



Credit risk modeling. Calibration and discretization of financial models

Aurélien Alfonsi

► To cite this version:

Aurélien Alfonsi. Credit risk modeling. Calibration and discretization of financial models. Mathematics [math]. Ecole des Ponts ParisTech, 2006. English. NNT : . pastel-00001859

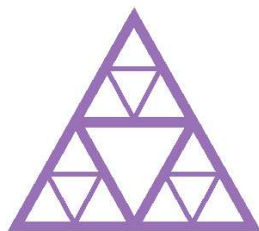
HAL Id: pastel-00001859

<https://pastel.hal.science/pastel-00001859>

Submitted on 28 Jul 2006

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

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



Thèse présentée pour obtenir le titre de
**DOCTEUR DE L'ÉCOLE NATIONALE DES
PONTES ET CHAUSSEES**

Spécialité : Mathématiques appliquées

par

Aurélien ALFONSI

**Modélisation en risque de crédit.
Calibration et discrétisation de
modèles financiers.**

Thèse soutenue le 27 juin 2006 devant le jury composé de :

Mireille BOSSY	<i>Examineur</i>
Damiano BRIGO	<i>Examineur</i>
Nicole EL KAROUI	<i>Rapporteur</i>
Benjamin JOURDAIN	<i>Directeur de thèse</i>
Damien LAMBERTON	<i>Président</i>
Marek RUTKOWSKI	<i>Rapporteur</i>

Résumé

Le premier volet de cette thèse traite du marché du risque de crédit. Après un bref chapitre introductif à ce marché et à sa modélisation, nous introduisons un modèle à intensité de défaut baptisé SSRD pour Shifted Square-Root Diffusion. Ce modèle a pour qualité principale de pouvoir être automatiquement calibré aux prix des Credit Default Swaps observés sur le marché. En outre, il permet d’avoir une intensité de défaut et un taux d’intérêt dépendants entre eux. Ensuite, nous présentons une nouvelle classe de fonctions copules appelées “copules périodiques” car leur construction est basée sur des fonctions périodiques. Les copules interviennent en risque de crédit dans la modélisation jointe de plusieurs instants de défaut. Les copules périodiques permettent de balayer un large spectre de dépendances, de C^- à C^+ en passant par $C^\perp$ dans le cas bivarié, et s’étend facilement au cas multivarié. En outre, la simulation de ces copules est aisée.

Ensuite nous nous intéressons à la discrétisation de l’équation différentielle stochastique de Cox-Ingersoll-Ross qui intervient dans le modèle SSRD. Nous présentons de nouveaux schémas, implicites et explicites, et étudions leurs convergences forte et faible. On regarde également numériquement le comportement de ces schémas et les comparons aux schémas déjà proposés par Deelstra et Delbaen [27] et par Diop [28]. Cette étude permet d’identifier un schéma qui réunit, a priori, le plus de propriétés de convergence désirables.

La dernière partie de cette thèse concerne le marché action. Il est désormais bien connu que l’on peut, dans un modèle à volatilité locale, trouver un profil de volatilités qui calibre exactement les prix des options européennes observés. On s’intéresse ici au problème analogue pour les options américaines : peut-on trouver un profil de volatilités qui donne exactement les prix des options américaines ? Pour attaquer ce problème, nous considérons le cas des options américaines perpétuelles et d’une fonction de volatilité locale homogène en temps. Nous mettons alors en évidence une relation de dualité Call-Put qui permet d’interpréter le prix d’un call comme le prix d’un put où le strike et la valeur actuelle de l’action sont intervertis. Par le biais de cette dualité, on montre qu’à un profil de volatilité correspond de manière univoque l’ensemble des prix des calls et puts américains, ce qui apporte une première réponse positive au problème de calibration.

Abstract

The first part of this thesis deals with credit risk. After a short introduction (in French) to this market and its modelling, we present a reduced-form model for the default time called SSRD (Shifted Square-Root Diffusion). One interesting feature of this model is that it can be automatically calibrated to the Credit Default Swap prices that come from the market. Moreover, it leaves the possibility to have dependence between the default intensity and the interest short rate. Then, we present a new family of copula functions named “periodic copulas” because their construction relies on periodic functions. Copula functions are used in the credit risk area to model dependence between different default times. Periodic copulas allow to explore a large range of dependence and can also be extended to the multivariate case. Furthermore, they can be easily sampled.

Then, we focus on the discretization schemes for the stochastic differential equation of Cox-Ingersoll-Ross which is used in the SSRD model. We present several discretization schemes of both the implicit and explicit types. We study their strong and weak convergence. We also examine numerically their behaviour and compare them to the schemes already proposed by Deelstra and Delbaen [27] and Diop [28]. Gathering all the results obtained, we recommend, in the standard case, the use of one of our explicit schemes.

In the last part we turn to the equity market. It is well known that it is possible, in a local volatility function model, to grab from the European option prices a volatility function which is consistent with. We consider here the analogous problem for American options: can we find a volatility function that explains all the American option prices observed? To tackle this problem, we consider perpetual American options and time-homogeneous local volatility functions. We put in evidence a relation named Call-Put Duality that allows to interpret a Put price as a Call price where strike value and spot value have been interchanged. Thanks to this duality result, we design a theoretical calibration procedure of the local volatility function from the perpetual Call and Put prices for a fixed spot price.

Remerciements

Il va de soi que mes premiers remerciements s'adressent à Benjamin Jourdain pour tout le temps qu'il m'a consacré au cours de ces trois années de thèse. J'ai particulièrement apprécié sa rigueur en matière de rédaction et la qualité de ses relectures. Je le remercie également pour son soutien et ses encouragements, et notamment pour m'avoir incité de nombreuses fois à présenter mes travaux.

Ringrazio Damiano Brigo di avermi accolto nel ormai mitico Product Development Group della Banca IMI che saluto anche. In particolare, voglio ringraziarlo di avermi introdotto al rischio di credito e ai problemi concreti della finanza. Ha saputo fare del mio periodo a Milano un momento di lavoro gradevole e produttivo.

Ensuite, je tiens à remercier Nicole El Karoui d'avoir accepté d'être rapporteur sur ma thèse. Cela me fait tout particulièrement plaisir car c'est à travers son cours que j'ai découvert le calcul stochastique et la finance. Je suis très honoré que Marek Rutkowski ait aussi accepté d'écrire un rapport sur ma thèse. Je suis également très heureux que Mireille Bossy et Damien Lamberton aient bien voulu être membre du jury.

Naturellement, je voudrais saluer tous mes collègues du CERMICS et notamment Bernard, Bouhari, Eric, Jean-François, Jérôme, Julien, Nicola, Simone et Tony avec qui j'ai eu le plaisir de travailler. Je tiens également ici à exprimer ma gratitude envers Mr Wigneron (et à travers lui plusieurs de mes professeurs de mathématiques) qui à un moment charnière a, par son enseignement, développé et renforcé mon attrait pour les mathématiques.

E infine, vurria finì sti ringrazi avendu un pensieru pè i cari, famiglia e amici, chi m'anu sempre datu curagiu e appoghju. A stu puntu di cunclusione di i studii, vogliu appruffittà di stu spaziu di espressione pè testimonià ricunnuscenza versu elli. Perché in qualchi manera, anu scrittu anch'elli un pezzu di sta tesi.

Table des matières

Introduction	12
I Risque de crédit	31
1 Introduction au risque de crédit	33
1.1 Présentation du marché du risque de crédit	33
1.1.1 Les produits à un seul sous-jacent (single-name)	34
1.1.2 Les produits sur panier (multi-name)	36
1.2 Modélisation	40
1.2.1 Les modèles structurels	40
1.2.2 Les modèles à intensité de défaut	41
1.3 Copules et dépendance	44
1.3.1 Définition et propriétés	44
1.3.2 Mesures de dépendance	47
1.3.3 Quelques familles de copules	49
1.3.4 Modèles à intensité avec plusieurs défauts	52
1.3.5 Interprétation du prix d'un CDO en terme de "base correlation" . .	53
2 Le modèle SSRD	55
2.1 Introduction	55
2.2 CDS's and Implied Hazard Rates	56
2.3 The SSRD intensity and interest rates model	58
2.3.1 CIR++ short-rate model (Brigo and Mercurio (2001))	58
2.3.2 CIR++ intensity model (Brigo and Alfonsi (2003))	59
2.3.3 Joint SSRD model calibration to CDS: Separability	60
2.3.4 Numerical simulation of the SSRD	62
2.3.5 Gaussian dependence mapping: A tractable approximated SSRD and numerical tests	63
2.3.6 The impact of correlation and numerical tests	66
2.4 Examples of pricing with the SSRD model	67

2.5	Conclusions and further research	68
3	Les copules “périodiques”	69
3.1	Introduction and Motivation	69
3.2	Periodic Copulas: The Bivariate Basic Case	71
3.3	A Bivariate Smooth Periodic Copulas	73
3.3.1	A first intuitive (non-smooth) construction	73
3.3.2	Regularization	76
3.4	Beyond the Bivariate Case	78
3.5	The Simulation of Periodic Copulas	79
3.6	Conclusions	79

II Schémas de discrétisations pour le processus de Cox-Ingersoll-Ross 81

4	Discrétisation du processus CIR	83
4.1	Introduction	83
4.2	Notations and preliminary lemmas	87
4.2.1	Some results on the CIR process	87
4.2.2	Introduction of the notations $\mathcal{O}(1/n^\delta)$ and $O(1/n^\delta)$	88
4.2.3	On the moments of the discretization schemes	89
4.2.4	Study of the expansion of the different schemes	91
4.3	Strong convergence	94
4.4	Weak convergence	96
4.5	Numerical results	102
4.5.1	Numerical study of the strong convergence	103
4.5.2	Numerical study of the weak convergence	106
4.5.3	Computation time required by the schemes	107
4.6	Conclusion	110
4.7	Appendix	111

III Une dualité Call-Put pour les options américaines perpétuelles, et calibration de la volatilité. 115

5	Dualité Call-Put	117
5.1	Perpetual American put and call pricing	119
5.2	Framework and notations	125
5.3	ODE for the exercise boundary	127

5.4	The call-put duality	130
5.4.1	The main result	130
5.4.2	An analytic example of dual volatility functions	132
5.5	Calibration procedure for the volatility	133
5.5.1	The calibration procedure	134
5.5.2	Calibration to “real” call and put prices	137
5.6	Invariant volatility functions through the duality	140
5.7	Conclusions and further developments	144

Introduction

La thèse que je présente se décompose en trois parties assez distinctes, mais qui ont toutes attrait à la modélisation stochastique en finance. La première partie est consacrée à la modélisation dans le marché du risque de crédit. La seconde se penche sur la simulation et la discrétisation de l'équation différentielle stochastique (EDS) de Cox-Ingersoll-Ross (CIR). Cette EDS est très largement utilisée en finance, et intervient en particulier dans notre modélisation du risque de crédit. Enfin, la troisième traite d'une dualité Call-Put pour les options américaines perpétuelles. Cette dualité permet notamment de fournir une procédure de calibration de la volatilité. Cette introduction présente, pour chacune de ces parties, les principaux résultats obtenus dans cette thèse en expliquant leur motivation. Elle évite au maximum de développer les arguments techniques qui seront, eux, abordés par la suite et essaye de ne présenter que les grandes idées nécessaires à la compréhension des travaux.

Risque de crédit

Le marché du risque de crédit est depuis une dizaine d'années en pleine expansion. Ce marché a pour objet d'échanger des produits financiers qui sont essentiellement des protections contre le risque de défaut d'une ou de plusieurs entreprises. On attribue souvent le développement récent de ce marché à la prise de conscience qu'a constitué la crise asiatique à la fin des années 1990, puis aux faillites retentissantes de la société ENRON ou de l'Etat d'Argentine en 2001. En Europe, la faillite en décembre 2003 de Parmalat a également fait couler beaucoup d'encre. Cela étant, si l'apparition de produits liquides traitant du risque de faillite est récente, les faillites et le problème de recouvrement de la dette existent eux depuis bien longtemps. On peut donc légitimement se demander pourquoi ce type de marché traitant du risque de défaut n'est pas apparu plus tôt. Nous ne nous étendrons bien évidemment pas sur cette question qui sort du cadre d'une thèse de mathématiques financières. Cependant, il est certain que les idées de Black, Merton et Scholes (1973) qui ont permis de voir un marché financier comme une source d'information et un moyen de se couvrir du risque sous-jacent au marché ne sont pas étrangères à l'émergence de nouveaux marchés, et pas seulement celui du risque de crédit. Le temps qui s'est écoulé entre ces travaux et l'émergence des dérivés de crédit peut s'expliquer au moins par deux raisons. D'une part il a fallu le temps nécessaire à la diffusion et l'acceptation de ces idées au sein des acteurs du marché. D'autre part, la spécificité des produits dérivés de crédit, avec notamment leur profil de paiement "tout ou rien", a certainement retardé leur développement.

Ainsi, nous commencerons par présenter dans le chapitre 1 de manière précise ce marché du risque de crédit en donnant la définition des principaux produits qui sont traités aujourd'hui. On en distingue deux grandes classes. Tout d'abord, les produits traitant du

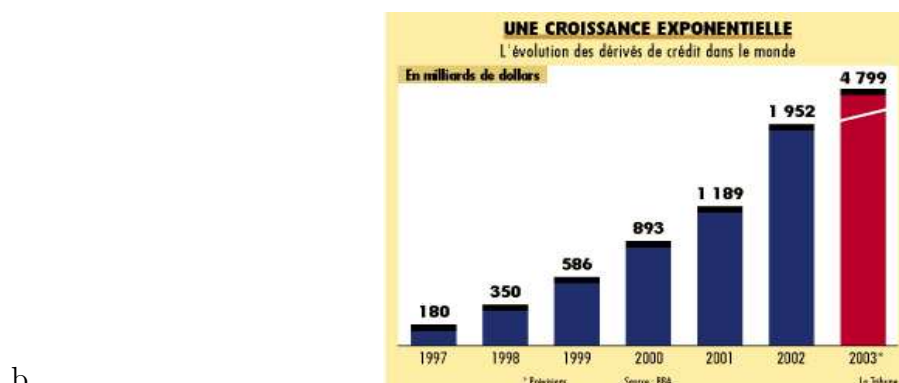


FIG. 1 – Evolution du volume des dérivés de crédit échangés dans le monde depuis 1997 (La Tribune 13/10/03).

défaut d’une seule entité (single-name). Le plus répandu est le Credit Default Swap qui est une protection contre le défaut d’une entité que nous nommons “C”. Imaginons qu’une autre entité “A” ait prêté une certaine somme d’argent à “C”, par exemple sous forme d’obligations, et souhaite ne pas avoir à supporter la perte de cet argent si “C” fait faillite. L’entité “A” va alors acheter un CDS qui l’engage à payer régulièrement au vendeur du CDS “B” une somme d’argent R . Il paye cette somme jusqu’à une certaine échéance fixée au préalable ou jusqu’au défaut de “C” s’il intervient avant. Si l’entité “C” fait défaut pendant cette période, l’acheteur reçoit de “B” une somme L_{GD} (Loss Given Default) qui idéalement correspond au préjudice subi par “A” à cause de la faillite de “C”. En revanche, si “C” ne fait pas défaut durant la période du contrat, le vendeur du CDS “B” aura reçu des paiements sans avoir rien à déboursier. Les CDS ont été “standardisés” afin de pouvoir être échangés sur le marché. Typiquement, les paiements réguliers de “A” à “B” ont lieu tous les trois mois, et les échéances du CDS peuvent être de un an, trois ans, cinq ans, sept ans ou dix ans. Ce type de produit permet bien évidemment à “A” de se couvrir du risque de défaut de l’entité “C” lui-même. Mais le fait qu’il soit coté et échangé permet également d’échanger le risque de détérioration de la qualité de crédit de “C”.

En revanche, ces produits ne permettent pas de se couvrir contre le risque de plusieurs faillites en cascades. Pourtant, il est fréquent que la faillite d’une grande entreprise entraîne la faillite d’autres entreprises. Pour citer un exemple, ENRON a entraîné dans sa chute le cabinet d’audit Arthur Andersen. Ce phénomène d’entraînement a certainement atteint son paroxysme durant la crise de 1929. Ce type de risque, appelé risque de contagion est en un certain sens plus dangereux car il a un impact financier plus grand pour les banques, mais également un impact économique et social pour les états. C’est afin de limiter ce risque que des règles plus contraignantes sur la gestion du risque de crédit ont été fixées lors des accords de Bâle I et II. Pour traiter ce risque de dépendance entre les différents défauts, plusieurs types de produits sont apparus, mais les plus répandus aujourd’hui sont certai-

nement les tranches de CDO (Collateralized Debt Obligation) appelés aussi simplement CDO. Le fonctionnement de ces tranches est le suivant. Supposons que “A” ait désormais prêté de l’argent à plusieurs entités “ C_1 ”, ..., “ C_n ” pour une période donnée : c’est le cas par exemple si “A” dispose d’un panier d’obligations. On suppose qu’il souhaite ne pas avoir à supporter le risque de faillite de ces entreprises. Pour cela, de manière analogue au CDS, “A” va entrer dans un CDO avec un vendeur “B”. Durant cette période, il s’engage alors à verser régulièrement une certaine somme R , et en échange “B” paye à “A” à chaque fois qu’une entité “ C_i ” fait faillite la somme qu’a perdu “A” en prêtant à “ C_i ”. A ce stade, un CDO ne serait rien d’autre qu’une somme de CDS. Mais en réalité, ce n’est pas ainsi que sont traités les CDO. Au lieu de faire supporter le risque de défaut à une seule entité “B”, on divise ce produit en tranches. Par exemple, “A” va acheter une tranche de CDO $[0, 5\%]$ à un vendeur “ B_1 ”, une tranche de CDO $[5\%, 15\%]$ à un vendeur “ B_2 ” et une tranche de CDO $[15\%, 1]$ à un vendeur “ B_3 ”. Il va leur payer régulièrement respectivement une somme R_1 , R_2 et R_3 . En échange “ B_1 ” paiera, à chaque fois qu’une entité “ C_i ” fait faillite la somme qu’a perdu “A” en prêtant à “ C_i ” et ce jusqu’à un montant représentant 5% de la somme initiale prêtée au total par “A”. Si la perte engendrée par les différents défauts dépasse les 5% de la valeur initiale, c’est alors “ B_2 ” qui prend le relais et assure la protection en cas de défaut et ce jusqu’à 15% de la valeur initiale. Enfin, si la perte va au-delà, ce sera à “ B_3 ” d’assurer la protection à “A”. Nous décrirons plus précisément les flux de paiements des tranches de CDO, mais les grandes lignes sont là. Pour être échangé et coté, ce type de produits est désormais standardisé. En Europe, le panier de référence est l’I-Traxx Europe qui est composé de 125 entreprises et est équipondéré, c’est à dire que la perte associée à chaque entreprise est la même. Les tranches de CDO sont également standardisées, la première ($[0, 5\%]$ dans notre exemple) est appelée tranche “equity” ou “junior”, les tranches intermédiaires ($[5\%, 15\%]$ dans notre exemple) sont appelées “mezzanine” et la tranche la plus élevée ($[15\%, 1]$ dans notre exemple) est appelée tranche “senior”. Ce découpage en tranches permet, contrairement à une somme de CDS ou à un découpage par noms de prendre en compte le risque de dépendance entre les défauts. De manière schématique, si toutes les entreprises faisaient faillite le même jour, la tranche “senior” aurait une grande valeur car dès qu’il y a défaut elle paie la totalité de la tranche (85% de la valeur initiale dans notre exemple). Si, au contraire, tous les défauts étaient indépendants, cette tranche senior aurait une plus faible valeur car il faudrait plusieurs défauts avant que cette tranche paie quelque chose et ensuite il faudrait que toutes les entreprises aient fait faillite pour qu’elle paie la totalité de la tranche. Ainsi, comme l’illustre cet exemple, le paiement régulier d’une tranche de CDO est sensible à la dépendance entre les défauts, et le fait d’avoir un marché sur les CDO permet de traiter de ce risque de dépendance entre les faillites.

Après une description précise des flux de paiements, nous présentons dans le chapitre 1



FIG. 2 – Evolution du volume des CDO échangés dans le monde depuis 1996 (Les Echos 30/01/06)

les deux grandes familles de modèles pour le défaut d'une entreprise. Tout d'abord nous regardons sans nous étendre quelques modèles structurels dont le principe est d'expliquer le défaut d'une entreprise à partir d'autres variables économiques. Par exemple, Merton [54] lie le défaut d'une entreprise à sa capacité à rembourser sa dette, et le modèle de Black et Cox [12] explique le temps de défaut d'une entreprise comme le premier temps de passage en dessous d'un certain seuil de la valeur de cette entreprise. Les modèles structurels sont aussi le point de vue utilisé par les agences de rating. Les principales agences sont Standard&Poor's, Moody's et Fitch. Elles fournissent régulièrement des notes sur la qualité de crédit de différentes entreprises ou collectivités publiques. La note d'une entité est établie en fonction de nombreux critères (son bilan, ses projets en cours,...). Elle est ainsi construite à partir d'autres variables et est en ce sens structurelle.

En revanche, toute une autre famille de modèles appelés modèles à intensité de défaut ou encore modèles à forme réduite cherchent à expliquer le défaut à partir d'un processus exogène appelé intensité de défaut. Ce processus décrit la probabilité instantanée qu'une entreprise donnée a de faire faillite. L'objet est alors d'extraire à partir des prix observés sur le marché du risque de crédit cette probabilité telle qu'elle est perçue par le marché. Bien évidemment, cette probabilité est influencée par les analyses structurelles puisque notamment les acteurs du marché tiennent compte des notations des agences de rating. Cependant, il peut y avoir quelques écarts entre la vision du marché et celle des analyses financières. Un exemple récent et très net de ce type d'écart a eu lieu sur le marché action lors de la "bulle internet" où le marché a anticipé des scénarios trop optimistes. En ce sens, on peut dire que les modèles à intensité sont par nature plus proche de l'opinion du marché tandis que les modèles structurels cherchent à donner une explication économique du défaut (endettement excessif, mauvais choix stratégiques...).

La littérature sur les modèles à intensité est assez vaste. Sans être exhaustifs, mentionnons ici les travaux de Jarrow et Turnbull [43] qui considèrent une intensité de défaut constante puis de Jarrow, Lando et Turnbull [42] qui ont modélisé l'intensité de défaut

comme un processus de Markov en temps continu avec un nombre fini d'état. Schönbucher [60] propose pour l'intensité un modèle de Vasicek étendu, et cette intensité est corrélée avec le taux d'intérêt qui suit également un modèle de Vasicek étendu. Enfin, Duffie et Singleton [29] proposent un modèle où l'intensité de défaut dépend d'une variable d'état.

Dans cette thèse, nous développons au chapitre 2 un modèle à intensité de défaut baptisé SSRD pour Shifted Square Root Diffusion. Il postule que l'intensité de défaut λ_t (appelée également parfois credit spread) s'écrit comme la somme d'une fonction déterministe $\psi(t)$ et d'un processus de Cox-Ingersoll-Ross

$$dy_t = \kappa(\mu - y_t)dt + \nu\sqrt{y_t}dZ_t.$$

Ici, $(Z_t, t \geq 0)$ désigne un mouvement brownien standard. Ce processus est bien défini et positif dès que l'on impose au paramètre $\beta = (y_0, \kappa, \mu, \nu)$ d'appartenir à $\mathbb{R}_+^4$. La fonction $\psi(t)$ est elle aussi supposée positive, afin que le taux instantané de faire défaut

$$\lambda_t = \psi(t) + y_t$$

soit positive. En appelant τ l'instant de défaut, la probabilité instantanée de faire défaut est décrite par ce taux : sur $\{\tau \geq t\}$,

$$\mathbb{P}(\tau \in [t, t+h] | \lambda_t) \underset{h \rightarrow 0}{\sim} \lambda_t h.$$

Ainsi, on définit pour avoir cette propriété l'instant de défaut par :

$$\tau = \Lambda^{-1}(\xi) \text{ avec } \Lambda(t) = \int_0^t \lambda_u du$$

où ξ est une variable aléatoire exponentielle de paramètre 1 indépendante de λ et des prix observés sur le marché. La classe de modèle "processus CIR + fonction déterministe" est appelée CIR++ (Brigo et Mercurio [21]) et a été introduite initialement pour modéliser le taux d'intérêt. Cette modélisation présente certains avantages pour la calibration que nous allons décrire puisque justement le modèle SSRD suppose que le taux d'intérêt court r_t suit également un modèle CIR++ :

$$r_t = \varphi(t) + x_t.$$

La fonction $\varphi(t)$ est supposée déterministe positive et $(x_t, t \geq 0)$ suit la dynamique $dx_t = k(\theta - x_t)dt + \sigma\sqrt{x_t}dW_t$. Comme précédemment, on suppose que $(W_t, t \geq 0)$ est un mouvement brownien standard et que le paramètre $\alpha = (x_0, k, \theta, \sigma)$ est dans $\mathbb{R}_+^4$ afin d'assurer l'existence de ce processus. Une particularité du modèle SSRD est de laisser la possibilité de mettre de la dépendance entre l'intensité de défaut et le taux d'intérêt à travers les mouvements browniens $(W_t, t \geq 0)$ et $(Z_t, t \geq 0)$ qui sont supposés être corrélés $\langle dW_t, dZ_t \rangle = \rho dt$ avec $\rho \in [-1, 1]$.

Un des points forts de ce modèle est qu'il peut être facilement calibré pour répliquer exactement les prix des CDS observés sur le marché, et qu'en outre cette calibration peut être effectuée indépendamment de la partie taux. Plus précisément, pour calibrer le taux d'intérêt, nous suivons la méthode indiquée par Brigo et Mercurio [21]. On commence par observer que, quelque soit le paramètre α , on peut trouver un shift $\varphi(t)$ qui calibre exactement les prix des zéro-coupons¹ obtenus à partir des prix du marché. On note $\varphi(t, \alpha)$ la fonction ainsi calibrée pour un paramètre α donné. Ensuite, α est calibré pour approcher de la meilleure façon la volatilité des cap observés sur le marché, typiquement en minimisant une distance avec ces volatilités. Cette optimisation est effectuée sous la contrainte que $(r_t, t \geq 0)$ doit être un processus positif. Les cap sont des produits financiers qui permettent de se protéger contre la hausse des taux d'intérêt. Comme souvent en finance, leurs prix sont interprétés en terme de volatilité implicite grâce à la formule de Black-Scholes, et ce sont ces volatilités que l'on calibre avec le modèle. Pour plus de détails, on se référera à [21]. De manière séparée à cette calibration, nous montrons que l'on peut en pratique, de manière analogue, trouver pour tout β une fonction $\psi(t)$ qui permet au modèle de donner les mêmes prix des CDS que ceux observés sur le marché. Pour justifier cette procédure, nous commençons par examiner le cas $\rho = 0$ où taux d'intérêt et intensité de défaut sont indépendants. Dans ce cadre, il est relativement simple de trouver pour tout β une fonction $\psi(t, \beta)$ qui réplique les prix des CDS. Nous expliquons alors pourquoi cette fonction convient toujours pour calibrer ces prix même quand $\rho \neq 0$. Cela est essentiellement dû au fait qu'ils sont très peu sensibles aux variations de ρ . Une fois la fonction $\psi(t, \beta)$ calibrée, les paramètres β et ρ restent libres pour approcher des prix de marché d'autres produits de risque de crédit, par exemple les prix des options d'achat de CDS. Au moment où le modèle a été développé, seuls les CDS étaient réellement liquides et le marché ne fournissait pas de tels prix, mais aujourd'hui ce type de calibration devient envisageable.

Avec ce modèle, nous abordons quelques aspects numériques. Le calcul de prix se faisant en général par méthode de Monte-Carlo, nous présentons une méthode de réduction de variance assez générique pour le risque de crédit. Il s'agit d'une méthode de stratification qui consiste à générer plus de simulations de scénarios où le défaut à lieu. En effet, beaucoup de produits en risque de crédit présentent des flux de paiements qui ont une faible variance lorsqu'il n'y a pas de défaut, mais qui sont très sensibles au temps de défaut. Dans un autre registre, nous donnons également des formules analytiques approchées pour certaines espérances qui apparaissent fréquemment dans le calcul de prix de produits dérivés de crédit.

Lorsque l'on aborde le problème du pricing de dérivés de crédit qui portent sur plusieurs noms comme les CDO, se pose le problème de la modélisation de la dépendance entre les

¹On rappelle que le prix d'un zéro-coupon d'échéance T est l'argent qu'il faut placer aujourd'hui pour disposer de la somme 1 en T .

défauts des différentes entreprises considérées. Etant donnée la structure des modèles à intensité de défaut, une modélisation très répandue est de mettre de la dépendance entre les variables exponentielles qui définissent les faillites de chaque entreprise. Cela permet d'utiliser le marché des CDS pour calibrer l'intensité de chacune, et le marché des CDO (et autres produits sur panier) pour capter une information sur la dépendance des instants de défaut. Pour modéliser cette dépendance entre les variables exponentielles appelées aussi seuils, on a recours à des fonctions appelées copules qui décrivent entièrement la structure de dépendance d'un vecteur aléatoire. Comme l'ensemble de ces fonctions est "très vaste", on postule en général que la copule qui décrit la dépendance entre les seuils appartient à une famille paramétrée par un "petit" nombre de paramètres. On cherche ensuite à en déduire à partir des prix du marché la valeur de ces paramètres. On fait face ainsi à deux objectifs contradictoires : d'une part on cherche à avoir une famille de copules assez riche pour modéliser correctement la dépendance, mais d'autre part ne disposant encore que de peu de prix, on est contraint de choisir une famille avec peu de paramètres afin de pouvoir les estimer correctement. Par conséquent, le choix de la famille que l'on prend est déterminant sur la qualité du modèle. La pratique du marché pour les CDO est de considérer la famille de copules gaussiennes à un seul facteur $\rho \in [0, 1]$. De manière grossière mais intuitive, cela revient à dire que dans la variable exponentielle qui gouverne le défaut d'une entreprise, une partie " $\sqrt{\rho}$ " vient de la santé générale de l'économie et une partie " $\sqrt{1-\rho}$ " est propre à l'entreprise. Par un mécanisme un peu compliqué (base correlation), le marché utilise ce modèle pour associer à chaque tranche de CDO une corrélation ρ implicite. Ce modèle est ainsi largement utilisé comme langage commun pour interpréter les prix des CDO en terme de dépendance ρ , de la même manière que le modèle de Black-Scholes est utilisé pour interpréter les prix en terme de volatilité implicite. Cela ne signifie pas en revanche qu'il soit utilisé effectivement (de même que le modèle de Black-Scholes) pour calculer des prix et des stratégies de couverture. Dans cette optique, il est légitime de s'intéresser à d'autres familles de copules.

Nous présentons ainsi dans le chapitre 1 les fonctions copules et quelques grandes familles généralement utilisées. Ensuite nous proposons dans cette thèse au chapitre 3 une nouvelle famille de copules qui a été baptisée "copules périodiques" car elles sont construites à partir de fonctions périodiques. On construit ainsi une famille de copules qui balaie un large spectre de dépendances comme c'est le cas pour les familles généralement utilisées. Une particularité intéressante des ces copules est qu'elles permettent éventuellement de donner un rôle asymétrique aux instants de défauts alors que beaucoup de familles de copules utilisées comme les copules gaussiennes donnent des rôles deux à deux symétriques ("2-exchangeable"). Ce type d'asymétrie peut être intéressant notamment pour la modélisation des défauts d'une grosse entreprise et d'un de ses fournisseurs. La faillite de la grosse entreprise risque de mettre en péril le fournisseur alors que la faillite du fournisseur a un impact moins décisif sur la santé de la grosse entreprise.

Simulation du processus de Cox-Ingersoll-Ross

La seconde partie (chapitre 4) traite de la discrétisation de l'équation différentielle stochastique de Cox-Ingersoll-Ross. Nous avons déjà utilisé ce type de processus lors de la modélisation du taux d'intérêt et de l'intensité de défaut dans la partie précédente. Ici, nous choisirons une paramétrisation légèrement différente de celle employée pour présenter le modèle SSRD. On considère ainsi l'EDS :

$$\begin{cases} X_t = x_0 + \int_0^t (a - kX_s)ds + \sigma \int_0^t \sqrt{X_s}dW_s, t \in [0, T] \\ x_0, \sigma, a \geq 0, k \in \mathbb{R}. \end{cases}$$

Le coefficient de diffusion étant $1/2$ -hölderien sur $\mathbb{R}_+^*$ et celui de dérive étant lipschitzien, nous avons unicité trajectorielle pour cette équation différentielle stochastique, et le processus solution est continu presque sûrement. Lorsque $\sigma = 2$ et $k = 0$, il s'agit d'un carré de Bessel de dimension a . D'un point de vue de la modélisation, ce processus est très utilisé car il possède quelques propriétés intéressantes. Tout d'abord, il est positif et lorsque $k > 0$ et il a un comportement de retour à la moyenne a/k avec la vitesse k . La positivité se comprend facilement : soit le processus (X_t) ne touche jamais zéro et reste strictement positif, soit il touche zéro et à cet instant le coefficient de diffusion est nul, tandis que le coefficient de dérive vaut $a \geq 0$ et “ramène” le processus vers les valeurs positives. Ensuite on connaît de manière analytique la transformée de Laplace de $(X_t, \int_0^t X_s ds)$ (voir e.g. Lamberton et Lapeyre [50]) ce qui peut être commode lorsqu'il faut calibrer les paramètres. Cette EDS a été regardée pour la première fois en finance par Cox-Ingersoll-Ross [25] pour modéliser le taux d'intérêt court. Dans un autre champ de la finance, Heston [39] a utilisé cette famille de processus pour modéliser la volatilité du cours d'une action.

La difficulté de discrétiser ce processus réside dans le caractère non lipschitzien du coefficient de diffusion. Etant donné que c'est en 0 que la racine carrée n'est pas lipschitzienne, la qualité de la discrétisation sera moins bonne lorsque l'approximation discrète se trouve proche de zéro. Or, on peut montrer (Lamberton et Lapeyre [50]) que lorsque $\sigma^2 \leq 2a$, le processus continu ne touche jamais zéro, tandis que si $\sigma^2 > 2a$, le processus a une probabilité positive de toucher 0 qui est égale à un quand $k \geq 0$. D'un point de vue heuristique, cela dit que plus σ est grand, plus le processus continu est susceptible de passer du temps au voisinage de 0, zone “non-lipschitzienne”. On verra en effet que la qualité des schémas de discrétisation dépend des paramètres, et elle a tendance à se dégrader lorsque σ augmente.

Au préalable, il est légitime d'avancer quelques motivations pour chercher à discrétiser cette EDS. Tout d'abord, de manière générale la discrétisation d'EDS est essentiellement utile pour calculer des espérances par la méthode de Monte-Carlo. Il faut commencer par mentionner que dans le cas du CIR, on sait simuler exactement les incréments de l'EDS puisqu'ils suivent des lois de chi-deux décentrés (voir e.g. Glassermann [37]). Cependant simuler exactement en loi plusieurs trajectoires s'avère coûteux en temps de calcul, et

l'emploi de techniques de discrétisation n'est donc pas superflu dès que l'on calcule des espérances faisant intervenir la trajectoire de $(X_t, t \geq 0)$. En outre, lorsque l'on souhaite corréler un processus CIR avec un autre processus, comme c'est le cas dans le modèle de Heston [39] ou dans le modèle SSRD présenté dans cette thèse, il n'y a pas de méthode facile pour simuler exactement le processus couplé alors qu'il est extrêmement simple de simuler les discrétisations avec des mouvements browniens corrélés. Pour ces deux raisons au moins, il est intéressant de disposer de schémas de discrétisation pour le CIR. Mentionnons tout de même ici que dans le cadre du modèle d'Heston, Broadie et Kaya [22] ont proposé une méthode pour simuler exactement les incréments du couple constitué du prix de l'action et de sa volatilité.

Avant de présenter les différents schémas étudiés dans cette thèse, il est d'abord utile de remarquer que les schémas usuels ne sont pas bien définis dans le cadre du CIR. On considère un intervalle de temps $[0, T]$ et la grille de discrétisation régulière $t_i = iT/n$ pour $n \in \mathbb{N}^*$ et $0 \leq i \leq n$. Le schéma d'Euler-Maruyama "naïf" pour le CIR

$$\hat{X}_{t_{i+1}}^n = \hat{X}_{t_i}^n + \frac{T}{n}(a - k\hat{X}_{t_i}^n) + \sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i})$$

avec $\hat{X}_{t_0}^n = x_0$ peut aboutir à des valeurs négatives (ce qui arrive en pratique) car l'incrément brownien peut prendre des valeurs fortement négatives. Ainsi, à l'étape suivante, la racine carrée n'est plus correctement définie. Il en est de même pour le schéma de Milstein qui peut amener à des valeurs négatives (au moins pour $k > 0$).

Parmi les schémas existants dans la littérature sur ce type de diffusion, Deelstra et Delbaen [27] ont proposé le schéma suivant

$$\hat{X}_{t_{i+1}}^n = \hat{X}_{t_i}^n + \frac{T}{n}(a - k\hat{X}_{t_i}^n) + \sigma\sqrt{\hat{X}_{t_i}^n \mathbf{1}_{\hat{X}_{t_i}^n > 0}}(W_{t_{i+1}} - W_{t_i}).$$

Il peut certes prendre des valeurs négatives mais reste toujours bien défini. Diop considère elle dans sa thèse [28] le schéma réfléchi

$$\hat{X}_{t_{i+1}}^n = |\hat{X}_{t_i}^n + \frac{T}{n}(a - k\hat{X}_{t_i}^n) + \sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i})|$$

qui est également bien défini et conserve la positivité.

Pour analyser mathématiquement la convergence des schémas numériques, on étudie en général deux types d'erreurs. La première, appelée erreur forte, mesure l'écart entre la trajectoire du processus et sa discrétisation pour un même aléa. Il y a plusieurs conventions possibles pour définir cette erreur. Nous choisirons dans cette thèse de prendre comme définition de l'erreur forte :

$$\mathbb{E} \left[\sup_{0 \leq i \leq n} |\hat{X}_{t_i}^n - X_{t_i}| \right].$$

Cependant cette définition de l'erreur peut paraître restrictive. En effet, elle mesure une convergence ω par ω alors qu'en général on utilise la discrétisation d'EDS pour calculer

une espérance par la méthode de Monte-Carlo. Dans ce cas, il est plus judicieux d'estimer l'erreur en loi induite par la discrétisation. Ce type d'estimation est appelé erreur faible. Toutefois, l'étude de l'erreur commise sur l'espérance d'une fonctionnelle générale de la trajectoire n'est pas simple. Ainsi, nous étudierons dans cette thèse l'erreur faible commise sur la loi de X_T , c'est à dire que l'on s'intéressera au comportement asymptotique de

$$\mathbb{E}(f(X_T)) - \mathbb{E}(f(\hat{X}_T^n)) = \mathbb{E}(f(X_T)) - \mathbb{E}(f(\hat{X}_T^n))$$

pour une classe de fonction f caractérisant la loi de X_T la plus large possible.

Concernant les schémas présentés, Deelstra et Delbaen [27] ont établi que l'erreur forte de leur schéma convergeait bien vers 0 lorsque le pas de temps tend vers 0. Berkaoui, Bossy et Diop [28, 9] ont montré que sous certaines conditions sur les coefficients du processus, l'erreur forte du schéma réfléchi est en $O(1/\sqrt{n})$. Pour l'erreur faible, Bossy et Diop [28, 14] ont montré que si $\sigma^2 \leq a$, et pour des fonctions $\mathcal{C}^4$ bornées à dérivées bornées, $\mathbb{E}(f(X_T)) - \mathbb{E}(f(\hat{X}_T^n)) = O(1/n)$. Ces vitesses de convergences sont les mêmes que pour le schéma d'Euler-Maruyama pour les EDS dont les coefficients sont lipschitziens.

Dans cette thèse nous considérons également de nouveaux schémas pour le processus CIR :

- En implicitant le coefficient de dérive et de diffusion de l'EDS, on obtient le schéma

$$\hat{X}_{t_{i+1}}^n = \left(\frac{\sigma(W_{t_{i+1}} - W_{t_i}) + \sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}_{t_i}^n + (a - \frac{\sigma^2}{2})\frac{T}{n})(1 + k\frac{T}{n})}}{2(1 + k\frac{T}{n})} \right)^2$$

qui est bien défini et reste positif lorsque $\sigma^2 \leq 2a$ et le pas de temps est assez petit.

- En implicitant le coefficient de dérive dans l'EDS satisfaite par $\sqrt{X_t}$, on obtient de la même manière le schéma

$$\hat{X}_{t_{i+1}}^n = \left(\frac{\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}_{t_i}^n} + \sqrt{(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}_{t_i}^n})^2 + 4(1 + \frac{kT}{2n})\frac{a - \sigma^2/4}{2}\frac{T}{n}}}{2(1 + \frac{kT}{2n})} \right)^2.$$

Pour $\sigma^2 \leq 4a$ et un pas de temps assez petit, ce schéma est bien défini et reste positif.

- Enfin, à partir des développements limités de ces deux schémas en $\frac{T}{n}$ et $W_{t_{i+1}} - W_{t_i}$, on construit une famille de schémas explicites $E(\lambda)$

$$\begin{aligned} \hat{X}_{t_{i+1}}^n &= \left(\left(1 - \frac{kT}{2n}\right) \sqrt{\hat{X}_{t_i}^n} + \frac{\sigma(W_{t_{i+1}} - W_{t_i})}{2(1 - \frac{kT}{2n})} \right)^2 \\ &\quad + (a - \sigma^2/4)T/n + \lambda[(W_{t_{i+1}} - W_{t_i})^2 - T/n]. \end{aligned}$$

Ces schémas sont bien définis dès lors que le pas de temps est assez fin et que l'on a $0 \leq \lambda \leq a - \sigma^2/4$.

Nous commençons par étudier la convergence forte de ces schémas, et nous montrons que le premier et le troisième ont une erreur forte qui tend vers zéro quand $n \rightarrow +\infty$. Nous étudions également d'un point de vue numérique l'erreur forte de ces schémas, ainsi que ceux de Deelstra et Delbaen, et de Diop. Ainsi, nous observons notamment l'influence des paramètres sur la vitesse de convergence forte. Ensuite, nous effectuons une analyse de l'erreur faible et obtenons, toujours pour le premier et le troisième schémas, un comportement en $O(1/n)$ pour des fonctions f de classe $\mathcal{C}^4$ dont les dérivées sont à croissance polynomiale. Pour les schémas explicites $E(\lambda)$, nous montrons que l'on peut même développer l'erreur faible sous la forme $\mathbb{E}(f(X_T)) - \mathbb{E}(f(\hat{X}_T^n)) = c_1/n + \dots + c_p/n^p + O(1/n^{p+1})$ quelque soit $p \in \mathbb{N}^*$ pourvu que f soit $\mathcal{C}^\infty$ avec des dérivées à croissance polynomiale. Ce développement de l'erreur faible est utile pour pouvoir utiliser la méthode de Romberg (ou Richardson). En effet, on sait qu'avec un tel développement $2\mathbb{E}[f(\hat{X}_T^{2n})] - \mathbb{E}[f(\hat{X}_T^n)]$ converge vers $\mathbb{E}(f(X_T))$ avec une vitesse en $O(1/n^2)$, alors que le temps de calcul demandé pour calculer $2\mathbb{E}[f(\hat{X}_T^{2n})] - \mathbb{E}[f(\hat{X}_T^n)]$ est seulement environ 3 fois supérieur à celui nécessaire pour calculer $\mathbb{E}[f(\hat{X}_T^n)]$. Nous illustrons ces résultats d'un point de vue numérique et nous observons des qualités de convergence bien différentes selon les schémas. Nous concluons cette étude par des tables comparatives de ces schémas récapitulant pour chaque schéma les propriétés démontrées ainsi que les propriétés observées lors des simulations numériques. Cela permet de mettre en évidence le schéma $E(0)$ qui réunit le plus de propriétés désirées en général pour un schéma de discrétisation.

Une Dualité Call-Put pour les options américaines perpétuelles

La troisième partie (chapitre 5) traite de la calibration de la volatilité à partir des prix d'options américaines. Il est bien connu désormais que la formule de Dupire [30] permet de calibrer automatiquement la fonction de volatilité locale aux prix des options européennes. Cette formule est également utile pour extrapoler, à partir des prix de marché des calls et puts européens, les prix des calls et puts européens pour des échéances et des maturités non cotées. De manière idéale, on souhaiterait développer une méthode analogue qui permette de trouver une fonction de volatilité qui réplique les prix des options américaines. Afin de commencer à traiter ce problème difficile, nous regardons dans cette thèse les options américaines perpétuelles et nous obtenons une méthode pour calibrer la volatilité locale lorsque celle-ci est supposée ne pas dépendre du temps.

Commençons par rappeler comment fonctionne la formule de Dupire dans le cadre Européen. On suppose que le prix d'une action suit la dynamique sous la probabilité risque-neutre

$$dS_t = (r - \delta)S_t dt + \sigma(t, S_t)S_t dW_t$$

où $(W_t, t \geq 0)$ désigne un mouvement brownien standard. Le taux d'intérêt r est supposé constant et positif tandis que δ , supposé également positif et constant, représente le taux de dividende distribué par action. La fonction $\sigma : \mathbb{R}_+ \times \mathbb{R}_+^* \rightarrow \mathbb{R}_+$ est appelée volatilité locale. On suppose que cette fonction est continue et est telle que S_t admette une densité $p(t, x)$. C'est le cas par exemple si $t \mapsto \sigma(t, 0)$ est strictement positive et

$$\forall t \leq T, \exists K_T > 0, \forall x, z \geq 0, |x\sigma(t, x) - z\sigma(t, z)| \leq K_T|x - z|$$

en utilisant le Théorème 2.1.3 de Bouleau et Hirsch [16]. Pour $T > 0$ et $K > 0$, on note $P(T, K) = \mathbb{E}[e^{-rT}(K - S_T)^+]$ (resp. $C(T, K) = \mathbb{E}[e^{-rT}(S_T - K)^+]$) le prix d'un put (resp. call) européen de maturité T et de strike K . Dans ce contexte Dupire [30] puis dans un cadre un peu plus général Klebaner [48] ont montré que le prix d'un put résout l'EDP suivante :

$$\begin{cases} \partial_T P(T, K) - \frac{1}{2}K^2\sigma^2(T, K)\partial_K^2 P(T, K) + (r - \delta)K\partial_K P(T, K) + \delta P(T, K) = 0, T > 0 \\ P(0, K) = (K - S_0)^+. \end{cases}$$

ce qui permet de déduire la volatilité locale à partir de la nappe des prix de put. Rappelons les arguments qui permettent d'aboutir à cette EDP. On applique la formule d'Itô-Meyer-Tanaka à $(K - S_t)^+$ sur $[0, T]$:

$$(K - S_T)^+ = (K - S_0)^+ + \int_0^T (\delta - r)S_t \mathbf{1}_{S_t \leq K} dt - \int_0^T \sigma(t, S_t)S_t \mathbf{1}_{S_t \leq K} dW_t + \frac{1}{2}L_T^K(S)$$

où $L_T^K(S)$ désigne le temps local en K à l'instant T du processus $(S_t, t \geq 0)$. En prenant l'espérance, on observe que le terme martingale est bien d'espérance nulle $((t, x) \mapsto x\sigma(t, x))$ est bornée sur $[0, T] \times [0, K]$, et on obtient alors par Fubini,

$$\mathbb{E}[(K - S_T)^+] = (K - S_0)^+ + (\delta - r) \int_0^T \mathbb{E}[S_t \mathbf{1}_{S_t \leq K}] dt + \frac{1}{2} \mathbb{E}[L_T^K(S)].$$

Par définition du temps local, pour toute fonction mesurable bornée positive, on a

$$\int_{-\infty}^{+\infty} f(x) L_T^x(S) dx = \int_0^T f(S_t) S_t^2 \sigma^2(t, S_t) dt.$$

et par le théorème de Fubini, on obtient $\int_{-\infty}^{+\infty} f(x) \mathbb{E}[L_T^x(S)] dx = \int_0^T \mathbb{E}[f(S_t) S_t^2 \sigma^2(t, S_t)] dt = \int_0^T \int_{-\infty}^{+\infty} f(x) x^2 \sigma^2(t, x) p(t, x) dx$. En utilisant une nouvelle fois Fubini, il vient alors que $\mathbb{E}[L_T^x(S)] = \int_0^T x^2 \sigma^2(t, x) p(t, x) dt$. Par conséquent, on obtient

$$\partial_T \mathbb{E}[(K - S_T)^+] = (\delta - r) \mathbb{E}[S_T \mathbf{1}_{S_T \leq K}] + \frac{1}{2} K^2 \sigma^2(T, K) p(T, K).$$

En remarquant enfin que $S_T \mathbf{1}_{S_T \leq K} = -(K - S_T)^+ + K \mathbf{1}_{S_T \leq K}$ et $e^{-rT} \mathbb{P}(S_T \leq K) = \partial_K P(T, K)$, on en déduit

$$\partial_T P(T, K) = -\delta P(T, K) + (\delta - r)K \partial_K P(T, K) + \frac{1}{2}K^2 \sigma^2(T, K) \partial_K^2 P(T, K).$$

Notons ici que si l'on est dans le cadre de Bouleau-Hirsch pour l'existence de la densité, cette EDP est valable presque partout en utilisant le théorème de dérivation de Lebesgue, dans la mesure où l'on n'a pas d'information sur la continuité de la densité (c'est juste une fonction L^1 Lebesgue mesurable). On remarque en passant qu'en utilisant la parité call-put $C(T, K) - P(T, K) = S_0 e^{-\delta T} - K e^{-rT}$, on trouve que le call satisfait aussi la même EDP.

Outre le fait que cette EDP permette de retrouver la fonction de volatilité, elle fait apparaître également une EDP de pricing (Feynman-Kac). En effet, si l'on pose pour une maturité $\mathbf{T} > 0$ $\tilde{C}(T, K) := P(\mathbf{T} - T, K)$, on observe que

$$\begin{cases} \partial_T \tilde{C}(T, K) + \frac{1}{2}K^2 \sigma^2(\mathbf{T} - T, K) \partial_K^2 \tilde{C}(T, K) + (\delta - r)K \partial_K \tilde{C}(T, K) - \delta \tilde{C}(T, K) = 0, T \in (0, \mathbf{T}) \\ \tilde{C}(\mathbf{T}, K) = (K - S_0)^+. \end{cases}$$

Ainsi, en définissant $\tilde{S}^{T,K}$ la solution de

$$d\tilde{S}_t^{T,K} = (\delta - r)\tilde{S}_t^{T,K} dt + \sigma(\mathbf{T} - t, \tilde{S}_t^{T,K}) \tilde{S}_t^{T,K} dW_t, t \geq T$$

telle que $\tilde{S}_T^{T,K} = K$, on peut interpréter le prix d'un put $P(\mathbf{T}, K)$ comme le prix d'un call sur $\tilde{S}$:

$$P(\mathbf{T}, K) = \mathbb{E}[e^{-\delta \mathbf{T}} (\tilde{S}_{\mathbf{T}}^{0,K} - S_0)^+].$$

Cette relation, qui sera appelée dualité Call-Put, fait donc apparaître le prix d'un put de spot S_0 , strike K , taux d'intérêt r , taux de dividende δ et volatilité locale $(\sigma(t, x), t \in [0, \mathbf{T}])$ comme celui d'un call de spot K , strike S_0 , taux d'intérêt δ , taux de dividende r et volatilité locale $(\sigma(\mathbf{T} - t, x), t \in [0, \mathbf{T}])$. D'un point de vue de la calibration, le point crucial est l'interversion $S_0 \leftrightarrow K$. En effet le marché cote ces options pour plusieurs strikes K . Grâce à cette dualité, on peut interpréter ces prix comme des prix de put avec plusieurs spots K , puis utiliser l'EDP de Feynman-Kac pour obtenir des informations sur la volatilité locale. Pour les options européennes, ce point de vue est équivalent à la formule de Dupire et, à la limite obscurcit le discours. Cependant, cette vision s'avère plus fructueuse quand il s'agit de développer une calibration analogue pour d'autres types d'options comme les options américaines. En effet, le moyen présenté ci-dessus pour obtenir l'EDP de Dupire est en quelque sorte spécifique au cas Européen. En revanche, on dispose d'EDP de pricing de type Feynman-Kac pour une plus grande classe d'options. C'est pour cela qu'interpréter l'EDP de Dupire comme l'EDP de pricing d'une option duale (i.e. où l'on a intervertit S_0 et K) laisse ouverte la perspective d'étendre ce type de résultat à d'autres options.

Dans le cadre américain, on souhaiterait ainsi pouvoir interpréter le prix d'un call ou d'un put de strike K comme le prix, sous un autre modèle, d'une option dont le sous-jacent

a pour valeur initiale K . Si le prix de l'action est modélisé comme précédemment avec une fonction de volatilité locale $\sigma(t, x)$, le prix d'un put américain d'échéance $\mathbf{T}$ et de strike K est donné par

$$\sup_{\tau \in \mathcal{T}_{0, \mathbf{T}}} \mathbb{E} \left[e^{-r\tau} (K - S_\tau)^+ \right].$$

Ici, $\mathcal{T}_{0, \mathbf{T}}$ désigne l'ensemble des temps d'arrêt par rapport à la filtration naturelle associée au processus $(S_t, t \geq 0)$ à valeurs dans $[0, \mathbf{T}]$. Guidés par l'exemple européen, il est légitime de s'interroger sur une éventuelle dualité et voir si on a une égalité du type :

$$\sup_{\tau \in \mathcal{T}_{0, \mathbf{T}}} \mathbb{E} \left[e^{-r\tau} (K - S_\tau)^+ \right] \stackrel{?}{=} \sup_{\tau \in \mathcal{T}_{0, \mathbf{T}}} \mathbb{E} \left[e^{-\delta\tau} (\tilde{S}_\tau^{0, K} - S_0)^+ \right]$$

Il est d'ailleurs facile de vérifier que dans le cas du modèle de Black-Scholes où $\sigma(t, x)$ est constante, cette égalité est vraie.

Pour aborder ce problème, nous regardons dans cette thèse le cas des options américaines perpétuelles ($\mathbf{T} = +\infty$), c'est à dire que l'on peut exercer l'option sans limite dans le temps. Nous supposons également que la volatilité locale ne dépend pas du temps et on la note $\sigma(x)$. Cette hypothèse simplifie le problème dans la mesure où le prix d'un put ou d'un call américain ne dépendent plus de la date du début de contrat. On note ainsi désormais

$$P(S_0, K) = \sup_{\tau \in \mathcal{T}_{0, +\infty}} \mathbb{E} \left[e^{-r\tau} (K - S_\tau)^+ \right]$$

le prix d'un put américain de strike K dont le cours initial du sous-jacent est S_0 . On peut montrer alors l'existence d'une frontière d'exercice $K \mapsto x^*(K)$ telle que

$$P(S_0, K) = K - S_0 \text{ si } S_0 \leq x^*(K), \text{ et } P(S_0, K) = \left[e^{-r\tau_{x^*(K)}} (K - S_{\tau_{x^*(K)}})^+ \right]$$

où $\tau_{x^*(K)}$ est le premier temps d'atteinte de $x^*(K)$ par le processus $(S_t, t \geq 0)$. On appelle zone d'exercice l'ensemble $\{S_0 \leq x^*(K)\}$ et zone de continuation l'ensemble $\{S_0 \geq x^*(K)\}$. De même, on peut montrer que dans un modèle à volatilité locale indépendante du temps, le prix d'un call américain où l'on a intervertit strike et spot ($S_0 \leftrightarrow K$), et taux d'intérêt et taux de dividende ($r \leftrightarrow \delta$) qui est défini par²

$$c(K, S_0) = \sup_{\tau \in \mathcal{T}_{0, +\infty}} \mathbb{E} \left[e^{-\delta\tau} (\tilde{S}_\tau^{0, K} - S_0)^+ \right]$$

vaut :

$$c(K, S_0) = K - S_0 \text{ si } K \geq y^*(S_0), \text{ et } c(K, S_0) = \mathbb{E} \left[e^{-\delta\tilde{\tau}_{y^*(S_0)}} (\tilde{S}_{\tilde{\tau}_{y^*(S_0)}}^{0, K} - S_0)^+ \right],$$

²Dans toute la partie concernant la dualité call put, la première variable désigne le spot, c'est à dire la valeur initiale du sous-jacent, et la seconde le strike de l'option.

où $\tilde{\tau}_{y^*(S_0)}$ est le premier instant où le processus $(\tilde{S}_t^{0,K}, t \geq 0)$ touche $y^*(S_0)$. La fonction y^* est appelée frontière d'exercice relative au call, et quand elle existe, dépend bien évidemment de la dynamique de $\tilde{S}$. La question que l'on se posait dans le cas général de l'existence d'une formule de dualité s'écrit dans ce le cas perpétuel :

$$\text{A-t-on : } \forall S_0, K > 0, \quad P(S_0, K) = c(K, S_0) ?$$

D'emblée, on observe que si l'on a une telle relation, les zones d'exercices doivent coïncider $\{(S_0, K) \in (\mathbb{R}_+^*)^2, K \geq y^*(S_0)\} = \{(S_0, K) \in (\mathbb{R}_+^*)^2, S_0 \leq x^*(K)\}$ et par conséquent les deux frontières d'exercice sont nécessairement réciproques l'une de l'autre. Ce que nous montrons dans cette thèse, c'est que dans un cadre que nous allons maintenant préciser, cette condition est également suffisante.

