avant le saut avec la même drift et volatilité et le deuxième état à la partie de processus après le saut dont le niveau du drift est augmenté en fonction de la magnitude du saut mais le paramètre de la volatilité demeure le même. On peut généraliser cette idée lors de la présence de plusieurs sauts. Cependant, la distribution exacte du processus de diffusion à sauts JCIR peut être calculé à partir de la convolution d'une variable normal et une variable Poisson composé.

Il est possible de trouver d'une manière fermée les moments conditionnels de premier et deuxième ordre du processus CIR et JCIR tel que montré par (Zhu *et al.*, 2014). Dans cet article le processus de diffusion JCIR avec un processus à sauts Poisson composé et un cas particulier du processus de diffusion JCIR avec un processus à sauts Hawkes.

3.4 Représentation d'état-espace à multiples sources d'information

On dénote $\mathbf{y}_t = \{\mathbf{y}_t^a, \mathbf{y}_t^b\}$ les ensembles de variables aléatoires indépendantes qui suivent une loi normale et qui représentent les logarithmes des instruments financiers susceptibles de défaut avec $\mathbf{y}_t^a = \{y_{t,1}^a, \dots, y_{t,n_a}^a\}$ et $\mathbf{y}_t^b = \{y_{t,1}^b, \dots, y_{t,n_b}^b\}$ et avec $y_{t,j}^a$ la j -ème obligation corporative risquée pour $j = 1, \dots, n_a$ et $y_{t,k}^b$ le k -ème CDS pour $k = 1, \dots, n_b$. On suppose qu'on observe $y_{t,j}^a$ avec un bruit qui suit une loi normale centrée et de variance δ_j^2 et $y_{t,k}^b$ avec un bruit qui suit une loi normale centrée et de variance ζ_j^2 .

Remarquons que d'après le schéma de discrétisation d'Euler-Maruyama, on peut approcher la solution d'équation différentielle stochastique du processus CIR ou JCIR, dont on a le système.

$$\begin{aligned}\lambda_t &\approx \lambda_{t-\Delta} + \kappa(\theta - \lambda_{t-\Delta})\Delta + \sigma\sqrt{\lambda_{t-\Delta}\Delta}\varepsilon_{t-1} + I_t^{\alpha,\gamma} \\ y_{t,j}^a &= \log(P_j(\lambda_t)) + \delta_j \cdot \nu_{t,j} \\ y_{t,k}^b &= \log(R_k(\lambda_t)) + \zeta_j \cdot \varepsilon_{t,k}\end{aligned}$$

avec $\varepsilon_t, \nu_{t,j}, \varepsilon_{t,k} \sim \mathcal{N}(0, 1)$. Si $I_t^{\alpha,\gamma} = 0$ pour $t = 0, 1, 2, \dots, T$, les densités de transition et de mesure de la représentation d'état-espace sont :

$$\begin{aligned}\lambda_t \mid \lambda_{t-\Delta} &\sim \mathcal{N}(\lambda_{t-\Delta} + \kappa(\theta - \lambda_{t-\Delta})\Delta, \sigma^2\lambda_{t-\Delta}dt) \\ \mathbf{y}_t \mid \lambda_t &\sim \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)\end{aligned}\tag{3.5}$$

avec

$$\boldsymbol{\mu}_t^T = \left[\log(P_1(\lambda_t)), \dots, \log(P_{n_a}(\lambda_t)), \log(R_1(\lambda_t)), \dots, \log(R_{n_b}(\lambda_t)) \right]^T$$

et

$$\boldsymbol{\Sigma}_t = \begin{bmatrix} \boldsymbol{\Sigma}_t^{(a)} & 0 \\ 0 & \boldsymbol{\Sigma}_t^{(b)} \end{bmatrix}$$

avec

$$\Sigma_t^{(a)} = \begin{bmatrix} \delta_1^2 & 0 & \dots & 0 \\ 0 & \delta_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \delta_{n_a}^2 \end{bmatrix}$$

On définit $\Sigma_t^{(b)}$ de manière semblable. Grâce à l'indépendance de $y_{t,j}^a$ et $y_{t,k}^b$ pour $j = 1, \dots, n_a$ et $k = 1, \dots, n_b$ on peut écrire la densité de mesure à multiples sources d'information comme le produit des densités de mesure à une source d'information :

$$p_{\Theta}(\mathbf{y}_t | \lambda_t) = \prod_{j=1}^{n_a} p_{\Theta}(y_{t,j}^a | \lambda_t) \prod_{k=1}^{n_b} p_{\Theta}(y_{t,k}^b | \lambda_t) \quad (3.6)$$

CHAPITRE IV

FILTRAGE

Les techniques de filtrage et de lissage nous servent à estimer les variables latentes des modèles sous des représentations d'état-espace. Bien que le filtrage calcule des estimations de l'état actuel de la représentation d'état-espace compte tenu de l'historique des observations données au même temps, le lissage peut être utilisé pour reconstruire les états qui se sont produits avant de la date actuelle.

L'un des objectifs du filtrage est de construire d'une manière séquentielle la densité conjointe des variables d'état tout en sachant les observations disponibles jusqu'au temps t , $p_{\Theta}(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t})$. Cette densité est connue comme la *densité conditionnelle a posteriori*.

D'abord, avec la règle de Bayes, on peut l'écrire comme :

$$p_{\Theta}(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t}) = \frac{p_{\Theta}(\mathbf{y}_{1:t} \mid \mathbf{x}_{0:t})p_{\Theta}(\mathbf{x}_{0:t})}{p_{\Theta}(\mathbf{y}_{1:t})} \quad (4.1)$$

mais elle peut être aussi exprimée récursivement en utilisant l'hypothèse de Markov d'ordre 1 que l'on a imposée sur les variables latentes dans la représentation

d'état-espace.

$$\begin{aligned}
p_{\Theta}(\mathbf{x}_{0:t} \mid \mathbf{y}_{1:t}) &= \frac{p_{\Theta}(\mathbf{y}_t, \mathbf{y}_{1:t-1}, \mathbf{x}_{0:t})}{p_{\Theta}(\mathbf{y}_t, \mathbf{y}_{1:t-1})} \\
&= \frac{p_{\Theta}(\mathbf{y}_t \mid \mathbf{y}_{1:t-1}, \mathbf{x}_{0:t}) [p_{\Theta}(\mathbf{y}_{1:t-1} \mid \mathbf{x}_{0:t})] p_{\Theta}(\mathbf{x}_{0:t})}{p_{\Theta}(\mathbf{y}_t \mid \mathbf{y}_{1:t-1}) p_{\Theta}(\mathbf{y}_{1:t-1})} \\
&= \frac{p_{\Theta}(\mathbf{y}_t \mid \mathbf{y}_{1:t-1}, \mathbf{x}_{0:t}) p_{\Theta}(\mathbf{x}_{0:t})}{p_{\Theta}(\mathbf{y}_t \mid \mathbf{y}_{1:t-1}) p_{\Theta}(\mathbf{y}_{1:t-1})} \left[\frac{p_{\Theta}(\mathbf{x}_{0:t}, \mathbf{y}_{1:t-1})}{p_{\Theta}(\mathbf{x}_{0:t})} \right] \\
&= \frac{p_{\Theta}(\mathbf{y}_t \mid \mathbf{y}_{1:t-1}, \mathbf{x}_{0:t}) p_{\Theta}(\mathbf{x}_{0:t})}{p_{\Theta}(\mathbf{y}_t \mid \mathbf{y}_{1:t-1}) p_{\Theta}(\mathbf{y}_{1:t-1})} \left[\frac{p_{\Theta}(\mathbf{x}_t \mid \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1}) p_{\Theta}(\mathbf{x}_{0:t-1} \mid \mathbf{y}_{1:t-1}) p_{\Theta}(\mathbf{y}_{1:t-1})}{p_{\Theta}(\mathbf{x}_{0:t})} \right] \\
&= p_{\Theta}(\mathbf{x}_{0:t-1} \mid \mathbf{y}_{1:t-1}) \frac{p_{\Theta}(\mathbf{y}_t \mid \mathbf{y}_{1:t-1}, \mathbf{x}_{0:t}) p_{\Theta}(\mathbf{x}_t \mid \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1})}{p_{\Theta}(\mathbf{y}_t \mid \mathbf{y}_{1:t-1})} \\
&\stackrel{\text{Markov}}{=} p_{\Theta}(\mathbf{x}_{0:t-1} \mid \mathbf{y}_{1:t-1}) \frac{p_{\Theta}(\mathbf{y}_t \mid \mathbf{x}_t) p_{\Theta}(\mathbf{x}_t \mid \mathbf{x}_{t-1})}{p_{\Theta}(\mathbf{y}_t \mid \mathbf{y}_{1:t-1})} \tag{4.2}
\end{aligned}$$

Malgré que cette relation donne une densité a posteriori multidimensionnelle complète, une expression analytique peut être obtenue seulement quand toutes les densités sont connues impliquant la résolution d'intégrales multiples.

Parce que le calcul de la densité conditionnelle a posteriori est inefficace et parfois inutile dans les applications en temps réel, les densités marginales sont plus souvent utilisées.

La *densité de filtrage* $p_{\Theta}(\mathbf{x}_t \mid \mathbf{y}_{1:t})$ est une densité marginale qui est obtenue récursivement en deux étapes : prévision et actualisation. On utilise la notation proposée entre autres par (Gordon *et al.*, 1993).

Dans la première étape, on suppose qu'au temps $t - 1$ on connaît la densité de filtrage $p_{\Theta}(\mathbf{x}_{t-1} \mid \mathbf{y}_{1:t-1})$, donc, d'après l'équation de Chapman-Kolmogorov, on peut calculer la *densité de prévision* au temps t comme :

$$p_{\Theta}(\mathbf{x}_t \mid \mathbf{y}_{1:t-1}) = \int p_{\Theta}(\mathbf{x}_t \mid \mathbf{x}_{t-1}) p_{\Theta}(\mathbf{x}_{t-1} \mid \mathbf{y}_{1:t-1}) d\mathbf{x}_{t-1} \tag{4.3}$$

Notons que $p_{\Theta}(x_t | x_{t-1})$ est la densité de transition de la représentation d'état-espace (3.1).

Dans la deuxième étape, l'observation y_t devient disponible, alors on actualise la densité de prévision pour obtenir la densité de filtrage au temps t . On s'appuie sur la règle de Bayes et sur la propriété markovienne d'ordre 1 pour montrer que la densité de filtrage est proportionnelle au produit entre la densité de mesure $p_{\Theta}(y_t | x_t)$ et la densité de prévision $p_{\Theta}(x_t | y_{1:t-1})$, comme cela est montré dans l'équation (4.4).

$$\begin{aligned}
 p_{\Theta}(x_t | y_{1:t}) &= \frac{p_{\Theta}(y_{1:t}, x_t)}{p_{\Theta}(y_{1:t})} \\
 &= \frac{p_{\Theta}(y_t, y_{1:t-1}, x_t)}{p_{\Theta}(y_{1:t})} \\
 &= \frac{p_{\Theta}(y_t | x_t, y_{1:t-1})p_{\Theta}(x_t | y_{1:t-1})p_{\Theta}(y_{1:t-1})}{p_{\Theta}(y_t | y_{1:t-1})p_{\Theta}(y_{1:t-1})} \\
 &= \frac{p_{\Theta}(y_t | x_t)p_{\Theta}(x_t | y_{1:t-1})}{p_{\Theta}(y_t | y_{1:t-1})} \\
 &\propto p_{\Theta}(y_t | x_t)p_{\Theta}(x_t | y_{1:t-1}) \tag{4.4}
 \end{aligned}$$

D'autres densités marginales à souligner sont la *densité de lissage* $p_{\Theta}(x_t | y_{1:T})$ qui est définie en termes de la densité de filtrage $p_{\Theta}(x_t | y_{1:t})$ et la densité de prévision un pas en avance $p_{\Theta}(x_{t+1} | y_{1:t})$. Cette densité sera très utile lorsque l'on fera des estimations de paramètres par la méthode d'espérance-maximisation. Cette méthode d'estimation sera approfondie dans le chapitre suivant.

En utilisant les propriétés de probabilités conditionnelles, ainsi que la propriété

markovienne et la règle de Bayes, la densité de lissage peut être déduite comme :

$$\begin{aligned}
p_{\Theta}(x_t | \mathbf{y}_{1:T}) &= \int p_{\Theta}(x_t, x_{t+1} | \mathbf{y}_{1:T}) dx_{t+1} \\
&= \int p_{\Theta}(x_t | x_{t+1}, \mathbf{y}_{1:T}) p_{\Theta}(x_{t+1} | \mathbf{y}_{1:T}) dx_{t+1} \\
&= \int \frac{p_{\Theta}(x_t, x_{t+1}, \mathbf{y}_{1:t}, \mathbf{y}_{t+1:T})}{p_{\Theta}(x_{t+1}, \mathbf{y}_{1:t}, \mathbf{y}_{t+1:T})} p_{\Theta}(x_{t+1} | \mathbf{y}_{1:T}) dx_{t+1} \\
&= \int \frac{p_{\Theta}(\mathbf{y}_{t+1:T} | x_t, x_{t+1}, \mathbf{y}_{1:t}) p_{\Theta}(x_t | x_{t+1}, \mathbf{y}_{1:t})}{p_{\Theta}(\mathbf{y}_{t+1:T} | x_{t+1}, \mathbf{y}_{1:t})} p_{\Theta}(x_{t+1} | \mathbf{y}_{1:T}) dx_{t+1} \\
&= \int p_{\Theta}(x_t | x_{t+1}, \mathbf{y}_{1:t}) p_{\Theta}(x_{t+1} | \mathbf{y}_{1:T}) dx_{t+1} \\
&= \int \frac{p_{\Theta}(x_{t+1} | x_t) p_{\Theta}(x_t | \mathbf{y}_{1:t})}{p_{\Theta}(x_{t+1} | \mathbf{y}_{1:t})} p_{\Theta}(x_{t+1} | \mathbf{y}_{1:T}) dx_{t+1} \\
&= p_{\Theta}(x_t | \mathbf{y}_{1:t}) \int \frac{p_{\Theta}(x_{t+1} | x_t) p_{\Theta}(x_{t+1} | \mathbf{y}_{1:T})}{p_{\Theta}(x_{t+1} | \mathbf{y}_{1:t})} dx_{t+1} \tag{4.5}
\end{aligned}$$

Dans le passage de la quatrième à la cinquième ligne, on a utilisé le fait que :

$$p_{\Theta}(\mathbf{y}_{t+1:T} | x_t, x_{t+1}, \mathbf{y}_{1:t}) = p_{\Theta}(\mathbf{y}_{t+1:T} | x_{t+1}, \mathbf{y}_{1:t}).$$

On a cette égalité parce que, étant donné x_{t+1} , il n'y a pas plus d'information additionnelle sur $\mathbf{y}_{t+1:T}$ disponible dans x_t . L'équation (4.5) représente une récursion arrière de $p_{\Theta}(x_{t+1} | \mathbf{y}_{1:T})$ à $p_{\Theta}(x_t | \mathbf{y}_{1:T})$, sachant les densités de filtrage et de prévision une étape en avance.

4.1 Filtre à particules

Notre principal intérêt avec la méthode de filtre à particules est de trouver des estimations des densités de prévision (4.3), de filtrage (4.4) et de lissage (4.5) de la section précédente.

Remarquons que des solutions analytiques à ces équations peuvent être obtenues

seulement pour un choix restrictif des représentations d'état-espace. Le filtre à particules est une des techniques qui ont été proposées pour les obtenir.

On trouve différentes versions du filtre à particules appliquées en finance, par exemple, (Johannes et Polson, 2003) utilise des échantillonneurs MCMC¹ comme Metropolis-Hastings ou Gibbs présentés dans l'article de (Geman et Geman, 1984). Une autre version du filtre à particules est l'échantillonnage préférentiel séquentiel (*sequential importance sampling*) laquelle est décrite par (Doucet *et al.*, 2000). Ce dernier version est l'objet de notre intérêt.

4.1.1 Échantillonnage préférentiel séquentiel

Le but de la méthode d'échantillonnage préférentiel séquentiel est d'estimer la densité de filtrage $p_{\Theta}(x_t | \mathbf{y}_{1:t})$ à partir de la densité conjointe a posteriori $p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t})$. Cette méthode permet aussi de calculer des espérances de la forme :

$$I(f_t) = \mathbb{E}(f_t(\mathbf{x}_{0:t})) = \int f_t(\mathbf{x}_{0:t}) p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) d\mathbf{x}_{0:t} \quad (4.6)$$

En effet, une estimation de cette espérance est obtenue en supposant que l'on dispose d'un ensemble de variables d'état jusqu'au temps t , $\{\mathbf{x}_{0:t}^{(i)}\}_{i=1}^N$ nommé *l'ensemble des particules*. Cet ensemble représente un échantillon de la densité $p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t})$.

D'abord, la densité conjointe a posteriori peut être approximée en utilisant la fonction delta de Dirac $\delta(\cdot)$. Il s'agit d'une abstraction mathématique qui permet de représenter une densité continue comme densité discrète. Des propriétés de cette fonction sont données dans (Aratyn et Rasinariu, 2006).

$$p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) \approx \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_{0:t} - \mathbf{x}_{0:t}^{(i)}) \quad (4.7)$$

1. en français : Méthode de Monte-Carlo par chaînes de Markov

D'après la propriété de filtrage de la fonction delta de Dirac, on obtient la première estimation de $I(f_t)$. Cette estimation peut être vue comme une intégration par la méthode de Monte Carlo.

$$\begin{aligned}
 I_N^{(1)}(f_t) &= \int f_t(\mathbf{x}_{0:t}) p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) d\mathbf{x}_{0:t} \\
 &= \int f_t(\mathbf{x}_{0:t}) \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_{0:t} - \mathbf{x}_{0:t}^{(i)}) d\mathbf{x}_{0:t} \\
 &= \frac{1}{N} \sum_{i=1}^N f_t(\mathbf{x}_{0:t}^{(i)})
 \end{aligned} \tag{4.8}$$

Si, pour une raison quelconque, des échantillons de $p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t})$ ne sont pas facilement pris, une alternative est de se servir du principe d'échantillonnage préférentiel dont on peut prendre des échantillons à partir d'une densité auxiliaire $q_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t})$ appelée la *densité d'importance*. Le principe d'échantillonnage préférentiel est décrit dans l'appendice (A).

Notons que la relation suivante est équivalente à celle définie dans (4.6).

$$I(f_t) = \frac{\int f_t(\mathbf{x}_{0:t}) \omega_t q_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) d\mathbf{x}_{0:t}}{\int \omega_t q_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) d\mathbf{x}_{0:t}} \tag{4.9}$$

avec ω_t connues comme les *poids d'importance* et définies comme :

$$\omega_t = \frac{p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t})}{q_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t})} \tag{4.10}$$

En termes de la fonction delta de Dirac, si $\{\mathbf{x}_{0:t}^{(i)}\}_{i=1}^N$ représente un échantillon de la densité $q_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t})$, une approximation de cette densité est donnée par :

$$q_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) \approx \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_{0:t} - \mathbf{x}_{0:t}^{(i)}) \tag{4.11}$$

Donc, d'après la propriété de filtrage de la fonction delta de Dirac, une deuxième

estimation de $I(f_t)$ est donnée par :

$$\begin{aligned}
 I_N^{(2)}(f_t) &= \frac{\int f_t(\mathbf{x}_{0:t}) \omega_t \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_{0:t} - \mathbf{x}_{0:t}^{(i)}) d\mathbf{x}_{0:t}}{\int \omega_t \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_{0:t} - \mathbf{x}_{0:t}^{(i)}) d\mathbf{x}_{0:t}} \\
 &= \frac{\frac{1}{N} \sum_{i=1}^N f_t(\mathbf{x}_{0:t}^{(i)}) \omega_t^{(i)}}{\frac{1}{N} \sum_{i=1}^N \omega_t^{(i)}} \\
 &= \sum f_t(\mathbf{x}_{0:t}^{(i)}) \tilde{\omega}_t^{(i)} \tag{4.12}
 \end{aligned}$$

avec $\tilde{\omega}_t^{(i)}$ les poids d'importance normalisés :

$$\tilde{\omega}_t^{(i)} = \frac{\omega_t^{(i)}}{\sum_{i=1}^N \omega_t^{(i)}} \tag{4.13}$$

Le résultat (4.12) peut être aussi vu comme une estimation de l'intégrale par la méthode de Monte Carlo ; en plus, notons qu'une autre approximation de la densité a priori conditionnelle $p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t})$ peut être obtenue de l'ensemble à particules pondérée $\{\mathbf{x}_{0:t}^{(i)}, \omega_t^{(i)}\}_{i=1}^N$ comme :

$$p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) \approx \sum_{i=1}^N \tilde{\omega}_t^{(i)} \delta(\mathbf{x}_{0:t} - \mathbf{x}_{0:t}^{(i)}) \tag{4.14}$$

Jusqu'ici les poids d'importance non normalisés de la relation (4.10) sont calculés directement à partir d'un échantillon, mais souvent il est préférable de les calculer d'une manière itérative.

