

Rapport de Stage de Master, PFE
Deux modèles d'évaluation de produits financiers

Ronan COSTAOUEC
ENPC, Arbitragis

Septembre 2007

Résumé

Ce document traite des divers modèles d'évaluation des produits financiers que nous avons été amenés à étudier au cours de notre stage au sein d'Arbitragis. Il s'intéresse plus précisément à deux d'entre eux : le modèle à volatilité locale et les modèles Lévy-exponentiels (en particulier le modèle à Variance Gamma). Une littérature abondante existe qui aborde ces deux sujets et nous tenterons dans ce rapport, en s'appuyant sur cette dernière, d'offrir une présentation claire des enjeux et outils mathématiques liés à ces deux classes très utilisées de modèles. Le plan de ce travail est le suivant. Dans une première partie nous ferons quelques remarques générales sur la modélisation des cours des actifs financiers, et introduirons quelques concepts et résultats fondamentaux de la théorie du pricing d'options. Dans une seconde partie, nous aborderons en détail le modèle à volatilité locale. Enfin dans une troisième et dernière partie, nous nous intéresserons aux modèles Lévy-exponentiels et plus particulièrement à l'un d'entre-eux : le modèle à Variance Gamma. Au fil des deuxième et troisième parties, nous distinguerons plusieurs problématiques et enjeux de la modélisation financière, dont nous préciserons l'importance et la nature.

Remerciements

Je tiens à remercier l'équipe d'Arbitragis, Momentum qui m'a accueilli au cours de ce stage, riche en enseignements. Je remercie particulièrement MM Tuan Nguyen et Mourad Benali, pour leurs précieux conseils et leur disponibilité. Merci également à Alban, Paul, Julien et Khiemm, dont les connaissances en programmation me furent d'un grand secours ; à Mickael, Nabil et Meriem avec qui il fût très agréable de travailler. Je garderai un très bon souvenir de mon passage à Arbitragis.

Introduction

Le présent rapport est le fruit d'un stage de quatre mois réalisé au sein d'Arbitragis. Il traite des principales problématiques auxquelles nous avons été confrontés durant cette période : le pricing d'options et la calibration dans la cadre des modèles à volatilité locale et Lévy-exponentiels. Ces sujets de travail font l'objet d'une littérature abondante et dense, dont nous n'avons malheureusement pu qu'effleurer une maigre partie. Néanmoins, nous tenterons de présenter de manière aussi rigoureuse que possible les outils mathématiques que chacun des deux modèles sus-cités fait intervenir. Ceci implique en particulier, pour chaque modèle, un nécessaire préalable contextuel et théorique au développement des méthodes numériques d'approximation et d'optimisation, plus directement liées à la "pratique".

Notre développement suit un ordre relativement classique. Dans notre première partie, nous formulons quelques remarques générales sur les marchés financiers, la modélisation du cours des actifs financiers et rappelons quelques concepts et principes fondamentaux relatifs à la théorie générale du pricing d'options. Cette section, bien qu'elle demeure très générale, nous semble une ouverture intéressante aux parties qui suivent en ce qu'elle permet une manière de mise en perspective. Son objectif est un rappel du cadre général de la théorie de valorisation des produits dérivés, et une discussion des liens qu'entretient cette valorisation avec les propriétés statistiques constatées des marchés financiers. Dans une deuxième partie, nous traitons en détail du modèle à volatilité locale, en insistant très fortement sur ce qui fut et demeure un de nos sujets d'investigation, à savoir la mise en place de méthodes de calibration précises, stables et rapides. Ceci nous conduit à éclaircir et préciser théoriquement ces notions de stabilité et précision, notamment au travers d'une évocation de la régularisation des problèmes inverses mal posés. Nous présentons plusieurs approches proposées tirées de l'importante quantité d'articles dédiés à ce problème et expliquons pourquoi nous nous sommes plus particulièrement intéressés à deux d'entre-elles dont nous expliquons en longueur le principe et l'implémentation. Enfin notre troisième et dernière partie aura pour objet les modèles Lévy-exponentiels. Ici encore il s'agit d'un thème particulièrement vaste, et qui réclame une ample introduction théorique, à laquelle nous ne nous sommes pas soustraits. Nous nous intéressons plus particulièrement aux différentes méthodes de pricing dans ces modèles, détaillons leur implémentation, et tentons d'expliquer les attraits et inconvénients de cette classe de modèle.

Table des matières

1	Principes généraux du pricing d'options	5
1.1	Ouverture	5
1.2	Théorie générale du pricing d'options	6
1.3	Analyse statistique des marchés financiers	10
1.4	Discussion	11
2	Le modèle à Volatilité Locale	13
2.1	Origine	13
2.2	L'EDP de Dupire	14
2.3	Calibration du modèle à volatilité locale	18
2.3.1	La nature des données	18
2.3.2	Formulation mathématique du problème de calibration	19
2.3.3	Introduction aux principales méthodes de calibration	20
2.4	Différentes méthodes de calibration	22
2.4.1	Calcul explicite du gradient	28
2.4.2	Description de l'algorithme	29
3	Le modèle à Variance Gamma	30
3.1	Introduction	30
3.2	Processus de Lévy	31
3.2.1	Définitions	31
3.2.2	Décomposition et représentation	31
3.2.3	Subordination	32
3.2.4	Processus de Lévy et martingales	34
3.3	Pricing d'options dans les modèles exp-Lévy	35
3.3.1	Changement de mesure pour les processus de Lévy	35
3.3.2	Pricing risque-neutre	37
3.4	Le modèle à Variance Gamma	39
3.4.1	Construction	39
3.4.2	Modélisation risque neutre	42
3.4.3	Pricing par Monte-Carlo	45
3.4.4	Pricing par EID	46

Chapitre 1

Principes généraux du pricing d'options

1.1 Ouverture

Ce premier chapitre regroupe un certain nombre de définitions et résultats relatifs à la modélisation financière en général, à la théorie du pricing d'options en particulier. Son objectif est de rappeler, le contexte général dans le cadre duquel ce rapport s'inscrit. Il rassemble de manière organisée diverses remarques, approches et énoncés fondamentaux, que j'ai croisé au fil de mon stage. Deux questions fondamentales, liées d'ailleurs, en constituent selon moi la trame : comment juger de la qualité d'un modèle mathématique d'évolution pour le cours des actifs ?, que signifie précisément pricer une option lorsqu'on s'intéresse à des modèles plus généraux que le modèle de Black Scholes ? Il va sans dire que je ne prétends pas ici donner une réponse à ces deux questions, mais plus modestement je tenterai de fournir quelques éléments qui permettent, selon moi de clarifier ces interrogations.

Très grossièrement, on peut juger de la qualité d'un modèle selon deux critères principaux : sa capacité à rendre compte d'une certaine réalité et sa simplicité (intelligibilité et implémentation aisée). Ceci conduit naturellement à se demander : quelles informations constituent dans le contexte présent cette réalité que les modèles d'évolution tentent d'expliquer ? Succinctement, elles sont de deux natures. Les unes proviennent de l'analyse statistique des marchés financiers et correspondent à l'estimation empirique des lois (lois marginales) des processus sur un espace de probabilité filtré $(\Omega, F, (F_t)_{t \geq 0}, P)$. On appelle P la probabilité historique du processus ; elle est donc liée aux propriétés statistiques, telles que l'on peut les observer sur les marchés. Les autres émanent des cotation, disponibles sur les marchés, d'instruments financiers dérivés suffisamment liquides (les options d'achat et de vente dans le cas des actions). Ce deuxième type d'informations est traditionnellement associé à une probabilité dite risque-neutre notée Q , dont on a parfois coutûme de dire qu'elle reflète les anticipations du marché, mais dont nous verrons qu'elle s'interprète, théoriquement du moins, plus exactement comme le choix d'un principe d'évaluation d'une famille précise de contrats sur le sous-jacent, vérifiant certaines propriétés intuitives certes mais néanmoins a priori.

Le chapitre est organisé suivant le plan suivant : dans une première partie, nous nous intéresserons à la théorie générale du pricing d'options dont la grande difficulté nous contraint à un traitement partiel. Nous nous efforcerons néanmoins d'en éclairer les hypothèses, principes et théorèmes cruciaux, afin d'alimenter la réflexion générale dont l'objet a été précisé ci-dessus. Dans une deuxième partie, nous aborderons les résultats fondamentaux de l'analyse statistique des marchés financiers fruits de nombreux travaux de recherche au cours des vingt dernières années et dont les points principaux ont déjà été recensés par de nombreux auteurs. On pourra par exemple consulter [2]. Enfin, notons que la présence de ce paragraphe n'est nullement gratuite, et que ce dernier doit permettre de mieux saisir les enjeux et limites des modèles évoqués dans ce rapport.

1.2 Théorie générale du pricing d'options

Dans ce paragraphe, nous formulons un certain nombre de rappels et précisions. Notre objectif est de donner un cadre théorique général à la problématique du pricing d'options, dans lequel s'inscrivent les modèles étudiés par la suite et dont la généralité suffise à illustrer les potentielles difficultés et ambiguïtés intervenant dès lors que l'on souhaite définir clairement la notion de prix.

Commençons par introduire quelques notations dont nous userons tout au long de ce chapitre. On considère donné, un espace de probabilité filtré $(\Omega, \mathcal{F}, (F_t)_{t \geq 0}, P)$ et un processus stochastique S càdlàg à valeurs dans \mathfrak{R}^{d+1} , qui modélisent respectivement l'ensemble des scénarii de marché et l'évolution des cours des différents actifs :

$$\begin{aligned} S : [0, T] \times \Omega &\rightarrow \mathfrak{R}^{d+1} \\ (t, \omega) &\rightarrow (S_t^i(\omega))_{0 \leq i \leq d} \end{aligned}$$

L'actif S_t^0 est appelé le numéraire. Il sert de référence à la valorisation des différents instruments financiers. On fait l'hypothèse dans la suite que sa valeur au temps t est $S_t^0 = e^{rt}$. La valeur actualisée de tout portefeuille V_t constitué des titres précédents s'exprimera donc :

$$\tilde{V}_t = \frac{V_t}{S_t^0} = V_t e^{-rt}$$

Dans ce cadre on définit la notion d'actif contingent de la manière suivante :

Définition 1.1 *Un actif contingent est une variable aléatoire $H \in Z \subset M(F_T)$ à valeurs dans \mathfrak{R} , où Z désigne un sous-espace de $M(F_T)$, ensemble des variables aléatoires F_T -mesurables.*

La spécification judicieuse de Z , dépend généralement de la loi du processus S fixée a priori. Cependant, on remarque que certaines variables aléatoires devront toujours appartenir à Z , à défaut de quoi le modèle est d'un intérêt limité. C'est notamment le cas des valeurs in fine

des actifs, des options d'achat et de vente, et d'une gamme suffisamment large d'options dites "path-dependant", qui correspondent respectivement aux payoffs suivants :

$$\begin{aligned} H(\omega) &= S_T^i(\omega) \\ H(\omega) &= (S_T^i(\omega) - K)^+, \quad H(\omega) = (K - S_T^i(\omega))^+ \\ H(\omega) &= h((S_t^i(\omega))_{t \in [0, T]}) \end{aligned}$$

Remarque 1.1 *On peut raisonnablement supposer que pour tout $A \in F$, l'indicatrice 1_A est dans Z . On considère en outre dans la suite que c'est effectivement le cas.*

L'objectif principal de la théorie générale du pricing d'option est de définir une notion cohérente de prix, pour les variables aléatoires du sous-espace Z . Ceci conduit à se demander quelles propriétés nous souhaitons voir vérifiées par cette notion. La définition suivante précise cela :

Définition 1.2 *On appellera règle de pricing cohérente, toute application :*

$$\Pi : [0, T] \times Z \rightarrow \mathfrak{R}$$

qui vérifie les propriétés suivantes :

– caractère non-anticipant :

$$\forall t \in [0, T], \Pi_t \text{ est } F_t\text{-mesurable}$$

– positivité :

$$H \geq 0 \Rightarrow \forall t \in [0, T], \Pi_t(H) \geq 0$$

– additivité :

$$\forall t \in [0, T], \quad \forall H_1, H_2 \in Z, \quad \Pi_t(H_1 + H_2) = \Pi_t(H_1) + \Pi_t(H_2)$$

Remarque 1.2 *Les hypothèses de linéarité peuvent être relâchées. Ceci est du aux différents effets d'échelle, qui interviennent éventuellement dans le cadre de la gestion d'un book de taille importante.*

Nous allons voir à présent comment on peut heuristiquement construire une nouvelle mesure de probabilité Q liée à toute règle de pricing vérifiant les hypothèses précédentes. Pour cela, on commence par remarquer que $1_\Omega \in Z$ et $1_\Omega = 1$. Or il est clair que $\Pi_t(1) = e^{-r(T-t)}$, c'est la définition même d'un zéro-coupon. On définit alors une application

$$\begin{aligned} Q : F &\rightarrow \mathfrak{R} \\ A &\rightarrow Q(A) = \frac{\Pi_0(1_A)}{\Pi_0(1)} = e^{rT} \Pi_0(1_A) \end{aligned}$$

Les propriétés de Π impliquent alors les propriétés suivantes pour Q

$$\forall A \in F, \quad 0 \leq Q(A) \leq 1$$

$$\forall A, B \in F, \quad A \cap B = \emptyset, \quad Q(A \cup B) = Q(A) + Q(B)$$

En d'autres termes, moyennant une extension de la seconde propriété au cas d'une union dénombrable d'événements, on constate que Q est une mesure sur (Ω, F) . Inversement, si l'on se donne une application Q vérifiant les propriétés précédentes, alors peut définir une règle de pricing de la manière suivante :

$$\forall H \in Z, \quad H = \sum_{i=1}^n \alpha_i 1_{A_i}, \quad \Pi_0(H) = \sum_{i=1}^n \alpha_i Q(A_i)$$

Si l'on suppose de plus que Π vérifie une hypothèse de continuité, alors plus généralement on peut définir :

$$\forall H \in Z, \quad \Pi_0(H) = E^Q[H]$$

L'enseignement principal de ce paragraphe est qu'il est équivalent, sous certaines hypothèses, de se donner un règle de pricing vérifiant quelques propriétés intuitives et se donner une mesure Q sur l'ensemble des événements de marché. Q est une mesure de pricing, appelée probabilité risque neutre : elle attribue une valeur aux événements de F , et n'a a priori rien à voir avec une quelconque perception de marché. Notons enfin qu'on peut reproduire le même raisonnement à tout instant.