Nous supposons que $dS_t = S_t((r-\delta)dt + \sigma(S_t)dW_t)$ et $d\tilde{S}_t^{0,K} = ((\delta-r)dt + \eta(\tilde{S}_t^{0,K})dW_t)$, où les fonctions de volatilité σ et η sont supposées être continues, minorées et majorées par deux constantes strictement positives. Dans ce cadre, nous sommes alors capable de montrer l'existence des frontières d'exercice x^* et y^* . Nous montrons également qu'en ces frontières, les prix sont des fonctions $\mathcal{C}^1$ de la valeur initiale du sous-jacent (Principe de Smooth-fit). Grâce aux relations de continuités induites par ce principe, nous établissons que la condition de réciprocity sur les frontières d'exercice $x^* = (y^*)^{-1}$ est équivalente à la relation de dualité $\forall S_0, K > 0, \quad P(S_0, K) = c(K, S_0)$. Pour cela, nous exploitons la forme particulière des prix des puts et calls américains perpétuels : ils s'écrivent dans la zone de continuation comme le produit d'une fonction du strike et d'une fonction du spot. Enfin, le principe de Smooth-fit et les équations différentielles satisfaites par les prix permettent d'obtenir que x^* et y^* sont chacune unique solution d'une équation différentielle ordinaire faisant intervenir respectivement σ et η . Ces EDO permettent ainsi facilement d'écrire et d'interpréter la condition $x^* = (y^*)^{-1}$ en terme des volatilités σ et η . Nous obtenons par ce biais une condition entre σ et η qui est équivalente à la relation de dualité. Contrairement au cas Européen, nous observons qu'en général les fonctions de volatilité σ et η qui satisfont cette condition ne sont pas les mêmes. Nous montrons plus précisément que si on a la relation de dualité pour une même fonction de volatilité σ , alors sous certaines conditions techniques ($\delta \leq r$ et σ analytique au voisinage de 0), nécessairement σ est une fonction constante, c'est à dire qu'il s'agit du modèle de Black-Scholes.

Du point de vue de la calibration, supposons que nous soyons dans un marché virtuel où nous connaîtrions pour tous les strikes $K > 0$ le prix des puts (resp. des calls) américains. Par cette dualité, nous interprétons ces prix comme des prix de calls (resp. de puts) américains de strike S_0 partant de la valeur initiale K . Nous pouvons dès lors utiliser l'équation de pricing pour récupérer (une partie de) la fonction de volatilité η_C (resp. η_P) pour les calls (resp. puts) duaux. En utilisant la relation liant la volatilité initiale σ avec la volatilité duale, nous pouvons alors obtenir la valeur de $\sigma(x)$ pour $x \in [0, S_0]$ (resp. $x \in [S_0, +\infty)$). Dans le cas où les prix observés sont supposés provenir d'un modèle à vo-

latilité locale homogène en temps, cette procédure fonctionne sans encombre. Elle nous dit que la donnée de la fonction $\sigma(x)$ est équivalente à celle des prix des put et call américains. Dans le cas où les prix proviendraient d'un marché "réel", nous donnons des conditions suffisantes pour pouvoir interpréter ces prix dans un modèle à volatilité $\sigma(x)$. Un point tout à fait remarquable de cette procédure de calibration est que les prix des calls et des puts américains perpétuels donnent dans ce cadre des informations complémentaires sur σ , alors que par exemple dans le cas Européen ces informations sont identiques en vertu de la parité call-put.

Première partie

Risque de crédit

Chapitre 1

Introduction au risque de crédit

Ce chapitre introductif a pour objet de présenter succinctement le marché du risque de crédit, ainsi que sa modélisation afin de motiver les travaux de cette thèse qui portent sur ces thématiques. La première partie présente les principaux produits traités sur le marché du risque de crédit tels que les CDS et les CDO. Pour une description financière détaillée de ce marché, on pourra se référer au livre de Bruyère et al. [23]. La seconde partie donne un bref aperçu de la modélisation en risque de crédit et présente le cadre des modèles à intensité de défaut dans lequel le modèle SSRD s'inscrit. Comme référence sur la modélisation du marché du risque de crédit, citons ici le livre de Bielecki et Rutkowski [11]. Enfin, la troisième partie est une introduction aux fonctions copules, qui sont utilisées pour modéliser la dépendance entre plusieurs défauts. Cela permet de situer la famille des copules dites “périodiques” introduites dans cette thèse. Pour un ouvrage détaillé sur ce sujet, on pourra se référer au livre de Nelsen [57].

1.1 Présentation du marché du risque de crédit

Dans cette partie nous allons nous attacher à décrire précisément les flux de paiement (payoff) des produits les plus courants. La description des flux n'est pas toujours aisée car les produits financiers deviennent de plus en plus complexes. Il faut traduire une description financière du produit comme par exemple celle qui suit du CDO (Fig. 1.1) en langage mathématique. On peut être amené lors de cette transcription à négliger certains termes ou faire quelques approximations afin d'obtenir un payoff relativement synthétique. Les payoffs présentés ici sont les plus communément utilisés dans les modèles mathématiques et comportent les caractéristiques essentielles des produits, mais il peut cependant y avoir quelques écarts par rapport aux payoffs réels.

Nous prendrons comme convention de décrire les flux par leur valeur actualisée. Ainsi nous noterons, pour $t < T$, $D(t, T)$ le coefficient aléatoire d'actualisation entre t et T . Ce

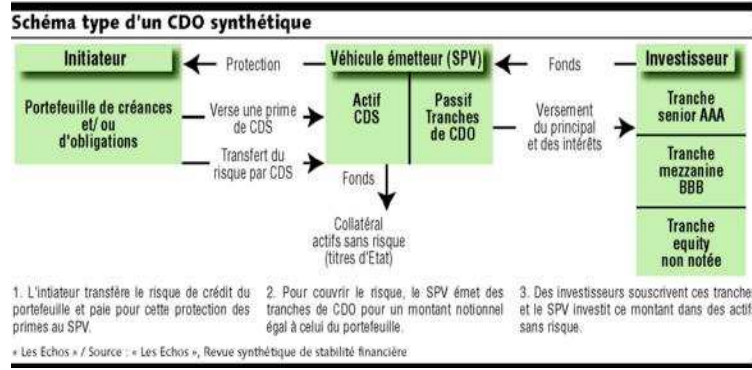


FIG. 1.1 – Description du fonctionnement des CDO (Les Echos 30/01/06)

coefficient vaut l'inverse de la valeur en T d'un portefeuille composé uniquement d'actif sans-risque avec une valeur initiale 1 en t . On supposera que l'on est à l'instant $t = 0$ où commencent chacun des contrats et considérera un échéancier $\mathcal{T} = \{T_1, \dots, T_n\}$ avec $0 < T_1 < \dots < T_n$ et $n \in \mathbb{N}^*$. On définit également $T_0 = 0$ et pour $i \in \{1, \dots, n\}$ on note $\alpha_i = T_i - T_{i-1}$ le temps écoulé entre T_{i-1} et T_i . Enfin, pour $t \in [0, T_n)$, on notera $\beta(t)$ l'indice appartenant à $\{1, \dots, n\}$ tel que $t \in [T_{\beta(t)-1}, T_{\beta(t)})$.

1.1.1 Les produits à un seul sous-jacent (single-name)

Dans ce paragraphe, τ désignera le temps de défaut de l'entité considérée que nous appelons "C". On le suppose positif pour qu'au moment où commence le contrat, la faillite de l'entité sous-jacente n'ait pas encore eu lieu.

Le zéro-coupon avec défaut

C'est le produit le plus simple à décrire, mais il s'agit plus d'un produit théorique qui apparaît fréquemment comme sous-composante de produits plus complexes, et il n'est pas en pratique très échangé. C'est un contrat qui paye à une échéance $T > 0$ prédéfinie la somme 1 si "C" n'a pas fait faillite. Son payoff s'écrit donc

$$D(0, T) \mathbf{1}_{\tau > T}.$$

Une obligation (avec défaut)

On considère une obligation émise par "C" du point de vue de son acheteur que nous appelons "A". "A" paye à "C" une somme nominale 1 à l'instant 0 qui en retour s'engage à payer des intérêts à chaque échéance d'un échéancier prédéfini $\mathcal{T}$ et à rembourser à la dernière échéance T_n la somme donnée par "A". Les intérêts payés à chaque échéance T_i

sont définis au début du contrat, nous supposons ici qu'ils sont proportionnels à la durée écoulée depuis la dernière échéance, c'est à dire qu'ils sont de la forme $X\alpha_i$ où $X > 0$. En cas de défaut de "C", "A" cesse de recevoir ces intérêts et en général récupère une partie $Rec \in [0, 1]$ (Recovery) de la somme initiale fixée. Nous écrirons ici le payoff en supposant que ce recouvrement a lieu à l'instant de défaut, ce qui est "optimiste" pour "A" puisqu'il y a en général un certain délai entre la faillite et la liquidation des biens. Le flux de paiements que recevra "A" après l'achat de l'obligation s'écrit :

$$\sum_{i=1}^n \alpha_i X D(0, T_i) \mathbf{1}_{\tau > T_i} + D(0, T_n) \mathbf{1}_{\tau > T_n} + Rec D(0, \tau) \mathbf{1}_{\tau \leq T_n}.$$

Le Credit Default Swap (CDS)

Supposons que "A" ait souscrit une certaine somme en obligations auprès de "C" et souhaite ne pas perdre sa créance si ce dernier fait défaut. Pour ce faire, il va entrer dans un Credit Default Swap auprès d'une autre entité "B". Pour une somme nominale 1, cela l'engage, selon un échéancier qui peut être différent de celui des obligations mais que nous notons toujours $\mathcal{T}$, à payer à chaque échéance une somme proportionnelle au temps écoulé depuis la dernière échéance $R\alpha_i$ où $R > 0$. Ces paiements cessent si le défaut de "C" intervient avant T_n . En échange, "B" s'engage à lui verser à l'instant de défaut de "C" la partie $L_{GD} = 1 - Rec$ (Loss Given Default) qu'il aura perdu, si le défaut intervient pendant la période $[0, T_n]$. "A" sera donc couvert contre le défaut de "C" sur cette période. Pour "B", le flux des paiements du CDS s'écrit

$$\sum_{i=1}^n \alpha_i R D(0, T_i) \mathbf{1}_{\tau > T_i} + R(\tau - T_{\beta(\tau)-1}) D(0, \tau) \mathbf{1}_{\tau \leq T_n} - L_{GD} D(0, \tau) \mathbf{1}_{\tau \leq T_n}.$$

La quantité R est appelée le *taux du CDS*. Le terme $R(\tau - T_{\beta(\tau)-1}) D(0, \tau) \mathbf{1}_{\tau \leq T_n}$ correspond au paiement de la protection entre la dernière échéance avant le défaut et le défaut. Il est parfois omis pour simplifier certains calculs dans la mesure où il ne change pas fondamentalement la structure du produit.

Notons ici que pour "A", le fait que l'échéancier du CDS coïncide avec celui de l'obligation peut être commode dans la mesure où la protection contre le défaut est vue comme un reversement d'une partie R de ses intérêts X . L'achat simultané d'une obligation avec défaut et d'un CDS protégeant contre ce défaut peut alors être vu comme l'achat d'une obligation sans défaut d'échéance finale $\min(\tau, T_n)$ reversant des intérêts selon le taux $X - R$. C'est pour cela que l'on appelle R également *credit spread* ou *spread de CDS*. Mais le terme de credit spread n'est pas exclusivement employé pour désigner ce taux. On l'emploie également comme autre nom de l'intensité de défaut, et de manière générale pour appeler toute variable qui "mesure" l'écart entre le prix d'un produit tenant compte du risque de défaut et le prix de ce même produit sans risque de défaut (i.e. avec $\tau = \infty$).

Cependant, l'échéancier du CDS n'a en général aucun lien avec l'échéancier d'une obligation. Ici nous avons présenté l'acheteur de CDS comme détenteur d'obligations pour justifier l'intérêt d'un produit comme le CDS. Désormais, les CDS sur les défauts de grandes entreprises ont été standardisés pour être échangés sur un marché. Typiquement, les échéances tombent tous les trois mois et les durées standard sont : $T_n \in \{1\text{an}, 3\text{ans}, 5\text{ans}, 7\text{ans}, 10\text{ans}\}$.

La pratique du marché n'est pas de coter directement le prix du CDS c'est à dire la valeur du flux de paiements. La valeur cotée sur le marché est le taux R du CDS qui rend le contrat équitable pour "A" et "B". C'est le taux pour lequel le prix du CDS est nul et qui fait que la jambe de protection $L_{GD}D(0, \tau)\mathbf{1}_{\tau \leq T_n}$ a la même valeur que la jambe de paiement $\sum_{i=1}^n \alpha_i R D(0, T_i)\mathbf{1}_{\tau > T_i} + R(\tau - T_{\beta(\tau)-1})D(0, \tau)\mathbf{1}_{\tau \leq T_n}$. Ce mécanisme de cotation est plus naturel du point de vue du vendeur de protection "B", car ce dernier se pose la question "à partir de quel taux R est-il raisonnable de vendre une protection?". Si un vendeur estime que le marché fait payer trop cher cette protection, il vendra un CDS sur le marché et achètera un CDS avec un taux inférieur au marché ; et inversement s'il estime le marché trop bas.

Option d'achat sur CDS

Une option d'achat d'un CDS est simplement la possibilité, pour un acheteur de protection, d'entrer à la date T_o dans un CDS d'échéancier $\mathcal{T}$ à un taux $K > 0$ prédéfini. Le payoff d'une telle option est donc :

$$D(0, T_o)\mathbf{1}_{\tau > T_o} \times \begin{cases} \left(L_{GD}D(T_o, \tau)\mathbf{1}_{\tau \leq T_n} - K(T_{\beta(T_o)} - T_o)D(T_o, T_{\beta(T_o)})\mathbf{1}_{\tau > T_{\beta(T_o)}} \right. \\ \left. - \sum_{i=\beta(T_o)+1}^n \alpha_i K D(T_o, T_i)\mathbf{1}_{\tau > T_i} - K(\tau - \max(T_{\beta(\tau)-1}, T_o))D(T_o, \tau)\mathbf{1}_{\tau \leq T_n} \right) \\ \text{ou } 0 \text{ si l'option n'est pas exercée.} \end{cases}$$

Supposons que le CDS qui débute en T_o et d'échéancier $\{T_{\beta(T_o)}, \dots, T_n\}$ soit coté et qu'il ait en T_o la valeur R_{T_o} . Alors, la règle de décision sera la suivante. On exercera l'option si $K < R_{T_o}$. Dans le cas contraire, il sera moins cher d'acheter directement une protection sur le marché.

1.1.2 Les produits sur panier (multi-name)

Dans cette partie nous nous intéresserons aux produits sur plusieurs défauts. Nous considérerons ainsi m entités susceptibles de faire défaut " C_1 ", ..., " C_m " et noterons respectivement par $\tau_1, \dots, \tau_m$ leurs instants de défaut qui sont supposés positifs. Comme nous l'avons expliqué en introduction, la caractéristique des produits sur panier est de faire intervenir la dépendance entre ces défauts et non seulement leurs lois marginales. En effet, l'objectif de ces produits est de permettre d'échanger le risque de contagion des défauts, et de pouvoir ainsi se couvrir de ce risque.

Le first-to-default swap

Il s'agit d'un contrat analogue au CDS mais qui considère comme événement de défaut le premier défaut intervenu dans le panier considéré. On définit ainsi $\tau^{(1)} = \min(\tau_1, \dots, \tau_m)$ le premier instant de défaut. Pour le vendeur de protection, le first-to-default swap consiste donc à recevoir un taux fixe R à chaque échéance. Il s'engage en échange à verser une protection (que nous prenons ici égale à 1) si le premier défaut intervient avant T_n . On considère comme précédemment un échéancier $\mathcal{T} = \{T_1, \dots, T_n\}$ et le payoff s'écrit donc

$$\sum_{i=1}^n \alpha_i R D(0, T_i) \mathbf{1}_{\tau^{(1)} > T_i} + R(\tau^{(1)} - T_{\beta(\tau^{(1)})-1}) D(0, \tau^{(1)}) \mathbf{1}_{\tau^{(1)} \leq T_n} - D(0, \tau^{(1)}) \mathbf{1}_{\tau^{(1)} \leq T_n}.$$

Il existe également des default-swaps sur le k -ième défaut ($2 \leq k \leq m$). Le fonctionnement est identique : le vendeur paie une protection si k entreprises au moins font faillites dans le panier considéré au moment de la k -ième faillite, et en échange reçoit des paiements réguliers à chaque échéance.

Les CDO (Collateralized Debt Obligation)

Partons du point de vue d'un vendeur de protection "A" qui a vendu des CDS sur " C_1 ", ..., " C_m ", et supposons que ces CDS soient sur un même échéancier $\mathcal{T}$ qui expire en T_n . Notons N_i le nombre de protections qu'il a vendu sur l'entreprise i . La somme, non actualisée, qu'il aura eu à payer pour assurer ces protections à l'instant $t \in [0, T_n]$ sera égale à :

$$L(t) = \sum_{j=1}^m N_j L_{GD}^j \mathbf{1}_{\tau_j \leq t}.$$

Ici, L_{GD}^j désigne la protection qu'il assure pour le défaut de l'entreprise " C_j ". Le processus (càdlàg) $L(t)$ est appelé processus de perte (Loss), et la perte maximale que peut avoir à subir "A" est $L_{\max} = \sum_{j=1}^m N_j L_{GD}^j$. Supposons par exemple que "A" soit prêt à subir une perte jusqu'à hauteur de 5% de $L_{\max}$ sur les premiers défauts qui arrivent, mais qu'il ne souhaite pas prendre le risque d'avoir une perte supérieure. Il va pour cela acheter un CDO auprès d'un autre vendeur de protection "B", à qui il va effectuer des paiements réguliers en l'échange d'une protection qui lui permettra de couvrir exactement les CDS dans lesquels il est engagé dès que $L(t) \geq 0.05 L_{\max}$. Le paiement régulier qu'il effectue est comme précédemment proportionnel au temps écoulé depuis la dernière échéance. Mais dans un CDO, il est aussi proportionnel à la perte maximale qui reste à couvrir par "B".

Pour “B”, le flux des paiements s’écrit sur ce CDO de tranche $[5\%, 1]$:

$$\begin{aligned} & \sum_{i=1}^n \alpha_i R(L_{\max} - \max(L(T_i), 0.05L_{\max})) D(0, T_i) \\ & + \sum_{j=1}^m \mathbf{1}_{\tau_j \leq T_n} R(\max(L(\tau_j), 0.05L_{\max}) - \max(L(\tau_j-), 0.05L_{\max})) (\tau_j - T_{\beta(\tau_j)-1}) D(0, \tau_j) \\ & - \sum_{j=1}^m \mathbf{1}_{\tau_j \leq T_n} \mathbf{1}_{L(\tau_j) \geq 0.05L_{\max}} D(0, \tau_j) (L(\tau_j) - \max(L(\tau_j-), 0.05L_{\max})). \end{aligned}$$

Le deuxième terme correspond à la somme sur les entités des paiements de la protection entre le défaut et l’échéance qui précède, et peut être omis en approximation. Supposons désormais qu’en fait, “A” ne croit pas que la perte totale puisse excéder 15% de $L_{\max}$, et qu’il ne souhaite donc pas payer une protection pour être couvert au-delà de cette perte. Dans ce cas là, il contractera avec “B” un CDO de tranche $[5\%, 15\%]$. Le flux de paiements sont analogues à celui présenté ci-dessus à ceci près qu’ils cessent dès que la perte $L(t)$ excède $0.15L_{\max}$. Ici, nous venons de justifier le recours aux CDO lorsque “A” a vendu différents CDS. Dans l’introduction de cette thèse, on avait présenté l’acheteur de CDO comme un détenteur d’obligations. De manière générale, dès que l’on dispose d’un portefeuille auquel on peut associer un processus de perte due aux faillites $L(t)$, on peut définir des tranches de CDO lui correspondant. Lorsque ce portefeuille est uniquement constitué de CDO, on parle de “CDO de CDO” ou de CDO².

Nous venons de décrire comment un CDO est contracté de gré à gré, nous allons maintenant nous attacher à décrire les CDO standardisés qui apparaissent sur le marché. Les échéances sont trimestrielles et les durées sont de 3, 5, 7 et 10 ans. Typiquement, en Europe le panier de référence est l’I-Traxx Europe qui compte $m = 125$ noms. La perte considérée est une perte équipondérée et normalisée afin que la perte maximale soit égale à 1 :

$$L(t) = \frac{1}{m} \sum_{j=1}^m \mathbf{1}_{\tau_j \leq t}.$$

Pour $0 \leq a < b \leq 1$, on définit la fonction

$$\forall x \in [0, 1], \quad H_a^b(x) = (x - a)^+ - (x - b)^+.$$

Elle nous permet de décrire les paiements du CDO de la tranche $[a, b]$:

$$\begin{aligned} & \sum_{i=1}^n \alpha_i R[b - a - H_a^b(L(T_i))] D(0, T_i) + \sum_{j=1}^m \mathbf{1}_{\tau_j \leq T_n} R[H_a^b(L(\tau_j)) - H_a^b(L(\tau_j-))] (\tau_j - T_{\beta(\tau_j)-1}) D(0, \tau_j) \\ & - \sum_{j=1}^m \mathbf{1}_{\tau_j \leq T_n} D(0, \tau_j) [H_a^b(L(\tau_j)) - H_a^b(L(\tau_j-))]. \end{aligned}$$

Encore une fois, la jambe de paiement définie par les deux premiers termes est essentiellement caractérisée par le premier tandis que la jambe de défaut est donnée par le troisième terme. Les tranches cotées sur l'I-Traxx sont $[0, 3\%]$, $[3\%, 6\%]$, $[6\%, 9\%]$, $[9\%, 12\%]$, $[12\%, 22\%]$ et $[22\%, 1]$. La première tranche ($[0, 3\%]$) est appelée *junior* ou *equity*, les intermédiaires sont appelées *mezzanine* et la dernière ($[22\%, 1]$) est appelée *senior*. Le mécanisme de cotation des CDO est analogue à celui des CDS. Ce n'est pas la valeur du flux de paiement qui est coté mais le taux de paiement R qui rend le contrat équitable pour l'acheteur et le vendeur de protection. Nous appelons ici R_a^b ce taux qui, à l'instant 0 est considéré comme équitable par le marché. Nous allons donner une relation d'arbitrage qui lie ces différents taux. Considérons tout d'abord pour simplifier un CDO "fictif" de tranche $[a, b]$, dont les paiements réguliers ne dépendraient plus de ce qui reste à couvrir :

$$\sum_{i=1}^n \alpha_i \tilde{R} D(0, T_i) + \sum_{j=1}^m \mathbf{1}_{\tau_j \leq T_n} \tilde{R} (\tau_j - T_{\beta(\tau_j)-1}) D(0, \tau_j) - \sum_{j=1}^m \mathbf{1}_{\tau_j \leq T_n} D(0, \tau_j) [H_a^b(L(\tau_j)) - H_a^b(L(\tau_j-))].$$

On note $\tilde{R}_a^b$ le taux de ce produit qui est équitable pour les deux parties. Etant donné que pour $0 \leq a < b < c \leq 1$, détenir un CDO de tranche $[a, c]$ est équivalent à détenir un CDO de tranche $[a, b]$ et un autre CDO de tranche $[b, c]$ ($H_a^b(x) + H_b^c(x) = H_a^c(x)$), l'absence d'opportunité d'arbitrage assure que :

$$\tilde{R}_a^b + \tilde{R}_b^c = \tilde{R}_a^c.$$

Pour les taux de CDO R_a^b on a une relation analogue, à savoir que le prix de

$$\begin{aligned} & \sum_{i=1}^n \alpha_i R_a^b [b - a - H_a^b(L(T_i))] D(0, T_i) + \sum_{j=1}^m \mathbf{1}_{\tau_j \leq T_n} R_a^b [H_a^b(L(\tau_j)) - H_a^b(L(\tau_j-))] (\tau_j - T_{\beta(\tau_j)-1}) D(0, \tau_j) \\ & + \sum_{i=1}^n \alpha_i R_b^c [c - b - H_b^c(L(T_i))] D(0, T_i) + \sum_{j=1}^m \mathbf{1}_{\tau_j \leq T_n} R_b^c [H_b^c(L(\tau_j)) - H_b^c(L(\tau_j-))] (\tau_j - T_{\beta(\tau_j)-1}) D(0, \tau_j) \\ & - \sum_{i=1}^n \alpha_i R_a^c [c - a - H_a^c(L(T_i))] D(0, T_i) + \sum_{j=1}^m \mathbf{1}_{\tau_j \leq T_n} R_a^c [H_a^c(L(\tau_j)) - H_a^c(L(\tau_j-))] (\tau_j - T_{\beta(\tau_j)-1}) D(0, \tau_j) \end{aligned}$$

est nul. Contrairement à la relation $\tilde{R}_a^b + \tilde{R}_b^c = \tilde{R}_a^c$, il faut spécifier ici un modèle pour calculer R_a^c à partir de R_a^b et R_b^c .

Une pratique du marché consiste à interpréter les prix des CDO en terme de corrélation implicite entre les défauts. Le principe est analogue à l'interprétation en terme de volatilité implicite des prix des options européennes à travers le modèle de Black-Scholes. Nous présenterons plus tard le mécanisme de cette interprétation appelée *base correlation* après avoir au préalable présenté les fonctions copules et les modèles à intensité.

1.2 Modélisation

Dans cette section, nous allons présenter succinctement la modélisation du marché du risque de crédit. Comme il a été dit en introduction, on distingue deux grandes classes de modèles : les modèles structurels et les modèles à intensité de défaut. Nous aborderons rapidement les premiers, puis nous définirons le cadre général des modèles à intensité de défaut. Cela nous permettra de situer dans cette classification le modèle SSRD, qui est introduit au chapitre suivant, comme un modèle à intensité. La modélisation prenant en compte plusieurs défauts sera abordée brièvement à la section suivante après avoir présenté les fonctions copules.

1.2.1 Les modèles structurels

Les modèles structurels sont les modèles qui cherchent à expliquer l'événement de défaut à partir de variables économiques. Nous allons présenter ici les modèles de Merton [54] et de Black-Cox [12] qui ont été les premiers introduits au milieu des années 1970.

Le modèle de Merton suppose que la valeur V_t d'une entreprise "C" à un instant t suit sous une probabilité martingale équivalente $\mathbb{P}$ (i.e. sous laquelle les actifs actualisés sont des martingales) la dynamique

$$dV_t = V_t((r - k)dt + \sigma dW_t).$$

Ici, $r \geq 0$ représente le taux d'intérêt sans risque, $k \geq 0$ le taux de dépense de l'entreprise (salaires, amortissement...) et $\sigma > 0$ la volatilité de la valeur de l'entreprise. On écrit alors la valeur de "C", V_t , comme la somme de son actif S_t (la valeur de ses actions) et de son passif D_t (dette contractée pour financer son activité). On se place sur un intervalle $[0, T]$, et on suppose que la dette a été contractée auprès d'une seule autre entité "A", et que "C" doit rembourser en T la somme $L > 0$ à "A". Le modèle de Merton considère alors deux cas de figure au moment du remboursement : soit $V_T \geq L$ et alors "C" rembourse sa dette, soit $V_T < L$ et alors "C" fait faillite et "A" récupère de "C" la somme V_T . L'instant de défaut s'écrit ainsi :

$$\tau = T\mathbf{1}_{V_T < L} + \infty\mathbf{1}_{V_T \geq L},$$

et la dette à l'instant T vaut $D_T = \min(V_T, L)$. Sous ce modèle la valeur de la dette à l'instant 0 s'écrit, à l'aide de la formule de Black-Scholes :

$$\begin{aligned} D_0 = \mathbb{E}[e^{-rT} \min(V_T, L)] &= e^{-rT}L - \mathbb{E}[e^{-rT}(L - V_T)^+] \quad (< e^{-rT}L) \\ &= e^{-rT}L \left[1 - \Phi \left(d_1 \left(\frac{Le^{(k-r)T}}{V_0} \right) \right) \right] + V_0 e^{-kT} \Phi \left(d_0 \left(\frac{Le^{(k-r)T}}{V_0} \right) \right), \end{aligned}$$

où $d_0(y) = \frac{1}{\sigma\sqrt{T}} \ln(y) - \frac{1}{2}\sigma\sqrt{T}$, $d_1(y) = d_0(y) + \sigma\sqrt{T}$ et $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$. En notant $\lambda = \frac{1}{T} \ln(L/D_0) - r > 0$, le remboursement, vu à l'instant 0 par "A", vaut $D_0 = Le^{-(r+\lambda)T}$.

En anticipant avec le paragraphe suivant, λ peut ainsi être vu comme l'intensité de défaut moyenne sur $[0, T]$ de l'entreprise "C". Pour finir cette présentation du modèle de Merton, on remarque que la valeur des actions est donnée à l'instant t par $S_t = V_t - D_t$. Ainsi, la valeurs des actions peut être vue comme un call sur la valeur de l'entreprise

$$S_0 = \mathbb{E}[e^{-rT}(V_T - \min(V_T, L))] = \mathbb{E}[e^{-rT}(V_T - L)^+]$$

et ainsi évaluée par la formule de Black-Scholes.

Un inconvénient de ce modèle est que le défaut d'une entreprise ne peut intervenir qu'à une seule date, connue à l'avance, ce qui n'est pas très réaliste. Pour pallier ce problème, Black et Cox [12] ont proposé de définir le défaut de l'entreprise "C" par le premier temps de passage en dessous d'une barrière $H(t)$:

$$\tau = \inf\{t \in [0, T], V_t \leq H(t)\},$$

avec $\inf \emptyset = +\infty$. Ils proposent de prendre $H(T) = L$ et pour $t \in [0, T[$, $H(t) = \tilde{L}e^{-\gamma(T-t)}$. Le paramètre γ est supposé positif, et on suppose également $\tilde{L} < L$ ce qui signifie que l'entreprise "C" a plus de marge de manoeuvre avant le remboursement de la dette.

Depuis ce modèle, plusieurs modèles expliquant le défaut comme le passage en dessous d'une barrière de la valeur (ou de l'actif) de l'entreprise ont été développés. Nous n'entrerons ici pas plus en détail sur les modèles structurels car le modèle SSRD présenté dans cette thèse appartient à l'autre catégorie de modèles : les modèles à intensité de défaut.

1.2.2 Les modèles à intensité de défaut

Dans cette partie nous considérons un espace de probabilité $(\Omega, \mathcal{E}, \mathbb{P})$ qui décrit les aléas observés sur le marché. Nous supposons que cet espace est muni d'une filtration $(\mathcal{F}_t, t \geq 0)$ de $\mathcal{E}$ qui contient schématiquement l'information de tous les actifs "sans risque de défaut". En particulier ici, nous supposerons que $B(t)$, la valeur à l'instant t d'une somme unité placée à l'instant 0 dans un placement sans risque, est adapté à la filtration $(\mathcal{F}_t)$. Notons en passant que le coefficient d'actualisation utilisé pour décrire les produits s'écrit :

$$\forall t < T, D(t, T) = B(t)/B(T).$$

On considère ici un seul instant de défaut qui est supposé être une variable aléatoire positive $\mathcal{E}$ -mesurable :

$$\tau : \Omega \rightarrow [0, +\infty].$$

On définit alors $(\mathcal{H}_t, t \geq 0) = (\sigma(\{\tau \leq u\}, u \leq t), t \geq 0)$, qui est également la filtration engendrée par $(\min(\tau, t), t \geq 0)$. Puis enfin on définit la filtration $(\mathcal{G}_t, t \geq 0)$ comme la filtration engendrée par $(\mathcal{F}_t, t \geq 0)$ et $(\mathcal{H}_t, t \geq 0)$: $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$. Intuitivement, l'information donnée par $\mathcal{G}_t$ est contenue dans celle donnée par le marché puisque la connaissance de $\mathcal{G}_t$

décrit d'une part les cours d'actifs sans risque de défaut jusqu'à t ainsi que la valeur de τ si $\tau \leq t$, et la certitude que $\{\tau > t\}$ sinon. Nous ferons ici l'hypothèse que la probabilité $\mathbb{P}$ sous laquelle nous travaillons est une probabilité risque neutre. Cela signifie que le processus des prix actualisés à l'instant t de tout payoff qui s'écrit comme variable aléatoire $\mathcal{G}_T$ -mesurable (pour un certain $T > 0$) est une $\mathcal{G}_t$ -martingale. Si on suppose L_{GD} déterministe, on remarque que tous les produits sur un seul défaut présentés avant rentrent dans ce cadre puisque le coefficient d'actualisation est mesurable par rapport à la filtration $(\mathcal{G}_t, t \geq 0)$. Ainsi, par exemple, le prix d'un CDS d'échéances $\mathcal{T} = \{T_1, \dots, T_n\}$ qui débute en $T_0 > 0$ aura comme prix en $t \in [0, T_0]$

$$\mathbb{E} \left[D(t, T_0) \mathbf{1}_{\tau \geq T_0} \left(\sum_{i=1}^n \alpha_i R D(T_0, T_i) \mathbf{1}_{\tau > T_i} + R(\tau - T_{\beta(\tau)-1}) D(T_0, \tau) \mathbf{1}_{\tau \leq T_n} - L_{GD} D(T_0, \tau) \mathbf{1}_{\tau \leq T_n} \right) \middle| \mathcal{G}_t \right].$$

Cela nous permet notamment de définir le taux équitable du CDS sur $\{\tau > t\}$,

$$R_t = \frac{\mathbb{E} \left[L_{GD} D(t, \tau) \mathbf{1}_{T_0 \leq \tau \leq T_n} \middle| \mathcal{G}_t \right]}{\mathbb{E} \left[\sum_{i=1}^n \alpha_i D(t, T_i) \mathbf{1}_{\tau > T_i} + (\tau - T_{\beta(\tau)-1}) D(t, \tau) \mathbf{1}_{T_0 \leq \tau \leq T_n} \middle| \mathcal{G}_t \right]}$$

qui apparaît ainsi comme un processus $\mathcal{G}_t$ -mesurable. Pour plus de détails sur l'évaluation risque-neutre dans le cadre du risque de crédit, on se référera au livre de Bielecki et Rutkowski [11].

Maintenant que nous avons bien défini l'espace de probabilité sur lequel nous travaillons, nous pouvons définir ce qu'est un modèle de défaut à intensité.

Définition 1.2.1. *On dira que τ suit un modèle à intensité de défaut s'il existe un processus càdlàg positif $(\lambda_t, t \geq 0)$ adapté à la filtration $(\mathcal{F}_t, t \geq 0)$ et une variable aléatoire ξ $\mathcal{E}$ -mesurable, indépendante de $\mathcal{F}_\infty = \bigvee_{t \geq 0} \mathcal{F}_t$ et de loi exponentielle de paramètre 1 sous $\mathbb{P}$ tels que :*

$$\tau = \inf \left\{ t \geq 0, \int_0^t \lambda_s ds \geq \xi \right\}$$

Le processus $(\lambda_t, t \geq 0)$ est alors appelé intensité de défaut ou credit spread.

On pose alors dans ce cas $\Lambda(t) = \int_0^t \lambda_s ds$ qui est un processus croissant. Remarquons ici que $\{\Lambda(t) = +\infty\} \subset \{\tau \leq t\}$, et on peut ainsi obtenir de l'information sur le défaut à partir de $(\mathcal{F}_t, t \geq 0)$. Pour éviter cela, on ajoute souvent pour définir les modèles à intensité de défaut l'hypothèse suivante :

$$\forall t \in \mathbb{R}, \mathbb{P}(\Lambda(t) < +\infty) = 1 \text{ et } \mathbb{P}(\lim_{t \rightarrow +\infty} \Lambda(t) = +\infty) = 1.$$

Cette hypothèse est satisfaite en général par les modèles à intensité que l'on utilise et en particulier par le modèle SSRD que nous allons présenter. Cependant, nous n'en avons techniquement pas besoin pour la suite de cette partie.

Par définition de τ , on a pour $t > 0$,

$$\{\tau > t\} = \left\{ \int_0^t \lambda_s ds < \xi \right\} = \{\Lambda(t) < \xi\}.$$

Par conséquent, on a pour toute variable aléatoire Y intégrable et $\mathcal{F}_\infty$ -mesurable :

$$\mathbb{E}[Y \mathbf{1}_{\{\tau > t\}}] = \mathbb{E}[Y \mathbf{1}_{\{\Lambda(t) < \xi\}}] = \mathbb{E}[Y \mathbb{E}[\mathbf{1}_{\{\Lambda(t) < \xi\}} | \mathcal{F}_\infty]] = \mathbb{E}[e^{-\Lambda(t)} Y]. \quad (1.1)$$

Appliquons ceci pour calculer le prix en 0 d'un zéro-coupon d'échéance $T > 0$ avec défaut. Puisque $D(0, T)$ est $\mathcal{F}_T$ -mesurable, on a :

$$\mathbb{E}[D(0, T) \mathbf{1}_{\tau > T}] = \mathbb{E}[D(0, T) e^{-\Lambda(T)}].$$

Supposons un instant que l'actif sans risque soit modélisé par un taux d'intérêt court $(r_t, t \geq 0)$ adapté à $(\mathcal{F}_t, t \geq 0)$: $B(t) = \exp(\int_0^t r_s ds)$. Ce prix est alors égal à $\mathbb{E}[\exp(-\int_0^T (r_s + \lambda_s) ds)]$ qui représente également le prix d'un zéro-coupon sans risque de défaut avec un taux d'intérêt court égal à $r_t + \lambda_t$ au lieu de r_t . L'intensité $(\lambda_t, t \geq 0)$ représente bien ainsi un écart entre le prix d'un zéro-coupon sans défaut et avec défaut. C'est pour cette raison qu'on l'appelle également credit spread.

Nous allons désormais présenter un lemme qui joue un rôle clef dans l'évaluation risque-neutre.

Lemme 1.2.2. *On se place dans le cadre présenté ci-dessus où τ suit un modèle à intensité de défaut. Soit Y une variable aléatoire $\mathcal{E}$ -mesurable intégrable. On a pour $t \geq 0$,*

$$\mathbf{1}_{\{\tau > t\}} \mathbb{E}[Y | \mathcal{G}_t] = \mathbf{1}_{\{\tau > t\}} e^{\Lambda(t)} \mathbb{E}[Y \mathbf{1}_{\{\tau > t\}} | \mathcal{F}_t].$$

Preuve. Il suffit de vérifier que l'on a pour tout ensemble $A \in \mathcal{G}_t$,

$$\mathbb{E}[\mathbf{1}_A e^{-\Lambda(t)} \mathbf{1}_{\{\tau > t\}} Y] = \mathbb{E}[\mathbf{1}_A \mathbf{1}_{\{\tau > t\}} \mathbb{E}[Y \mathbf{1}_{\{\tau > t\}} | \mathcal{F}_t]].$$

Il est en fait même suffisant de vérifier cette propriété pour une famille d'ensembles de $\mathcal{G}_t$ qui engendre $\mathcal{G}_t$. Or, par notre construction de $\mathcal{G}_t$, il est clair que $\mathcal{G}_t$ est la tribu engendrée par $\{B \cap \{\tau > u\}, B \in \mathcal{F}_t, u \leq t\}$. Soit $B \in \mathcal{F}_t$ et $u \leq t$, on a :

$$\begin{aligned} \mathbb{E}[\mathbf{1}_{B \cap \{\tau > u\}} e^{-\Lambda(t)} \mathbf{1}_{\{\tau > t\}} Y] &= \mathbb{E}[\mathbf{1}_B e^{-\Lambda(t)} \mathbf{1}_{\{\tau > t\}} Y] \text{ car } u \leq t \\ &= \mathbb{E}[\mathbf{1}_B e^{-\Lambda(t)} \mathbb{E}[\mathbf{1}_{\{\tau > t\}} Y | \mathcal{F}_t]] \\ &= \mathbb{E}[\mathbf{1}_B \mathbf{1}_{\{\tau > t\}} \mathbb{E}[\mathbf{1}_{\{\tau > t\}} Y | \mathcal{F}_t]] \text{ en utilisant (1.1)} \\ &= \mathbb{E}[\mathbf{1}_{B \cap \{\tau > u\}} \mathbf{1}_{\{\tau > t\}} \mathbb{E}[\mathbf{1}_{\{\tau > t\}} Y | \mathcal{F}_t]] \text{ car } u \leq t. \quad \square \end{aligned}$$

Ce lemme nous donne en particulier pour $t < s$, $\mathbb{P}(\tau > s | \mathcal{G}_t) = \mathbf{1}_{\{\tau > t\}} \mathbb{E}[\exp(-\int_t^s \lambda_u du) | \mathcal{F}_t]$. Ainsi, en faisant l'approximation $1 - \exp(-\int_t^s \lambda_u du) \underset{s \rightarrow t^+}{\approx} \lambda_t(s - t)$, on a :

$$\mathbb{P}(\tau \in (t, s] | \mathcal{G}_t) \underset{s \rightarrow t^+}{\approx} \lambda_t(s - t)$$

ce qui permet de voir l'intensité de défaut comme le taux instantané de défaut.

Pour modéliser plusieurs instants de défaut suivant chacun un modèle à intensité, il est utile d'introduire la notion de copule.

1.3 Copules et dépendance

1.3.1 Définition et propriétés

Définition 1.3.1. On appelle copule de dimension $m \geq 2$ la fonction de répartition

$$(u_1, \dots, u_m) \in [0, 1]^m \mapsto \mathbb{P}(U_1 \leq u_1, \dots, U_m \leq u_m)$$

d'un vecteur aléatoire $(U_1, \dots, U_m)$ dont les lois marginales sont uniformes sur $[0, 1]$ (i.e. $\forall i \in \{1, \dots, m\}, \forall u_i \in [0, 1], \mathbb{P}(U_i \leq u_i) = u_i$). Nous noterons $\mathfrak{C}^m$ l'ensemble des copules de dimension m .

Proposition 1.3.2. Une fonction $C : [0, 1]^m \rightarrow [0, 1]$ est une copule (de dimension m) si et seulement si elle satisfait les trois propriétés suivantes :

1. $\forall k \in \{1, \dots, m\}, C(u_1, \dots, u_{k-1}, 0, u_{k+1}, \dots, u_m) = 0$
2. $\forall k \in \{1, \dots, m\}, C(1, \dots, 1, u_k, 1, \dots, 1) = u_k$
3. $\partial_{u_1} \partial_{u_2} \dots \partial_{u_m} C$ est une mesure positive au sens des distributions.

Preuve. Il est immédiat de voir que ces trois conditions sont nécessaires : montrons qu'elles sont suffisantes. La mesure $\mu = \partial_{u_1} \dots \partial_{u_m} C$ est positive et de masse 1 sur $[0, 1]^m$ car $\int_0^1 \dots \int_0^1 \mu(du_1, \dots, du_m) = C(1, \dots, 1) = 1$ grâce aux propriétés 1 et 2. Soit $U = (U_1, \dots, U_m)$ une variable aléatoire de loi μ sur $[0, 1]^m$. Les lois marginales de U sont uniformément réparties sur $[0, 1]$ grâce à la propriété 2, et C est par construction la fonction de répartition de U . $\square$

Remarque 1.3.3. La propriété 3 est souvent remplacée par la propriété de m -croissance qui signifie concrètement que pour tout hypercube $\mathcal{C} = [a_1, b_1] \times \dots \times [a_m, b_m] \subset [0, 1]^m$,

$$\mathbb{P}[(U_1, \dots, U_m) \in \mathcal{C}] \geq 0.$$

Pour $m = 2$ la 2-croissance s'écrit : $C(b_1, b_2) - C(a_1, b_2) - C(b_1, a_2) + C(a_1, a_2) \geq 0$.

Une conséquence immédiate de la Proposition 1.3.2 est que l'ensemble des copules $\mathfrak{C}^m$ est convexe. Donnons une preuve plus probabiliste de ce résultat qui a également un intérêt pour simuler un vecteur de variable aléatoire dont la fonction de répartition est obtenue par combinaison linéaire convexe de deux copules. Soient $(U_1, \dots, U_m)$ et $(U'_1, \dots, U'_m)$ deux vecteurs aléatoires dont les marginales sont uniformes sur $[0, 1]$ et de fonctions de répartition

respectives C et C' . Alors, si B est une variable aléatoire de Bernoulli de paramètre $p \in [0, 1]$ indépendante de ces deux vecteurs, $(BU_1 + (1 - B)U'_1, \dots, BU_m + (1 - B)U'_m)$ est un vecteur aléatoire dont les marginales sont uniformes sur $[0, 1]$ et de fonction de répartition $pC + (1 - p)C'$.

Considérons désormais m variables aléatoires réelles $X_1, \dots, X_m$ de fonctions de répartition respectives $F_1, \dots, F_m$ que l'on suppose continues. Alors, $U_1 = F_1(X_1), \dots, U_m = F_m(X_m)$ sont des variables aléatoires uniformes sur $[0, 1]$ et on note C la fonction de répartition de $(U_1, \dots, U_m)$:

$$\forall u_1, \dots, u_m \in [0, 1], C(u_1, \dots, u_m) = \mathbb{P}(U_1 \leq u_1, \dots, U_m \leq u_m)$$

Alors, pour $x_1, \dots, x_m \in \mathbb{R}$, la fonction de répartition du vecteur $(X_1, \dots, X_m)$ vaut :

$$\mathbb{P}(X_1 \leq x_1, \dots, X_m \leq x_m) = \mathbb{P}(U_1 \leq F_1(x_1), \dots, U_m \leq F_m(x_m)) = C(F_1(x_1), \dots, F_m(x_m))$$

On a ainsi établi l'existence d'une fonction $C : [0, 1]^m \rightarrow [0, 1]$ telle que :

$$\mathbb{P}(X_1 \leq x_1, \dots, X_m \leq x_m) = C(F_1(x_1), \dots, F_m(x_m))$$

appelée copule de $(X_1, \dots, X_m)$. Il est facile de voir que cette fonction C est unique puisque les fonctions de répartition sont surjectives sur $[0, 1]$.

Théorème 1.3.4 (Sklar). *Soit $X = (X_1, \dots, X_m) : \Omega \rightarrow \mathbb{R}^m$ un vecteur aléatoire dont les lois marginales $X_1, \dots, X_m$ ont des fonctions de répartition $F_1, \dots, F_m$ continues. Alors il existe une unique fonction $C : [0, 1]^m \rightarrow [0, 1]$ appelée copule du vecteur X telle que :*

$$\forall x_1, \dots, x_m \in \mathbb{R}, \quad \mathbb{P}(X_1 \leq x_1, \dots, X_m \leq x_m) = C(F_1(x_1), \dots, F_m(x_m)). \quad (1.2)$$

Cette fonction C est la fonction de répartition du vecteur $(F_1(X_1), \dots, F_m(X_m))$.

En réalité, le théorème de Sklar est plus général que celui énoncé ci-dessus car il donne également l'existence d'une fonction C satisfaisant la propriété (1.2) lorsque les variables aléatoires ont des fonctions de répartition discontinues. D'un point de vue intuitif, ce théorème nous dit que la loi de $(X_1, \dots, X_m)$ est entièrement caractérisée par les lois marginales F_i et la copule C , qui décrit elle toutes les interactions entre les variables X_i . De façon analogue et toujours avec $U_i = F_i(X_i)$, le vecteur $(1 - U_1, \dots, 1 - U_m)$ est également un vecteur de variables aléatoires uniformes sur $[0, 1]$. On note $\check{C}$ sa fonction de répartition, et on a alors

$$\begin{aligned} \mathbb{P}(X_1 \geq x_1, \dots, X_m \geq x_m) &= \mathbb{P}(1 - U_1 \leq 1 - F_1(x_1), \dots, 1 - U_m \leq 1 - F_m(x_m)) \\ &= \check{C}(\mathbb{P}(X_1 \geq x_1), \dots, \mathbb{P}(X_m \geq x_m)) \end{aligned}$$

puisque les fonctions $1 - F_i(x)$ sont continues décroissantes. De la même manière $\check{C} : [0, 1]^m \rightarrow [0, 1]$ est l'unique fonction qui satisfait cette relation et est appelée *copule de survie* du vecteur X .

Nous allons maintenant énoncer quelques propriétés générales sur les copules.

Proposition 1.3.5. Soient $X_1, \dots, X_m$ m variables aléatoires réelles de fonctions de répartition respectives $F_1, \dots, F_m$ continues, et $\Phi_1, \dots, \Phi_m$ m fonctions continues strictement croissantes et déterministes. On pose pour $i \in \{1, \dots, m\}$,

$$Y_i = \Phi_i(X_i).$$

Alors les fonctions de répartition des variables aléatoires $Y_1, \dots, Y_m$ sont continues, et le vecteur $Y = (Y_1, \dots, Y_m)$ a la même copule que $X = (X_1, \dots, X_m)$.

Preuve. Il suffit de voir que pour $y \in \mathbb{R}$, $\mathbb{P}(Y_i \leq y) = \mathbb{P}(X_i \leq \Phi_i^{-1}(y)) = F_i(\Phi_i^{-1}(y))$ pour obtenir la continuité des fonctions de répartition. Ensuite, on remarque en utilisant le théorème de Sklar que la copule de X est la fonction de répartition du vecteur $(F_1(X_1), \dots, F_m(X_m)) = (F_1(\Phi_1^{-1}(Y_1)), \dots, F_m(\Phi_m^{-1}(Y_m)))$. $\square$

Avant d'avancer plus loin, nous allons donner quelques copules remarquables. Soient $U_1, \dots, U_m$ m variables aléatoires uniformes sur $[0, 1]$. Lorsque ces variables sont indépendantes, on note

$$C^\perp(u_1, \dots, u_m) = \mathbb{P}(U_1 \leq u_1, \dots, U_m \leq u_m) = u_1 u_2 \dots u_m$$

la fonction de répartition de $(U_1, \dots, U_m)$. Lorsque au contraire $U_1 = U_2 = \dots = U_m$ on pose

$$\begin{aligned} C^+(u_1, \dots, u_m) &= \mathbb{P}(U_1 \leq u_1, \dots, U_1 \leq u_m) \\ &= \min(u_1, \dots, u_m) \end{aligned}$$

et on parle de dépendance positive. Enfin, on définit

$$C^-(u_1, \dots, u_m) = [u_1 + \dots + u_m - (m-1)]^+.$$

Il s'agit d'une copule seulement lorsque $m = 2$: c'est la fonction de répartition du vecteur $(U_1, 1 - U_1)$. Pour $m \geq 3$, on peut vérifier qu'au sens des distributions, $\partial_{u_1} \dots \partial_{u_m} C$ n'est pas une mesure positive mais seulement la dérivée $(m-2)^{\text{ème}}$ d'une mesure positive. La remarque suivante nous montre cependant que C^- peut décrire la dépendance d'un vecteur aléatoire discret au sens de (1.2).

Remarque 1.3.6. Soit U une variable aléatoire uniforme sur $[0, 1]$ et $m \geq 2$. On pose $X_i = \mathbf{1}_{[(i-1)/m, i/m]}(U)$ pour $i \in \{1, \dots, m\}$. On peut vérifier que l'on a comme en (1.2) :

$$\forall x_1, \dots, x_m \in \mathbb{R}, \mathbb{P}(X_1 \leq x_1, \dots, X_m \leq x_m) = C^-(\mathbb{P}(X_1 \leq x_1), \dots, \mathbb{P}(X_m \leq x_m)).$$

Proposition 1.3.7 (bornes de Fréchet-Hoeffding). Pour toute fonction copule C de dimension $m \geq 2$, on a :

$$\forall u_1, \dots, u_m \in [0, 1], C^-(u_1, \dots, u_m) \leq C(u_1, \dots, u_m) \leq C^+(u_1, \dots, u_m).$$

Preuve. Soit $U = (U_1, \dots, U_m)$ un vecteur aléatoire dont les marginales suivent des lois uniformes sur $[0, 1]$. Pour $(u_1, \dots, u_m) \in [0, 1]^m$ et $i \in \{1, \dots, m\}$, on a évidemment $\mathbb{P}(U_1 \leq u_1, \dots, U_m \leq u_m) \leq \mathbb{P}(U_i \leq u_i) = u_i$ ce qui donne la borne supérieure. D'autre part, $\mathbb{P}(U_1 \leq u_1, \dots, U_m \leq u_m) = \mathbb{P}(U_1 \leq u_1, \dots, U_{m-1} \leq u_{m-1}) - \mathbb{P}(U_1 \leq u_1, \dots, U_m > u_m) \geq \mathbb{P}(U_1 \leq u_1, \dots, U_{m-1} \leq u_{m-1}) - \mathbb{P}(U_m > u_m) = \mathbb{P}(U_1 \leq u_1, \dots, U_{m-1} \leq u_{m-1}) + u_m - 1$. Par récurrence, on obtient que $\mathbb{P}(U_1 \leq u_1, \dots, U_m \leq u_m) \geq u_1 + \dots + u_m - (m-1)$ ce qui donne la borne inférieure. $\square$

1.3.2 Mesures de dépendance

Commençons ce paragraphe par une remarque et considérons U une variable aléatoire uniforme sur $[0, 1]$. Grâce à la propriété d'invariance donnée par la Proposition 1.3.5, Les vecteurs (U, U) et (U^2, U) ont la même copule $C^+(u_1, u_2) = \min(u_1, u_2)$. Néanmoins, la corrélation entre U et U est égale à 1, alors que celle entre U^2 et U vaut $\sqrt{\frac{45}{48}} \neq 1$ après calculs. Par conséquent, deux vecteurs aléatoires peuvent avoir une même structure de dépendance (i.e. une même copule C), alors que leurs coordonnées ont des corrélations différentes. La corrélation dépend ainsi à la fois des marginales et de la copule. Cela motive l'introduction de *mesures de dépendance* d'un vecteur aléatoire (X, Y) , c'est à dire de mesures qui permettent de quantifier l'influence de la copule de (X, Y) sur les interactions entre deux variables aléatoires réelles X et Y sans tenir compte des lois marginales.

Nous donnons ici la définition de mesure de concordance introduite par Scarsini [59].

Définition 1.3.8. Une application $\kappa : \mathfrak{C}^2 \rightarrow [-1, 1]$ est appelée mesure de concordance si elle satisfait les propriétés suivantes :

1. $\kappa(C^-) = -1, \kappa(C^\perp) = 0, \kappa(C^+) = 1$.
2. symétrie : $\forall C \in \mathfrak{C}^2, \kappa(C) = \kappa(C')$ où $C'(u, v) = C(v, u)$.
(Remarque : si C est la copule de (X, Y) , C' est la copule de (Y, X) .)
3. $\forall C \in \mathfrak{C}^2, \kappa(C) = -\kappa(C')$ où $C'(u, v) = v - C(1 - u, v)$.
(Remarque : si C est la copule de (X, Y) , C' est la copule de $(-X, Y)$.)
4. croissance : $\forall C, C' \in \mathfrak{C}^2, (\forall u, v \in [0, 1], C(u, v) \leq C'(u, v) \implies \kappa(C) \leq \kappa(C'))$.
5. continuité : Soit $C \in \mathfrak{C}^2$. Si C_n est une suite de copules qui converge simplement vers C (i.e. $\forall u, v \in [0, 1], C_n(u, v) \xrightarrow{n \rightarrow +\infty} C(u, v)$), alors $\kappa(C_n) \xrightarrow{n \rightarrow +\infty} \kappa(C)$.