D'une part, on suppose que la densité d'importance admet la factorisation suivante :

$$q_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) = q_{\Theta}(x_t | \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t}) q_{\Theta}(\mathbf{x}_{0:t-1} | \mathbf{y}_{1:t-1}) \tag{4.15}$$

D'une autre part, la densité conditionnelle a posteriori peut être écrite à l'aide de

la règle de Bayes comme :

$$\begin{aligned}
 p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) &= \frac{p_{\Theta}(\mathbf{y}_t, \mathbf{y}_{1:t-1}, \mathbf{x}_{0:t})}{p_{\Theta}(\mathbf{y}_t | \mathbf{y}_{1:t-1})p_{\Theta}(\mathbf{y}_{1:t-1})} \\
 &= \frac{p_{\Theta}(\mathbf{y}_t | \mathbf{x}_{0:t}, \mathbf{y}_{1:t-1})p_{\Theta}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t-1})p_{\Theta}(\mathbf{y}_{1:t-1})}{p_{\Theta}(\mathbf{y}_t | \mathbf{y}_{1:t-1})p_{\Theta}(\mathbf{y}_{1:t-1})} \\
 &= \frac{p_{\Theta}(\mathbf{y}_t | \mathbf{x}_{0:t}, \mathbf{y}_{1:t-1}) [p_{\Theta}(x_t | \mathbf{x}_{0:t-1}, \mathbf{y}_{1:t-1})p_{\Theta}(\mathbf{x}_{0:t-1} | \mathbf{y}_{1:t-1})] p_{\Theta}(\mathbf{y}_{1:t-1})}{p_{\Theta}(\mathbf{y}_t | \mathbf{y}_{1:t-1})p_{\Theta}(\mathbf{y}_{1:t-1})} \\
 &\propto p_{\Theta}(\mathbf{y}_t | x_t)p_{\Theta}(x_t | x_{t-1})p_{\Theta}(\mathbf{x}_{0:t-1} | \mathbf{y}_{1:t-1}) \quad (4.16)
 \end{aligned}$$

en remplaçant les résultats (4.15) et (4.16) dans (4.10), on obtient la façon itérative de calculer les poids d'importance que l'on souhaitait :

$$\omega_t^{(i)} \propto \frac{p_{\Theta}(\mathbf{y}_t | x_t^{(i)})p_{\Theta}(x_t^{(i)} | x_{t-1}^{(i)})p_{\Theta}(\mathbf{x}_{0:t-1}^{(i)} | \mathbf{y}_{1:t-1})}{q_{\Theta}(x_t^{(i)} | \mathbf{x}_{0:t-1}^{(i)}, \mathbf{y}_{1:t})} \frac{p_{\Theta}(\mathbf{x}_{0:t-1}^{(i)} | \mathbf{y}_{1:t-1})}{q_{\Theta}(\mathbf{x}_{0:t-1}^{(i)} | \mathbf{y}_{1:t-1})} \quad (4.17)$$

$$= \frac{p_{\Theta}(\mathbf{y}_t | x_t^{(i)})p_{\Theta}(x_t^{(i)} | x_{t-1}^{(i)})}{q_{\Theta}(x_t^{(i)} | \mathbf{x}_{0:t-1}^{(i)}, \mathbf{y}_{1:t})} \omega_{t-1}^{(i)} \quad (4.18)$$

Afin d'avoir une représentation de la densité de filtrage, il est utile de supposer que la densité d'importance ne dépend que de x_{t-1} et y_t , ainsi les poids d'importance deviennent :

$$\omega_t^{(i)} \propto \frac{p_{\Theta}(\mathbf{y}_t | x_t^{(i)})p_{\Theta}(x_t^{(i)} | x_{t-1}^{(i)})}{q(x_t^{(i)} | x_{t-1}^{(i)}, y_t)} \omega_{t-1}^{(i)} \quad (4.19)$$

En normalisant ces poids, on peut maintenant supposer que l'on dispose d'un échantillon pondéré de taille N qui caractérise la densité de filtrage au temps $t - 1$, posons $\{x_{t-1}^{(i)}, \tilde{\omega}_{t-1}^{(i)}\}_{i=1}^N$. À partir de cet ensemble-ci, la densité de prévision

(4.3) est estimée comme un mélange des densités de transition.

$$\begin{aligned}
 \hat{p}_{\Theta}(x_t | \mathbf{y}_{1:t-1}) &= \int p_{\Theta}(x_t | x_{t-1}) \hat{p}_{\Theta}(x_{t-1} | \mathbf{y}_{1:t-1}) dx_{t-1} \\
 &= \int p_{\Theta}(x_t | x_{t-1}) \sum_{i=1}^N \tilde{\omega}_{t-1}^{(i)} \delta(x_{t-1} - x_{t-1}^{(i)}) dx_{t-1} \\
 &= \sum_{i=1}^N p_{\Theta}(x_t | x_{t-1}^{(i)}) \tilde{\omega}_{t-1}^{(i)} \tag{4.20}
 \end{aligned}$$

En combinant la densité de mesure avec le résultat précédent, on obtient une estimation de la densité de filtrage (4.4), comme ci-dessous :

$$\hat{p}_{\Theta}(x_t | \mathbf{y}_{1:t}) = p_{\Theta}(y_t | x_t) \sum_{i=1}^N p_{\Theta}(x_t | x_{t-1}^{(i)}) \tilde{\omega}_{t-1}^{(i)} \tag{4.21}$$

Finalement, l'algorithme qui décrit comment mettre en œuvre l'échantillonnage préférentiel séquentiel se résume dans le cadre suivant :

Algorithme d'échantillonnage préférentiel séquentiel

1. Initialiser l'algorithme avec $t = 0$
 - Pour $i = 1, \dots, N$ générer des valeurs aléatoires telles que $x_0^{(i)} \sim p(x_0)$
2. Pour $t = 1, 2, \dots, T$
 - Pour $i = 1, \dots, N$ générer des valeurs aléatoires telles que

$$x_t^{(i)} \sim q_{\Theta}(x_t | x_{t-1}^{(i)}, y_t)$$

et poser $\mathbf{x}_{0:t}^{(i)} = (\mathbf{x}_{0:t-1}^{(i)}, x_t^{(i)})$

- Pour $i = 1, \dots, N$ calculer les poids d'importance $\omega_t^{(i)}$, avec l'expression donnée dans (4.19)
- Pour $i = 1, \dots, N$ normaliser les poids d'importance pour obtenir $\tilde{\omega}_t^{(i)}$

4.1.2 Le problème de dégénérescence

La méthode d'échantillonnage préférentiel séquentiel induit une importante limitation. Après plusieurs itérations, toutes les particules, sauf une, ont des poids proches de zéro. On connaît ceci comme le problème de *dégénérescence*.

La taille effective de l'échantillon N_{eff} , aussi connue comme diagnostic de survie des particules, est une mesure de la dégénérescence attendue, laquelle est fonction de la variance du poids d'importance. Parce que la variance du poids n'est pas exactement calculable, (Arulampalam *et al.*, 2002) et (Kong *et al.*, 1994) utilisent une estimation définie en termes des poids d'importance normalisées au carré.

$$\widehat{N_{eff}} \approx \frac{1}{\sum_{i=1}^N (\tilde{\omega}_t^{(i)})^2} \quad (4.22)$$

Avec cette mesure, on peut avoir deux cas extrêmes. Dans le premier cas (et sans perte de généralité), si la première particule a survécu, alors son poids est $\tilde{\omega}_t^{(1)} = 1$ impliquant que tous les autres poids sont zéro $\tilde{\omega}_t^{(i)} = 0$ pour $i = 2, \dots, N$, alors $\widehat{N_{eff}} = 1$. Dans le deuxième cas, si tous les poids sont égaux à $1/N$, alors $\widehat{N_{eff}} = N$. Ainsi, chaque particule a été sélectionnée exactement une fois, autrement dit toutes les particules ont survécu au temps t .

4.1.3 Échantillonnage préférentiel séquentiel avec rééchantillonnage

On fait deux modifications à l'algorithme introduit précédemment. La première pour corriger le problème de dégénérescence et la deuxième par rapport au choix de la densité d'importance.

En ce qui concerne le problème de dégénérescence, une solution est d'introduire une étape de rééchantillonnage dont le principal objectif est de maintenir les particules qui ont un poids suffisamment élevé et d'éliminer celles qui ont un poids trop faible.

D'après (Arulampalam *et al.*, 2002) il y a plusieurs algorithmes pour faire le rééchantillonnage, par exemple, le rééchantillonnage stratifié ou bien rééchantillonnage systématique. Sous le même critère d'(Arulampalam *et al.*, 2002), on utilisera le rééchantillonnage systématique parce qu'il est simple à implémenter, qu'il minimise la variance des estimateurs de Monte-Carlo et que le temps de calcul est favorable.

Par rapport à la densité d'importance, on adopte celle utilisée par (Doucet *et al.*, 2001) qu'il a défini comme :

$$q_{\Theta}(x_t | x_{t-1}^{(i)}, y_t) = p_{\Theta}(x_t | x_{t-1}^{(i)}) \quad (4.23)$$

ce qui implique que les poids d'importance définis dans (4.19) deviennent :

$$\omega_t^{(i)} = p_{\Theta}(y_t | x_t^{(i)}) \omega_{t-1}^{(i)} \quad (4.24)$$

De plus, si l'étape de rééchantillonnage est incorporée, les poids d'importance au temps $t - 1$ n'affectent pas ceux du temps t rendant les particules équipondérées, en conséquence, les poids d'importance au temps t sont proportionnels à la densité de mesure :

$$\omega_t^{(i)} \propto p_{\Theta}(y_t | x_t^{(i)}) \quad (4.25)$$

On a finalement un algorithme qui inclut d'une part la correction au problème de dégénérescence et d'une autre le choix d'une densité d'importance qui facilite le processus de filtrage :

Algorithme (SIR)

1. Soit d'abord $t = 0$

— Pour $i = 1, \dots, N$ générer des valeurs aléatoires telles que

$$x_0^{(i)} \sim p(x_0)$$

— Définir $t = 1$

2. Échantillonnage d'importance : Pour $i = 1, \dots, N$

— générer des valeurs aléatoires telles que

$$\tilde{x}_t^{(i)} \sim p_{\Theta}(x_t | x_{t-1}^{(i)})$$

— calculer les poids d'importance

$$\omega_t^{(i)} = p_{\Theta}(y_t | \tilde{x}_t^{(i)})$$

— normaliser les poids d'importance pour obtenir $\tilde{\omega}_t^{(i)}$

3. Rééchantillonnage :

— À partir du vecteur $\tilde{x}_t^{(i)}$, rééchantillonner avec remplacement N particules et définir $x_t^{(i)}$

— On peut stocker d'une manière récursive le vecteur d'états rééchantillonnés

$$\mathbf{x}_{0:t}^{(i)} = (\mathbf{x}_{0:t-1}^{(i)}, x_t^{(i)})$$

— Définir $t = t + 1$ et retourner à l'étape 2.

4.2 Lissage

Le but de cette section est de donner une estimation de la densité de lissage ayant la forme suivante :

$$p_{\Theta}(x_t | \mathbf{y}_{1:T}) = \sum_{i=1}^N \tilde{\omega}_{t|T}^{(i)} \delta(x_t - x_t^{(i)}) \quad (4.26)$$

Remarquons que cette densité a le même support que la densité de filtrage, mais les poids sont égaux à $\tilde{\omega}_t^{(i)}$ au lieu de $\tilde{\omega}_t^{(i)}$. Ces poids sont connus comme les *poids lisses* et ont été proposés par (Doucet *et al.*, 2000). On va voir qu'ils sont obtenus à partir du processus de filtrage et que chacun est calculé avec une série de passes vers l'arrière en préservant les valeurs des particules.

On commence le développement en se rappelant que l'on a déduit l'expression (4.5) qui représente la densité de lissage :

$$p_{\Theta}(x_t | \mathbf{y}_{1:T}) = p_{\Theta}(x_t | \mathbf{y}_{1:t}) \int \frac{p_{\Theta}(x_{t+1} | x_t) p_{\Theta}(x_{t+1} | \mathbf{y}_{1:T})}{p_{\Theta}(x_{t+1} | \mathbf{y}_{1:t})} dx_{t+1}$$

Premièrement, on peut noter que la partie $p_{\Theta}(x_t | \mathbf{y}_{1:t})$ est la densité de filtrage qui peut être estimée par la méthode de filtre à particules par les poids d'importance normalisés $\{\tilde{\omega}_t^{(i)}\}_{i=1}^N$.

Deuxièmement, par rapport à la partie de l'intégral, notons que le dénominateur correspond à la densité de prévision qui, d'après la loi de probabilité totale, peut s'exprimer comme :

$$p_{\Theta}(x_{t+1} | \mathbf{y}_{1:t}) = \int p_{\Theta}(x_{t+1} | x_t) p_{\Theta}(x_t | \mathbf{y}_{1:t}) dx_t$$

Cette expression peut être aussi estimée en utilisant le filtre à particules comme l'on a fait avec l'expression (4.20) :

$$\hat{p}_{\Theta}(x_{t+1} | \mathbf{y}_{1:t}) = \sum_{i=1}^N p_{\Theta}(x_{t+1} | x_t^{(i)}) \tilde{\omega}_t^{(i)}$$

En ce qui concerne le numérateur, on remarque que dans le cas particulier où $t = T$ les densités de lissage et de filtrage sont les mêmes, on peut ainsi définir

$\tilde{\omega}_{T|T}^{(i)} = \tilde{\omega}_T^{(i)}$, pour $i = 1, \dots, N$. Donc, l'estimation de l'intégrale au temps t est alors :

$$\int \frac{p_{\Theta}(x_{t+1} | x_t) p_{\Theta}(x_{t+1} | \mathbf{y}_{1:T})}{p_{\Theta}(x_{t+1} | \mathbf{y}_{1:t})} dx_{t+1} \approx \sum_{j=1}^N \frac{p_{\Theta}(x_{t+1}^{(j)} | x_t^{(i)}) \tilde{\omega}_{t+1|T}^{(j)}}{\sum_{l=1}^N p_{\Theta}(x_{t+1}^{(j)} | x_t^{(l)}) \tilde{\omega}_t^{(l)}} \quad (4.27)$$

et finalement, l'estimation de la densité de lissage est donnée par :

$$\hat{p}_{\Theta}(x_t | \mathbf{y}_{1:T}) = \sum_{i=1}^N \tilde{\omega}_t^{(i)} \sum_{j=1}^N \frac{p_{\Theta}(x_{t+1}^{(j)} | x_t^{(i)}) \tilde{\omega}_{t+1|T}^{(j)}}{\sum_{l=1}^N p_{\Theta}(x_{t+1}^{(j)} | x_t^{(l)}) \tilde{\omega}_t^{(l)}} \quad (4.28)$$

Comme on le souhaitait, la dernière expression est une estimation de la densité de lissage (4.26) avec les poids lisses définis comme :

$$\tilde{\omega}_{t|T}^{(i)} = \sum_{j=1}^N \tilde{\omega}_{t+1|T}^{(j)} \frac{\tilde{\omega}_t^{(i)} p_{\Theta}(x_{t+1}^{(j)} | x_t^{(i)})}{\sum_{l=1}^N \tilde{\omega}_t^{(l)} p_{\Theta}(x_{t+1}^{(j)} | x_t^{(l)})} \quad (4.29)$$

et avec $\tilde{\omega}_{T|T}^{(i)} = \tilde{\omega}_T^{(i)}$.

Lors de l'estimation de paramètres par la méthode d'espérance-maximisation, il va nous falloir estimer la densité conjointe d'une paire de particules $(x_t^{(i)}, x_{t-1}^{(j)})$, représentée par $p_{\Theta}(x_t, x_{t-1} | \mathbf{y}_{1:T}, \theta_k)$. D'après (Schön *et al.*, 2011), cette densité est reliée aux poids suivants :

$$\tilde{\omega}_{t|T}^{(ij)} = \tilde{\omega}_{t+1|T}^{(j)} \frac{\tilde{\omega}_t^{(i)} p_{\Theta}(x_{t+1}^{(j)} | x_t^{(i)})}{\sum_{l=1}^N \tilde{\omega}_t^{(l)} p_{\Theta}(x_{t+1}^{(j)} | x_t^{(l)})} \quad (4.30)$$

Finalement, afin d'obtenir les poids lisses, il faut tout d'abord exécuter l'algorithme SIR, décrit dans la section 4.1.3. L'algorithme qui décrit comment faire le calcul avec une série de passes vers l'arrière est donné ci-dessous :

Algorithme de lissage

1. Exécuter l'algorithme SIR et stocker les particules $\{x_t^{(i)}\}$ et ses poids d'importance normalisés $\{\tilde{\omega}_t^{(i)}\}$ pour $i = 1, 2, \dots, N$ et pour $t = 1, 2, \dots, T$

2. Initialiser les poids lissés avec les poids terminaux $\tilde{\omega}_t^{(i)}$ pour $t = T$

$$\tilde{\omega}_{T|T}^{(i)} = \tilde{\omega}_T^{(i)} \text{ pour } i = 1, \dots, N$$

définir $t = T - 1$.

3. Calculer les poids lissés $\tilde{\omega}_{t|T}^{(i)}$ en utilisant les poids $\tilde{\omega}_t^{(i)}$ et les particules $\{x_t^{(i)}, x_{t-1}^{(i)}\}$ via les équations (4.29) et (4.30).

4. Définir $t = t - 1$. Si $t > 0$ retourner au pas 3, sinon terminer le processus.

4.3 Filtrage proposée pour multiples sources d'information

Dans la section 4.1.3, on a mentionné que, pour $t \geq 0$, l'ensemble de particules $\{x_t^{(i)}, \omega_t^{(i)}\}$ est une bonne approximation de la densité de filtrage, en outre, si on se trouve avec le problème de dégénérescence, une deuxième estimation de la densité de filtrage est reliée à l'ensemble de particules rééchantillonnées dont les particules sont équipondérées $\{x_t^{(i)}, \frac{1}{N}\}$.