On considère à présent une règle de pricing au sens de la définition (1.2), équivalente à la donnée d'une mesure Q . On cherche à préciser le lien entre deux notions fondamentales des mathématiques financières : celle d'absence d'opportunité d'arbitrage et celle de mesure martingale équivalente. Nous allons voir que cela revient essentiellement à répondre à la question : quelles contraintes doit vérifier Q (règle de pricing) afin d'être compatible avec la donnée initiale de la mesure P qui reflète l'évolution historique ou objective du cours des actifs. A cet effet, on commence par rappeler la définition suivante :

Définition 1.3 *Le modèle d'évolution historique du sous-jacent P est dit sans opportunité d'arbitrage s'il n'existe aucune stratégie auto-finançante admissible $\phi = \{(\phi_t^i)_{1 \leq i \leq d+1}\}_{t \in [0, T]}$, dont la valeur $(V_t(\phi))_{t \in [0, T]}$ vérifie :*

$$P(\forall t \in [0, T], V_t(\phi) \geq 0) = 1, \quad \text{et} \quad P(V_T(\phi) > V_0(\phi)) > 0$$

Remarque 1.3 *Cette notion dépend donc du modèle initial, stipulé via P . Mais elle dépend aussi fortement de l'ensemble des stratégies dites admissibles, sous-espace de la classe des processus progressivement mesurables.*

Remarque 1.4 *Notons qu'une conséquence importante de cette hypothèse est la suivante : deux stratégies de même valeur terminale appartenant à la classe des processus pour laquelle l'AOA est vérifiée ont même valeur en chaque instant.*

On suppose le modèle sans opportunité d'arbitrage. On considère un élément A de F tel que $P(A) = 0$. La valeur du contrat de payoff 1_A doit donc être nulle : c'est le prix d'un contrat qui paie 1, lorsqu'un évènement impossible se produit. Or des paragraphes précédents découle que $\Pi_0(1_A) = e^{-rT}Q(A)$, et on en déduit que $Q(A) = 0$. Réciproquement, si on suppose que $Q(A) = 0$, et $P(A) > 0$, alors l'achat de l'option constitue un arbitrage, ce qui contredit notre hypothèse départ. Ce qui finalement conduit au résultat principal suivant :

Proposition 1.1 *Dans un modèle de marché $(\Omega, F, (F_t)_{t \geq 0}, P)$ sans opportunité d'arbitrage, toute règle de pricing Π (au sens de la définition précédente), peut être représentée sous la forme suivante :*

$$\forall H \in Z, \quad \Pi_t(H) = e^{-r(T-t)} E^Q[H|F_t]$$

ou Q est une mesure dite martingale équivalente, telle que

$$- P \sim Q : \quad \forall A \in F \quad Q(A) = 0 \Leftrightarrow P(A) = 0$$

$$- \forall i = 1 \dots d, \quad E^Q[\tilde{S}_T^i | F_t] = \tilde{S}_t^i$$

Les remarques précédentes postulent l'existence d'un règle cohérente de pricing exempte d'opportunités d'arbitrage (d'une probabilité martingale équivalente), mais ce résultat n'a a priori rien d'évident dans un modèle donné. Ce qu'elles illustrent c'est que lorsqu'une mesure martingale existe alors le marché est sans opportunité d'arbitrage. Le Théorème fondamental du pricing d'actifs établit un résultat plus fort : en fait il y a équivalence.

Théorème 1.1 *Le modèle de marché définit par $(\Omega, F, (F_t)_{t \geq 0}, P)$ et le processus $(S_t)_{t \geq 0}$ est sans opportunité d'arbitrage si et seulement si il existe une probabilité $Q \sim P$ tel que le prix des actifs actualisés est une martingale sous Q*

Remarque 1.5 *Il est très difficile de donner une version définitive de ce théorème, dont les hypothèses et le cadre dépendent beaucoup de la nature du processus $(S_t)_{t \geq 0}$ et de la mesure P . Typiquement, dans le cadre des modèles Lévy-exponentiels dont la mesure de Lévy charge \mathfrak{R} tout entier, qui appartiennent à la classe plus générale des semi-martingales non bornées, on doit substituer à la notion d'absence d'opportunité celle d'absence de "Free lunch with vanishing risk" et à celle de mesure martingale celle de mesure σ -martingale, et la démonstration est beaucoup plus lourde que dans le cadre par exemple d'un modèle de diffusion markovienne. La démonstration dans le cas général se trouve dans [5].*

Autre notion fondamentale des mathématiques financières, la complétude exprime le fait que l'on puisse répliquer une certaine gamme d'actifs contingents.

Définition 1.4 *On dit qu'un marché est complet lorsque tout $H \in \tilde{Z} \subset M(F_T)$ est répliquable à l'aide d'une stratégie autofinçante admissible $\phi = \{(\phi_t^i)_{1 \leq i \leq d+1}\}_{t \in [0, T]}$, i.e*

$$\forall H \in \tilde{Z}, \quad \exists \phi = \{(\phi_t^i)_{1 \leq i \leq d+1}\}_{t \in [0, T]} : P(V_T(\phi) = H) = 1$$

Remarque 1.6 Ici encore on remarque que cette notion dépend de l'espace \tilde{Z} et de ce qu'on entend par une stratégie admissible. Comme dans le cas de la propriété d'absence d'opportunité d'arbitrage, la formulation rigoureuse et pertinente de la notion de complétude dépend de la nature des modèles

Avec les mêmes réserves que dans le cas précédent, on peut formuler un deuxième résultat d'importance qui lie l'unicité de la mesure de probabilité martingale équivalente à la complétude.

Théorème 1.2 Le modèle de marché définit par $(\Omega, F, (F_t)_{t \geq 0}, P)$ et le processus $(S_t)_{t \geq 0}$ est complet si et seulement si il existe une unique probabilité $Q \sim P$ telle que le prix des actifs actualisés est une martingale sous Q

Lorsque le marché est complet, on peut définir le prix des actifs contingents de manière non ambiguë, et cette valorisation est par nature liée à une stratégie de couverture. Le problème du pricing est plus délicat lorsque l'on est dans le cas d'un marché incomplet, car on a alors le choix entre plusieurs mesures de pricing, et l'on doit donc ajouter un critère supplémentaire pour définir la notion de prix.

De ce constat découle différentes méthodes de pricing, qui passent outre cette indétermination.

- La technique de sur-réplication consiste en une définition du prix comme valeur initial d'un portefeuille non plus simplement répliquant mais sur-répliquant. C'est à dire d'un portefeuille d'actifs dont la valeur in fine soit strictement supérieure à celle du payoff considéré. Le problème principal d'une telle démarche est qu'elle permet de calculer des prix généralement trop élevés.
- La minimisation du risque quadratique est une méthode très utilisée. Elle réside dans le principe suivant : parmi toutes les mesures martingales équivalentes, on choisit celle qui minimise la variance de l'erreur de couverture. Cette technique fournit d'ailleurs dans le cadre des processus Lévy-exponentiels des résultats satisfaisants. On dispose dans ce cadre de formules explicites pour les quantités à détenir dans le portefeuille de pseudo-couverture (voir [1])
- La minimisation entropique consiste à choisir parmi tous les prix possibles celui qui est issu de la mesure de probabilité qui minimise une certaine distance à la mesure historique. Ce recours à une notion d'entropie est également utilisé par certaines techniques de calibration. Sa principale difficulté est que l'on ne dispose pas a priori d'une bonne connaissance de cette mesure historique.

1.3 Analyse statistique des marchés financiers

L'objectif de cette courte section est d'identifier les caractéristiques principales de la dynamique des actifs financiers sous la probabilité objective P , qu'a révélées l'analyse statistique des données empiriques entreprises depuis de nombreuses années.

- *Épaisseur des queues de distribution* : Les distributions des rendements des différents actifs possèdent des queues épaisses de distribution (on a donc l'excess kurtosis). Ce constat invalide en particulier l'hypothèse selon laquelle le cours des actifs suivraient une diffusion brownienne. Notons que cette non-coïncidence s'accroît à mesure que l'on considère des échelles de temps de plus en plus courtes.
- *Absence d'autocorrélation* : sauf pour des très hautes fréquences (bruit de microstructure)
- *Asymétrie hausse, baisse* : On observe plus de chutes violentes des cours que de hausses violentes.
- *Normalité basse-fréquence* : Lorsqu'on diminue la fréquence à laquelle on échantillonne, la distribution des log-rendements tend vers une distribution gaussienne. Autrement dit, l'hypothèse de la diffusion brownienne est acceptable pour des basses fréquences.
- *Volatility Clustering* : auto-corrélation des valeurs absolues des rendements. Ce constat confirme l'intuition selon laquelle les amplitudes des variations du prix des actifs (à la hausse ou à la baisse) d'un instant à l'autre sont liées.
- *Épaisseur des queues des distributions conditionnelles* : même les distributions conditionnelles, corrigées du phénomène de "Volatility Clustering" possèdent des queues épaisses.
- *Lente décroissance de l'auto-corrélation des valeurs absolues*
- *Effet de levier* : la volatilité d'un actif est négativement corrélée aux rendements de cet actif. Ce qui signifie en somme qu'un marché haussier est moins volatil qu'un marché à la baisse.
- *Corrélation positive entre le volume et volatilité*. Ce phénomène est très intéressant puisqu'il implique une variable dont peu de modèles de pricing tiennent compte : le volume de titres échangés. Il ne signifie somme toute rien d'autre que : plus d'opérations ont lieu sur un titre, plus ce titre est volatil.

1.4 Discussion

A la lumière des paragraphes précédents, nous sommes en mesure de distinguer les difficultés majeures que recelle l'évaluation d'options. Elles sont de deux natures. La première d'entre elle provient de la nature même des données de marché, dont l'analyse statistique a mis en évidence un certain nombre de propriétés qu'il est parfois difficile d'intégrer au sein d'un même modèle historique. La seconde tient à la difficulté à définir une notion claire de prix lorsque le processus historique sous-jacent est suffisamment général ; lorsque l'on souhaite par exemple intégrer la possibilité des sauts. La difficulté dernière consistant à établir dans ce cadre général

le lien entre mesures historiques et d'évaluation. Dans ce contexte, il est très important de bien connaître et garder à l'esprit, les hypothèses simplificatrices à la base d'un modèle donné et qui en constituent le risque propre. Mais les modèles d'évaluations, les équations et formules qu'ils produisent sont à la base d'autres difficultés : numériques par exemple. La suite de ce rapport s'intéresse à deux modèles différents. Dans chacune des deux parties j'ai tenté de bien préciser les hypothèses sous-jacentes, et d'aborder les problèmes tant théoriques que numériques liés à l'évaluation des produits dérivés dans chacun de ces cadres.

Chapitre 2

Le modèle à Volatilité Locale

2.1 Origine

Comme nous l'avons déjà noté précédemment, bien que le modèle de Black Scholes se soit rapidement imposé comme le cadre de référence pour la valorisation des options (à tel point que le prix des calls et des puts négociés sur les marchés financiers s'expriment à présent naturellement en point de volatilité), les praticiens ont très rapidement constaté la non coïncidence entre les prix théoriques fournis par ce modèle et la valeur réelle des options d'achat et de vente, portant sur un sous-jacent donné, cotées sur les marchés. Le phénomène du smile est la preuve manifeste de cette inadéquation. En effet, la courbure de la nappe des volatilités implicites qui donne son nom à ce phénomène est en contradiction avec le modèle de Black Scholes. Si ce modèle correspondait à la réalité des prix d'options, la nappe de volatilité implicite devrait être plane, constante égale à la valeur de la volatilité Black Scholes : ce n'est pas le cas.