On peut montrer (Nelsen [57]) que les deux mesures que nous allons présenter, le rho de Spearman et le tau de Kendall, sont des mesures de concordance au sens donné ci-dessus.

Définition 1.3.9 (tau de Kendall). Soient X et Y deux variables aléatoires réelles dont les fonctions de répartition sont continues, et C la copule du vecteur (X, Y) . Alors, le tau

de Kendall est défini par

$$\tau(C) = 4 \int_0^1 \int_0^1 C(u, v) \partial_u \partial_v C(du, dv) - 1.$$

En outre, si (X', Y') est une copie indépendante de (X, Y) , on a

$$\tau(C) = \mathbb{P}((X - X')(Y - Y') > 0) - \mathbb{P}((X - X')(Y - Y') < 0).$$

Démontrons cette dernière égalité. On a :

$$\begin{aligned} \mathbb{P}((X - X')(Y - Y') > 0) &= \mathbb{P}((X - X')(Y - Y') < 0) \\ &= 2\mathbb{P}((X - X')(Y - Y') > 0) - 1 \\ &= 2(\mathbb{P}(X > X', Y > Y') + \mathbb{P}(X' > X, Y' > Y)) - 1 \\ &= 4\mathbb{P}(X' > X, Y' > Y) - 1 \text{ car } (X, Y) \stackrel{\text{loi}}{=} (X', Y') \\ &= 4\mathbb{E}(C(F_X(X'), F_Y(Y'))) - 1. \end{aligned}$$

en notant F_X (resp. F_Y) la fonction de répartition de X (resp. Y).

Définition 1.3.10 (rho de Spearman). *Soient X et Y deux variables aléatoires réelles dont les fonctions de répartition sont continues, et C la copule du vecteur (X, Y) . Alors, le rho de Spearman est défini par*

$$\rho_S(C) = 12 \int_0^1 \int_0^1 uv \partial_u \partial_v C(du, dv) - 3.$$

En outre, si (X', Y') et (X'', Y'') sont deux copies indépendantes de (X, Y) , on a

$$\rho_S(C) = 3[\mathbb{P}((X - X')(Y - Y'') > 0) - \mathbb{P}((X - X')(Y - Y'') < 0)].$$

Donnons la preuve de cette dernière égalité. Comme précédemment pour le tau de Kendall, on a :

$$\begin{aligned} \mathbb{P}((X - X')(Y - Y'') > 0) &= \mathbb{P}((X - X')(Y - Y'') < 0) \\ &= 2(\mathbb{P}(X > X', Y > Y'') + \mathbb{P}(X < X', Y < Y'')) - 1 \\ &= 2(\mathbb{E}[(1 - F_X(X))(1 - F_Y(Y))] + \mathbb{E}[F_X(X)F_Y(Y)]) - 1 \\ &= 4\mathbb{E}[F_X(X)F_Y(Y)] - 1. \end{aligned}$$

Pour terminer ce bref paragraphe sur les mesures de dépendance, nous allons présenter des mesures qui permettent de quantifier la dépendance des valeurs extrêmes prises par deux variables aléatoires X et Y . On considère comme précédemment deux variables aléatoires X et Y dont les fonctions de répartition sont continues et on note C la copule de

(X, Y) et $\check{C}$ sa copule de survie. On montre aisément que $\check{C}(u, v) = C(1-u, 1-v) + u + v - 1$. En posant $U = F_X(X)$ et $V = F_Y(Y)$, on définit l'*upper-tail dependence* par :

$$\lim_{u \rightarrow 0^+} \frac{\mathbb{P}(U > 1-u, V > 1-u)}{u} = \lim_{u \rightarrow 0^+} \frac{\check{C}(u, u)}{u},$$

lorsque cette limite existe. Comme $\mathbb{P}(U > 1-u, V > 1-u) \leq \mathbb{P}(U > 1-u) = u$, cette limite appartient à $[0, 1]$. Cela mesure la dépendance asymptotique entre les grandes valeurs de X et celles de Y . De même, on définit la *lower-tail dependence* par

$$\lim_{u \rightarrow 0^+} \frac{\mathbb{P}(U < u, V < u)}{u} = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}$$

lorsque cette limite existe, et cette limite appartient alors nécessairement à $[0, 1]$. La lower-tail dependence mesure la dépendance asymptotique entre les petites valeurs de X et celles de Y . Il est facile de voir que pour la copule C^+ l'*upper-tail* et la *lower-tail* dependence sont égales à 1 alors qu'elles sont égales à 0 pour les copules $C^\perp$ et C^- . Concernant le risque de crédit, ce type de dépendance peut être intéressant pour tenir compte du risque de contagion. En effet, un défaut peut être vu comme un événement extrême, et ce type de dépendance permet de quantifier le risque que cet événement engendre un autre événement extrême, c'est à dire une autre faillite.

1.3.3 Quelques familles de copules

Nous concluons cette partie introductive sur les copules en présentant quelques familles fréquemment utilisées. Pour une présentation plus exhaustive de ces familles, on se référera au livre de Nelsen [57] qui présente également des méthodes pour construire des copules. Rappelons ici que l'ensemble des copules étant convexe, on peut considérer des combinaisons linéaires convexes de ces familles afin d'obtenir une famille encore plus large. La famille de copules "périodiques" présentée dans cette thèse vient s'ajouter aux nombreuses familles existantes et présente au moins deux caractéristiques intéressantes. La première est qu'elles s'étendent facilement au cas multivarié $m \geq 2$. La seconde est qu'elle permet de décrire des interactions deux à deux non symétriques. Ce point sera discuté plus en détail lors de la présentation de ces copules au chapitre 3.

Les copules gaussiennes

On appelle copule gaussienne la copule d'un vecteur gaussien $X = (X_1, \dots, X_m)$. Lorsque $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ avec $\mu_i \in \mathbb{R}$ et $\sigma_i > 0$, on voit en utilisant la proposition 1.3.5 que X a la même copule que $(\frac{X_1 - \mu_1}{\sigma_1}, \dots, \frac{X_m - \mu_m}{\sigma_m})$. Ainsi, quitte à effectuer cette transformation, on peut supposer que X a ses marginales qui suivent la loi gaussienne centrée réduite. Sa loi est ainsi entièrement caractérisée par sa matrice de covariance Γ qui est symétrique, positive

et dont les coefficients diagonaux sont égaux à 1. Si on note Φ la fonction de répartition de la loi normale centrée réduite et Φ_Γ la fonction de répartition de X , la copule gaussienne s'écrit :

$$\forall u_1, \dots, u_m \in [0, 1], C_\Gamma(u_1, \dots, u_m) = \Phi_\Gamma(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_m)).$$

Cette formule n'est guère exploitable pour mener des calculs analytiques. En revanche, il est très facile de simuler des variables aléatoires uniformes distribuées suivant des copules gaussiennes puisqu'il suffit de simuler le vecteur gaussien X en utilisant par exemple une décomposition de Cholesky de la matrice Γ , et ensuite calculer $U = (\Phi(X_1), \dots, \Phi(X_n))$.

Les copules gaussiennes à un facteur constituent un cas particulier important des copules gaussiennes. Il s'agit des copules gaussiennes dont la matrice de covariance s'écrit $\Gamma = (1 - \rho)I + \rho J$ où I est la matrice identité et J la matrice carrée $m \times m$ dont tous les coefficients sont égaux à 1. En décomposant J dans une base propre, on voit que la condition de positivité de la matrice Γ équivaut à

$$\rho \in [-1/(m-1), 1].$$

Lorsque $\rho \geq 0$, la copule gaussienne de facteur ρ est la copule du vecteur

$$(\sqrt{\rho}G + \sqrt{1-\rho}G_1, \dots, \sqrt{\rho}G + \sqrt{1-\rho}G_m)$$

où $G, G_1, \dots, G_m$ sont $m+1$ gaussiennes centrées réduites indépendantes.

La famille des copules gaussiennes permet d'atteindre les copules $C^\perp$, C^+ en dimension $m \geq 2$ et C^- en dimension 2. En effet, $C^\perp$ est atteinte si $\Gamma = I$, C^+ est atteinte lorsque $\Gamma = J$, et C^- est atteinte en dimension 2 avec $\Gamma = 2I - J$.

Les t -copules

On considère un vecteur gaussien $X = (X_1, \dots, X_m)$ dont les marginales suivent la gaussienne centrée réduite de matrice de covariance Γ . Pour un entier n strictement positif, on se donne une variable aléatoire $\Theta \stackrel{\text{loi}}{=} \sqrt{n/\chi^2(n)}$ indépendante de X . On appelle alors t -copule de paramètre (n, Γ) la copule du vecteur ΘX , dont les marginales suivent une loi de Student de paramètre n . Cette famille peut être vue comme une extension de la famille des copules gaussiennes qui correspond au cas limite $n = +\infty$. De la même manière $C^\perp$, C^+ et C^- en dimension 2 appartiennent à cette famille. Contrairement aux copules gaussiennes, ces copules permettent d'avoir des mesures de tail-dependence non nulles.

Les copules archimédiennes

Nous allons commencer par présenter le cas des copules archimédiennes bivariées. On considère une fonction $\varphi : [0, 1] \rightarrow [0, +\infty]$ que l'on suppose convexe, strictement

décroissante et telle que $\varphi(1) = 0$. On définit alors la fonction

$$\forall u_1, u_2 \in [0, 1], C(u_1, u_2) = \mathbf{1}_{\{\varphi(u_1) + \varphi(u_2) \leq \varphi(0)\}} \varphi^{-1}(\varphi(u_1) + \varphi(u_2))$$

qui est bien une copule (cf. Nelsen [57]). Voici quelques familles de fonction φ qui satisfont ces conditions, ainsi que les copules en dimension 2 qu'elles induisent.

- Les copules de Gumbel :

$$\varphi(x) = (-\ln x)^\alpha, \alpha \in [1, +\infty[$$

$$C(u_1, u_2) = \exp(-((-\ln u_1)^\alpha + (-\ln u_2)^\alpha)^{\frac{1}{\alpha}}).$$

- Les copules de Clayton :

$$\varphi(x) = \alpha(x^{-\alpha} - 1), \alpha \in [-1, +\infty[\setminus \{0\}$$

$$C(u_1, u_2) = (\max(u_1^{-\alpha} + u_2^{-\alpha} - 1, 0))^{-\frac{1}{\alpha}}.$$

- Les copules de Frank :

$$\varphi(x) = -\ln \frac{\exp(-\alpha x) - 1}{\exp(-\alpha) - 1}, \alpha \in]-\infty, +\infty[$$

$$C(u_1, u_2) = -\frac{1}{\alpha} \ln(1 + \frac{(\exp(-\alpha u_1) - 1)(\exp(-\alpha u_2) - 1)}{\exp(-\alpha) - 1}).$$

- Les copules de Joe :

$$\varphi(x) = -\ln(1 - (1 - x)^\alpha), \alpha \in [1, +\infty[$$

$$C(u_1, u_2) = 1 - ((1 - u_1)^\alpha + (1 - u_2)^\alpha - (1 - u_1)^\alpha(1 - u_2)^\alpha)^{\frac{1}{\alpha}}.$$

Pour d'autres familles archimédiennes, on pourra se référer à Nelsen [57]. Chacune de ces familles s'approche des copules remarquables pour certains paramètres (voir le tableau 1.1 ci dessous).

	C^-	$C^\perp$	C^+
Frank	$-\infty$	0	$+\infty$
Cook-Johnson	-1	0	$+\infty$
Gumbel	/	1	$+\infty$
Joe	/	1	$+\infty$

TAB. 1.1 – Valeurs du paramètre α pour lesquelles les copules remarquables sont atteintes

Dans le cadre multivarié $m \geq 2$, on peut étendre la construction des copules archimédiennes. On considère comme précédemment une fonction $\varphi : [0, 1] \rightarrow [0, +\infty]$ strictement décroissante et telle que $\varphi(1) = 0$. On suppose désormais que $\varphi(0) = +\infty$, et de

plus on suppose que φ est de classe C^∞ et est *complètement monotone* sur $[0, 1]$, c'est à dire :

$$\forall k \in \mathbb{N}, \forall x \in [0, 1], \quad (-1)^k \varphi^{(k)}(x) \geq 0.$$

On définit alors :

$$C(u_1, \dots, u_m) = \varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_m)).$$

La fonction C est bien une copule (cf. Nelsen [57]). On peut vérifier que les fonctions φ qui définissent les copules de Gumbel sont complètement monotones, et celles qui définissent les copules de Clayton et Frank sont complètement monotones pour $\alpha > 0$.

1.3.4 Modèles à intensité avec plusieurs défauts

Nous reprenons ici le cadre utilisé pour présenter les modèles à intensité : $(\Omega, \mathcal{E}, \mathbb{P})$ désigne l'espace de probabilité dans lequel sont décrits les actifs et $(\mathcal{F}_t, t \geq 0)$ est la filtration engendrée par les actifs sans risque de défaut. Désormais, on considère m instants de défaut notés $\tau_1, \dots, \tau_m$ qui sont des variables aléatoires $\mathcal{E}$ -mesurables à valeurs dans $[0, +\infty]$. On définit, pour $j \in \{1, \dots, m\}$, $(\mathcal{H}_t^j, t \geq 0)$ par $\mathcal{H}_t^j = \sigma(\{\tau_j \leq u\}, u \leq t)$, la filtration qui représente l'information sur $\tau_j \wedge t$. Enfin, comme précédemment, on définit la filtration $(\mathcal{G}_t, t \geq 0)$ par $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t^1 \vee \dots \vee \mathcal{H}_t^m$, et on suppose que $\mathbb{P}$ est la probabilité risque-neutre, c'est à dire que sous $\mathbb{P}$, le prix d'un produit financier dont le payoff est $\mathcal{G}_T$ -mesurable pour un certain $T > 0$ est une $\mathcal{G}_t$ -martingale.

Définition 1.3.11. *On dira que $\tau_1, \dots, \tau_m$ suivent conjointement un modèle à intensité de défaut s'il existe m processus càdlàg positifs $(\lambda_t^j, t \geq 0)$, $j \in \{1, \dots, m\}$, adaptés à la filtration $(\mathcal{F}_t, t \geq 0)$ et m variables aléatoires $\xi^1, \dots, \xi^m$ $\mathcal{E}$ -mesurables, indépendantes de $\mathcal{F}_\infty = \bigvee_{t \geq 0} \mathcal{F}_t$ et de loi exponentielle de paramètre 1 sous $\mathbb{P}$ tels que :*

$$\forall j \in \{1, \dots, m\}, \quad \tau^j = \inf\{t \geq 0, \int_0^t \lambda_s^j ds \geq \xi^j\}.$$

Avec cette définition, les intensités de défaut peuvent être dépendantes entre elles, mais également les seuils $\xi^1, \dots, \xi^m$. Etant donné que les seuils suivent la même loi sous $\mathbb{P}$, la loi du vecteur $(\xi^1, \dots, \xi^m)$ est entièrement caractérisée par sa copule C (cf. relation (1.2)). On note $\Lambda^j(t) = \int_0^t \lambda_s^j ds$. Sous ce modèle la loi jointe des défauts est donnée pour par :

$$\begin{aligned} \forall t_1, \dots, t_m \geq 0, \quad \mathbb{P}(\tau_1 \leq t_1, \dots, \tau_m \leq t_m) &= \mathbb{E}[\mathbb{E}[\mathbf{1}_{\xi^1 \leq \Lambda^1(t_1), \dots, \xi^m \leq \Lambda^m(t_m)} | \mathcal{F}_\infty]] \\ &= \mathbb{E}[C(1 - e^{-\Lambda^1(t_1)}, \dots, 1 - e^{-\Lambda^m(t_m)})]. \end{aligned}$$

De manière plus générale, le fait de connaître la loi jointe des défauts conditionnellement à $\mathcal{F}_\infty$ permet d'exprimer les prix des dérivés de crédit sur panier sous la forme de l'espérance d'une variable $\mathcal{F}_\infty$ -mesurable. Si on dispose d'une expression analytique de la copule, cela

permet d'éviter la simulation des instants de défaut lorsqu'on calcule cette espérance par méthode de Monte-Carlo.

Le grand intérêt de ce modèle est qu'il s'agit d'une extension du modèle à intensité de défaut avec un seul défaut présenté à la section précédente. Cela permet ainsi de réutiliser facilement la calibration de chaque intensité faite à partir des prix du marché des dérivés de crédit sur un seul nom comme les prix de CDS. La calibration se fait ainsi en deux temps comme suit. Tout d'abord, on se donne un modèle pour chaque intensité de défaut que l'on calibre en utilisant les prix de CDS sur chaque nom. Ensuite, on utilise les prix des dérivés sur paniers comme les CDO pour estimer la copule C .

La grande difficulté de cette calibration réside dans l'estimation de la copule C . La plupart du temps, on postule que la copule appartient à une certaine famille paramétrique, et on cherche numériquement les paramètres qui donnent les mêmes prix théoriques que les prix observés sur le marché. Dans cette optique, nous allons décrire à la section suivante comment le marché interprète les prix des CDO en terme du coefficient, plus intuitif, de corrélation ρ de la famille des copules gaussiennes à un facteur. Lorsque l'on souhaite ensuite utiliser la copule ainsi estimée pour évaluer des prix ou des couvertures, la question de la qualité et de la stabilité de l'estimation se pose. Nous ne nous pencherons pas sur cette question délicate, mais nous dirons simplement ici que la difficulté d'estimer la copule motive le développement de modèles proprement spécifiques aux produits sur paniers, comme par exemple le modèle "loss transition rates" introduit par Schönbucher [61].

1.3.5 Interprétation du prix d'un CDO en terme de "base correlation"

On considère un panier de m entreprises " C_1 ", ..., " C_m ", sur lequel le marché cote des tranches de CDO. Supposons que pour chaque entreprise " C_j ", on soit capable de se donner un modèle à intensité de défaut $(\lambda_t^j, t \geq 0)$ qui calibre les CDS sur " C_j ". C'est à dire que les prix théoriques des CDS donnés par ce modèle sont "proches" ou même égaux (c'est le cas d'un modèle à intensité déterministe ou du modèle SSRD) à ceux observés sur le marché. On postule que la dépendance des seuils $\xi_1, \dots, \xi_m$ est donnée par une copule gaussienne à un facteur $\rho \in [-1/(m-1), 1]$. Comme la santé économique des entreprises du panier considéré est liée au mouvement de fond de l'économie, on suppose le plus souvent $\rho \geq 0$. D'un point de vue heuristique, cela revient à dire que chaque seuil s'explique par une part $(\sqrt{1-\rho})$ propre à l'entreprise, et une part $(\sqrt{\rho})$ qui reflète l'influence de la santé générale de l'économie. On cherche alors la corrélation ρ_a^b telle que le prix théorique avec ce modèle du CDO de tranche $[a, b]$ soit égal au prix R_a^b du marché. Lorsque $a = 0$ (tranche equity), le prix théorique est en pratique généralement croissant en ρ , et on peut par dichotomie trouver ρ_0^b . Par contre, lorsque $a > 0$ (tranches mezzanine), le prix théorique n'est pas en pratique une fonction monotone de ρ , mais il est généralement croissant puis décroissant

en ρ . On trouve en général ainsi deux candidats pour ρ_a^b qui peuvent être très différents, ce qui est gênant. Pour éviter cela, la pratique est alors de calculer le prix théorique R_0^b de la tranche non cotée $[0, b]$ à partir de R_0^a et R_a^b en utilisant la relation de non arbitrage $\tilde{R}_0^a + \tilde{R}_a^b = \tilde{R}_0^b$ entre les trois tranches $[0, a]$, $[a, b]$ et $[0, b]$. Ensuite, on trouve la corrélation implicite ρ_0^b de la tranche $[0, b]$. Ce mécanisme, qui au prix d'une tranche $[a, b]$ associe une corrélation implicite de la tranche (théorique) $[0, b]$ est appelé *base correlation*. Pour l'I-Traxx Europe, cela revient à calculer $\rho_0^{3\%}$, $\rho_0^{6\%}$, $\rho_0^{9\%}$, $\rho_0^{12\%}$ et $\rho_0^{22\%}$. Pour la dernière tranche senior ($[22\%, 1]$), le prix est en général une fonction décroissante de ρ et on obtient une seule corrélation implicite notée $\rho_{22\%}^1$. De façon systématique, comme le prix d'un CDO de tranche $[0, 1]$ est une simple somme de CDS, son prix ne fait intervenir que les lois marginales des défauts et ne dépend pas de ρ . Or, nous avons vu que $\tilde{R}_0^1 - \tilde{R}_0^{22\%} = \tilde{R}_{22\%}^1$, et par conséquent la corrélation implicite sur la tranche $[22\%, 1]$ est la même que celle sur la tranche $[0, 22\%]$: $\rho_{22\%}^1 = \rho_0^{22\%}$. Les $\rho_0^{3\%}$, $\rho_0^{6\%}$, $\rho_0^{9\%}$, $\rho_0^{12\%}$ et $\rho_0^{22\%}$ obtenus ainsi avec les prix de marché sont en général très sensiblement différents ce qui invalide le modèle en tant que tel. A la différence du modèle de Black-Scholes où l'interprétation en terme de volatilité implicite est justifiée par la robustesse du modèle par rapport au paramètre de volatilité, on ne dispose pas a priori dans ce modèle de résultat de robustesse par rapport au paramètre ρ .

Chapter 2

Le modèle SSRD

Ce chapitre est un article écrit avec Damiano Brigo [18] publié dans la revue *Finance and Stochastics*.

Abstract. We introduce the two-dimensional shifted square-root diffusion (SSRD) model for interest-rate and credit derivatives with (positive) stochastic intensity. The SSRD is the unique explicit diffusion model allowing an automatic and separated calibration of the term structure of interest rates and of credit default swaps (CDS's), and retaining free dynamics parameters that can be used to calibrate option data. We propose a new positivity preserving implicit Euler scheme for Monte Carlo simulation. We discuss the impact of interest-rate and default-intensity correlation and develop an analytical approximation to price some basic credit derivatives terms involving correlated CIR processes. We hint at a formula for CDS options under CIR++ CDS-calibrated stochastic intensity.

2.1 Introduction

In this paper we introduce the two-dimensional shifted square-root diffusion (SSRD) model for interest-rate and credit derivatives in presence of stochastic intensity. One important feature of this model is the possible correlation between the interest rate and the default intensity (related to the credit spread). It is also the unique explicit model, to the best of our knowledge, providing positive intensities and interest rates that can be exactly and analytically calibrated to the term structure of interest rates and credit default swaps (CDS's). Some parameters remain to be used to calibrate cap options data in the interest-rate market and possibly options on CDS's on the credit derivatives market. Calibration to the two markets can be kept separate, which is of practical value and ensures consistency between different desks of a financial institution.

Earlier general works on intensity models appeared for example in Lando (1998) and Duffie and Singleton (1999). In the present paper we explain what we deem to be the most convenient choice within this class and give a detailed discussion of how the chosen model can be calibrated to the market and subsequently used for pricing defaultable derivatives payoffs. An extended version of this paper with more details is available in Brigo and Alfonsi (2003).

2.2 CDS's and Implied Hazard Rates

A CDS is a contract ensuring protection against default. Consider two companies “A” and “B” who agree on the following: If a third reference company “C” defaults at time $\tau < T$, before the maturity T , “B” pays to “A” a certain cash amount Z , supposed to be deterministic in the present paper, at the default time τ itself. This cash amount is a *protection* for “A” in case “C” defaults. A typical case occurs when “A” holds a corporate bond issued from “C” and waits for the coupons and final notional payment from this bond: If “C” defaults before the corporate bond maturity, “A” does not receive such payments. “A” then goes to “B” and buys some protection against this danger, asking “B” a payment that roughly amounts to the bond notional in case “C” defaults. Typically Z is equal to a notional amount, or to a notional amount minus a recovery rate. In exchange for this protection, company “A” agrees to pay periodically to “B” a fixed amount R . Payments occur at times $\mathcal{T} = \{T_1, \dots, T_n\}$ until $\tau \wedge T$. Set $\alpha_i = T_i - T_{i-1}$, $T_0 = 0$, and R is fixed in advance at time 0. We assume $T_n \leq T$, typically $T_n = T$. We may write the CDS discounted value to “B” at time t as $\mathbf{1}_{\{\tau > t\}} \times$

$$\times \left[D(t, \tau)(\tau - T_{\beta(\tau)-1})R\mathbf{1}_{\{\tau < T_n\}} + \sum_{i=\beta(t)}^n D(t, T_i)\alpha_i R\mathbf{1}_{\{\tau > T_i\}} - \mathbf{1}_{\{\tau < T\}}D(t, \tau)Z \right]$$

where $t \in [T_{\beta(t)-1}, T_{\beta(t)})$, i.e. $T_{\beta(t)}$ is the first date of $T_1, \dots, T_n$ following t . The stochastic discount factor at time t for maturity T is denoted by $D(t, T) = B(t)/B(T)$, where $B(t) = \exp(\int_0^t r_u du)$ denotes the bank-account numeraire, r being the instantaneous short interest rate. We denote by $\text{CDS}(t, \mathcal{T}, T, R, Z)$ the price at time t of the above CDS.

In general, we can compute the CDS price according to risk-neutral valuation (see for example Bielecki and Rutkowski (2002)):

$$\begin{aligned} \text{CDS}(t, \mathcal{T}, T, R, Z) &= \mathbf{1}_{\{\tau > t\}} \mathbb{E} \left\{ D(t, \tau)(\tau - T_{\beta(\tau)-1})R\mathbf{1}_{\{\tau < T_n\}} \right. \\ &\quad \left. + \sum_{i=\beta(t)}^n D(t, T_i)\alpha_i R\mathbf{1}_{\{\tau > T_i\}} - \mathbf{1}_{\{\tau < T\}}D(t, \tau)Z \middle| \mathcal{G}_t \right\} \end{aligned} \quad (2.1)$$

where $\mathcal{G}_t = \mathcal{F}_t \vee \sigma(\{\tau < u\}, u \leq t)$, $\mathcal{F}_t$ is the default-free filtration, representing the information flow of interest rates, intensities and possibly other default-free market quantities,

and $\mathbb{E}$ denotes the risk-neutral expectation in the enlarged probability space supporting τ , corresponding to the risk neutral measure $\mathbb{Q}$ under which we work in this paper. We abbreviate $\mathbb{E}\{\cdot|\mathcal{G}_0\} = \mathbb{E}\{\cdot|\mathcal{F}_0\}$ by $\mathbb{E}\{\cdot\}$.

Now we explain shortly how the market quotes CDS prices. Usually at time 0, provided default has not yet occurred, the market sets R to a value $R^{\text{MID}}(0, T)$ that makes the CDS fair, i.e. such that

$$\text{CDS}(0, \mathcal{T}, T, R^{\text{MID}}(0, T), Z) = 0. \quad (2.2)$$

In fact, in the market CDS's are quoted through a bid and an ask value for this "fair" $R^{\text{MID}}(0, T)$, for a set of canonical maturities $T = 1y$ up to $T = 10y$.

Consider now a positive deterministic function γ (the deterministic intensity or hazard rate) and $\Gamma(T) = \int_0^T \gamma(t)dt$. Consider a standard unit-mean exponential random variable ξ independent of any default-free financial quantity (and thus of $\mathcal{F}$), and define the default time τ as $\tau = \Gamma^{-1}(\xi)$. In particular, we have $\mathbb{Q}\{s < \tau \leq t\} = \exp(-\Gamma(s)) - \exp(-\Gamma(t))$. This amounts to modeling τ as the first jump of a Poisson process with intensity γ . We assume the interest rate dynamics to be driven by $\mathcal{F}_t$ -Brownian motion(s) W on the same probability space, and to be calibrated to the observed zero coupon bonds $P(t, T)$ at time $t = 0$ for the quoted maturities T . In this context, the CDS price (2.1) is easily computed by a careful use of iterated expectations.

$$\begin{aligned} \text{CDS}(t, \mathcal{T}, T, R, Z; \Gamma(\cdot)) = \mathbf{1}_{\{t < \tau\}} & \left[R \int_t^{T_n} P(t, u) (T_{\beta(u)-1} - u) d_u (e^{-(\Gamma(u)-\Gamma(t))}) + \right. \\ & \left. \sum_{i=\beta(t)}^n P(t, T_i) R \alpha_i e^{\Gamma(t)-\Gamma(T_i)} + Z \int_t^T P(t, u) d_u (e^{-(\Gamma(u)-\Gamma(t))}) \right]. \quad (2.3) \end{aligned}$$

One may wish to calibrate the deterministic-intensity model to CDS quotes $R^{\text{MID}}(0, T)$ in order to value different payoffs. To do so, one has to invert the model formula and find the Γ 's that match the given CDS market quotes, by solving in Γ a set of equations for different T : $\text{CDS}(0, \mathcal{T}, T, R^{\text{MID}}(0, T), Z; \Gamma(\cdot)) = 0$. We can assume a piecewise constant hazard rate γ , constant among different maturities T , and invert prices in an iterative way as T increases. Other possibilities include a piecewise *linear* γ or some parametric forms for γ such as Nelson and Siegel or extensions thereof. In the following we denote by γ^{mkt} and Γ^{mkt} respectively the (implied) hazard rate and hazard function that are obtained in a deterministic model when calibrating CDS market data as above. We close this section by giving as an example, in Figure 2.1, the risk-neutral default probabilities associated with a piecewise linear hazard rate $\gamma^{\text{mkt}}(t)$ obtained by calibrating the 1y, 3y, 5y, 7y and 10y CDS's on Merrill-Lynch on October 2002. These are equal, first order in the hazard function, to the hazard function $\Gamma(t)$ itself, since $\mathbb{Q}\{\tau < t\} = 1 - \exp(-\Gamma(t)) \approx \Gamma(t)$ for small Γ .

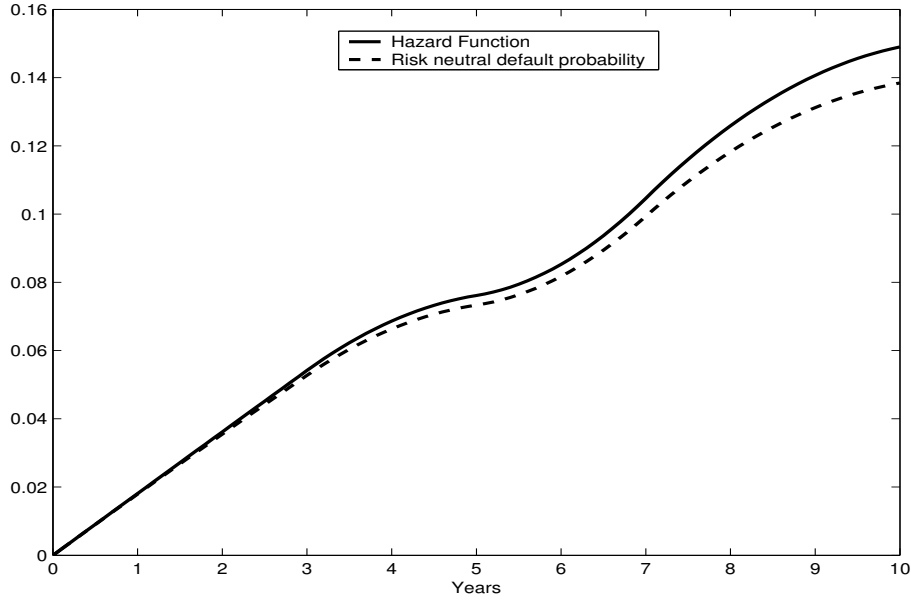


Figure 2.1: Graph of the implied hazard function $t \mapsto \Gamma^{\text{mkt}}(t)$ and implied risk-neutral default probability for Merrill-Lynch CDS's of several maturities on October 25, 2002

2.3 The SSRD intensity and interest rates model

In this section we consider a model with stochastic default intensity and interest rates. We denote by λ the default intensity in this case, instead of γ . We assume λ to be positive and we consider $\Lambda(t) = \int_0^t \lambda_u du$. The standard unit-mean exponential variable ξ introduced in the previous section is still supposed to be independent of the interest rates r (and of any default-free quantity), and is also taken independent of λ , which in turn is assumed to be $\mathcal{F}_t$ -adapted. We can then define the default time as in the deterministic intensity case by

$$\tau = \Lambda^{-1}(\xi).$$

This amounts to modeling τ as the first jump of a Cox process with intensity λ .

We now describe our assumptions on the short-rate process r and on the intensity λ dynamics. For more details on shifted r diffusion dynamics see also Avellaneda and Newman (1998), or Brigo and Mercurio (2001, 2001b).

2.3.1 CIR++ short-rate model (Brigo and Mercurio (2001))

We write the short-rate r_t as the sum of a deterministic function φ and of a Markovian process x_t^α :

$$r_t = x_t^\alpha + \varphi(t; \alpha), \quad t \geq 0, \quad (2.4)$$

where φ depends on the parameter vector α (which includes x_0^α) and is integrable on closed intervals. Notice that x_0^α is indeed one more parameter at our disposal: we are free to select its value as long as $\varphi(0; \alpha) = r_0 - x_0$. We take as reference model for x the Cox-Ingersoll-Ross (CIR) process:

$$dx_t^\alpha = k(\theta - x_t^\alpha)dt + \sigma\sqrt{x_t^\alpha}dW_t,$$

where the parameter vector is $\alpha = (k, \theta, \sigma, x_0^\alpha)$, with $k, \theta, \sigma, x_0^\alpha$ positive deterministic constants. The condition $2k\theta > \sigma^2$ ensures that the origin is inaccessible to the reference model, so that the process x^α is well defined and remains positive. As is well known, this process x^α features a noncentral *chi-square* distribution, and yields an affine term-structure of interest rates. Denote by f instantaneous forward rates, i.e. $f(t, T) = -\partial \ln P(t, T)/\partial T$. The initial market zero-coupon interest-rate curve $T \mapsto P^M(0, T)$ is automatically calibrated by our model if we set $\varphi(t; \alpha) = \varphi^{\text{CIR}}(t; \alpha)$ where $\varphi^{\text{CIR}}(t; \alpha) = f^M(0, t) - f^{\text{CIR}}(0, t; \alpha)$,

$$f^{\text{CIR}}(0, t; \alpha) = \frac{2k\theta(e^{th} - 1)}{2h + (k + h)(e^{th} - 1)} + x_0 \frac{4h^2 e^{th}}{[2h + (k + h)(e^{th} - 1)]^2}$$

with $h = \sqrt{k^2 + 2\sigma^2}$. For restrictions on the α 's that keep r positive see Brigo and Mercurio (2001, 2001b). Moreover, the price at time t of a zero-coupon bond maturing at time T is

$$P(t, T) = \frac{P^M(0, T)A(0, t; \alpha) \exp\{-B(0, t; \alpha)x_0\}}{P^M(0, t)A(0, T; \alpha) \exp\{-B(0, T; \alpha)x_0\}} P^{\text{CIR}}(t, T, r_t - \varphi^{\text{CIR}}(t; \alpha); \alpha) \quad (2.5)$$

where $P^{\text{CIR}}(t, T, x_t; \alpha) = \mathbb{E}(e^{-\int_t^T x^\alpha(u)du} | \mathcal{F}_t) = A(t, T; \alpha) \exp\{-B(t, T; \alpha)x_t\}$ is the bond price formula for the basic CIR model with the classical expressions for A and B given for example in (3.25) of Brigo and Mercurio (2001b). From P 's the spot LIBOR rate $L(t, T)$ at t for maturity T , the forward LIBOR rates $F(t, T, S)$ at t for maturity T and expiry S , and all other rates can be computed as explicit functions of r_t .

The cap price formula for the CIR++ model can be derived in closed form from the corresponding formula for the basic CIR model. This formula is a function of the parameters α . In our application we will calibrate the parameters α to cap prices, by inverting the analytical CIR++ formula, so that our interest rate model is calibrated to the initial zero coupon curve through φ and to the cap market through α , as in Brigo and Mercurio (2001, 2001b).

2.3.2 CIR++ intensity model (Brigo and Alfonsi (2003))

For the intensity model we adopt a similar approach, in that we set

$$\lambda_t = y_t^\beta + \psi(t), \quad t \geq 0, \quad (2.6)$$

where ψ is a positive deterministic function that is integrable on closed intervals. As before, the parameter vector is $\beta = (\kappa, \mu, \nu, y_0^\beta)$, with $\kappa, \mu, \nu, y_0^\beta$ positive deterministic constants such that $2\kappa\mu > \nu^2$, and we take y again of the form:

$$dy_t^\beta = \kappa(\mu - y_t^\beta)dt + \nu\sqrt{y_t^\beta}dZ_t,$$

where the process Z is a $\mathcal{F}_t$ -Brownian motion. This ensures that λ be strictly positive, as should be for an intensity process. Notice incidentally that this basically forces the choice of a tractable y to the CIR model among all one factor short-rate diffusion models. Dependence of ψ on β and possibly on other parameters will be specified later when dealing with CDS calibration. We will often use the integrated process, that is $\Lambda(t) = \int_0^t \lambda_s ds$, and also $Y^\beta(t) = \int_0^t y_s^\beta ds$ and $\Psi(t) = \int_0^t \psi(s)ds$. We assume the short rate r and the intensity λ processes to be correlated, by assuming the driving Brownian motions W and Z to be instantaneously correlated according to $dW_t dZ_t = \rho dt$.

2.3.3 Joint SSRD model calibration to CDS: Separability

The SSRD model is characterized by the terms $\mathcal{P} = (\alpha, \varphi, \beta, \psi, \rho)$ and can be seen as an extension of the CIR++ model for interest rates. As we explained before, φ is chosen to fit exactly the default-free zero-coupon bonds and α is then selected to have the better approximation of the cap prices. This procedure for the CIR++ interest rate part of the model r still remains valid in presence of a correlated λ since the products used for that calibration do not depend on the dynamics of λ and of (β, ρ, ψ) .

Once α and φ are fixed, we would like to fit the three remaining terms to the credit derivatives market. To do so, we need first to calculate the price of CDS's in the SSRD model. We find easily, through iterated expectations and the definition of τ , that (see Brigo and Alfonsi (2003) for the details)

$$\begin{aligned} \text{CDS}(0, T, T, R, Z; \mathcal{P}) &= R \int_0^{T_n} \mathbb{E}[\exp\left(-\int_0^u (r_s + \lambda_s)ds\right) \lambda_u] (u - T_{\beta(u)-1}) du \quad (2.7) \\ &+ \sum_{i=1}^n \alpha_i R \mathbb{E}[\exp\left(-\int_0^{T_i} (r_s + \lambda_s)ds\right)] - Z \int_0^T \mathbb{E}[\exp\left(-\int_0^u (r_s + \lambda_s)ds\right) \lambda_u] du. \end{aligned}$$

We plan to use ψ to calibrate exactly the market CDS quotes. More precisely, we want to find, for each (β, ρ) , a function $\psi_\alpha(\cdot; \beta, \rho)$ that makes the CDS present values null, in agreement with (2.2). This could be done if we were able to calculate analytically the above expectations in general, taking as in the deterministic case a specific shape for ψ_α . Since these expectations are known only when $\rho = 0$, we first restrict ourselves to calibrate the subclass of models with $\rho = 0$. Interest rates and default intensities are independent with $\rho = 0$. By switching expectation and differentiation with respect to u and Fubini's

theorem it is easy to see that the price of the CDS satisfies the deterministic case formula (2.3) when replacing terms such as $\exp(\Gamma(t) - \Gamma(u))$ by $\mathbb{E}(\exp(\Lambda(t) - \Lambda(u)))$ (with $u \geq t$). Therefore, at time $t = 0$, for any β we can calibrate automatically our model to the CDS by choosing ψ such that

$$e^{-\Gamma^{\text{mkt}}(u)} = \mathbb{E}(e^{-\Lambda(u)}) = e^{-\Psi(u)} \mathbb{E}(e^{-Y^\beta(u)}) = e^{-\Psi(u)} P^{\text{CIR}}(0, u, y_0; \beta).$$

The remarkable point is that ψ does not depend on α (the zero-coupon bonds have been calibrated exactly earlier), so that this calibration to CDS can be done independently of the interest rate calibration. This “separability” is of practical interest. We thus denote by $\psi(\cdot; \beta)$ the obtained ψ function, given by

$$\psi(u; \beta) = \gamma^{\text{mkt}}(u) + \frac{d}{du} \ln(\mathbb{E}(e^{-Y^\beta(u)})) = \gamma^{\text{mkt}}(u) + \frac{d}{du} \ln(P^{\text{CIR}}(0, u, y_0; \beta)). \quad (2.8)$$

The shape of ψ is partly implicitly specified by our choice for γ^{mkt} (piecewise linear or otherwise). Now that CDS's are automatically calibrated, we would like to calibrate the parameters β to some *options* on the credit derivatives market in the same way as α is used to fit cap prices. Unfortunately, even as we write there are no options on credit derivatives that are liquid enough for this. Quotes for CDS options appeared in 2003 but they featured too large bid/ask spreads to be taken into account for calibration. It seems thus more reasonable, for the time being, to change the point of view. Since λ is a time-homogeneous CIR process y^β shifted by a deterministic function $\psi(\cdot; \beta)$, we try to keep $\psi(\cdot; \beta)$ “uniformly” at a minimum. Lacking any other information, we decide to privilege time-homogeneity and we look for the λ model that is as close as possible to a time-homogeneous CIR process y . We minimize the criterion $\int_0^T \psi(u; \beta)^2 du$ under the constraint $\psi(u; \beta) \geq 0$, $u \in [0, T]$. The calibration to the same CDS data as at the end of Section 2.2 up to a ten years maturity leads to the following results :

$$\begin{aligned} \beta & : \quad \kappa = 0.354201, \mu = 0.00121853, \nu = 0.0238186; y_0 = 0.0181, \\ (\alpha & : \quad k = 0.528905, \theta = 0.0319904, \sigma = 0.130035, x_0 = 8.32349 \times 10^{-5}). \end{aligned}$$

Let us go back to the calibration in the general case with $\rho \in [-1, 1]$. We should find for any (β, ρ) a function $\psi_\alpha(\cdot; \beta, \rho)$ that makes the CDS's present value null, and then find (β, ρ) . *The key point is that with our SSRD model, the prices of CDSs essentially do not depend on ρ : their dependence on ρ is negligible with respect to the CDS bid/ask spread. Thus, we can still take ψ as in (2.8) even if $\rho \neq 0$.* Indeed, let us take α , $\varphi(\cdot; \alpha)$ and $\psi(\cdot; \beta)$ coming from our $\rho = 0$ calibration, and let β be any parameter that ensures $\psi(\cdot; \beta) \geq 0$. With these $(\varphi, \alpha, \psi, \beta)$ terms, we can calculate the values of the CDS prices for the extreme values $\rho = -1$ and $\rho = 1$ (and their differences) with a Monte-Carlo (MC) algorithm. We observe (see Table 2.3) for the 5y CDS (with $Z = 0.593$) differences smaller

than $3E - 4$, and we are well within the bid/ask window

$$\text{CDS}(0, \mathcal{T}, T, R^{\text{BID}}, Z) = -17.14E - 4, \quad \text{CDS}(0, \mathcal{T}, T, R^{\text{ASK}}, Z) = 17.16E - 4, \quad (2.9)$$

corresponding to quoted $R^{\text{BID}} = 0.009$ and $R^{\text{ASK}} = 0.0098$. Thus, we can consider that the prices of CDS obtained are null for any ρ , so that the function $\psi(\cdot; \beta)$ calibrated under the assumption $\rho = 0$ makes also the CDS's prices null when $\rho \neq 0$ as desired. Then, β can be calibrated as before, while we must wait that a liquid credit derivative asset appears to calibrate ρ (and possibly to better calibrate β). However, it is interesting to use this α and β calibrated model to have an idea of the impact of dependence between λ and r on credit derivatives prices. As a different possibility, one might estimate ρ historically from time series of proxies of λ 's and r 's (for example from CDS stripped 3 months default probabilities one may obtain a proxy of λ and from three months LIBOR rates a proxy of r).

2.3.4 Numerical simulation of the SSRD

To calculate the price of credit derivatives with a MC algorithm, we need to simulate paths of (r, λ) and the difficulty lies in simulating a CIR process. It is well known that a standard Euler scheme can lead to negative values and then to complex values even if $2k\theta > \sigma^2$, which we assume to hold. We present here briefly the implicit positivity-preserving Euler scheme (IPPES) introduced in Brigo and Alfonsi (2003) (to which we refer for the details) that ensures positivity and preserves the following positivity property : “for a given path $(W_{t_i}(\omega))_i$, $x_0 \leq \bar{x}_0$ implies $\tilde{x}_{t_i}^\alpha(\omega) \leq \tilde{x}_{t_i}^{\bar{\alpha}}(\omega)$ for all t_i 's” (where $\bar{\alpha} = (k, \theta, \sigma, \bar{x}_0)$ and $\tilde{x}$ is the scheme process). This positivity preserving property holds for the original process in continuous time.

Let us consider a time interval $[0, T]$ with the standard time-discretization $(t_i = i \cdot T/n, 0 \leq i \leq n)$. The CIR process x^α satisfies, in L^2 ,

$$x_t^\alpha = \lim_{n \rightarrow \infty} \left[x_0^\alpha + \sum_{i; t_i < t} (k\theta - \frac{\sigma^2}{2} - kx_{t_{i+1}}^\alpha)T/n + \sigma \sum_{i; t_i < t} \sqrt{x_{t_{i+1}}^\alpha} (W_{t_{i+1}} - W_{t_i}) \right] \quad \text{We will focus}$$

then on the following IPPES (Brigo and Alfonsi (2003)):

$$\tilde{x}_{t_{i+1}}^\alpha = \tilde{x}_{t_i}^\alpha + (k\theta - \frac{\sigma^2}{2} - k\tilde{x}_{t_{i+1}}^\alpha)T/n + \sigma \sqrt{\tilde{x}_{t_{i+1}}^\alpha} (W_{t_{i+1}} - W_{t_i}).$$

It follows that $\sqrt{\tilde{x}_{t_{i+1}}^\alpha}$ is the unique positive root of the polynomial $P(X) = (1 + kT/n)X^2 - \sigma(W_{t_{i+1}} - W_{t_i})X - (\tilde{x}_{t_i}^\alpha + (k\theta - \frac{\sigma^2}{2})T/n)$, and we get

$$\tilde{x}_{t_{i+1}}^\alpha = \left(\frac{\sigma(W_{t_{i+1}} - W_{t_i}) + \sqrt{\Delta(\tilde{x}_{t_i}^\alpha, W_{t_{i+1}} - W_{t_i})}}{2(1 + kT/n)} \right)^2, \quad (2.10)$$

where $\Delta(\tilde{x}_{t_i}^\alpha, W_{t_{i+1}} - W_{t_i}) = \sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\tilde{x}_{t_i}^\alpha + (k\theta - \frac{\sigma^2}{2})T/n)(1 + kT/n)$. Since this expression is clearly increasing in $\tilde{x}_{t_i}^\alpha$, we obtain the positivity preserving property above, and in particular the positivity of $\tilde{x}_{t_{i+1}}^\alpha$ is guaranteed.

Now, through the IPPEs we are able to run a MC algorithm for pricing. For products whose price is known conditional on τ exceeding the maturity (no default), like CDS's, we may speed up this algorithm through a “threshold barrier” $\bar{B}$ such that $\mathbb{Q}(\Lambda(T) < \bar{B}) \simeq 1$. We thus assume that default may occur only when $\Lambda(\tau) < \bar{B}$. We simulate default times conditional on $\xi := \Lambda(\tau) < \bar{B}$. Indeed, we see that if “DCDS” is the CDS discounted payoff, recalling that $\Lambda(\tau)$ is exponential with parameter 1 independent of $\mathcal{F}$, we have that

$$\mathbb{E} \text{DCDS} = \mathbb{E}[\text{DCDS} | \Lambda(\tau) < \bar{B}] (1 - e^{-\bar{B}}) + \mathbb{E}[\text{DCDS} | \Lambda(\tau) \geq \bar{B}] e^{-\bar{B}}.$$

The CDS value is known in case $\xi > \bar{B}$, since we assume no default in this case. Our simulations then need concern only the first term, and we just simulate $\xi | \xi < \bar{B}$, whose density is $p_{\xi | \xi < \bar{B}}(u) = 1_{\{u < \bar{B}\}} e^{-u} / (1 - e^{-\bar{B}})$. Simulating N scenarios for ξ amounts to simulate $N(1 - e^{-\bar{B}})$ scenarios with $\xi < \bar{B}$ and $Ne^{-\bar{B}}$ with $\xi \geq \bar{B}$. So in turn simulating $M = N(1 - e^{-\bar{B}})$ scenarios for $\xi < \bar{B}$, as we will do, amounts to simulate in total $N = M/(1 - e^{-\bar{B}})$ scenarios, the extra scenarios corresponding to the known value $R \sum_{i=1}^n P(0, T_i) \alpha_i$ of the CDS in case of no default. Dividing by $1 - e^{-\bar{B}}$ may help us increase efficiency (in our examples typically it increases the number of scenarios by a factor 10), but the time needed for MC simulation remains large, due also to the presence of indicator functions in the payoff.

2.3.5 Gaussian dependence mapping: A tractable approximated SSRD and numerical tests

The SSRD model pricing formulas typically involve terms such as

$$h_i^{\text{CIR}}(T; \alpha, \beta, \rho) = \mathbb{E} \left[\exp \left(- \int_0^T (x_s^\alpha + y_s^\beta) ds \right) (y_T^\beta)^{i-1} \right], i \in \{1, 2\},$$

as in the CDS's case (2.7). These terms admit no exact closed form when $\rho \neq 0$, and in this case it would be convenient to have an analytical approximation. Our idea is to “map” the two-dimensional CIR dynamics in an analogous tractable two-dimensional Vasicek dynamics that preserves as much as possible of the original CIR structure, and then do calculations with the Vasicek Gaussian model using the following lemmas.

Lemma 2.3.1. *Let $A = m_A + \sigma_A N_A$ and $B = m_B + \sigma_B N_B$ be two random variables such that N_A and N_B are two correlated standard Gaussian random variables with $[N_A, N_B]$ jointly Gaussian vector with correlation $\bar{\rho}$. Then,*

$$\mathbb{E}(e^{-A} B) = m_B e^{-m_A + \frac{1}{2} \sigma_A^2} - \bar{\rho} \sigma_A \sigma_B e^{-m_A + \frac{1 - \bar{\rho}^2}{2} \sigma_A^2} \quad (2.11)$$

Lemma 2.3.2. *Let $x_t^{\alpha,V}$ and $y_t^{\beta,V}$ be two Vasicek processes as follows:*

$$dy_t^{\beta,V} = \kappa(\mu - y_t^{\beta,V})dt + \nu dZ_t, \quad dx_t^{\alpha,V} = k(\theta - x_t^{\alpha,V})dt + \sigma dW_t \quad (2.12)$$

with $dW_t dZ_t = \rho dt$. Then $A = \int_0^T (x_t^{\alpha,V} + y_t^{\beta,V})dt$ and $B = y_T^{\beta,V}$ are Gaussian random variables with respective means and variances:

$$\begin{aligned} m_A &= (\mu + \theta)T - [(\theta - x_0)g(k, T) + (\mu - y_0)g(\kappa, T)] \\ m_B &= \mu - (\mu - y_0)e^{-\kappa T} \\ \sigma_A^2 &= \left(\frac{\nu}{\kappa}\right)^2 (T - 2g(\kappa, T) + g(2\kappa, T)) + \left(\frac{\sigma}{k}\right)^2 (T - 2g(k, T) + g(2k, T)) \\ &\quad + \frac{2\rho\nu\sigma}{k\kappa} (T - g(\kappa, T) - g(k, T) + g(\kappa + k, T)) \\ \sigma_B^2 &= \nu^2 g(2\kappa, T) \end{aligned}$$

and correlation $\bar{\rho} = \frac{1}{\sigma_A \sigma_B} \left[\frac{\nu^2}{\kappa} (g(\kappa, T) - g(2\kappa, T)) + \frac{\rho \sigma \nu}{k} (g(\kappa, T) - g(\kappa + k, T)) \right]$ where $g(k, T) = (1 - e^{-kT})/k$.

Thus, we are able to calculate, for any ρ , $h_i^V(T; \alpha, \beta, \rho) = \mathbb{E}[\exp(-\int_0^T (x_s^{\alpha,V} + y_s^{\beta,V})ds)(y_T^{\beta,V})^{i-1}]$ with $i \in \{1, 2\}$, and we use the following approximation:

$$h_i^{\text{CIR}}(T; \alpha, \beta, \rho) \approx h_i^{\text{CIR}}(T; \alpha, \beta, 0) + h_i^V(T; \alpha_T^V, \beta_T^V, \rho) - h_i^V(T; \alpha_T^V, \beta_T^V, 0) \quad (2.13)$$

for suitable parameters $\alpha_T^V = (k_T^V, \theta_T^V, \sigma_T^V, x_0^V)$ and $\beta_T^V = (\kappa_T^V, \mu_T^V, \nu_T^V, y_0^V)$. Since all these parameters but volatility play an analogous role in both models, we take $\alpha_T^V = (k, \theta, \sigma_T^V, x_0)$ and $\beta_T^V = (\kappa, \mu, \nu_T^V, y_0)$, and we decide to take σ_T^V and ν_T^V such that $\mathbb{E} \left[\exp \left(-\int_0^T x_s^{\alpha_T^V, V} ds \right) \right] = \mathbb{E} \left[\exp \left(-\int_0^T x_s^\alpha ds \right) \right]$ (and analogously for y). In practical cases, these volatilities exist and can be seen as some sort of means of time-averages of $\sigma \sqrt{x_s^\alpha}$ (resp. $\nu \sqrt{y_s^\beta}$) on $[0, T]$. We perform numerical tests for (2.13) and the related CDS prices, based on MC simulations of the left-hand sides. In Table 2.1, we take the α and β parameters as from Section 2.3.3, and assume $T = 5y$. The Vasicek mapped volatilities are $\sigma_{5y}^V = 0.016580$ and $\nu_{5y}^V = 0.0025675$. To check the quality of the approximation under stress, we multiply all parameters k, θ, σ and κ, μ, ν by three and check again the approximation. The results are shown in Table 2.2, and now the Vasicek mapped volatilities are $\sigma_{5y}^V = 0.108596$ and $\nu_{5y}^V = 0.0060675$.

Since the values h_1 in Table 2.1 can be interpreted as defaultable-bond prices, the corresponding continuously compounded spot rates would be $-\ln(0.86191)/5 = 0.02972$ and $-\ln(0.861762)/5 = 0.029755$, respectively, giving a small difference. With Table 2.2 we obtain $-\ln(0.64232)/5 = 0.088534$ and $-\ln(0.641989)/5 = 0.088637$, with a larger difference than before, ranging around 1 basis point, which is however still contained. So we may trust the approximation to work well within the typical market bid-ask spreads

	$\rho = -1$ (95% MC window)	$\rho = 1$ (95% MC window)
h_1 , LHS of (2.13)	0.86191 (0.861815 0.862004)	0.8624 (0.862272 0.862529)
h_1 , RHS of (2.13)	0.861762,	0.862554
h_2 , LHS of (2.13)	3.5848E-3 (3.57946 3.59014)	3.44852E-3 (3.44408 3.45295)
h_2 , RHS of (2.13)	3.59831E-3	3.43174E-3

Table 2.1: MC simulation for the quality of the approximation (2.13)

	$\rho = -1$ (95% MC window)	$\rho = 1$ (95% MC window)
h_1 , LHS of (2.13)	0.64232 (0.642106 0.642534)	0.644151 (0.643909 0.644393)
h_1 , RHS of (2.13)	0.641989	0.643904
h_2 , LHS of (2.13)	2.4757E-3 (2.46991 2.48149)	2.27465E-3 (2.27018 2.27913)
h_2 , RHS of (2.13)	2.53527	2.24435

Table 2.2: MC simulation for the quality of the approximation (2.13) under stress

for CDS's. Indeed, we consider the valuation of CDS's both by MC simulation and by the Gaussian dependence mapped model, where we apply formula (2.13) each time with the most convenient maturity T for that part of the CDS payoff we are evaluating.