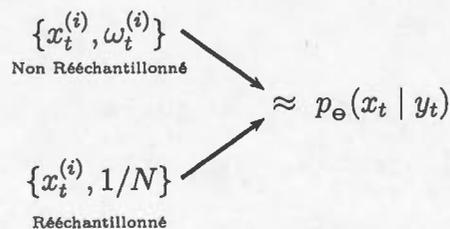


Figure 4.1 Estimation de la densité des particules

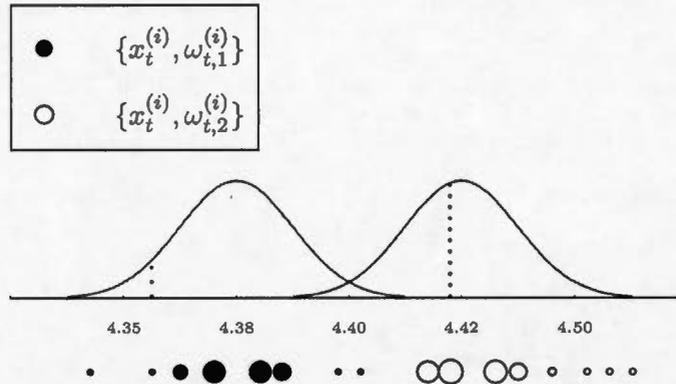
D'après la relation (3.6), la façon naturelle de calculer les poids d'importance lorsque l'on compte avec n sources d'information est le produit des densités de mesure de chaque source d'information. Les sources d'information sont représentées par $\mathbf{y}_t = \{y_{t,1}, \dots, y_{t,J}\}$

$$\omega_t^{(i)} = p_{\Theta}(\mathbf{y}_t | x_t^{(i)}) = \prod_{j=1}^J p_{\Theta}(y_{t,j} | x_t^{(i)}) \quad (4.31)$$

La manière dont on calcule les poids d'importance amène la discussion suivante. Pendant que l'on calcule le produit des densités de mesure, on élimine implicitement et systématiquement quelques particules qui peuvent améliorer le filtrage. On s'aperçoit que chacune de densités de mesure représente un véritable tamis, servant d'une manière indépendante à sélectionner les meilleurs particules qui décrivent la probabilité instantanée de défaut, mais lorsqu'on les combine (par le produit) le tamis résultant peut être trop rigoureux.

Pour mieux comprendre ce phénomène, considérons un exemple. Supposons qu'à l'instant t , on connaît huit particules $\{x_t^{(i)}\}$ et qu'elles sont ordonnées, en plus on suppose que l'on observe deux log-prix $y_{t,1}$ et $y_{t,2}$ qui proviennent d'une loi normale bivariée. Dans la figure (4.2), on observe à gauche la densité de mesure $p_{\Theta}(y_{t,1} | x_t^{(i)})$ et à droite celle de $p_{\Theta}(y_{t,2} | x_t^{(i)})$. Les points noirs (respec., blancs) sont des poids liés au sous-ensemble de particules $\{x_t^{(i)}, \omega_{t,j}^{(i)}\}$ qui proviennent de la densité de mesure de $y_{t,j}$, pour $j = 1, 2$.

Par rapport à la deuxième particule $x_t^{(2)}$ (de gauche à droite), on remarque que sous la densité de mesure de $y_{t,1}$ elle est négligeable, contrairement à ce que l'on observe sous la densité de mesure de $y_{t,2}$. Cela nous dit qu'on sélectionne les particules en fonction de chaque mesure, mais, en faisant le produit des mesures, le poids d'importance relié à cette particule deviendra nul dû à la première mesure donc à l'étape de rééchantillonnage la probabilité de la sélectionner est aussi nulle.



À gauche : La fonction de mesure de la source $y_{t,1}$. À droite : La fonction de mesure de la source $y_{t,2}$. Les points noirs (resp. blancs) représentent les poids d'importance sous la première (resp. deuxième) mesure

Figure 4.2 Maximun de vraisemblance et poids d'importance de particules

Avec l'objectif d'avoir une estimation de la densité de filtrage au temps t , $p_e(x_t | y_t)$ mais surtout de sélectionner des particules qui nous permettent une meilleure malléabilité en fonction des sources d'information de telle sorte que les particules sélectionnées donnent plus d'information sur cette densité, on propose une variation à la technique de filtrage dans l'étape rééchantillonnage.

D'habitude, à l'étape de rééchantillonnage du filtre à particules, on sélectionne N particules proportionnelles aux poids $\omega_t^{(i)}$. On propose de créer des ensembles à particules reliés à chaque source en leur donnant une pondération différente.

En premier, définissons $E_{t,j} = \{x_t^{(i)}, W_j \omega_{t,j}^{(i)}\}$ comme le sous-ensemble relié à la j -ème source au temps t , avec W_j la pondération assignée aux particules de la source d'information j , pour $j = 1, \dots, J$ et $t = 1, 2, \dots, N$ et avec $\omega_{t,j}^{(i)}$ les poids normalisés reliés au poids d'importance $\tilde{\omega}_{t,j}^{(i)} = p_e(y_{t,j} | x_t)$ au temps $t \geq 0$.

De plus, soit $\mathbf{E}_t = \{E_{t,j}\}_{j=1}^J$ l'ensemble de particules qui contient tous les sous-

ensembles à particules de chaque source $E_{t,j}$. Notons que $\sum_{j=1}^J \sum_{i=1}^N W_j \omega_{t,j}^{(i)} = 1$.

Finalement, on sélectionne N particules proportionnellement aux poids $W_j \omega_{t,j}^{(i)}$ de l'ensemble \mathbf{E}_t dont le nombre d'éléments est $N \times J$. La figure (4.3) illustre la modification proposée à la méthode de filtrage.

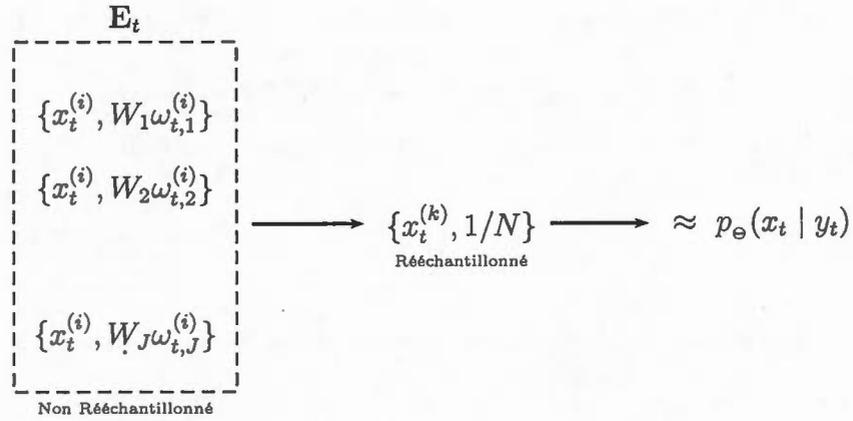


Figure 4.3 Nouvelle estimation de la densité à particule

CHAPITRE V

ESTIMATION DE PARAMÈTRES DE MODÈLES DE RISQUE DE CRÉDIT

L'estimation du vecteur de paramètres Θ implicite dans la représentation d'état-espace est un des sujets d'étude dans ce document. Ce chapitre présente deux méthodes d'estimation de paramètres, à savoir par maximum de vraisemblance et par espérance-maximisation.

5.1 Estimation de paramètres par maximum de vraisemblance

Des articles comme (Pitt, 2002) et (Creal, 2012) décrivent la manière d'estimer les paramètres via maximum de vraisemblance dans le cadre de modèles d'état-espace, plus précisément pour le filtre à particules. Soit $y_{1:T} = \{y_t, t = 1, 2, \dots, T\}$ les valeurs observées d'une série temporelle. Remarquons $\{y_t | x_t\}$ représente la dépendance de y_t par rapport à x_t avec x_t la variable d'état supposée markovienne d'ordre 1. On suppose aussi que la densité de mesure $p_{\Theta}(y_t | x_t)$ et la densité de transition $p_{\Theta}(x_t | x_{t-1})$ sont connues. L'objectif est alors d'estimer la fonction de log-vraisemblance pour un modèle d'état-espace :

$$\begin{aligned} l(\Theta) = \log L(\Theta) &= \log p_{\Theta}(\mathbf{y}_{1:T}) \\ &= \sum_{t=1}^T \log p_{\Theta}(y_t | \mathbf{y}_{1:t-1}) \end{aligned} \quad (5.1)$$

On va exploiter la relation, qui représente la contribution au log-vraisemblance.

D'après (Chen, 2003), on a deux manières de calculer cette contribution :

$$\begin{aligned} p_{\Theta}(y_t | \mathbf{y}_{1:t-1}) &= \int p_{\Theta}(y_t | x_t) p_{\Theta}(x_t | \mathbf{y}_{1:t-1}) dx_t \\ &= \int p_{\Theta}(y_t | x_{t-1}) p_{\Theta}(x_{t-1} | \mathbf{y}_{1:t-1}) dx_{t-1} \end{aligned}$$

La première égalité utilise la densité de prévision $p_{\Theta}(x_t | \mathbf{y}_{1:t-1})$ sachant la densité de filtrage au temps $t - 1$, $p_{\Theta}(x_{t-1} | \mathbf{y}_{1:t-1})$. La deuxième utilise directement la densité de filtrage au temps $t - 1$. Des estimations de la contribution à la vraisemblance sont :

$$\begin{aligned} \hat{p}_{\Theta}(y_t | \mathbf{y}_{1:t-1}) &= \sum_{i=1}^N p_{\Theta}(y_t | x_t^{(i)}) \tilde{\omega}_{t-1}^{(i)} \\ \hat{p}_{\Theta}(y_t | \mathbf{y}_{1:t-1}) &= \sum_{i=1}^N p_{\Theta}(y_t | x_{t-1}^{(i)}) \tilde{\omega}_{t-1}^{(i)} \end{aligned}$$

Dans notre cas, il est convenable de prendre la première estimation parce que l'on connaît $p_{\Theta}(y_t | x_t)$ de la représentation d'état-espace. Donc, la fonction de log-vraisemblance peut être calculée comme :

$$l(\Theta) \approx \sum_{t=1}^T \log \left[\sum_{i=1}^N p_{\Theta}(y_t | x_t^{(i)}) \tilde{\omega}_{t-1}^{(i)} \right] \quad (5.2)$$

Lorsque l'on fait le ré-échantillonnage dans toutes les étapes de $t = 1, 2, \dots, T$ on sait que $\omega_t^{(i)} = \frac{1}{N}$.

$$l(\Theta) \approx \sum_{t=1}^T \log \left[\frac{1}{N} \sum_{i=1}^N p_{\Theta}(y_t | x_t^{(i)}) \right] \quad (5.3)$$

5.2 Estimation de paramètres par espérance-maximisation

La méthode d'estimation espérance-maximisation (EM) trouve des estimations de paramètres du maximum de vraisemblance dans les problèmes où certaines

variables sont non-observées. Cette méthode est considérée comme une généralisation de la méthode de maximum de vraisemblance traditionnelle et le but est d'estimer les paramètres du modèle de manière itérative, à partir d'une certaine estimation initiale .

D'après (Neal et Hinton, 1998), chaque itération consiste en deux étapes : espérance (E) et maximisation (M). Dans la première étape, on trouve la distribution des variables non-observées étant donné les valeurs connues (les variables observées) et les estimations actuelles des paramètres. Dans la deuxième étape, on ré-estime les paramètres avec la méthode de maximum de vraisemblance, en supposant que la distribution trouvée dans l'étape (E) est correcte.

À partir de la fonction de log-vraisemblance, on peut déduire la borne inférieure suivante qui nous permet d'introduire les estimations de façon itérative.

Soit $I_i \subseteq \mathbb{R}$ sous-ensemble un ouvert borné de la variable d'état x_i pour $i = 1, 2, \dots, T$ et $\chi = I_0 \times I_1 \times I_2 \times \dots \times I_T \subseteq \mathbb{R}^{T+1}$ un sous-ensemble ouvert borné des variables d'état $\mathbf{x}_{0:T}$.

$$\begin{aligned}
l(\Theta) &= \log p_{\Theta}(\mathbf{y}_{1:T}) \\
&= \log \int_{\mathcal{X}} p_{\Theta}(\mathbf{y}_{1:T}, \mathbf{x}_{1:T}) d\mathbf{x}_{1:T} \\
&= \log \int_{\mathcal{X}} U_{\Theta_k}(\mathbf{x}_{0:T}) \frac{p_{\Theta}(\mathbf{y}_{1:T}, \mathbf{x}_{1:T})}{U_{\Theta_k}(\mathbf{x}_{0:T} | \theta_k)} d\mathbf{x}_{1:T} \\
&= \log \left[\mathbb{E}^U \left(\frac{p_{\Theta}(\mathbf{y}_{1:T}, \mathbf{x}_{1:T})}{U_{\Theta_k}(\mathbf{x}_{0:T})} \right) \right] \\
&\stackrel{\text{Jensen}}{\geq} \mathbb{E}^U \left[\log \left(\frac{p_{\Theta}(\mathbf{y}_{1:T}, \mathbf{x}_{1:T})}{U_{\Theta_k}(\mathbf{x}_{0:T})} \right) \right] \\
&= \int_{\mathcal{X}} U_{\Theta_k}(\mathbf{x}_{0:T}) \log \left(\frac{p_{\Theta}(\mathbf{y}_{1:T}, \mathbf{x}_{1:T})}{U_{\Theta_k}(\mathbf{x}_{0:T})} \right) d\mathbf{x}_{1:T} \\
&= \int_{\mathcal{X}} U_{\Theta_k}(\mathbf{x}_{0:T}) \log p_{\Theta}(\mathbf{y}_{1:T}, \mathbf{x}_{1:T}) d\mathbf{x}_{1:T} - \int_{\mathcal{X}} U_{\Theta_k}(\mathbf{x}_{0:T}) \log U_{\Theta_k}(\mathbf{x}_{0:T}) d\mathbf{x}_{1:T} \\
&= \int_{\mathcal{X}} U_{\Theta_k}(\mathbf{x}_{0:T}) \log p_{\Theta}(\mathbf{y}_{1:T}, \mathbf{x}_{1:T}) d\mathbf{x}_{1:T} + H(U_{\Theta_k}(\mathbf{x}_{0:T})) \tag{5.4}
\end{aligned}$$

Premièrement, notons que l'inégalité précédente est obtenue grâce à l'inégalité de Jensen ; deuxièmement, remarquons que dans le résultat obtenu la fonction $H(U_{\Theta_k}(\mathbf{x}_{0:T}))$ ne dépend pas Θ , donc on ne la considérera pas dans l'algorithme de maximisation (EM) ; finalement, on peut montrer que le maximum dans l'étape (E) est atteint lorsque la fonction U est exactement la densité de filtrage paramétrisée en Θ_k , c'est-à-dire $U_{\Theta_k}(\mathbf{x}_{0:T}) = p_{\Theta_k}(\mathbf{x}_{0:T} | \mathbf{y}_{1:T})$ tel que mentionné dans l'article (Roweis et Ghahramani, 2001). On définit l'espérance conditionnelle comme :

$$\begin{aligned}
Q(\Theta, \Theta_k) &= \int_{\mathcal{X}} U_{\Theta_k}(\mathbf{x}_{0:T} | \theta_k) \log p_{\Theta}(\mathbf{y}_{1:T}, \mathbf{x}_{1:T}) d\mathbf{x}_{1:T} \\
&= \mathbb{E}^U [\log p_{\Theta}(\mathbf{y}_{1:T}, \mathbf{x}_{1:T}) | \Theta_k, \mathbf{y}_{1:T}] \tag{5.5}
\end{aligned}$$

De cette façon, dans l'étape (E) de l'algorithme, il faut calculer la fonction $Q(\Theta, \Theta_k)$ et dans l'étape (M) il faut la maximiser par rapport à Θ pour obtenir une nouvelle estimation du paramètre dans l'itération k , Θ_{k+1} . En peu de mots :

Étape E : Calculer $Q(\Theta, \Theta_k)$

Étape M : $\Theta_{k+1} = \arg \max_{\Theta} Q(\Theta, \Theta_k)$

Le premier défi est de calculer $Q(\Theta, \Theta_k)$. Commençons par voir que l'on a supposé que x_t est un processus markovien d'ordre 1 et que y_t ne dépend que de x_t , donc on peut réécrire $Q(\Theta, \Theta_k)$ (voir l'appendice B) comme :

$$Q(\Theta, \Theta_k) = I_1(\Theta, \Theta_k) + I_2(\Theta, \Theta_k) + I_3(\Theta, \Theta_k) \quad (5.6)$$

avec

$$\begin{aligned} I_1(\Theta, \Theta_k) &= \int p_{\Theta_k}(x_0 | y_{1:T}) \log p_{\Theta}(x_0) dx_0 \\ &= \mathbb{E}[\log p_{\Theta}(x_0) | y_{1:T}, \Theta_k] \end{aligned}$$

$$\begin{aligned} I_2(\Theta, \Theta_k) &= \int p_{\Theta_k}(x_t | y_{1:T}) \log p_{\Theta}(y_t | x_t) dx_t \\ &= \sum_{t=1}^T \mathbb{E}[\log p_{\Theta}(y_t | x_t) | y_{1:T}, \Theta_k] \end{aligned}$$

$$\begin{aligned} I_3(\Theta, \Theta_k) &= \int \int p_{\Theta_k}(x_t, x_{t-1} | y_{1:T}) \log p_{\Theta}(x_t | x_{t-1}) dx_t dx_{t-1} \\ &= \sum_{t=1}^T \mathbb{E}[\log p_{\Theta}(x_t | x_{t-1}) | y_{1:T}, \Theta_k] \end{aligned}$$

Afin d'estimer $Q(\Theta, \Theta_k)$, on a besoin de la densité de lissage $p_{\Theta_k}(x_t | y_{1:T})$ et la densité de lissage conjointe d'une paire de particules, $p_{\Theta_k}(x_t, x_{t-1} | y_{1:T})$, mais ces densités ne sont pas disponibles analytiquement, en conséquence on a besoin d'une approximation. D'après l'article de (Kokkala *et al.*, 2014), cette approximation peut être fait de deux façons, à savoir le lissage sigma-point et le lissage de particules. La technique de lissage de particules a été proposée par (Schön *et al.*, 2011)

dont les estimations de I_1 , I_2 et I_3 sont données par les expressions suivantes :

$$\begin{aligned}\hat{I}_1(\Theta, \Theta_k) &= \sum_{i=1}^N w_{0|T}^{(i)} \log p_{\Theta}(x_0^{(i)}) \\ \hat{I}_2(\Theta, \Theta_k) &= \sum_{t=0}^{T-1} \sum_{i=1}^N \sum_{j=1}^N w_{t|T}^{(ij)} \log p_{\Theta}(x_{t+1}^{(j)} | x_t^{(i)}) \\ \hat{I}_3(\Theta, \Theta_k) &= \sum_{t=1}^T \sum_{i=1}^N w_{t|T}^{(i)} \log p_{\Theta}(y_t | x_t^{(i)})\end{aligned}$$

Tous les poids et les particules de cette approche sont calculés dans les sections (4.1.3) et (4.2). Il est important de noter que tous les calculs sont faits en utilisant la représentation d'état-espace paramétrisée par Θ_k . Quant à l'étape de maximisation, on utilise la technique descente de gradient.

On a finalement l'algorithme d'espérance-maximisation qui nous permet d'estimer le vecteur de paramètres Θ_k d'une manière itérative.

Algorithme d'espérance-maximisation (EM)

1. Assigner $k = 1$, Θ_k et α_k
2. Obtenir $\{x_t^{(i)}, w_t^{(i)}\}$ pour $i = 1, 2, \dots, N$ avec le filtre à particules.
3. Calculer $Q(\Theta, \Theta_k)$
4. Calculer le gradient $g_k = \nabla Q(\Theta, \Theta_k) |_{\Theta=\Theta_k}$
 - (a) si $Q(\Theta_k - \alpha_1 g_k, \Theta_k) > Q(\Theta_k, \Theta_k)$ poser $\alpha_{k+1} = 2^k \alpha_k$
 - (b) sinon $\alpha_{k+1} = \alpha_k / 2^k$
5. Assigner $k = k + 1$ jusqu'à convergence.