Ce constat fut à l'origine de l'élaboration de nouveaux modèles, destinés à expliquer ce phénomène du smile constaté sur les marchés financiers. On définit le modèle à volatilité locale de la manière suivante. On se donne un espace de probabilité filtré $(\Omega, F, (F_t)_{t \geq 0})$ et deux actifs, un risqué l'autre non dont les dynamiques respectives, sous la probabilité historique, sont les suivantes :

$$\begin{aligned}dS_t &= S_t (\mu_t dt + \sigma(t, S_t) dW_t) \\dB_t &= B_t r_t dt\end{aligned}$$

où le taux d'intérêt r_t est supposé déterministe, et la tendance μ_t est aussi supposée déterministe (on peut même faire l'hypothèse supplémentaire μ_t processus prévisible). On utilise dès lors la même méthodologie que dans le cadre du modèle de Black Scholes. On démontre que, moyennant quelques hypothèses supplémentaires sur μ_t , r_t et $\sigma(t, S_t)$ le processus :

$$Z_t = \left(\frac{dQ}{dP} \right)_{F_t} = \exp \left(\int_0^t \frac{\mu_s - r_s}{\sigma(s, S_s)} dW_s - \int_0^t \frac{(\mu_s - r_s)^2}{\sigma(s, S_s)^2} ds \right) \quad (2.1)$$

est bien une martingale. En pratique, si on note $\mu_t = r_t + \lambda_t \sigma(t, S_t)$ (quantité qui s'interprète, en termes financiers, comme une prime de risque), il suffit que la condition suivante d'intégrabilité soit vérifiée (critère de Novikov) :

$$\forall t \geq 0 \quad E \left(\exp \left(\frac{1}{2} \int_0^t \lambda_s^2 ds \right) \right) < \infty \quad (2.2)$$

Cette propriété nous permet donc de conclure à l'existence d'une probabilité martingale équivalente Q . Sous cette nouvelle mesure de probabilité, la dynamique de l'actif risqué devient :

$$dS_t = S_t \left(r_t dt + \sigma(t, S_t) dW_t^Q \right)$$

où $(W_t^Q)_{t \geq 0} = (W_t + \lambda_t)_{t \geq 0}$ est un mouvement brownien sous Q . Cette probabilité martingale équivalente est de plus unique. Cela signifie que comme dans le cas du modèle de Black Scholes, nous disposons d'un modèle de marché complet. Intuitivement, il existe un unique facteur de risque dans notre marché, totalement capté par la dynamique du sous-jacent. La preuve de ce résultat repose sur le théorème de représentation des martingales browniennes. Notons enfin que la donnée a priori d'une dynamique de l'actif risqué sous la probabilité historique impose dans ce modèle le prix des actifs contingents. En d'autres termes, $\sigma(t, S_t)$ ayant été fixée sous P , il ne reste aucun "degré de liberté" sur lequel on puisse jouer en vue d'une quelconque calibration.

Toujours par analogie avec le modèle de Black Scholes, on démontre que le prix $P_t = P(t, S_t)$ (nous sommes dans un cadre markovien) d'une option européenne de payoff $h(S_T)$, où T désigne la maturité de l'option, est solution d'une équation au dérivée partielle, de la forme suivante :

$$\begin{aligned} \frac{\partial P}{\partial t} + r_t S \frac{\partial P}{\partial S} + \frac{1}{2} \sigma(t, S)^2 \frac{\partial^2 P}{\partial S^2} - r_t P &= 0 \\ P(T, S) &= h(S) \end{aligned}$$

2.2 L'EDP de Dupire

La section précédente constitue une présentation succincte du modèle à volatilité locale, mais elle ne répond pas à la question initiale posée par les praticiens, à savoir : quid d'un modèle qui rende compte des prix de marché effectivement observés. L'équation différentielle ci-dessus permet certes de calculer les prix d'options européennes de payoff et maturité fixés, mais elle n'est pas directement liée à la problématique de calibration. Le résultat crucial de ce point de vue est l'oeuvre de Bruno Dupire qui en 1993 dans [6], a le premier proposé une méthode de construction de la volatilité locale à partir des prix des options d'achats fournis par le marché. Sa méthode se base sur une équation aux dérivées partielles, vérifiée uniquement par le prix des calls, et dont les variables sont désormais le strike K et la maturité T . L'énoncé du résultat fondamental est donné ci-dessous (pour simplifier les notations, on suppose le taux d'intérêt constant égal à r).

Théorème 2.1 On considère le processus $(S_t)_{t \geq 0}$ solution de l'EDS :

$$\begin{aligned} dS_t &= S_t(rdt + \sigma(t, S_t)dW_t) \\ S_0 &= x \end{aligned}$$

On suppose que :

- 1. $(S_t)_{t \geq 0}$ est de carré intégrable
- 2. $\forall t > 0$, S_t possède une fonction de densité de probabilité $p(t, x)$ continue sur $]0, +\infty[\times]0, +\infty[$
- 3. Le terme de $x\sigma(t, x)$ est localement Hölder continu : $\forall x > 0$ et pour δ suffisamment petit, il existe $\alpha > 0$ et une fonction continue $C(t)$ telle que $|x - y| < \delta$ implique :

$$|x\sigma(t, x) - y\sigma(t, y)| \leq C(t)|x - y|^\alpha, \quad \forall t > 0$$

Alors le prix des options call de strike K et maturité T , donné par la formule :

$$P(T, K) = e^{-rT} E((S_T - K)^+)$$

est solution de l'équation aux dérivées partielles suivante :

$$\frac{\partial P}{\partial T} = \frac{1}{2}\sigma^2(K, T)K^2 \frac{\partial^2 P}{\partial K^2} - rK \frac{\partial P}{\partial K}, \quad \forall (K, T) \in [0, +\infty[\times]0, +\infty[$$

avec pour condition initiale $C(0, K) = (x - K)^+$

Preuve :

L'idée de la démonstration consiste à appliquer le lemme d'Ito à la fonction $f(x) = x^+$. Malheureusement, cette fonction n'est pas de classe C^2 . Soit alors $\epsilon > 0$, on introduit la fonction :

$$f_\epsilon(x) = \frac{(x + \epsilon)^2}{2\epsilon} 1_{-\epsilon/2 \leq x \leq \epsilon/2} + x 1_{x > \epsilon/2}$$

f_ϵ est de classe C^2 , coïncide avec f sur $|x| > \epsilon/2$ et on a de plus :

$$f'_\epsilon = \frac{x + \epsilon/2}{\epsilon} 1_{-\epsilon/2 \leq x \leq \epsilon/2} + 1_{x > \epsilon/2}, \quad f''_\epsilon = \frac{1}{\epsilon} 1_{-\epsilon/2 \leq x \leq \epsilon/2}$$

On applique alors la formule d'Ito au processus $e^{-rt} f_\epsilon(S_t - K)$ entre les instants T et $T + \theta$. On obtient :

$$\begin{aligned} e^{-r(T+\theta)} f_\epsilon(S_{T+\theta} - K) - e^{-rT} f_\epsilon(S_T - K) &= -r \int_T^{T+\theta} e^{-rt} f_\epsilon(S_t - K) dt \\ &\quad + \int_T^{T+\theta} e^{-rt} f'_\epsilon(S_t - K) dS_t \\ &\quad + \frac{1}{2} \int_T^{T+\theta} f''_\epsilon(S_t - K) \sigma^2(t, S_t) S_t^2 dt \end{aligned}$$

En utilisant les résultats précédents, on a l'égalité suivante pour le dernier terme :

$$\begin{aligned} \int_T^{T+\theta} f_\epsilon''(S_t - K) \sigma^2(t, S_t) S_t^2 dt &= \int_T^{T+\theta} -rt K^2 \sigma^2(t, K) \frac{1}{\epsilon} 1_{K-\epsilon/2 \leq S_t \leq K+\epsilon/2} \\ &+ \int_T^{T+\theta} -rt (S_t \sigma^2(t, S_t) - K^2 \sigma^2(t, K)) \frac{1}{\epsilon} 1_{K-\epsilon/2 \leq S_t \leq K+\epsilon/2} \end{aligned}$$

L'hypothèse 3 du théorème, nous permet de majorer le dernier terme dans la formule précédente par

$$\int_T^{T+\theta} e^{-rt} C(t) \frac{\epsilon^\alpha}{\epsilon} 1_{K-\epsilon/2 \leq S_t \leq K+\epsilon/2} dt$$

Par passage à l'espérance dans la formule d'Itô ci-dessus, on obtient :

$$\begin{aligned} e^{-r(T+\theta)} E(f_\epsilon(S_{T+\theta} - K)) - e^{-rT} E(f_\epsilon(S_T - K)) &= \\ -r \int_T^{T+\theta} e^{-rt} E(f_\epsilon(S_t - K)) dt + \int_T^{T+\theta} e^{-rt} E(f_\epsilon'(S_t - K)) r dt \\ + \frac{1}{2} \int_T^{T+\theta} e^{-rt} K^2 \sigma^2(t, K) \frac{1}{\epsilon} E(1_{K-\epsilon/2 \leq S_t \leq K+\epsilon/2}) dt + O(\epsilon) \end{aligned}$$

L'estimation du terme en $O(\epsilon)$ est obtenue grace à l'approximation ci-dessus, et à l'hypothèse de continuité de la densité. On utilise alors à nouveau la deuxième hypothèse du théorème et on fait tendre ϵ vers 0.

$$\begin{aligned} C(T + \theta, K) - C(T, K) &= -r \int_T^{T+\theta} e^{-rt} C(t, K) dt + r \int_T^{T+\theta} E(S_t 1_{S_t \geq K}) \\ &+ \frac{1}{2} \int_T^{T+\theta} e^{-rt} \sigma^2(t, K) K^2 p(t, K) dt \\ &= rK \int_T^{T+\theta} e^{-rt} P[S_t \geq K] dt + \frac{1}{2} \int_T^{T+\theta} e^{-rt} \sigma^2(t, K) K^2 p(t, K) dt \end{aligned}$$

Par suite on utilise le théorème de la moyenne (en la variable θ) et on obtient alors :

$$\frac{\partial C}{\partial t} = rK e^{-rT} P[S_T \geq K] + \frac{1}{2} e^{-rT} \sigma^2(T, K) K^2 p(T, K)$$

Or grace aux hypothèses de régularité formulées sur la densité de S_t , on peut remarquer :

$$\begin{aligned} e^{-rT} P[S_T \geq K] &= -\frac{\partial C}{\partial K} \\ e^{-rT} p(T, K) &= \frac{\partial^2 C}{\partial K^2} (\sigma^2(t, S) S^2 p(t, S)) + \end{aligned}$$

Et on retrouve bien finalement l'équation aux dérivées partielles de Dupire.

Remarque 2.1 Cette EDP s'interprète également, comme une forme intégrée de l'équation de Fokker-Planck vérifiée par la mesure de probabilité μ_t du processus S_t , moyennant néanmoins (et cela est d'une grande importance) des hypothèses de régularité plus fortes. L'équation de Fokker-Planck s'écrit sous la forme suivante :

$$\frac{\partial \mu_t}{\partial t} - \frac{1}{2} \frac{\partial^2}{\partial S^2} (\sigma^2(t, S) S^2 \mu_t) + \frac{\partial}{\partial t} (r S \mu_t) = 0, \quad \mu_0 = \delta_{S_0}$$

Notons que $\forall t > t_0$, μ_t possède bien une densité et on peut donc écrire $\mu_t(dS) = p(t, S) dS$. On observe en outre, qu'une des conditions pour que cette équation ait un sens, est que la fonction de densité soit C^2 , ce qui constitue bien une hypothèse supplémentaire relativement forte. On peut, à partir de cette équation, retrouver rapidement l'EDP de Dupire. Pour ce faire, on commence par exprimer le prix d'un call à l'aide de la densité du processus :

$$C(0, S_0, T, K) = e^{-rT} \int_K^{+\infty} dS_1 \int_{S_1}^{+\infty} dS_2 \mu_T(S_2)$$

Ensuite, on multiplie l'équation de Fokker-Planck par e^{-rt} et on intègre deux fois. Il est alors facile de remarquer que l'on a les résultats suivants. Pour le premier terme de l'équation :

$$e^{-rT} \int_K^{+\infty} \int_{S_1}^{+\infty} \frac{\partial \mu_T(S_2)}{\partial T} \mu_T(S_2) dS_2 dS_1 = rC(T, K) + \frac{\partial C}{\partial T}$$

Pour le deuxième terme de l'équation

$$e^{-rT} \int_K^{+\infty} \int_{S_1}^{+\infty} \frac{1}{2} \frac{\partial^2 \mu_T(S_2)}{\partial S^2} (\sigma^2 S_2^2 \mu_T(S_2)) dS_2 dS_1 = -\frac{1}{2} \sigma^2(T, K) K^2 \frac{\partial^2 C}{\partial K^2}$$

Et pour le troisième terme de l'équation :

$$e^{-rT} \int_K^{+\infty} \int_{S_1}^{+\infty} \frac{\partial}{\partial S} (r S_2 \mu_T(S_2)) dS_2 dS_1 = rK \frac{\partial C}{\partial K} - rC(T, K)$$

En combinant ces trois égalités, on retrouve l'équation de Dupire.