In Table 2.3 we present the related results for the extreme cases $\rho = -1$ and $\rho = 1$ and the known case $\rho = 0$, both under the usual parameters of Section 2.3.3 and under some amplified λ parameters, increasing stochasticity. The Gaussian mapping approximation, even in the case of increased randomness, remains well within a small fraction of the CDS bid-ask spread (2.9).

CDS prices	Gaussian Mapping	Monte Carlo value and 95% window
$\rho = -1$	-1.12E-4	-1.48625E-4 (-1.79586 -1.17664)
$\rho = 0$	0.012E-4	0.17708E-4 (-0.142444 0.496605)
$\rho = 1$	1.14E-4	1.25475E-4 (0.922997 1.5865)

Same run with κ, ν increased by a factor 5 and μ by a factor 3 :

CDS prices	Gaussian Mapping	Monte Carlo value and 95% window
$\rho = -1$	-1.03E-4	-1.77E-4 (-2.02 -1.51)
$\rho = 0$	0.021E-4	0.143E-4 (-0.138 0.424)
$\rho = 1$	1.07E-4	1.08E-4 (0.78 1.37)

Table 2.3: 5y CDS prices as functions of ρ : MC vs Gaussian Mapping

2.3.6 The impact of correlation and numerical tests

It can be interesting to study the main terms that appear in basic payoffs of the credit derivatives world from the point of view of the impact of the correlation ρ between r and λ . Consider the following payoffs

$$\begin{aligned} A &= L(T - 1y, T)D(0, T)\mathbf{1}_{\{\tau < T\}}, & B &= D(0, \tau)\mathbf{1}_{\{\tau < T\}} \\ C &= D(0, \tau \wedge T), & D &= D(0, T)L(T - 1y, T)\mathbf{1}_{\{\tau \in [T - 1y, T]\}}, \end{aligned} \quad (2.14)$$

under our SSRD correlated model. We will see that in all cases even high correlations ρ between r and λ induce a small dependence effect on the particular functional forms of $D(0, \cdot)$ in r and of indicators of the default times τ in λ . Higher effects are observed, in relative terms, when terms such as $L(T - 1y, T)$ and $\mathbf{1}_{\{\tau \in [T - 1y, T]\}}$ are included in the payoff. Indeed, the indicator isolates λ between $T - 1y$ and T , while L isolates r between $T - 1y$ and T , so that the direct correlation between r and λ has a more direct impact than in other cases. This explains the highest percentage influence of correlations in payoff D . Results are summarized in Table 2.4. We have used the same W and Z paths when changing ρ from -1 to 1 , and we have taken $T = 5y$ and the same parameters in the dynamics as in Section 2.3.3. To check that indeed it is the “localization” of λ and r in the same interval $[T - 1y, T] = [4y, 5y]$ that generates the high relative influence of ρ , we consider also the terms E, F, G and H, defined by

$$D(0, 5)L(4, 5)\mathbf{I}_4, \quad D(0, 5)L(4, 5)\mathbf{I}_3, \quad D(0, 5)L(4, 5)\mathbf{I}_2, \quad D(0, 5)L(4, 5)\mathbf{I}_1, \quad (2.15)$$

respectively, where $\mathbf{I}_i = \mathbf{1}_{\{\tau \in [i-1, i]\}}$, and check that the correlation decreases as τ gets far from the 4y LIBOR reset date. Table 2.4 shows that this happens.

	$\rho = -1$	$\rho = 1$	relative variation	absolute variation
A	30.3672E-4	31.1962E-4	+2.73%	+0.829E-4
B	679.197E-4	676.208E-4	-0.44%	-2.989E-4
C	8207.23E-4	8209.61E-4	+0.03%	+2.38E-4
D	2.77376E-4	3.10889E-4	+12.08%	+0.34E-4
E	5.6E-4	5.88E-4	+5.010%	+0.281E-4
F	7.16E-4	7.31 E-4	+2.09%	+0.149E-4
G	7.41E-4	7.44E-4	+0.36%	2.66E-6
H	7.55E-4	7.56E-4	+0.056%	4.26 E-7

Table 2.4: Influence of ρ on the miscellaneous terms defined above

2.4 Examples of pricing with the SSRD model

In this final section we present an example of payoff that can be valued with the calibrated (λ, r) model. The example we consider is a sort of cancellable swap with a recovery value. A first company “A” owns a bond issued by “C” (say Merrill Lynch), and receives from “C” once an year at time T_i a payment consisting of $L(T_i - 1, T_i) + s$, where s is a spread, up to a final date $T = T_n = 5y$. We assume unit year fractions. In turn, “A” has a swap with a bank “B”, where “A” turns the payment $L(T_i - 1y, T_i) + s$ to “B”, and, in exchange for this, the company “A” receives from “B” some fixed payments that we express as the percentages of the unit nominal value given in (2.16).

Year	$T_1 = 1$	$T_2 = 2$	$T_3 = 3$	$T_4 = 4$	$T_5 = 5$
%	$\alpha_1 = 4.20$	$\alpha_2 = 3.75$	$\alpha_3 = 3.25$	$\alpha_4 = 0.50$	$\alpha_5 = 0.50$

(2.16)

However, if “C” defaults, “A” receives a recovery rate $\tilde{Z}$ from “C” (typically one recovers from $\tilde{Z} = 0$ to 0.5 out of 1), and still has to pay the remaining payments $L(T_i - 1, T_i) + s$ to “B”. “A” wishes to have the possibility to cancel the swap with “B” in case both “C” defaults and the recovery rate $\tilde{Z}$ is not enough to close the swap with “B” without incurring in a loss. Continuing the swap after the default τ implies for “A” to pay cash flows whose total discounted value at time τ is:

$$X(\tau) := -\tilde{Z} + \sum_{i=\beta(\tau)}^n P(\tau, T_i) (-\alpha_i + s + F(\tau; T_{i-1}, T_i)). \quad (2.17)$$

“A” wishes to cancel this payment when it is positive. This cancellation has as value at time 0 the expectation of $D(0, \tau) 1_{\{\tau < T_n\}} (X(\tau))^+$. By a joint simulation of (λ, r) this payoff can be easily valued. Our results, with the same interest-rate and default-intensity dynamics (r, λ) as in Section 2.3.3 are reported in Table 2.5 (recovery $\tilde{Z} = 0$) and show a relevant impact of ρ .

In case one assumes deterministic r , one can value analytically CDS options under our CDS-calibrated stochastic λ , thanks to Jamshidian’s decomposition. This would allow us to calibrate a CIR++ stochastic intensity model, with deterministic interest rates, to CDS (ψ) and CDS options (β) . The calibration of a model with stochastic and correlated r could be attempted by combining Jamshidian’s decomposition with one-dimensional integration and the Gaussian dependence mapping method given earlier. We would obtain a formula similar in spirit to formula (4.31) in Brigo and Mercurio (2001b) for swaptions in the G2++ model. The accuracy of such a formula has to be tested. Details on Jamshidian’s decomposition for CDS options and further numerical examples with the cancellable structure are in Brigo and Alfonsi (2003).

$s \backslash \rho$	-1	0	1	Det
-100	32.56 (32.15, 32.97)	34.26 (33.83, 34.69)	36.24 (35.78, 36.70)	34.38
-50	43.48 (42.96, 44.00)	45.19 (44.65, 45.74)	47.03 (46.46, 47.59)	45.08
0	54.351 (53.71, 54.99)	55.59 (54.94, 56.25)	57.40 (56.72, 58.08)	55.79
+50	64.91 (64.15, 65.67)	66.26 (65.48, 67.04)	68.25 (67.45, 69.05)	66.49
+100	75.64 (74.76, 76.53)	76.78 (75.88, 77.68)	78.81 (77.89, 79.73)	77.20

Table 2.5: Cancellable swap price in basis points (10^{-4}) as a function of ρ and s with MC simulation, $\tilde{Z} = 0$, “Det” for deterministic model

2.5 Conclusions and further research

We have introduced the two-dimensional SSRD model, providing an automatic calibration of the term structure of interest rates and of CDS's. Additional parameters can be set so as to calibrate option data from the interest rate market and option data on the credit market, although we do not use the latter, due to the fact that the related products appeared recently and are characterized by wide bid-ask spreads. The interest-rate calibration and the credit market calibration can be kept separate, which can be helpful in splitting competence while maintaining consistency. We discussed numerically the impact of interest-rate and default-intensity correlation on calibration and pricing. We also introduced an analytical approximation based on a Gaussian dependence mapping for some basic credit derivatives terms involving correlated CIR processes. We hinted at a formula for CDS options under CIR++ CDS-calibrated intensity based on Jamshidian's decomposition. Further research includes checking future default structures implied by a calibration and related diagnostics, the impact of correlation on more involved payoffs, possible extensions to multiline credit derivatives, and analyzing hedging strategies associated with the model, for example in the framework of Blanchet-Scalliet and Jeanblanc (2004).

Chapter 3

Les copules “périodiques”

Ce chapitre est un article écrit avec Damiano Brigo [3] publié dans la revue *Communications in Statistics: Theory and Methods*.

Abstract. Although there exists a large variety of copula functions, only a few are practically manageable, and often the choice in dependence modeling falls on the Gaussian copula. Further, most copulas are exchangeable, thus implying symmetric dependence. We introduce a way to construct copulas based on periodic functions. We study the two-dimensional case based on one dependence parameter and then provide a way to extend the construction to the n -dimensional framework. We can thus construct families of copulas in dimension n and parameterized by $n - 1$ parameters, implying possibly asymmetric relations. Such “periodic” copulas can be simulated easily.

3.1 Introduction and Motivation

Consider a random vector $X = (X_1, \dots, X_n)$, and suppose that we wish to analyze the dependence between its components. The law of X is then characterized by its joint cumulative distribution function (CDF) $(x_1, \dots, x_n) \mapsto \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n)$. However, this function mixes information on the dependence between the different components of the vector with information on the distribution of the single components themselves. Copula functions have been introduced in order to allow for a separation between the marginal CDF and the dependence structure. The former concerns single components, taken one at the time, and is given by the CDF's $F_i(x) := \mathbb{P}(X_i \leq x)$, $i = 1, \dots, n$, which we assume to be continuous. The latter is entirely represented by the copula function we introduce now. It is well known that $U_1 = F_1(X_1), \dots, U_n = F_n(X_n)$ are uniformly distributed random variables on $[0, 1]$. The joint cumulative distribution function of $(U_1, \dots, U_n)$, that we denote by $C(u_1, \dots, u_n) = \mathbb{P}(U_1 \leq u_1, \dots, U_n \leq u_n)$, is called the copula function of $(X_1, \dots, X_n)$ and

satisfies:

$$\mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = C(\mathbb{P}(X_1 \leq x_1), \dots, \mathbb{P}(X_n \leq x_n)). \quad (3.1)$$

One can easily check that a copula has the following properties:

- i) $C(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_n) = 0$;
- ii) $C(1, \dots, 1, u_k, 1, \dots, 1) = u_k$;
- iii) $\partial_{u_1 \dots u_n} C$ is a positive measure in the sense of Schwartz distributions. This means concretely that for any hypercube $H = [a_1, b_1] \times \dots \times [a_n, b_n] \subset [0, 1]^n$ we have $\mathbb{P}[(U_1, \dots, U_n) \in H] \geq 0$.

Conversely, one can show that any function that satisfies these three conditions can be viewed as the joint CDF of a vector of uniform variables on $[0, 1]$ and is thus a copula. This is known as Sklar’s theorem, see for example Joe [44] or Nelsen [57]. In this spirit, *in the following, the expression “simulating a copula C ” will denote the simulation of a random vector of uniform variables $(U_1, \dots, U_n)$ on $[0, 1]$ whose joint CDF is C .*

Among the different ways to define specific copula functions, there are following two. The first one consists in seeking functions C satisfying the three above properties. Archimedean copulas are an example of this approach. Indeed, Archimedean copulas come from the remark that if φ is a completely monotone function such that $\varphi(1) = 0$ and $\varphi(0) = +\infty$, $C(u_1, \dots, u_n) = \varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_n))$ has the above three properties and is thus a copula. Therefore, by specifying families of completely monotone functions, we specify families of copulas (e.g. Gumbel, Joe, Frank...), see Joe [44], Nelsen [57], Bouyé et al. [17], and Genest and Rivest [34]. The second method consists in working directly with joint CDF’s $F(x_1, \dots, x_n)$ and the related marginal CDF’s F_i , providing the copula $F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n))$. Even if this method does not always lead to analytically tractable copulas, it can provide us with copulas that are easy to simulate. Indeed, the main example of this kind of construction is the fundamental family of Gaussian copulas. A Gaussian copula is defined as the copula of a joint Gaussian random vector X with standard Gaussian marginals and correlation matrix ρ , and is thus given by $N_\rho(N^{-1}(u_1), \dots, N^{-1}(u_n))$ where N is the CDF of a scalar standard Gaussian variable and N_ρ is the joint CDF of X . A possible drawback of Archimedean and Gaussian copulas is symmetry, and is related to the following definition.

Definition 3.1.1. (k -exchangeability). *Let us consider a copula C that is the CDF of the random vector $(U_1, \dots, U_n)$ (with standard uniform margins). The copula C is said to be k -exchangeable ($2 \leq k \leq n$) if, for any $1 \leq i_1 < i_2 < \dots < i_k \leq n$ and any permutation σ on $\{1, \dots, k\}$, $(U_{i_1}, \dots, U_{i_k})$ and $(U_{i_{\sigma(1)}}, \dots, U_{i_{\sigma(k)}})$ have the same law.*

It is clear, with this definition, that a k' -exchangeable copula is also a k -exchangeable copula whenever $k' > k$, and it is also clear that an Archimedean copula is n -exchangeable. Let us now examine the case of the Gaussian copulas and consider a Gaussian random vector $(N_1, \dots, N_n)$ of n standard Gaussian random variables with a general correlation matrix ρ . Its copula is then the CDF of $(U_1, \dots, U_n) = (N(N_1), \dots, N(N_n))$. Taking $1 \leq i_1 < i_2 \leq n$, we see that $(N_{i_1}, N_{i_2}) \stackrel{\text{law}}{=} (N_{i_2}, N_{i_1})$ since this law is characterized by $\mathbb{E}(N_{i_1}N_{i_2})$, and so $(U_{i_1}, U_{i_2}) \stackrel{\text{law}}{=} (U_{i_2}, U_{i_1})$. Therefore, a Gaussian copula is 2-exchangeable. In the two-dimensional case we resort directly to the term “exchangeable” rather than “2-exchangeable”. When dealing with defaultable bonds as in Jouanin and al. [45], n -exchangeable copulas cannot model situations where the dependence is asymmetric and based on the assets themselves. With 2-exchangeable copulas (such as for example Gaussian or Archimedean copulas), we cannot model asymmetric relations featuring a first entity that influences a second one more than the latter influences the former.

In recent years, copula functions have received a great deal of attention, see for example the papers of Genz and Bretz [35], Hürlimann [41], Juri and Wüthrich [46], Wei and Hu [65], and the books of Joe [44] and Nelsen [57]. For financial and insurance applications, recent applications on copulas include for example Bouyé et al. [17], Cherubini et al. [24], Embrechts et al. [32], Jouanin et al. [45], Klugman and Parsa [49], Schönbucher and Schubert [62]. In this paper, we build new families of copulas based on the first approach, using periodic functions following Alfonsi [1]. We first begin to work in the two-dimensional case, obtaining a one-parameter copula, and then give a way to extend the result to the n -dimensional case with $n > 2$, getting a family with $n - 1$ parameters. Finally, we explain how such copulas can be simulated.

3.2 Periodic Copulas: The Bivariate Basic Case

We begin by defining our new copula functions for bivariate dependence, i.e. for possible dependence structures between two random variables. It is helpful to first recall three particular “limit” copulas. The “middle” one, typically denoted by $C^\perp$, is the copula obtained with two independent uniform variables on $[0, 1]$: $C^\perp(u_1, u_2) = u_1 u_2$. The two other ones, denoted by C_F^- and C_F^+ respectively, are the two Frechet bounds of the convex subset of copulas: $C_F^+(u_1, u_2) = \min(u_1, u_2)$ and $C_F^-(u_1, u_2) = (u_1 + u_2 - 1)^+$, where $x^+ = \max(x, 0)$ denotes the positive part operator. Naming U a uniform random variable on $[0, 1]$, C_F^+ can be obtained as the copula of (U, U) and corresponds obviously to perfect positive dependence, whereas C_F^- is obtained as the copula of $(U, 1 - U)$ and describes total negative dependence.

We will say in the following that a copula admits a *density* when $\frac{\partial^2 C}{\partial u_1 \partial u_2} = c(u_1, u_2)$ exists in the ordinary sense. In this paper we propose copulas that have a density that can

be written in the form

$$\forall (u_1, u_2) \in [0, 1]^2, \quad c(u_1, u_2) = \tilde{c}(u_1 + u_2) \quad (\text{resp. } c(u_1, u_2) = \tilde{c}(u_1 - u_2))$$

for a suitable function $\tilde{c} : \mathbb{R} \rightarrow \mathbb{R}$. To satisfy properties i), ii) and iii), $\tilde{c}$ must be nonnegative and verify for $u_1, u_2 \in [0, 1]$:

$$\int_0^{u_1} \int_0^1 \tilde{c}(x_1 \pm x_2) dx_1 dx_2 = u_1 \quad \text{and} \quad \int_0^1 \int_0^{u_2} \tilde{c}(x_1 \pm x_2) dx_1 dx_2 = u_2.$$

Differentiating with respect to u_1 and u_2 respectively, we obtain $\int_0^1 \tilde{c}(u_1 \pm x_2) dx_2 = 1$ and $\int_0^1 \tilde{c}(x_1 \pm u_2) dx_1 = 1$. Differentiating further, since $\int_0^1 \tilde{c}(u_1 + x_2) dx_2 = \int_{u_1}^{u_1+1} \tilde{c}(x_2) dx_2$ (resp. $\int_0^1 \tilde{c}(u_1 - x_2) dx_2 = \int_{u_1-1}^{u_1} \tilde{c}(x_2) dx_2$), we obtain:

$$\tilde{c}(u_1 + 1) = \tilde{c}(u_1) \quad \forall u_1 \in [0, 1], \quad (\text{resp. } \tilde{c}(u_1 - 1) = \tilde{c}(u_1) \quad \forall u_1 \in [0, 1]).$$

Thus, a consequence of requiring $c(u_1, u_2) := \tilde{c}(u_1 \pm u_2)$ to be the density of a copula is that $\tilde{c}$ has to be 1-periodic (at least on $[-1, 2]$, but its value outside this interval is irrelevant) and that $\int_0^1 \tilde{c}(u) du = 1$. Conversely, it is easy to see that if $\tilde{c}$ is a nonnegative 1-periodic function such that $\int_0^1 \tilde{c}(u) du = 1$, then

$$\tilde{C}^-(u_1, u_2) := \int_0^{u_1} \int_0^{u_2} \tilde{c}(x_1 + x_2) dx_1 dx_2 \quad (\text{resp. } \tilde{C}^+(u_1, u_2) := \int_0^{u_1} \int_0^{u_2} \tilde{c}(x_1 - x_2) dx_1 dx_2)$$

satisfies conditions i), ii) and iii), and so is a copula function that we call, with a slight abuse of language, *periodic copula*. We note here that copulas obtained with these densities form a convex set since a convex combination of 1-periodic nonnegative functions satisfying $\int_0^1 \tilde{c}(u) du = 1$ is also a 1-periodic nonnegative function with integral 1 on a period. The following proposition will be useful for calculations:

Proposition 3.2.1. *Within the above setting, let us denote by φ the primitive of $\tilde{c}$ that vanishes at 0, and $\Phi(x) := \int_0^x \varphi(u) du$, so that Φ is a double primitive of $\tilde{c}$. The above periodic copulas can be written:*

$$\begin{aligned} \tilde{C}^-(u_1, u_2) &= \int_0^{u_1} \int_0^{u_2} \tilde{c}(x_1 + x_2) dx_1 dx_2 = \Phi(u_1 + u_2) - \Phi(u_1) - \Phi(u_2), \\ \tilde{C}^+(u_1, u_2) &= \int_0^{u_1} \int_0^{u_2} \tilde{c}(x_1 - x_2) dx_1 dx_2 = \Phi(u_1) + \Phi(-u_2) - \Phi(u_1 - u_2) \end{aligned} \quad (3.2)$$

Hence, $\tilde{C}^-$ is always exchangeable while $\tilde{C}^+$ is exchangeable if, and only if Φ is even.

Using the 1-periodicity, we have also $\tilde{C}^-(u_1, u_2) = \Phi(u_1 + u_2 - 1) - \Phi(u_1 - 1) - \Phi(u_2 - 1) + \Phi(-1)$, and it is then sufficient to know Φ on $[-1, 1]$.

The construction above leads to copula densities that have remarkable Schur properties. Let us recall that the *majorization order* $\prec$ (see Marshall and Olkin [53]) is defined in the bivariate case by:

$$(x_1, x_2) \prec (x'_1, x'_2) \text{ if } x_1 + x_2 = x'_1 + x'_2 \text{ and } \max(x_1, x_2) \leq \max(x'_1, x'_2).$$

A scalar function $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ is said to be Schur-convex (resp. Schur-concave) if it is nondecreasing (resp. nonincreasing) respect to this order, and f is Schur-constant when it is both Schur-convex and Schur-concave. Through this framework, we have shown that if a copula has a Schur-constant density $\tilde{c}(u_1 + u_2)$ for $(u_1, u_2) \in [0, 1]^2$, $\tilde{c}$ is 1-periodic.

Typically, Schur properties are used for the survival functions of lifetimes, especially to model aging problems. Indeed, the Schur properties of survival function are linked to the main notions of aging such as Increasing/Decreasing Failure Rates and New Better/Worse than Used problems (see Spizzichino [63]). Unfortunately, there are no straightforward correspondences between the Schur property of a survival function and the Schur property of the density of its copula. More specifically, let us consider two lifetimes τ_1 and τ_2 having the same law of density f and CDF F . We denote by C the copula of (τ_1, τ_2) and we assume it to have a density $c(u_1, u_2) = \frac{\partial^2 C}{\partial u_1 \partial u_2}(u_1, u_2)$ for $(u_1, u_2) \in [0, 1]^2$. We also denote by S the survival function $S(t_1, t_2) = \mathbb{P}(\tau_1 > t_1, \tau_2 > t_2)$ and let the joint density be $p(t_1, t_2) = \frac{\partial^2 S}{\partial t_1 \partial t_2}(t_1, t_2)$ for $t_1, t_2 \geq 0$. A sufficient condition (Theorem 4.30 in Spizzichino [63]) to get the Schur concavity (resp. the Schur convexity) of the survival function S is requiring p to be Schur-concave (resp. Schur-convex). Since the expression linking p with c is

$$\forall t_1, t_2 \geq 0, \quad p(t_1, t_2) = f(t_1)f(t_2)c(F(t_1), F(t_2)),$$

we see that there is no immediate relationship between the Schur properties of p and c , due to the nontrivial presence on the marginal laws of the lifetimes (through f and F).

3.3 A Bivariate Smooth Periodic Copulas Family Attaining the Frechet Bounds and Independence

3.3.1 A first intuitive (non-smooth) construction

A first example of 1-periodic function which arises naturally is $\tilde{c}(x) = 1 + \sin(2\pi x + a)$ where a is a parameter that we can take in $[0, 2\pi)$. This method provides families that cannot model strong positive or negative dependence, since they approach neither C_F^- nor C_F^+ . *On the contrary, it might be interesting to have a family of copulas which attains the copulas C_F^+ , $C^\perp$ and C_F^- as limit cases in order to be able to describe a large range of*

dependence structures. By expressing copulas by means of $\frac{\partial^2 C}{\partial u_1 \partial u_2}$, attaining C_F^+ , C_F^- and $C^\perp$ amounts to attaining

$$\mu^+ = \delta_{x_1}(dx_2) \otimes dx_1, \quad \mu^- = \delta_{1-x_1}(dx_2) \otimes dx_1, \quad c^\perp = 1$$

with the copula density. Since $\frac{\partial^2 C^\perp}{\partial u_1 \partial u_2}$ exists in the ordinary sense of differentiation, it has a density $c^\perp = 1$ that corresponds to the Lebesgue measure on the square $[0, 1]^2$, $\mu^\perp = dx_1 \otimes dx_2$. Instead, μ^+ and μ^- are to be interpreted in the generalized sense. Moreover, μ^+ and μ^- charge only the diagonals of the square $[0, 1]^2$, i.e. $\Delta^+ = \{(x, x), x \in [0, 1]\}$ and $\Delta^- = \{(x, 1-x), x \in [0, 1]\}$ respectively. The idea is then to find a family of periodic functions $\tilde{c}_\gamma$ indexed by a parameter γ and such that the density $\tilde{c}_\gamma(x_1 - x_2)$ (resp. $\tilde{c}_\gamma(x_1 + x_2)$) concentrates on Δ^+ (resp. Δ^-) for some values of γ . Thus, if we define the piecewise 1-periodic function $\tilde{c}_\gamma$ for $0 < \gamma \leq \frac{1}{2}$ by

$$\tilde{c}_\gamma(x) := \frac{1}{2\gamma} (\mathbf{1}_{[0, \gamma]}(x) + \mathbf{1}_{(1-\gamma, 1)}(x)) \quad \text{for } x \in [0, 1], \quad (3.3)$$

we see that the family of densities defined as $c_\gamma^+(x_1, x_2) := \tilde{c}_\gamma(x_1 - x_2)$ verifies:

$$c_{1/2}^+ = c^\perp, \quad c_\gamma^+(x_1, x_2) dx_1 dx_2 \xrightarrow[\gamma \rightarrow 0]{\mathcal{D}} \mu^+$$

$\xrightarrow{\mathcal{D}}$ denoting convergence in distribution. The corresponding convergence in law for random variables is denoted by $\mathcal{L}$. To calculate the associated copula $C_\gamma^+ := \tilde{C}^+$, it is best to try a drawing and see that its value in (u_1, u_2) is the area in the intersection of the rectangle delimited by $(0, 0)$ and (u_1, u_2) with $\{(x_1, x_2) \in [0, 1]^2, -\gamma \leq x_1 - x_2 \leq \gamma \text{ or } x_1 - x_2 \leq \gamma - 1 \text{ or } x_1 - x_2 \geq 1 - \gamma\}$. We obtain, for $u_1 \leq u_2$ which is not restrictive since $C_\gamma^+(u_1, u_2) = C_\gamma^+(u_2, u_1)$,

$$C_\gamma^+(u_1, u_2) = \frac{1}{2\gamma} [u_1 u_2 + \frac{1}{2} [-((u_2 - u_1 - \gamma)^+ + (u_2 - \gamma)^+)(\min(u_1, u_2 - \gamma))^+ - ((u_1 - \gamma)^+)^2 + ((u_2 - 1 + \gamma)^+ + (u_2 - 1 + \gamma - u_1)^+) \cdot (\min(u_1, u_2 - 1 + \gamma))^+ + ((u_1 - 1 + \gamma)^+)^2]] \quad (3.4)$$

In order to obtain a family that reaches C_F^- instead, we use the other family, precisely $c_\gamma^-(x_1, x_2) = \tilde{c}_\gamma(x_1 + x_2)$. We have $c_{1/2}^- = c^\perp$ and $c_\gamma^-(x_1, x_2) dx_1 dx_2 \xrightarrow[\gamma \rightarrow 0]{\mathcal{D}} \mu^-$. Fortunately, a simple geometric remark links the related copula C_γ^- to C_γ^+ , and gives $C_\gamma^-(u_1, u_2) = u_2 - C_\gamma^+(1 - u_1, u_2)$.

Through this method, we have obtained a family of copula quite “exhaustive”, going from C_F^- to C_F^+ and taking the in-between value $C^\perp$. Incidentally, we see why we chose to name $\tilde{C}^+$ the copula coming from $\tilde{c}(x_1 - x_2)$ and $\tilde{C}^-$ the copula coming from $\tilde{c}(x_1 + x_2)$: this is done because in our case the former attains C_F^+ and the latter C_F^- .

At times it can be handy to have a single number measuring some stylized aspects of a given copula. The Spearman’s rho is such a number and is a well known measure of concordance, see for example Embrechts et al. [32]. When defined in terms of copula functions, it is

given by the following integral in the copula density c : $\rho := 12 \int_0^1 \int_0^1 u_1 u_2 c(u_1, u_2) du_1 du_2 - 3$. We obtain for the C_γ^+ and C_γ^- copulas respectively,

$$\rho_\gamma^+ = (2\gamma - 1)(\gamma - 1), \quad \rho_\gamma^- = (1 - 2\gamma)(\gamma - 1).$$

An interesting remark concerns the construction of non-exchangeable (non-symmetric) copulas ($C(u_1, u_2) \neq C(u_2, u_1)$) through this method. This can be relevant for modelling two life durations in case of symbiosis or parasitism. As a different application, we may consider expressing the default dependence between two firms with asymmetric relations in credit risk. One may have a first firm depending more on a second one than the latter depends on the former. This could be the case of a little firm that mainly provides goods to a large one (symbiosis). A default of the large company could induce a dramatic effect on the smaller one, whereas a default of the small firm could have little relevance to the large one. Let us name τ_1 the default time of the large company and τ_2 the default time of the small company, and assume moreover to clarify our statement that τ_1 and τ_2 have the same law. We would like to have $\mathbb{P}(\tau_1 \leq \tau_2) < \mathbb{P}(\tau_2 \leq \tau_1)$, reflecting the fact that the large company has more chances to survive the default of the small firm than the small firm has of surviving the default of the large one. This is impossible with 2-exchangeable copulas such as Archimedean and Gaussian copulas, because they imply the symmetric relation $(\tau_1, \tau_2) \stackrel{\text{law}}{=} (\tau_2, \tau_1)$ and in particular $\mathbb{P}(\tau_1 \leq \tau_2) = \mathbb{P}(\tau_2 \leq \tau_1)$. In order to provide an example of non-exchangeable copula obtained from our family, we see from (3.2) that our only chance is to select a periodic function $\tilde{c}$ whose double primitive Φ is not even and then take the related $\tilde{C}^+$. The simplest such function is $\tilde{c} := \bar{c}_\gamma$ defined, for $\gamma \in [0, 1]$, as

$$\bar{c}_\gamma(x) := (1/\gamma) \mathbf{1}_{[0, \gamma]}(x), \quad \text{for } x \in [0, 1]. \quad (3.5)$$

We thus obtain the copulas $\bar{C}_\gamma^+$ and $\bar{C}_\gamma^-$ that have respectively the densities $\bar{c}_\gamma(x_1 - x_2)$ and $\bar{c}_\gamma(x_1 + x_2)$, and observing that $\bar{c}_\gamma(x) = \tilde{c}_{\frac{\gamma}{2}}(x - \frac{\gamma}{2})$ we get:

$$\begin{aligned} \bar{C}_\gamma^+(u_1, u_2) &= C_{\gamma/2}^+(u_1, \min(u_2 + \gamma/2, 1)) - C_{\gamma/2}^+(u_1, \gamma/2) + C_{\gamma/2}^+(u_1, (u_2 + \gamma/2 - 1)^+) \\ \bar{C}_\gamma^-(u_1, u_2) &= C_{\gamma/2}^-(u_1, (u_2 - \gamma/2)^+) + C_{\gamma/2}^-(u_1, 1 - (\gamma/2 - u_2)^+) - C_{\gamma/2}^-(u_1, 1 - \gamma/2). \end{aligned}$$

With the asymmetric periodic copula $\bar{C}_\gamma^+$ we still have good asymptotic properties, in that $\bar{C}_1^+ = C^\perp$ and $\bar{C}_\gamma^+ \rightarrow C_F^+$ when $\gamma \rightarrow 0$. $\bar{C}_\gamma^-$, which is symmetric, satisfies analogous properties. Calculating the Spearman's rho, we find again

$\rho(\bar{C}_\gamma^+) = (2\gamma - 1)(\gamma - 1)$ and $\rho(\bar{C}_\gamma^-) = -\rho(\bar{C}_\gamma^+)$ for $\gamma \in [0, 1]$. Let us remark that $\rho(\bar{C}_\gamma^+)$ takes negative value between $\gamma = 1/2$ and $\gamma = 1$, and vanishes at $1/2$ for a copula different from $C^\perp$. Thus, if we wish to describe a negative asymmetric dependence, it is best to use C_γ^+ with $1/2 < \gamma < 1$.

Another interesting synthetic quantity concerning copulas is the upper-tail dependence which is defined as $\lambda := \lim_{u \rightarrow 0} \frac{1}{u} \int_{1-u}^1 \int_{1-u}^1 c(x_1, x_2) dx_1 dx_2$ when the copula has a density c . We

can easily show that periodic copulas have no upper-tail dependence (i.e. $\lambda = 0$). However, if one wishes to obtain a copula with an upper-tail dependence equal to $\lambda > 0$, it is still possible to consider the convex combination $(1 - \lambda)C + \lambda C_F^+$ where C is a preferred periodic copula. This combination can be simulated easily when one knows how to simulate the basic C , as we do for the periodic copulas (with invertible φ , i.e. with a strictly positive periodic function $\tilde{c}$) we introduced here.

3.3.2 Regularization

A drawback of the families C_γ^+ and C_γ^- is that these copulas are constant on some intervals and comes from the 0-1 nature of the density, and more precisely from the existence of a domain (with positive measure) where the density vanishes. This causes problems, especially when in need of simulating the copula. In order to avoid this drawback, the idea is then to replace γ with a random variable Γ and then take the expectation of $\tilde{c}_\Gamma$, using the convexity of the subset of the periodic copulas. Indeed, if $\Gamma \sim p$ where p is the density of a probability measure on $[0, 1/2]$ such that $p(\gamma) > 0 \forall \gamma$, then $\hat{c}_p(x) := E[\tilde{c}_\Gamma(x)] = \int_0^{\frac{1}{2}} \tilde{c}_\gamma(x) p(\gamma) d\gamma$ is a positive 1-periodic function. Thus, if we have a family of random variables $(\Gamma_\alpha)_{\alpha \geq 0}$ concentrated on $[0, 1/2]$ with densities $\{p_\alpha, \alpha \in]0, +\infty[$ on $[0, 1/2]$ and such that $\Gamma_\alpha \xrightarrow[\alpha \rightarrow 0]{\mathcal{L}} 1/2$ and $\Gamma_\alpha \xrightarrow[\alpha \rightarrow \infty]{\mathcal{L}} 0$, we can define

$$\hat{C}_\alpha^+ := E[C_{\Gamma_\alpha}^+] = \int_0^{\frac{1}{2}} C_\gamma^+ p_\alpha(\gamma) d\gamma, \quad \hat{C}_\alpha^- := E[C_{\Gamma_\alpha}^-] = \int_0^{\frac{1}{2}} C_\gamma^- p_\alpha(\gamma) d\gamma \quad (3.6)$$

(that correspond respectively to the periodic densities $\tilde{c}(x_1 - x_2) = \hat{c}_{p_\alpha}(x_1 - x_2)$ and $\tilde{c}(x_1 + x_2) = \hat{c}_{p_\alpha}(x_1 + x_2)$). We have obtained a family of copulas with good asymptotic properties, in that $\hat{C}_\alpha^+ \xrightarrow[\alpha \rightarrow 0]{} C^\perp$, $\hat{C}_\alpha^+ \xrightarrow[\alpha \rightarrow +\infty]{} C_F^+$, $\hat{C}_\alpha^- \xrightarrow[\alpha \rightarrow 0]{} C^\perp$, and $\hat{C}_\alpha^- \xrightarrow[\alpha \rightarrow +\infty]{} C_F^-$.

We can build easily a random variable with suitable density on $[0, \frac{1}{2}]$ by transforming a uniform variable U on $[0, 1]$ according to a homeomorphism. Indeed, consider $\Gamma_\alpha := \frac{1}{2}U^\alpha$, $\alpha \in]0, +\infty[$, so that we get a family of densities p_α on $[0, \frac{1}{2}]$ that feature the desired asymptotic properties in 0 and $+\infty$: $p_\alpha(u) = (2^{\frac{1}{\alpha}}/\alpha)u^{\frac{1-\alpha}{\alpha}}$.

The calculation of $\hat{C}_\alpha^+$ does not present difficulties either. We first calculate the periodic function $\hat{c}_\alpha := \hat{c}_{p_\alpha}$, obtaining

$$\hat{c}_\alpha(x) = \frac{1}{1-\alpha} [1 - (2x)^{\frac{1-\alpha}{\alpha}}], \quad \alpha \neq 1, \quad \hat{c}_1(x) = -\ln(2x)$$

for $0 \leq x \leq 1/2$, and $\hat{c}_\alpha(x) = \hat{c}_\alpha(1-x)$ for $1/2 \leq x \leq 1$, since the same property holds for the basic $\tilde{c}_\gamma$'s. Let us compute the primitive ψ^α of $\hat{c}_\alpha$ that vanishes at $x = 0$. We obtain, for $0 \leq x \leq 1/2$:

$$\psi_\alpha(x) = \frac{1}{2(1-\alpha)} [2x - \alpha(2x)^{\frac{1}{\alpha}}], \quad \alpha \neq 1, \quad \psi_1(x) = x - x \ln(2x).$$

Using the symmetry property $\widehat{c}_\alpha(x) = \widehat{c}_\alpha(1-x)$ we obtain, for $1/2 \leq x \leq 1$, $\psi_\alpha(x) = \psi_\alpha(1/2) + (\psi_\alpha(1/2) - \psi_\alpha(1-x)) = 1 - \psi_\alpha(1-x)$, since $\psi_\alpha(1/2) = 1/2$. Instead, for $x \in [-1, 0]$ we use the periodicity of $\widehat{c}$ to get $\psi_\alpha(x) = -\psi_\alpha(-x)$. To proceed further, we need to know also the primitive Ψ_α of ψ_α , i.e. the double primitive of $\widehat{c}_\alpha$. We find:

$$\begin{aligned} x \in [0, \frac{1}{2}], \quad \Psi_\alpha(x) &= \frac{1}{2(1-\alpha)} \left[x^2 - \frac{\alpha^2}{\alpha+1} 2^{\frac{1}{\alpha}} x^{\frac{1+\alpha}{\alpha}} \right], \quad \alpha \neq 1, \quad \Psi_1(x) = \frac{3}{4}x^2 - \frac{x^2}{2} \ln(2x), \\ x \in [\frac{1}{2}, 1], \quad \Psi_\alpha(x) &= x - \frac{1}{2} + \Psi_\alpha(1-x) \end{aligned}$$

and $\Psi_\alpha(x) = \Psi_\alpha(-x)$ for $x \in [-1, 0]$, since ψ_α is an odd function. The copulas $\widehat{C}_\alpha^+$ and $\widehat{C}_\alpha^-$ are then given by the Proposition (3.2.1). We can also easily calculate the Spearman's rho of $\widehat{C}_\alpha^+$, since $\rho(\widehat{C}_\alpha^+) = \int_0^{1/2} (2\gamma - 1)(\gamma - 1) 2^{1/\alpha} \gamma^{(1-\alpha)/\alpha} / \alpha d\gamma$ so that $\rho(\widehat{C}_\alpha^+) = 1 - \frac{3}{2(1+\alpha)} + \frac{1}{2(1+2\alpha)}$ and we have $\rho(\widehat{C}_\alpha^-) = -\rho(\widehat{C}_\alpha^+)$ (this is a general relation between the rho of the periodic copulas with density $\tilde{c}(x_1 + x_2)$ and $\tilde{c}(x_1 - x_2)$). The families built previously are exchangeable, since the related $\tilde{c}$ are expectations of functions leading to even double primitives and therefore lead themselves to even double primitives, so that (3.2) yields symmetry.

However, we can also construct a smooth family of non symmetric copulas by defining $\tilde{c}$ as the expectation of the previous non symmetric function (3.5) with a random γ . Indeed, if $Z \sim q$ where q is a positive density of a probability measure on $[0, 1]$, then $\bar{c}_q := E[\bar{c}_Z] = \int_0^1 \bar{c}_\gamma(x) q(\gamma) d\gamma$ is a positive 1-periodic function. Thus, as before, if we have a family of random variables $(Z_\alpha)_{\alpha \geq 0}$ on $[0, 1]$ with densities $\{q_\alpha, \alpha \in]0, +\infty[\}$ on $[0, 1]$ and such that $Z_\alpha \xrightarrow[\alpha \rightarrow 0]{\mathcal{L}} 1$ and $Z_\alpha \xrightarrow[\alpha \rightarrow \infty]{\mathcal{L}} 0$, we can define

$$\widehat{C}_\alpha^+ := E[\bar{C}_{Z_\alpha}^+] = \int_0^1 \bar{C}_\gamma^+ q_\alpha(\gamma) d\gamma, \quad \widehat{C}_\alpha^- := E[\bar{C}_{Z_\alpha}^-] = \int_0^1 \bar{C}_\gamma^- q_\alpha(\gamma) d\gamma \quad (3.7)$$

that correspond respectively to the periodic densities $\tilde{c}(x_1 - x_2) = \bar{c}_{q_\alpha}(x_1 - x_2)$ and $\tilde{c}(x_1 + x_2) = \bar{c}_{q_\alpha}(x_1 + x_2)$, and have good asymptotic properties.

We take $Z_\alpha := U^\alpha$, where U is a uniform random variable on $[0, 1]$. Its density is $q_\alpha(u) = \frac{1}{\alpha} u^{\frac{1-\alpha}{\alpha}}$, so that the associated periodic function $\tilde{c} = \bar{c}_\alpha$ is given in $[0, 1]$ by

$$\widehat{c}_\alpha(x) := \bar{c}_{q_\alpha} = \frac{1}{1-\alpha} [1 - x^{\frac{1-\alpha}{\alpha}}], \quad \alpha \neq 1, \quad \widehat{c}_1(x) := \bar{c}_{q_1} = -\ln(x).$$

In order to find an expression for the copula, we compute the primitive g_α of $\widehat{c}_\alpha$ that vanishes at $x = 0$. We have, for $x \in [0, 1]$:

$$g_\alpha(x) = \frac{1}{1-\alpha} [x - \alpha x^{\frac{1}{\alpha}}], \quad \alpha \neq 1, \quad g_1(x) = x - x \ln(x),$$

and for $x \in [-1, 0]$ we have $g_\alpha(x) = g_\alpha(1+x) - 1$. Denote by G_α the primitive of g_α vanishing at 0. For $x \in [0, 1]$ G_α is given by

$$G_\alpha(x) = \frac{1}{1-\alpha} \left[\frac{1}{2}x^2 - \frac{\alpha^2}{1+\alpha}x^{\frac{1+\alpha}{\alpha}} \right], \quad \alpha \neq 1, \quad G_1(x) = \frac{3}{4}x^2 - \frac{1}{2}x^2 \ln(x),$$

whereas for $x \in [-1, 0]$ we have $G_\alpha(x) = G_\alpha(1+x) - x - G_\alpha(1)$. By Fubini's theorem, the copulas defined by (3.7) are the same as the copulas associated with the periodic functions $\tilde{c} = \widehat{c}_\alpha$ and defined by (3.2). The Proposition (3.2.1) gives then $\widehat{C}_\alpha^+(u_1, u_2) = G_\alpha(u_1) + G_\alpha(-u_2) - G_\alpha(u_1 - u_2)$ and $\widehat{C}_\alpha^-(u_1, u_2) = G_\alpha(u_1 + u_2 - 1) - G_\alpha(u_2 - 1) - G_\alpha(u_1 - 1) + G_\alpha(-1)$. Their Spearman rho's are $\rho(\widehat{C}_\alpha^+) = 1 - 3/(1+\alpha) + 2/(1+2\alpha) = \frac{\alpha(2\alpha-1)}{(1+\alpha)(1+2\alpha)}$ (negative for $\alpha \in (0, 1/2)$) and $\rho(\widehat{C}_\alpha^-) = -\rho(\widehat{C}_\alpha^+)$.

3.4 Beyond the Bivariate Case

There are several ways to extend the previous construction to build a copula in dimension $n > 2$. In dimension $n > 2$, we see that if $\tilde{c}$ is a 1-periodic function such that $\int_0^1 \tilde{c}(x)dx = 1$, then $\tilde{c}(\sum_{i=1}^n \varepsilon_i x_i)$ are densities of copulas when $\varepsilon_i \in \{-1, 1\}$. However, these copulas cannot be obtained directly in terms of bivariate copulas and require new calculations. To keep the analysis simple, we work instead with the densities already defined for the bivariate case. Consider the following proposition.

Proposition 3.4.1. *Assume that $c_1, \dots, c_{n-1}$ are densities of two dimensional copulas built through periodic densities $\tilde{c}_1, \dots, \tilde{c}_{n-1}$, i.e. $c_j(x, y) = \tilde{c}_j(x + \varepsilon_j y)$ with $\varepsilon_j \in \{-1, 1\}$ and $\tilde{c}_j$ a nonnegative 1-periodic function with unit integral on a period. Set $\check{c} := (c_1, \dots, c_{n-1})$. Then the following two functions are copulas:*

$$\begin{aligned} \check{C}_1(u_1, \dots, u_n) &:= \int_0^{u_1} \dots \int_0^{u_n} c_1(x_1, x_2) c_2(x_2, x_3) \dots c_{n-1}(x_{n-1}, x_n) dx_1 \dots dx_n \\ \check{C}_2(u_1, \dots, u_n) &:= \int_0^{u_1} \dots \int_0^{u_n} c_1(x_1, x_2) c_2(x_1, x_3) \dots c_{n-1}(x_1, x_n) dx_1 \dots dx_n. \end{aligned}$$

The proof is quite immediate. Properties 1 and 3 in Section 3.1 are satisfied by construction. It remains to observe that $\check{C}(1, \dots, 1, u_k, 1, \dots, 1) = \int_0^{u_k} dx_k = u_k$, by using Fubini's theorem, integrating first with respect to the x_i 's with $i \neq k$, and using then the property $\int_x^{1+x} \tilde{c}(u)du = 1$. The first copula $\check{C}_1$ is convenient if we wish to express the n -dependence in terms of dependences of two consecutive variables, whereas the second one $\check{C}_2$ allows us to express the n -dependence in terms of the dependence of a preferred variable with all other variables. The second method could be referred to as a “preferred-” or “main-factor” approach.

3.5 The Simulation of Periodic Copulas

Let us begin by recalling how to simulate a copula that admits a density $p(x_1, \dots, x_n)$. We need simulating a vector of uniform variables $(U_1, \dots, U_n)$ that has $C(u_1, \dots, u_n) = \int_0^{u_1} \dots \int_0^{u_n} p(x_1, \dots, x_n) dx_1 \dots dx_n$ as joint CDF. This can be done according to the following steps: (i) To simulate the first variable U_1 , it suffices to sample from a uniform random variable $\tilde{U}_1$ in $[0, 1]$. Let us call u_1 the simulated sample; (ii) To simulate U_k consistently with the earlier sampled $u_1, \dots, u_{k-1}$, we need the law of U_k conditional on $U_i = u_i$ for $i < k$. Denoting as usual by $F_k(\cdot | u_1, \dots, u_{k-1})$ the CDF of this law,

$$\begin{aligned} F_k(u_k | u_1, \dots, u_{k-1}) &= \mathbb{P}(U_k \leq u_k | U_1 = u_1, \dots, U_{k-1} = u_{k-1}) \\ &= \frac{\partial_{u_1, \dots, u_{k-1}} C(u_1, \dots, u_k, 1, \dots, 1)}{\partial_{u_1, \dots, u_{k-1}} C(u_1, \dots, u_{k-1}, 1, \dots, 1)} = \frac{\int_0^{u_k} \int_0^1 \dots \int_0^1 p(u_1, \dots, u_{k-1}, x_k, \dots, x_n) dx_k dx_{k+1} \dots dx_n}{\int_0^1 \int_0^1 \dots \int_0^1 p(u_1, \dots, u_{k-1}, x_k, \dots, x_n) dx_k dx_{k+1} \dots dx_n} \end{aligned}$$

and we can take $U_k = F_k^{-1}(\tilde{U}_k | u_1, \dots, u_{k-1})$ where $\tilde{U}_k$ is a uniform- $[0, 1]$ variable independent of $(\tilde{U}_1, \dots, \tilde{U}_{k-1})$.

In the case of the periodic copulas $\check{C}_{1,2}$, with the notation of Proposition 3.4.1, we have respectively $F_k^1(u_k | u_1, \dots, u_{k-1}) = \int_0^{u_k} c_{k-1}(u_{k-1}, x_k) dx_k$ and $F_k^2(u_k | u_1, \dots, u_{k-1}) = \int_0^{u_k} c_{k-1}(u_1, x_k) dx_k$, where the upper index refers to the copula we are considering. Taking the smooth families of the previous section, these F functions can be expressed in terms of ψ_α and g_α (for example $\int_0^{u_k} \hat{c}_\alpha(u \pm x) dx = \pm(\psi_\alpha(u \pm u_k) - \psi_\alpha(u))$). Moreover they are strictly increasing, and can therefore be inverted easily numerically. We note here that if we choose the “non smooth” copulas $C^\pm$ and $\overline{C}^\pm$ as two-dimensional building blocks, this inversion is not feasible since the densities vanish on some intervals. By choosing instead $\hat{C}_\alpha^\pm$ and $\hat{\overline{C}}_\alpha^\pm$ we have obtained families of n -dimensional copulas essentially characterized by $n - 1$ parameters α_i plus the flags sgn_i, sym_i , for $i = 1, \dots, n - 1$, where sym_i is set according to whether we take a symmetric family or not (symbolized here by the $\overline{}$), and where sgn_i is taken from the set $\{-, +\}$.

3.6 Conclusions

The new family of “periodic” copulas introduced in this paper offers a possibility for practically manageable and possibly asymmetric copulas. We have studied in detail the two-dimensional case, based on a single dependence parameter, also in connection with classical dependence notions such as upper tail dependence, Spearman’s rho and Frechet bounds for Copula functions. We then provide a means to construct an n -dimensional copula building on the two-dimensional case. We obtain families of copulas in dimension n and parameterized by $n - 1$ parameters, implying possibly asymmetric relations. We explain how periodic copulas can be simulated.

Part II

Schémas de discrétisations pour le processus de Cox-Ingersoll-Ross

Chapter 4

Discrétisation du processus CIR

Ce chapitre est un article [2] publié dans la revue *Monte Carlo Methods and Applications*.

Abstract. In this paper, we focus on the simulation of the CIR processes and present several discretization schemes of both the implicit and explicit types. We study their strong and weak convergence. We also examine numerically their behaviour and compare them to the schemes already proposed by Deelstra and Delbaen [27] and Diop [28]. Finally, we gather all the results obtained and recommend, in the standard case, the use of one of our explicit schemes.

4.1 Introduction

The aim of this paper is to present an overview on the discretization schemes that can be used for the simulation of the square-root diffusions of Cox-Ingersoll-Ross type. These processes, initially introduced to model the short interest rate (Cox, Ingersoll and Ross [25]), are now widely used in modelling because they present interesting features like the nonnegativity and the mean reversion. Moreover, some standard expectations can be analytically calculated which can be useful especially for calibrating the parameters. Thus, they have also been used in finance to model the stochastic volatility of the stock price (Heston [39]) or the credit spread (Brigo and Alfonsi [18]). We will use in this paper the following notation for this diffusion: (X_t) will denote a Cox-Ingersoll-Ross (CIR for short) process of parameter (k, a, σ, x_0) if

$$\begin{cases} X_t = x_0 + \int_0^t (a - kX_s)ds + \sigma \int_0^t \sqrt{X_s}dW_s, t \in [0, T] \\ x_0, \sigma, a \geq 0, k \in \mathbb{R}. \end{cases} \quad (4.1)$$

Under the above assumption on the parameters that we will suppose valid through all the paper, it is well known that this SDE has a nonnegative solution, and this solution is

pathwise unique (see for example Rogers and Williams [58]). Let us recall here that under the assumption (see for example Lamberton and Lapeyre [50])

$$2a > \sigma^2 \text{ and } x_0 > 0 \quad (4.2)$$

the process is always positive.

When $k > 0$, it is common to define $\theta = a/k$ and rewrite the SDE $dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t$. Indeed, θ appears as the asymptotic mean of X_t toward which the process is attracted. In practice, this more intuitive parametrization is preferred.

In the sequel, $(\mathcal{F}_t, t \geq 0)$ will denote the natural filtration of the Brownian motion W , and we will consider the regular grid $t_i^n = \frac{iT}{n}$. Except in cases where it is important to remind the dependency in n , we will write t_i rather than t_i^n . It is well known that the increments of the CIR process are non-central chi-squared random variables that can be simulated exactly. Thus, we can inductively simulate a random vector distributed according to the law of $(X_{t_0}, \dots, X_{t_n})$ (see Glasserman [37], pp. 120-134). However, the exact simulation in general requires more time than a simulation with approximation schemes. It may also be restrictive if one wishes to correlate this diffusion with another diffusion via the Brownian motions as in Brigo and Alfonsi [18] where two correlated CIR processes are considered. At least for both these reasons, studying approximation schemes is relevant.

It is important to remark first that the natural way to simulate this process, that is the explicit Euler-Maruyama scheme

$$\hat{X}_{t_{i+1}}^n = \hat{X}_{t_i}^n + \frac{T}{n}(a - k\hat{X}_{t_i}^n) + \sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i})$$

with $\hat{X}_{t_0}^n = x_0$ can lead to negative values since the Gaussian increment is not bounded from below. Thus, this scheme is not well defined. To correct this problem, Deelstra and Delbaen [27] have proposed to consider:

$$\hat{X}_{t_{i+1}}^n = \hat{X}_{t_i}^n + \frac{T}{n}(a - k\hat{X}_{t_i}^n) + \sigma\sqrt{\hat{X}_{t_i}^n \mathbf{1}_{\hat{X}_{t_i}^n > 0}}(W_{t_{i+1}} - W_{t_i})$$

while Diop proposes in [28]:

$$\hat{X}_{t_{i+1}}^n = |\hat{X}_{t_i}^n + \frac{T}{n}(a - k\hat{X}_{t_i}^n) + \sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i})|.$$

However, we can as proposed in Brigo and Alfonsi [18] obtain the positivity using an implicit scheme. More precisely, if we rewrite the CIR process with the posticipated

stochastic integral, we get, since $d\langle\sqrt{X}, W\rangle_s = \frac{\sigma}{2}ds$:

$$\begin{aligned}
X_t &= x_0 + \int_0^t (a - kX_s)ds + \sigma \int_0^t \sqrt{X_s}dW_s \\
&= x_0 + \lim_{n \rightarrow \infty} \left\{ \sum_{i; t_i < t} (a - kX_{t_{i+1}}) \frac{T}{n} + \sigma \sum_{i; t_i < t} \sqrt{X_{t_{i+1}}} (W_{t_{i+1}} - W_{t_i}) \right. \\
&\quad \left. - \sigma \sum_{i; t_i < t} (\sqrt{X_{t_{i+1}}} - \sqrt{X_{t_i}}) (W_{t_{i+1}} - W_{t_i}) \right\} \\
&= x_0 + \lim_{n \rightarrow \infty} \left\{ \sum_{i; t_i < t} (a - \frac{\sigma^2}{2} - kX_{t_{i+1}}) \frac{T}{n} + \sigma \sum_{i; t_i < t} \sqrt{X_{t_{i+1}}} (W_{t_{i+1}} - W_{t_i}) \right\}.
\end{aligned}$$

It is then natural to consider the following implicit scheme that is well defined under the hypothesis (4.2) at least when the time step is small enough:

$$\hat{X}_{t_{i+1}}^n = \hat{X}_{t_i}^n + (a - \frac{\sigma^2}{2} - k\hat{X}_{t_{i+1}}^n) \frac{T}{n} + \sigma \sqrt{\hat{X}_{t_{i+1}}^n} (W_{t_{i+1}} - W_{t_i}).$$

More precisely, when $\hat{X}_{t_i}^n \geq 0$ and $\frac{T}{n} \leq 1/k^-$ (where $y^- = \max(-y, 0)$), $\sqrt{\hat{X}_{t_{i+1}}^n}$ can then be chosen as the unique positive root (since $2a > \sigma^2$, $P(0) < 0$) of the second-degree polynomial $P(x) = (1 + k\frac{T}{n})x^2 - \sigma(W_{t_{i+1}} - W_{t_i})x - (\hat{X}_{t_i}^n + (a - \frac{\sigma^2}{2})\frac{T}{n})$, and we get

$$\hat{X}_{t_{i+1}}^n = \left(\frac{\sigma(W_{t_{i+1}} - W_{t_i}) + \sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}_{t_i}^n + (a - \frac{\sigma^2}{2})\frac{T}{n})(1 + k\frac{T}{n})}}{2(1 + k\frac{T}{n})} \right)^2. \quad (4.3)$$

This scheme is well defined and it is also easy to check that it preserves the monotonicity property satisfied by the CIR process: if $x_0 < x'_0$ are two initial conditions, the scheme satisfies $\hat{X}_{t_i}^n < \hat{X}_{t_i}^{n'}$. This is an interesting example of implicit scheme on the diffusion coefficient whose general form is given by Milstein and al. (2002) since it leads to an analytical formula. In the same spirit, we can look at the SDE that drives the square-root:

$$d\sqrt{X_t} = \frac{a - \sigma^2/4}{2\sqrt{X_t}}dt - \frac{k}{2}\sqrt{X_t}dt + \frac{\sigma}{2}dW_t$$

and consider the scheme obtained by impliciting the drift. This gives also a second-degree equation in $\sqrt{\hat{X}_{t_{i+1}}^n}$:

$$\left(1 + \frac{kT}{2n}\right) \hat{X}_{t_{i+1}}^n - \left[\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}_{t_i}^n}\right] \sqrt{\hat{X}_{t_{i+1}}^n} - \frac{a - \sigma^2/4}{2} \frac{T}{n} = 0$$

that has also only one positive root when $\sigma^2 < 4a$ and $\frac{T}{n} < 2/k^-$, and it gives:

$$\hat{X}_{t_{i+1}}^n = \left(\frac{\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}_{t_i}^n} + \sqrt{(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}_{t_i}^n})^2 + 4(1 + \frac{kT}{2n})\frac{a - \sigma^2/4}{2}\frac{T}{n}}}{2(1 + \frac{kT}{2n})} \right)^2. \quad (4.4)$$

In this case $\hat{X}_{t_{i+1}}^n$ is still an increasing function of $\hat{X}_{t_i}^n$ so that the monotonicity property is satisfied. One can wonder whether we can get other schemes looking at the implicit scheme (implicit on the drift and the diffusion coefficients) with the SDE satisfied by X^α . It is not hard to see that the only two values of α that give a second-degree equation are 1 and 1/2. The other powers do not lead to analytical formulas and require a numerical resolution.