Les critères de convergence peuvent être sur un point critique ou bien sur une valeur de Θ_k et selon un critère de tolérance.

CHAPITRE VI

ANALYSE PAR SIMULATION

Dans le présent chapitre, on analyse les méthodes d'estimation de paramètres et la performance de la méthode de filtrage dans le cadre de l'approche structurelle et celle à forme réduite, via des simulations.

On commence par la définition de paramètres que l'on utilisera dans les simulations et ensuite on définira la mesure de performance.

Sous l'approche structurelle, on s'intéresse à savoir quels sont les attributs des fonds propres qui nous renseignent le plus à propos de la valeur des actifs. Le modèle de Merton et le modèle à taux d'intérêt stochastique sont utilisés pour obtenir ces valeurs.

De même, sous l'approche à forme réduite, on analyse la performance du filtrage et les attributs de deux instruments financiers. On s'intéresse à savoir lequel d'entre eux nous renseigne le plus à propos des intensités. Premièrement, on analyse les attributs des obligations corporatives risquées, puis ceux des CDS et finalement on analyse le filtrage résultant en mettant les deux instruments ensemble. Dans chaque analyse, on évalue le pourcentage de gain ou de perte de faire le filtrage en utilisant conjointement plusieurs sources d'information plutôt que d'utiliser seulement une. On analyse aussi l'effet d'inclure des sauts dans les intensités, et

finalement, on compare la méthode de filtrage proposée par rapport à la méthode standard de filtrage dans le cadre de plusieurs sources d'information.

Dans la dernière partie de ce chapitre, on analyse les propriétés des deux méthodes d'estimation de paramètres, à savoir maximum de vraisemblance (MV) et espérance-maximisation (EM).

6.1 Analyse du filtrage

6.1.1 Paramètres du modèle structurel

On s'appuie sur les représentations d'état-espace définies dans le chapitre 3 pour simuler la valeur des actifs V_t et la valeur de fonds propres S_t . Ces valeurs sont calculées sous la mesure risque-neutre \mathbb{Q} .

En ce qui concerne la valeur des actifs des deux modèles, on définit sa valeur initiale comme $V_0 = 100$ et sa volatilité en trois niveaux $\sigma = 0.10, 0.16, 0.20$. Le taux d'intérêt sans risque du modèle de Merton est $r = 0.05$. On fixe deux des trois paramètres de la dynamique du modèle à taux d'intérêt stochastique, r_t , ainsi $\kappa_r = 0.9$, $\theta_r = 0.05$ et on fait varier sa volatilité $\sigma_r = 0.001, 0.01, 0.3$. Encore sous ce dernier modèle, on suppose que la corrélation entre la valeur des actifs et le taux d'intérêt est nulle, $\rho = 0$.

En plus, on suppose que la compagnie émet quatre obligations zéro coupon de maturités $T_i = 7, 10, 15, 30$, respectivement. La valeur faciale est la même pour les quatre obligations, $L = 80$ ¹.

La valeur des fonds propres S_t , qui est fonction de V_t , L et T_i , peut prendre la forme de l'équation (2.12) ou bien de l'équation (2.15), selon le modèle choisi pour la représenter. En plus, on suppose que S_t est observée avec un bruit transactionnel

1. Notons que le choix de la valeur faciale ne change pas la valeur des actifs V_t

de $\delta = 0.001, 0.01, 0.1$.

6.1.2 Paramètres du modèle à forme réduite

Tout d'abord, on va définir les paramètres des instruments financiers, ensuite, on va définir les paramètres des processus décrivant les intensités.

On dispose de quatre obligations corporatives risquée à coupons avec temps de maturités $T_o = 7, 10, 15, 30$ et six CDS avec temps de maturité $T_c = 0.5, 1, 3, 5, 7, 10$. Pour les deux instruments financiers, on définit un taux d'intérêt sans risque de $r = 0.02$ et un pourcentage de recouvrement de $R_{EC} = 0.40$.

Les quatre obligations risquée ont le même taux de coupon semi-annuel de $c = 0.1$ et la même valeur faciale $L = 100$. En plus, afin de calculer les CDS, on suppose que l'acheteur de protection fait des paiements trimestriels et que le vendeur de protection suppose une partition mensuelle pour calculer leur *payoff*, ainsi $\Delta(T_{i-1}, T_i) = \frac{1}{12}$.

On suppose que les instruments financiers sont observés avec un bruit transactionnel de $\delta = 0.01, 0.1, 0.5$, bien que des bruits de l'ordre de 0.5 soient peu réalistes, on explore et analyse ce cas extrême afin de déterminer le pourcentage de gain ou perte d'utiliser plusieurs sources d'information.

Notons que le prix d'une obligation risquée à coupons est appelé « *dirty price* » lorsqu'il inclut les intérêts courus du paiement du coupon suivant. On fait le filtrage sur le « *clean price* » qui n'inclut pas les intérêts courus.

Jusqu'ici, les paramètres reliés aux instruments financiers ont été définis, il nous reste les paramètres des processus d'intensités. Dans le tableau 6.1, on considère quatre scénarios de paramètres du processus CIR où la valeur initiale des intensités est $\lambda_0 = 0.03$ pour chacun d'eux.

Tableau 6.1 Scénarios des paramètres des intensités CIR

	κ_r	θ_r	σ_r
Scénario 1	0.3	0.01	0.02
Scénario 2	0.3	0.05	0.02
Scénario 3	0.8	0.01	0.02
Scénario 4	0.8	0.05	0.02

Quant aux paramètres du processus d'intensité avec sauts (JCIR), on fixe ceux de la partie de diffusion comme $\lambda_0 = 0.03$, $\kappa_r = 0.8$, $\theta_r = 0.03$ et $\sigma_r = 0.02$ mais on fait varier les paramètres de taux d'arrivée (fréquence) de sauts α et de la magnitude des sauts γ tel que montré dans les quatre scénarios du tableau 6.2.

Tableau 6.2 Scénarios des paramètres des intensités avec sauts - JCIR

	α	γ
Scénario 1	0.2	0.005
Scénario 2	0.2	0.02
Scénario 3	0.5	0.005
Scénario 4	0.5	0.02

6.1.3 Paramètres de la méthode de filtrage proposée

Le seul paramètre à modifier par rapport à la méthode standard de filtrage est la pondération W_i assignée à chaque source d'information qui est directement relié aux ensembles à particules. En effet, on évalue trois possibilités : dans la première, on assigne la même pondération aux sources, dans la deuxième on assigne une pondération plus grande aux sources qui ont une maturité plus longue, et dans la troisième on assigne une pondération plus petite aux sources qui ont une maturité plus longue.

Par exemple, si on compte avec quatre obligations risquées à coupons de maturités $T_o = 7, 10, 15, 30$, dans le deuxième cas les pondérations respectives seront : $W_1 = \frac{1}{10}$, $W_2 = \frac{2}{10}$, $W_3 = \frac{3}{10}$ et $W_4 = \frac{4}{10}$. Notons que le dénominateur est la somme des numérateurs. On assigne les pondérations aux six CDS d'une manière très semblable, étant le dénominateur $\sum_{i=1}^6 i = 21$. Les tableaux 6.3 et 6.4 montrent les pondérations assignées à chacun des produits financiers.

Tableau 6.3 Pondération assignée aux obligations risquées à coupons pour la méthode de filtrage proposée

	$T_o = 7$	$T_o = 10$	$T_o = 15$	$T_o = 30$
	W_1	W_2	W_3	W_4
<i>Prop 1</i>	0.25	0.25	0.25	0.25
<i>Prop 2</i>	0.10	0.20	0.30	0.40
<i>Prop 3</i>	0.40	0.30	0.20	0.10

Tableau 6.4 Pondération assignée aux CDS pour la méthode de filtrage proposée

	$T_c = 0.5$	$T_c = 1$	$T_c = 3$	$T_c = 5$	$T_c = 7$	$T_c = 10$
	W_1	W_2	W_3	W_4	W_5	W_6
<i>Prop 1</i>	0.17	0.17	0.17	0.17	0.17	0.17
<i>Prop 2</i>	0.05	0.10	0.14	0.19	0.24	0.29
<i>Prop 3</i>	0.29	0.24	0.19	0.14	0.10	0.05

Lorsqu'on met les deux produits financiers ensemble on aura 10 sources d'information observées. Donc, les pondérations pour *Prop 2* seront de la forme $W_{i,1} = \frac{i}{M}$ et $W_{j,2} = \frac{j}{M}$ pour $i = 1, 2, 3, 4$ et $j = 1, 2, \dots, 6$ qui correspondent aux quatre obligations risquées à coupons de maturités $T_o = 7, 10, 15, 30$ et aux six CDS de maturités $T_c = 0.5, 1, 3, 5, 7, 10$. Pour *Prop 3* les pondérations seront

$W_{i,1} = \frac{4-i+1}{M}$ et $W_{j,2} = \frac{6-j+1}{M}$. Notons que $M = \sum_{i=1}^4 i + \sum_{j=2}^6 j = 31$. Bien évidemment $W_{i,1} = W_{j,2} = \frac{1}{10}$ pour *Prop* 1.

6.1.4 Mesures de la performance du filtrage

Sans perte de généralité, soit X_t une variable latente qui peut prendre la forme de la valeur des actifs V_t sous l'approche de modèles structurels ou bien la forme des intensités λ_t sous l'approches de modèles à forme réduite.

On suppose que le nombre de jours ouvrables par mois est 21, faisant une année de 252 jours. On génère 200 trajectoires journalières des variables latentes $\{X_{t,k}\}_{k=1}^{200}$ dans une période d'observation de T^* ans, donc $t = 1, 2, \dots, 252T^*$.

À partir de chacune des trajectoires de variables latentes simulées, on calcule la valeur de fonds propres S_t sous l'approche structurelle ou bien des deux instruments financiers (obligations risquées et CDS) sous l'approche à forme réduite, qui incluent le bruit transactionnel δ .

Ensuite, on obtient un ensemble de variables latentes estimées $\{\hat{X}_{t,k}\}_{k=1}^{200}$ qui sont extraites en prenant soit les fonds propres ou bien les instruments financiers comme étant observés, tout ceci en utilisant la méthode de filtre à particules.

Donc, la mesure de la performance du filtrage est donnée par la racine de l'erreur quadratique moyenne (REQM) :

$$REQM_k = \sqrt{\frac{1}{252T^* - 1} \sum_{t=1}^{252T^*} (X_{t,k} - \hat{X}_{t,k})^2} \quad (6.1)$$

pour $k = 1, 2, \dots, 200$ trajectoires.

6.1.5 Performance du filtrage des modèles structurels

Le filtrage de la valeur estimée des actifs \widehat{V}_t est obtenue avec $N = 1000$ particules. On définit trois niveaux de volatilité de la valeur des actifs $\sigma = 0.10, 0.16, 0.20$ et trois niveaux de bruit transactionnel $\delta = 0.001, 0.01, 0.1$. La combinaison des niveaux représentent ainsi neuf scénarios que l'on va analyser.

Dans chacun des scénarios, on calcule la moyenne de la REQM des quatre obligations zéro-coupon de maturités $T_i = 7, 10, 15, 30$. On fait un test F de Fisher pour comparer ces moyennes. Le tableau 6.5 rapporte les résultats par maturité, volatilité des actifs et bruit transactionnel.

On remarque qu'il n'y a pas de différences significative entre les moyennes des REQM dans les trois scénarios où le bruit transactionnel est de l'ordre de $\delta = 0.001$, c'est-à-dire que les quatre obligations zéro-coupon fournissent la même information de la valeur des actifs, ce qui est très raisonnable parce que le bruit est assez petit.

Par contre, on trouve de différences significatives dans les six scénarios restants où le bruit transactionnel est $\delta = 0.01$ et 0.1 . On peut affirmer, avec un niveau de confiance de 95%, que les obligations qui nous offrent plus d'information à propos des actifs sont celles de maturités plus courtes.

Tableau 6.5 Moyennes des REQM par maturité, bruit transactionnel et volatilité de la valeur des actifs du modèle de Merton

Bruit trans.	Maturité	σ		
		0.10	0.16	0.20
$\delta = 0.001$	$T_1 = 7$	0.6375	1.0199	1.2754
	$T_2 = 10$	0.6361	1.0187	1.2750
	$T_3 = 15$	0.6443	1.0339	1.2963
	$T_4 = 30$	0.6439	1.0303	1.2900
$\delta = 0.01$	$T_1 = 7$	0.7172*	1.0942*	1.3494*
	$T_2 = 10$	0.7366	1.1129	1.3688
	$T_3 = 15$	0.7721	1.1534	1.4139
	$T_4 = 30$	0.8274	1.2004	1.4536
$\delta = 0.1$	$T_1 = 7$	1.5780*	2.1802*	2.5793*
	$T_2 = 10$	1.6970	2.3366	2.7548
	$T_3 = 15$	1.8468	2.5260	2.9651
	$T_4 = 30$	2.0854	2.7776	3.2083

*Le calcul est basé sur 200 trajectoires et $N = 1000$ particules. Les valeurs avec * sont celles dont la moyenne de la REQM est significativement plus petite avec un niveau de confiance de 95%*

À l'égard du modèle de taux d'intérêt stochastique, on a configuré des scénarios en fonction du paramètre de la volatilité du taux $\sigma_r = 0.001, 0.01, 0.03$ et du bruit $\delta = 0.001, 0.01, 0.1$. Quatre obligations zéro-coupon de maturités $T_i = 7, 10, 15, 30$ sont disponibles dans chacun des scénarios.

Le tableau 6.6 rapporte la moyenne de la REQM par maturité, volatilité du taux d'intérêt et bruit transactionnel δ . D'après le test F de Fisher, on trouve deux

scénarios extrêmes où les moyennes ne sont pas statistiquement différentes.

Le premier de ces scénarios, $[\delta = 0.001, \sigma_r = 0.001]$, représente celui dont la performance du filtrage est la meilleure avec la REQM d'environ 0.0822. On ne trouve pas de différences significatives parce que le bruit et la volatilité du taux sont assez petits et donc le filtrage estime d'une manière précise la valeur des actifs. Le deuxième scénario, $[\delta = 0.1, \sigma_r = 0.3]$, est celui où la REQM est la plus élevée (d'environ 2.773), ainsi les obligations fournissent la même information mais de mauvaise qualité.

D'un autre côté, notons que l'on a configuré les paramètres du taux stochastique de telle sorte que les scénarios où $\sigma_r = 0.001$ ressemblent aux scénarios du modèle de Merton. C'est pourquoi on affirme que, sauf lorsque le bruit est de l'ordre de $\delta = 0.001$, les obligations de maturité plus courte offrent plus d'information de la valeur des actifs, ce qui est cohérent avec les résultats de la performance du modèle de Merton.

En revanche, les quatre scénarios, qui résultent des combinaisons des niveaux suivants : $[\delta = 0.001, 0.01] \times [\sigma_r = 0.01, 0.03]$, nous suggèrent qu'avec des obligations de plus longue maturité le filtrage est significativement meilleure.

Tableau 6.6 Moyennes des REQ_M par maturité, bruit transactionnel et volatilité du taux d'intérêt du modèle à taux d'intérêt stochastique

Bruit trans.	σ_r		0.001	0.01	0.03
	Maturité				
$\delta = 0.001$	$T_1 = 7$		0.0836	0.7090	2.0789
	$T_2 = 10$		0.0806	0.6224	1.8195
	$T_3 = 15$		0.0808	0.5436	1.6075
	$T_4 = 30$		0.0837	0.3455*	1.0395*
$\delta = 0.01$	$T_1 = 7$		0.4135*	0.8001	2.0557
	$T_2 = 10$		0.4634	0.7552	1.8093
	$T_3 = 15$		0.5287	0.7481	1.6383
	$T_4 = 30$		0.6275	0.7169*	1.1892*
$\delta = 0.1$	$T_1 = 7$		1.9736*	2.0896	2.7644
	$T_2 = 10$		2.1385	2.2110	2.6900
	$T_3 = 15$		2.3526	2.4111	2.7901
	$T_4 = 30$		2.6650	2.6921	2.8473

*Le calcul est basé sur 200 trajectoires et $N = 1000$ particules. Les valeurs avec * sont celles dont la moyenne de la REQ_M est significativement plus petite avec un niveau de confiance de 95%*

Bien évidemment, dans les deux modèles, on constate le fait que la performance du filtrage de la valeur des actifs est de mauvaise qualité lorsque le bruit transactionnelle δ , la volatilité σ et/ou σ_r augmentent.

Globalement, les fonds propres qui sont reliés aux obligations zéro-coupons de maturités plus courtes fournissent plus information de la valeur des actifs sous le modèle de Merton, tandis que les fonds propres qui sont reliés aux obligations zéro-

coupons de maturités plus longues fournissent plus information sous le modèle à taux d'intérêt stochastique.

6.1.6 Performance du filtrage de modèle à forme réduite

Dans cette analyse, on dispose de quatre obligations risquées à coupons de maturités $T_o = 7, 10, 15, 30$ et de six CDS de maturités $T_c = 0.5, 1, 3, 5, 7, 10$. On considère ces dix instruments financiers comme une source d'information indépendante, mais contenant la même information sur les intensités de défaut λ_t étant donné que les sources d'information appartiennent à la même compagnie. Ainsi, on est capable d'estimer les intensités soit en utilisant une source à la fois ou bien en utilisant plusieurs sources simultanément.

Comme on l'a déjà mentionné, on divise cette analyse en trois parties. Premièrement, on montre la performance du filtrage résultante des obligations risquées à coupons, ensuite on analyse la performance des CDS et finalement on met ensemble les deux types d'instruments financiers. La discussion est basée sur la mesure de performance REQM, tout ceci en considérant les scénarios définis dans les tableaux 6.1 et 6.2.

Dans le cadre de plusieurs sources d'information, on compare la performance de la méthode de filtrage standard par rapport à la méthode de filtrage proposée (voir 4.3).

Remarquons que dans tous les scénarios, on définit trois niveaux de bruit transactionnel $\delta = 0.01, 0.1, 0.5$ faisant 12 groupes par instrument.

6.1.6.1 Performance du filtrage des obligations risquées à coupons

Les résultats de la performance du filtrage des obligations risquées à coupons se trouvent dans l'annexe C. Dans cette section, on souligne les résultats les plus

importants.

D'après le test F de Fisher, les moyennes des REQM, calculées sur chaque source d'information de ce produit financier, ne sont pas statistiquement différentes dans les scénarios 3 et 4 du tableau 6.1, mais dans les scénarios 1 et 2, où la vitesse de retour à la moyenne des intensités est $\kappa = 0.3$, les obligations de maturité plus longue donnent significativement plus d'information sur les intensités par rapport à celles de maturité courte.

De manière globale, si le bruit transactionnel s'élève de $\delta = 0.01$ à $\delta = 0.1$, on constate une augmentation de la moyenne de la REQM d'environ 2.23 fois. S'il s'élève de $\delta = 0.01$ à $\delta = 0.5$ l'augmentation est d'environ 2.68 fois. Ce dernier nous donne une idée de la robustesse de la méthode de filtrage appliquée sur ce type de produit financier.

Par ailleurs, on trouve un gain statistiquement significatif, si l'on fait le filtrage en considérant plusieurs sources d'information conjointement.

En particulier, sous le scénario 1 avec des intensités sans sauts (voir le tableau C.1 de l'annexe C), si le bruit transactionnel est $\delta = 0.01$, les REQM moyennes des obligations risquées de maturités $T_o = 7, 10, 15, 30$ sont respectivement : 8.67×10^4 , 8.27×10^4 , 7.92×10^4 et 7.81×10^4 . En faisant le filtrage avec les quatre obligations conjointement, la REQM moyenne est égale à 5.75×10^4 (voir le tableau C.5 de l'annexe C) en obtenant un gain d'environ 29.52%.