L'interprétation du théorème 1 est la suivante : lorsque la valeur S_t du sous-jacent suit sous la probabilité risque neutre Q le modèle à volatilité locale (sous les bonnes hypothèses de régularité), alors à partir d'un ensemble complet (continu) de prix de calls, on peut reconstruire la fonction de volatilité du modèle. En effet, la valeur de cette fonction en chaque point est dans ce cas déterminée par la formule suivante :

$$\sigma(T, K) = \sqrt{2 \frac{\frac{\partial C}{\partial T} + rK \frac{\partial C}{\partial K}}{K^2 \frac{\partial^2 C}{\partial K^2}}}$$

Ce résultat est en effet une conséquence directe de l'EDP de Dupire. Mais comme le laisse entendre la remarque précédente, bien qu'il s'agisse d'un résultat puissant, il n'offre qu'une réponse partielle au problème de calibration. On peut énoncer ses principales limites :

- La possibilité de reconstruire de manière unique un processus markovien à partir d'un continuum de prix d'options pour une date et une valeur du spot donnée n'implique pas que le modèle qui produit effectivement ces prix d'options (s'il en existe un) soit nécessairement celui que la méthode évoquée ci-dessus nous permet de déterminer. En d'autres termes, les prix d'options s'ils nous permettent d'avoir accès aux lois marginales du processus ne nous autorisent pas à trancher entre les différents modèles (dont certains d'ailleurs peuvent n'être ni markoviens ni continus) qui peuvent produire les mêmes prix.
- Le Théorème 1 ne nous permet pas de répondre effectivement à la question suivante : étant donné un ensemble continu de prix de call $\{C(T, K)\}_{T \geq 0, K \geq 0}$ existe-t-il un processus markovien continu qui en rende compte. Car en effet, il est tout à fait possible que la formule de Dupire, dont l'application est possible dès que le numérateur et le dénominateur sont positifs (ce qui correspond à des contraintes d'arbitrage sur les Butterfly spread), produisent une fonction de volatilité $\sigma(t, S)$ qui ne vérifie pas les hypothèses du théorème 1. Le résultat suivant permet de préciser et mieux cerner ce cas de figure.

2.3 Calibration du modèle à volatilité locale

Le résultat précédent, comme nous l'avons constaté, est certes très important en ce qu'il a établi le premier un lien entre des prix d'options "observés" sur les marchés et le coefficient d'une diffusion supposée les avoir produit. Mais, force est de constater qu'outre sa limitation intrinsèque (tenant à la nature des hypothèses requises), son utilisation est également compliquée par un certain nombre de faits de marché. Nous nous proposons dans cette partie de recenser une partie de ces "faits de marché" afin d'illustrer les limites du cadre précédent, et d'aboutir à une formulation plus précise du problème de calibration qui tienne compte de la réalité des marchés.

2.3.1 La nature des données

Le théorème 1 suppose que nous disposons, à une date t_0 fixée telle que $S_{t_0} = S_0$ d'un continuum de prix de call $\{C(T, K)\}_{T \geq 0, K \geq 0}$; or ce n'est évidemment jamais le cas. En pratique on dispose d'un nombre fini N de prix d'options $\{C(T_i, K_I)\}_{1 \leq i \leq N}$. La valeur de N varie suivant le sous-jacent, mais il est en moyenne de l'ordre de 20, 25 prix d'options. Notons de plus, pour être parfaitement exact, que l'on dispose en réalité de deux séries de prix : une première série correspondant au prix de vente des options $\{C^{bid}(T_i, K_I)\}_{1 \leq i \leq N}$ et une deuxième série correspondant au prix d'achat $\{C^{ask}(T_i, K_I)\}_{1 \leq i \leq N}$. La conséquence de ces deux constats est une complication importante du problème de calibration.

En effet, pratiquement, le caractère discret des données implique l'impossibilité d'un recours direct à la formule de Dupire, puisque les dérivées qui y figurent perdent leur sens (du reste, comme les hypothèses du théorème 1). Nous sommes en fait face à un nouveau problème. Intuitivement, l'incertitude est plus importante : si dans le cadre continu de l'EDP de Dupire

on était poussé à se demander s'il existait un modèle de diffusion markovienne vérifiant les hypothèses du théorème, capable d'expliquer les prix d'options, on doit à présent se demander s'il n'en existe pas plusieurs (et comment trancher alors ?) et s'interroger sur la continuité ("en un sens à préciser") d'une telle reconstruction. En résumé, nous entrons plus clairement dans la classe des problèmes inverses mal-posés. Nous allons revenir sur cette terminologie et ses implications.

L'existence de deux séries de prix pour les options côtées a également des conséquences importantes. Tout d'abord, en ce qu'elle induit une incertitude sur les données de référence, elle implique une nécessaire stabilité ou régularité des différentes méthodes de calibration. En effet, si une faible variation des données de calibration est susceptible d'induire une variation importante des fonctions de volatilité reconstruites, alors à l'intérieur même de l'intervalle d'erreur lié au spread, on pourra obtenir selon la valeur retenue des fonctions de volatilités très différentes et en conséquence des prix d'options exotiques très différents. Ceci constitue un affre très sérieux. D'autre part une classe importante de méthodes (basées sur l'interpolation de la nappe des prix de marché par exemple), utilisent directement les prix de marché et dans cette optique négligent la distinction entre prix d'achat et de vente (en considérant par exemple une seule série de prix égaux à la moyenne bid-ask); cette approximation est une source potentielle d'introduction d'opportunités d'arbitrage au sein même des données de référence.

2.3.2 Formulation mathématique du problème de calibration

Les techniques de calibration proposées par la littérature qui traite de ce sujet ne tiennent pas toutes compte de cette distinction entre prix d'achat et de vente, qui est néanmoins intégrée le cas échéant dans une procédure de pré-traitement des données de marché afin d'en exclure tout arbitrage. Cette étape est d'une importance cruciale dans le cadre des méthodes basées sur l'interpolation des prix de marché couplée avec la formule de Dupire. Néanmoins dans le cadre des méthodes reposant sur l'adjonction d'un terme de régularisation, cette étape n'est pas toujours nécessaire.

D'une manière générale, on peut formuler le problème de calibration de la manière suivante :

Définition 2.1 *On appelle solution du problème de calibration du modèle à volatilité locale toute fonction $\sigma(t, S)$, telle que :*

$$\forall i \in [1, N], \quad P(\sigma, T_i, K_i) = P^m(T_i, K_i)$$

avec :

- $\{P^m(T_i, K_i)\}_{1 \leq i \leq N}$ qui désigne les prix des options côtées sur S_t
- $\sigma \in H$, espace fonctionnel (paramétrisé ou non suivant la méthode)
- $P(\sigma, T_i, K_i)$ désigne le prix du call de maturité T_i et de strike K_i dans le modèle à volatilité locale défini par le coefficient de diffusion $\sigma = \sigma(t, S)$

Evidemment, le problème ainsi posé n'a pas toujours de solution (c'est le sens de nos précédentes remarques), mais à partir de cette écriture nous verrons comme il est possible et parfois nécessaire de la modifier afin de se ramener à des classes de problèmes dont les solutions possèdent de "bonnes propriétés".

2.3.3 Introduction aux principales méthodes de calibration

Il existe un nombre important d'articles qui traitent de la calibration du modèle à volatilité locale, et une très grande variété de méthodes existe de ce fait. L'objectif de ce paragraphe est d'introduire les deux grandes familles d'approches qui émergent de cette littérature.

Une première famille d'approche se base sur des méthodes d'interpolation des données de marché. Leur principe est globalement le suivant : à partir des données discrètes sur les prix ou volatilités implicites des options liquides cotées sur le marché, on interpole un ensemble complet, au sens de continu, de prix d'options ou de volatilités implicites. On utilise ensuite la formule de Dupire, conséquence du théorème 1, afin d'obtenir la fonction de volatilité locale.

Ce qui distingue ces méthodes les unes des autres, c'est alors principalement l'objet et la méthode d'interpolation. Certains auteurs choisissent par exemple de travailler à partir de la nappe des volatilités implicites, tandis que d'autres préfèrent s'appuyer sur la surface des prix d'options call. Les familles de fonctions choisies en vue de l'interpolation du type de données choisies varient également, bien qu'il s'agisse dans la grande majorité des cas de splines cubiques ou bi-cubiques. L'avantage principal de ces méthodes est que la "régularisation" (même si elle est dans le cas présent très souvent ad hoc) s'opère directement sur les données de marché, c'est à dire sur des variables au sujet desquelles les praticiens conservent une intuition, avec lesquelles ils ont l'habitude de travailler. Le sens de l'approximation semble dans ce cas plus intuitif.

Néanmoins, ces approches présentent de nombreux risques, et un soin tout particulier doit être accordé à l'interpolation. En effet, si cette dernière est effectuée négligemment, de sorte que les surfaces interpolées ne vérifient pas certaines bornes ou propriétés d'arbitrage, il est possible de générer des densités de loi marginale négatives pour le processus de diffusion, qui rendent inapplicable la formule de Dupire. De ce point de vue, il est souvent préférable de raisonner en terme de prix d'options, puisque dans ce cadre on dispose d'expressions relativement simples des contraintes d'arbitrage, qu'on trouvera par exemple dans [8].

Une deuxième famille d'approches repose sur la régularisation du problème de calibration, qui appartient à la classe générale des problèmes inverses mal-posés. L'idée générale sous-jacente à ces méthodes, est de ne plus chercher une solution au problème de calibration au sens de la définition précédente, mais de lui substituer un problème de minimisation d'une fonctionnelle. Cette fonctionnelle s'écrit de manière générale comme la somme de l'erreur de pricing, différence entre les prix de marché et les prix produits par une fonction de volatilité donnée, et un terme de pénalisation, mesure de l'écart relativement à un coefficient de diffusion fixé a priori ou estimation de la régularité de la fonction de volatilité utilisée pour calculer l'erreur

de pricing.

L'avantage de ces méthodes est qu'elles s'inscrivent pour la plupart dans le cadre classique de la régularisation de Tikhonov des problèmes inverses mal-posés. Ainsi, la procédure de calibration, pour ce type d'approches, possède en général de bonnes propriétés : existence de solutions au problème régularisé, continuité (en un sens à préciser selon la nature des espaces considérés) de la solution du problème en les données de marché, convergence de la solution lorsque l'erreur tend vers 0. Notons enfin que pour se ramener à des problèmes d'optimisation fini-dimensionnels, les auteurs font souvent l'hypothèse selon laquelle la volatilité locale que l'on cherche à déterminer appartient à un espace fonctionnel paramétrisable. La définition suivante traduit la forme générale des problèmes régularisés.

Définition 2.2 *On appelle solution du problème régularisé de calibration du modèle à volatilité locale toute fonction $\sigma^\alpha(t, S)$, telle que :*

$$\forall \sigma \in H, \quad J(\sigma^\alpha) \leq J(\sigma)$$

où $J^\alpha(\sigma)$ défini de la manière suivante :

$$J^\alpha(\sigma) = \|P(\sigma) - P^m\|_2^2 + \alpha R(\sigma)$$

avec :

- $\sigma \in H$, espace fonctionnel (paramétrisé ou non suivant la méthode)
- $\|P(\sigma) - P^m\|_2^2 = \sum_{i=1}^N (P(\sigma, T_i, K_i) - P^m(T_i, K_i))^2$
- $R(\sigma)$ terme de régularisation (ex : $R(\sigma) = \|\sigma - \sigma_0\|$, $R(\sigma) = \|\nabla \sigma\|^2$)
- $P(\sigma, T_i, K_i)$ désigne le prix du call de maturité T_i et de strike K_i dans le modèle à volatilité locale défini par le coefficient de diffusion $\sigma = \sigma(t, S)$

Cette approche, même si elle présente de nombreux avantages sus-cités, offre néanmoins un inconvénient majeur. En effet, comme nous l'avons noté précédemment, elle repose sur le choix d'un critère de régularisation, et d'un espace H paramétrisable pour la volatilité locale. De ce fait, l'approximation s'opère directement sur le coefficient de diffusion $\sigma(t, S)$. Or il est délicat de se forger a priori une intuition quant à la forme de cet objet. C'est ce qui peut conduire à une certaine méfiance des praticiens relativement à ce type d'approche, bien moins naturelle a priori que la précédente, mais pourtant plus sûre.