It is then interesting to make a rough Taylor expansion of order 1 of these schemes, i.e. we fix $\hat{X}_{t_i}^n$ and only conserve the terms in $\frac{T}{n}$, $(W_{t_{i+1}} - W_{t_i})$ and $(W_{t_{i+1}} - W_{t_i})^2$. We get respectively for the first scheme (4.3) and the second (4.4):

$$\begin{aligned} \hat{X}_{t_{i+1}}^n &\approx \hat{X}_{t_i}^n \left(1 - k\frac{T}{n} \right) + \sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i}) + \sigma^2/2(W_{t_{i+1}} - W_{t_i})^2 + (a - \sigma^2/2)\frac{T}{n} \\ \hat{X}_{t_{i+1}}^n &\approx \hat{X}_{t_i}^n \left(1 - k\frac{T}{n} \right) + \sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i}) + \sigma^2/4(W_{t_{i+1}} - W_{t_i})^2 + (a - \sigma^2/4)\frac{T}{n} \end{aligned}$$

This indicates us a family of explicit schemes $E(\lambda)$ for $0 \leq \lambda \leq a - \sigma^2/4$ that ensure nonnegative values but not the property of monotonicity:

$$\begin{aligned} \hat{X}_{t_{i+1}}^n &= \left(\left(1 - \frac{kT}{2n} \right) \sqrt{\hat{X}_{t_i}^n} + \frac{\sigma(W_{t_{i+1}} - W_{t_i})}{2(1 - \frac{kT}{2n})} \right)^2 \\ &\quad + (a - \sigma^2/4)T/n + \lambda[(W_{t_{i+1}} - W_{t_i})^2 - T/n]. \end{aligned} \quad (4.5)$$

It is well defined for $kT/n \neq 2$. The expansion of the scheme (4.3) corresponds then to $\lambda = \sigma^2/4$ while the scheme (4.4) to $\lambda = 0$. It is interesting here to notice that the implicit scheme on the square-root and the explicit scheme $E(0)$ have the same expansion (up to order 1) as the Milstein scheme for (4.1) (which can lead to negative values like the Euler scheme when $k > 0$) and for $k = 0$, $E(0)$ is exactly the Milstein scheme. Let us mention also that we could have considered as well the schemes obtained by replacing the factor $1 - \frac{kT}{2n}$ by $\sqrt{1 - kT/n}$ in (4.5).

This paper aims to get results on the weak and strong convergence of these schemes. Let us mention here that Deelstra and Deelbaen have proven in [27] a strong convergence result for their scheme. Diop also gets a strong convergence result in [28] but under some strong assumptions on the coefficients. She also obtains a weak convergence rate that depends on parameters. We introduce a framework in Section 4.2 that will allow us to study

simultaneously several schemes presented above. In Section 4.3, we will thus establish a result of strong convergence for the schemes that satisfy an hypothesis denoted by $(\mathcal{H}_S)$. Then we analyze the weak error in Section 4.4, establishing a convergence result with a $1/n$ rate for schemes satisfying an hypothesis denoted by $(\mathcal{H}_W)$. Moreover, an expansion of the weak error is given for the schemes $E(\lambda)$. Section 4.5 presents numerical results. We study in particular the strong convergence speed numerically and also calculate the computing time required by the several schemes. All the properties put in evidence by our analysis are listed in the conclusion, and $E(0)$ seems to be the scheme that gathers the most interesting properties.

4.2 Notations and preliminary lemmas

4.2.1 Some results on the CIR process

Lemma 4.2.1. *The moments of $(X_t)_{t \in [0, T]}$ are uniformly bounded by a constant that depends only on the parameters (k, a, σ, x_0) , T , and the order of the moment $p \in \mathbb{N}^*$. More precisely, setting $\tilde{u}_p(t, x_0) = \mathbb{E}[X_t^p]$, there exists smooth functions $\tilde{u}_{j,p}(t)$ that depend on (k, a, σ) such that:*

$$\tilde{u}_p(t, x_0) = \sum_{j=0}^p \tilde{u}_{j,p}(t) x_0^j.$$

Proof: We have $\tilde{u}_0(t, x_0) = 1$ and in the case $p = 1$, $\tilde{u}_1(t, x_0) = x + \int_0^t (a - k\tilde{u}_1(s, x_0)) ds$ than can be solved:

$$\tilde{u}_1(t, x_0) = x_0 e^{-kt} + a \frac{1 - e^{-kt}}{k}$$

with the convention that $\frac{1 - e^{-kt}}{k} = t$ for $k = 0$. Let us consider $p \geq 2$ and assume the result true for $1 \leq j \leq p - 1$. One has $\frac{d(\tilde{u}_p(t, x_0))}{dt} = [ap + \frac{1}{2}p(p - 1)\sigma^2]\tilde{u}_{p-1}(t, x_0) - kp\tilde{u}_p(t, x_0)$. Hence, we have

$$\tilde{u}_p(t, x_0) = (e^{-kt})^p \left(x_0^p + \int_0^t [ap + \frac{1}{2}p(p - 1)\sigma^2] (e^{ks})^p \tilde{u}_{p-1}(s, x_0) ds \right)$$

and we get the induction relations

$$\begin{aligned} \forall j \leq p - 1, \quad \tilde{u}_{j,p}(t) &= (e^{-kt})^p \int_0^t [ap + \frac{1}{2}p(p - 1)\sigma^2] (e^{ks})^p \tilde{u}_{j,p-1}(s) ds \\ \tilde{u}_{p,p}(t) &= (e^{-kt})^p. \end{aligned}$$

This gives the desired result, and we remark incidentally that $\tilde{u}_{j,p}(t)$ can be written as a polynomial of e^{-kt} or t depending on whether we are in the case $k \neq 0$ or $k = 0$. $\square$

4.2.2 Introduction of the notations $\mathcal{O}(1/n^\delta)$ and $O(1/n^\delta)$

In this section, we introduce Landau type notations for sequences of random variables that will considerably simplify formulas later. To allow the multiplication of two $\mathcal{O}$, we suppose the existence of moments of any order. The results presented here are elementary, and will largely be used later.

Definition 4.2.2. *Let us consider a doubly indexed family of random variables $Z = (Z_\gamma^n)_{n,\gamma}$ with $n \in \mathbb{N}$ and $\gamma \in \Gamma_n$ a nonempty set. We will say that Z is of order $\delta \in \mathbb{R}$ - and use the notation $Z_\gamma^n = \mathcal{O}(1/n^\delta)$ - if there exists a family of positive random variables $(A_\gamma^n)_{\gamma,n}$ that have moments of any order uniformly bounded (i.e. $\forall p \in \mathbb{N}^*, \exists \kappa(A, p) > 0, \forall n \in \mathbb{N}^*, \sup_{\gamma \in \Gamma_n} \mathbb{E}[(A_\gamma^n)^p] \leq \kappa(A, p)$) and such that:*

$$|Z_\gamma^n| \leq A_\gamma^n / n^\delta$$

This is clearly equivalent to the following property:

$$\forall p \in \mathbb{N}^*, \exists \kappa(p) > 0, \forall n \in \mathbb{N}^*, \sup_{\gamma \in \Gamma_n} \mathbb{E}[(n^\delta |Z_\gamma^n|)^p] \leq \kappa(p)$$

When in particular the $(Z_\gamma^n)_{\gamma,n}$ are deterministic, this is equivalent to the boundedness of $(n^\delta Z_\gamma^n)_{\gamma,n}$ and we use the standard notation $Z_\gamma^n = O(1/n^\delta)$.

Remarks 4.2.3. 1. *It is obvious but important to observe that $Z_\gamma^n = \mathcal{O}(1/n^\delta)$ implies that $\mathbb{E}[Z_\gamma^n] = O(1/n^\delta)$.*

2. *Typically we will use in the paper this definition for $\Gamma_n = \{t_0, t_1, \dots, t_n\}$.*

3. *A simple but fundamental example is $W_{t_{i+1}^n} - W_{t_i^n} = \mathcal{O}(1/\sqrt{n})$ which is clear since $\sqrt{n}|W_{t_{i+1}^n} - W_{t_i^n}| \stackrel{\text{law}}{=} |\mathcal{N}(0, T)|$ has moments of any order.*

Proposition 4.2.4. *If $(Z_\gamma^n)_{n \in \mathbb{N}, \gamma \in \Gamma_n}$ and $(Z_{\gamma'}^n)_{n \in \mathbb{N}, \gamma' \in \Gamma'_n}$ are two families such that $Z_\gamma^n = \mathcal{O}(1/n^\delta)$ and $Z_{\gamma'}^n = \mathcal{O}(1/n^{\delta'})$, we have:*

$$\begin{aligned} 1) \quad & \forall c \in \mathbb{R}^*, cZ_\gamma^n = \mathcal{O}(1/n^\delta) & 2) \quad & \forall d \in \mathbb{R}, Z_\gamma^n / n^d = \mathcal{O}(1/n^{\delta+d}) \\ 3) \quad & Z_\gamma^n + Z_{\gamma'}^n = \mathcal{O}(1/n^{\inf(\delta, \delta')}) & 4) \quad & \forall d > 0, (Z_\gamma^n)^d = \mathcal{O}(1/n^{d\delta}) \\ 5) \quad & Z_\gamma^n Z_{\gamma'}^n = \mathcal{O}(1/n^{\delta+\delta'}) \end{aligned}$$

where the families in 3) and 5) are indexed in $\Gamma_n \times \Gamma'_n$. In particular, if we have a family of functions $h_n : \Gamma_n \rightarrow \Gamma'_n$, we have also:

$$3') \quad Z_\gamma^n + Z_{h_n(\gamma)}^n = \mathcal{O}(1/n^{\inf(\delta, \delta')}) \quad 5') \quad Z_\gamma^n Z_{h_n(\gamma)}^n = \mathcal{O}(1/n^{\delta+\delta'}).$$

Proof : 1) and 2) are obvious. To prove 3), let us assume for example that $\delta \leq \delta'$. Then, it is not hard to see that $Z_i^n = \mathcal{O}(1/n^\delta)$. Since a sum of L^p random variables is L^p , we conclude easily. 4) comes immediately from the definition while 5) requires the use of Cauchy-Schwarz inequality to get the boundedness of the moments. $\square$

By Jensen's inequality, we also easily check the following result.

Lemma 4.2.5. *Let us consider a family $(\mathcal{G}_{\gamma'})_{\gamma' \in \Gamma'_n}$ of σ -algebras and $(Z_\gamma^n)_{n \in \mathbb{N}, \gamma \in \Gamma_n}$ a family of random variables such that $Z_\gamma^n = \mathcal{O}(1/n^\delta)$, then $\mathbb{E}(Z_\gamma^n | \mathcal{G}_{\gamma'}) = \mathcal{O}(1/n^\delta)$.*

4.2.3 On the moments of the discretization schemes

First of all, we need the following lemma to control the moments of the schemes presented here.

Lemma 4.2.6. *Let us suppose that $(\hat{X}_{t_i}^n)$ is an nonnegative adapted scheme (i.e. $\hat{X}_{t_i}^n$ is $\mathcal{F}_{t_i}$ -measurable) such that for all $n \in \mathbb{N}$,*

$$\begin{aligned} \hat{X}_{t_0}^n &= x_0 \\ \forall i \leq n-1, \hat{X}_{t_{i+1}}^n &\leq (1+b/n)\hat{X}_{t_i}^n + \sigma_{t_i}^n \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) + \mathcal{O}(1/n) \end{aligned}$$

where $(\sigma_{t_i}^n)$ is also supposed to be adapted with $\sigma_{t_i}^n = \mathcal{O}(1)$ and $b > 0$. Then, $(\hat{X}_{t_i}^n)$ has uniformly bounded moments, that is $\hat{X}_{t_i}^n = \mathcal{O}(1)$.

Proof : Let us first remark that it is sufficient to study the case $b = 0$. Indeed, $(1+b/n)^{-i}\hat{X}_{t_i}^n$ satisfies the condition above with $b = 0$: we have for $x \in [0, n]$, $1 \leq (1+b/n)^x \leq e^b$ and thus on the one hand, $(1+b/n)^{-1-i/2}\sigma_{t_i}^n$ is adapted and thanks to Proposition 4.2.4 is a $\mathcal{O}(1)$, and on the other hand $(1+b/n)^{-i-1}\mathcal{O}(1/n) = \mathcal{O}(1/n)$. We observe then that $\hat{X}_{t_i}^n = \mathcal{O}(1) \iff (1+b/n)^{-i}\hat{X}_{t_i}^n = \mathcal{O}(1)$.

By Definition 4.2.2, there is $A_i^n = \mathcal{O}(1)$ such that we can rewrite the inequality (with $b = 0$) as follows :

$$\hat{X}_{t_{i+1}}^n \leq \hat{X}_{t_i}^n + \sigma_{t_i}^n \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) + A_i^n/n.$$

We denote in this proof $\kappa(A, p) = \sup_{i, n} \mathbb{E}[|A_i^n|^p]$. We are going to check by on p that

$\forall p \in \mathbb{N}, \sup_{i, n} \mathbb{E}[(\hat{X}_{t_i}^n)^p] < \infty$. It is easy to check that $\mathbb{E}[\hat{X}_{t_i}^n] \leq x_0 + \kappa(A, 1)$ since we have $\mathbb{E}[\hat{X}_{t_{i+1}}^n] \leq \mathbb{E}[\hat{X}_{t_i}^n] + \kappa(A, 1)/n$. Let us assume for any $q \leq p-1$, there is a positive constant $\kappa(q)$ such that

$$\mathbb{E}[(\hat{X}_{t_i}^n)^q] \leq \kappa(q).$$

Since $(\hat{X}_{t_{i+1}}^n)^p \leq \sum_{l_1+l_2+l_3=p} \frac{p!}{l_1!l_2!l_3!} (\hat{X}_{t_i}^n)^{l_1+l_2/2} (\sigma_{t_i}^n (W_{t_{i+1}} - W_{t_i}))^{l_2} (A_i^n/n)^{l_3}$, it is sufficient to control $E(l_1, l_2, l_3) = \mathbb{E}[(\hat{X}_{t_i}^n)^{l_1+l_2/2} (\sigma_{t_i}^n (W_{t_{i+1}} - W_{t_i}))^{l_2} (A_i^n/n)^{l_3}]$ for $l_1 + l_2 + l_3 = p$. If

$l_1 + l_2/2 \leq p - 3/2$, we have necessary $l_3 + l_2/2 \geq 3/2$ and Hölder inequality gives

$$E(l_1, l_2, l_3) \leq (\kappa(p-1))^{1/\alpha} \mathbb{E} \left[\left((\sigma_{t_i}^n(W_{t_{i+1}} - W_{t_i}))^{l_2} (A_i^n/n)^{l_3} \right)^\beta \right]^{1/\beta} \leq \frac{C(l_1, l_2, l_3)}{n^{3/2}}$$

where $\alpha = \frac{p-1}{l_1+l_2/2}$ and $1/\alpha + 1/\beta = 1$. Thus, there is a positive constant Cte such that :

$$\begin{aligned} \mathbb{E}[(\hat{X}_{t_{i+1}}^n)^p] &\leq \mathbb{E}[(\hat{X}_{t_i}^n)^p] + \frac{p}{n} \mathbb{E}[(\hat{X}_{t_i}^n)^{p-1} A_i^n] + \frac{p}{n} \mathbb{E}[(\hat{X}_{t_i}^n)^{p-1/2} \sigma_{t_i}^n] \mathbb{E}(W_{t_{i+1}} - W_{t_i}) \\ &\quad + \frac{p(p-1)}{2n} \mathbb{E}[(\hat{X}_{t_i}^n)^{p-1} (\sigma_{t_i}^n)^2] + Cte/n. \end{aligned}$$

Using once again the Hölder inequality to bound $\mathbb{E}[(\hat{X}_{t_i}^n)^{p-1} A_i^n]$ from above, have for a constant $C > 0$

$$\mathbb{E}[(\hat{X}_{t_{i+1}}^n)^p] \leq \mathbb{E}[(\hat{X}_{t_i}^n)^p] + \frac{C}{n} (\mathbb{E}[(\hat{X}_{t_i}^n)^p] + 1) = \mathbb{E}[(\hat{X}_{t_i}^n)^p] (1 + \frac{C}{n}) + \frac{C}{n}$$

and then we easily conclude that $\mathbb{E}[(\hat{X}_{t_i}^n)^p] + 1 \leq (x_0^p + 1)e^C$. $\square$

Now, we present a quite general framework that includes, as we will see, the implicit scheme (4.3) and the explicit schemes $E(\lambda)$. The hypotheses that are stated below will be useful later to get results of strong and weak convergence.

Hypothesis ($\mathcal{H}_S$) *We will say that $(\hat{X}_{t_i}^n)$ satisfies $(\mathcal{H}_S)$ if it is a nonnegative adapted scheme such that:*

$$\hat{X}_{t_{i+1}}^n = \hat{X}_{t_i}^n + \frac{T}{n} (a - k\hat{X}_{t_i}^n) + \sigma \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) + m_{t_{i+1}}^n - m_{t_i}^n + \mathcal{O}(1/n^{3/2}) \quad (4.6)$$

where $m_{t_{i+1}}^n - m_{t_i}^n$ is a martingale increment (i.e. $\mathbb{E}[m_{t_{i+1}}^n - m_{t_i}^n | \mathcal{F}_{t_i}] = 0$) of order 1:

$$m_{t_{i+1}}^n - m_{t_i}^n = \mathcal{O}(1/n). \quad (4.7)$$

If it is satisfied, we get immediately that $\hat{X}_{t_{i+1}}^n \leq \hat{X}_{t_i}^n + \frac{|k|T}{n} \hat{X}_{t_i}^n + \sigma \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) + \mathcal{O}(1/n)$ using that $|a - k\hat{X}_{t_i}^n| \leq a + |k|\hat{X}_{t_i}^n$. Therefore, we can apply the Lemma 4.2.6 and deduce that $\hat{X}_{t_i}^n = \mathcal{O}(1)$. We define in that case the discrete martingale $(M_{t_i}^n)$ by

$$\begin{cases} M_{t_0}^n = 0 \\ M_{t_{i+1}}^n - M_{t_i}^n = \sigma \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) + m_{t_{i+1}}^n - m_{t_i}^n. \end{cases} \quad (4.8)$$

Thanks to Proposition 4.2.4 and Remark 4.2.3, we get

Corollary 4.2.7. *Under hypothesis $(\mathcal{H}_S)$, $\hat{X}_{t_i}^n$ has uniformly bounded moments, and we have:*

$$\begin{aligned} (M_{t_{i+1}}^n - M_{t_i}^n)^2 &= \sigma^2 \hat{X}_{t_i}^n (W_{t_{i+1}} - W_{t_i})^2 + \mathcal{O}(1/n^{3/2}) \\ \hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n &= \mathcal{O}(1/\sqrt{n}). \end{aligned}$$

However, as we will see when studying the weak error, it can be useful to make a stronger assumption to get a faster convergence.

Hypothesis $(\mathcal{H}_W)$ We say that a scheme $(\hat{X}_{t_i}^n)$ satisfies $(\mathcal{H}_W)$ if it already satisfies $(\mathcal{H}_S)$ and moreover

$$\hat{X}_{t_{i+1}}^n = \hat{X}_{t_i}^n + \frac{T}{n}(a - k\hat{X}_{t_i}^n) + \sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i}) + m_{t_{i+1}}^n - m_{t_i}^n + \mathcal{O}(1/n^2) \quad (4.9)$$

$$\mathbb{E} \left[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^2 | \mathcal{F}_{t_i} \right] = \sigma^2 \hat{X}_{t_i}^n T/n + \mathcal{O}(1/n^2). \quad (4.10)$$

The absence of term of order $3/2$ in (4.10) and the knowledge of the expansion of the scheme (4.9) up to order 2 play a key role to get a weak error at most proportional to the time step.

Remark 4.2.8. Let us suppose that there is a function $\psi^n(x, w)$ which is even with respect to its second argument w such that:

$$m_{t_{i+1}}^n - m_{t_i}^n = \psi^n(\hat{X}_{t_i}^n, (W_{t_{i+1}} - W_{t_i})) + \mathcal{O}(1/n^{3/2}).$$

Then, $(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^2 = \sigma^2 \hat{X}_{t_i}^n (W_{t_{i+1}} - W_{t_i})^2 + 2\sigma\sqrt{\hat{X}_{t_i}^n} \psi^n(\hat{X}_{t_i}^n, (W_{t_{i+1}} - W_{t_i}))(W_{t_{i+1}} - W_{t_i}) + \mathcal{O}(1/n^2)$, and therefore condition (4.10) is automatically satisfied thanks to Lemma 4.2.5.

4.2.4 Study of the expansion of the different schemes

In this section we examine each scheme presented in the introduction and our aim is to discuss whether it satisfies or not Hypotheses $(\mathcal{H}_S)$ and $(\mathcal{H}_W)$ defined before.

Expansion of the implicit scheme (4.3)

We assume here that $2a > \sigma^2$, and expand the relation that defines the implicit scheme (4.3):

$$\begin{aligned} \hat{X}_{t_{i+1}}^n &= \frac{1}{4(1 + kT/n)^2} \left(2\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}_{t_i}^n + (a - \frac{\sigma^2}{2})T/n)(1 + kT/n) \right. \\ &\quad \left. + 2\sigma(W_{t_{i+1}} - W_{t_i})\sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}_{t_i}^n + (a - \frac{\sigma^2}{2})T/n)(1 + kT/n)} \right) \end{aligned} \quad (4.11)$$

Let us now observe that

$$\begin{aligned} & \left| \sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}_{t_i}^n + (a - \frac{\sigma^2}{2})T/n)(1 + kT/n)} - 2\sqrt{\hat{X}_{t_i}^n(1 + kT/n)} \right| \\ & \leq \sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(a - \frac{\sigma^2}{2})(1 + kT/n)T/n} = \mathcal{O}(1/\sqrt{n}), \end{aligned} \quad (4.12)$$

using Proposition 4.2.4. Thus, we have

$$\hat{X}_{t_{i+1}}^n = \frac{1}{1+kT/n} \hat{X}_{t_i}^n + \frac{1}{(1+kT/n)^{3/2}} \sigma \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) + \mathcal{O}(1/n)$$

which gives that $\hat{X}_{t_i}^n = \mathcal{O}(1)$ using Lemma 4.2.6. Once we know this, we can continue the expansion thanks to Proposition 4.2.4 and it is not hard to get:

$$\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n = \frac{T}{n} (a - k\hat{X}_{t_i}^n) + \frac{\sigma^2}{2} [(W_{t_{i+1}} - W_{t_i})^2 - T/n] + \dot{M}_{t_{i+1}}^n - \dot{M}_{t_i}^n + \mathcal{O}(1/n^2) \quad (4.13)$$

where $\dot{M}_{t_i}^n$ is a discrete $\mathcal{F}_{t_i}$ -martingale defined by $\dot{M}_0^n = 0$ and

$$\dot{M}_{t_{i+1}}^n = \dot{M}_{t_i}^n + \frac{\sigma(W_{t_{i+1}} - W_{t_i})}{2(1+kT/n)^2} \sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4 \left(\hat{X}_{t_i}^n + (a - \frac{\sigma^2}{2}) \frac{T}{n} \right) (1+kT/n)}.$$

Indeed, we have $\mathbb{E}(\dot{M}_{t_{i+1}}^n | \mathcal{F}_{t_i}) =$

$$\dot{M}_{t_i}^n + \frac{\sigma\sqrt{T}}{2\sqrt{2\pi n}(1+kT/n)^2} \underbrace{\int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}} \sqrt{\frac{\sigma^2 T}{n} x^2 + 4(\hat{X}_{t_i}^n + (a - \frac{\sigma^2}{2})T/n)(1+kT/n)} dx}_0 = \dot{M}_{t_i}^n.$$

Moreover, we have $(\dot{M}_{t_{i+1}}^n - \dot{M}_{t_i}^n)^2 = \sigma^2(W_{t_{i+1}} - W_{t_i})^2 \hat{X}_{t_i}^n + \mathcal{O}(1/n^2)$ and in particular $\dot{M}_{t_{i+1}}^n - \dot{M}_{t_i}^n = \mathcal{O}(1/\sqrt{n})$. Now, we can define the martingale $(m_{t_i}^n)$ by

$$m_{t_{i+1}}^n - m_{t_i}^n = \frac{\sigma^2}{2} [(W_{t_{i+1}} - W_{t_i})^2 - T/n] + \dot{M}_{t_{i+1}}^n - \dot{M}_{t_i}^n - \sigma \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i})$$

and it is easy from (4.13) to see that the properties (4.6) and (4.9) are satisfied. Inequality (4.12) gives us that $\dot{M}_{t_{i+1}}^n - \dot{M}_{t_i}^n - \sigma \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) = \mathcal{O}(1/n)$ and therefore property (4.7) is satisfied by m^n since $\frac{\sigma^2}{2} [(W_{t_{i+1}} - W_{t_i})^2 - T/n] = \mathcal{O}(1/n)$. We have first shown thus that $(\mathcal{H}_S)$ is satisfied. Now, using the Proposition 4.2.4, we get that:

$$\begin{aligned} (\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^2 &= \sigma^2(W_{t_{i+1}} - W_{t_i})^2 \hat{X}_{t_i}^n \\ &\quad + [\sigma^2((W_{t_{i+1}} - W_{t_i})^2 - T/n) + 2(a - k\hat{X}_{t_i}^n)T/n](\dot{M}_{t_{i+1}}^n - \dot{M}_{t_i}^n) + \mathcal{O}(1/n^2) \end{aligned}$$

and that the term of order 3/2, $[\sigma^2((W_{t_{i+1}} - W_{t_i})^2 - T/n) + 2(a - k\hat{X}_{t_i}^n)T/n](\dot{M}_{t_{i+1}}^n - \dot{M}_{t_i}^n)$, has a null conditional expectation respect to $\mathcal{F}_{t_i}$ since it can be written as an odd function respect to the Brownian increment. This shows that we have (4.10) and $(\mathcal{H}_W)$ is also satisfied by this implicit scheme.

Expansion of the implicit scheme (4.4)

Let us assume here that $4a > \sigma^2$. Expanding (4.4), we get:

$$\begin{aligned} \hat{X}_{t_{i+1}}^n &= \frac{1}{4(1 + \frac{kT}{2n})^2} \left[2 \left(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}_{t_i}^n} \right)^2 + 4 \left(1 + \frac{kT}{2n} \right) \frac{a - \sigma^2/4}{2} T/n \right. \\ &\quad \left. + 2 \left(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}_{t_i}^n} \right) \sqrt{\left(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}_{t_i}^n} \right)^2 + 4 \left(1 + \frac{kT}{2n} \right) \frac{a - \sigma^2/4}{2} \frac{T}{n}} \right]. \end{aligned}$$

Thus, using the inequality $x^2 + x\sqrt{x^2 + y} \leq \begin{cases} 2x^2 + y/2 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$ for $y \geq 0$, we get that

$$\hat{X}_{t_{i+1}}^n \leq \frac{1}{(1 + \frac{kT}{2n})^2} \left[\left(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}_{t_i}^n} \right)^2 + \left(1 + \frac{kT}{2n} \right) \frac{a - \sigma^2/4}{2} T/n \right]$$

and we can therefore apply Proposition 4.2.6 to deduce that $\hat{X}_{t_i}^n$ has bounded moments. Unfortunately, if we try now to get an expansion of $\hat{X}_{t_i}^n$ up to order $3/2$ by expanding the square-root, we get a term in $\frac{1}{\sqrt{\hat{X}_{t_i}^n}} \mathcal{O}(1/n^{3/2})$ which is hard to manage. Despite the good numerical convergence of this scheme, our approach in this paper did not enable us to obtain theoretical results for it.

Expansion of the explicit scheme $E(\lambda)$

Let us assume here that $4a \geq \sigma^2$ and consider $\lambda \in [0, a - \sigma^2/4]$. Expanding (4.5), we get

$$\begin{aligned} \hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n &= (a - k\hat{X}_{t_i}^n) \frac{T}{n} + \frac{k^2}{4} \hat{X}_{t_i}^n \left(\frac{T}{n} \right)^2 + \frac{k\sigma^2}{8} \frac{2 - kT/(2n)}{(1 - kT/(2n))^2} \left(\frac{T}{n} \right)^2 \\ &\quad + \sigma \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) + \left(\frac{\sigma^2}{4(1 - kT/(2n))^2} + \lambda \right) ((W_{t_{i+1}} - W_{t_i})^2 - T/n) \\ &\leq \sigma \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) + \left(k^- + \frac{k^2 T}{4n} \right) \hat{X}_{t_i}^n \frac{T}{n} + \mathcal{O}(1/n). \end{aligned}$$

We can then apply Lemma 4.2.6 to deduce that $\hat{X}_{t_i}^n = \mathcal{O}(1)$. We have then an expansion analogous to that obtained for the implicit scheme, that is

$$\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n = \frac{T}{n} (a - k\hat{X}_{t_i}^n) + \sigma \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) + m_{t_{i+1}}^n - m_{t_i}^n + \mathcal{O}(1/n^2) \quad (4.14)$$

where $m_{t_i}^n$ is a $\mathcal{F}_{t_i}$ -martingale defined by $m_{t_0}^n = 0$ and $m_{t_{i+1}}^n - m_{t_i}^n = (\frac{\sigma^2}{4} + \lambda)[(W_{t_{i+1}} - W_{t_i})^2 - T/n]$. It is in this case straightforward to see that we have the properties (4.6) and (4.9) and that the martingale increments satisfy (4.7) and (4.10) thanks to Remark 4.2.8. Hence, explicit scheme $E(\lambda)$ fulfills the conditions of $(\mathcal{H}_S)$ and $(\mathcal{H}_W)$.

4.3 Strong convergence

In all this section, we consider a scheme $(\hat{X}_{t_i}^n)$ that satisfies the hypothesis $(\mathcal{H}_S)$. We will prove the strong convergence for it, following the method proposed by Deelstra and Delbaen [27] that relies on Yamada's functions. Thus, we first need to build a continuous adapted extension of our scheme in order to use then Itô's formula. For that purpose, we need to explicit the $\mathcal{O}$ terms and first define $Z_{t_i}^n = \mathcal{O}(1/n^{3/2})$ as:

$$\hat{X}_{t_{i+1}}^n = \hat{X}_{t_i}^n + \frac{T}{n}(a - k\hat{X}_{t_i}^n) + M_{t_{i+1}}^n - M_{t_i}^n + Z_{t_i}^n$$

We can suppose that $Z_{t_i}^n$ is $\mathcal{F}_{t_i}$ -measurable. Indeed, if it were not the case, it would be sufficient then to consider the martingale increment

$$\tilde{M}_{t_{i+1}}^n - \tilde{M}_{t_i}^n = M_{t_{i+1}}^n - M_{t_i}^n + Z_{t_i}^n - \mathbb{E}[Z_{t_i}^n | \mathcal{F}_{t_i}]$$

and $\tilde{Z}_{t_i}^n = \mathbb{E}[Z_{t_i}^n | \mathcal{F}_{t_i}]$ instead of respectively $M_{t_{i+1}}^n - M_{t_i}^n$ and $Z_{t_i}^n$. Thus, we have $\tilde{M}_{t_{i+1}}^n - \tilde{M}_{t_i}^n + \tilde{Z}_{t_i}^n = M_{t_{i+1}}^n - M_{t_i}^n + Z_{t_i}^n$ and, thanks to Lemma 4.2.5, we get that $\tilde{Z}_{t_i}^n = \mathcal{O}(1/n^{3/2})$, and also $\tilde{M}_{t_{i+1}}^n - \tilde{M}_{t_i}^n = \sigma \sqrt{\hat{X}_{t_i}^n} (W_{t_{i+1}} - W_{t_i}) + \mathcal{O}(1/n)$.

Now, we apply the martingale representation theorem to the martingales $\{\mathbb{E}(m_{t_{i+1}}^n | \mathcal{F}_t) - m_{t_i}^n, t \in [t_i, t_{i+1}]\}$ to get the existence of an $\mathcal{F}_t$ -adapted process $(R_t^n, 0 \leq t \leq T)$ such that

$$\mathbb{E}(m_{t_{i+1}}^n | \mathcal{F}_t) - m_{t_i}^n = \int_{t_i}^t R_s^n dW_s.$$

In particular, we know that $\int_{t_i}^t R_s^n dW_s = \mathcal{O}(1/n)$ and so $(\int_{t_i}^t R_s^n dW_s)^2 = \mathcal{O}(1/n^2)$ which gives us that, for $t \in [t_i, t_{i+1}]$:

$$\int_{t_i}^t \mathbb{E}[(R_s^n)^2] ds = \mathcal{O}(1/n^2). \quad (4.15)$$

Now, we are able to build a continuous extension $(\hat{X}_t^n, 0 \leq t \leq T)$ $\mathcal{F}_t$ -adapted of our discretization scheme. Indeed, we define for $t \in [t_i, t_{i+1}]$:

$$\hat{X}_t^n = \hat{X}_{t_i}^n + (t - t_i)(a - k\hat{X}_{t_i}^n + \frac{n}{T}Z_{t_i}^n) + \int_{t_i}^t (\sigma \sqrt{\hat{X}_{t_i}^n} + R_s^n) dW_s.$$

Thus, naming $\eta(t)$ the function defined on $[0, T]$ by $\eta(t) = t_i$ for $t \in [t_i, t_{i+1})$, we can rewrite our scheme as follows:

$$\hat{X}_t^n = x_0 + \int_0^t (a - k\hat{X}_{\eta(s)}^n + \frac{n}{T}Z_{\eta(s)}^n) ds + \int_0^t (\sigma \sqrt{\hat{X}_{\eta(s)}^n} + R_s^n) dW_s. \quad (4.16)$$

Let us now introduce a family of Yamada's functions (see Karatzas and Shreve [47]) $\psi_{\epsilon, m}$ parametrized by two positive numbers ϵ and m . Since we have $\int_{\epsilon e^{-\sigma^2 m}}^{\epsilon} \frac{1}{\sigma^2 u} du = m$, there

exists a continuous function $\rho_{\epsilon,m}$ with a compact support in $]e^{-\sigma^2 m}, \epsilon[$ such that $\rho_{\epsilon,m}(x) \leq \frac{2}{\sigma^2 x m}$ for $x > 0$ and $\int_{\epsilon e^{-\sigma^2 m}}^{\epsilon} \rho_{\epsilon,m}(u) du = 1$. We then consider

$$\psi_{\epsilon,m}(x) = \int_0^{|x|} \int_0^y \rho_{\epsilon,m}(u) du dy$$

that can be viewed as a sequence of smooth approximation of $x \rightarrow |x|$ when m is large and ϵ tends to 0. Indeed functions $\psi_{\epsilon,m}$ thus satisfies:

$$|x| - \epsilon \leq \psi_{\epsilon,m}(x) \leq |x|, |\psi'_{\epsilon,m}(x)| \leq 1, 0 \leq \psi''_{\epsilon,m}(x) = \rho_{\epsilon,m}(|x|) \leq \frac{2}{\sigma^2 |x| m}.$$

Following the method used by Deelstra and Delbaen [27], we first write

$$|\hat{X}_t^n - X_t| \leq \epsilon + \psi_{\epsilon,m}(\hat{X}_t^n - X_t) \quad (4.17)$$

and then apply Itô's formula :

$$\begin{aligned} \psi_{\epsilon,m}(\hat{X}_t^n - X_t) &= \int_0^t (kX_s - k\hat{X}_{\eta(s)}^n + \frac{n}{T} Z_{\eta(s)}^n) \psi'_{\epsilon,m}(\hat{X}_s^n - X_s) ds \\ &\quad + \int_0^t (\sigma \sqrt{\hat{X}_{\eta(s)}^n} + R_s^n - \sigma \sqrt{X_s}) \psi'_{\epsilon,m}(\hat{X}_s^n - X_s) dW_s \\ &\quad + \frac{1}{2} \int_0^t (\sigma \sqrt{\hat{X}_{\eta(s)}^n} + R_s^n - \sigma \sqrt{X_s})^2 \psi''_{\epsilon,m}(\hat{X}_s^n - X_s) ds \\ &=: I_1(t, n) + I_2(t, n) + I_3(t, n). \end{aligned}$$

The absolute value of the first integral can be bounded using that $\|\psi'_{\epsilon,m}\|_{\infty} \leq 1$:

$$I_1(t, n) \leq |k| \int_0^t (|X_s - \hat{X}_s^n| + |\hat{X}_s^n - \hat{X}_{\eta(s)}^n|) ds + \int_0^t \frac{n}{T} |Z_{\eta(s)}^n| ds.$$

For the third integral, we have that

$$\begin{aligned} (\sigma \sqrt{\hat{X}_{\eta(s)}^n} + R_s^n - \sigma \sqrt{X_s})^2 &\leq 2(\sigma^2 |X_s - \hat{X}_{\eta(s)}^n| + (R_s^n)^2) \\ &\leq 2(\sigma^2 |X_s - \hat{X}_s^n| + \sigma^2 |\hat{X}_s^n - \hat{X}_{\eta(s)}^n| + (R_s^n)^2). \end{aligned} \quad (4.18)$$

Therefore, using that $\psi''_{\epsilon,m}(x)|x| \leq \frac{2}{\sigma^2 m}$ and $\|\psi''_{\epsilon,m}\|_{\infty} \leq \frac{2e^{\sigma^2 m}}{\sigma^2 \epsilon m}$ we get:

$$I_3(t, n) \leq \frac{2t}{m} + \frac{2e^{\sigma^2 m}}{\epsilon m} \int_0^t (|\hat{X}_s^n - \hat{X}_{\eta(s)}^n| + \frac{1}{\sigma^2} (R_s^n)^2) ds.$$

Using Lemma 4.2.1, Corollary 4.2.7 and (4.15), we check that $\mathbb{E}[I_2(t, n)] = 0$. Now, taking the expectation in (4.17), we get

$$\forall t \in [0, T], \mathbb{E}(|\hat{X}_t^n - X_t|) \leq \epsilon + |k| \int_0^t \mathbb{E}(|\hat{X}_s^n - X_s|) ds + \frac{2T}{m} + \left(\frac{2e^{\sigma^2 m}}{\sigma^2 \epsilon m} + |k| \right) \frac{Cte}{\sqrt{n}}$$

for some $Cte > 0$, using that $|\hat{X}_s^n - \hat{X}_{\eta(s)}^n| = \mathcal{O}(1/\sqrt{n})$ and $\frac{n}{T}Z_{\eta(s)}^n = \mathcal{O}(1/\sqrt{n})$. Gronwall's lemma leads then to

$$\forall t \in [0, T], \mathbb{E}(|\hat{X}_t^n - X_t|) \leq e^{|k|T} \left[\epsilon + \frac{2T}{m} + \left(\frac{2e^{\sigma^2 m}}{\sigma^2 \epsilon m} + |k| \right) \frac{Cte}{\sqrt{n}} \right]. \quad (4.19)$$

Now, taking $m = \frac{1}{4\sigma^2} \ln(n)$ and $\epsilon = 1/\ln(n)$, we get that

$$\sup_{0 \leq t \leq T} \mathbb{E}(|\hat{X}_t^n - X_t|) = O(1/\ln(n)). \quad (4.20)$$

Now, we would like to exchange the supremum and the expectation. Doob's inequality gives $\mathbb{E}[\sup_{0 \leq s \leq t} |I_2(s, n)|] \leq C \sqrt{\mathbb{E} \left[\int_0^t (\sigma \sqrt{\hat{X}_{\eta(s)}^n} + R_s^n - \sigma \sqrt{X_s})^2 (\psi'_{\epsilon, m}(\hat{X}_s^n - X_s))^2 ds \right]}$. We use that $\|\psi'_{\epsilon, m}\|_\infty \leq 1$ and the inequality (4.18), and then control each terms thanks to relations (4.20) and (4.15), and observing that $|\hat{X}_s^n - \hat{X}_{\eta(s)}^n| = \mathcal{O}(1/\sqrt{n})$:

$$\mathbb{E}[\sup_{0 \leq s \leq t} |I_2(s, n)|] = O(1/\sqrt{\ln(n)}).$$

We can then use the same controls as before for I_1 and I_3 to conclude that

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |\hat{X}_t^n - X_t| \right) = O(1/\sqrt{\ln(n)}). \quad (4.21)$$

We sum up our results in the proposition that follows.

Proposition 4.3.1. *Let us consider a discretization scheme $(\hat{X}^n)$ that satisfies the hypothesis $(\mathcal{H}_S)$. Then, there exists a positive constant C depending on T and on the parameters (k, a, σ, x_0) but not on n such that:*

$$\begin{aligned} \sup_{0 \leq i \leq n} \mathbb{E}(|\hat{X}_{t_i}^n - X_{t_i}|) &\leq C/\ln(n) \\ \mathbb{E} \left(\sup_{0 \leq i \leq n} |\hat{X}_{t_i}^n - X_{t_i}| \right) &= C/\sqrt{\ln(n)}. \end{aligned}$$

4.4 Weak convergence

In this section, we will establish a result that gives the convergence rate of $\mathbb{E}[f(\hat{X}_T^n)]$ to $\mathbb{E}[f(X_T)]$. We will use the method introduced by Talay and Tubaro (1990) to study that weak error and get also a convergence rate in $1/n$ provided that f is regular enough. We thus introduce the notation X_t^x to denote the CIR process with initial value x , and we first need to establish the following technical result.

Proposition 4.4.1. *Let us consider $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ a $\mathcal{C}^q$ function with $q \geq 2$, such that there is $A > 0$ and $m \geq q$, $m \in \mathbb{N}$ such that*

$$\forall x \geq 0, |f^{(q)}(x)| \leq A(1 + x^m).$$

Then $u : [0, T] \times \mathbb{R}_+ \rightarrow \mathbb{R}$ defined by $u(t, x) = \mathbb{E}[f(X_{T-t}^x)]$ has successive derivatives $\partial_x^l \partial_t^{l'} u(t, x)$ for $l, l' \in \mathbb{N}$ and $l + 2l' \leq q$, that satisfy the following property:

$$\exists C > 0, \forall (t, x) \in [0, T] \times \mathbb{R}_+, \max_{l+2l' \leq q} |\partial_x^l \partial_t^{l'} u(t, x)| \leq C(1 + x^{m+q+l'}) \quad (4.22)$$

and is a classical solution of the PDE:

$$\begin{cases} \partial_t u(t, x) + (a - kx) \partial_x u(t, x) + \frac{\sigma^2}{2} x \partial_x^2 u(t, x) = 0 \\ u(T, x) = f(x). \end{cases} \quad (4.23)$$

More generally, let us assume that $(f_\theta, \theta \in \Theta)$ is a family of $\mathcal{C}^q$ functions with $q \geq 2$, such that there is $A > 0$ and $m \geq q$, $m \in \mathbb{N}$ such that

$$\forall \theta \in \Theta, \forall x \geq 0, |f_\theta^{(q)}(x)| \leq A(1 + x^m) \text{ and } \forall l < q, |f_\theta^{(l)}(0)| \leq A. \quad (4.24)$$

For $0 \leq \tau \leq T$, we consider $u_{\theta, \tau}(t, x) = \mathbb{E}[f_\theta(X_{\tau-t}^x)]$ for $0 \leq t \leq \tau$ and $x \geq 0$. Then there is a constant $C > 0$ that does not depend on τ and θ such that

$$\forall \theta \in \Theta, \tau \in [0, T], \forall (t, x) \in [0, \tau] \times \mathbb{R}_+, \max_{l+2l' \leq q} |\partial_x^l \partial_t^{l'} u_{\theta, \tau}(t, x)| \leq C(1 + x^{m+q+l'}) \quad (4.25)$$

The proof of this proposition, mainly based on the analytical formula available for the transition density of the CIR process is made in the Appendix 4.7.

We are now able to prove the main results of this section:

Proposition 4.4.2. *Let $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ be a $\mathcal{C}^4$ function such that $\exists A, m > 0, \forall x \geq 0, |f^{(4)}(x)| \leq A(1 + x^m)$. Let us suppose moreover that the scheme $(\hat{X}^n)$ satisfies the hypothesis $(\mathcal{H}_W)$. Then, the weak error is in $1/n$:*

$$\mathbb{E}[f(\hat{X}_T^n)] = \mathbb{E}[f(X_T)] + O(1/n).$$

More generally, if $(f_\theta, \theta \in \Theta)$ is a family of $\mathcal{C}^4$ functions satisfying condition (4.24) for $q = 4$,

$$\mathbb{E}[f_\theta(\hat{X}_{t_j^n}^n)] = \mathbb{E}[f_\theta(X_{t_j^n})] + O(1/n)$$

where $O(1/n)$ has to be understood in the sense of Definition 4.2.2 with $(\theta, t_j^n) \in \Gamma_n = \Theta \times \{t_0^n, \dots, t_n^n\}$.

Proof : We have $\mathbb{E}[f(\hat{X}_T^n)] = \mathbb{E}[u(T, \hat{X}_T^n)]$ and $\mathbb{E}[f(X_T)] = u(0, x_0)$ so that:

$$\mathbb{E}[f(\hat{X}_T^n)] - \mathbb{E}[f(X_T)] = \mathbb{E}[u(T, \hat{X}_T^n) - u(0, x_0)] = \sum_{i=0}^{n-1} \mathbb{E}[u(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u(t_i, \hat{X}_{t_i}^n)].$$

Let us consider (t, x) and (s, y) in $[0, T] \times \mathbb{R}_+$. We can apply the Taylor formula to $t \mapsto u(t, y)$ up to order 2 and get:

$$u(s, y) = u(t, y) + (s - t)\partial_t u(t, y) + (s - t)^2 \int_0^1 (1 - \tau)\partial_t^2 u(t + \tau(s - t), y)d\tau.$$

Now, we apply Taylor formula to $y \mapsto u(t, y)$ and $y \mapsto \partial_t u(t, y)$ and we finally get

$$\begin{aligned} u(s, y) &= \sum_{0 \leq l+2l' < 4} \partial_x^l \partial_t^{l'} u(t, x) \frac{(s-t)^{l'}(y-x)^l}{l!l'!} + (s-t)^2 \int_0^1 (1-\tau)\partial_t^2 u(t + \tau(s-t), y)d\tau \\ &\quad + (s-t)(y-x)^2 \int_0^1 (1-\xi)\partial_x^2 \partial_t u(t, x + \xi(y-x))d\xi \\ &\quad + \frac{(y-x)^4}{3!} \int_0^1 (1-\xi)^3 \partial_x^4 u(t, x + \xi(y-x))d\xi. \end{aligned}$$

Proposition 4.4.1 allows us then to get:

$$\begin{aligned} \left| u(s, y) - \sum_{0 \leq l+2l' < 4} \partial_x^l \partial_t^{l'} u(t, x) \frac{(s-t)^{l'}(y-x)^l}{l!l'!} \right| \\ \leq C(1 + \max(x, y)^{6+m}) [(s-t)^2 + |s-t|(y-x)^2 + (y-x)^4] \end{aligned}$$

and we apply this bound to $(t_i, \hat{X}_{t_i}^n)$ and $(t_{i+1}, \hat{X}_{t_{i+1}}^n)$. Proposition 4.2.4 and Corollary 4.2.7 give immediately that $C(1 + \max(\hat{X}_{t_i}^n, \hat{X}_{t_{i+1}}^n)^{6+m})[(T/n)^2 + (\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^2 T/n + (\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^4] = \mathcal{O}(1/n^2)$ and therefore:

$$u(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u(t_i, \hat{X}_{t_i}^n) = \sum_{0 < l+2l' < 4} \partial_x^l \partial_t^{l'} u(t_i, \hat{X}_{t_i}^n) \frac{(T/n)^{l'}(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^l}{l!l'!} + \mathcal{O}(1/n^2). \quad (4.26)$$

Now we expand the powers of $(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)$ up to order 2 using the Hypothesis $(\mathcal{H}_W)$:

$$\begin{aligned} \hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n &= \frac{T}{n}(a - k\hat{X}_{t_i}^n) + \sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i}) + m_{t_{i+1}}^n - m_{t_i}^n + \mathcal{O}(1/n^2) \\ (\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^3 &= \sigma^3(\hat{X}_{t_i}^n)^{3/2}(W_{t_{i+1}} - W_{t_i})^3 + \mathcal{O}(1/n^2) \end{aligned}$$

Therefore, we get that

$$\begin{aligned} \mathbb{E}[\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n | \mathcal{F}_{t_i}] &= \frac{T}{n}(a - k\hat{X}_{t_i}^n) + \mathcal{O}(1/n^2) \\ \mathbb{E}[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^3 | \mathcal{F}_{t_i}] &= \mathcal{O}(1/n^2) \\ \text{and according to (4.10), } \mathbb{E}[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^2 | \mathcal{F}_{t_i}] &= \sigma^2 \hat{X}_{t_i}^n T/n + \mathcal{O}(1/n^2). \end{aligned}$$

The bound (4.22) and Lemma 4.2.6 ensure that $\partial_x^l \partial_t^{l'} u(t_i, \hat{X}_{t_i}^n) = \mathcal{O}(1)$ for $l + 2l' < 4$. Thus, using Lemma 4.2.5, we can deduce from (4.26) :

$$\begin{aligned} & \mathbb{E} \left[u(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u(t_i, \hat{X}_{t_i}^n) | \mathcal{F}_{t_i} \right] \\ &= \sum_{0 < l + 2l' < 4} \partial_x^l \partial_t^{l'} u(t_i, \hat{X}_{t_i}^n) \frac{(T/n)^{l'} \mathbb{E} \left[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^l | \mathcal{F}_{t_i} \right]}{l!l'!} + \mathcal{O}(1/n^2) \\ &= \partial_t u(t_i, \hat{X}_{t_i}^n) T/n + \partial_x u(t_i, \hat{X}_{t_i}^n) \frac{T}{n} (a - k \hat{X}_{t_i}^n) + \partial_x^2 u(t_i, \hat{X}_{t_i}^n) \frac{\sigma^2}{2} \hat{X}_{t_i}^n T/n + \mathcal{O}(1/n^2) \\ &= \mathcal{O}(1/n^2) \end{aligned}$$

since u solves the PDE (4.23). Therefore, there is a constant $C > 0$ that does not depend on i such that $\left| \mathbb{E} \left[u(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u(t_i, \hat{X}_{t_i}^n) \right] \right| \leq C/n^2$ and so, we finally get that

$$\left| \mathbb{E}[f(\hat{X}_T^n)] - \mathbb{E}[f(X_T)] \right| \leq C/n$$

which is the desired result.

Now let us explain why this proof can be generalized easily to the case of the family of functions f_θ that satisfy (4.24) and all times t_j^n . We apply as before Taylor formula to functions u_{θ, t_j^n} and thanks to Proposition 4.4.1, the bounds we have on its derivatives do not depend on (θ, t_j^n) and we get for $0 \leq i < j \leq n$ as in (4.26)

$$u_{\theta, t_j^n}(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u_{\theta, t_j^n}(t_i, \hat{X}_{t_i}^n) = \sum_{0 < l + 2l' < 4} \partial_x^l \partial_t^{l'} u_{\theta, t_j^n}(t_i, \hat{X}_{t_i}^n) \frac{(T/n)^{l'} (\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^l}{l!l'!} + \mathcal{O}(1/n^2)$$

with the difference that the $\mathcal{O}$ symbol is now meant with $\Gamma_n = \{(t_i^n, t_j^n), 0 \leq i < j \leq n\} \times \Theta$ instead of $\{t_i^n, 0 \leq i \leq n\}$ before. Then, the proof is the same, noticing that we still have $\partial_x^l \partial_t^{l'} u_{\theta, t_j^n}(t_i, \hat{X}_{t_i}^n) = \mathcal{O}(1)$ for $l + 2l' \leq q$. $\square$

Remark 4.4.3. *We desired to get a weak error in $1/n$ as in the case of the Euler scheme for stochastic differential equations with coefficients regular enough (C^4 and bounded derivatives). Using the argument of Talay and Tubaro, we need then a control on $u(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u(t_i, \hat{X}_{t_i}^n)$ up to order 2. This is why we assume to know the relation (4.9) between $\hat{X}_{t_{i+1}}^n$ and $\hat{X}_{t_i}^n$ up to order 2. Expanding $u(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u(t_i, \hat{X}_{t_i}^n)$, we see that the term of order $1/2$ has a null expectation, the term of order $1/n$ is null since u solves the PDE (4.23), but we need to require condition (4.10) so that the term of order $3/2$ has a null expectation. If we had only assumed that the scheme satisfies $(\mathcal{H}_S)$, we would have obtained a weak error in $1/\sqrt{n}$.*

Now, we would like to expand further the weak error, in particular to justify the use the Romberg method that mainly relies on the following remark: if we know that there is

$c_1 \in \mathbb{R}$ such that $\mathbb{E}[f(\hat{X}_T^n)] = \mathbb{E}[f(X_T)] + c_1/n + O(1/n^2)$, then $2\mathbb{E}[f(\hat{X}_T^{2n})] - \mathbb{E}[f(\hat{X}_T^n)] = \mathbb{E}[f(X_T)] + O(1/n^2)$ converges thus faster toward the desired expectation. If we want to adapt the previous proof, we see that we need to add the following assumptions to get a weak error up to order $\nu \in \mathbb{N}^*$:

- f is regular enough ($\mathcal{C}^{4\nu}$) and its derivatives have a polynomial growth.
- We know the relation between $\hat{X}_{t_{i+1}}^n$ and $\hat{X}_{t_i}^n$ up to order $\nu + 1$.