Encore sous le scénario 1, le gain de faire le filtrage à partir de plusieurs obligations risquées conjointement est 22.86% si le bruit est $\delta = 0.1$. Finalement, si le bruit transactionnel est de l'ordre de $\delta = 0.5$, le gain est tout juste de 3.77%.

Les résultats aussi suggèrent que le gain de filtrage diminue dans la mesure où le bruit transactionnel augmente, en d'autres mots, si les quatre obligations risquées

observées contiennent beaucoup de bruit (par exemple $\delta = 0.5$), les quatre REQMs des respectives sources sont similaires à la REQM calculée en considérant toutes les sources conjointement.

Le gain moyen en pourcentage de faire le filtrage des intensités en considérant conjointement les quatre obligations risquées plutôt que de le faire individuellement est donné dans le tableau 6.7. Les résultats sont montrés par scénarios et bruit transactionnelle.

Tableau 6.7 Gain en pourcentage des 4 obligations conjointement par scénario et bruit transactionnel : Intensités sans sauts.

	Scénario 1	Scénario 2	Scénario 3	Scénario 4
$\delta = 0.01$	29.52%	29.65%	26.14%	27.03%
$\delta = 0.1$	22.86%	23.45%	9.24%	11.91%
$\delta = 0.5$	3.77%	8.21%	1.65%	3.19%

Quant à la méthode de filtrage proposée, on remarque qu'elle n'améliore pas significativement la méthode standard de filtrage lorsque l'on utilise de multiples sources d'information de type de produits financiers.

Comme illustration de ce fait, la figure 6.1 montre des distributions des REQMs du scénario 1 selon le bruit transactionnelle sous la forme de *Boxplot*. Dans chaque panel de cette figure, les quatre premiers *Boxplots* de gauche à droite correspondent aux distributions reliés aux obligations de maturité $T_o = 5, 10, 15, 30$. Les quatre derniers *Boxplots* sont reliés aux distributions de la REQM lorsqu'on utilise les quatre sources simultanément dont la première d'entre elles correspond à la méthode standard de filtrage et les trois restantes à la méthode de filtrage proposée.

Rappelons que l'étiquette *Prop 1* signifie que les pondérations assignées aux quatre

sources sont les mêmes, de plus la distribution sous l'étiquette *Prop 2* représente une assignation des pondérations plus grandes aux sources de maturités plus longues et finalement la distribution sous l'étiquette *Prop 3* représente une assignation des pondérations plus grandes aux sources de maturités plus courtes (voir les tableaux 6.3 et 6.4 pour plus de détails).

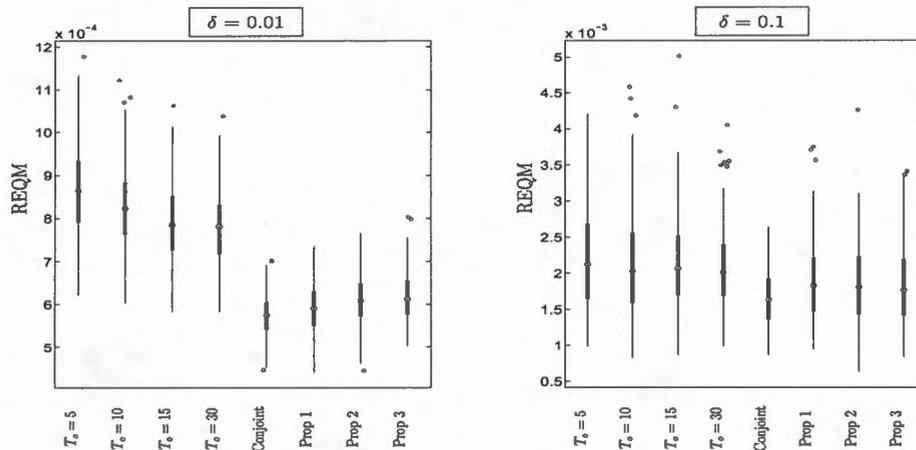


Figure 6.1 Boxplot pour les REQM pour le scénario 1 CIR Obligations

Les distributions empiriques des REQM sont symétriques dans tous les scénarios avec intensités sans sauts. En revanche, on constate que si les intensités présentent des sauts, ces distributions seront à queues lourdes parce que, dans certains cas, les particules ne sont pas capables de capturer avec précision le changement de niveau généré par la magnitude du saut.

Dans tous les scénarios avec des intensités avec sauts, les moyennes des REQM des obligations risquées à coupons ne varient pas significativement en fonction des maturités. Le résultats sont disponibles dans le tableau C.2 de l'annexe C. Notons que les mêmes considérations qui ont été faites ci-dessus quant à la figure 6.1 s'appliquent à la figure 6.2 dont on considère des intensités avec sauts .

Le gain/perte moyenne en pourcentage de faire le filtrage des intensités avec saut

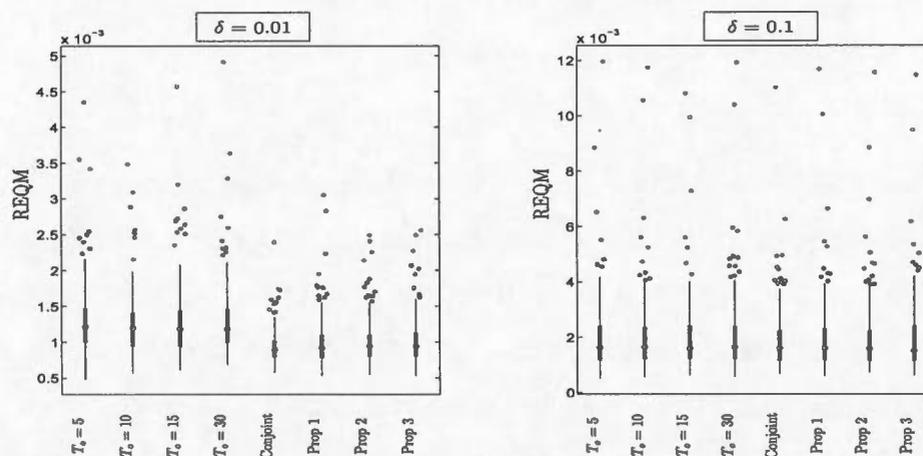


Figure 6.2 Boxplot pour les REQM pour le scénario 1 CIR Obligations avec sauts

en considérant les quatre obligations risquées de manière conjointe est donné dans le tableau 6.8.

Les résultats de la simulation portant sur les scénario 1 et 3, où la magnitude des sauts est relativement basse nous permettent de souligner que dans le cas extrême où le bruit transactionnel était de l'ordre de 50%, il est préférable de faire le filtrage des intensités avec les obligations individuellement, bien qu'en pratique, il pourrait être peu probable de trouver un bruit de cet ordre.

Tableau 6.8 Gain en pourcentage des 4 obligations conjointement par scénario et bruit transactionnel : Intensités avec sauts.

	Scénario 1	Scénario 2	Scénario 3	Scénario 4
$\delta = 0.01$	25.16%	24.13%	26.2%	28.75%
$\delta = 0.1$	1.18%	15.51%	3.27%	17.85%
$\delta = 0.5$	-3.09%	8.21%	-3.37%	11.17%

6.1.6.2 Performance du filtrage des CDS

Par rapport à ce type de dérivé de crédit, on trouve que, sous les scénarios sans sauts, la moyenne de la REQM est significativement inférieure quand les maturités sont plus courtes. De plus, on ne trouve pas de différences significatives des moyennes des REQM entre des CDS de maturités 0.5 et 1, signifiant ces deux sources fournissent statistiquement la même information à propos des intensités. Le résultats de ces simulations se trouvent dans le tableau C.3 de l'annexe C.

On observe une augmentation de la moyenne de la REQM si le bruit transactionnel augmente. En effet, les résultats de la simulation nous montrent que si le bruit transactionnel passe de 0.01 à 0.1, on s'attend à une augmentation d'environ 3.28 fois sur la moyenne de la REQM et si on passe de 0.01 à 0.5, une augmentation d'environ 6.19 fois.

D'un autre côté, on obtient une amélioration significative sur la qualité du filtrage lorsque l'on met ensemble les 6 CDS de différentes maturités avec ce dérivé de crédit. En effet, en considérant les quatre scénarios, la qualité du filtrage s'améliore d'environ 25.3% par rapport aux maturités 6 mois et 1 an, et d'environ 50.1% par rapport aux maturités de 7 et 10 ans. De plus, on remarque, encore une fois, que la méthode de filtrage proposée est statistiquement égale à la méthode standard de filtrage dans le cadre de multiples sources d'information avec des intensités avec sauts. La figure 6.3 montre les résultats obtenus pour le scénario 1.

Le gain de faire le filtrage des intensités en considérant conjointement les six CDS par rapport au filtrage de chacune des CDS est donné dans les tableau 6.9. Ces résultats sont montrés par scénario et bruit transactionnelle. Si $\delta = 0.01$ le gain est d'environ 49%, et d'environ 38% et 30% si $\delta = 0.1$ ou $\delta = 0.5$, respectivement.

Comme c'était le cas avec les obligations risquées, sous les scénarios avec sauts

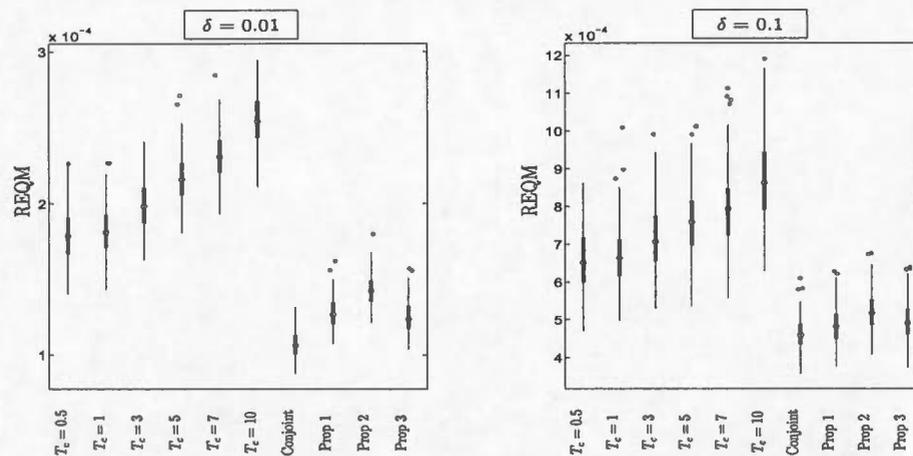


Figure 6.3 Boxplot pour les REQM pour le scénario 1 CIR CDS

Tableau 6.9 Gain en pourcentage d'utiliser les six CDS conjointement par scénario et bruit transactionnel : Intensités sans sauts.

	Scénario 1	Scénario 2	Scénario 3	Scénario 4
$\delta = 0.01$	48.64%	47.23%	50.45%	48.83%
$\delta = 0.1$	37.36%	38.72%	39.77%	39.62%
$\delta = 0.5$	33.11%	33.29%	31.59%	23.73%

dans les intensités, les distributions des REQMs sont à queues lourdes. Pareillement aux scénarios sans sauts dans les intensités, les CDS de maturité plus courte nous renseigne le plus sur le filtrage, bien évidemment on peut affirmer ceci non seulement en termes des moyennes des REQM mais aussi en termes de la précision, donnée par l'écart-type des REQMs. La figure 6.4 montre comme exemple le scénario 1 avec un bruit transactionnel égal à $\delta = 0.01, 0.1$.

Le gain de faire le filtrage des intensités en considérant conjointement les six CDS par rapport au filtrage de chacune des CDS donnée dans le tableau 6.10.

On faire ressortir que les CDS de maturités plus courtes (6 mois ou 1 an) contiennent

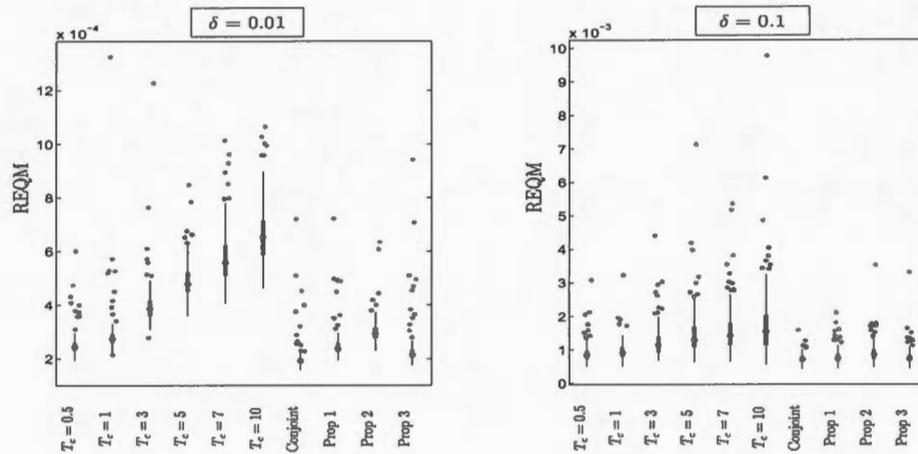


Figure 6.4 Boxplot pour les REQM pour le scénario 1 CIR CDS avec sauts

Tableau 6.10 Gain en pourcentage d'utiliser les six CDS conjointement par scénario et bruit transactionnel : Intensités avec des sauts.

	Scénario 1	Scénario 2	Scénario 3	Scénario 4
$\delta = 0.01$	49.6%	50.68%	49.6%	50.68%
$\delta = 0.1$	37.91%	40.43%	37.91%	40.43%
$\delta = 0.5$	15.48%	20.86%	15.48%	20.86%

plus d'information à propos des intensités (avec ou sans sauts), cependant l'estimation des intensités est significativement supérieure lorsque le filtrage est fait en considérant plusieurs sources d'information à la fois.

Dans le cadre de plusieurs sources d'information de ce type de produit financière, on remarque que la méthode de filtrage proposé montre de caractéristiques semblables en termes de REQM et la précision des estimations des intensités (sans et avec sauts) par rapport à la méthode de filtrage standard.

6.1.6.3 Performance du filtrage joint des deux produits financiers

Les résultats de la performance du filtrage peut se consulter dans les tableaux C.5 et C.6 de l'annexe C. Le premier est relié aux intensités sans sauts et le deuxième aux intensités avec sauts. Dans ces tableaux, on compare les moyennes des REQM, qui on été obtenues à partir de la méthode de filtrage standard par rapport à la méthode de filtrage proposée. Les résultats sont montrés selon le bruit translationnelle, le type de produit financière et les scénarios.

On remarque que la méthode de filtrage standard et la proposée sont statistiquement égales en termes de la REQM dans tous les scénarios testés.

On met en évidence le fait que les CDS nous renseignent plus que les obligations risquées à coupons. Premièrement, sous les scénarios sans sauts, la REQM moyenne des CDS est 3.34 fois plus petite que celle des obligations. Deuxièmement, sous les scénarios avec sauts, la REQM moyenne des CDS est 2.95 fois plus petite que celui des obligations.

En plus, si l'on met ensemble les deux produits financiers pour faire du filtrage, on ne trouve aucune amélioration significative dans la REQM moyenne par rapport au filtrage obtenue en utilisant seulement des CDS, cela veut dire que les obligations risquées ne donnent pas plus d'information et qu'il suffit de prendre seulement de CDS.

6.2 Analyse d'estimation de paramètres par simulation

6.2.1 Surface de la fonction de log-vraisemblance et temps de calcul

Il est à noter que la fonction de log-vraisemblance des modèles état-espace est stochastique, dû au manière comme les estimations sont obtenues : le filtre à particules.

Le fait que la surface de maximisation soit irrégulière est illustré à partir d'un exemple simple en utilisant la représentation d'état-espace du modèle de Merton avec paramètres σ et δ .

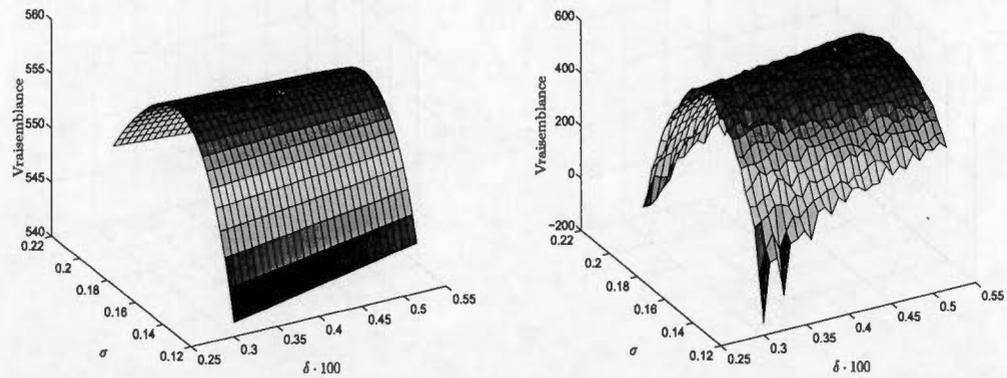
D'abord, on simule une trajectoire de la valeur d'actifs V_t avec $V_0 = 100$, $r = 0.03$, $\sigma = 0.16$, $\Delta t = \frac{1}{252}$, ensuite, on définit une l'obligation zéro-coupon de la valeur faciale $L = 80$ et de maturité $T = 1$ et que les fonds propres S_t ont un bruit transactionnel de $\delta = 0.004$.

On évalue les deux fonction de log-vraisemblance, $l_{UKF}(\sigma, \delta)$ et $l_{FP}(\sigma, \delta)$, reliés à la méthode de filtrage de Kalman « *Unscented* » (UKF) et au filtrage à particules, respectivement, sur les intervalles $\sigma \in (0.12, 0.20)$ et $\delta \in (0.0030, 0.0052)$.

En ce qui concerne le filtrage de Kalman « *Unscented* », on a définit $\alpha_{UKF} = 0.001$, $\beta_{UKF} = 2$ et $\kappa_{UKF} = 0$ et donc $\lambda_{UKF} = \alpha_{UKF}^2(3 + \kappa_{UKF}) - 3 = -3$ et les valeur initiales $\hat{x}_{0|0} = \log(V_0)$ et $P_{0|0} = \sigma^2 \Delta t$. La notation correnpond à celle utilisé par (Pilon, 2012). À l'égard du filtrage à particules, on utilise $N = 200$ particules .

Le panel à gauche de la figure 6.5 correspond à la surface de log-vraisemblance du filtrage UKF tandis que le panel à droite correspond la surface de log-vraisemblance du filtrage à particules tous les deux construits d'après la représentation d'état-espace du modèle de Merton (3.3). Notons que le filtrage UKF est une méthode de quasi-vraisemblance à surface lisse.

Ainsi, on conclut l'exemple en disant qu'il faut faire attention au fait qu'avec la surface stochastique du filtrage à particules, on est sujet à avoir des estimations qui tombent dans des maximums locaux. De plus, on met en évidence l'importance du nombre de particules pour avoir une surface moins rugueuse qui nous permet d'avoir des estimations de paramètres plus fiables. Selon l'article de (Duan et Fulop, 2009) un nombre de particules de 1000 est suffisant à cet effet.



À gauche : La fonction de log-vraisemblance du filtrage de Kalman « Unscented ». À droite : La fonction de log-vraisemblance du filtrage à particules. Les paramètres de la représentation état-espace du modèle de Merton sont $\sigma = 0.16$ et $\delta = 0.004$.