Remarque 2.2 *Les démonstrations des résultats d'existence, de convergence et de continuité des solutions du problème régularisé dépendent des espaces fonctionnels choisis et de la définition du terme de régularisation. Elles sont le plus souvent occultées. On leur substitue des tests pratiques de validation des algorithmes. Néanmoins, un certain nombre d'articles difficiles, nécessitant des bases solides en analyse fonctionnelle examinent ces questions dans des contextes plus ou moins généraux. On pourra par exemple consulter [7] et [8] pour plus de détails.*

2.4 Différentes méthodes de calibration

Dans le cadre de ce stage, nous nous sommes intéressés plus particulièrement à deux méthodes de calibration, qui appartiennent à la seconde des deux grandes familles de méthodes que nous avons identifiées dans le paragraphe précédent. La première d'entre elle est détaillée dans [13], et repose sur une régularisation du problème de calibration selon la méthode sus-citée (définition (2.2)). L'espace fonctionnel paramétrisé est celui des splines cubiques. La procédure de calibration basée sur cet espace fournit des résultats satisfaisants pour des nappes de volatilités implicites "artificielles". Cependant pour des données moins académiques, ses performances sont moins bonnes. Il faut de plus noter qu'elle est relativement instable. Nous allons détailler une procédure alternative de calibration, qui semble fournir de meilleurs résultats. Nous tenterons d'en exposer avec clarté l'esprit, d'en expliquer avec précision les différentes étapes et enfin d'en discuter les avantages et inconvénients.

L'article [9] propose une méthode de régularisation du problème inverse mal-posé qui repose sur un compromis entre l'adéquation aux données de marché et la stabilité (régularité de la fonction de volatilité locale). Ceci se traduit ici également par l'adjonction, dans un premier temps, d'un terme de régularisation de type $\alpha R(\sigma)$. Encore une fois, la détermination du coefficient α de régularisation s'avère être le point délicat de la méthode. Mais à la différence de la méthode précédente, les auteurs optent pour une méthode différente de régularisation dite multi-échelle, dont nous détaillons le principe dans la suite. Notons enfin que la méthode proposée diffère des approches classiques de type "problèmes inverses", auxquelles appartient la méthode précédente, de part les choix suivants d'implémentation et de modélisation :

- La formulation du problème de pricing (problème dit direct) à l'aide de l'EDP de Dupire, et non de l'EDP de Black-Scholes généralisée. L'avantage principal de ce point de vue est que l'évaluation de la fonctionnelle de pricing à chaque itération de l'algorithme d'optimisation ne requiert qu'une seule résolution d'EDP. Ceci permet d'accélérer de manière importante l'algorithme
- La fonction σ appartient à un espace plus gros que dans le cas précédent, celui des splines bicubiques par morceaux $S_{2,m}^3$ ce qui autorise une plus grande variété de solutions au problème de calibration. Le choix de cette famille paramétrique plus "grosse" de fonction rend possible l'approche multi-échelle, détaillée ci-dessous
- La méthode d'optimisation qui repose sur un calcul exact du gradient de la fonctionnelle régularisée en les degrés de liberté, qui définissent de manière unique une fonction de volatilité dans l'espace $H = S_{2,m}^3$ de dimension finie.

Nous présentons dans la suite de cette partie, les étapes, calculs et résultats d'où provient la singularité de cette méthode. Ainsi nous évoquerons tout d'abord la résolution numérique du problème direct (EDP de Dupire)

Reformulation et résolution numérique du problème direct

Afin d'avoir une condition de stabilité moins contraignante, on effectue le changement de variable suivant :

$$\begin{aligned} y &= \ln(K) \\ U(y, T) &= P(K, T) \\ \hat{\sigma}(y, T) &= \sigma(K, T) \end{aligned}$$

L'EDP de Dupire, avec ces notations s'écrit dès lors :

$$\frac{\partial U}{\partial T} = -qU - (r - q + \frac{1}{2}\hat{\sigma}(y, T))\frac{\partial U}{\partial y} + \frac{1}{2}\hat{\sigma}^2(y, T)\frac{\partial^2 U}{\partial y^2}$$

avec pour conditions aux limites :

$$\begin{aligned} \forall y \in \mathfrak{R}, \quad U(t_0, y) &= (S_0 - e^y)^+ \\ \forall t \geq t_0, \quad \lim_{y \rightarrow -\infty} U(t, y) &= 0 \\ \forall t \geq t_0, \quad \lim_{y \rightarrow +\infty} U(t, y) &= S_0 e^{-q(T-t_0)} \end{aligned}$$

Afin d'alléger les notations on définit deux opérateurs différentiels A et \tilde{A} de la manière suivante :

$$\begin{aligned} AU &= -\frac{1}{2}\hat{\sigma}^2(y, T)\frac{\partial^2 U}{\partial y^2} + (r - q + \frac{1}{2}\hat{\sigma}(y, T))\frac{\partial U}{\partial y} \\ \tilde{A}U &= AU + qU \end{aligned}$$

On localise ensuite le problème en se restreignant au domaine $[y_{min}, y_{max}] \times [t_0, T_{max}]$, et on obtient une nouvelle formulation :

$$\begin{aligned} \forall (y, T) \in]y_{min}, y_{max}[\times]t_0, T_{max}[\quad \frac{\partial U}{\partial T} + \tilde{A}U &= 0 \\ \forall T \in]t_0, T_{max}[, \quad U(y_{min}, T) &= S_0 e^{-q(T-t_0)} \\ \forall T \in]t_0, T_{max}[, \quad U(y_{max}, T) &= 0 \\ \forall y \in]y_{min}, y_{max}[, \quad U(y, t_0) = f(y) &= (S_0 - e^y)^+ \end{aligned}$$

On discrétise ensuite l'EDP (et l'opérateur en espace). On adopte les notations suivantes :

$$\begin{aligned} h &= \frac{y_{max} - y_{min}}{N} \\ y_i &= y_{min} + ih, \quad i = 1 \dots N \\ f_i &= f(y_i) \end{aligned}$$

En chaque instant T , on note $u(T) = (U(y_i, T))_{1 \leq i \leq N-1}$, le vecteur de \mathfrak{R}^{N-1} qui contient les prix aux points y_i de la grille d'espace discrétisée. On construit également à cet instant une version discrète de l'opérateur \tilde{A} , que l'on note :

$$\begin{aligned} \tilde{A}_T : \mathfrak{R}^{N-1} &\rightarrow \mathfrak{R}^{N-1} \\ u(T) &\rightarrow \tilde{A}_T u(T) \end{aligned}$$

On s'appuie sur les schéma d'approximation qui suivent pour les dérivées :

$$\begin{aligned}\frac{\partial U}{\partial y}(y_i, T) &\approx \frac{u_{i+1}(T) - u_{i-1}(T)}{2h} \\ \frac{\partial^2 U}{\partial y^2}(y_i, T) &\approx \frac{u_{i+1}(T) - 2u_i(T) + u_{i-1}(T)}{h^2}\end{aligned}$$

Enfin, on pose :

$$\begin{aligned}\alpha_{i,T} &= -\frac{\hat{\sigma}^2(y_i, T)}{2h^2} - \frac{1}{2h}(r - q + \frac{\hat{\sigma}^2(y_i, T)}{2}) \\ \beta_{i,T} &= \frac{\hat{\sigma}^2(y_i, T)}{h^2} + q \\ \gamma_{i,T} &= -\frac{\hat{\sigma}^2(y_i, T)}{2h^2} + \frac{1}{2h}(r - q + \frac{\hat{\sigma}^2(y_i, T)}{2})\end{aligned}$$

On peut dès lors exprimer chaque composante de l'opérateur discrétisé \tilde{A}_T , à l'aide de ces quantités. En effet pour $i = 1 \dots N - 1$ et pour $T \in]t_0, T_{max}]$:

$$(\tilde{A}_T u(T))_i = \alpha_{i,T} u_{i-1}(T) + \beta_{i,T} u_i(T) + \gamma_{i,T} u_{i+1}(T)$$

Restent à traduire les conditions aux limites :

- Pour $i = 0$, on $\forall T$ une condition de Dirichlet $u_0(T) = S_0 e^{-q(T-t_0)}$. On en déduit donc

$$(\tilde{A}_T u(T))_1 = \alpha_{1,T} S_0 e^{-q(T-t_0)} + \beta_{1,T} u_1(T) + \gamma_{1,T} u_2(T)$$

- Pour $i = N$, on $\forall T$ une condition de Dirichlet $u_N(T) = 0$. On en déduit donc

$$(\tilde{A}_T u(T))_{N-1} = \alpha_{N-1,T} u_{N-2}(T) + \beta_{N-1,T} u_{N-1}(T)$$

- Pour $T = t_0$, on a la condition au bord imposée par le payoff :

$$\begin{aligned}f_0 &= S_0 e^{-q(T-t_0)} \\ f_N &= 0\end{aligned}$$

Afin d'alléger les notations, et d'arriver à une formulation matricielle relativement simple du problème discrétisé, on introduit l'opérateur intermédiaire \hat{A} défini ainsi :

$$\begin{aligned}(\hat{A}u(T))_1 &= (\tilde{A}u(T))_1 - \alpha_{1,T} S_0 e^{-q(T-t_0)} \\ (\hat{A}u(T))_i &= (\tilde{A}u(T))_i, \quad \text{pour } i = 2 \dots N - 1\end{aligned}$$

Cette discrétisation en espace nous permet de nous ramener à la résolution d'une équation différentielle ordinaire dont la formulation est la suivante :

$$\begin{aligned}\frac{du}{dT} + \left(\hat{A}u + \alpha_{1,T} S_0 e^{-q(T-t_0)} e_1 \right) &= 0, \quad \forall T \in]t_0, T_{max}] \\ u_0 &= f\end{aligned}$$

avec $f = (f(y_i))_{1 \leq i \leq N-1}$ et e_1 vecteur colonne de taille $N-1$ dont toutes les coordonnées sont nulles à l'exception de la première (cf définition de l'opérateur auxiliaire)

La dernière étape consiste à présent à se donner un schéma de discrétisation en temps de l'EDO, obtenue ci-dessus. L'article de référence choisit à cet effet la méthode des θ -schéma. On fixe les notations suivantes :

$$\begin{aligned} k &= \frac{T_{max} - t_0}{M} \\ T_i &= t_0 + ik, \quad \forall i \in 0 \dots M \\ \hat{A}^i &= \hat{A}_{T_i} \text{ et } u^i = u(T_i) \\ \theta &\in [0, 1] \end{aligned}$$

On calcule alors les valeurs approchées de la solution u^i en résolvant les équations :

$$\begin{aligned} \frac{u^{n+1} - u^n}{k} + \theta \tilde{A}^n u^n + (1 - \theta) \tilde{A}^{n+1} u^{n+1} &= 0, \text{ pour } n = 0 \dots M - 1 \\ u^0 &= f \end{aligned} \quad (2.4)$$

Où on rappelle que $\tilde{A}^n = \hat{A}^n + \alpha_{1, T_n} S_0 e^{-q(T_n - t_0)}$. Cette formulation montre en particulier qu'à chaque itération on est amené à résoudre un système linéaire de type :

$$H^n u^{n+1} = b^n \quad (2.5)$$

avec :

$$\begin{aligned} b^n &= (I^{N-1} - \theta k \hat{A}^n) u^n - S_0 k \left(\theta \alpha_{1, T_n} e^{-q(T_n - t_0)} - (1 - \theta) \alpha_{1, T_{n+1}} e^{-q(T_{n+1} - t_0)} \right) \\ H^n &= I^{N-1} + (1 - \theta) k \hat{A}^{n+1} \end{aligned} \quad (2.6)$$

Remarque 2.3 *Les notations utilisées peuvent sembler lourdes. Cependant, les reformulations successives offrent l'avantage de fournir un cadre clair, nécessaire à l'obtention d'une formule fermée pour le gradient, qui est une des caractéristiques importantes de cette méthode.*

Remarque 2.4 *Lorsqu'on teste la méthode décrite ci-dessus numériquement, i.e lorsqu'on calcule la surface de prix pour une fonction de volatilité donnée, on s'aperçoit que l'erreur de pricing (on se place dans un modèle pour lequel on dispose d'une formule fermée par exemple) est centrée autour du point $K = S_0$ et est d'autant plus importante que l'on est proche de t_0 (dans la zone qui nous intéresse de plus particulièrement). Or, en vue de la calibration, il est indispensable de disposer d'une méthode de pricing aussi précise que possible, puisque la fonctionnelle que l'on cherche à minimiser dépend de manière critique de cette méthode. Ceci conduit à affiner le schéma de discrétisation précédent, notamment en ayant recours à un maillage non uniforme.*