Moreover, if we wish to have as for the Euler scheme an error that expands only on the integer orders: $\mathbb{E}[f(\hat{X}_T^n)] = \mathbb{E}[f(X_T)] + c_1/n + c_2/n^2 + \dots + c_{\nu-1}/n^{\nu-1} + O(1/n^\nu)$, we need to make assumptions of the same kind as (4.10) for any power of $(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)$ to get terms of order “integer + one half” with null expectation. However these assumptions would be hardly readable, and practically, they would be clearly satisfied only by the explicit schemes $E(\lambda)$. That’s why we prefer to state here directly the result for the explicit schemes $E(\lambda)$.

Proposition 4.4.4. *Let $\nu \in \mathbb{N}^*$ and $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ that we suppose $\mathcal{C}^\infty$ and such that $\forall q, \exists A_q > 0, m_q \in \mathbb{N}, |f^{(q)}(x)| \leq A_q(1 + x^{m_q})$. Let $(\hat{X}^n)$ be the explicit scheme $E(\lambda)$ with $0 \leq \lambda \leq a - \sigma^2/4$. Then, the weak error has an expansion up to order ν :*

$$\mathbb{E}[f(\hat{X}_T^n)] = \mathbb{E}[f(X_T)] + c_1/n + c_2/n^2 + \dots + c_{\nu-1}/n^{\nu-1} + O(1/n^\nu)$$

where $c_1 = T \int_0^T \mathbb{E}[\psi_{E(\lambda)}(t, X_t)] dt$ with $\psi_{E(\lambda)}$ defined below in (4.28).

Proof : With the same argument as in Proposition 4.4.2, first using the Taylor expansion respect to t and then to x , we get that there is $C(\nu) > 0$ and $M(\nu) \in \mathbb{N}$:

$$\begin{aligned} & \left| u(s, y) - \sum_{0 \leq l+2l' < 2\nu+2} \partial_x^l \partial_t^{l'} u(t, x) \frac{(s-t)^{l'} (y-x)^l}{l!l'} \right| \\ & \leq C(\nu) (1 + \max(x, y)^{M(\nu)}) \sum_{j=0}^{\nu+1} |s-t|^{\nu+1-j} (y-x)^{2j}. \end{aligned}$$

Similarly, we get that

$$u(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u(t_i, \hat{X}_{t_i}^n) = \sum_{0 < l+2l' < 2\nu+2} \partial_x^l \partial_t^{l'} u(t_i, \hat{X}_{t_i}^n) \frac{(T/n)^{l'} (\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^l}{l!l'} + \mathcal{O}(1/n^{\nu+1}).$$

and then

$$\begin{aligned} \mathbb{E} \left[u(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u(t_i, \hat{X}_{t_i}^n) | \mathcal{F}_{t_i} \right] &= \sum_{0 < l+2l' < 2\nu+2} \partial_x^l \partial_t^{l'} u(t_i, \hat{X}_{t_i}^n) \frac{(T/n)^{l'} \mathbb{E} \left[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^l | \mathcal{F}_{t_i} \right]}{l!l'} \\ &+ \mathcal{O}(1/n^{\nu+1}). \end{aligned}$$

Let us first expand (4.5) to get:

$$\begin{aligned}\hat{X}_{t_{i+1}}^n &= \hat{X}_{t_i}^n + \sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i}) + (a - k\hat{X}_{t_i}^n)\frac{T}{n} + \left(\lambda + \frac{\sigma^2}{4}\right)((W_{t_{i+1}} - W_{t_i})^2 - T/n) \\ &\quad + \frac{k^2}{4}\hat{X}_{t_i}^n(T/n)^2 + \frac{\sigma^2}{4}(W_{t_{i+1}} - W_{t_i})^2 \left(\frac{1}{(1 - \frac{kT}{2n})^2} - 1\right).\end{aligned}$$

Since $\frac{1}{(1 - \frac{kT}{2n})^2} - 1 = \sum_{j \geq 1} (j+1)(k/2)^j(T/n)^j$, we get that

$$\begin{aligned}\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n &= \sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i}) + (a - k\hat{X}_{t_i}^n)\frac{T}{n} + \left(\lambda + \frac{\sigma^2}{4}\right)((W_{t_{i+1}} - W_{t_i})^2 - T/n) \\ &\quad + \frac{k^2}{4}\hat{X}_{t_i}^n(T/n)^2 + \frac{\sigma^2}{4}(W_{t_{i+1}} - W_{t_i})^2 \sum_{j=1}^{\nu-1} (j+1)(k/2)^j(T/n)^j + \mathcal{O}(1/n^{\nu+1}).\end{aligned}$$

All the terms here are of integer order but $\sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i})$ that is of order $1/2$. Now, taking the power l of these expansion, we get using Proposition 4.2.4 an expansion of $(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^l$ up to order $\nu+1$ (even $\nu+1+(l-1)/2$). What is important to remark is that the term of order “integer + one half” comes from an odd power of $\sigma\sqrt{\hat{X}_{t_i}^n}(W_{t_{i+1}} - W_{t_i})$ and a product of the other terms. Since all these other terms are even respect to $(W_{t_{i+1}} - W_{t_i})$, we finally get that all the terms of order “integer + one half” have a null conditional expectation. Thus, we see that we can write for $l \in \mathbb{N}$

$$\mathbb{E} \left[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^l | \mathcal{F}_{t_i} \right] = \sum_{j \geq l/2}^{\nu} \phi^{l,j}(\hat{X}_{t_i}^n)(T/n)^j + \mathcal{O}(1/n^{\nu+1})$$

where $\phi^{l,j}$ are polynomial functions that we do not explicit and satisfy $\phi^{l,j}(\hat{X}_{t_i}^n) = \mathcal{O}(1)$.

Thus, $(T/n)^{l'} \mathbb{E} \left[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^l | \mathcal{F}_{t_i} \right] = \sum_{l/2 \leq j \leq \nu} \phi^{l,j}(\hat{X}_{t_i}^n)(T/n)^{j+l'} + \mathcal{O}(1/n^{\nu+1})$

$= \sum_{l+2l' \leq 2j < 2\nu+2} \phi^{l,j-l'}(\hat{X}_{t_i}^n)(T/n)^j + \mathcal{O}(1/n^{\nu+1})$ and so, $\mathbb{E} \left[u(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u(t_i, \hat{X}_{t_i}^n) | \mathcal{F}_{t_i} \right] =$

$$\sum_{j=1}^{\nu} (T/n)^j \left(\sum_{0 < l+2l' \leq 2j} \partial_x^l \partial_t^{l'} u(t_i, \hat{X}_{t_i}^n) \frac{\phi^{l,j-l'}(\hat{X}_{t_i}^n)}{l!l'!} \right) + \mathcal{O}(1/n^{\nu+1}). \quad (4.27)$$

For $\nu = 2$, one obtains:

$$\begin{aligned}
\mathbb{E} \left[\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n | \mathcal{F}_{t_i} \right] &= (a - k\hat{X}_{t_i}^n)T/n + \frac{k^2\hat{X}_{t_i}^n + k\sigma^2}{4}(T/n)^2 + \mathcal{O}(1/n^3) \\
\mathbb{E} \left[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^2 | \mathcal{F}_{t_i} \right] &= \sigma^2\hat{X}_{t_i}^n T/n + \left[(a - k\hat{X}_{t_i}^n)^2 + 2(\lambda + \sigma^2/4)^2 \right] (T/n)^2 + \mathcal{O}(1/n^3) \\
\mathbb{E} \left[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^3 | \mathcal{F}_{t_i} \right] &= 3\sigma^2\hat{X}_{t_i}^n \left[a - k\hat{X}_{t_i}^n + 2(\lambda + \sigma^2/4) \right] (T/n)^2 + \mathcal{O}(1/n^3) \\
\mathbb{E} \left[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^4 | \mathcal{F}_{t_i} \right] &= 3\sigma^4(\hat{X}_{t_i}^n)^2(T/n)^2 + \mathcal{O}(1/n^3) \\
\mathbb{E} \left[(\hat{X}_{t_{i+1}}^n - \hat{X}_{t_i}^n)^5 | \mathcal{F}_{t_i} \right] &= \mathcal{O}(1/n^3)
\end{aligned}$$

and so:

$$\mathbb{E} \left[u(t_{i+1}, \hat{X}_{t_{i+1}}^n) - u(t_i, \hat{X}_{t_i}^n) | \mathcal{F}_{t_i} \right] = (T/n)^2 \psi_{E(\lambda)}(t_i, \hat{X}_{t_i}^n) + \mathcal{O}(1/n^3)$$

where

$$\begin{aligned}
\psi_{E(\lambda)}(t, x) &= \frac{1}{2} \partial_t^2 u(t, x) + \frac{k^2 x + k\sigma^2}{4} \partial_x u(t, x) + (a - kx) \partial_x \partial_t u(t, x) \\
&+ \frac{1}{2} \left[(a - kx)^2 + 2(\lambda + \sigma^2/4)^2 \right] \partial_x^2 u(t, x) + \frac{\sigma^2}{2} x \partial_x^2 \partial_t u(t, x) \\
&+ \frac{\sigma^2}{2} x \left[a - kx + 2(\lambda + \sigma^2/4) \right] \partial_x^3 u(t, x) + \frac{\sigma^4}{8} x^2 \partial_x^4 u(t, x).
\end{aligned} \tag{4.28}$$

Therefore, summing and taking the expectation, we get that $\mathbb{E}[f(\hat{X}_T^n)] = \mathbb{E}[f(X_T)] + (T/n)^2 \sum_{i=0}^{n-1} \mathbb{E}[\psi_{E(\lambda)}(t_i, \hat{X}_{t_i}^n)] + \mathcal{O}(1/n^2)$. We then apply Proposition 4.4.2 to the family of functions $x \mapsto \psi_{E(\lambda)}(t_i, x)$ which satisfies condition (4.24) thanks to (4.22) (we incidentally remark that it is sufficient to have $f \in \mathcal{C}^8$ to get the expansion with $\nu = 2$). It gives that $\mathbb{E}[\psi_{E(\lambda)}(t_i, \hat{X}_{t_i}^n)] = \mathbb{E}[\psi_{E(\lambda)}(t_i, X_{t_i})] + \mathcal{O}(1/n)$. Since $\partial_t \mathbb{E}[\psi_{E(\lambda)}(t, X_t)]$ is bounded on $[0, T]$, we have that $(T/n) \sum_{i=0}^{n-1} \mathbb{E}[\psi_{E(\lambda)}(t_i, X_{t_i})] = \int_0^T \mathbb{E}[\psi_{E(\lambda)}(t, X_t)] dt + \mathcal{O}(1/n)$ and then:

$$\mathbb{E}[f(\hat{X}_T^n)] = \mathbb{E}[f(X_T)] + (T/n) \int_0^T \mathbb{E}[\psi_{E(\lambda)}(t, X_t)] dt + \mathcal{O}(1/n^2). \tag{4.29}$$

To get the expansion for $\nu = 3$ and further, one has to check by induction the desired result for any ν using the same methodology. $\square$

4.5 Numerical results

In this section, we will analyze numerically the convergence of the discretization schemes. For the theoretical study, an interesting feature of the implicit schemes (4.3) and (4.4)

and of the explicit schemes $E(\lambda)$, is their “automatic” nonnegativity for the following parameters:

Scheme	Condition on (a, σ)
Implicit (4.3)	$\sigma^2 \leq 2a$
Implicit (4.4)	$\sigma^2 \leq 4a$
$E(\lambda)$	$0 \leq \lambda \leq a - \sigma^2/4$

(4.30)

Indeed, contrary to the schemes using a reflection technique as those proposed by Deelstra-Delbaen or Diop, there is no need to control the reflection. However, we can use the following trick to extend schemes (4.3), (4.4) and $E(\lambda)$ to all the values of the parameters (k, a, σ) :

- For the implicit schemes which are defined with second-degree polynomials, we will set $\hat{X}_{t_{i+1}}^n = 0$ when the discriminant is negative and else use formulas (4.3) and (4.4).
- For the explicit schemes $E(\lambda)$, we simply define $\hat{X}_{t_{i+1}}^n$ as the positive part of the left-hand side of (4.5)

We will use these extensions when needed for the simulations presented in this section.

4.5.1 Numerical study of the strong convergence

In this paragraph we present a numerical analysis of the strong convergence of various schemes. It does not seem possible to compute the limit process on the same probability space, and we overcome this difficulty using the following lemma that says that it is sufficient to study the difference between the values obtained with a scheme for a given time step and the ones obtained with the same scheme and a time step twice smaller. Let us recall here that $t_n^i = iT/n = t_{2n}^{2i}$.

Lemma 4.5.1. *Let us consider a scheme $(\hat{X}_{t_i}^n)$ that converges toward a continuous process X_t in the following sense:*

$$\mathbb{E} \left[\sup_{0 \leq i \leq n} |\hat{X}_{t_i^n}^n - X_{t_i^n}| \right] \xrightarrow{n \rightarrow \infty} 0. \quad (4.31)$$

Then, for any $\alpha > 0$ and $\beta \geq 0$,

$$\mathbb{E} \left[\sup_{0 \leq i \leq n} |\hat{X}_{t_i^n}^n - X_{t_i^n}| \right] = O\left(\frac{(\ln n)^\beta}{n^\alpha}\right) \iff \mathbb{E} \left[\sup_{0 \leq i \leq n} |\hat{X}_{t_i^n}^n - \hat{X}_{t_{2i}^{2n}}^{2n}| \right] = O\left(\frac{(\ln n)^\beta}{n^\alpha}\right).$$

The condition (4.31) has been established in this paper for the explicit schemes $E(\lambda)$ and for the implicit scheme (4.3) and it has also been proved for the scheme of Deelstra-Delbaen [27]. Under some restrictive conditions of the parameters, the scheme proposed by Diop converges with a $1/\sqrt{n}$ rate [28]. For the other parameters and for the Implicit

scheme on the square-root (4.4), we can check numerically the condition (4.31) doing the comparison with a scheme on which this comparison has been proved.

Proof of the Lemma. If there is $K > 0$ such that $\forall n \in \mathbb{N}^*, \mathbb{E} \left[\sup_{0 \leq i \leq n} |\hat{X}_{t_i^n}^n - X_{t_i^n}| \right] \leq K \frac{(\ln n)^\beta}{n^\alpha}$, then

$$\mathbb{E} \left[\sup_{0 \leq i \leq n} |\hat{X}_{t_i^n}^n - \hat{X}_{t_{2i}^{2n}}^{2n}| \right] \leq K \left(\frac{(\ln n)^\beta}{n^\alpha} + \frac{(\ln n + \ln 2)^\beta}{(2n)^\alpha} \right) \leq K' \frac{(\ln n)^\beta}{n^\alpha}.$$

Reciprocally, since $\sup_{0 \leq i \leq n} |\hat{X}_{t_i^n}^n - X_{t_i^n}| \leq \sum_{k=0}^l \sup_{0 \leq i \leq n} |\hat{X}_{t_{2^k i}^{2^k n}}^{2^k n} - \hat{X}_{t_{2^{k+1} i}^{2^{k+1} n}}^{2^{k+1} n}| + \sup_{0 \leq i \leq n} |\hat{X}_{t_{2^{l+1} i}^{2^{l+1} n}}^{2^{l+1} n} - X_{t_i^n}|$, we get $\mathbb{E} \left[\sup_{0 \leq i \leq n} |\hat{X}_{t_i^n}^n - X_{t_i^n}| \right] \leq K \sum_{k=0}^l \frac{(\ln n + k \ln 2)^\beta}{(2^k n)^\alpha} + \mathbb{E} \left[\sup_{0 \leq i \leq n} |\hat{X}_{t_{2^{l+1} i}^{2^{l+1} n}}^{2^{l+1} n} - X_{t_i^n}| \right]$ and with $l \rightarrow \infty$,

$$\mathbb{E} \left[\sup_{0 \leq i \leq n} |\hat{X}_{t_i^n}^n - X_{t_i^n}| \right] \leq K \sum_{k=0}^{\infty} C_\beta \frac{(\ln n)^\beta + (k \ln 2)^\beta}{(2^k n)^\alpha} \leq K' \frac{(\ln n)^\beta}{n^\alpha}$$

for some constant $K' > 0$, using that $\sum_{k=0}^{\infty} k^\beta / 2^k < \infty$. $\square$

Now for the numerical study, we consider the standard time interval $[0, 1]$ ($T = 1$) and set $S_n = \mathbb{E} \left[\sup_{0 \leq i \leq n} |\hat{X}_{t_i^n}^n - \hat{X}_{t_{2i}^{2n}}^{2n}| \right]$. The figures below show the convergence of S_n in function of the time-step $1/n$ for different parameters. Let us first observe that the implicit scheme (4.4) and the explicit scheme $E(0)$ give errors smaller than the others, for all the values of the parameters tested. Which is also interesting and nontrivial is that the behaviour of the convergence depends on the parameters.

We notice that for the case $2a > \sigma^2$ the schemes (4.4) and $E(0)$ present an error which looks linear respect to the time-step while the others give a square-root shape (see Fig. 4.1). This is not totally surprising because we have seen that these schemes correspond to the Milstein expansion, and we also know that under this hypothesis, X_t never reaches 0 so that the non-lipschitzian behaviour of the square-root is less important.

When $2a < \sigma^2 < 4a$, the schemes (4.3), $E(\sigma^2/8)$, $E(\sigma^2/4)$, Deelstra-Delbaen and Diop, S_n still has a square-root behaviour (see Fig. 4.2). Finally, let us mention that for the last case $\sigma^2 > 4a$, the schemes (4.4) and $E(0)$ still give the smaller value of S_n . However, we have to say that when $\sigma^2 \gg 4a$, the convergence is really slow.

Lastly, concerning the impact of λ for the explicit schemes $E(\lambda)$, we see (Fig. 4.1 and 4.2) that $\lambda = \sigma^2/4$ is the parameter that gives a strong convergence analogous to the schemes of Diop, Deelstra-Delbaen and implicit (4.3); and the value of S_n for $E(\sigma^2/8)$ is as one can expect between those of $E(0)$ and $E(\sigma^2/4)$.

To get an idea of the speed of convergence in function of the parameters, we postulate that $S_n \sim C/n^\alpha$ with $\alpha > 0$. Thanks to the lemma, this is equivalent to a strong convergence speed in $1/n^\alpha$. To estimate α , we remark that

$$\log_{10}(S_n) - \log_{10}(S_{10n}) \xrightarrow{n \rightarrow +\infty} \alpha,$$

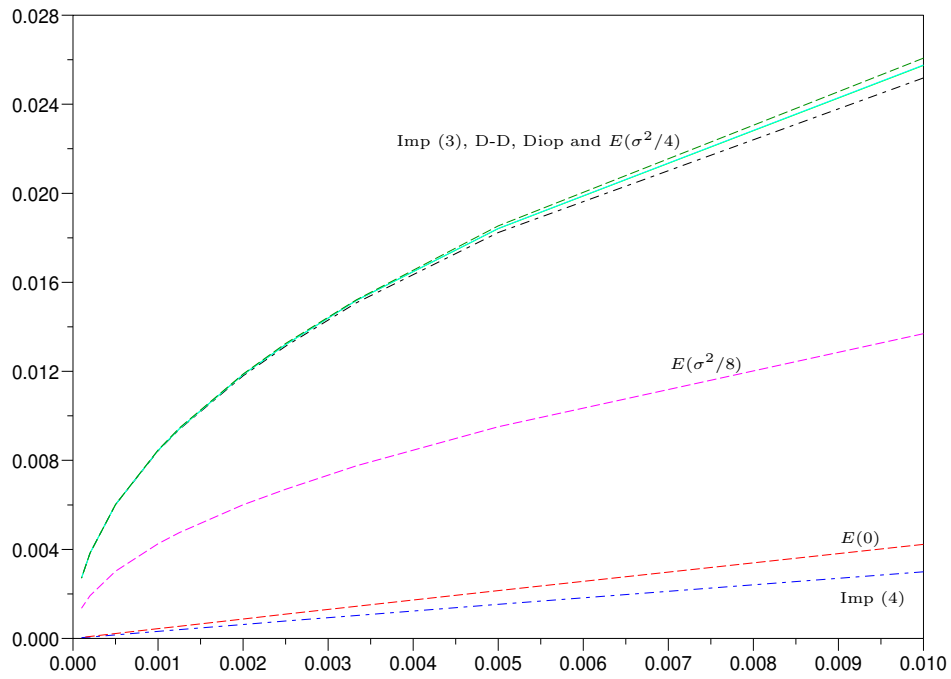


Figure 4.1: S_n in function of the time-step $1/n$ for $x_0 = 1$, $k = 1$, $a = 1$ and $\sigma = 1$.

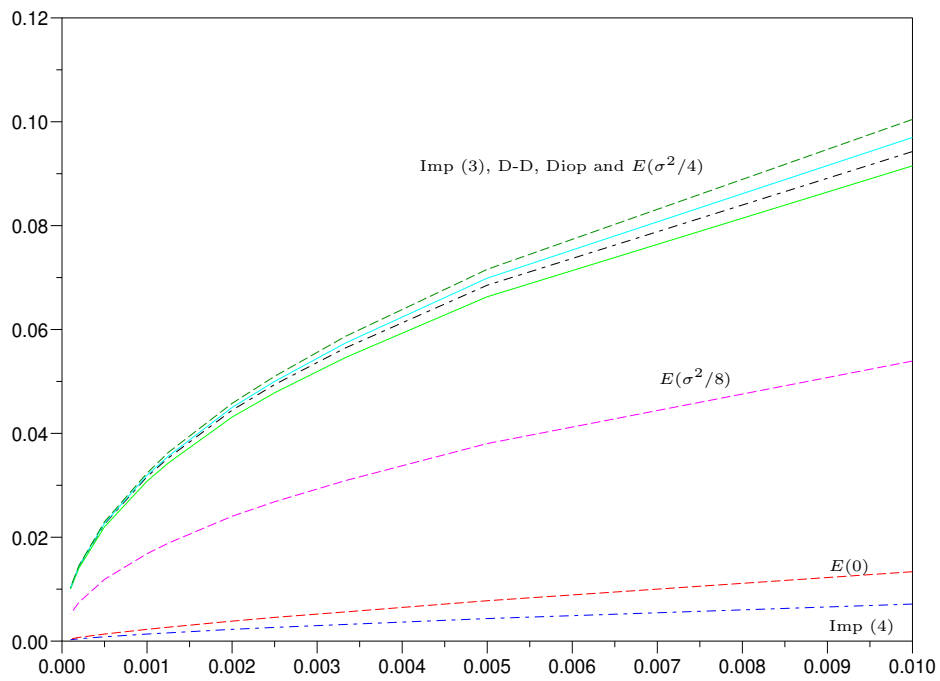


Figure 4.2: S_n in function of the time-step $1/n$ for $x_0 = 1$, $k = 1$, $a = 1$ and $\sigma = \sqrt{3}$.

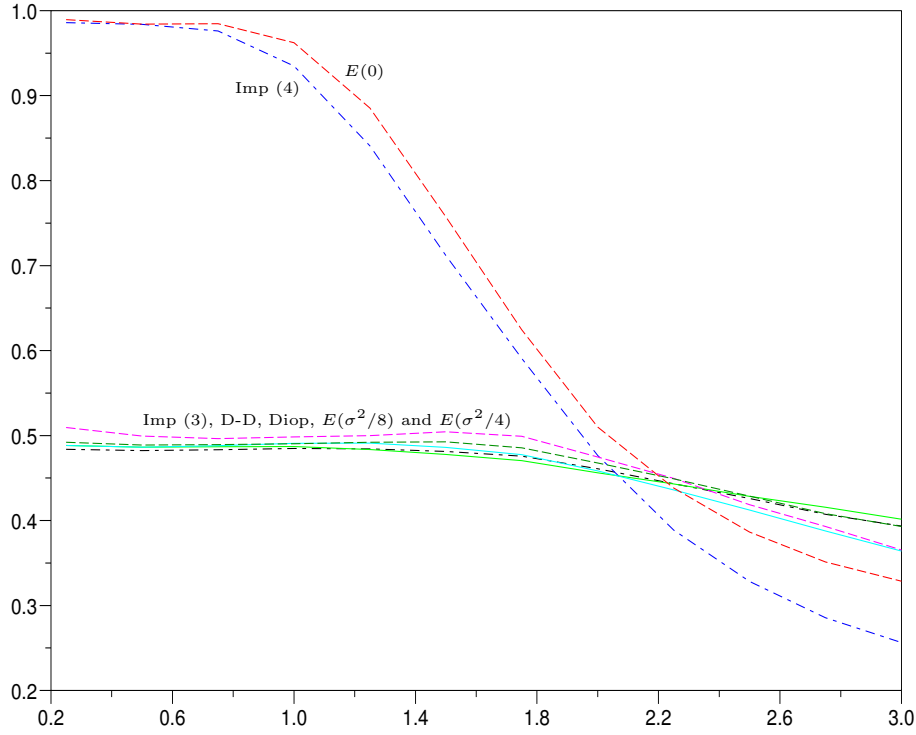


Figure 4.3: Speed convergence of S_n : estimation of the α parameter in function of $\sigma^2/(2a)$ for $x_0 = 1$, $k = 1$ and $a = 1$.

and we have reported $\log_{10}(S_n) - \log_{10}(S_{10n})$ for $n = 200$ in Figure 4.3. We have plotted the result in function of the parameter $\sigma^2/(2a)$ since it is the one that plays a key role. This can be understood easily with a time-scaling. For the schemes (4.4) and $E(0)$, the estimated α is close to 1 for $\sigma^2 < 2a$ and decreases from 1 to $1/2$ for $2a < \sigma^2 < 4a$ while for the other schemes, the estimated value of α is close to $1/2$ for $\sigma^2 < 4a$. Intuitively, we can understand this decrease because for $\sigma^2 > 2a$, X_t can reach the origin, and a non negligible time is spent in the neighbourhood of 0 where the square root is non Lipschitz. Obviously, the speed of convergence may have a more complicated form than the one postulated, but our method gives nonetheless a good idea of its behaviour.

4.5.2 Numerical study of the weak convergence

We have plotted in figures 4.4, 4.5, 4.6 and 4.7, for fixed parameters of the CIR process, the approximation given by the scheme or a Romberg extrapolation of the expected value

$\mathbb{E}[f(X_1)]$ for the function $f(x) = \frac{5+3x^4}{2+5x}$. This function has been chosen to be sensitive to variation for large and small values so that it catches the defaults of the schemes near 0 and ∞ . We have taken two sets of parameters that illustrate the cases $\sigma^2 \leq 2a$ and $2a \leq \sigma^2 \leq 4a$.

Let us recall here that we have proved here the $O(1/n)$ convergence only for regular functions and for the schemes satisfying $(\mathcal{H}_W)$, that is (4.3) with $\sigma^2 < 2a$ and $E(\lambda)$ with $0 \leq \lambda \leq a - \sigma^2/4$. What comes out from the computations (see Figures 4.4 and 4.5) is that for the small values of σ ($\sigma^2 \leq 2a$, Fig. 4.4) all the schemes seem to have a behaviour in $O(1/n)$ while for the large values ($\sigma^2 > 2a$, Fig. 4.5), only the Explicit schemes and the Deelstra-Delbaen scheme give shapes compatible with a behaviour in $O(1/n)$. On the contrary, the scheme of Diop shows clearly a root shape while the implicit schemes (4.3) and (4.4) seem to converge a little bit slower than K/n .

Concerning the Romberg method to calculate $\mathbb{E}(f(\hat{X}_1))$, the figure 4.6 show that in the both cases $\sigma^2 \leq 2a$ and $\sigma^2 > 2a$, Diop's and implicit schemes (4.3) and (4.4) do not show a quadratic convergence. As expected, Explicit schemes have a quadratic shape in all the cases even if, strictly speaking, we have not proved the speed convergence observed for $E(\sigma^2/8)$, $k = 1$, $a = 1$ and $\sigma = \sqrt{3}$ since $\lambda = \sigma^2/8 > a - \sigma^2/4$. Concerning the Deelstra-Delbaen scheme, let us first say that for large time-steps, negative values may be frequent which explains the strange behaviour observed. However, for time-steps small enough, the convergence seems compatible with a quadratic convergence.

4.5.3 Computation time required by the schemes

In this paragraph we compare the time required by the schemes and the exact method to simulate 10^6 paths with a time step equal to 10^{-3} on the time interval $[0, 1]$ (see Table 4.1). Concerning the exact simulation of the increment of the CIR process, we have used the method proposed by Glasserman in [37] (see p. 120-128). As we could expect, this method is more time-consuming (up to a factor 10). Thus, it should be used to compute expectations that depend on the values of the process (X_t) at a few fixed times. On the contrary, for expectations that depends on all the path (such as integrals), discretization schemes should be preferred. As we see in Table 4.1, the time required by the schemes presented are of the same order. Let us mention here that for the implicit scheme (4.4), one has to be careful and store at each step the value of $\sqrt{\hat{X}_{t_i}^n}$ so that only one square-root has to be computed at each time step.

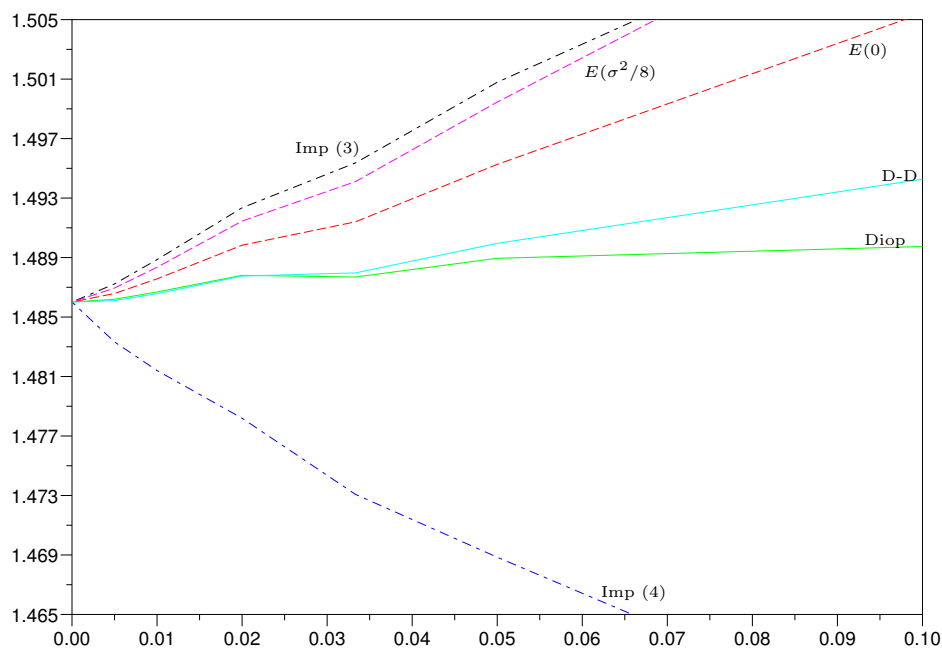


Figure 4.4: $\mathbb{E}(f(\hat{X}_1^{2n}))$ in function of $1/n$ with $f(x) = \frac{5+3x^4}{2+5x}$ for $x_0 = 0$, $k = 1$, $a = 1$ and $\sigma = 1$.

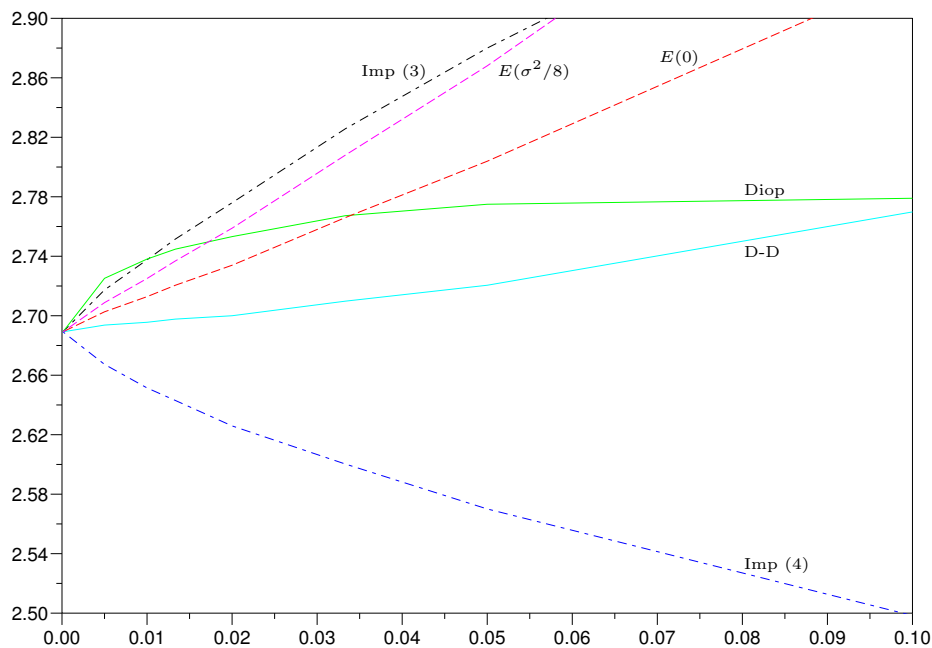


Figure 4.5: $\mathbb{E}(f(\hat{X}_1^{2n}))$ in function of $1/n$ with $f(x) = \frac{5+3x^4}{2+5x}$ for $x_0 = 0$, $k = 1$, $a = 1$ and $\sigma = \sqrt{3}$.

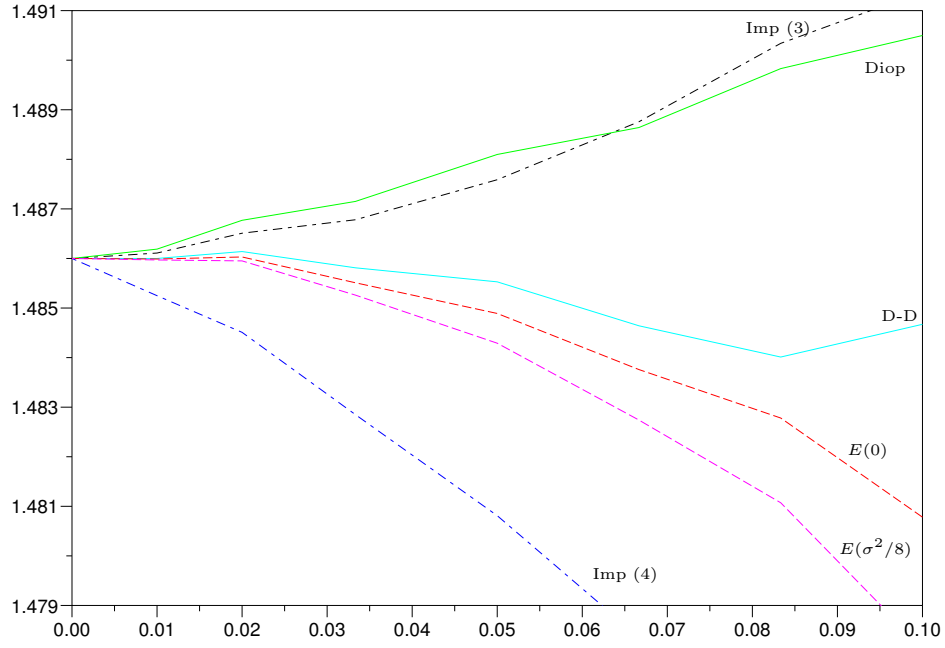


Figure 4.6: $2\mathbb{E}(f(\hat{X}_1^{2n})) - \mathbb{E}(f(\hat{X}_1^n))$ in function of $1/n$ with $f(x) = (5+3x^4)/(2+5x)$ for $x_0 = 0$, $k = 1$, $a = 1$ and $\sigma = 1$.

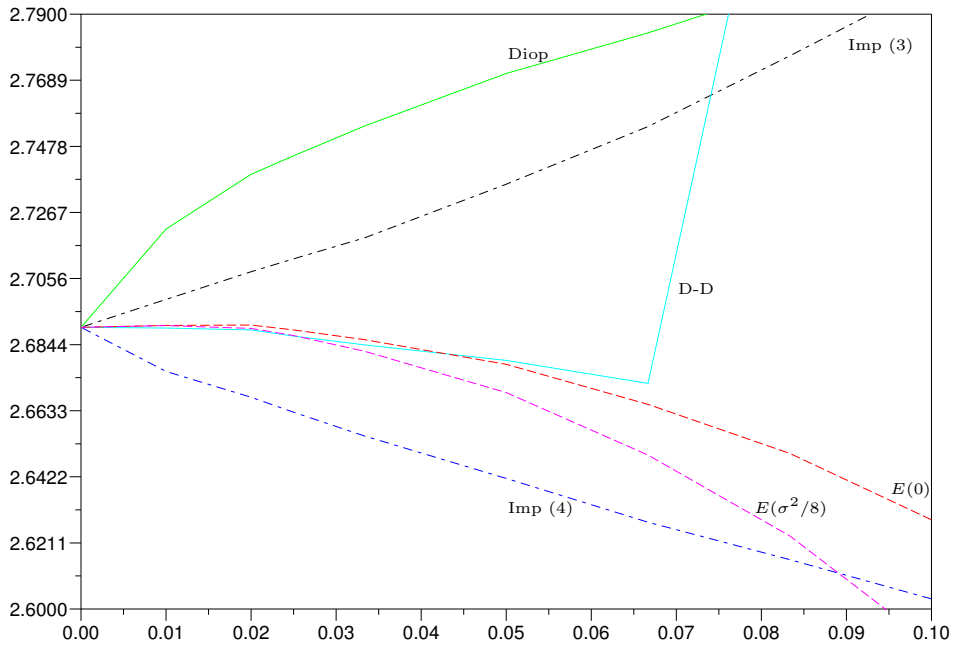


Figure 4.7: $2\mathbb{E}(f(\hat{X}_1^{2n})) - \mathbb{E}(f(\hat{X}_1^n))$ in function of $1/n$ with $f(x) = (5+3x^4)/(2+5x)$ for $x_0 = 0$, $k = 1$, $a = 1$ and $\sigma = \sqrt{3}$.

	Implicit (4.3)	Implicit (4.4)	Diop	Deelstra-Delbaen	$E(0)$	$E(\sigma^2/4)$	Exact
$\sigma = 1$	72	64	65	67	67	68	668
$\sigma = \sqrt{2}$	77	64	67	67	66	70	1092

Table 4.1: Simulation time (in s) for 10^6 paths with a time step equal to 10^{-3} and parameters $k = 1$, $a = 1$ and $x_0 = 1$.

	Implicit (4.3), $\sigma^2 \leq 2a$	Implicit (4.4), $\sigma^2 \leq 4a$	Diop	Deelstra Delbaen	$E(0)$, $\sigma^2 \leq 4a$	$E(\lambda)$, $0 < \lambda, \lambda \leq a - \sigma^2/4$
Nonnegativity	Y	Y	Y	N	Y	Y
Monotonicity	Y	Y	N	N	N	N
Strong CV	Y	?	Y*	Y	Y	Y
Weak CV rate in $1/n$	Y	?	Y*	?	Y	Y
Weak error expansion	?	?	?	?	Y	Y

Table 4.2: Theoretical results

4.6 Conclusion

We have sum up in Table 4.2 the theoretical results obtained in this paper and those of Diop, Deelstra and Delbaen [28, 27]. We first point out which scheme satisfy the algebraic properties of positivity and monotonicity. Then, we examine among the several schemes whether it has been proved

- a result of strong convergence,
- a weak convergence rate in $1/n$,
- an expansion of the weak error along the powers of $1/n$.

The star (Y*) means that the result has been established under some assumption on the parameters while the question mark indicates that no result has been shown yet. Let us mention here that Diop in [28] has also obtained a strong convergence speed in $1/\sqrt{n}$ under some restrictive conditions on parameters. Table 4.3 presents the results of the numerical tests of Section 4.5.

All these results tend to show that the explicit scheme $E(0)$ is the one that gathers the most interesting properties. Moreover, it is really easy to implement and is not more time consuming than the other schemes. That is why in the general case, it is recommended to use this scheme, at least for $\sigma^2 \leq 4a$.

		Implicit (4.3)	Implicit (4.4)	Diop	Deelstra Delbaen	$E(0)$	$E(\lambda), 0 < \lambda, \frac{\sigma^2}{\lambda} \leq a - \sigma^2/4$
$\sigma^2 \in [0, 2a]$	Strong CV order	$\approx 1/2$	≈ 1	$\approx 1/2$	$\approx 1/2$	≈ 1	$\approx 1/2$
	Weak CV rate in $1/n$	Y	Y	Y	Y	Y	Y
	Romberg in $1/n^2$	N	N	N	Y	Y	Y
$\sigma^2 \in [2a, 4a]$	Strong CV order	$\approx 1/2$	$\gtrsim 1/2$	$\approx 1/2$	$\approx 1/2$	$\gtrsim 1/2$	$\approx 1/2$
	Weak CV rate in $1/n$	?	?	N	Y	Y	Y
	Romberg in $1/n^2$	N	N	N	Y?	Y	Y

Table 4.3: Numerical results

As a further work, it would be interesting to get an accurate mathematical study on the dependence of the strong convergence of $E(0)$ on $\frac{\sigma^2}{2a}$ (see Fig. 4.3). It would be also interesting to study the behaviour of the convergence of the various schemes for large values of σ , ($\sigma^2 \geq 4a$). Since none of the scheme studied in this paper seems to be efficient for these large values of σ , designing a relevant scheme appears to be an interesting challenge. Lastly, in a different direction, it would be nice to relax the condition of regularity on f for the weak error and prove estimates on the cumulated distribution function and the density of X_T , as in Bally and Talay [7, 8] or more recently Guyon [38].

Acknowledgement. I am grateful to Benjamin Jourdain (ENPC-CERMICS) for his numerous and helpful comments. I also thank Chalinène Bassinah (Paris 13-Institut Galilée) for having double checked some numerical results.

4.7 Appendix : Proof of the Proposition 4.4.1

We will focus for sake of simplicity on the case of one function and one time T before explaining how to extend the results to the case of a family of functions that satisfy (4.24).

We will first prove

$$\max_{0 \leq l \leq q} |\partial_x^l u(t, x)| \leq C(1 + x^{q+m}) \quad (4.32)$$

for some constant $C > 0$, and then (4.23), so that (4.22) will outcome automatically by an induction on l' , using that for $l' \geq 1$ such that $l + 2l' \leq q$,

$$\begin{aligned} \partial_x^l \partial_t^{l'} u(t, x) &= -\partial_x^l \left((a - kx) \partial_x \partial_t^{l'-1} u(t, x) + \frac{\sigma^2}{2} x \partial_x^2 \partial_t^{l'-1} u(t, x) \right) \\ &= -\frac{\sigma^2}{2} x \partial_x^{l+2} \partial_t^{l'-1} u(t, x) - (l \frac{\sigma^2}{2} + a - kx) \partial_x^{l+1} \partial_t^{l'-1} u(t, x) + lk \partial_x^l \partial_t^{l'-1} u(t, x). \end{aligned}$$

Let us set $\tilde{u}(t, x) = u(T - t, x) = \mathbb{E}(f(X_t^x))$. By Lemma 4.2.1, (4.32) holds for $f(x) = x^p$ ($p \in \mathbb{N}$) and therefore for any polynomial. Now, using the decomposition $f(x) = f(x) - P(x) + P(x)$ with $P(x) = \sum_{l=0}^q f^{(l)}(0)x^l/l!$, we deduce that it is enough to prove (4.32) for $f \in \mathcal{C}^q$ such that $|f(x)| \leq A(1 + x^m)$ and $f^{(l)}(0) = 0$ for $l \leq q$.

Integrating successively, we get easily that $|f^{(l)}(x)| \leq A(1 + x^{m+q-l})$ and so, $\forall l \leq q$, $|f^{(l)}(x)| \leq A(1 + x^{m+q})$. The density of X_t^x is known and is given by:

$$p(t, x, z) = \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} \frac{c_t/2}{\Gamma(i + v/2)} \left(\frac{c_t z}{2}\right)^{i-1+v/2} e^{-c_t z/2}$$

where $c_t = \frac{4k}{\sigma^2(1-e^{-kt})}$, $v = 4a/\sigma^2$ and $\lambda_t = c_t e^{-kt}$. Let us remark here that

$$c_t \geq c_{\min} := \begin{cases} \frac{4k}{\sigma^2}, & k > 0 \\ \frac{4}{\sigma^2 T}, & k = 0 \\ \frac{4|k|}{\sigma^2(e^{|k|T}-1)}, & k < 0. \end{cases}$$

We have for $t > 0$:

$$\tilde{u}(t, x) = \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} I_i(f, c_t)$$

where

$$I_i(f, c_t) = \int_0^{\infty} f(z) \frac{c_t/2}{\Gamma(i + v/2)} \left(\frac{c_t z}{2}\right)^{i-1+v/2} e^{-c_t z/2} dz.$$

Since for $l \leq q$, $|f^{(l)}(z)| \leq A(1 + z^{m+q})$, we have

$$\forall i \in \mathbb{N}, |I_i(f^{(l)}, c_t)| \leq A \left(1 + \left(\frac{2}{c_t}\right)^{m+q} \frac{\Gamma(i + m + q + v/2)}{\Gamma(i + v/2)} \right). \quad (4.33)$$

Taking $l = 0$, the convergence of the above series is ensured. Derivating successively in x , we get that for $l \leq q$,

$$\forall t \in (0, T], x \in \mathbb{R}^+, \partial_x^l \tilde{u}(t, x) = \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} \Delta_t^l(I_i(f, c_t)) \quad (4.34)$$

where $\Delta_t : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ is an the operator defined on sequences $(I_i)_{i \geq 0} \in \mathbb{R}^{\mathbb{N}}$ by $\Delta_t(I_i) = \frac{\lambda_t}{2}(I_{i+1} - I_i) = \frac{e^{-kt}}{2} c_t (I_{i+1} - I_i)$. Let us remark now that, since $f^{(l-1)}(0) = 0$, an integration by part gives for $0 < l \leq q$ and $i \geq 1$

$$\begin{aligned} I_i(f^{(l)}, c_t) &= \int_0^{\infty} f^{(l-1)}(z) \frac{(c_t/2)^2}{\Gamma(i + v/2)} \left(\frac{c_t z}{2}\right)^{i-1+v/2} e^{-c_t z/2} dz \\ &\quad - \int_0^{\infty} f^{(l-1)}(z) \frac{(c_t/2)^2 (i-1+v/2)}{\Gamma(i + v/2)} \left(\frac{c_t z}{2}\right)^{i-2+v/2} e^{-c_t z/2} dz \\ &= \frac{c_t}{2} (I_i(f^{(l-1)}, c_t) - I_{i-1}(f^{(l-1)}, c_t)). \end{aligned}$$

Therefore, we get that $\Delta_t(I_i(f, c_t)) = e^{-kt} I_{i+1}(f^{(1)}, c_t)$ and finally:

$$\forall t \in (0, T], x \in \mathbb{R}^+, \partial_x^l \tilde{u}(t, x) = \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} I_{i+l}(f^{(l)}, c_t) e^{-klt}.$$

Using (4.33), it gives immediately that $|\partial_x^l \tilde{u}(t, x)| \leq A \left(1 + \frac{2^{m+q}}{c_t^{m+q}} \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} \frac{\Gamma(i+l+m+q+v/2)}{\Gamma(i+l+v/2)} \right)$.

The quotient $\frac{\Gamma(i+l+m+q+v/2)}{\Gamma(i+l+v/2)}$ is a polynomial of degree $m+q$ in i , and we note $\beta_0, \dots, \beta_{m+q}$ its coefficients in the basis $\{1, i, i(i-1), \dots, i(i-1) \cdots (i-(m+q)+1)\}$. Thus, we get that $|\partial_x^l \tilde{u}(t, x)| \leq A + \frac{A 2^{m+q}}{c_t^{m+q}} (\beta_0 + \beta_1 \lambda_t x + \dots + \beta_{m+q} (\lambda_t x)^{m+q})$ and since $|\lambda_t| \leq c_t e^{|k|T}$,

$$|\partial_x^l \tilde{u}(t, x)| \leq A + A e^{(m+q)|k|T} (|\beta_0|/c_{\min}^{m+q} + |\beta_1|/c_{\min}^{m+q-1} x + \dots + |\beta_{m+q}| x^{m+q}).$$

This allows us to conclude that there is a constant $C > 0$ (that depends only on A, T and the parameters (x_0, k, a, σ)) such that

$$\forall l \leq q, \forall t \in (0, T], x > 0, |\partial_x^l \tilde{u}(t, x)| \leq C(1 + x^{m+q}).$$

Proof of (4.23). We deduce from Lemma 4.2.1 that $\tilde{u}_0(T-t, x)$ and $\tilde{u}_1(T-t, x)$ solve the PDE (4.23) and it is therefore sufficient to prove the result for functions $f \in \mathcal{C}^2$ that satisfy $|f(x)| \leq A(x^2 + x^m)$. Let us now observe that $\frac{dc_t}{dt} = -\sigma^2 c_t \lambda_t / 4$ and $\frac{d\lambda_t}{dt} = -(\sigma^2 \lambda_t / 4 + k) \lambda_t$. Then, it is no hard to get $\frac{dI_i(f, c_t)}{dt} = (\sigma^2 i / 2 + a) \Delta_t(I_i(f, c_t))$ and that for any bounded sequence I_i , $\frac{d}{dt} \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} I_i = -(\frac{\sigma^2 \lambda_t}{4} + k) x \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} \Delta_t(I_i)$. Combining these results, we get using relation (4.34):

$$\begin{aligned} \partial_t \tilde{u}(t, x) &= -\left(\frac{\sigma^2 \lambda_t}{4} + k\right) x \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} \Delta_t(I_i(f, c_t)) \\ &\quad + \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} \left(\frac{\sigma^2 i}{2} + a\right) \Delta_t(I_i(f, c_t)) \\ &= (a - kx) \partial_x \tilde{u}(t, x) + \frac{\sigma^2}{2} \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} \left(-\frac{\lambda_t x}{2} + i\right) \Delta_t(I_i(f, c_t)) \\ &= (a - kx) \partial_x \tilde{u}(t, x) + \frac{\sigma^2}{2} x \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2} (\lambda_t x/2)^i}{i!} \frac{\lambda_t}{2} (\Delta_t(I_{i+1}(f, c_t)) - \Delta_t(I_i(f, c_t))) \\ &= (a - kx) \partial_x \tilde{u}(t, x) + \frac{\sigma^2}{2} x \partial_x^2 \tilde{u}(t, x). \end{aligned}$$

Finally, the continuity of f ensures that $\tilde{u}(t, x) = \mathbb{E}(f(X_t^x)) \rightarrow f(x)$ when $t \rightarrow 0$ thanks to Lebesgue's theorem.

Let us explain now how to extend the result to a family of functions f_θ and get (4.25). Let us denote $P_\theta(x) = \sum_{l=0}^q \frac{1}{l!} f_\theta^{(l)}(0) x^l$. Condition (4.24) ensures that the coefficients of P_θ are uniformly bounded in θ . Writing $f_\theta(x) = P_\theta(x) + (f_\theta(x) - P_\theta(x))$, one obtains (4.25) in the same way as (4.22). $\square$

Part III

Une dualité Call-Put pour les options
américaines perpétuelles, et
calibration de la volatilité.

Chapter 5

Une dualité Call-Put pour les options américaines perpétuelles

Ce chapitre provient d'un papier [4] écrit avec Benjamin Jourdain.

Abstract. It is well known [30], [5] that in models with time-homogeneous local volatility functions and constant interest and dividend rates, the European Put prices are transformed into European Call prices by the simultaneous exchanges of the interest and dividend rates and of the strike and spot price of the underlying. This paper investigates such a Call Put duality for perpetual American options. It turns out that the perpetual American Put price is equal to the perpetual American Call price in a model where, in addition to the previous exchanges between the spot price and the strike and between the interest and dividend rates, the local volatility function is modified. We prove that equality of the dual volatility functions only holds in the standard Black-Scholes model with constant volatility. Thanks to these duality results, we design a theoretical calibration procedure of the local volatility function from the perpetual Call and Put prices for a fixed spot price x_0 . The knowledge of the Put (resp. Call) prices for all strikes enables to recover the local volatility function on the interval $(0, x_0)$ (resp. $(x_0, +\infty)$).

Introduction

In a model with local volatility function $\varsigma(t, x)$, interest rate r and dividend rate δ

$$\begin{cases} dS_t^x = \varsigma(t, S_t^x) S_t^x dW_t + (r - \delta) S_t^x dt, & t \geq 0 \\ S_0^x = x \end{cases} \quad (5.1)$$

the initial price

$$h(T, y) = \mathbb{E} \left[e^{-rT} (y - S_T^x)^+ \right]$$

of the European Put option considered as a function of the maturity $T > 0$ and the Strike $y > 0$ solves Dupire's partial differential equation [30] :

$$\begin{cases} \partial_T h(T, y) = \frac{\varsigma^2(T, y)y^2}{2} \partial_{yy}^2 h(T, y) + (\delta - r)y \partial_y h(T, y) - \delta h(T, y), & T, y > 0 \\ h(0, y) = (y - x)^+, & y > 0 \end{cases}$$

One easily deduces that the function $h(T - t, y)$ for $(t, y) \in [0, T] \times \mathbb{R}_+^*$ satisfies the pricing partial differential equation for the Call option with strike x and maturity T in the model

$$\begin{cases} d\bar{S}_t^{y,T} = \varsigma(T - t, \bar{S}_t^{y,T}) \bar{S}_t^{y,T} dW_t + (\delta - r) \bar{S}_t^{y,T} dt, & t \in [0, T] \\ \bar{S}_0^{y,T} = y \end{cases} \quad (5.2)$$

with local volatility function $\varsigma(T - t, y)$, interest rate δ and dividend rate r . Therefore $h(T, y) = \mathbb{E} \left[e^{-\delta T} (\bar{S}_T^{y,T} - x)^+ \right]$ and one deduces the following Call-Put duality relation which is also a consequence of [5]

$$\forall T \geq 0, \forall x, y > 0, \mathbb{E} \left[e^{-rT} (y - S_T^x)^+ \right] = \mathbb{E} \left[e^{-\delta T} (\bar{S}_T^{y,T} - x)^+ \right].$$

Since it derives from Dupire's formula, this Call-Put duality equality is closely related to calibration issues. One remarks that in the particular case of a time-homogeneous volatility function ($\varsigma(t, x) = \sigma(x)$), then $\bar{S}_t^{y,T}$ also evolves according to the same time-homogeneous volatility function.

In this work, we are interested in deriving such a Call-Put duality relation in the case of American options and in investigating consequences in terms of calibration. In the Black-Scholes model with constant volatility $\varsigma(t, x) = \sigma$, when τ denotes a bounded stopping-time of the natural filtration of the Brownian motion $(W_t)_{t \geq 0}$, one has

$$\begin{aligned} \mathbb{E} \left[e^{-r\tau} \left(y - x e^{\sigma W_\tau + (r - \delta - \frac{\sigma^2}{2})\tau} \right)^+ \right] &= \mathbb{E} \left[e^{-\delta\tau} e^{\sigma W_\tau - \frac{\sigma^2}{2}\tau} \left(y e^{-\sigma W_\tau + (\delta - r + \frac{\sigma^2}{2})\tau} - x \right)^+ \right] \\ &= \mathbb{E} \left[e^{-\delta\tau} \left(y e^{-\sigma W_\tau + (\delta - r - \frac{\sigma^2}{2})\tau} - x \right)^+ \right] \end{aligned}$$

where the second equality follows from Girsanov theorem. Taking the supremum over all stopping-times τ smaller than T one deduces that the price of the American Put option with maturity T is equal to the price of the American Call option with the same maturity up to the simultaneous exchange between the underlying spot price and the strike and between the interest and dividend rates. Extensions of this result when the underlying evolves according to the exponential of a Lévy process have been obtained in [33]. But, to our knowledge, no study has been devoted to the case of models with local volatility functions like (5.1).

In the present paper, we consider the case of perpetual ($T = +\infty$) American options in models with time-homogeneous local volatility functions $\varsigma(t, x) = \sigma(x)$. In the first part, we recover well-known properties of the perpetual American call and put pricing functions by extending an approach recently developed by Beibel and Lerche [10] in the Black-Scholes case. This makes the paper self-contained.

In the second part, we introduce the framework used in the remaining of the paper.