Figure 6.5 Surfaces de log-vraisemblance

Bien évidemment, on constate qu'augmenter le nombre de particules diminue la rugosité de la surface. Pour illustrer cela, on propose de mesurer la rugosité d'une fonction $f(x, y)$ comme la moyenne de la variance quadratique marginale, définie comme :

$$VQM_y = \frac{1}{M} \sum_{j=1}^M \sum_{i=1}^{N-1} (f(x_{i+1}, y_j) - f(x_i, y_j))^2 \quad (6.2)$$

$$VQM_x = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{M-1} (f(x_i, y_{j+1}) - f(x_i, y_j))^2 \quad (6.3)$$

Cette mesure fonctionne lorsque l'on compare deux surfaces ayant la même échelle. Dans notre cas particulier, on calcule la rugosité de la fonction de log-vraisemblance $l(\Theta)$ en fonction du nombre de particules. Le tableau 6.11 révèle que la rugosité diminue 49.37% dans le sens du paramètre σ et 5.39% dans le sens du paramètre δ si l'on utilise 5000 particules plutôt que 400 particules.

Tableau 6.11 Rugosité de la fonction de log-vraisemblance en fonction du nombre de particules

Mesure de Rugosité	Nombre de particules			
	400	1000	2000	5000
VQM_σ	26.7092	26.1198	24.2607	13.5207
VQM_δ	53.0475	55.5257	55.0352	50.1836

En outre, avec le modèle à taux d'intérêt stochastique, on vérifie que la rugosité de la fonction de log-vraisemblance est augmentée de nouveau. Pour illustrer concrètement ce propos, on calcule la rugosité sur quatre scénarios de cette représentation d'état-espace. Ces scénarios sont définis en termes du nombre de particules et de la volatilité du taux d'intérêt, σ_r .

L'idée ici est de générer une surface et de calculer la rugosité en fonction des paramètres σ et δ tout en sachant les paramètres κ_r , θ_r et σ_r , dans chacun des scénarios.

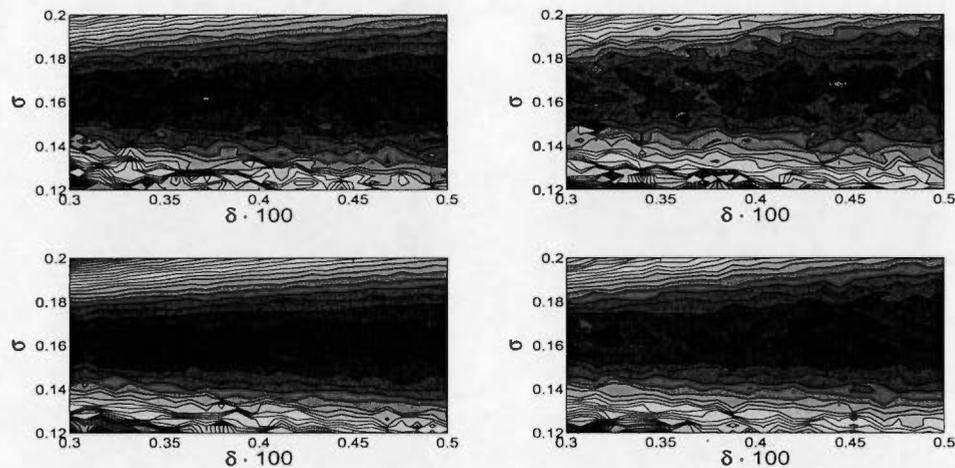
On pose $N = 400$ et 1000 particules et $\sigma_r = 0.0001$ et 0.08 . Les deux autres paramètres reliés à la dynamique du taux, à savoir, $\kappa_r = 0.5$ et $\theta_r = 0.03$ sont les mêmes dans les quatre scénarios.

La figure 6.6 montre que dans les deux premiers panels en haut, les scénarios dont le nombre de particules est $N = 400$ et les paramètres de la volatilité du taux d'intérêt sont $\sigma_r = 0.001$ et $\sigma_r = 0.08$, respectivement. Dans les deux panels en bas, le nombre de particules est augmenté à 1000 et les volatilités du taux d'intérêt sont $\sigma_r = 0.001$ et $\sigma_r = 0.08$, respectivement.

Tel que l'on attendait, la surface de vraisemblance qui a le plus de rugosité est celle qui a le moins de particules et dont la volatilité est la plus grande (deuxième panel en haut), montrant de cette façon l'impact d'avoir un taux stochastique et

l'importance d'avoir un nombre de particules assez grand.

Donc en bref, le choix du nombre de particules très important, cependant, il est recommandé de définir plusieurs points de départ Θ_0 afin de garantir la convergence des estimations qui se trouvent dans la région du maximum de la fonction de log-vraisemblance.



L'axe vertical représente le paramètre σ , l'axe horizontal le paramètre $\delta \cdot 100$. Dans les deux premiers panels en haut, on utilise $N = 400$ particules et le paramètre de volatilité du taux stochastique est $\sigma_r = 0.0001$ et $\sigma_r = 0.08$ respectivement. Pour les troisième et quatrième panels, on utilise $N = 1000$. La surface de vraisemblance qui a plus grande rugosité est celle qui a le moins de particules et dont la volatilité est plus grande (deuxième panel). Les rugosités sont : Panel(1)=91.08, Panel(2)=101.94, Panel(3)=66.01 et Panel(4)=67.77. On fixe les paramètres $\kappa_r = 0.5$ et $\theta_r = 0.03$ dans tous les scénarios.

Figure 6.6 Contour de la vraisemblance du modèle Vasicek

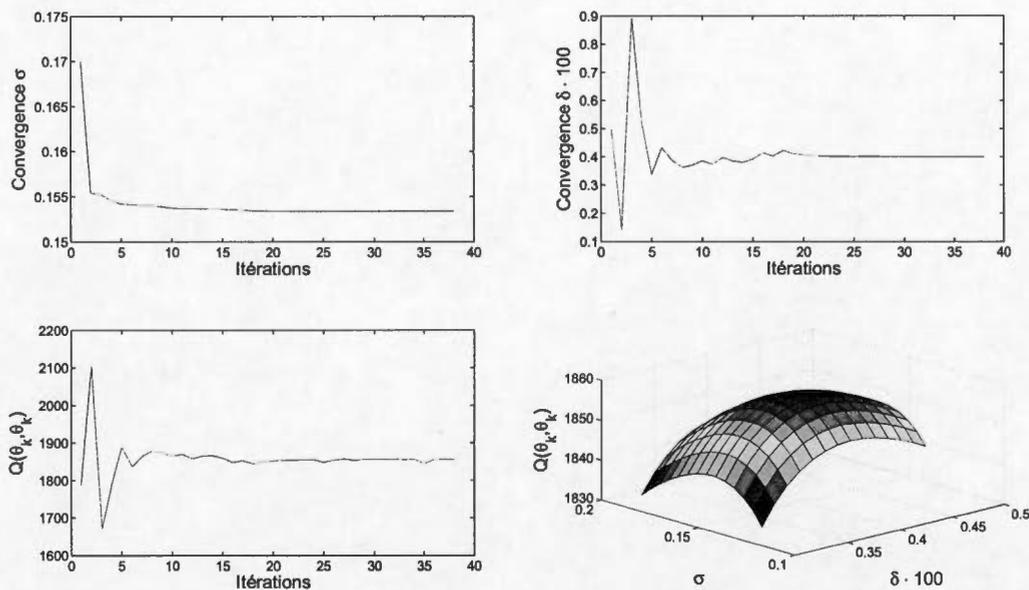
En ce qui concerne la méthode d'estimation par espérance-maximisation (EM), rappelons que l'on construit une fonction auxiliaire $Q(\Theta, \Theta_k)$ qui est une borne inférieure de la fonction de maximum de vraisemblance. L'une des caractéristiques

de cette fonction est qu'elle est déterministe sachant les particules et les poids lisses de l'itération k , ainsi au moment de l'estimation, on s'approche d'une manière itérative à la véritable valeur de Θ en fonction du nombre de particules et du maximum local de l'itération précédente.

On constate que le choix de la valeur α_1 de l'algorithme (EM) décrit à la fin de la section 5.2 et aussi du point de départ Θ_0 est très important en termes de la vitesse de la convergence. Pour donner un exemple, avec $\alpha_1 = 0.03$ le temps de calcul a été 514.0 secondes mais avec $\alpha_1 = 0.01$ le temps du calcul est 78% plus grand (= 774.6 secondes). Dans l'exemple, on a fixé $\Theta_0 = (\sigma_0, \delta_0) = (0.17, 0.005)$.

La figure 6.7 montre les résultats des itérations dans le cas où on établit $N = 400$ particules, $\alpha_1 = 0.03$ et $\Theta_0 = (0.17, 0.005)$. Les deux premiers panels correspondent aux estimations de paramètres σ et δ à chaque itération. Le troisième panel est la valeur de la fonction $Q(\Theta, \Theta_k)$ et le dernier panel montre la surface de la $Q(\Theta, \Theta_k)$ au moment de convergence.

Le temps de calcul entre les deux méthodes d'estimation de paramètres est très sensible au choix du nombre de particules, surtout pour la méthode (EM). Le temps de calcul de la méthode (MV) ne varie pas significativement lorsque le nombre de particules passe de 400 à 1000, cependant, pour la méthode (EM), un grand nombre de particules implique que le temps de calcul reflète une croissance exponentielle parce que des estimations sont obtenues itérativement et qu'à chaque étape, on estime la densité de lissage (les poids lisses $\omega_{t|T}^{(i)}$ et $\omega_{t|T}^{(ij)}$) provenant du lissage définie dans la section 4.2. Donc, on se permet d'affirmer que la méthode (EM) a un désavantage en termes du temps de calcul pour obtenir des estimations. Ceci est illustré dans le tableau 6.12.



Les deux panels en haut montrent la convergence de l'estimation de paramètres

Figure 6.7 Itérations de la méthode (EM) et surface de la fonction $Q(\theta_k, \theta_k)$

Tableau 6.12 Temps de calcul pour estimer les paramètres par les méthodes MV et EM en fonction du nombre de particules

Nombre Particules	Maximum de Vraisemblance			Espérance-Maximisation		
	$\hat{\sigma}$	$\hat{\delta} \cdot 100$	Temp(sec)	$\hat{\sigma}$	$\hat{\delta} \cdot 100$	Temps(sec)
100	0.1638	0.8000	3.64	0.1600	0.6751	31.2
200	0.1575	0.3936	7.36	0.1621	0.5278	91.3
300	0.1539	0.2797	7.87	0.1553	0.4239	201.3
400	0.1650	0.4125	8.52	0.1534	0.3999	434.7
500	0.1570	0.4232	7.63	0.1585	0.3856	514.0
1000	0.1578	0.4185	9.97	0.1587	0.4275	1012.7

CHAPITRE VII

ANALYSE EMPIRIQUE

Dans ce chapitre, on va tester si le modèle à taux d'intérêt stochastique représente une meilleure alternative concernant à l'indice industriel Dow Jones sur la période d'étude 2003. L'analyse empirique est basée sur l'article de (Duan et Fulop, 2009). Ils font des analyses sous l'approche structurelle en considérant le modèle de Merton. On fait donc une analyse similaire en incluant le modèle à taux d'intérêt stochastique.

7.1 Dow Jones 2003

On estime les paramètres du modèle de Merton et du modèle à taux d'intérêt stochastique à partir des 30 compagnies qui constituent l'indice industriel Dow Jones pour l'année 2003. En sachant ces paramètres, on estime les variables latentes, c'est-à-dire la valeur des actifs des 30 compagnies.

Dans le premier modèle, on utilise les deux méthodes d'estimation de paramètres, c'est-à-dire, maximum de vraisemblance et espérance-maximisation, mais dans le deuxième, on n'utilise que la méthode de maximum de vraisemblance étant donné que le temps de calcul pour les obtenir est notablement plus favorable.

Les données sont journalières et ont été prises du site Web <http://www.rmi.nus.edu.sg/DuanJC/>. Il s'agit des mêmes données que (Duan et Fulop, 2009) ont

utilisées pour estimer leur modèle et elles ont été prises du CRSP (*The Center for Research in Security Prices*).

La date de maturité de la dette est $T = 10$ ans. La valeur faciale de la dette de chacune des compagnies est calculée avec leur valeur aux livres respective, accumulée avec un taux d'intérêt sans risque r . La valeur initiale V_0 de chaque compagnie a été prise des estimations obtenues par (Duan et Fulop, 2009).

Avant de procéder avec les estimations, il faut tenir compte que dans ce même article, ils ont montré que parmi les compagnies qui constituent l'index, 6 compagnies admettent un bruit significativement plus grand que zéro.

Dans cette perspective, on ne teste pas la signification du bruit transactionnel, cependant on s'attend à trouver des estimations très semblables dans le modèle de Merton. Au-delà de ceci, on va tester si le modèle à taux d'intérêt stochastique représente une meilleure alternative sur la période d'étude.

7.1.1 Estimation des paramètres du modèle de Merton

Dans ce modèle, le taux d'intérêt sans risque correspond à un taux constant de maturité 1 an, obtenue de la Réserve Fédérale Américaine de l'année 2003. Les paramètres à estimer sont reliés à la volatilité de la valeur des actifs de la compagnie σ et au bruit transactionnel δ .

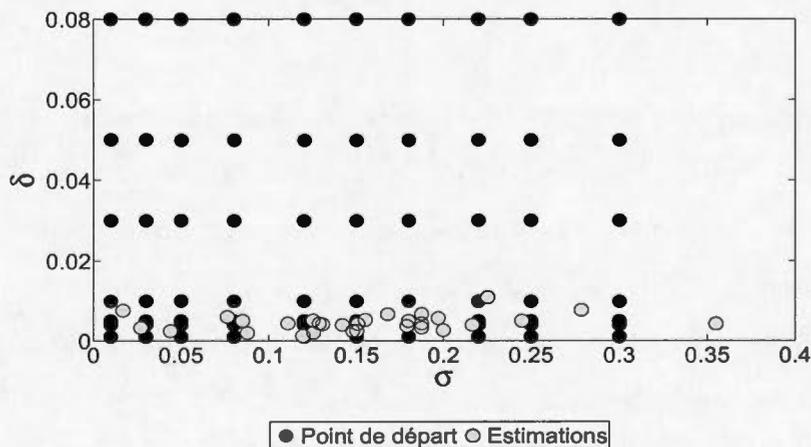
D'après l'analyse par simulation, il est recommandé de définir plusieurs points de départ Θ_0 afin de garantir la qualité des estimations. Donc, on définit une grille de points initiaux $\Theta_0^{(ij)} = [\sigma_0^{(i)}, \delta_0^{(j)}]$, en plus, dans le processus d'estimation, on utilise $N = 1000$ particules.

Pour chaque point de la grille, on va avoir une estimation du paramètre $\hat{\Theta}^{(ij)}$ et sa valeur correspondante de log-vraisemblance. Donc, on choisit l'estimation dont

la valeur de la log-vraisemblance est la plus élevée.

Bien évidemment, on tient compte de la région où les estimations présentées par (Duan et Fulop, 2009) ont été obtenues. La figure 7.1 montre les points de départ ainsi que les coordonnées des estimations que l'on a obtenues via maximum de vraisemblance. La même grille a été considérée pour des estimations obtenues via espérance-maximisation.

Figure 7.1 Points de départ et des estimations du modèle de Merton



Grille de points de départ $\Theta_0^{(ij)}$ et des estimations obtenues $\hat{\Theta}_0^{(ij)}$ via maximum de vraisemblance.

Les estimations obtenues avec les deux méthodes sont rapportées dans le tableau 7.1. L'erreur-type de chaque estimation est montrée entre parenthèses. Notons que les estimations obtenues par les deux méthodes se ressemblent. Comme on s'attendait, ces estimations sont très similaires à celles obtenues dans l'article de (Duan et Fulop, 2009).

7.1.2 Estimation des paramètres du modèle à taux d'intérêt stochastique

Dans ce deuxième modèle, on a trois paramètres additionnels reliés au processus du taux d'intérêt stochastique : κ_r , θ_r et σ_r . On peut envisager deux possibilités pour les estimer.

Premièrement, on peut se servir d'une source d'information externe qui nous donne une indication de la valeur de ces paramètres. Une fois estimés, ils seront une entrée de la représentation d'état-espace, et ensuite, on procède à l'estimation des deux autres paramètres : σ et δ , cependant les estimations obtenues peuvent être biaisées.

Deuxièmement, on peut estimer tous les paramètres simultanément. Cette souhaitable possibilité, et celle que l'on choisit, implique une exploration exhaustive des régions où les estimations sont optimales. C'est pourquoi on définit des points de départ la forme $\Theta_0^{(ijklm)} = [\sigma_0^{(i)}, \delta_0^{(j)}, \kappa_{r,0}^{(k)}, \theta_{r,0}^{(l)}, \sigma_{r,0}^{(m)}]$ afin de trouver la meilleure combinaison des paramètres qui optimise la fonction de log-vraisemblance. En effet, on fait l'exploration dans les intervalles $\sigma \in (0.005, 0.50)$; $\delta \in (0.001, 0.5)$; $\kappa_r \in (0.05, 0.9)$; $\theta_r \in (0.005, 0.07)$ et $\sigma_r \in (0.005, 0.07)$.

Le tableau 7.2 rapporte les estimations et l'erreur-type de chaque estimation entre parenthèses. D'après le test du rapport de vraisemblance le modèle à taux stochastique est significativement meilleur que le modèle de Merton pour 21 des 30 compagnies avec un niveau de confiance de 95%.

On remarque que, parmi les six compagnies dont le bruit est significativement plus grand que zéro, *Exxon Mobil Corp* est la seule qui accepte le modèle à taux stochastique. Sachant que la statistique du test du rapport de vraisemblance (RV) est asymptotiquement distribuée selon une loi khi-deux de Pearson de 3 degrés de liberté, le tableau 7.3 rapporte les vraisemblances des modèles par compagnie et

la correspondante p-valeur. De plus, on calcule les Erreurs Absolues Moyennes en Pourcentage (EAMP) entre les valeurs des actifs estimées par les deux modèles. On remarque les pourcentages ne dépassent pas 3.5% ce qui montre la similitude des estimations de l'année 2003.