Afin de remédier au problème évoqué dans la remarque précédente, on construit un nouveau maillage non-uniforme en la variable d'espace, tel que le pas soit plus petit dans la zone d'intérêt (autour du point $K = S_0$), même si cela conduit à une moins grande précision pour des valeurs extrêmes de y . En supposant que $y_{min} = -y_{max}$, on peut procéder ainsi :

- on construit une grille uniforme sur $[y_{min}, y_{max}]$
- on calcule l'image de cette grille par la fonction

$$f(x) = y_{max} \tanh(x - y_0)$$

- , où $y_0 = \log(S_0)$
- on note que l'inverse de cette fonction est

$$g(y) = \frac{1}{2} \log((y_{max} + y)/(y_{max} - y))$$

On choisit cette fonction, car ses images inverse sont plus denses autour de S_0 . Pratiquement, ceci nous conduit à définir un nouveau maillage sur $[-y_{max}, y_{max}]$, de la manière suivante :

- on pose $h = 2y_{max}/N$
- On note $i_0 \in 0 \dots N$ le plus petit indice tel que $g(-y_{max} + i_0 h) > -y_{max}$. Entre $-y_{max}$ et $g(-y_{max} + i_0 h)$, l'espace est discrétisé à l'aide de i_0 points : y_0, \dots, y_{i_0-1}
- On note $i_1 \in 0 \dots N$ le plus grand indice tel que $g(-y_{max} + i_1 h) < y_{max}$. Entre $g(y_{max} + i_1 h)$ et y_{max} , l'espace est discrétisé à l'aide de $N - i_1 + 1$ points : y_{i_1+1}, \dots, y_N
- on pose $\forall i \in i_0 \dots i_1, y_i = g(-y_{max} + ih)$

- on définit ensuite pour $i \in 0, \dots, N$:

$$\begin{aligned} f_i &= f(y_i) \\ h_i &= y_{i+1} - y_i \end{aligned}$$

Ce nouveau maillage entraîne des modifications dans la discrétisation de l'EDP, déjà amplement détaillée dans les paragraphes précédents. Ainsi, on approximera désormais les dérivées en les variables d'espaces à l'aide des formules suivantes :

$$\begin{aligned} \frac{\partial U}{\partial y}(y_i, T) &\approx \frac{1}{2} \left(\frac{u_{i+1}(T) - u_i(T)}{h_i} + \frac{u_i(T) - u_{i-1}(T)}{h_{i-1}} \right) \\ \frac{\partial^2 U}{\partial y^2}(y_i, T) &\approx \frac{2}{h_i + h_{i+1}} \left(\frac{u_{i+1}(T) - u_i(T)}{h_i} - \frac{u_i(T) - u_{i-1}(T)}{h_{i-1}} \right) \end{aligned}$$

Les coefficients, qui interviennent dans la définition de l'opérateur discrétisé \tilde{A} , changent également :

$$\begin{aligned} \alpha_{i,T} &= -\frac{\hat{\sigma}^2(y_i, T)}{h_{i-1}(h_i + h_{i-1})} - \frac{1}{2h_{i-1}}(r - q + \frac{\hat{\sigma}^2(y_i, T)}{2}) \\ \beta_{i,T} &= \frac{\hat{\sigma}^2(y_i, T)}{h_i + h_{i+1}} \left(\frac{1}{h_i} + \frac{1}{h_{i-1}} \right) + \frac{1}{2}(r - q + \frac{\hat{\sigma}^2(y_i, T)}{2}) \left(\frac{1}{h_{i-1}} - \frac{1}{h_i} \right) + q \\ \gamma_{i,T} &= -\frac{\hat{\sigma}^2(y_i, T)}{h_i(h_i + h_{i-1})} + \frac{1}{2h_i}(r - q + \frac{\hat{\sigma}^2(y_i, T)}{2}) \end{aligned}$$

Le principe de résolution, la formulation matricielle du problème sont à ce changements près strictement les mêmes que dans le cas d'un maillage uniforme.

Les fonctions splines bi-cubiques par morceaux

Une des caractéristiques de cette méthode tient à l'espace des fonctions choisi pour la détermination de la volatilité locale. En effet, cette dernière est ici supposée appartenir à l'ensemble des splines bi-cubiques par morceaux non-contraints. Cette espace est, nous l'avons déjà noté plus gros que les familles de splines généralement proposées dans la littérature (splines bicubiques simples, naturels) et il autorise un "raffinement" ou une "relache" des hypothèses de régularisation à la base de l'approche multi-échelle. Nous donnons dans ce paragraphe les définitions et résultats principaux relatifs à ces fonctions. L'article de référence est [11]

On choisit dans le cas présent de discrétiser σ , à l'aide des splines bi-cubiques sur une grille bien moins fine que celle utilisée pour la résolution de l'EDP de Dupire. Un tel choix offre deux avantages majeurs : il réduit le nombre de variables associées au problème d'optimisation, et il impose une régularité C^2 à la solution de ce problème.

La grille de discrétisation de $[t_0, T] \times [y_{min}, y_{max}]$ associée au spline bi-cubique divise les intervalles $[y_{min}, y_{max}]$ et $[t_0, T]$ en respectivement n et m sous-intervalles. On définit cette grille de la manière suivante. On pose $y_0 = y_{min}$, $\delta y = (y_{max} - y_{min})/N$, $\delta t = (T - t_0)/M$ et

$$\begin{aligned} \forall i \in 0, \dots, N, \quad y_i &= y_0 + i\delta y \\ \forall j \in 0, \dots, M, \quad t_j &= t_0 + j\delta t \end{aligned}$$

La définition précise d'un spline bicubique par morceaux pour cette grille est la suivante :

Définition 2.3 Soit $R_{i,j} = \{(y, t) : y_{i-1} \leq y < y_i, t_{i-1} \leq t < t_i\}$. Une fonction σ est polynomiale bicubique par morceaux si elle est définie sur chaque rectangle $R_{i,j}$ de la grille par une fonction polynomiale bicubique, i.e

$$\sigma(y, t) = c_{i,j}(y, t) = \sum_{p=0}^3 \sum_{q=0}^3 \alpha_{p,q}^{i,j} (y - y_{i-1})^p (t - t_{i-1})^q, \quad (y, t) \in R_{i,j}$$

On adopte également des notations particulières pour préciser la valeur du spline et de ses dérivées en tous les points de la grille :

$$\begin{aligned} \forall i \in 0, \dots, N, \quad \forall j \in 0, \dots, M, \quad u_{i,j} &= \sigma(y_i, t_j) \\ \forall i \in 0, \dots, N, \quad \forall j \in 0, \dots, M, \quad p_{i,j} &= \frac{\partial \sigma}{\partial y}(y_i, t_j) \\ \forall i \in 0, \dots, N, \quad \forall j \in 0, \dots, M, \quad q_{i,j} &= \frac{\partial \sigma}{\partial t}(y_i, t_j) \\ \forall i \in 0, \dots, N, \quad \forall j \in 0, \dots, M, \quad s_{i,j} &= \frac{\partial \sigma}{\partial y \partial t}(y_i, t_j) \end{aligned}$$

Les degrés de liberté de cette fonction, qui la définissent complètement sont alors :

- les $u_{i,j}$ pour $i \in 0, \dots, N$ et $j \in 0, \dots, M$
- les $p_{i,j}$ pour $i = 0, N$ et $j \in 0, \dots, M$

- les $q_{i,j}$ pour $j = 0, M$ et $i \in 0, \dots, N$
- les $s_{i,j}$ pour $i = 0, N$ et $j = 0, N$

Ce qui fait au total $(n+3) \times (m+3)$ paramètres libres. Une preuve de ce résultat se trouve dans [11]. C'est cette paramétrisation qui sera retenue dans le cadre de cette méthode. Deux propositions importantes sont également démontrées dans [11] : la première fournit une méthode de calcul des dérivées simples et croisées en tous les points de la grille à partir des degrés de liberté grâce à une série de systèmes linéaires, la seconde précise la relation entre les coefficients du polynôme bicubique local sur le rectangle $R_{i,j}$ et les valeurs du spline et de toutes ses dérivées aux coins de $R_{i,j}$. Combinées ces deux propositions nous fournissent une procédure d'évaluation du spline bicubique par morceaux à partir de la donnée des seuls degrés de liberté.

2.4.1 Calcul explicite du gradient

Comme nous l'avons déjà noté plus haut, cette méthode repose sur une approche de type régularisation des problèmes inverses mal posés. Le terme de régularisation $R(\sigma)$ est égal comme dans l'exemple précédent à $\|\nabla\sigma\|_2^2$. On a donc une fonctionnelle $J(\sigma)$ de la forme :

$$\begin{aligned} J(\sigma) &= \|P(\sigma) - P^m\|_2^2 + \alpha\|\nabla\|_2^2 \\ &= G(\sigma) + \alpha F(\sigma) \end{aligned}$$

En tenant compte de la paramétrisation précédente, on peut reformuler le problème de minimisation en terme de degrés de liberté du spline bicubique par morceaux. On notera Σ le vecteur de taille $(n+3)(m+3)$ qui contient leurs valeurs correctement ordonnées

$$J(\Sigma) = G(\Sigma) + \alpha F(\Sigma)$$

Une des idées importantes de l'article original est de parvenir à des formules exactes pour le gradient de chacun des termes de ces fonctionnelles de coût. Cela nécessite une reformulation fastidieuse de la fonctionnelle J en terme de calcul matriciel. Nous ne détaillons pas ici les calculs, qui sont faits quasi-intégralement dans les annexes de [9]. On se contentera de deux constats d'importance.

On note tout d'abord, qu'afin de simplifier le problème, les auteurs substituent à F la fonction de pénalisation :

$$\begin{aligned} F_1(\Sigma) &= \sum_{i=0}^N \sum_{j=0}^M \left(\left(\frac{\partial\sigma}{\partial y}(y_i, t_j) \right)^2 + \left(\frac{\partial\sigma}{\partial t}(y_i, t_j) \right)^2 \right) \\ &= \sum_{i=0}^N \sum_{j=0}^M (p_{i,j}^2 + q_{i,j}^2) \end{aligned}$$

C'est à dire que l'on considère uniquement les valeurs du gradient aux points de la grille. Cela simplifie énormément les calculs, mais rend possible l'apparition de légères irrégularités. Pour remédier à ce travers les auteurs proposent un autre terme de pénalisation simplifié faisant

intervenir les dérivées croisées.

La formulation matricielle de G est superflue, puisqu'on dispose avec le schéma de résolution détaillé ci-dessus d'une méthode précise d'estimation pour le prix des options, lorsque la fonction de volatilité est connue. Or on dispose bien d'une procédure d'interpolation (cf paragraphe précédent). Le calcul du gradient est plus difficile. Il nécessite en particulier un recours à la méthode de l'état adjoint. Une des difficultés majeures provient du fait que ce calcul fait intervenir la dépendance des H^n et b^n par rapport aux degrés de liberté qui définissent le spline. Or la dépendance des opérateurs différentiels, qui constituent ce système, relativement aux degrés de liberté du spline a lieu en chaque points de la grille via la procédure d'interpolation décrite ci-dessus, dont la traduction dans le formalisme du calcul matriciel s'avère très technique. Nous renvoyons à nouveau aux annexes de l'article fondateur pour les détails du calcul.

2.4.2 Description de l'algorithme

Le principe de l'algorithme de calibration dans cette méthode est relativement simple. On se donne des critères d'arrêt : une tolérance d'erreur ϵ et un nombre maximal d'itération $N_{iterations}$. La volatilité sera représentée par un spline bicubique par morceaux.

- On initialise la volatilité $\sigma^{n \times m}$ en fixant, le nombre de points de la grille de discrétisation $(n+1) \times (m+1)$ (associée au spline) et des valeurs cohérentes pour les degrés de liberté.
- Ensuite, on utilise un algorithme de quasi-Newton (de préférence avec contraintes) pour résoudre le problème de minimisation discret, dont nous avons précisé l'obtention plus haut. C'est lors de cette étape que les formules fermées pour la fonctionnelle et son gradient sont particulièrement utiles.
- Si après $N_{iterations}$, la valeur du gradient est inférieure à ϵ , on stoppe l'algorithme. Si ce n'est pas le cas, alors on augmente n et (ou) m , et on recommence toute la procédure, et ce jusqu'à ce qu'on obtienne la convergence de l'algorithme.

Remarque 2.5 *Notons que nous avons dans ce paragraphe postulé que la grille de discrétisation associée au spline était uniforme. Si l'on relâche cette hypothèse, peut-être est-il possible de trouver des techniques de régularisation plus efficace que celle consistant à raffiner globalement le maillage, en privilégiant les zones où la norme du gradient est la plus élevée par exemple.*

Chapitre 3

Le modèle à Variance Gamma

3.1 Introduction

Comme nous l'avons déjà souligné, les praticiens se sont très tôt aperçus des insuffisances du modèle de Black Scholes, notamment du fait de son incapacité à expliquer le phénomène du smile. Mais il existe bien d'autres raisons de s'intéresser à des modèles alternatifs. Une de ces raisons fondamentales provient de l'analyse statistique des cours des différents actifs financiers. En effet, celle-ci montre qu'à des fréquences élevées, l'hypothèse selon laquelle les log-rendements des actifs financiers suivent une loi normale est tout simplement fautive. Elle révèle en outre, que l'épaisseur des queues de distributions empiriques est incompatible avec ce postulat. A partir de ce constat, on comprend la nécessité de s'intéresser à des modèles capables d'expliquer ce phénomène.