In the third part of the paper, we consider the exercise boundaries as functions of the strike variable and characterize them as the unique solutions of some non-autonomous ordinary differential equations.

The fourth part is dedicated to our main result. We prove that the perpetual American Put prices are equal to the perpetual American Call prices in a model where, in addition to the exchanges between the spot price of the underlying and the strike and between the interest and dividend rates, the volatility function is modified. We also derive an expression of this modified volatility function. Notice that in the European case presented above, time-homogeneous volatility functions are not modified.

The fifth part addresses calibration issues. It turns out that for a given initial value $x_0 > 0$ of the underlying one recovers the restriction of the time-homogeneous volatility function $\sigma(x)$ to $(0, x_0]$ (resp. $[x_0, +\infty)$) from the perpetual Put (resp. Call) prices for all strikes.

In the last part, we show that at least when $\delta < r$, in the class of volatility functions analytic in a neighbourhood of the origin, the only ones invariant by our duality result are the constants. This means that the case of the standard Black-Scholes model presented above is very specific.

Acknowledgements. We thank Damien Lamberton (Univ. Marne-la-vallée) and Mihail Zervos (King's College) for interesting discussions. We also thank Alexander Schied (TU Berlin) for pointing out the work of Beibel and Lerche [10] to us and Antonino Zanette (University of Udine) for providing us with the routine that calculates American option prices.

5.1 Perpetual American put and call pricing

We consider a constant interest spot-rate r that is assumed to be nonnegative and an asset S_t which pays a constant dividend rate $\delta \geq 0$ and is driven by a homogeneous volatility function $\sigma : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ that satisfies the following hypothesis.

Hypothesis ($\mathcal{H}_{\text{vol}}$): σ is continuous on $\mathbb{R}_+^*$ and there are $0 < \underline{\sigma} < \bar{\sigma} < +\infty$ such that:

$$\forall x > 0, \underline{\sigma} \leq \sigma(x) \leq \bar{\sigma}.$$

In other words, S_t is assumed to follow under the risk-neutral measure the SDE:

$$dS_t = S_t((r - \delta)dt + \sigma(S_t)dW_t). \quad (5.3)$$

With the assumption made on σ , we know that for any initial condition $x \in \mathbb{R}_+^*$, there is a unique solution in the sense of probability law (see for example Theorem 5.15 in [47], using a log transformation) denoted by $(S_t^x, t \geq 0)$. Moreover, Theorem 4.20 ensures that the strong Markov property holds for $(S_t^x, t \geq 0)$. Under that model, we denote by

$$P_\sigma(x, y) = \sup_{\tau \in \mathcal{T}_{0, \infty}} \mathbb{E} [e^{-r\tau} (y - S_\tau^x)^+] \text{ and } C_\sigma(x, y) = \sup_{\tau \in \mathcal{T}_{0, \infty}} \mathbb{E} [e^{-r\tau} (S_\tau^x - y)^+]$$

respectively the prices of the American perpetual put and call options with strike $y > 0$ and spot x . Here, $\mathcal{T}_{0, \infty}$ simply denotes the set of the stopping times with respect to the natural filtration of $(S_t^x, t \geq 0)$. Since $e^{-rt} S_t^x = x \exp \left(-\delta t + \int_0^t \sigma(S_u^x) dW_u - \frac{1}{2} \int_0^t \sigma^2(S_u^x) du \right)$ and $\int_0^t \sigma^2(S_u^x) du \geq \underline{\sigma}^2 t \xrightarrow{t \rightarrow +\infty} +\infty$, it follows from the Dubins-Schwarz theorem that

$$e^{-rt} S_t^x \xrightarrow{t \rightarrow +\infty} 0 \text{ a.s.} \quad (5.4)$$

As a consequence, $e^{-rt} (y - S_t^x)^+ \xrightarrow{t \rightarrow +\infty} \mathbf{1}_{\{r=0\}} y$ and $e^{-rt} (S_t^x - y)^+ \xrightarrow{t \rightarrow +\infty} 0$ a.s. On $\{\tau = \infty\}$, we thus set

$$e^{-r\tau} (y - S_\tau^x)^+ = \mathbf{1}_{\{r=0\}} y \text{ and } e^{-r\tau} (S_\tau^x - y)^+ = 0. \quad (5.5)$$

Let us consider the second-order ordinary differential equation

$$\frac{1}{2} \sigma^2(x) x^2 f''(x) + (r - \delta) x f'(x) - r f(x) = 0, \quad x > 0. \quad (5.6)$$

According to Borodin and Salminen ([15], chap. 2) the functions

$$\forall x > 0, \quad f_\uparrow(x) = \begin{cases} \mathbb{E}[e^{-r\tau_1^x}], & \text{if } x \leq 1 \\ 1/\mathbb{E}[e^{-r\tau_1^1}], & \text{if } x > 1 \end{cases} \quad \text{and} \quad f_\downarrow(x) = \begin{cases} 1/\mathbb{E}[e^{-r\tau_1^1}], & \text{if } x \leq 1 \\ \mathbb{E}[e^{-r\tau_1^x}], & \text{if } x > 1 \end{cases} \quad (5.7)$$

where for $x, y > 0$, $\tau_y^x = \inf\{t \geq 0, S_t^x = y\}$ ($\inf \emptyset = +\infty$), are the unique solutions (up to a multiplicative constant) that are positive and respectively increasing and decreasing. The volatility function σ being continuous, (5.7) ensures that these functions are $\mathcal{C}^2$ on $\mathbb{R}_+^*$.

Remark 5.1.1. *It is easy using the strong Markov property to get:*

$$\forall x, y > 0, \quad \mathbb{E}[e^{-r\tau_y^x}] = \begin{cases} f_\uparrow(x)/f_\uparrow(y), & \text{if } x \leq y \\ f_\downarrow(x)/f_\downarrow(y), & \text{if } x \geq y. \end{cases}$$

Remark 5.1.2. *Assuming $r > 0$, one has (see Borodin and Salminen [15])*

$$\begin{aligned} \lim_{x \rightarrow 0} f_\downarrow(x) &= +\infty, & \lim_{x \rightarrow +\infty} f_\downarrow(x) &= 0 \\ \lim_{x \rightarrow 0} f_\uparrow(x) &= 0, & \lim_{x \rightarrow +\infty} f_\uparrow(x) &= +\infty. \end{aligned}$$

The function $f_\downarrow$ (resp. $f_\uparrow$) is thus, up to a multiplicative constant, the unique solution to (5.6) such that $\lim_{x \rightarrow +\infty} f(x) = 0$ (resp. $\lim_{x \rightarrow 0} f(x) = 0$).

Remark 5.1.3. In the case $\delta = 0$, we have the analytical solutions :

$$f_{\downarrow}(x) = \frac{\varphi(x)}{\varphi(1)} \text{ where } \varphi(x) = x \int_x^{+\infty} \left(\frac{1}{v^2} \exp \left[- \int_1^v \frac{2r}{u\sigma^2(u)} du \right] \right) dv, \quad f_{\uparrow}(x) = x.$$

Indeed, since $f(x) = x$ is solution of $\frac{1}{2}\sigma^2(x)x^2f''(x) + rxf'(x) - rf(x) = 0$, we search a general solution that can be written $f(x) = x\tilde{f}(x)$. This leads to $\frac{1}{2}\sigma^2(x)x\tilde{f}''(x) + (r + \sigma^2(x))\tilde{f}'(x) = 0$ and then $\tilde{f}'(x) = \frac{C_1}{x^2} \exp \left[- \int_1^x \frac{2r}{u\sigma^2(u)} du \right]$. Therefore, $\exists C_1, C_2 \in \mathbb{R}$, $f(x) = C_2x + C_1x \int_x^{+\infty} \left(\frac{1}{v^2} \exp \left[- \int_1^v \frac{2r}{u\sigma^2(u)} du \right] \right) dv$.

Now, we are in position to show the existence of an optimal stopping time and give the call and put prices. Let us mention here that the problem of perpetual optimal stopping is treated in the paper of Dayanik and Karatzas [26], for a general payoff function and an underlying evolving according to a general one-dimensional time homogeneous diffusion process. Villeneuve [64] considers a model where the constant dividend rate δ in (5.3) is replaced by a function $\delta(S_t)$ and gives sufficient condition on the payoff function ensuring that a threshold strategy is optimal. Here, we give a direct proof that generalizes the approach developed by Beibel and Lerche [10] in the Black-Scholes case.

Theorem 5.1.4. Assume $r > 0$. For any strike $y > 0$, there is a unique $x_{\sigma}^*(y) \in (0, y) \cap (0, \frac{r}{\delta}y]$ such that $\tau_x^P = \inf\{t \geq 0, S_t^x \leq x_{\sigma}^*(y)\}$ (convention $\inf \emptyset = +\infty$) is an optimal stopping time for the put and:

$$\forall x \leq x_{\sigma}^*(y), P_{\sigma}(x, y) = (y - x)^+, \quad \forall x > x_{\sigma}^*(y), P_{\sigma}(x, y) = \frac{y - x_{\sigma}^*(y)}{f_{\downarrow}(x_{\sigma}^*(y))} f_{\downarrow}(x) > (y - x)^+. \quad (5.8)$$

In addition, we have $f'_{\downarrow}(x_{\sigma}^*(y)) < 0$ and:

$$x_{\sigma}^*(y) - y = \frac{f_{\downarrow}(x_{\sigma}^*(y))}{f'_{\downarrow}(x_{\sigma}^*(y))}. \quad (5.9)$$

Last, the smooth-fit principle holds: $\partial_x P_{\sigma}(x_{\sigma}^*(y), y) = -1$.

Remark 5.1.5. If $r = 0$, $\forall x, y > 0$, $P_{\sigma}(x, y) = y$ since for any stopping time τ , $(y - S_{\tau}^x)^+ \leq y$ and equality holds for $\tau = +\infty$ by (5.5).

Proof. Let us define:

$$\forall z > 0, h(z) = \frac{(y - z)^+}{f_{\downarrow}(z)} \text{ and } h^* = \sup_{z > 0} h(z).$$

Since the function h is continuous such that $h(y) = 0$ and $h(0+) = 0$ (Remark 5.1.2), $x_{\sigma}^*(y) = \sup\{z > 0, h(z) = h^*\}$ belongs to $(0, y)$ and is such that $h(x_{\sigma}^*(y)) = h^*$. Since the

function h is $\mathcal{C}^2$ on $(0, y)$, we have $h'(x_\sigma^*(y)) = 0$ and $h''(x_\sigma^*(y)) \leq 0$. These conditions give easily

$$f_\downarrow(x_\sigma^*(y)) + (y - x_\sigma^*(y))f'_\downarrow(x_\sigma^*(y)) = 0 \text{ and } f''_\downarrow(x_\sigma^*(y)) \geq 0.$$

Since $f_\downarrow$ is positive and $x_\sigma^*(y) < y$, we have $f'_\downarrow(x_\sigma^*(y)) < 0$ and deduce (5.9). The second order condition and equation (5.6) then give $x_\sigma^*(y)(r - \delta)f'_\downarrow(x_\sigma^*(y)) - rf_\downarrow(x_\sigma^*(y)) \leq 0$ and so

$$ry - \delta x_\sigma^*(y) \geq 0.$$

Now let us check the optimality of τ_x^P and consider $\tau \in \mathcal{T}_{0,\infty}$. By Fatou's lemma and Doob's optional sampling theorem, we have

$$\begin{aligned} \mathbb{E}[e^{-r\tau}(y - S_\tau^x)^+] &\leq \liminf_{t \rightarrow +\infty} \mathbb{E}[e^{-r\tau \wedge t}(y - S_{\tau \wedge t}^x)^+] \\ &= \liminf_{t \rightarrow +\infty} \mathbb{E}[e^{-r\tau \wedge t} f_\downarrow(S_{\tau \wedge t}^x) h(S_{\tau \wedge t}^x)] \\ &\leq h(x_\sigma^*(y)) \liminf_{t \rightarrow +\infty} \mathbb{E}[e^{-r\tau \wedge t} f_\downarrow(S_{\tau \wedge t}^x)] \leq h(x_\sigma^*(y)) f_\downarrow(x) \end{aligned}$$

since $e^{-rt}f_\downarrow(S_t^x) = f_\downarrow(x) + \int_0^t e^{-ru}\sigma(S_u^x)S_u^x f'_\downarrow(S_u^x)du$ is a nonnegative local martingale and therefore a supermartingale. If $x \geq x_\sigma^*(y)$, we have using Remark 5.1.1:

$$\mathbb{E}[e^{-r\tau_x^P}(y - S_{\tau_x^P}^x)^+] = \mathbb{E}[e^{-r\tau_{x_\sigma^*(y)}^x}(y - S_{\tau_{x_\sigma^*(y)}^x}^x)^+] = (y - x_\sigma^*(y))\mathbb{E}[e^{-r\tau_{x_\sigma^*(y)}^x}] = h(x_\sigma^*(y))f_\downarrow(x)$$

and τ_x^P is optimal for $x \geq x_\sigma^*(y)$. Since $x_\sigma^*(y) = \sup\{z > 0, h(z) = h^*\}$, we have $(y - x)^+ = h(x)f_\downarrow(x) < f_\downarrow(x)h(x_\sigma^*(y))$ for $x > x_\sigma^*(y)$, and finally deduces (5.8) for $x \geq x_\sigma^*(y)$.

We consider now the complementary case $x \in (0, x_\sigma^*(y))$, and set $\tau \in \mathcal{T}_{0,\infty}$. Using the strong Markov property and the optimality result when the initial spot is $x_\sigma^*(y)$, we get

$$\mathbb{E}[e^{-r\tau}(y - S_\tau^x)^+] \leq \mathbb{E}[e^{-r\tau \wedge \tau_{x_\sigma^*(y)}^x}(y - S_{\tau \wedge \tau_{x_\sigma^*(y)}^x}^x)^+].$$

On $\{t < \tau_{x_\sigma^*(y)}^x\}$, we have $S_t^x < x_\sigma^*(y)$, $de^{-rt}(y - S_t^x) = e^{-rt}(\underbrace{\delta S_t^x - ry}_{\leq 0})dt - e^{-rt}\sigma(S_t^x)S_t^x dW_t$

and so $\mathbb{E}[e^{-r\tau \wedge \tau_{x_\sigma^*(y)}^x}(y - S_{\tau \wedge \tau_{x_\sigma^*(y)}^x}^x)^+] \leq \liminf_{t \rightarrow +\infty} \mathbb{E}[e^{-r\tau \wedge \tau_{x_\sigma^*(y)}^x}^t(y - S_{\tau \wedge \tau_{x_\sigma^*(y)}^x}^x)^+] \leq (y - x)$. $\square$

Now, we state the similar result for the call prices.

Theorem 5.1.6. *Assume $\delta > 0$. For any strike $y > 0$, there is a unique $\Upsilon_\sigma^*(y) \in (y, \infty) \cap [\frac{r}{\delta}y, +\infty)$ such that $\tau_x^C = \inf\{t \geq 0, S_t^x \geq \Upsilon_\sigma^*(y)\}$ is an optimal stopping time for the call and:*

$$\forall x \geq \Upsilon_\sigma^*(y), C_\sigma(x, y) = (x - y)^+, \quad \forall x < \Upsilon_\sigma^*(y), C_\sigma(x, y) = \frac{\Upsilon_\sigma^*(y) - y}{f_\uparrow(\Upsilon_\sigma^*(y))} f_\uparrow(x) > (x - y)^+. \quad (5.10)$$

In addition, we have $f'_\uparrow(\Upsilon_\sigma^*(y)) > 0$ and:

$$\Upsilon_\sigma^*(y) - y = \frac{f'_\uparrow(\Upsilon_\sigma^*(y))}{f'_\uparrow(\Upsilon_\sigma^*(y))}. \quad (5.11)$$

Last, the smooth-fit principle holds: $\partial_x C_\sigma(\Upsilon_\sigma^*(y), y) = 1$.

Remark 5.1.7. If $\delta = 0$, $\forall x, y > 0$, $C_\sigma(x, y) = x$. Indeed, the Call-Put parity $\mathbb{E}[e^{-rt}(S_t^x - y)^+] = x - ye^{-rt} + \mathbb{E}[e^{-rt}(y - S_t^x)^+]$ gives the convergence to x in both cases $r > 0$ and $r = 0$ when $t \rightarrow +\infty$. Now, thanks to the Fatou lemma, we have for $\tau \in \mathcal{T}_{0,\infty}$: $\mathbb{E}[e^{-r\tau}(S_\tau^x - y)^+] \leq \liminf_{t \rightarrow +\infty} \mathbb{E}[e^{-r\tau \wedge t}(S_{\tau \wedge t}^x - y)^+] \leq \liminf_{t \rightarrow +\infty} \mathbb{E}[e^{-r\tau \wedge t} S_{\tau \wedge t}^x] = x$.

Proof. The proof works as for the put, and we just hint the differences. We define

$$\forall z > 0, h(z) = \frac{(z - y)^+}{f_\uparrow(z)} \text{ and } h^* = \sup_{z > 0} h(z).$$

Let us admit for a while that $\lim_{z \rightarrow +\infty} h(z) = 0$. Then, since $h(y) = 0$ and h is continuous, h reaches its maximum in $\Upsilon_\sigma^*(y) = \inf\{z > 0, h(z) = h^*\}$, and $\Upsilon_\sigma^*(y) \in (y, \infty)$. This gives (5.11). We then consider the case $x \leq \Upsilon_\sigma^*(y)$ and show that $\tau_x^C = \tau_{\Upsilon_\sigma^*(y)}^C$ is optimal. Note that in the special case $r = 0$, we have to use Proposition 5.1.8 which is stated below. Finally, we prove that τ_x^C is optimal when $x > \Upsilon_\sigma^*(y)$ using that $\delta \Upsilon_\sigma^*(y) - ry \geq 0$.

Now, let us check that $\lim_{z \rightarrow +\infty} h(z) = 0$. In the case $r = 0$, it is straightforward using the explicit form given in Proposition 5.1.8 below that $f_\uparrow(x) \geq \frac{1}{1+\frac{2\delta}{\sigma^2}} x^{1+2\delta/\sigma^2} - 1$ for $x \geq 1$, and we have then $\lim_{z \rightarrow +\infty} h(z) = 0$. When $r > 0$, Itô's Formula gives

$$de^{-rt}(S_t^1)^{1+a} = e^{-rt}(S_t^1)^{1+a} \left\{ (a+1)\sigma(S_t^1)dW_t + [a(r + (a+1)\sigma^2(S_t^1)/2) - (a+1)\delta]dt \right\}.$$

When $a > 0$, the drift term is bounded from above by $a(r + (a+1)\sigma^2/2) - (a+1)\delta$ and we can find $a > 0$ such that this bound is negative since $a(r + (a+1)\sigma^2/2) - (a+1)\delta \xrightarrow{a \rightarrow 0} -\delta < 0$. Then, for $x \geq 1$, we have $\mathbb{E}[e^{-r\tau_x^1 \wedge t}(S_{\tau_x^1 \wedge t}^1)^{1+a}] \leq 1$ thanks to Doob's optional sampling theorem. The Fatou lemma gives then $\mathbb{E}[e^{-r\tau_x^1}(S_{\tau_x^1}^1)^{1+a}] \leq 1$, and therefore we get $f_\uparrow(x) = 1/\mathbb{E}[e^{-r\tau_x^1}] \geq x^{1+a}$. This shows $\lim_{z \rightarrow +\infty} h(z) = 0$. $\square$

Proposition 5.1.8. In the case $r = 0$, the unique nonincreasing and increasing solution of (5.6) starting from 1 in 1 are respectively:

$$f_\downarrow(x) = 1, \quad f_\uparrow(x) = \frac{\psi(x)}{\psi(1)} \text{ where } \psi(x) = \int_0^x \exp \left[\int_1^v \frac{2\delta}{u\sigma^2(u)} du \right] dv.$$

Moreover, we have $f_\uparrow(x) = 1/\mathbb{P}(\tau_x^1 < +\infty)$ for $x \geq 1$ and $f_\uparrow(x) = \mathbb{P}(\tau_1^x < +\infty)$ for $x \in (0, 1]$.

Proof. When $r = 0$, the differential equation $\frac{1}{2}\sigma^2(y)y^2f''(y) - \delta yf'(y) = 0$ is easy to integrate: $f'(y) = C_3 \exp \left[\int_1^y \frac{2\delta}{u\sigma^2(u)} du \right]$ and then

$$g(y) = C_4 + C_3 \int_0^y \exp \left[\int_1^v \frac{2\delta}{u\sigma^2(u)} du \right] dv$$

for $C_3, C_4 \in \mathbb{R}$. For $x \geq 1$, $f_\uparrow(S_{\tau_x^1 \wedge t}^1)$ is a bounded martingale that converges almost surely to $f_\uparrow(x)\mathbf{1}_{\{\tau_x^1 < +\infty\}}$ thanks to (5.4) and $\lim_{x \rightarrow 0} f_\uparrow(x) = 0$. Therefore $f_\uparrow(x)\mathbb{P}(\tau_x^1 < +\infty) = 1$, and the proof is the same for $x \in (0, 1]$. $\square$

To conclude this section, we state a comparison result which will enables us to compare $x_\sigma^*(y)$ and $\Upsilon_\sigma^*(y)$ with the exercise boundaries obtained in the Black-Scholes model with constant volatility $\underline{\sigma}$ (resp. $\bar{\sigma}$) where $\underline{\sigma}$ (resp. $\bar{\sigma}$) bounds the function σ from below (resp. above).

Proposition 5.1.9. *Let us consider two volatility functions σ_1 and σ_2 such that $\forall x > 0$, $\sigma_1(x) \leq \sigma_2(x)$ and that satisfy $(\mathcal{H}_{\text{vol}})$. We also assume that either $f''_{\downarrow, \sigma_1}$ or $f''_{\downarrow, \sigma_2}$ (resp. either $f''_{\uparrow, \sigma_1}$ or $f''_{\uparrow, \sigma_2}$) are nonnegative functions and $r > 0$ (resp. $\delta > 0$). Then, we have*

$$\forall x, y > 0, P_{\sigma_1}(x, y) \leq P_{\sigma_2}(x, y) \quad (\text{resp. } C_{\sigma_1}(x, y) \leq C_{\sigma_2}(x, y))$$

and we can compare the exercise boundaries:

$$\forall y > 0, x_{\sigma_1}^*(y) \geq x_{\sigma_2}^*(y) \quad (\text{resp. } \Upsilon_{\sigma_1}^*(y) \leq \Upsilon_{\sigma_2}^*(y)).$$

Here and in the proof below, we add in the notation for each mathematical object the volatility function to which it refers. El Karoui and al. [31] and Hobson [40] prove that for a convex payoff function, the price of an American option with finite maturity is a convex function of the underlying spot price. They deduce monotonicity with respect to the local volatility function. Their results imply at the same time the convexity assumption made in the above proposition and its conclusion. In this paper, we prefer to give autonomous proofs of these results in our simple framework. And we will first use proposition 5.1.9 to compare with the Black-Scholes case where convexity is obvious. We can then deduce (Lemma 5.3.1) that $f''_\downarrow$ and $f''_\uparrow$ are positive for any σ satisfying $(\mathcal{H}_{\text{vol}})$.

Proof. Let us consider for example the put case with $f''_{\downarrow, \sigma_1} \geq 0$. Let $x \geq z > 0$. Ito's formula gives:

$$\begin{aligned} de^{-rt} f_{\downarrow, \sigma_1}(S_t^{x, \sigma_2}) &= e^{-rt} f'_{\downarrow, \sigma_1}(S_t^{x, \sigma_2}) \sigma_2(S_t^{x, \sigma_2}) S_t^{x, \sigma_2} dW_t + e^{-rt} \left[\frac{\sigma_2^2(S_t^{x, \sigma_2})}{2} (S_t^{x, \sigma_2})^2 f''_{\downarrow, \sigma_1}(S_t^{x, \sigma_2}) \right. \\ &\quad \left. + (r - \delta) S_t^{x, \sigma_2} f'_{\downarrow, \sigma_1}(S_t^{x, \sigma_2}) - r f_{\downarrow, \sigma_1}(S_t^{x, \sigma_2}) \right] dt. \end{aligned}$$

The term between brackets is nonnegative since we have $\sigma_1 \leq \sigma_2$ and $f_{\downarrow, \sigma_1}$ is a convex function solving (5.6). Therefore we get $\mathbb{E}[e^{-r\nu_n \wedge \tau_z^{x, \sigma_2}} f_{\downarrow, \sigma_1}(S_{\nu_n \wedge \tau_z^{x, \sigma_2}}^{x, \sigma_2})] \geq f_{\downarrow, \sigma_1}(x)$ using the optional sampling theorem, where $\nu_n = \inf\{t \geq 0, S_t^{x, \sigma_2} \geq n\} \wedge n$. Since $\nu_n \xrightarrow{a.s.} +\infty$ and $f_{\downarrow, \sigma_1}(S_{\nu_n \wedge \tau_z^{x, \sigma_2}}^{x, \sigma_2})$ is bounded by $f_{\downarrow, \sigma_1}(z)$, Lebesgue's dominated convergence theorem gives then $\mathbb{E}[e^{-r\tau_z^{x, \sigma_2}}] \geq f_{\downarrow, \sigma_1}(x)/f_{\downarrow, \sigma_1}(z)$ and so $\mathbb{E}[e^{-r\tau_z^{x, \sigma_2}}] \geq \mathbb{E}[e^{-r\tau_z^{x, \sigma_1}}]$ using Remark 5.1.1. The same conclusion holds when $f_{\downarrow, \sigma_2}$ is convex by estimating $\mathbb{E}[e^{-r\tau_z^{x, \sigma_1}} f_{\downarrow, \sigma_2}(S_{\tau_z^{x, \sigma_1}}^{x, \sigma_1})]$. We then get the result: if $P_{\sigma_1}(x, y) > (x - y)^+$, $P_{\sigma_1}(x, y) = (y - x_{\sigma_1}^*(y))\mathbb{E}[e^{-r\tau_{x_{\sigma_1}^*(y)}^{x, \sigma_1}}] \leq (y - x_{\sigma_1}^*(y))\mathbb{E}[e^{-r\tau_{x_{\sigma_1}^*(y)}^{x, \sigma_2}}] = \mathbb{E}[(y - S_{\tau_{x_{\sigma_1}^*(y)}^{x, \sigma_2}}^{x, \sigma_2})^+ e^{-r\tau_{x_{\sigma_1}^*(y)}^{x, \sigma_2}}] \leq P_{\sigma_2}(x, y)$. Now we just observe that $\{x > 0, P_{\sigma_1}(x, y) > (x - y)^+\} \subset \{x > 0, P_{\sigma_2}(x, y) > (x - y)^+\}$ and thus $x_{\sigma_1}^*(y) = \inf\{x > 0, P_{\sigma_1}(x, y) > (x - y)^+\} \geq \inf\{x > 0, P_{\sigma_2}(x, y) > (x - y)^+\} = x_{\sigma_2}^*(y)$. $\square$

5.2 Framework and notations

We will present in this section the framework that we will consider in all the paper. To clarify the duality, we will use names that implicitly refer either to the primal (or “real”) world, or to the dual world. This denomination has no mathematical meaning since, as we will see, there are no difference between them. On the contrary, from a financial point of view, natural variables such as the interest rate, the dividend rate have their true meaning in the primal world, while in the dual world they interchange their role. This is the reason why we also name the primal world “real” world.

The primal (“real”) world

The primal world is the framework we just have described. The spot interest rate r is constant and nonnegative, and S_t is an asset which pays a constant dividend rate $\delta \geq 0$ and is driven by a homogeneous volatility function $\sigma : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ that satisfies $(\mathcal{H}_{\text{vol}})$ under a risk-neutral measure. The prices of the perpetual American put and call are respectively denoted by $P_\sigma(x, y)$ and $C_\sigma(x, y)$, and their exercise boundary by x_σ^* and Υ_σ^* .

The dual world

In the dual world, δ plays the role of the interest rate and r of the dividend rate; x plays the role of the strike and y is the spot value of the share. Let $\eta : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ be an homogeneous volatility function that is also assumed to satisfy $(\mathcal{H}_{\text{vol}})$. We consider then $(\bar{S}_t^y, t \geq 0)$ the solution of $d\bar{S}_t = \bar{S}_t((\delta - r)dt + \eta(\bar{S}_t)dW_t)$ that starts from y at time 0. Under that model,

we denote respectively by

$$p_\eta(y, x) = \sup_{\tau \in \mathcal{T}_{0, \infty}} \mathbb{E} [e^{-\delta\tau} (x - \overline{S}_\tau^y)^+] \quad \text{and} \quad c_\eta(y, x) = \sup_{\tau \in \mathcal{T}_{0, \infty}} \mathbb{E} [e^{-\delta\tau} (\overline{S}_\tau^y - x)^+]$$

the prices of the perpetual put and call with strike $x > 0$ and spot y . We can define, as in the primal world, $g_\downarrow$ and $g_\uparrow$ as the unique decreasing (non-increasing when $\delta = 0$) and increasing positive solution to

$$\frac{1}{2}\eta^2(x)x^2g''(x) + (\delta - r)xg'(x) - \delta g(x) = 0, \quad (5.12)$$

and we name the exercise boundaries $\xi_\eta^*(x) < x$ and $y_\eta^*(x) > x$ that are respectively associated with the put $p_\eta(y, x)$ and the call $c_\eta(y, x)$.

Notations

The aim of this paper is to put in evidence a duality relation and interpret put (resp. call) prices in the primal world as call (resp. put) prices in the dual world for a specific volatility function $\eta = \tilde{\sigma}$ (resp. $\eta = \hat{\sigma}$). When $r = 0$ (resp. $\delta = 0$), this is trivial since $P_\sigma(x, y) = c_\eta(y, x) = y$ (resp. $C_\sigma(x, y) = p_\eta(y, x) = x$) but not really fruitful, and we take thus the following convention in the sequel.

Convention 5.2.1. *We will always assume $r > 0$ (resp. $\delta > 0$) to state properties on P_σ and c_η (resp. C_σ and p_η).*

Both worlds being mathematically equivalent, *we will work with the put price in the primal world and the call price in the dual world* in order not to do the things twice. Following the Convention 5.2.1, we will consider a positive interest rate. Let us then denote from now:

$$\boxed{f = f_\downarrow \text{ and } g = g_\uparrow,}$$

and define $\alpha(y) = \frac{y - x_\sigma^*(y)}{f(x_\sigma^*(y))}$ and $\beta(x) = \frac{y_\eta^*(x) - x}{g(y_\eta^*(x))}$. The functions α and β are positive functions and it follows from the present section that :

$$\forall y > 0, \forall x \geq x_\sigma^*(y), \quad P_\sigma(x, y) = \alpha(y)f(x) \quad (5.13)$$

$$\forall x > 0, \forall y \leq y_\eta^*(x), \quad c_\eta(y, x) = \beta(x)g(y). \quad (5.14)$$

That product form will play an important role for the duality. Let us finally introduce notations relative the the Black-Scholes model. We define for $\varsigma > 0$

$$a(\varsigma) = \frac{\delta - r + \varsigma^2/2 - \sqrt{(\delta - r + \varsigma^2/2)^2 + 2r\varsigma^2}}{\varsigma^2} < 0,$$

$$b(\varsigma) = \frac{r - \delta + \varsigma^2/2 + \sqrt{(\delta - r - \varsigma^2/2)^2 + 2\delta\varsigma^2}}{\varsigma^2} = 1 - a(\varsigma) > 1,$$

and we can easily check that $f(x) = x^{a(\varsigma)}$ (resp. $g = x^{b(\varsigma)}$) when $\sigma(x) = \varsigma$ (resp. $\eta(x) = \varsigma$). In that case, the unique solution to (5.9) (resp. (5.11)) is:

$$x_\varsigma^*(y) = \frac{a(\varsigma)}{a(\varsigma) - 1}y \quad (\text{resp. } y_\varsigma^*(x) = \frac{b(\varsigma)}{b(\varsigma) - 1}x). \quad (5.15)$$

With Proposition 5.1.9, we deduce very useful estimations on the exercise boundaries.

Lemma 5.2.2. *If, $\forall x > 0$, $\underline{\sigma} \leq \sigma(x) \leq \bar{\sigma}$ and $\underline{\sigma} \leq \eta(x) \leq \bar{\sigma}$, then we have :*

$$\frac{a(\bar{\sigma})}{a(\bar{\sigma}) - 1}y \leq x_\sigma^*(y) \leq \frac{a(\underline{\sigma})}{a(\underline{\sigma}) - 1}y \quad \text{with} \quad \frac{a(\underline{\sigma})}{a(\underline{\sigma}) - 1} < \min(1, r/\delta) \quad (5.16)$$

$$\frac{b(\underline{\sigma})}{b(\underline{\sigma}) - 1}x \leq y_\eta^*(x) \leq \frac{b(\bar{\sigma})}{b(\bar{\sigma}) - 1}x \quad \text{with} \quad \max(1, \delta/r) < \frac{b(\underline{\sigma})}{b(\underline{\sigma}) - 1}. \quad (5.17)$$

Proof. It is straightforward from Proposition 5.1.9 and (5.15) to get the bound on the exercise boundaries since $x \mapsto x^\gamma$ is convex for $\gamma \notin]0, 1[$. We have to show that for $\varsigma > 0$, $\frac{a(\varsigma)}{a(\varsigma) - 1} < \min(1, r/\delta)$ and $\frac{b(\varsigma)}{b(\varsigma) - 1} > \max(1, \delta/r)$. Since $a(\varsigma) < 0$ and $b(\varsigma) > 1$, we get that $a(\varsigma)/(a(\varsigma) - 1) \in (0, 1)$ and $b(\varsigma)/(b(\varsigma) - 1) \in (1, +\infty)$. We can also check that $a(\varsigma)/(a(\varsigma) - 1)$ is a root of the polynomial $Q(X) = \delta X^2 - (r + \delta + \varsigma^2/2)X + r$. As $Q(x) = 0 \iff \sigma^2 x = 2(1 - x)(r - \delta x)$ and since $a(\varsigma)/(a(\varsigma) - 1) \in (0, 1)$, we then deduce that $a(\varsigma)/(a(\varsigma) - 1) < r/\delta$. In the same way, we have $b(\varsigma)/(b(\varsigma) - 1) > \delta/r$. $\square$

5.3 ODE for the exercise boundary

We have seen previously that the exercise boundaries satisfy

$$x_\sigma^*(y) - y = f(x_\sigma^*(y))/f'(x_\sigma^*(y)) \quad (\text{resp. } y_\eta^*(x) - x = g(y_\eta^*(x))/g'(y_\eta^*(x))). \quad (5.18)$$

We will soon prove that for fixed $y > 0$ (resp. $x > 0$), (5.18) admits a unique solution $x_\sigma^*(y)$ (resp. $y_\eta^*(x)$).

Lemma 5.3.1. *The function f' (resp. g') is negative (resp. positive) and f'' (resp. g'') is positive on $(0, +\infty)$. Moreover, the boundaries $x_\sigma^*(y)$ and $y_\eta^*(x)$ are respectively the unique solution to $y - x + f(x)/f'(x) = 0$ and $y - x - g(x)/g'(x) = 0$. Last, $x_\sigma^*(y)$, $\alpha(y)$, $y_\eta^*(x)$ and $\beta(x)$ are $\mathcal{C}^1$ functions on $\mathbb{R}_+^*$.*

Remark 5.3.2. *Positivity of f'' and g'' and (5.13) and (5.14) imply positivity of $\partial_x^2 P_\sigma(x, y)$ and $\partial_y^2 c_\eta(y, x)$ in the continuation regions.*

Differentiating (5.18) with respect to y (resp. x), one obtains $1 = (x_\sigma^*)'(y) \frac{f(x_\sigma^*(y))f''(x_\sigma^*(y))}{f'(x_\sigma^*(y))^2}$ (resp. $1 = (y_\eta^*)'(y) \frac{g(y_\eta^*(y))g''(y_\eta^*(y))}{g'(y_\eta^*(y))^2}$). Using (5.18) and equation (5.6) (resp. (5.12)) one deduces the following result (see equation (5.22) below).

Proposition 5.3.3. *Let us assume that the volatility functions σ and η satisfy $(\mathcal{H}_{\text{vol}})$. Then, the boundaries $x_\sigma^*(y)$ and $y_\eta^*(x)$ satisfy the following ODEs:*

$$(x_\sigma^*)'(y) = \frac{x_\sigma^*(y)^2 \sigma(x_\sigma^*(y))^2}{2(y - x_\sigma^*(y))(ry - \delta x_\sigma^*(y))}, \quad (5.19)$$

$$(y_\eta^*)'(x) = \frac{\eta^2(y_\eta^*(x))y_\eta^*(x)^2}{2(y_\eta^*(x) - x)(ry_\eta^*(x) - \delta x)}. \quad (5.20)$$

Proof of Lemma 5.3.1. We only give the proof in the put case, the argument being similar for the call. By (5.6), for $x > 0$, $f''(x)$ has the same sign as $h(x) = rf(x) + (\delta - r)xf'(x)$. If for some $x > 0$, $f'(x) = 0$, then since f is positive, $f''(x) > 0$. Therefore x is a local minimum point of f which contradicts the decreasing property of this function. Hence f' is a negative function.

When $\delta \leq r$, h and therefore f'' are positive functions. When $\delta > r$, we remark that if $f''(x) = 0$ then $h'(x) = \delta f'(x) < 0$. Since the continuous function f'' and h have the same sign, this implies that

$$\forall x > \inf\{z > 0 : f''(z) \leq 0\}, \quad f''(x) < 0. \quad (5.21)$$

Now for $y > 0$, by (5.6) then (5.18), we have

$$\frac{x_\sigma^*(y)^2 \sigma(x_\sigma^*(y))^2}{2} \frac{f''(x_\sigma^*(y))}{f'(x_\sigma^*(y))} = r \frac{f(x_\sigma^*(y))}{f'(x_\sigma^*(y))} - (r - \delta)x_\sigma^*(y) = \delta x_\sigma^*(y) - ry. \quad (5.22)$$

By (5.16), the right-hand-side is negative and moreover $\lim_{y \rightarrow +\infty} x_\sigma^*(y) = +\infty$. Hence $\sup\{z > 0 : f''(z) > 0\} = +\infty$ and with (5.21), we conclude that $f'' > 0$.

According to (5.18), $F(x_\sigma^*(y), y) = 0$ where

$$F(x, y) = y - x + f(x)/f'(x).$$

The function F is $\mathcal{C}^1$ on $(0, +\infty) \times (0, +\infty)$ and such that

$$\forall x, y > 0, \quad \partial_x F(x, y) = -f(x)f''(x)/f'(x)^2 < 0.$$

Therefore for fixed $y > 0$, $x^*(y)$ is the unique solution to $F(x, y) = 0$. Moreover, $y \rightarrow x^*(y)$ is C^1 by the implicit function theorem. Last, one deduces from (5.18) that $\alpha(y)$ is a C^1 function. $\square$

The positivity of f'' and g'' gives the following result.

Corollary 5.3.4. *The comparison result stated in Proposition 5.1.9 holds for any $\sigma_1 \leq \sigma_2$ satisfying $(\mathcal{H}_{\text{vol}})$.*

Let us now give a uniqueness result for the ODEs (5.19) and (5.20).

Proposition 5.3.5. *There is only one solution x_σ^* of (5.19) (resp. y_η^* of (5.20)) defined on $(0, +\infty)$ that satisfies $\forall y > 0, c_1 y \leq x_\sigma^*(y) \leq c_2 y$ with $0 < c_1 \leq c_2 < \min(1, r/\delta)$ (resp. $\forall x > 0, d_1 x \leq y_\eta^*(x) \leq d_2 x$ with $d_1 > \max(1, \delta/r)$).*

Proof. Let us first remark that the uniqueness result for (5.19) is equivalent to the uniqueness result for (5.20). Indeed, it is easy to see that $x_\sigma^*(y)$ is solution of (5.19) if and only if $\hat{y}(x) := 1/(x_\sigma^*(1/x))$ is solution of (5.20) with the volatility function $\eta(x) = \sigma(1/x)$. This new volatility also satisfies $(\mathcal{H}_{\text{vol}})$. Moreover, $d_1 x \leq \hat{y}(x) \leq d_2 x$ with $d_1 > \max(1, \delta/r)$ if, and only if $0 \leq c_1 y \leq x_\sigma^*(y) \leq c_2 y$ with $0 < c_1 \leq c_2 < \min(1, r/\delta)$.

Let us suppose then that there are two solutions of (5.20), $y_1(x)$ and $y_2(x)$, that are defined on $\mathbb{R}_+$ and satisfy $d_2 x \geq y_j(x) \geq d_1 x$ for some $d_2 > d_1 > \max(1, \delta/r)$. Since $y_1'(x) > 0$ for $x > 0$, y_1 is invertible and we have:

$$\begin{aligned} \frac{d}{dx} y_1^{-1}(y_2(x)) &= \frac{y_2(x)^2 \eta(y_2(x))^2}{2(y_2(x) - x)(ry_2(x) - \delta x)} \frac{2(y_2(x) - y_1^{-1}(y_2(x)))(ry_2(x) - \delta y_1^{-1}(y_2(x)))}{y_2(x)^2 \eta(y_2(x))^2} \\ &= \frac{(y_2(x) - y_1^{-1}(y_2(x)))(ry_2(x) - \delta y_1^{-1}(y_2(x)))}{(y_2(x) - x)(ry_2(x) - \delta x)}. \end{aligned}$$

Thus, the function $\psi(x) = y_1^{-1}(y_2(x))/x$ solves

$$\begin{aligned} \psi'(x) &= \frac{1}{x} \left[\frac{y_2(x) - \psi(x)x}{y_2(x) - x} \times \frac{ry_2(x) - \delta x\psi(x)}{ry_2(x) - \delta x} - \psi(x) \right] \\ &= \frac{1}{x} \left[\left(1 - \frac{\psi(x) - 1}{y_2(x)/x - 1} \right) \left(1 - \frac{\psi(x) - 1}{ry_2(x)/(\delta x) - 1} \right) - \psi(x) \right]. \end{aligned} \quad (5.23)$$

The estimation $d_2 x \geq y_j(x) \geq d_1 x$ for $j \in \{1, 2\}$ with $d_1 > \max(1, \delta/r)$ implies that :

$$\exists A > 0, \forall x > 0, 1/A \leq \psi(x) \leq A, \quad (5.24)$$

$$\forall x > 0, \psi(x) < \min \left(\frac{y_2(x)}{x}, \frac{ry_2(x)}{\delta x} \right), \frac{y_2(x)}{x} - 1 > 0 \text{ and } \frac{ry_2(x)}{\delta x} > 0. \quad (5.25)$$

Since local uniqueness holds for (5.23) by the Cauchy Lipschitz theorem, the only solution φ such that $\varphi(1) = 1$ is the constant $\varphi \equiv 1$. Therefore checking that (5.24) does not hold for solutions φ satisfying (5.25) and such that $\varphi(1) \neq 1$ is enough to conclude that $\psi \equiv 1$.

Let φ be a solution to (5.23) satisfying (5.25). If $\varphi(1) > 1$, by local uniqueness for (5.23), for all $x \in \mathbb{R}_+$, $\varphi(x) > 1$. By (5.25), one deduces that for all $x \in \mathbb{R}_+$, $\varphi'(x) < \frac{1-\varphi(x)}{x} < 0$. Therefore, $\varphi'(x) \leq (1 - \varphi(1))/x$ for $x \in (0, 1]$, and we have

$$\varphi(x) \geq \varphi(1) + (1 - \varphi(1)) \ln(x) \xrightarrow{x \rightarrow 0} +\infty$$

which is contradictory to (5.24). In the same manner, if $\varphi(1) < 1$, $\varphi(x) < 1$ for $x \in \mathbb{R}_+^*$ and φ is strictly increasing. In particular, for $x \leq 1$, $\varphi'(x) \geq (1 - \varphi(1))/x$ and therefore $\varphi(1) - \varphi(x) \geq (1 - \varphi(1)) \ln(1/x) \xrightarrow{x \rightarrow 0} +\infty$ and this yields another contradiction. $\square$

Corollary 5.3.6. *Let us denote $\tilde{\mathcal{C}} = \{f \in \mathcal{C}^1(\mathbb{R}_+^*), \text{ s.t. } f(0) = 0, \exists 0 < a < b, \forall x \geq 0, a \leq f'(x) \leq b\}$. The application $\sigma \mapsto x_\sigma^*$ (resp. $\eta \mapsto y_\eta^*$) is one-to-one between the set $\{\sigma \in \mathcal{C}(\mathbb{R}_+^*) \text{ that satisfies } (\mathcal{H}_{\text{vol}})\}$ and the set of function $\tilde{\mathcal{C}}_x = \{x \in \tilde{\mathcal{C}}, \text{ s.t. } \exists 0 < c_1 \leq c_2 < \min(1, r/\delta), \forall y > 0, c_1 y \leq x(y) \leq c_2 y\}$ (resp. $\tilde{\mathcal{C}}_y = \{y \in \tilde{\mathcal{C}}, \text{ s.t. } \exists \max(1, \delta/r) < d_1 \leq d_2, \forall x > 0, d_1 x \leq y(x) \leq d_2 x\}$.)*

Proof. If σ is a continuous function satisfying $(\mathcal{H}_{\text{vol}})$, by (5.19) and (5.16), x_σ^* belongs to $\tilde{\mathcal{C}}_x$. The one to one property is easy to get. If $x_{\sigma_1}^* \equiv x_{\sigma_2}^*$ with σ_1 and σ_2 satisfying $(\mathcal{H}_{\text{vol}})$, the ODE (5.19) ensures that $\sigma_1^2(x_{\sigma_1}^*(y)) = \sigma_2^2(x_{\sigma_2}^*(y))$ for $y > 0$. Therefore $\sigma_1 \equiv \sigma_2$. Let us check the onto property and consider $x^*(y) \in \tilde{\mathcal{C}}_x$. The function σ defined by

$$\sigma(x^*(y)) = \frac{\sqrt{2(y - x^*(y))(ry - \delta x^*(y))x^{*'}(y)}}{x^*(y)} \quad (5.26)$$

is well defined thanks to the hypothesis made on x^* . As x_σ^* satisfies (5.16) and solves the same ODE (5.19) as x^* , we have $x^* \equiv x_\sigma^*$ using Proposition 5.3.5.

The proof for $\eta \mapsto y_\eta^*$ is the same and gives incidentally the expression of η in function of the exercise boundary $y^*(x)$:

$$\eta(y^*(x)) = \frac{\sqrt{2(y^*(x) - x)(ry^*(x) - \delta x)y^{*'}(x)}}{y^*(x)}. \quad (5.27)$$

$\square$

5.4 The call-put duality

This section is devoted to the key result of the paper : for related local volatility functions σ and η , we can interpret a put price in the primal world as a call price in the dual world.

5.4.1 The main result

Theorem 5.4.1 (Duality). *The following conditions are equivalent:*

1.

$$\forall x, y > 0, P_\sigma(x, y) = c_\eta(y, x). \quad (5.28)$$

2. x_σ^* and y_η^* are reciprocal functions: $\forall x > 0, x_\sigma^*(y_\eta^*(x)) = x$.

3. $\eta \equiv \tilde{\sigma}$ where

$$\tilde{\sigma}(y) = \frac{2(y - x_\sigma^*(y))(ry - \delta x_\sigma^*(y))}{yx_\sigma^*(y)\sigma(x_\sigma^*(y))}. \quad (5.29)$$

4. $\sigma \equiv \eta$ where

$$\eta(x) = \frac{2(y_\eta^*(x) - x)(ry_\eta^*(x) - \delta x)}{y_\eta^*(x)x\eta(y_\eta^*(x))}. \quad (5.30)$$

Remark 5.4.2. Thanks to relation (5.16) (resp. (5.17)), if σ (resp. η) satisfies $(\mathcal{H}_{\text{vol}})$ then the dual volatility function $\tilde{\sigma}$ defined by (5.29) (resp. η defined by (5.30)) satisfies $(\mathcal{H}_{\text{vol}})$.

Proof. 1 $\implies$ 2 : We have on the one hand $P_\sigma(x, y) = y - x$ on $\{(x, y), x \leq x_\sigma^*(y)\}$ and $P_\sigma(x, y) > y - x$ outside, and on the other hand $c_\eta(y, x) = y - x$ on $\{(x, y), y \geq y_\eta^*(x)\}$ and $c_\eta(y, x) > y - x$ outside. The duality relation (5.28) imposes then that $\{(x, y), x \leq x_\sigma^*(y)\} = \{(x, y), y \geq y_\eta^*(x)\}$ and so $y_\eta^*(x_\sigma^*(y)) = y$.

2 $\implies$ 3, 4 : Taking the derivative of the last relation, we get thanks to (5.19) and (5.20) $\frac{x_\sigma^*(y)^2 \sigma(x_\sigma^*(y))^2}{2(y - x_\sigma^*(y))(ry - \delta x_\sigma^*(y))} \frac{\eta^2(y)y^2}{2(y - x_\sigma^*(y))(ry - \delta x_\sigma^*(y))} = 1$ and deduce (5.29) and (5.30).

3 $\implies$ 2 (resp. 4 $\implies$ 2) : By (5.19) (resp. (5.20)) and (5.29) (resp. (5.30)), x_σ^{*-1} (resp. y_η^{*-1}) satisfies (5.20) (resp. (5.19)). Since by (5.16) (resp. (5.17)) this function satisfies and (5.17) (resp. (5.16)), one concludes by Proposition 5.3.5.

2 $\implies$ 1 : The equality (5.28) is clear in the exercise region since $\{(x, y), x \leq x_\sigma^*(y)\} = \{(x, y), y \geq y_\eta^*(x)\}$. Let us check that it also holds in the continuation region. Using the product form (5.14), and the smooth-fit principle (Theorem 5.1.6) we get for all $y \in \mathbb{R}_+^*$

$$\begin{cases} y - x_\sigma^*(y) = \beta(x_\sigma^*(y))g(y) \\ 1 = -\beta(x_\sigma^*(y))g'(y). \end{cases}$$

Differentiating the first equality with respect to y , one gets $1 - x_\sigma^*(y)' = x_\sigma^*(y)'\beta'(x_\sigma^*(y))g(y) + \beta(x_\sigma^*(y))g'(y)$, which combined with the second equality gives

$$-1 = \beta'(x_\sigma^*(y))g(y).$$

Dividing by the first equality and using (5.18), one deduces $\frac{\beta'}{\beta}(x_\sigma^*(y)) = \frac{f'}{f}(x_\sigma^*(y))$. Since $x_\sigma^* : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ is a bijection, there is a constant $C \neq 0$ such that $\beta \equiv Cf$. Since $\forall y > 0, \alpha(y)f(x_\sigma^*(y)) = y - x_\sigma^*(y) = \beta(x_\sigma^*(y))g(y)$, one has $\alpha \equiv g/C$. From (5.13) and (5.14), one concludes that (5.28) holds. $\square$

In this proof, we have shown that α is proportional to g , and so there is a constant $C > 0$ such that $P_\sigma(x, y) = Cf(x)g(y)$ for $x \leq x_\sigma^*(y)$. In the Black-Scholes' case, we have $P_\sigma(x, y) = Cx^{a(\sigma)}y^{b(\sigma)}$ for $x \leq x_\sigma^*(y)$, and we are able to calculate C using the boundary condition $P_\sigma(x_\sigma^*(y), y) = y - x_\sigma^*(y)$. We get that $C = b(\sigma)^{-b(\sigma)}/(-a(\sigma))^{a(\sigma)}$. We retrieve then the already known analytical formulae for the put prices (e.g. Gerber and Shiu [36]).

5.4.2 An analytic example of dual volatility functions

By (5.30) and (5.20), if $y^* \in \tilde{\mathcal{C}}_y$ (where $\tilde{\mathcal{C}}_y$ is defined in Corollary 5.3.6), then the reciprocal function of y^* is the put exercise boundary x_σ^* associated to the local volatility function

$$\sigma(x) = \frac{\sqrt{2(ry^*(x) - \delta x)(y^*(x) - x)}}{x\sqrt{y^*(x)'}}.$$

Now by (5.27), y^* is the call exercise boundary associated with the dual volatility function :

$$\tilde{\sigma}(y) = \frac{\sqrt{2(y - x_\sigma^*(y))(ry - \delta x_\sigma^*(y))y'(x_\sigma^*(y))}}{y}.$$

Let us consider the family of exercise boundaries

$$y^*(x) = x \frac{x + a}{bx + c}$$

where a, b, c are positive constants such that $\max(c/a, b) < \min(1, r/\delta)$ (condition ensuring $y^* \in \tilde{\mathcal{C}}_y$). Since $y^*(x)' = (bx^2 + 2cx + ac)/(bx + c)^2$, one has

$$\sigma(x) = \sqrt{2 \frac{((r - \delta b)x + ra - \delta c)((1 - b)x + a - c)}{bx^2 + 2cx + ac}}, x > 0.$$

Moreover, the function $x_\sigma^*(y)$ is the only positive root of the polynomial function: $X^2 + X(a - by) - cy$, that is:

$$x_\sigma^*(y) = \frac{1}{2} \left(by - a + \sqrt{(by - a)^2 + 4cy} \right)$$

and

$$\forall y > 0, \tilde{\sigma}(y) = \frac{\sqrt{2(y - x_\sigma^*(y))(ry - \delta x_\sigma^*(y))(bx_\sigma^*(y)^2 + 2cx_\sigma^*(y) + ac)}}{y(bx_\sigma^*(y) + c)}.$$

This example enables us to check numerically the duality. We have plotted in Figure (5.1, the prices of an American put $P_\sigma(T, x, y)$ in the primal world for the local volatility $\sigma(x)$ and an American call $c_{\tilde{\sigma}}(T, y, x)$ in the dual world for the local volatility $\tilde{\sigma}(x)$ as functions of the maturity T . These prices have been computed using a finite difference method. We can see at $T = 10$ that the limit value is quite reached and both prices are equal. The plots are nonetheless distinct which means that the duality does not hold for finite maturities. We have also plotted, in function of T , $C_\sigma(T, x, y)$ in the primal world and $p_{\tilde{\sigma}}(T, y, x)$ in the dual world to check numerically whether the volatility function $\tilde{\sigma}$ is such as $C_\sigma(x, y) = p_{\tilde{\sigma}}(y, x)$. As we can see, the curves do not seem to converge toward the same limit when T is large. This means that the volatility function $\hat{\sigma}$ such that $\forall x, y > 0, C_\sigma(x, y) = p_{\hat{\sigma}}(y, x)$ (obtained from σ as $\hat{\sigma}$ is obtained from η but with exchange of r and δ) is different from $\tilde{\sigma}$.

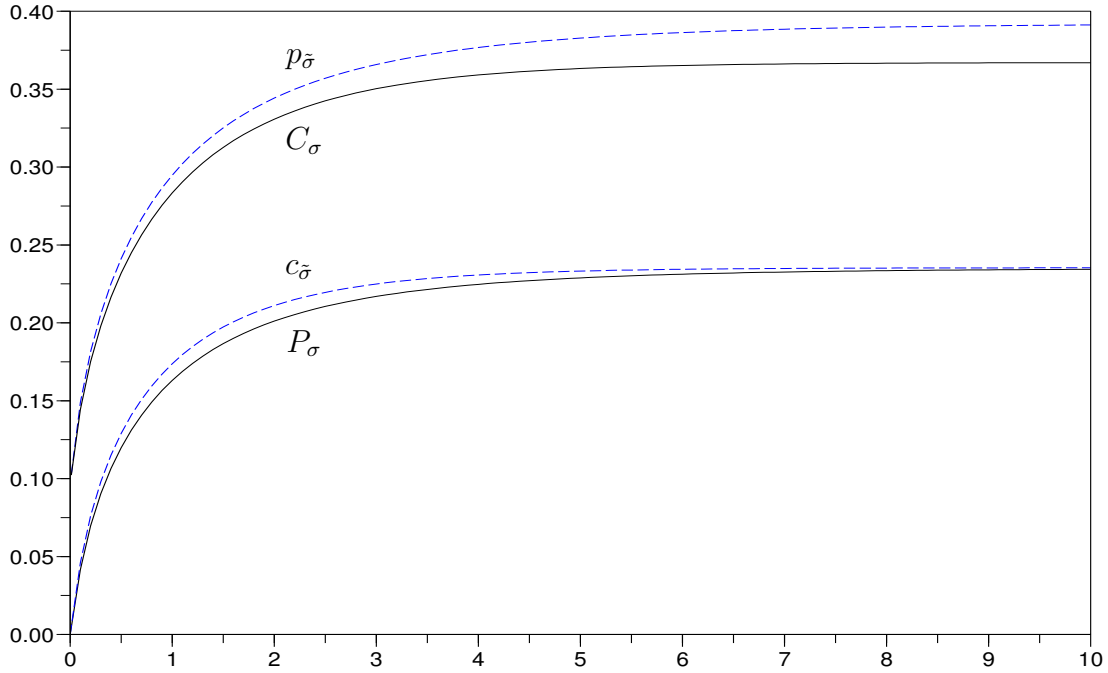


Figure 5.1: $P_\sigma(T, x, y)$ and $c_{\tilde{\sigma}}(T, y, x)$, and $C_\sigma(T, x, y)$ and $p_{\tilde{\sigma}}(T, y, x)$ as functions of the maturity T for $x = 0.5$, $y = 0.4$, $r = 0.2$, $\delta = 0.1$ and the volatility parameters $(a, b, c) = (1, 0.4, 0.1)$.