Tableau 7.1 Estimations des paramètres du modèle de Merton

Compagnie	Maximum de vraisemblance		espérance-maximisation	
	$\hat{\sigma}$	$\hat{\delta} \cdot 100$	$\hat{\sigma}$	$\hat{\delta} \cdot 100$
3M Corporation	0.1479(0.0077)	0.2537(0.1173)	0.1524(0.0078)	0.0997(0.0218)
Alcoa Inc	0.1873(0.0054)	0.4392(0.1099)	0.1929(0.0021)	0.2407(0.0219)
Altria Group inc	0.1875(0.003)	0.6785(0.0198)	0.184(0.0004)	0.3193(0.0044)
American Express Company	0.0761(0.0211)	0.6109(0.1742)	0.0745(0.0024)	0.1678(0.0703)
American Intl Group Inc	0.085(0.0073)	0.5059(0.1187)	0.0781(0.0036)	0.2626(0.0065)
Boeing Co.	0.1109(0.0169)	0.4379(0.0392)	0.1141(0.0087)	0.1669(0.0317)
Caterpillar Inc.	0.125(0.0054)	0.5172(0.0428)	0.1284(0.0018)	0.3244(0.0667)
Citigroup, Inc.	0.044(0.0119)	0.2502(0.127)	0.0458(0.0066)	0.1916(0.109)
E.I. du Pont de Nemours and Company	0.1287(0.0134)	0.4377(0.1014)	0.1383(0.0006)	0.1045(0.0178)
Exxon Mobil Corp	0.1254(0.0066)	0.2058(0.0469)	0.1303(0.0039)	0.1027(0.0226)
General Electric Company	0.0877(0.0053)	0.2004(0.0619)	0.0831(0.0013)	0.1439(0.0606)
General Motors Corporation	0.0169(0.0158)	0.7661(0.1174)	0.0129(0.0078)	0.8976(0.1497)
Hewlett Packard Co	0.2785(0.0116)	0.7796(0.0287)	0.0057(0.0836)	15.1071(1.8249)
Honeywell Intl Inc	0.1507(0.0128)	0.4264(0.1205)	0.1576(0.0043)	0.1241(0.0577)
IBM	0.1503(0.0106)	0.2538(0.023)	0.1513(0.0052)	0.1236(0.0225)
Intel Corporation	0.3548(0.0128)	0.4309(0.1267)	0.3357(0.0046)	0.1455(0.0553)
J.P. Morgan & Co. Incorporated	0.027(0.001)	0.3341(0.1343)	0.0285(0.0007)	0.2464(0.1136)
Johnson & Johnson	0.1551(0.0147)	0.5297(0.1435)	0.1872(0.0035)	0.101(0.0573)
McDonald's Corporation	0.197(0.0029)	0.5739(0.0678)	0.2137(0.0007)	0.2741(0.0058)
Merck & Co., Inc.	0.2166(0.0126)	0.3999(0.0569)	0.201(0.0002)	0.1112(0.0033)
Microsoft Corporation	0.2447(0.0028)	0.5065(0.0208)	0.005(0.1302)	6.3653(0.5609)
Pfizer Inc	0.1999(0.0043)	0.2675(0.0774)	0.2013(0.0006)	0.1007(0.036)
SBC Communications Inc.	0.1681(0.0024)	0.6786(0.0219)	0.198(0.0003)	0.1833(0.0039)
The Coca-Cola Company	0.1874(0.0048)	0.3313(0.0737)	0.1782(0.0002)	0.1088(0.0034)
The Home Depot, Inc.	0.2253(0.0102)	1.0996(0.0804)	0.2439(0.0038)	0.6782(0.021)
The Procter & Gamble Company	0.1192(0.0105)	0.1265(0.118)	0.1057(0.0036)	0.1001(0.1033)
United Technologies Corporation	0.1306(0.0051)	0.4178(0.1641)	0.1441(0.0053)	0.1189(0.0601)
Verizon Communications	0.1419(0.0054)	0.3987(0.1163)	0.1329(0.0019)	0.2166(0.0563)
Wal-Mart Stores, Inc.	0.1789(0.0098)	0.3751(0.176)	0.1822(0.0037)	0.1043(0.0178)
Walt Disney Company	0.1797(0.0026)	0.5009(0.0801)	0.1867(0.0005)	0.2003(0.0199)

Les estimations sont obtenues via la méthodes de maximum de vraisemblance et espérance-maximisation. L'erreur type de chaque estimation est entre parenthèses.

Tableau 7.2 Estimations des paramètres du modèle à taux d'intérêt stochastique

Compagnie	$\hat{\sigma}$	$\hat{\delta} \cdot 100$	$\hat{\kappa}_r$	$\hat{\theta}_r$	$\hat{\sigma}_r$
3M Corporation	0.1317(0.0116)	0.3089(0.0075)	0.4484(0.0133)	0.0155(0.0117)	0.0111(0.0009)
Alcoa Inc	0.1682(0.0205)	0.9013(0.0224)	0.5852(0.0122)	0.0159(0.0187)	0.001(0.0024)
Altria Group inc	0.1762(0.0081)	1.175(0.0099)	0.4375(0.013)	0.0176(0.011)	0.0117(0.0012)
American Express Co.	0.059(0.0213)	0.77(0.0185)	0.4427(0.0161)	0.0142(0.0205)	0.0011(0.0021)
American Intl Group Inc	0.0948(0.0208)	0.6826(0.0175)	0.5104(0.0189)	0.0093(0.0115)	0.0009(0.0022)
Boeing Co.	0.0969(0.0058)	0.2908(0.0095)	0.0987(0.0094)	0.0145(0.0102)	0.001(0.0009)
Caterpillar Inc.	0.1051(0.024)	0.788(0.0105)	0.5444(0.0179)	0.0169(0.0141)	0.0011(0.0014)
Citigroup, Inc.	0.0506(0.0145)	1.7(0.0292)	0.7441(0.0149)	0.017(0.0142)	0.0017(0.0028)
E.I. du Pont de Nemours Co.	0.1497(0.0131)	0.2994(0.0096)	0.1497(0.0145)	0.01(0.0131)	0.001(0.0015)
Exxon Mobil Corp	0.0992(0.0095)	0.2977(0.0067)	0.0992(0.0094)	0.0099(0.009)	0.001(0.0009)
General Electric Company	0.1011(0.0136)	0.3033(0.0135)	0.1488(0.0072)	0.0101(0.0128)	0.001(0.0015)
General Motors Corporation	0.0111(0.0077)	0.332(0.0076)	0.0504(0.0071)	0.0143(0.004)	0.001(0.0007)
Hewlett Packard Co	0.255(0.0079)	1.1(0.0085)	0.025(0.0235)	0.085(0.0235)	0.0017(0.0022)
Honeywell Intl Inc	0.15(0.0089)	0.3(0.012)	0.1(0.0095)	0.015(0.0091)	0.001(0.001)
IBM	0.1352(0.0057)	0.3846(0.0139)	0.0641(0.0099)	0.0205(0.0143)	0.0014(0.0008)
Intel Corporation	0.3417(0.0098)	0.2839(0.0101)	0.0941(0.0108)	0.0139(0.0104)	0.008(0.001)
J.P. Morgan & Co. Incorp.	0.0095(0.0024)	0.2984(0.0015)	0.4995(0.0008)	0.0143(0.0012)	0.001(0.0001)
Johnson & Johnson	0.1429(0.0236)	0.6204(0.0212)	0.4847(0.0124)	0.0133(0.0252)	0.0166(0.0019)
McDonald's Corporation	0.1461(0.0136)	0.8189(0.0081)	0.4425(0.009)	0.0145(0.0079)	0.0142(0.0006)
Merck & Co., Inc.	0.15(0.0134)	0.7(0.0129)	0.5(0.0067)	0.015(0.0107)	0.01(0.0011)
Microsoft Corporation	0.1442(0.0122)	1.0522(0.0107)	0.1524(0.0121)	0.0157(0.0104)	0.0155(0.0009)
Pfizer Inc	0.1785(0.0067)	0.4499(0.0153)	0.4672(0.0161)	0.0064(0.0152)	0.0096(0.001)
SBC Communications Inc.	0.2068(0.0071)	0.5(0.0081)	0.075(0.0117)	0.0075(0.0111)	0.0005(0.0014)
The Coca-Cola Company	0.1508(0.0163)	0.6095(0.0156)	0.1181(0.0155)	0.0177(0.0179)	0.0009(0.0016)
The Home Depot, Inc.	0.1486(0.0142)	0.9913(0.0164)	0.0991(0.0146)	0.0099(0.0129)	0.0495(0.0017)
The Procter & Gamble Co.	0.0884(0.0152)	0.2611(0.0136)	0.097(0.0142)	0.0088(0.0131)	0.0009(0.0013)
United Technologies Co.	0.1649(0.0127)	0.3236(0.0156)	0.1059(0.0139)	0.0107(0.0168)	0.0011(0.0017)
Verizon Communications	0.139(0.0123)	0.5961(0.0123)	0.6026(0.0216)	0.0113(0.018)	0.0009(0.0012)
Wal-Mart Stores, Inc.	0.1405(0.0184)	0.618(0.0136)	0.0883(0.0179)	0.0088(0.0183)	0.0009(0.0018)
Walt Disney Company	0.157(0.0233)	0.907(0.0123)	0.467(0.0137)	0.0147(0.0268)	0.0011(0.002)

Les estimations sont obtenues via la méthodes de maximum de vraisemblance. L'erreur type de chaque estimation est entre parenthèses.

Tableau 7.3 Rapport de vraisemblance (RV) et EAMP

Compagnie	l_1 (Merton)	l_2 (Vasicek)	p-value (RV)	EAMP
3M Corporation	773.5947	775.3306	0.1622	0.33%
Alcoa Inc	620.5459	624.2686	0.0295 *	0.54%
Altria Group inc	593.8895	624.8915	0.0000 *	1.23%
American Express Co.	683.1725	684.8016	0.1767	0.37%
American Intl Group Inc	632.4636	640.6108	0.0005 *	0.87%
Boeing Co.	658.1906	662.4067	0.0189 *	0.15%
Caterpillar Inc.	654.3979	666.1797	0.0000 *	1.4%
Citigroup, Inc.	528.3531	708.1335	0.0000 *	3.42%
E.I. du Pont de Nemours Co.	714.5611	720.4560	0.0041 *	0.3%
Exxon Mobil Corp	776.9770	786.3874	0.0001 *	0.21%
General Electric Company	693.6482	708.9437	0.0000 *	0.21%
General Motors Corporation	388.7475	659.8904	0.0000 *	0.21%
Hewlett Packard Co	559.2712	562.9455	0.0308 *	1.69%
Honeywell Intl Inc	659.2195	660.0155	0.3306	0.39%
IBM	711.1551	714.4788	0.0420 *	0.65%
Intel Corporation	597.2031	598.8476	0.1746	0.04%
J.P. Morgan & Co. Incorp.	477.6284	649.8752	0.0000 *	0.59%
Johnson & Johnson	736.6122	738.3268	0.1650	0.09%
McDonald's Corporation	610.7897	625.2145	0.0000 *	0.42%
Merck & Co., Inc.	678.7827	692.3601	0.0000 *	0.32%
Microsoft Corporation	630.9883	658.7985	0.0000 *	0.49%
Pfizer Inc	713.1761	716.6704	0.0361 *	0.58%
SBC Communications Inc.	619.9609	621.7135	0.1601	0.6%
The Coca-Cola Company	741.6155	745.3739	0.0286 *	0.25%
The Home Depot, Inc.	569.5300	605.7600	0.0000 *	1.4%
The Procter & Gamble Co.	848.8539	849.4863	0.3687	0.24%
United Technologies Co.	703.3667	710.2458	0.0016 *	0.11%
Verizon Communications	646.8693	652.8166	0.0039 *	0.62%
Wal-Mart Stores, Inc.	726.1573	728.3803	0.1086	0.23%
Walt Disney Company	636.1264	638.0440	0.1399	0.33%

Les deuxième et troisième colonnes représentent les log-vraisemblances des modèles ; la quatrième montre la p-valeur reliée au test de rapport de vraisemblance ; La dernier colonne montre les Erreurs Absolues Moyennes en Pourcentage (EAMP) entre les valeurs des actifs estimées par les deux modèles

CONCLUSION

Dans ce mémoire, on a analysé les propriétés des modèles de risque de crédit sous les deux principales approches. On a introduit le concept de filtrage pour estimer les variables latentes inhérentes dans chaque approche.

On a étudié minutieusement l'estimation de modèles de risque de crédit tant univariés que multivariés, tant au niveau de variables d'état que du nombre d'instruments financiers.

On a aussi étudié deux méthodes d'estimation de paramètres. Finalement, on a fait une analyse empirique en prenant les 30 compagnies qui composent l'indice Dow Jones.

D'après l'analyse par simulation, dans le modèle de Merton, les fonds propres qui nous offrent le plus d'information à propos des actifs sont ceux de maturité plus courte, Mais dans le modèle à taux d'intérêt stochastique, ceux qui nous offrent le plus d'information sont ceux de maturité plus longue.

Pour les modèles à forme réduite, on peut souligner que les CDS donnent plus d'information sur les intensités par rapport aux obligations risquées à coupons. De plus les CDS de maturités courtes donnent plus d'informations. On remarque que la qualité du filtrage est notablement améliorée lorsque l'on considère plusieurs sources d'information dans tous les scénarios proposés.

Le filtrage à multiples sources d'information proposée est statistiquement égale à la méthode standard en termes de la REQ. Elle se présente comme une autre alternative de filtrage.

On met en évidence le fait que l'une des difficultés du filtrage lors de la présence des sauts est qu'il y a une probabilité que les particules ne soient pas capables d'estimer (filtrer) le saut, donc l'utilisation de plus des particules est nécessaire.

D'un autre côté, on a constaté que la méthode d'estimation de paramètres par maximum de vraisemblance est plus efficiente en termes de temps de calcul, cependant la méthode d'estimation de paramètres par espérance-maximisation est une alternative très intéressante théoriquement, qui nous a servis comme un outil de validation.

Finalement, dans notre analyse empirique de l'indice Dow Jones, la qualité de l'ajustement du modèle de Merton est améliorée par le modèle à taux d'intérêts stochastique dans 26 des 30 compagnies.

Il serait intéressant d'étudier les modèles à volatilité stochastique, comme celui de Heston-Nandi, pour améliorer la qualité de l'ajustement, cependant le filtrage peut être complexe étant donné le calcul des fonds propres implique le calcul numérique des intégrales. Certainement, des méthodes d'interpolation seront nécessaires afin d'optimiser le processus de filtrage.

APPENDICE A

ÉCHANTILLONNAGE PRÉFÉRENTIEL

L'échantillonnage préférentiel (*importance sampling*) est une technique de réduction de variance utilisée pour calculer un estimateur non biaisé à partir de la méthode de simulation de Monte-Carlo. Dans la méthode de réduction de variance, on s'aperçoit qu'il y a des valeurs qui sont plus importantes que d'autres, alors si l'on peut faire que la réalisation des valeurs, qui sont plus importantes, soit plus fréquentes, on aura finalement un estimateur d'une variance plus petite.

Une bonne définition de cette méthode se trouve dans le papier de (Bergman, 1999). Soit $p(x) \propto \pi(x)$ une fonction de densité pour laquelle on a la difficulté de générer des nombres aléatoires, mais pour laquelle $\pi(x)$ peut être évaluée.

Soit $q(x)$ une fonction avec laquelle on peut générer des nombres aléatoires facilement, donc on pourrait avoir un échantillon de taille n , c'est-à-dire, $x^{(i)} \sim q(x)$ pour $i = 1, 2, \dots, n$. Cette fonction doit satisfaire que son support couvre celui de $\pi(x)$, par exemple, si $\pi(x) > 0$ donc $q(x) > 0$ pour tout $x \in \mathbb{R}^n$. On trouve que $q(x)$ est communément appelée la densité d'importance *proposal density*.

En considérant la condition décrite ci-haut, on peut définir l'égalité suivante :

$$I = \int_{\mathbb{R}^n} f(x)\pi(x)dx = \int_{\mathbb{R}^n} f(x)\frac{\pi(x)}{q(x)}q(x)dx \quad (\text{A.1})$$

D'après la méthode de Monte-Carlo, un estimateur asymptotiquement sans biais

de cette intégrale est obtenu à partir de la somme pondérée suivante :

$$f_N = \frac{1}{N} \sum_{i=1}^N f(x_i) w_i \quad \text{pour} \quad w_i = \frac{\pi(x_i)}{q(x_i)} \quad (\text{A.2})$$

Les poids w_i sont connus comme les *poids d'importance*. Si le facteur de normalisation de la fonction objective $\pi(x)$ est inconnu, les poids d'importance peuvent jouer ce rôle avec lequel on obtient l'estimateur suivant :

$$f_N = \frac{\sum_{i=1}^N f(x_i) w_i}{\sum_{i=1}^N w_i} \quad \text{pour} \quad w_i = \frac{\pi(x_i)}{q(x_i)}$$

APPENDICE B

DÉVELOPPEMENT DE L'ESPÉRANCE CONDITIONNELLE : ESPÉRANCE-MAXIMISATION

On a supposé que x_t est un processus markovien d'ordre 1 et que y_t ne dépend que de x_t . En plus, soit $I_i \subseteq \mathbb{R}$ sous-ensemble un ouvert borné de la variable d'état x_i pour $i = 1, 2, \dots, T$ et $\chi = I_0 \times I_1 \times I_2 \times \dots \times I_T \subseteq \mathbb{R}^{T+1}$ un sous-ensemble ouvert borné des variables d'état $\mathbf{x}_{0:T}$. On remarque que l'équation (5.5) peut s'écrire comme :

$$\begin{aligned}
 Q(\Theta, \Theta_k) &= \mathbb{E}^U [\log p_{\Theta}(\mathbf{y}_{1:T}, \mathbf{x}_{0:T}) \mid \Theta_k, \mathbf{y}_{1:T}] \\
 &= \int_{\chi} p_{\Theta_k}(\mathbf{x}_{0:T} \mid \mathbf{y}_{1:T}) \left[\log p_{\Theta}(x_0) + \sum_{t=1}^T \log p_{\Theta}(y_t \mid x_t) \right. \\
 &\quad \left. + \sum_{t=1}^T \log p_{\Theta}(x_t \mid x_{t-1}) \right] d\mathbf{x}_{1:T} \\
 &= \int_{\chi} p_{\Theta_k}(\mathbf{x}_{0:T} \mid \mathbf{y}_{1:T}) \log p_{\Theta}(x_0) d\mathbf{x}_{1:T} \\
 &\quad + \sum_{t=1}^T \int_{\chi} p_{\Theta_k}(\mathbf{x}_{0:T} \mid \mathbf{y}_{0:T}) \log p_{\Theta}(y_t \mid x_t) d\mathbf{x}_{1:T} \\
 &\quad + \sum_{t=1}^T \int_{\chi} p_{\Theta_k}(x_t, x_{t-1} \mid \mathbf{y}_{1:T}) \log p_{\Theta}(x_t \mid x_{t-1}) d\mathbf{x}_{1:T} \\
 &= I_1(\Theta, \Theta_k) + I_2(\Theta, \Theta_k) + I_3(\Theta, \Theta_k)
 \end{aligned}$$

$$\begin{aligned}
I_1(\Theta, \Theta_k) &= \int_{\chi_0} \int_{\chi_1} \cdots \int_{\chi_T} p_{\Theta_k}(\mathbf{x}_{0:T} | \mathbf{y}_{1:T}) \log p_{\Theta}(x_0) dx_0 dx_1 \dots dx_T \\
&= \int_{\chi_0} \log p_{\Theta}(x_0) \left(\int_{\chi_1} \cdots \int_{\chi_T} p_{\Theta_k}(\mathbf{x}_{0:T} | \mathbf{y}_{1:T}) dx_1 \dots dx_T \right) dx_0 \\
&= \int_{\chi_0} \log p_{\Theta}(x_0) p_{\Theta_k}(x_0 | \mathbf{y}_{1:T}) dx_0 \\
&= \mathbb{E}[\log p_{\Theta}(x_0) | \mathbf{y}_{1:T}, \Theta_k]
\end{aligned}$$

$$\begin{aligned}
I_2(\Theta, \Theta_k) &= \sum_{t=1}^T \int_{\chi_0} \int_{\chi_1} \cdots \int_{\chi_T} p_{\Theta_k}(\mathbf{x}_{0:T} | \mathbf{y}_{1:T}) \log p_{\Theta}(y_t | x_t) dx_0 dx_1 \dots dx_T \\
&= \sum_{t=1}^T \int_{\chi_t} \log p_{\Theta}(y_t | x_t) \left(\int_{\chi'} p_{\Theta_k}(\mathbf{x}_{0:T} | \mathbf{y}_{1:T}) d\mathbf{x}_{1:T'} \right) dx_t \\
&= \sum_{t=1}^T \int_{\chi_t} \log p_{\Theta}(y_t | x_t) p_{\Theta_k}(x_t | \mathbf{y}_{1:T}) dx_t \\
&= \sum_{t=1}^T \mathbb{E}[\log p_{\Theta}(y_t | x_t) | \mathbf{y}_{1:T}, \Theta_k]
\end{aligned}$$

Avec χ' un sous-ensemble ouvert borné de \mathbb{R}^T qui ne considère pas le sous-ensemble χ_t .

$$\begin{aligned}
I_3(\Theta, \Theta_k) &= \sum_{t=1}^T \int_{\chi_0} \int_{\chi_1} \cdots \int_{\chi_T} p_{\Theta_k}(\mathbf{x}_{0:T} | \mathbf{y}_{1:T}) \log p_{\Theta}(x_t | x_{t-1}) dx_0 dx_1 \dots dx_T \\
&= \sum_{t=1}^T \int_{\chi_{t-1}} \int_{\chi_t} \log p_{\Theta}(x_t | x_{t-1}) \left(\int_{\chi''} p_{\Theta_k}(\mathbf{x}_{0:T} | \mathbf{y}_{1:T}) d\mathbf{x}_{1:T''} \right) dx_t dx_{t-1} \\
&= \sum_{t=1}^T \int_{\chi_{t-1}} \int_{\chi_t} \log p_{\Theta}(x_t | x_{t-1}) p_{\Theta_k}(x_{t-1}, x_t | \mathbf{y}_{1:T}) dx_t dx_{t-1} \\
&= \sum_{t=1}^T \mathbb{E}[\log p_{\Theta}(x_t | x_{t-1}) | \mathbf{y}_{1:T}, \Theta_k]
\end{aligned}$$

Avec χ'' un sous-ensemble ouvert borné de \mathbb{R}^{T-1} qui ne considère pas le sous-ensemble χ_t ni χ_{t-1} .