Les modèles Lévy-exponentiels sont liés à cet effort d'intégration des résultats de l'analyse statistique des marchés financiers. Il en existe une très grande variété dont nous n'évoquons ici que quelques exemples ; pour un exposé complet et extrêmement intéressant nous renvoyons à [1], ainsi qu'aux thèses de Raible et Krause. Evidemment, ces nombreux modèles possèdent des propriétés souvent très différentes, et n'ont pas tous le même potentiel explicatif. Très grossièrement, leur trait commun réside dans leur capacité à intégrer d'éventuels sauts du sous-jacent. Cela complique en général singulièrement le pricing des options (incomplétude des marchés etc...)

Nous nous intéresserons plus en détail dans le cadre de ce rapport à un modèle Lévy-exponentiel précis que nous avons étudié lors de ce stage : le modèle dit à Variance Gamma. Ce modèle est dans l'esprit de ceux que nous évoquions précédemment, et son examen s'est avéré particulièrement enrichissant. Le chapitre est organisé de la manière suivante : dans un premier temps nous donnons quelques définitions et résultats généraux relatifs aux processus de Lévy, nous traitons ensuite dans un cadre général des modèles Lévy-exponentiels (définitions, difficultés...), pour enfin nous concentrer plus longuement sur le modèle à Variance Gamma, au travers duquel nous aborderons différentes méthodes de pricing (communes à de nombreux modèles Lévy-exp). La présentation et l'explication des différentes méthodes de pricing d'op-

tions européennes dans le cadre du modèle à Variance Gamma, ou dans celui plus général des modèles Lévy-exp constituera la partie la plus importante de cette section.

3.2 Processus de Lévy

3.2.1 Définitions

Nous commençons par rappeler la définition d'un processus de Lévy :

Définition 3.1 *Un processus càdlàg $(X_t)_{t \geq 0}$ sur un espace de probabilité (Ω, F, P) à valeurs dans \mathbb{R}^d , tel que $X_0 = 0$ est un processus de Lévy s'il possède les trois propriétés suivantes :*

- $\forall t_0 < t_1 < \dots < t_n$ les variables aléatoires $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ sont indépendantes
- $\forall h > 0$ la loi de $X_{t+h} - X_t$ est indépendante de t
- $\forall \epsilon > 0, \lim_{h \rightarrow 0} P(|X_{t+h} - X_t| \geq \epsilon) = 0$

Remarque 3.1 *Les processus de Lévy sont du fait de cette définition liés à la notion de distribution infiniment divisible, dont nous ne donnons pas ici de définition précise. On la trouve du reste dans tous les ouvrages de référence sur le sujet.*

Une notion fondamentale et absolument nécessaire à l'étude des processus de Lévy est celle de mesure de Lévy, dont nous donnons une définition ci-dessous.

Définition 3.2 *Soit $(X_t)_{t \geq 0}$ un processus de Lévy à valeurs dans \mathbb{R}^d alors la fonction ν définie sur $B(\mathbb{R}^d)$ par*

$$\nu(A) = E(\{t \in [0, 1], \Delta X_t \neq 0, X_t \in A\})$$

définit une mesure sur \mathbb{R}^d , appelée mesure de Lévy de \cdot . $\nu(A)$ représente le nombre moyen de sauts par unité de temps dont l'amplitude appartient à A

La mesure $\nu(A)$ est en fait une mesure de Radon sur \mathbb{R}^d , puisqu'elle est positive pour tout $A \in B(\mathbb{R}^d)$, et qu'on a nécessairement $\nu(A) < +\infty$ pour tout A compact tel que $0 \notin A$. Notons que si cette dernière propriété n'était pas vérifiée, l'hypothèse càdlàg se trouverait contredite. Cependant ν ne définit pas nécessairement une mesure finie, et il est tout à fait possible pour un processus de Lévy X de posséder une infinité de petits sauts sur un intervalle $[0, T]$. La mesure de Lévy doit néanmoins même dans ce cas vérifier certaines hypothèses d'intégrabilité.

3.2.2 Décomposition et représentation

Un résultat également très important afin de caractériser les processus de Lévy est donné par la proposition suivante appelée décomposition de Lévy-Itô. Elle est à la base de la notion de triplet caractéristique associée aux processus de Lévy. Son énoncé est le suivant :

Proposition 3.1 *Soit $(X_t)_{t \geq 0}$ un processus de Lévy à valeurs dans \mathbb{R}^d et ν sa mesure de Lévy alors on a :*

– ν est une mesure de Radon sur $\mathbb{R}^d - \{0\}$ qui vérifie :

$$\int_{|x| \leq 1} |x|^2 \nu(dx) < +\infty, \quad \int_{|x| \geq 1} \nu(dx) < +\infty$$

– La mesure de saut de X , notée J_X , est une mesure de Poisson aléatoire sur $[0, +\infty[\times \mathbb{R}^d$ d'intensité $\nu(dx)dt$

– Il existe un vecteur $\gamma \in \mathbb{R}^d$ et un mouvement brownien d -dimensionnel de matrice de variance-covariance $A \in \mathbb{R}^{d \times d}$ tels que l'on ait :

$$\begin{aligned} X_t &= \gamma t + B_t + X_t^l + \lim_{\epsilon \rightarrow 0} \tilde{X}_t^\epsilon \text{ avec} \\ X_t^l &= \int_{|x| \geq 1, s \in [0, t]} x J(dx \times ds) \\ \tilde{X}_t^\epsilon &= \int_{|x| \leq 1, s \in [0, t]} x (J(dx \times ds) - \nu(dx)ds) \end{aligned}$$

où les termes intervenant dans la première égalité sont indépendants, et où la convergence du dernier terme est presque sur et uniforme sur tout intervalle $[0, T]$

Le théorème suivant dit de Lévy-Khinchin précise la forme de la fonction caractéristique d'un processus de Lévy. Il s'agit d'une conséquence directe de la décomposition de Lévy-Itô.

Théorème 3.1 Soit $(X_t)_{t \geq 0}$ un processus de Lévy à valeurs dans \mathbb{R}^d de triplet caractéristique (A, ν, γ) alors on a :

$$\begin{aligned} E[e^{iz \cdot X_t}] &= e^{t\psi(z)}, \quad \forall z \in \mathbb{R}^d \\ \text{avec } \psi(z) &= -\frac{1}{2}z \cdot Az + i\gamma \cdot z + \int_{\mathbb{R}^d} (e^{iz \cdot x} - 1 - iz \cdot x 1_{|x| \leq 1}) \nu(dx) \end{aligned}$$

Les résultats précédents sont vraiment fondamentaux et il nous était tout bonnement impossible de ne pas les mentionner. Les définitions, propositions et théorèmes qui suivent sont également très importants et ils s'avèrent très utiles dans le cadre précis de l'étude des modèles Lévy-exponentiels. Un premier résultat intéressant fournit une condition nécessaire et suffisante pour qu'un processus de Lévy soit à variations finies.

Proposition 3.2 Un processus de Lévy $(X_t)_{t \geq 0}$ est à variation finie si et seulement si son triplet caractéristique (A, ν, γ) vérifie la propriété suivante :

$$A = 0 \quad \text{et} \quad \int_{|x| \geq 1} |x| \nu(dx) \leq +\infty$$

3.2.3 Subordination

La proposition suivante a valeur de définition pour une catégorie particulière de processus de Lévy appelés subordinateurs, dont la singularité est d'être croissant presque sûrement, et qui sont très utilisés pour construire de nouveaux processus de Lévy à partir de processus existants, par la technique du "changement de temps".

Proposition 3.3 Soit $(X_t)_{t \geq 0}$ un processus de Lévy à valeurs dans \mathfrak{R} . Les conditions suivantes sont équivalentes :

- 1. $\exists t > 0$ tel que $X_t \geq 0$ presque sûrement
- 2. $\forall t > 0$ $X_t \geq 0$ presque sûrement
- 3. Les trajectoires du processus X_t sont presque sûrement croissantes, ce qui signifie $t \geq s \Rightarrow X_t \geq X_s$ presque sûrement
- 4. Le triplet caractéristique du processus $(X_t)_{t \geq 0}$ satisfait les hypothèses suivantes

$$A = 0, \quad \nu(] - \infty, 0]) = 0, \quad b \geq 0$$

$$\int_0^{+\infty} (x \wedge 1) \nu(dx) < +\infty$$

ce qui signifie que X n'a pas de composante diffusive, seulement un drift positif et des sauts positifs de variations finies

Remarque 3.2 On construit de nombreux processus de Lévy, à partir de processus de Lévy préexistants par subordination. En particulier, le processus à Variance gamma est obtenu de cette manière, ce qui justifie que nous nous atardions ici quelque peu sur cette notion.

Un subordonateur $(S_t)_{t \geq 0}$ est donc un processus de Lévy satisfaisant une des conditions équivalentes de la proposition précédente. On sait donc que pour tout t , S_t est une variable aléatoire positive, ce qui incite à représenter sa loi en ayant recours non plus à la transformée de Fourier mais à la transformée de Laplace, et l'on sait également que son triplet caractéristique est de la forme $(0, \rho, b)$. Ceci nous permet de déduire que la fonction génératrice des moments pour S_t vérifie :

$$\forall u \leq 0, \quad E[e^{uS_t}] = e^{tl(u)} \quad \text{avec } l(u) = bu + \int_0^{+\infty} (e^{ux} - 1) \rho(dx)$$

On appelle $l(u)$ l'exposant de Laplace de S . Il intervient dans le théorème suivant qui expose le principe de la subordination

Théorème 3.2 On se donne un espace de probabilité (Ω, F, P) . Soit alors $(X_t)_{t \geq 0}$ un processus de Lévy d'exposant caractéristique $\psi(u)$, de triplet caractéristique (A, ν, γ) et soit $(S_t)_{t \geq 0}$ un subordonateur d'exposant de Laplace $l(u)$ et de triplet caractéristique $(0, \rho, b)$, alors le processus $(Y_t)_{t \geq 0}$ défini pour tout $\omega \in \Omega$ par $Y(t, \omega) = X(S(t, \omega), \omega)$ est un processus de Lévy et sa fonction caractéristique vérifie :

$$E[e^{iu \cdot Y_t}] = e^{\psi(l(u))}$$

En d'autres termes, celà revient à dire que l'exposant caractéristique de Y est obtenu par composition de l'exposant de Laplace de S avec l'exposant caractéristique de X . Le triplet

caractéristique (A^Y, γ^Y, ν^Y) est donnée par les expressions suivantes :

$$\begin{aligned} A^Y &= bA \\ \nu^Y(B) &= b\nu(B) + \int_0^{+\infty} p_s^X(B)\rho(ds) \\ \gamma^Y &= b\gamma + \int_0^{+\infty} \rho(ds) \int_{|x|\leq 1} xp_s^X(dx) \end{aligned}$$

où p_t^X représente la distribution de probabilité de X_t

3.2.4 Processus de Lévy et martingales

Nous avons vu précédemment que dans le contexte du pricing d'options les mesures de probabilité sous lesquelles les actifs financiers sont des martingales jouent un rôle fondamental puisqu'elles correspondent à des systèmes de prix cohérents vérifiant certaines contraintes (linéarité, additivité). Or puisque l'objectif de ce chapitre est précisément l'étude du pricing dans des modèles Lévy-exponentiels, il serait souhaitable de disposer de critères nous permettant de déterminer si un processus de Lévy X (ou l'exponentielle d'un processus de Lévy) est une martingale ou non. L'obtention de telles conditions est l'objectif de ce paragraphe. La proposition suivante apporte un élément de réponse à cette question.

Proposition 3.4 *Soit $(X_t)_{t \geq 0}$ un processus à valeurs dans \mathfrak{R} à accroissements indépendants. On considère les quatre assertions suivantes :*

- $\left(\frac{e^{iuX_t}}{E[e^{iuX_t}]} \right)_{t \geq 0}$ est une martingale $\forall u \in \mathfrak{R}$
- S'il existe $u \in \mathfrak{R}$, tel que $E[e^{uX_t}] < +\infty$ pour tout $t \geq 0$, alors $\left(\frac{e^{uX_t}}{E[e^{uX_t}]} \right)_{t \geq 0}$ est une martingale
- Si $E[X_t] < \infty$, $\forall t \geq 0$, alors $M_t = X_t - E[X_t]$ est une martingale
- Si $\text{Var}[X_t] < \infty$, $\forall t \geq 0$, alors $M_t^2 - E[X_t^2]$ est une martingale

Lorsque $(X_t)_{t \geq 0}$ est un processus de Lévy, pour que ces quatre assertions soient vraies, il suffit que les moments qu'impliquent chacune d'entre elles soient bien définis

Une conséquence intéressante de ce résultat est exprimée dans la proposition suivante. Cette fois le caractère martingale s'exprime à l'aide de conditions nécessaires et suffisantes sur le triplet caractéristique.