5.5 Consequence of the duality : A (theoretical) method of calibration for the volatility $\sigma(x)$

In that section, we will put in evidence the importance of the duality within the calibration scope. We suppose for this that we are on a (virtual) market where are traded perpetual securities, and where the short interest rate r and the dividend rate δ can be observed. This means that we know the price of the share x_0 , and the market quotes on that share the perpetual American puts and calls for all strikes $K > 0$. We name respectively $p(K)$ and $c(K)$ these prices and denote:

$$X = \sup\{K > 0, c(K) = x_0 - K\} \text{ and } Y = \inf\{K > 0, p(K) = K - x_0\}. \quad (5.31)$$

We will first suppose that the put and call prices derive from a time-homogeneous local volatility model before relaxing this assumption.

5.5.1 The calibration procedure

Let us assume that there is a volatility function σ satisfying $(\mathcal{H}_{\text{vol}})$ such that for all $K > 0$, $p(K) = P_\sigma(x_0, K)$ and $c(K) = C_\sigma(x_0, K)$. The following proposition says that these prices characterize σ and its proof gives a constructive way to retrieve the volatility function from the prices.

Proposition 5.5.1. *Let us consider $x_0 > 0$. The map*

$$\sigma \mapsto ((P_\sigma(x_0, K), C_\sigma(x_0, K)), K > 0)$$

is one-to-one on the set of volatility functions satisfying $(\mathcal{H}_{\text{vol}})$.

Proof. We first consider the put case. The differential equation satisfied by the put prices in the continuation region makes only appear the values and the derivatives in x , K being fixed. Hence, we cannot exploit directly the prices. But the duality relation enables to get a differential equation in the strike variable. Thanks to the Duality Theorem, we have $P_\sigma(x_0, K) = c_{\tilde{\sigma}}(K, x_0)$ for some $\tilde{\sigma}$ satisfying $(\mathcal{H}_{\text{vol}})$. It is then easy to calibrate $\tilde{\sigma}(\cdot)$. Indeed, one has $\frac{K^2 \tilde{\sigma}(K)^2}{2} p''(K) + K(\delta - r)p'(K) - \delta p(K) = 0$ for $K \leq Y = y_{\tilde{\sigma}}^*(x_0)$. Since the differential equation is valid only for $K < Y$, we only get $\tilde{\sigma}$ on $(0, Y]$ by continuity:

$$\forall K \leq Y, \tilde{\sigma}(K) = \frac{1}{K} \sqrt{\frac{2(\delta p(K) + K(r - \delta)p'(K))}{p''(K)}}$$

which is well defined since $p''(K) = \partial_K^2 c_{\tilde{\sigma}}(K, x) > 0$ (Remark 5.3.2). Then, we can calculate the exercise boundary $y_{\tilde{\sigma}}^*(x)$, for $x \in (0, x_0]$, solving (5.20) supplemented with the final condition $y_{\tilde{\sigma}}^*(x_0) = Y$ backward. This step only requires the knowledge of $\tilde{\sigma}$ only on the interval $(0, Y]$. Finally, we can recover the desired volatility $\sigma(x)$ for $x \leq x_0$ thanks to (5.29):

$$\forall x \in (0, x_0], \sigma(x) = \frac{2(y_{\tilde{\sigma}}^*(x) - x)(ry_{\tilde{\sigma}}^*(x) - \delta x)}{xy_{\tilde{\sigma}}^*(x)\tilde{\sigma}(y_{\tilde{\sigma}}^*(x))}. \quad (5.32)$$

Now let us consider the calibration to the call prices. This relies on the same principle, but we have to be careful because the Duality Theorem is stated given to the call interest rate δ and dividend rate r . So we have to interchange these variables when we apply that theorem. There is a function $\hat{\sigma}$ satisfying $(\mathcal{H}_{\text{vol}})$ such that : $\forall K > 0$, $C_\sigma(x_0, K) = p_{\hat{\sigma}}(K, x_0)$. We have

$$\frac{1}{2} K^2 \hat{\sigma}(x)^2 c''(K) + (\delta - r)Kc'(K) - \delta c(K) = 0$$

for $K > X = \xi_{\hat{\sigma}}^*(x_0)$. Thus, we get

$$\forall K \geq X, \hat{\sigma}(K) = \frac{1}{K} \sqrt{\frac{2(\delta c(K) + K(r - \delta)c'(K))}{c''(K)}}$$

which is well defined for analogous reasons. We can then obtain as before the exercise boundary solving (5.19) forward

$$\forall y \geq x_0, \xi_{\hat{\sigma}}^*(y)' = \frac{\xi_{\hat{\sigma}}^*(y)^2 \hat{\sigma}(\xi_{\hat{\sigma}}^*(y))^2}{2(y - \xi_{\hat{\sigma}}^*(y))(\delta y - r\xi_{\hat{\sigma}}^*(y))}, \xi_{\hat{\sigma}}^*(x_0) = X$$

and we finally get the volatility $\sigma(y)$ for $y \geq x_0$ using the Duality Theorem. More precisely, we interchange r and δ in (5.29) to get

$$\sigma(y) = \frac{2(y - \xi_{\hat{\sigma}}^*(y))(\delta y - r\xi_{\hat{\sigma}}^*(y))}{y\xi_{\hat{\sigma}}^*(y)\hat{\sigma}(\xi_{\hat{\sigma}}^*(y))}. \quad (5.33)$$

□

This calibration method, although being theoretical, sheds light on a striking and interesting result: the perpetual American put prices only give the restriction of $\sigma(x)$ to $(0, x_0]$ and the call prices only the restriction of $\sigma(x)$ to $[x_0, +\infty)$. This has the following economical interpretation : long-term American put prices mainly give information on the downward volatility while long-term American call prices give information on the upward volatility. This dichotomy is remarkable. In comparison, according to Dupire's formula [30], there is no such phenomenon for European options : the knowledge of the call prices gives the whole local volatility surface, not only one part. In other words, the European call and put prices give the same information on the volatility while the perpetual American call and put prices give complementary information.

Thus, one may think that the perpetual American call and put prices only depend on a part of the volatility curve. This is precised by the Proposition below that gives necessary and sufficient conditions on the volatility functions to observe the same put prices (resp. call prices).

Proposition 5.5.2. *Let us consider $x_0 > 0$ and $\sigma_1(\cdot), \sigma_2(\cdot)$ two volatility functions satisfying $(\mathcal{H}_{\text{vol}})$. Then, the following properties are equivalent:*

- (i) $\forall y > 0, P_{\sigma_1}(x_0, y) = P_{\sigma_2}(x_0, y)$ (resp. $\forall y > 0, C_{\sigma_1}(x_0, y) = C_{\sigma_2}(x_0, y)$)
- (ii) $\forall y \leq y_{\hat{\sigma}_2}^*(x_0), \tilde{\sigma}_1(y) = \tilde{\sigma}_2(y)$. (resp. $\forall x \geq \xi_{\hat{\sigma}_1}^*(x_0), \hat{\sigma}_1(x) = \hat{\sigma}_2(x)$ where $\hat{\sigma}_j$ denotes the local volatility function such that $\forall x, y > 0, C_{\sigma_j}(x, y) = p_{\hat{\sigma}_j}(y, x)$.)
- (iii) $\forall x \in (0, x_0], \sigma_1(x) = \sigma_2(x)$ and $y_{\hat{\sigma}_1}^*(x_0) = y_{\hat{\sigma}_2}^*(x_0)$. (resp. $\forall x \in [x_0, +\infty), \sigma_1(x) = \sigma_2(x)$ and $\xi_{\hat{\sigma}_1}^*(x_0) = \xi_{\hat{\sigma}_2}^*(x_0)$.)
- (iv) $\forall x \in (0, x_0], \sigma_1(x) = \sigma_2(x)$ and $\frac{f'_{\downarrow, \sigma_1}(x_0)}{f_{\downarrow, \sigma_1}(x_0)} = \frac{f'_{\downarrow, \sigma_2}(x_0)}{f_{\downarrow, \sigma_2}(x_0)}$. (resp. $\forall x \in [x_0, +\infty), \sigma_1(x) = \sigma_2(x)$ and $\frac{f'_{\uparrow, \sigma_1}(x_0)}{f_{\uparrow, \sigma_1}(x_0)} = \frac{f'_{\uparrow, \sigma_2}(x_0)}{f_{\uparrow, \sigma_2}(x_0)}$.)

- (v) $f_{\downarrow, \sigma_1}$ and $f_{\downarrow, \sigma_2}$ (resp. $f_{\uparrow, \sigma_1}$ and $f_{\uparrow, \sigma_2}$) are proportional on $(0, x_0]$ (resp. $[x_0, +\infty)$).
- (vi) $\forall x \leq x_0, \forall y > 0, P_{\sigma_1}(x, y) = P_{\sigma_2}(x, y)$ (resp. $\forall x \geq x_0, \forall y > 0, C_{\sigma_1}(x, y) = C_{\sigma_2}(x, y)$).

Remark 5.5.3. • Among these many conditions, let us remark that condition (ii) on the dual volatility is much simpler than condition (iii) on the primal volatility since the latter requires the equality of the dual exercise boundaries at x_0 .

- When $\delta = 0$, according to Remark 5.1.3, in the put case, condition (iv) also writes $\forall x \in (0, x_0], \sigma_1(x) = \sigma_2(x)$ and

$$\int_{x_0}^{+\infty} \left(\frac{1}{v^2} \exp \left[- \int_{x_0}^v \frac{2r}{u\sigma_1^2(u)} du \right] \right) dv = \int_{x_0}^{+\infty} \left(\frac{1}{v^2} \exp \left[- \int_{x_0}^v \frac{2r}{u\sigma_2^2(u)} du \right] \right) dv.$$

- Since, by definition of $f_{\downarrow, \sigma_j}$ (resp. $f_{\uparrow, \sigma_j}$) and the strong Markov property $\forall 0 < z \leq x$, $\mathbb{E}[e^{-r\tau_{\sigma_j}^x}] = f_{\downarrow, \sigma_j}(x)/f_{\downarrow, \sigma_j}(z)$ (resp. $\forall 0 < x \leq z$, $\mathbb{E}[e^{-r\tau_{\sigma_j}^x}] = f_{\uparrow, \sigma_j}(x)/f_{\uparrow, \sigma_j}(z)$), the probabilistic counterpart of assertion (v) is $\forall 0 < z \leq x \leq x_0$, $\mathbb{E}[e^{-r\tau_{\sigma_1}^x}] = \mathbb{E}[e^{-r\tau_{\sigma_2}^x}]$ (resp. $\forall x_0 \leq x \leq z$, $\mathbb{E}[e^{-r\tau_{\sigma_1}^x}] = \mathbb{E}[e^{-r\tau_{\sigma_2}^x}]$).

Proof. We consider for example the put case.

(i) $\implies$ (ii) : See the proof of Theorem 5.5.1.

(ii) $\implies$ (iii) : Let us define $\psi(x) = (y_{\tilde{\sigma}_1}^*)^{-1}(y_{\tilde{\sigma}_2}^*(x))/x$. We can show as in the proof of Proposition 5.3.5 that $\psi(x_0) = 1$ and then $\psi \equiv 1$ on $(0, x_0]$, otherwise it would go to 0 or $+\infty$ when $x \rightarrow 0$, which is not possible thanks to (5.17). We get then $\forall x \in (0, x_0], \sigma_1(x) = \sigma_2(x)$ using (5.30) that express σ_j in function of $y_{\tilde{\sigma}_j}^*$ and $\tilde{\sigma}_j$, $j \in \{1, 2\}$.

(iii) $\implies$ (iv) : Thanks to (5.18) and Theorem 5.4.1, we have $\frac{f'_{\downarrow, \sigma_1}(x_0)}{f_{\downarrow, \sigma_1}(x_0)} = \frac{-1}{y_{\tilde{\sigma}_1}^*(x_0) - x_0} = \frac{-1}{y_{\tilde{\sigma}_2}^*(x_0) - x_0} = \frac{f'_{\downarrow, \sigma_2}(x_0)}{f_{\downarrow, \sigma_2}(x_0)}$.

(iv) $\implies$ (v) : The set of solutions to $\frac{1}{2}\sigma_1^2(x)x^2f''(x) + (r - \delta)xf'(x) - rf(x) = 0$ on $(0, x_0]$ is a two-dimensional vector space, but thanks to the relation $\frac{f'_{\downarrow, \sigma_1}(x_0)}{f_{\downarrow, \sigma_1}(x_0)} = \frac{f'_{\downarrow, \sigma_2}(x_0)}{f_{\downarrow, \sigma_2}(x_0)}$, $f_{\downarrow, \sigma_1}$ and $f_{\downarrow, \sigma_2}$ are proportional on $(0, x_0]$:

$$\forall x \leq x_0, f_{\downarrow, \sigma_1}(x) = \frac{f_{\downarrow, \sigma_1}(x_0)}{f_{\downarrow, \sigma_2}(x_0)} f_{\downarrow, \sigma_2}(x). \quad (5.34)$$

(v) $\implies$ (vi) : The proportionality implies that $\forall x \in (0, x_0], \frac{f_{\downarrow, \sigma_1}(x)'}{f_{\downarrow, \sigma_1}(x)} = \frac{f_{\downarrow, \sigma_2}(x)'}{f_{\downarrow, \sigma_2}(x)}$, and then $(y_{\tilde{\sigma}_1}^*(x) - x)^{-1} = (y_{\tilde{\sigma}_2}^*(x) - x)^{-1}$ using (5.18) and Theorem 5.4.1. Therefore

$$\forall x \in (0, x_0], y_{\tilde{\sigma}_1}^*(x) = y_{\tilde{\sigma}_2}^*(x).$$

We have $\alpha_{\sigma_1}(y_{\tilde{\sigma}_1}^*(x))f_{\downarrow, \sigma_1}(x) = \alpha_{\sigma_2}(y_{\tilde{\sigma}_2}^*(x))f_{\downarrow, \sigma_2}(x)$ using (5.13), and obtain from (5.34) that

$$\forall x \leq x_0, \forall y \leq y_{\tilde{\sigma}_1}^*(x_0), \alpha_{\sigma_1}(y) = \frac{f_{\downarrow, \sigma_2}(x_0)}{f_{\downarrow, \sigma_1}(x_0)} \alpha_{\sigma_2}(y) = \frac{f_{\downarrow, \sigma_2}(x)}{f_{\downarrow, \sigma_1}(x)} \alpha_{\sigma_2}(y). \quad (5.35)$$

Thus, we deduce from (5.13), (5.34) and (5.35) the equality of the put prices for the low strikes

$$\forall x \leq x_0, \forall y \leq y_{\tilde{\sigma}_1}^*(x), P_{\sigma_1}(x, y) = P_{\sigma_2}(x, y).$$

For $y > y_{\tilde{\sigma}_1}^*(x) = y_{\tilde{\sigma}_2}^*(x)$, the equality is clear since both prices are equal to $y - x$.

(vi) $\implies$ (i) : clear. $\square$

Let us observe that the point (ii) of the last proposition allows to exhibit different volatility functions with analytic expressions that give the same put (or call) prices. Let us consider the same family as in subsection 5.4.2 coming from the call exercise boundary $y_1^*(x) = x \frac{x+a}{bx+c}$ (assuming $a, b, c > 0$ and $\max(c/a, b) < \min(1, r/\delta)$). For $x_0 > 0$, we introduce the exercise boundary:

$$y_2^*(x) = y_1^*(x) \text{ for } x \leq x_0 \text{ and } y_2^*(x) = y_1^*(x_0) + (y_1^*)'(x_0)(x - x_0) \text{ for } x \geq x_0.$$

The condition $(y_2^* \in \tilde{\mathcal{C}}_y)$ is satisfied provided that $(y_1^*)'(x_0) > \max(1, \delta/r)$. This is automatically ensured by the assumptions made on a, b, c since $(y_1^*)'(x_0) = (bx_0^2 + 2cx_0 + ac)/(b^2x_0^2 + 2bcx_0 + c^2)$. That family is such that $\tilde{\sigma}_1(y) = \tilde{\sigma}_2(y)$ for $y \leq y_2^*(x_0)$. We can then calculate σ_2 as in subsection 5.4.2 using the relation $\sigma_2(x) = \frac{\sqrt{2(r y_2^*(x) - \delta x)(y_2^*(x) - x)}}{x \sqrt{y_2^*(x)'}}$.

This gives $\sigma_2(x) = \sigma(x)$ for $x \leq x_0$ and for $x \geq x_0$,

$$\sigma_2(x) = \sqrt{2 \frac{[(r(y_1^*)'(x_0) - \delta)x + r(y_1^*(x_0) - x_0(y_1^*)'(x_0))][((y_1^*)'(x_0) - 1)x + y_1^*(x_0) - x_0(y_1^*)'(x_0)]}{x^2(y_1^*)'(x_0)}}}.$$

In Figure 5.2, we have plotted the same example as in Figure 5.1 ($x = 0.5$ and $y = 0.4$), adding the graph of $T \mapsto P_{\sigma_2}(T, x, y)$. The volatility function σ_2 has been calculated with the formula above with $x_0 = 0.5$. According to Proposition 5.5.2 and the Duality, the three prices are equal when T is large. In the second example ($x = 3$ and $y = 1$), we still observe that $P_{\sigma}(T, x, y)$ and $c_{\tilde{\sigma}}(T, y, x)$ converge toward the same value when T is large. On the contrary, the limit price of $P_{\sigma_2}(T, x, y)$ is significantly different. To observe the same price, we should have taken, according to Proposition 5.5.2, $x_0 \geq 3$.

5.5.2 Calibration to “real” call and put prices

In that subsection, we address some problems that arise if one tries to apply the calibration procedure when the prices $p(K)$ and $c(K)$ do not derive from a time-homogeneous model.

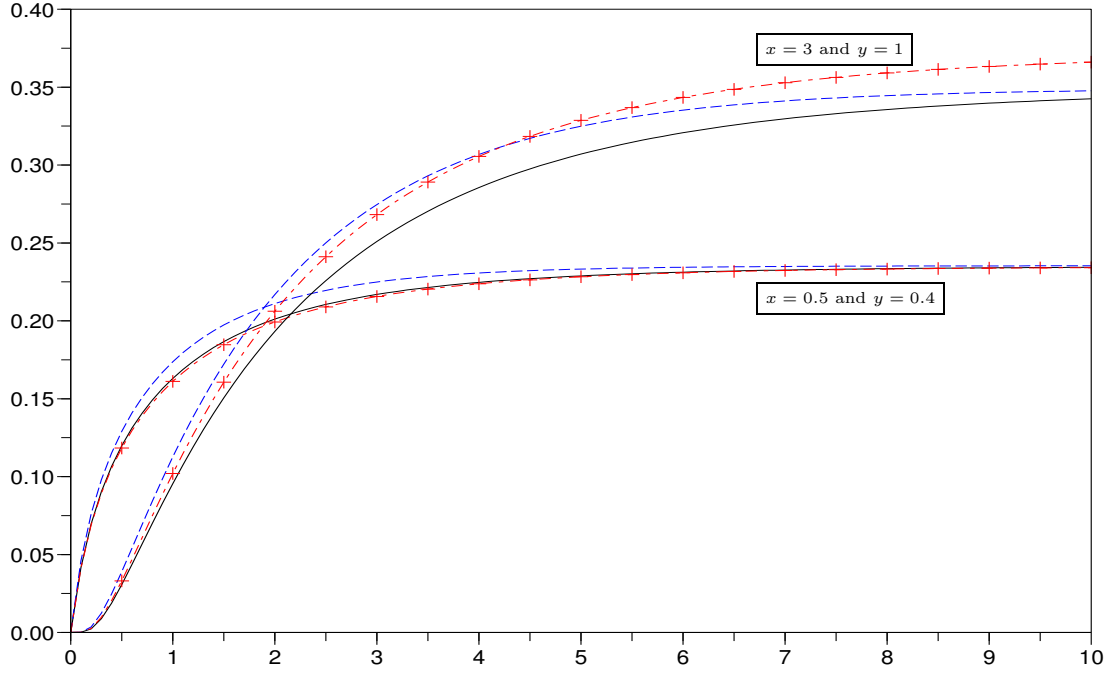


Figure 5.2: $P_\sigma(T, x, y)$ (solid line), $P_{\sigma_2}(T, x, y)$ (dashed line with crosses) and $c_{\bar{\sigma}}(T, y, x)$ (dashed line) in function of the time T for $a = 1$, $b = 0.4$, $c = 0.1$, $x_0 = 0.5$, $r = 0.2$ and $\delta = 0.1$.

We assume however that they are smooth functions of the strike K , and focus for example on the calibration to put prices.

Firstly, let us observe that the arbitrage-free theory allows to define a dual volatility as previously by $(0, Y]$:

$$\forall K < Y, \quad \eta_p(K) = \frac{1}{K} \sqrt{\frac{2(\delta p(K) + K(r - \delta)p'(K))}{p''(K)}}. \quad (5.36)$$

Indeed, the payoff convexity in K ensures the positivity of $p''(K)$ and the arbitrage-free assumption ensures that $\delta p(K) + K(r - \delta)p'(K)$ is nonnegative, so that the square-root is well defined. Let us prove the last point and suppose the contrary (i.e. $\exists y > 0$ such that $\frac{d}{dy} e^{\delta y} p(e^{(r-\delta)y}) < 0$) to exhibit an arbitrage opportunity. In that case, there is $z > y$ such that $e^{\delta y} p(e^{(r-\delta)y}) > e^{\delta z} p(e^{(r-\delta)z})$. We then sell one put with strike $e^{(r-\delta)y}$ and buy $e^{\delta(z-y)}$ puts with strike $e^{(r-\delta)z}$. This initial transaction generates a positive flow. The hedging works as follows: naming τ the time at which the put sold is exercised, we have to pay $e^{(r-\delta)y} - S_\tau$. In other words, we receive one share and borrow $e^{(r-\delta)y}$ in cash. We keep this

position until time $\tau + z - y$. At this time, we have exactly $e^{\delta(z-y)}$ shares and puts with strike $e^{(r-\delta)z}$. Thus, we obtain at least $e^{\delta(z-y)}e^{(r-\delta)z} = e^{(r-\delta)y}e^{r(z-y)}$ and we cancel the debt. The next proposition gives sufficient conditions that allow to construct an homogeneous volatility which is consistent to the observed prices.

Proposition 5.5.4. *Let us assume that $K \in \mathbb{R}_+^* \mapsto p(K)$ is a $\mathcal{C}^1$ function, $\mathcal{C}^2$ on $\mathbb{R}_+^* - \{Y\}$ with $Y = \inf\{K > 0 : p(K) = K - x_0\} < +\infty$. Let us also assume that η_p defined by (5.36) is bounded from below and above by two positive constants and admits a left-hand limit in Y . Then, if we extend η_p in any continuous function on $(0, +\infty)$ satisfying $(\mathcal{H}_{\text{vol}})$ still denoted by η_p , we have*

$$\forall K > 0, P_{\eta_p}(x_0, K) = p(K).$$

Notice that once we choose the extended function η_p , we obtain η_p by first solving (5.20) on $\mathbb{R}_+^*$ starting from x_0 with the condition $y_{\eta_p}^*(x_0) = Y$ and then using (5.30).

Proof. The functions $K \mapsto p(K)$ and $K \mapsto c_{\eta_p}(K, x_0)$ solve (5.12). Since we have $0 \leq p(K) \leq K$ for arbitrage-free reasons, both functions go to 0 when $K \rightarrow 0$. Thanks to Remark 5.1.2, they are proportional to $g_{\uparrow}$ and therefore there is $\lambda > 0$ such that:

$$\forall K \leq Y, p(K) = \lambda c_{\eta_p}(K, x_0).$$

The $\mathcal{C}^1$ assumption made on p ensures $p(Y) = Y - x_0$ and $p'(Y) = 1$. This gives $g_{\uparrow}(Y)/g'_{\uparrow}(Y) = Y - x_0$ and therefore $Y = y_{\eta_p}^*(x_0)$ using Lemma 5.3.1. Thus, $c_{\eta_p}(Y, x_0) = Y - x_0 = p(Y)$ and $\lambda = 1$. One concludes with Theorem 5.4.1. $\square$

For the call case, everything works in the same manner, but we need to assume moreover that $c(K) \rightarrow 0$ when $K \rightarrow +\infty$. This is a rather natural hypothesis that plays the same role as $p(K) \rightarrow 0$ when $K \rightarrow 0$.

Proposition 5.5.5. *Let us assume that $K \in \mathbb{R}_+^* \mapsto c(K)$ is a $\mathcal{C}^1$ function, $\mathcal{C}^2$ on $\mathbb{R}_+^* - \{X\}$ with $X = \sup\{K > 0, c(K) = x_0 - K\} > 0$ and $\lim_{K \rightarrow +\infty} c(K) = 0$. Let us also assume that η_c defined by*

$$\forall K > X, \eta_c(K) = \frac{1}{K} \sqrt{\frac{2(\delta c(K) + K(r - \delta)c'(K))}{c''(K)}}$$

is bounded from below and above by two positive constants and admits a right-hand limit in X . Then, if we extend η_c in any continuous function on $(0, +\infty)$ satisfying $(\mathcal{H}_{\text{vol}})$ still denoted by η_c , we have

$$\forall K > 0, C_{\eta_c}(x_0, K) = c(K)$$

where η_c is obtained from η_c like σ from $\hat{\sigma}$ in the end of the proof of Proposition 5.5.1.

Therefore, we are able to find volatility functions that give exactly the put prices and others that give exactly the call prices. Now, the natural question is whether one can find a volatility function σ that is consistent to both the put and call prices. According to Proposition 5.5.2, all the volatility functions η_p (resp. η_c) giving the put (resp. call) prices coincide on $(0, x_0)$ (resp. $(x_0, +\infty)$). The only volatility function possibly giving both the put and call prices is

$$\sigma(x) = \begin{cases} \eta_p(x) & \text{if } x < x_0 \\ \eta_c(x) & \text{if } x > x_0 \end{cases}.$$

We deduce from Proposition 5.5.2 :

Proposition 5.5.6. *Assume that $\eta_p(x_0^-) = \eta_c(x_0^+)$. Then,*

$$\forall K > 0, p(K) = P_\sigma(x_0, K) \text{ and } c(K) = C_\sigma(x_0, K) \text{ iff } x_\sigma^*(Y) = x_0 \text{ and } \Upsilon_\sigma^*(X) = x_0.$$

5.6 The Black-Scholes model: the unique model invariant through this duality

The purpose of that section is to put in evidence the particular role played by the Black-Scholes' model for the perpetual American call-put duality. We have recalled in the introduction that constant volatility functions are invariant by the duality. We have also mentioned that for the European case, the call-put duality holds for all maturities without any change of the volatility function. Here, on the contrary, we are going to prove that if the duality holds for the perpetual American options with the same volatility:

$$\forall x, y > 0 \quad P_\sigma(x, y) = c_\sigma(y, x) \tag{5.37}$$

then, under some technical assumptions, necessarily $\sigma(\cdot)$ is a constant function.

Proposition 5.6.1. *Let us consider a positive interest rate r and a nonnegative dividend rate $\delta < r$. We suppose that the volatility function σ satisfies $(\mathcal{H}_{\text{vol}})$, and is analytic in a neighborhood of 0, i.e.*

$$\exists \rho > 0, \forall x \in [0, \rho), \quad \sigma(x) = \sum_{k=0}^{\infty} \sigma_k x^k. \tag{5.38}$$

Then, (5.37) holds if and only if $\forall x \geq 0, \sigma(x) = \sigma_0$.

We have already shown in the introduction that (5.37) holds in the Black-Scholes' case. So we only have to prove the necessary condition. We decompose the proof into the three following lemmas.

Lemma 5.6.2. *Let us consider a volatility function that satisfies $(\mathcal{H}_{\text{vol}})$. If the dual volatility function $\tilde{\sigma}$ is analytic in a neighborhood of 0, then the boundaries x_σ^* and y_σ^* are also analytic in a neighborhood of 0.*

Lemma 5.6.3. *Let us suppose that σ satisfies $(\mathcal{H}_{\text{vol}})$ and is analytic in a neighborhood of 0. Let us assume moreover that $r > \delta$. If the equality (5.37) holds, σ is constant in a neighborhood of 0:*

$$\exists \rho > 0, \forall y \in [0, \rho], \sigma(y) = \sigma_0.$$

Lemma 5.6.4. *Let us suppose that σ is a constant function on $[0, \rho]$ for $\rho > 0$ satisfying $(\mathcal{H}_{\text{vol}})$ and (5.37). Then, σ is constant on $\mathbb{R}_+$ (and x_σ^* and y_σ^* are linear functions).*

Proof of Lemma 5.6.2. Let us first show that x_σ^* is analytic in 0. Thanks to the relation (5.18), we have $\frac{g(y_\sigma^*(x))}{g'(y_\sigma^*(x))} = y_\sigma^*(x) - x$, and therefore $\frac{g(y)}{g'(y)} = y - x_\sigma^*(y)$. Thus, $x_\sigma^*(y)$ is analytic in 0 iff $\phi(y) = \frac{g(y)}{g'(y)}$ is analytic in 0. Using the relation (5.12) and $\phi' = 1 - \frac{g''}{g'}\phi$, we get that ϕ is solution of

$$\phi'(y) = 1 + \frac{2}{\tilde{\sigma}^2(y)} ((\delta - r)\phi(y)/y - \delta(\phi(y)/y)^2). \quad (5.39)$$

Notice that $\phi(y) = y - x_\sigma^*(y)$ and (5.16) imply that if ϕ is analytic in 0 then the coefficient of order 0 in its expansion vanishes and the coefficient of order 1 belongs to $(0, 1)$.

To complete the proof we are first going to check that if $\psi(y) = \sum_{k=1}^{\infty} \phi_k y^k$ with $\phi_1 \in (0, 1)$ solves (5.39) in a neighborhood of 0 then $\phi \equiv \psi$ in this neighborhood. Then we will prove existence of such an analytic solution ψ . We have $\psi(0) = 0$, and the function ψ being analytic with $\phi_1 \neq 0$, its zeros are isolated points. There is therefore a neighborhood of 0, $(0, 2\epsilon)$ where ψ does not vanish. Let us consider γ a solution of $\gamma' - \frac{1}{\psi}\gamma = 0$ starting from $\gamma(\epsilon) \neq 0$ in $\epsilon : \gamma(x) = \gamma(\epsilon) \exp\left(\int_\epsilon^x \frac{1}{\psi(u)} du\right)$. Since ψ solves (5.39), it is not hard to check that γ is solution of (5.12) with $\eta = \tilde{\sigma}$. The limit condition $\gamma(x) \xrightarrow{x \rightarrow 0} 0$ (cf. Remark 5.1.2, still valid for $g_\uparrow$ when $\delta = 0$) is satisfied since we have $\frac{1}{\psi(u)} \underset{u \rightarrow 0}{\sim} \frac{1}{\phi_1 u}$ and so $\int_\epsilon^x \frac{1}{\psi(u)} du \xrightarrow{x \rightarrow 0} -\infty$. Thus we have $\gamma(y) = c g(y)$ with $c \neq 0$ and $\psi(y) = g(y)/g'(y) = \phi(y) = y - x_\sigma^*(y)$. We can then write $x_\sigma^*(y) = (1 - \phi_1)y - \sum_{k=2}^{\infty} \phi_k y^k$ in the neighborhood of 0 with $1 - \phi_1 > 0$. It is well-known that in that case, the reciprocal function y_σ^* is also analytic in 0.

Let us turn to the existence of ψ . Since $\sigma_0 \geq \underline{\sigma} > 0$, $y \rightarrow \frac{2}{\tilde{\sigma}^2(y)}$ is an analytic function in the neighborhood of 0. Thus, there is $\rho_0 > 0$ and $a_0 > 0$ such that

$$\forall y \in [0, \rho_0], \frac{2}{\tilde{\sigma}^2(y)} = \sum_{k=0}^{\infty} a_k y^k \text{ and } \sum_{k=0}^{\infty} |a_k| \rho_0^k < \infty.$$

The analytic function $\sum_{k \geq 1} \phi_k y^k$ solves (5.39) if and only if

$$\sum_{k=0}^{\infty} (k+1) \phi_{k+1} y^k = 1 + (\delta - r) \sum_{k=0}^{\infty} \left(\sum_{i+j=k} a_i \phi_{j+1} \right) y^k - \delta \sum_{k=0}^{\infty} \left(\sum_{i+j+l=k} a_i \phi_{j+1} \phi_{l+1} \right) y^k.$$

Identifying the terms of order 0, we get that ϕ_1 solves $P(\phi_1) = 0$ where $P(x) = \delta a_0 x^2 + (1 - (\delta - r)a_0)x - 1$. Since $P(0) = -1 < 0$ and $P(1) = r a_0 > 0$, the polynomial P admits a unique root on $(0, 1)$ and we choose ϕ_1 equal to this root. Then, by identification of the terms with order k , we define the sequence $(\phi_k)_{k \geq 1}$ inductively by

$$\phi_{k+1} = \frac{(\delta - r) \sum_{i+j=k, j \neq k} a_i \phi_{j+1} - \delta \sum_{i+j+l=k, j \neq k, l \neq k} a_i \phi_{j+1} \phi_{l+1}}{k+1 + (r - \delta)a_0 + 2\delta a_0 \phi_1}.$$

This ratio is well defined since $(r - \delta)a_0 + 2\delta a_0 \phi_1 = \delta a_0 \phi_1 + 1/\phi_1 - 1 > 0$.

We still have to check that the series $\sum_{k \geq 1} \phi_k y^k$ is defined in a neighborhood of 0. To do so, we are going to show that there is $\rho > 0$ such that the sequence $(|\phi_k| \rho^k)_{k \geq 1}$ is bounded. We have for $1 \leq k \leq n$:

$$|\phi_{k+1}| \rho^k \leq \frac{|\delta - r| \sum_{j=0}^{k-1} |a_{k-j}| \rho^{k-j} |\phi_{j+1}| \rho^j + \delta \sum_{i=0}^k \left(\sum_{\substack{j+l=k-i \\ j \neq k, l \neq k}} |\phi_{j+1}| \rho^j |\phi_{l+1}| \rho^l \right) |a_i| \rho^i}{k+1}.$$

Let us suppose that for $1 \leq j < k$, $|\phi_{j+1}| \rho^j \leq 1/(j+1)$. Then,

$$|\phi_{k+1}| \rho^k \leq \frac{|\delta - r| \rho \sum_{j=1}^k |a_j| \rho^{j-1} + \delta \sum_{i=0}^k \left(\sum_{j+l=k-i} \frac{1}{j+1} \frac{1}{l+1} \right) |a_i| \rho^i}{k+1}.$$

We remark that $\sum_{j+l=k-i} \frac{1}{j+1} \frac{1}{l+1} = \frac{1}{k-i+2} \sum_{j+l=k-i} \frac{1}{j+1} + \frac{1}{l+1} \leq 2 \frac{\ln(k-i+1)+1}{k-i+2}$, and we finally get:

$$|\phi_{k+1}| \rho^k \leq \frac{2\delta |a_0| \frac{\ln(k+1)+1}{k+2} + \rho(|\delta - r| + 2\delta) \sum_{j=1}^k |a_j| \rho^{j-1}}{k+1} \quad (5.40)$$

since $\frac{\ln(k-i+1)+1}{k-i+2} \leq 1$. Let us now consider k_0 such that $\forall k \geq k_0, 2\delta |a_0| \frac{\ln(k+1)+1}{k+2} < 1/2$. Now, we chose $\rho \in (0, \rho_0)$ small enough such that $\forall k \leq k_0, |\phi_{k+1}| \rho^k \leq 1/(k+1)$ and $\rho(|\delta - r| + 2\delta) \sum_{j=1}^{\infty} |a_j| \rho^{j-1} < 1/2$. Then we get by induction from (5.40) that $\forall k \geq k_0, |\phi_{k+1}| \rho^k \leq 1/(k+1)$. $\square$

Proof of Lemma 5.6.3. On the one hand, thanks to the assumption, $\sigma = \tilde{\sigma}$ is analytic in 0, and therefore x_σ^* is analytic in 0 thanks to Lemma 5.6.2:

$$\exists \rho > 0, \forall y \in [0, \rho), x_\sigma^*(y) = \sum_{i=1}^{\infty} x_i y^i \text{ and } \sigma(y) = \sum_{i=0}^{\infty} \sigma_i y^i.$$

On the other hand, it is not hard then to deduce from (5.29), $\sigma = \tilde{\sigma}$ and the differential equation (5.19) that

$$x_\sigma^*(y)' = \frac{2(y - x_\sigma^*(y))(ry - \delta x_\sigma^*(y))}{y^2 \sigma(y)^2}. \quad (5.41)$$

From Corollary 5.3.6 and (5.26), we get

$$x_\sigma^*(y)' = \frac{(y - x_\sigma^*(y))(ry - \delta x_\sigma^*(y))((x_\sigma^*)^{-1})'(y)}{((x_\sigma^*)^{-1}(y) - y)(r(x_\sigma^*)^{-1}(y) - \delta y)}. \quad (5.42)$$

Now, we consider $n = \inf\{i \geq 2, x_i \neq 0\}$ and suppose it finite. We can get easily that:

$$\begin{aligned} x_\sigma^*(y) &= x_1 y + x_n y^n + \dots & x_\sigma^*(y)' &= x_1 + n x_n y^{n-1} + \dots \\ (x_\sigma^*)^{-1}(y) &= \frac{1}{x_1} y - \frac{x_n}{x_1^{n+1}} y^n + \dots & ((x_\sigma^*)^{-1})'(y) &= \frac{1}{x_1} (1 - \frac{n x_n}{x_1^n} y^{n-1}) + \dots \end{aligned}$$

and then

$$\begin{aligned} (1 - x_\sigma^*(y)/y)(r - \delta x_\sigma^*(y)/y) &= (1 - x_1)(r - \delta x_1) + x_n(2\delta x_1 - (r + \delta))y^{n-1} + \dots \\ \left(\frac{(x_\sigma^*)^{-1}(y)}{y} - 1\right) \left(r \frac{(x_\sigma^*)^{-1}(y)}{y} - \delta\right) &= \frac{1}{x_1^2} \left\{ (1 - x_1)(r - \delta x_1) + \frac{x_n}{x_1^n} ((r + \delta)x_1 - 2r)y^{n-1} \right\} + \dots \end{aligned}$$

The right hand side of (5.42) has then the following expansion:

$$x_1 \left\{ 1 + \frac{x_n}{(1 - x_1)(r - \delta x_1)} \left[2\delta x_1 - (r + \delta) + \frac{2r}{x_1^n} - \frac{r + \delta}{x_1^{n-1}} \right] y^{n-1} - \frac{n x_n}{x_1^n} y^{n-1} \right\} + \dots$$

The equality of the terms of order $n - 1$ in (5.42) then leads to:

$$n x_n x_1^{n-1} = \frac{x_n}{(1 - x_1)(r - \delta x_1)} [2\delta x_1^{n+1} - (r + \delta)x_1^n - (r + \delta)x_1 + 2r] - n x_n.$$

Since $x_n \neq 0$ and with a simplification we get

$$n(1 + x_1^{n-1}) = \frac{1}{r - \delta x_1} \left[-2\delta x_1^n + (r - \delta) \sum_{k=1}^{n-1} x_1^k + 2r \right]. \quad (5.43)$$

In the case $\delta = 0$ this gives $n(1 + x_1^{n-1}) = x_1^{n-1} + \dots + x_1 + 2$ which is not possible because $x_1 \in (0, 1)$. When $0 < \delta < r$, we denote $\alpha = r/\delta > 1$ and rewrite (5.43):

$$n(1 + x_1^{n-1})(\alpha - x_1) = -2x_1^n + (\alpha - 1)x_1^{n-1} + \dots + (\alpha - 1)x_1 + 2\alpha = \alpha - x_1^n + (\alpha - x_1) \frac{1 - x_1^n}{1 - x_1}.$$

Therefore, $n(1 + x_1^{n-1}) = \frac{\alpha - x_1^n}{\alpha - x_1} + \frac{1 - x_1^n}{1 - x_1} < 2 \frac{1 - x_1^n}{1 - x_1}$ because $\beta \mapsto \frac{\beta - x_1^n}{\beta - x_1}$ is decreasing on $(1, \alpha)$ ($x_1^n < x_1$). To show that this is impossible, we consider $P_n(x) = n(1 + x^{n-1}) - 2 \sum_{k=0}^{n-1} x^k$. We have $P_n(1) = 0$ and for $x < 1$, $P_n'(x) = n(n-1)x^{n-2} - 2 \sum_{k=1}^{n-1} kx^{k-1} = 2 \sum_{k=1}^{n-1} k(x^{n-2} - x^{k-1}) < 0$. Thus P_n is positive on $[0, 1)$ and $P_n(x_1) > 0$ which is a contradiction. $\square$

Proof of Lemma 5.6.4. It is easy to get from (5.19) and $\sigma = \tilde{\sigma}$ that

$$x_\sigma^*(y)' = \frac{x_\sigma^*(y)\sigma(x_\sigma^*(y))}{y\sigma(y)}. \quad (5.44)$$

We have $\sigma(x) = \sigma_0$ for $x \in [0, \rho]$. Since $x_\sigma^*(y)$ solves (5.44) and $x_\sigma^*(y) \leq y$, $x_\sigma^*(y)' = x_\sigma^*(y)/y$ on $[0, \rho]$. Therefore, $x_\sigma^*(y) = x_1 y$ for $y \in [0, \rho]$. Thanks to (5.19), x_1 is the unique root in $(0, \min(1, r/\delta))$ of

$$x_1 \sigma_0^2 = 2(1 - x_1)(r - \delta x_1).$$

Now let us observe that (5.19) gives for $y \in (0, y_\sigma^*(\rho)]$, $x_\sigma^*(y)' = \frac{x_\sigma^*(y)^2 \sigma_0^2}{2(y - x_\sigma^*(y))(ry - \delta x_\sigma^*(y))}$ with $x_\sigma^*(\rho) = x_1 \rho$. Since $y \rightarrow x_1 y$ solves this ODE, for which local uniqueness holds thanks to the Cauchy Lipschitz theorem, we then have $x_\sigma^*(y) = x_1 y$ on $[\rho, y_\sigma^*(\rho)]$ and so $y_\sigma^*(\rho) = (x_\sigma^*)^{-1}(\rho) = \rho/x_1$. Then, (5.44) gives $\sigma_0/\sigma(y) = 1$ on $[\rho, \rho/x_1]$. Thus, we prove by induction on n that $x_\sigma^*(y) = x_1 y$ and $\sigma(y) = \sigma_0$ for $y \in [0, \rho/(x_1)^n]$. This shows the desired result. $\square$

5.7 Conclusions and further developments

Addressing Call-Put duality for American options with finite maturity in models with time-dependent local volatility functions like (5.1) would be of great interest. For the perpetual case treated in this paper, we could take advantage of a very nice feature : in the continuation region, the price of the option writes as the product of a function of the underlying spot price by another function of the strike price. Unfortunately, this product property no longer holds in the general case.

Next, according to our numerical experiments (see figure 5.2), American Put and Call prices computed in infinite maturity dual models may differ for finite maturities. This means that in the case of a time-homogeneous primal local volatility function $\varsigma(t, x) = \sigma(x)$, if there exists a dual local volatility function for some finite maturity T , then this volatility function is either time-dependent or depends on the maturity T . On the contrary, in the European case presented in the introduction, time-homogeneous volatility functions are preserved by the duality.

Let us nevertheless conclude on an encouraging remark. Let $P(T, x, y)$ denote the initial price of the American Put option with maturity T and strike y in the model (5.1) and $x^*(T, y)$ stand for the corresponding exercise boundary such that $P(T, x, y) = (y - x)^+$ if and only if $x \leq x^*(t, y)$. Then the smooth-fit principle writes

$$\begin{cases} P(T, x^*(T, y), y) = y - x^*(T, y) \\ \partial_x P(T, x^*(T, y), y) = -1 \end{cases}.$$

Differentiating the former equality with respect to y yields

$$\partial_x P(T, x^*(T, y), y) \partial_y x^*(T, y) + \partial_y P(T, x^*(T, y), y) = 1 - \partial_y x^*(T, y).$$

With the second equality, one deduces that $\partial_y P(T, x^*(T, y), y) = 1$. Therefore the smooth-fit principle automatically holds for the dual Call option if there exists any.

Bibliographie

- [1] Alfonsi, A. (2002). Construction of copulas with periodic densities, *working paper*.
- [2] Alfonsi, A. (2005). On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods and Applications Volume 11, No. 4, pp. 355-467*.
- [3] Alfonsi, A. and Brigo, D., (2005) New Families of Copulas Based on Periodic Functions. *Communications in Statistics : Theory and Methods, Vol. 34, No. 7, pp 1437-1447*.
- [4] Alfonsi, A. and Jourdain, B., (2006) A Call-Put Duality for Perpetual American Options, *rapport CERMICS [2006-307]*.
- [5] Andreasen, J. and Carr, P. (2002). Put Call Reversal. *Working paper*.
- [6] Avellaneda, M., and Newman, J. (1998). Positive interest rates and nonlinear term structure models. *Working Paper, Courant Institute New York*.
- [7] Bally, V. and Talay, D. (1996). The law of the Euler scheme for stochastic differential equations I : convergence rate of the distribution function, *Probab. Theory Related Fields, Vol. 104, pp. 43-60*.
- [8] Bally, V. and Talay, D. (1996). The law of the Euler scheme for stochastic differential equations : II. Convergence rate of the density, *Monte Carlo Methods Appl., Vol. 2, pp. 93-128*.
- [9] Berkaoui A., Bossy M. and Diop A. (2005). Euler scheme for SDEs with non-Lipschitz diffusion coefficient : strong convergence. *RR-5637, INRIA, Juillet 2005*.
- [10] Beibel, M. and Lerche, R. (1997). A New Look at Optimal Stopping Problems related to Mathematical Finance, *Statistica Sinica Vol. 7, pp. 93-108*.
- [11] Bielecki, T., and Rutkowski, M. (2001), Credit risk : Modeling, Valuation and Hedging. *Springer Verlag*
- [12] Black F., Cox J. C. (1976), Valuing corporate securities : Some effects of bond indenture provisions, *Journal of Finance, Vol. 31, pp. 351-367..*
- [13] Blanchet-Scalliet, C., and Jeanblanc, M. (2004). Hazard rate for credit risk and hedging defaultable contingent claims, *Finance and Stochastics Vol. 8, pp. 145-159*.

- [14] Bossy, M. and Diop, A. (2004). An efficient discretisation scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^a$, a in $[1/2, 1)$. *RR-5396, INRIA, Decembre 2004*.
- [15] Borodin, A.N. and Salminen, P. (2002). Handbook of Brownian Motion - Facts and Formulae, *2nd edition. Birkhauser Verlag, Basel-Boston-Berlin*.
- [16] Bouleau, N. and Hirsch, F. (1991). Dirichlet Forms and Analysis on Wiener Space, *De Gruyter, Berlin*.
- [17] Bouyé E., Durrleman V., Nikeghbali A., Riboulet G., Roncalli T. (2000). Copulas for Finance, A Reading Guide and Some Applications. *Groupe de Recherche Opérationnelle, Crédit Lyonnais, France*.
- [18] Brigo, D. and Alfonsi, A. (2005). Credit default swap calibration and derivatives pricing with the SSRD stochastic intensity model. *Finance and Stochastics, Vol. 9, No. 1, pp 29-42*.
- [19] Brigo, D. and Alfonsi, A. (2003). Credit Default Swaps Calibration and Option Pricing with the SSRD Stochastic Intensity and Interest-Rate Model. *Proceedings of the 6-th Columbia-JAFEE Conference, Tokyo, March 15-16, pp 563-585*.
- [20] Brigo, D., and Mercurio, F. (2001), A deterministic-shift extension of analytically tractable and time-homogeneous short rate models. *Finance and Stochastics Vol. 5, pp. 369-388*.
- [21] Brigo, D., and Mercurio, F. (2001b), Interest Rate Models : Theory and Practice. *Springer-Verlag*.
- [22] Broadie, M. and Kaya, Ö. (2003). Exact simulation of stochastic volatility and other ane jump diusion processes, *Working Paper*.
- [23] Bruyère, R., Cont, R., Copinot, R., Jaeck, C., Féry, L. and Spitz, T. (2004). Les produits dérivés de crédit, *Economica*.
- [24] Cherubini, U., and Luciano, E. (2002). Bivariate option pricing with copulas. *Applied Mathematical Finance, Vol 9 (2), pp. 69-86*.
- [25] Cox, J.C. Ingersoll, J.E. and Ross, S.A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica 53, pp. 385-407*.
- [26] Dayanik, S. and Karatzas, I. (2003). On the optimal stopping problem for one-dimensional diffusions. *Stochastic Processes and their Applications, Vol. 107, pp. 173-212*
- [27] Deelstra, G. and Delbaen, F (1998). Convergence of Discretized Stochastic (Interest Rate) Processes with Stochastic Drift Term, *Appl. Stochastic Models Data Anal. 14, pp. 77-84*.

- [28] Diop, A. (2003). Sur la discrétisation et le comportement à petit bruit d'EDS multidimensionnelles dont les coefficients sont à dérivées singulières, ph.D Thesis, INRIA. (available at <http://www.inria.fr/rrrt/tu-0785.html>)
- [29] Duffie D., and Singleton K. (1999), Modeling term structures of defaultable bonds. *Review of Financial Studies*, Vol. 12, pp. 687-720.
- [30] Dupire, B. (1994). Pricing with a smile. *Risk*, Vol. 7, No. 1, January 1994, pp. 18-20.
- [31] El Karoui, N. , Jeanblanc-Picqué, M., Shreve S. E. (1998). Robustness of the Black and Scholes Formula, *Mathematical Finance*, Vol. 8, No. 2, pp. 93-126.
- [32] Embrechts P., Lindskog F., McNeil A. (2001). Modelling Dependence with Copulas and Applications to Risk Management.
- [33] Fajardo, J. and Mordecki, E. (2003). Put-Call Duality and Symmetry, *Finance Lab Working Papers 54*, Finance Lab, Ibmecc São Paulo.
- [34] Genest C., Rivest L. (1993). Statistical Inference Procedures for Bivariate Archimedean Copulas, *Journal of the American Statistical Association*, 88, 1034-1043.
- [35] Genz, A. and Bretz, F. (2002) Methods for the computation of multivariate t-probabilities. *Journal of Computational and Graphical Statistics*, Vol. 11, pp. 950-971.
- [36] Gerber, H. U. and Shiu, E. S. W. (1994). Martingale Approach to Pricing Perpetual American Options, *Astin Bulletin*, Vol. 24, pp. 195-220.
- [37] Glasserman, P. (2003). Monte Carlo Methods in Financial Engineering *Springer, Series : Applications of Mathematics* , Vol. 53.
- [38] Guyon, J. (2005). Euler scheme and tempered distributions, *preprint CERMICS No. 277*.
- [39] Heston, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, Vol. 6, No. 2, pp. 327-343.
- [40] Hobson, D. G. (1998). Volatility Misspecification, Option Pricing and Superreplication via Coupling, *The Annals of Applied Probability*, Vol. 8, No. 1, pp. 193-205.
- [41] Hürlimann, W. (2004). Fitting bivariate cumulative returns with copulas. *Computational Statistics and Data Analysis*, Vol. 45, pp. 355-372.
- [42] Jarrow, R., Lando, D., and Turnbull, S. (1997). A markov model for the term structure of credit risk spreads, Pricing derivatives on financial securities subject to credit risk. *Review of Financial Studies*, Vol. 10, pp. 481-523.
- [43] Jarrow, R. and Turnbull, S. (1995). Pricing derivatives on financial securities subject to credit risk. *Journal of Finance*, Vol. 50, pp. 53-85.
- [44] Joe, H. (1997). Multivariate Models and Dependence Concepts. *Chapman & Hall*, London.

- [45] Jouanin J.-F., Rapuch G., Riboulet G., Roncalli T. (2001), Modelling dependence for credit derivatives with copulas, *Groupe de Recherche Opérationnelle, Crédit Lyonnais, France*.
- [46] Juri, A., and Wüthrich, M.V. (2002). Copula convergence theorems for tail events. *Insurance Mathematics and Economics*, Vol. 30(3), pp. 405-420.
- [47] Karatzas, I. and Shreve, S. E. (1991). Brownian Motion and Stochastic Calculus, 2nd edition. *Springer, Series : Graduate Texts in Mathematics*, Vol. 113.
- [48] Klebaner, F. (2002). Option price when the stock is a semimartingale. *Elec. Comm.Prob.* 7, pp. 89-93.
- [49] Klugman, S.A., and Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance Mathematics and Economics* Vol. 24(1-2), pp. 139-148.
- [50] Lamberton, D. and Lapeyre, B. (1992). Une introduction au calcul stochastique appliqué à la finance, *Ellipses*. English version (1995) : An Introduction to Stochastic Calculus Applied to Finance, *Chapman and Hall*.
- [51] Lando, D. (1998), On Cox processes and credit-risky securities. *Rev. Derivatives Res.* Vol. 2, pp. 99-120.
- [52] Martini C. (1999). Propagation of Convexity by Markovian and Martingalian Semi-groups, *Potential Analysis*, Vol. 10, pp. 133-175.
- [53] Marshall, A.W. and Olkin I. (1979). Inequalities : theory of majorization and its applications. *Academic Press, New York*.
- [54] Merton R. (1974). On the pricing of corporate debt : The risk structure of interest rates, *Journal of Finance*, Vol. 29, pp. 449-470.
- [55] Milstein, G. N., Repin, Y. M., and Tretyakov, M. V. (2002). Numerical methods for stochastic systems preserving symplectic structure. *SIAM J. Numer. Anal.*, Vol. 40, No. 4 pp. 1583-1604.
- [56] Myneni, R. (1992). The pricing of the American option, *The Annals of Applied Probability*, Vol. 2, pp. 1-23.
- [57] Nelsen, R. (1999). An Introduction to Copulas. *Springer*, New York.
- [58] Rogers, L.C.G. and Williams, D. (2000). Diffusions, Markov Processes and Martingales, 2nd edition. *Cambridge Mathematical Library*.
- [59] Scarsini, M. (1984). On measures of concordance. *Stochastica*, Vol. 8, pp. 201-218.
- [60] Schönbucher, P. (2002). A tree implementation of a credit spread model for credit derivatives, *Journal of Computational Finance*, 2002, Vol. 6, No. 2.
- [61] Schönbucher, P. (2005). Portfolio Losses and the Term Structure of Loss Transition Rates : A New Methodology for the Pricing of Portfolio Credit Derivatives. *Working Paper*.

- [62] Schönbucher, P., and Schubert, D. (2001). Copula-Dependent Default Risk in Intensity Models, *working paper*.
- [63] Spizzichino, F. (2001). Subjective Probability Models for Lifetimes, *Chapman & Hall/CRC*.
- [64] Villeneuve, S. (2004). On the Threshold strategies and smooth-fit principle for Optimal Stopping Problems. *Working paper*.
- [65] Wei, G. and Hu, T. (2002). Supermodular dependence ordering on a class of multivariate copulas. *Statistics and Probability Letters*, Vol. 57(4), pp. 375-385.