APPENDICE C

PERFORMANCE DU FILTRAGE : REQM

Tableau C.1 Comparaison des obligations risquées à coupons sans sauts dans les intensités

Bruit trans.	Maturité	Scénario 1	Scénario 2	Scénario 3	Scénario 4
$\delta = 0.01$	T_5	0.00087	0.00098	0.00107	0.00137
	T_{10}	0.00083	0.00094	0.00107	0.00138
	T_{15}	0.00079	0.00091	0.00105	0.00135
	T_{30}	0.00078*	0.00090*	0.00106	0.00136
$\delta = 0.1$	T_5	0.00223	0.00258	0.00183	0.00262
	T_{10}	0.00214	0.00251	0.00183	0.00259
	T_{15}	0.00214	0.00249	0.00181	0.00261
	T_{30}	0.00209	0.00244	0.00180	0.00260
$\delta = 0.5$	T_5	0.00279	0.00326	0.00185	0.00277
	T_{10}	0.00275	0.00332	0.00190	0.00277
	T_{15}	0.00279	0.00342	0.00184	0.00278
	T_{30}	0.00277	0.00329	0.00190	0.00275

*Les scénarios correspondent à ceux qui ont été définis dans le tableau 6.1. Les valeurs avec * montrent la valeur plus petite parmi le groupe étant significativement différente d'après le test F*

Tableau C.2 Comparaison des obligations risquées à coupons avec sauts dans les intensités

Bruit trans.	Maturité	Scénario 1	Scénario 2	Scénario 3	Scénario 4
$\delta = 0.01$	T_5	0.00826	0.00951	0.00909	0.00968
	T_{10}	0.00782	0.00927	0.00907	0.00991
	T_{15}	0.00805	0.00940	0.00916	0.00929
	T_{30}	0.00812	0.00897	0.00935	0.00951
$\delta = 0.1$	T_5	0.01349	0.02739	0.01799	0.03752
	T_{10}	0.01351	0.02609	0.01697	0.04152
	T_{15}	0.01320	0.02592	0.01714	0.04010
	T_{30}	0.01396	0.02600	0.01788	0.04322
$\delta = 0.5$	T_5	0.01372	0.03105	0.01732	0.05419
	T_{10}	0.01357	0.03078	0.01762	0.05793
	T_{15}	0.01379	0.02965	0.01780	0.05616
	T_{30}	0.01367	0.03134	0.01775	0.05590

*Les scénarios correspondent à ceux qui ont été définis dans le tableau 6.2. Les valeurs avec * montrent la valeur plus petite parmi le groupe étant significativement différente d'après le test F*

Tableau C.3 Comparaison des CDS sans sauts dans les intensités

Bruit trans.	Maturité	Scénario 1	Scénario 2	Scénario 3	Scénario 4
$\delta = 0.01$	$T_{0.5}$	0.000179*	0.000273*	0.000152*	0.000331*
	T_1	0.000182	0.000289	0.000162	0.000372
	T_3	0.000198	0.000351	0.000210	0.000525
	T_5	0.000216	0.000411	0.000257	0.000653
	T_7	0.000231	0.000458	0.000293	0.000741
	T_{10}	0.000255	0.000524	0.000342	0.000854
$\delta = 0.1$	$T_{0.5}$	0.000659*	0.000957*	0.000554*	0.001107*
	T_1	0.000669	0.001003	0.000585	0.001209
	T_3	0.000718	0.001181	0.000716	0.001582
	T_5	0.000759	0.001319	0.000822	0.001838
	T_7	0.000798	0.001440	0.000909	0.001987
	T_{10}	0.000872	0.001626	0.001032	0.002201
$\delta = 0.5$	$T_{0.5}$	0.001383*	0.001984*	0.001122*	0.002069*
	T_1	0.001425	0.002052	0.001163	0.002223
	T_3	0.001529	0.002402	0.001373	0.002545
	T_5	0.001562	0.002623	0.001488	0.002579
	T_7	0.001639	0.002714	0.001560	0.002682
	T_{10}	0.001771	0.002928	0.001689	0.002725

*Les scénarios correspondent à ceux qui ont été définis dans le tableau 6.1. Les valeurs avec * montrent la valeur plus petite parmi le groupe étant significativement différente d'après le test F*

Tableau C.4 Comparaison des CDS avec sauts dans les intensités

Bruit trans.	Maturité	Scénario 1	Scénario 2	Scénario 3	Scénario 4
$\delta = 0.01$	$T_{0.5}$	0.001551*	0.001858*	0.001959*	0.002311*
	T_1	0.001780	0.002049	0.002308	0.002423
	T_3	0.002441	0.002849	0.002912	0.003126
	T_5	0.003013	0.003475	0.003329	0.003893
	T_7	0.003569	0.003968	0.003951	0.004647
	T_{10}	0.004102	0.004500	0.004458	0.005142
$\delta = 0.1$	$T_{0.5}$	0.005628*	0.007088*	0.006182*	0.010093*
	T_1	0.006010	0.007410	0.006953	0.011034
	T_3	0.007906	0.009790	0.009072	0.014645
	T_5	0.009460	0.012435	0.010878	0.017918
	T_7	0.010235	0.014008	0.012231	0.021262
	T_{10}	0.011564	0.015914	0.013618	0.024449
$\delta = 0.5$	$T_{0.5}$	0.010499*	0.016791*	0.013192*	0.023782*
	T_1	0.011377	0.017276	0.013944	0.026272
	T_3	0.012619	0.021440	0.015562	0.033537
	T_5	0.013303	0.025451	0.016898	0.040676
	T_7	0.013708	0.026855	0.017463	0.044241
	T_{10}	0.013826	0.028637	0.017576	0.048788

Les scénarios correspondent à ceux qui ont été définis dans le tableau 6.2. Les valeurs avec * montrent la valeur plus petite parmi le groupe étant significativement différente d'après le test F

Tableau C.5 Comparaison des filtrages standard (Sd) et proposée (Pr) prenant de multiples sources d'information sans sauts dans les intensités

Bruit trans.		Scénario 1	Scénario 2	Scénario 3	Scénario 4
$\delta = 0.01$	OBR-Sd	0.00057	0.00066	0.00078	0.00100
	OBR-Pr	0.00059	0.00067	0.00082	0.00103
	CDS-Sd	0.00011	0.00019	0.00011	0.00026
	CDS-Pr	0.00013	0.00023	0.00014	0.00032
	PF-Sd	0.00011	0.00019	0.00011	0.00026
	PF-Pr	0.00017	0.00027	0.00017	0.00037
$\delta = 0.1$	OBR-Sd	0.00166	0.00192	0.00165	0.00230
	OBR-Pr	0.00187	0.00210	0.00176	0.00253
	CDS-Sd	0.00046	0.00074	0.00044	0.00094
	CDS-Pr	0.00048	0.00078	0.00047	0.00101
	PF-Sd	0.00046	0.00074	0.00044	0.00093
	PF-Pr	0.00051	0.00083	0.00051	0.00109
$\delta = 0.5$	OBR-Sd	0.00267	0.00305	0.00184	0.00268
	OBR-Pr	0.00280	0.00326	0.00182	0.00274
	CDS-Sd	0.00103	0.00160	0.00094	0.00187
	CDS-Pr	0.00111	0.00185	0.00105	0.00215
	PF-Sd	0.00103	0.00158	0.00093	0.00187
	PF-Pr	0.00122	0.00205	0.00117	0.00235

*Les scénarios correspondent à ceux qui ont été définis dans le tableau 6.2. Les valeurs avec * montrent la valeur plus petite parmi le groupe étant significativement différente d'après le test F. OBR : Des obligations risquées à coupons; CDS : Des swap sur défaillance de crédit; PF : Les deux produits financiers ensemble*

Tableau C.6 Comparaison des filtrages standard (Sd) et proposée (Pr) prenant de multiples sources d'information avec sauts dans les intensités

Bruit trans.		Scénario 1	Scénario 2	Scénario 3	Scénario 4
$\delta = 0.01$	OBR-Sd	0.00583	0.00654	0.00643	0.00659
	OBR-Pr	0.00624	0.00686	0.00694	0.00699
	CDS-Sd	0.00125	0.00149	0.00160	0.00189
	CDS-Pr	0.00155	0.00179	0.00182	0.00210
	PF-Sd	0.00128	0.00173	0.00160	0.00213
	PF-Pr	0.00191	0.00240	0.00212	0.00241
$\delta = 0.1$	OBR-Sd	0.01291	0.02106	0.01604	0.03009
	OBR-Pr	0.01320	0.02163	0.01656	0.03266
	CDS-Sd	0.00455	0.00553	0.00527	0.00908
	CDS-Pr	0.00493	0.00620	0.00585	0.01004
	PF-Sd	0.00455	0.00573	0.00515	0.00934
	PF-Pr	0.00560	0.00736	0.00627	0.01022
$\delta = 0.5$	OBR-Sd	0.01381	0.02914	0.01813	0.05164
	OBR-Pr	0.01363	0.02980	0.01828	0.05499
	CDS-Sd	0.00958	0.01251	0.01142	0.01901
	CDS-Pr	0.01080	0.01523	0.01313	0.02268
	PF-Sd	0.00944	0.01290	0.01098	0.01848
	PF-Pr	0.01138	0.01662	0.01448	0.02347

*Les scénarios correspondent à ceux qui ont été définis dans le tableau 6.2. Les valeurs avec * montrent la valeur plus petite parmi le groupe étant significativement différente d'après le test F. OBR : Des obligations risquées à coupons; CDS : Des swap sur défaillance de crédit; PF : Les deux produits financières ensemble*

RÉFÉRENCES

- Amin, K. I. et Jarrow, R. A. (1992). Pricing options on risky assets in a stochastic interest rate economy. *Mathematical Finance*, 2(4), 217–237.
- Anderson, B. et Moore, J. B. (1979). Optimal filtering. *Prentice-Hall Information and System Sciences Series, Englewood Cliffs : Prentice-Hall, 1979, 1.*
- Aratyn, H. et Rasinariu, C. (2006). *A short course in mathematical methods with Maple.* World Scientific.
- Arulampalam, M. S., Maskell, S., Gordon, N. et Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on*, 50(2), 174–188.
- Ballestra, L. V. et Pacelli, G. (2014). Valuing risky debt : A new model combining structural information with the reduced-form approach. *Insurance : Mathematics and Economics*, 55, 261–271.
- Bergman, N. (1999). Recursive bayesian estimation. *Department of Electrical Engineering, Linköping University, Linköping Studies in Science and Technology. Doctoral dissertation, 579.*
- Black, F. et Cox, J. C. (1976). Valuing corporate securities : Some effects of bond indenture provisions. *The Journal of Finance*, 31(2), 351–367.
- Boudreault, M., Gauthier, G. et Thomassin, T. (2013). Recovery rate risk and credit spreads in a hybrid credit risk model. *Journal of Credit Risk*, 9 Fall 2013.
- Brigo, D. et Alfonsi, A. (2005). Credit default swap calibration and derivatives pricing with the ssrd stochastic intensity model. *Finance and Stochastics*, 9(1), 29–42.
- Brigo, D. et El-Bachir, N. (2006). Credit derivatives pricing with a smile-extended jump stochastic intensity model. *Available at SSRN 950208.*
- Brigo, D. et Mercurio, F. (2007). *Interest rate models-theory and practice : with smile, inflation and credit.* Springer Science & Business Media.

- Bu, D. et Liao, Y. (2013). Structural credit risk model with stochastic volatility : a particle-filter approach. *Available at SSRN 2336452*.
- Cappé, O., Godsill, S. J. et Moulines, E. (2007). An overview of existing methods and recent advances in sequential Monte Carlo. *Proceedings of the IEEE*, 95(5), 899–924.
- Carr, P. et Wu, L. (2007). Stochastic skew in currency options. *Journal of Financial Economics*, 86(1), 213–247.
- Cetin, U., Jarrow, R. A., Protter, P. et Yildirim, Y. (2004). Modeling credit risk with partial information. *Annals*, 14(3), 1167–1178.
- Chen, L., Filipovic, D. et Poor, H. V. (2004). A mixed approach to modeling default risk. *Risk*, 17.
- Chen, Z. (2003). Bayesian filtering : From Kalman filters to particle filters, and beyond. *Statistics*, 182(1), 1–69.
- Creal, D. (2012). A survey of sequential Monte Carlo methods for economics and finance. *Econometric Reviews*, 31(3), 245–296.
- Dempster, A. P., Laird, N. M. et Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, 1–38.
- Doucet, A., De Freitas, N. et Gordon, N. (2001). An introduction to sequential Monte Carlo methods. In *Sequential Monte Carlo methods in practice* 3–14. Springer.
- Doucet, A., Godsill, S. et Andrieu, C. (2000). On sequential Monte Carlo sampling methods for bayesian filtering. *Statistics and computing*, 10(3), 197–208.
- Doucet, A. et Johansen, A. M. (2009). A tutorial on particle filtering and smoothing : Fifteen years later. *Handbook of Nonlinear Filtering*, 12, 656–704.
- Duan, J.-C. (1994). Maximum likelihood estimation using price data of the derivative contract. *Mathematical Finance*, 4(2), 155–167.
- Duan, J.-C. et Fulop, A. (2009). Estimating the structural credit risk model when equity prices are contaminated by trading noises. *Journal of Econometrics*, 150(2), 288–296.
- Duffie, D. et Lando, D. (2001). Term structures of credit spreads with incomplete accounting information. *Econometrica*, 633–664.

- Duffie, D. et Singleton, K. J. (1999). Modeling term structures of defaultable bonds. *Review of Financial studies*, 12(4), 687–720.
- Elizalde, A. (2005). Credit risk models III : reconciliation reduced-structural models. In *CEMFI and Universidad de Navarra Working Paper*.
- Ericsson, J. et Reneby, J. (2002). Estimating structural bond pricing models. *Available at SSRN 268786*.
- Geman, S. et Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6), 721–741.
- Glasserman, P. (2003). *Monte Carlo methods in financial engineering*, volume 53. Springer Science & Business Media.
- Goldberg, L. R. et Giesecke, K. (2004). Forecasting default in the face of uncertainty. *Journal of Derivatives*, 12(1), 14–25.
- Gordon, N. J., Salmond, D. J. et Smith, A. F. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. Dans *IEEE Proceedings F (Radar and Signal Processing)*, volume 140, 107–113. IET.
- Hamilton, J. D. (1995). *Time series analysis*, volume 1. Princeton university press Princeton.
- Heath, D., Jarrow, R. et Morton, A. (1990). Bond pricing and the term structure of interest rates : A discrete time approximation. *Journal of Financial and Quantitative analysis*, 25(04), 419–440.
- Heath, D., Jarrow, R. et Morton, A. (1992). Bond pricing and the term structure of interest rates : A new methodology for contingent claims valuation. *Econometrica : Journal of the Econometric Society*, 77–105.
- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2), 327–343.
- Heston, S. L. et Nandi, S. (1997). A closed-form GARCH option pricing model.
- Huang, J.-z. et Wu, L. (2004). Specification analysis of option pricing models based on time-changed lévy processes. *The Journal of Finance*, 59(3), 1405–1439.
- Huang, S. J. et Yu, J. (2010). Bayesian analysis of structural credit risk models with microstructure noises. *Journal of Economic Dynamics and Control*, 34(11), 2259–2272.

- Jarrow, R. et Protter, P. (2004). Structural versus reduced form models : a new information based perspective. *Journal of Investment management*, 2(2), 1–10.
- Jarrow, R. A., Lando, D. et Turnbull, S. M. (1997). A Markov model for the term structure of credit risk spreads. *Review of Financial studies*, 10(2), 481–523.
- Jarrow, R. A. et Turnbull, S. M. (1995). Pricing derivatives on financial securities subject to credit risk. *The Journal of Finance*, 50, 53–53.
- Jarrow, R. A. et Yu, F. (2001). Counterparty risk and the pricing of defaultable securities. *The Journal of Finance*, 56(5), 1765–1799.
- Johannes, M. S. et Polson, N. (2003). MCMC methods for continuous-time financial econometrics. Available at SSRN 480461.
- Johannes, M. S., Polson, N. G. et Stroud, J. R. (2009). Optimal filtering of jump diffusions : Extracting latent states from asset prices. *Review of Financial Studies*, 2759–2799.
- Julier, S. J. et Uhlmann, J. K. (1997). New extension of the Kalman filter to non-linear systems. Dans *AeroSense'97*, 182–193. International Society for Optics and Photonics.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1), 35–45.
- Kim, Y.-J. (2002). Option pricing under stochastic interest rates : an empirical investigation. *Asia-Pacific Financial Markets*, 9(1), 23–44.
- Kokkala, J., Solin, A. et Sarkka, S. (2014). Expectation maximization based parameter estimation by sigma-point and particle smoothing. Dans *Information Fusion (FUSION), 2014 17th International Conference on*, 1–8. IEEE.
- Kong, A., Liu, J. S. et Wong, W. H. (1994). Sequential imputations and Bayesian missing data problems. *Journal of the American statistical association*, 89(425), 278–288.
- Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives research*, 2(2-3), 99–120.
- Leland, H. E. (1994). Corporate debt value, bond covenants, and optimal capital structure. *The Journal of Finance*, 1213–1252.
- Leland, H. E. et Toft, K. B. (1996). Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *The Journal of Finance*, 987–1019.

- Longstaff, F. A. et Schwartz, E. S. (1995). A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance*, 50(3), 789–819.
- Madan, D. et Unal, H. (2000). A two-factor hazard rate model for pricing risky debt and the term structure of credit spreads. *Journal of Financial and Quantitative Analysis*, 35(01), 43–65.
- Madan, D. B. et Unal, H. (1998). Pricing the risks of default. *Review of Derivatives Research*, 2(2-3), 121–160.
- Merton, R. C. (1974). On the pricing of corporate debt : The risk structure of interest rates. *The Journal of Finance*, 29(2), 449–470.
- Mikulevičius, R. et Platen, E. (1988). Time discrete taylor approximations for itô processes with jump component. *Mathematische Nachrichten*, 138(1), 93–104.
- Neal, R. M. et Hinton, G. E. (1998). A view of the em algorithm that justifies incremental, sparse, and other variants. In *Learning in graphical models* 355–368. Springer.
- Pilon, D. (2012). Estimation des modèles structurels de risque de crédit en présence de bruit transactionnel et de biais de survie. *Rapport de recherche, Université de Québec à Montréal*.
- Pitt, M. K. (2002). Smooth particle filters for likelihood evaluation and maximisation. *Warwick Economic Research Papers*.
- Roweis, S. et Ghahramani, Z. (2001). Learning nonlinear dynamical systems using the expectation–maximization algorithm. *Kalman filtering and neural networks*, p. 175.
- Schön, T. B., Wills, A. et Ninness, B. (2011). System identification of nonlinear state-space models. *Automatica*, 47(1), 39–49.
- Zhu, L. et al. (2014). Limit theorems for a Cox-Ingersoll-Ross process with Hawkes jumps. *Journal of Applied Probability*, 51(3), 699–712.