Proposition 3.5 *Soit $(X_t)_{t \geq 0}$ un processus à valeurs dans \mathfrak{R} , de triplet caractéristique (A, ν, γ) . On a les équivalences suivantes :*

– $(X_t)_{t \geq 0}$ est une martingale si et seulement si $\int_{|x| \geq 1} |x| \nu(dx) < +\infty$ et

$$\gamma + \int_{|x| \geq 1} x \nu(dx) = 0$$

– $(e^{X_t})_{t \geq 0}$ est une martingale si et seulement si $\int_{|x| \geq 1} e^x \nu(dx) < +\infty$ et

$$\frac{A}{2} \gamma + \int_{-\infty}^{+\infty} (e^x - 1 - x 1_{|x| \leq 1} \nu(dx)) = 0$$

3.3 Pricing d'options dans les modèles exp-Lévy

3.3.1 Changement de mesure pour les processus de Lévy

Le résultat suivant tente de répondre à la question suivante : quand deux processus de Lévy (X_t, P) et (X'_t, P') définis sur un même espace de probabilité filtré $(\Omega, F, (F_t)_{t \geq 0})$ sont-ils équivalents ? L'interrogation sous-jacente est bien entendu celle de l'incomplétude des modèles financiers reposant sur de tels processus.

Théorème 3.3 Soient (X_t, P) et (X'_t, P') deux processus de Lévy à valeurs dans \mathfrak{R} , de triplets caractéristiques respectifs (σ^2, ν, γ) et $(\sigma'^2, \nu', \gamma')$, alors $P|_{F_t}$ et $P'|_{F_t}$ sont équivalentes si et seulement si les conditions suivantes sont vérifiées :

– $\sigma = \sigma'$

– Les mesures de Lévy des deux processus sont équivalentes et on a

$$\int_{-\infty}^{+\infty} (e^{\phi(x)/2} - 1)^2 \nu(dx) < +\infty, \quad \text{où } \phi(x) = \log \left(\frac{d\nu'}{d\nu} \right)$$

– Lorsque $\sigma = 0$, il faut de plus que

$$\gamma' - \gamma = \int_{-1}^1 x(\nu' - \nu)(dx)$$

Quand P et P' sont équivalentes, la dérivée de Radon-Nikodym associée à ces deux mesures s'écrit

$$\frac{dP'|_{F_t}}{dP|_{F_t}} = e^{U_t}$$

avec U_t s'écrivant de la manière suivante :

$$\begin{aligned} U_t &= \eta X_t^c - \frac{\eta^2 \sigma^2 t}{2} - \gamma t \\ &+ \lim_{\epsilon \rightarrow 0} \left(\sum_{s \leq t, |\Delta X_s| > \epsilon} \phi(\Delta X_s) - t \int_{|x| > \epsilon} (e^{\phi(x)} - 1) \nu(dx) \right) \end{aligned}$$

Dans l'équation précédente (X_t^c) désigne la partie continue de $(X_t)_{t \geq 0}$ et η vérifie

$$\gamma' - \gamma - \int_{-1}^1 x(\nu' - \nu)(dx) = \sigma^2 \eta$$

Enfin il faut noter que (U_t, P) est un processus de Lévy de triplet caractéristique (a^U, ν^U, γ^U) défini par

$$\begin{aligned} a^U &= \sigma^2 \eta \\ \nu^U &= \nu \phi^{-1} \\ \gamma^U &= -\frac{1}{2} \sigma^2 \eta - \int_{-\infty}^{+\infty} (e^x - 1 - x 1_{|x| \leq 1}(\nu \phi^{-1}))(dx) \end{aligned}$$

Ce théorème est riche en enseignements. Son premier message est que dans le cadre des processus de Lévy à sauts, la classe des probabilités équivalentes à une probabilité donnée et qui préservent la structure du processus (il reste un Lévy sous la probabilité équivalente) est "relativement vaste" (conséquence : incomplétude). Il est de plus possible de modifier très librement la mesure de Lévy tout en restant dans la classe des mesures équivalentes. Par contre en l'absence de terme diffusif sous la probabilité initiale, il n'est pas possible de modifier le drift γ .

Remarque 3.3 Notons que si l'on a trouvé, dans le cadre des processus de Lévy à sauts, une classe de probabilité équivalente relativement vaste, on s'est cependant restreint dans le cadre du théorème précédent à un sous-ensemble de mesures de probabilité (les probabilités associées également à des processus de Lévy). La classe des probabilités équivalentes à P lorsque (X_t, P) est un processus de Lévy n'est donc pas réduite à celles que nous permet de construire naturellement le théorème précédent.

Remarque 3.4 En pratique, on remarque en fait que s'il est possible de changer librement la distribution des sauts de grandes tailles, qui sont au passage les plus importants dans une optique de pricing d'options, il faut prendre plus de précautions au voisinage de l'origine.

Il existe de nombreuses techniques pour construire "simplement" des probabilités équivalentes à une probabilité donnée. Au nombre de celles-ci figure la transformée de Esscher, dont nous ne donnons pas ici le principe. Nous précisons simplement qu'elle s'avère particulièrement commode pour construire une proba martingale équivalente dans le cadre des modèles Lévy-exponentiels, et qu'elle peut servir de fait à justifier le résultat suivant, qui clôt cette section et fixe le contexte des suivants.

Proposition 3.6 Soit (X, P) un processus de Lévy. Si les trajectoires du processus de Lévy ne sont ni croissantes presque sûrement, ni décroissantes presque sûrement, alors le modèle Lévy-exponentiel donné par $S_t = e^{rt + X_t}$ est sans opportunité d'arbitrages et par conséquent, il existe une probabilité martingale Q équivalente à P , telle que le prix de l'actif actualisé sous cette probabilité est une martingale.

En d'autres termes un modèle Lévy-exponentiel est sans opportunité d'arbitrage dans les cas suivants :

- La composante diffusive est non-nulle : $\sigma > 0$
- X est un processus à variation infinies (c'est le cas du processus à variance gamma)
- Les sauts de X sont positifs et négatifs
- X possède simultanément un drift négatif et des sauts positifs (ou l'inverse)

3.3.2 Pricing risque-neutre

Dans cette section, on considèrera, sauf mention contraire, que l'on est sous une probabilité risque-neutre Q , donc de pricing, et nous mènerons les calculs sans plus nous soucier des liens qu'entretient cette dernière avec une probabilité objective P . On s'intéressera uniquement aux outils mathématiques qui permettent un pricing efficace.

Pricing par transformation de Fourier

Dans cette section, on supposera que $S_0 = 1$. On fait également l'hypothèse que sous la probabilité risque neutre, S_T possède un moment d'ordre $(1 + \alpha)$, avec $\alpha > 0$:

$$\exists \alpha > 0, \quad \int_{-\infty}^{+\infty} \rho_T(x) e^{(1+\alpha)x} dx < +\infty$$

où ρ_T désigne la densité risque-neutre de X_T . On peut s'arranger pour que cette contrainte soit vérifiée pour une large classe de processus de Lévy. L'idée est d'exprimer la transformée de Fourier en le strike du prix d'un call

$$C(k) = e^{-rT} E[(e^{rT+X_T} - e^k)^+]$$

à l'aide de la fonction caractéristique $\Phi_T(v)$ de X_T et d'obtenir par transformation de Fourier inverse les prix de call pour toute une gamme de strike. Le problème est que la fonction $C(k)$ n'est pas intégrable. On a donc besoin d'introduire un nouvelle fonction à laquelle on puisse appliquer cette méthode. Carr et Madan suggèrent d'utiliser la fonction de "time-value" :

$$z_T(k) = e^{-rT} E[(e^{rT+X_T} - e^k)^+] - (1 - e^{k-rT})^+$$

On note $\zeta_T(v)$ la transformée de Fourier de cette fonction :

$$\zeta_T(v) = \int_{-\infty}^{+\infty} e^{ivk} z_T(k) dk$$

On peut exprimer cette fonction à l'aide de la fonction caractéristique de X_T . Pour celà, on commence par

$$e^{-rT} \int_{-\infty}^{+\infty} \rho_T(x) dx e^{ivk} (e^{rT+x} - e^k) (\mathbf{1}_{k \leq x+rT} - \mathbf{1}_{k \leq rT})$$

Ceci découle du caractère martingale de S_t sous Q . L'hypothèse d'intégrabilité nous autorise alors à calculer $\zeta(v)$ en permutant les intégrales et on obtient :

$$\begin{aligned}\zeta_T(v) &= e^{-rT} \int_{-\infty}^{+\infty} dk \int_{-\infty}^{+\infty} dx e^{ivk} \rho_T(x) e^{ivk} (e^{rT+x} - e^k) (\mathbf{1}_{k \leq x+rT} - \mathbf{1}_{k \leq rT}) \\ &= e^{-rT} \int_{-\infty}^{+\infty} \rho_T(x) dx \int_{-\infty}^{+\infty} e^{ivk} (e^k - e^{rT+x}) dk \\ &= \int_{-\infty}^{+\infty} \rho_T(x) dx \left\{ \frac{e^{ivrT}(1 - e^x)}{iv + 1} - \frac{e^{x+ivrT}}{iv(1 + iv)} + \frac{e^{(iv+1)x+ivrT}}{iv(iv + 1)} \right\}\end{aligned}$$

A partir de ce résultat, on démontre que

$$\zeta_T(v) = e^{ivrT} \frac{\Phi_T(v - i) - 1}{iv(1 + iv)} \quad (3.1)$$

Et ainsi, moyennant quelques remarques d'analyse, on obtient

$$z_T(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ivrT} \zeta_T(v) dv \quad (3.2)$$

L'étape suivante consiste à se ramener à un problème de transformation de Fourier rapide. Pour cela on doit discrétiser l'espace des strike, tronquer et évaluer numériquement les intégrales.

Remarque 3.5 *On peut améliorer à moindres frais la méthode décrite ci-dessus en considérant non plus la fonction "time-value" mais une nouvelle fonction introduite par P.Tankov dans [3] définie de la manière suivante :*

$$\tilde{z}_T(k) = e^{-rT} E[(e^{rT+X_T} - e^k)^+] - C_{BS}^\sigma(k)$$

où $C_{BS}^\sigma(k)$ désigne le prix Black Scholes d'un call de strike e^k et de volatilité σ . Cette volatilité est un paramètre libre, que l'on doit chercher à fixer de manière optimale en vue de la convergence. Cette nouvelle définition modifie en conséquence la fonction ζ à considérer :

$$\tilde{\zeta}_T(v) = e^{ivrT} \frac{\Phi_T(v - i) - \phi_T^\sigma(v - i)}{iv(1 + iv)} \quad (3.3)$$

où $\phi_T^\sigma(v) = e^{-\frac{\sigma^2 T}{2}} (v^2 + iv)$

Remarque 3.6 *L'avantage majeure de cette méthode de pricing est sa rapidité, particulièrement adapté au contexte de la calibration. L'analyse rigoureuse des conditions de définition pour les différentes intégrales de ce chapitre. L'estimation de l'erreur en fonction des paramètres d'approximation (troncature et évaluation numérique de l'intégrale) sont abordés en détail dans [3]. Pratiquement cette méthode fonctionne bien et pour une grande variété de modèles.*

3.4 Le modèle à Variance Gamma

Dans cette section nous traitons d'un modèle Lévy-exponentiel particulier, qui donne d'ailleurs son nom à ce chapitre : le modèle dit à Variance Gamma. L'apparition de ce modèle remonte à l'article fondateur de Madan, Carr et Chang [12], qui les premiers établirent notamment une formule fermée pour le prix du call européen. D'autres articles ont suivi depuis qui s'intéressent à ce modèle d'évaluation des actifs financiers contingents. Nous présentons ici en détail ce modèle, évoquons sa construction, son interprétation, et enfin nous lui appliquons les différentes techniques de pricing développées dans le cadre générale des modèles Lévy-exponentiels, en veillant à noter les méthodes et difficultés rencontrées dans le cadre de leurs implémentations.

3.4.1 Construction

Le processus à Variance Gamma est un processus de Lévy, obtenu par subordination d'un mouvement brownien standard de drift θ , noté dans la suite $(B_t^\theta)_{t \geq 0}$, à un processus gamma. Ce processus de Lévy, caractérisé par sa moyenne μ et sa variance ν , que nous noterons $(\gamma_t^{\mu, \nu})_{t \geq 0}$ appartient à la classe des processus de Lévy dits stables et tempérés. Nous ne revenons pas sur ces notions et renvoyons le lecteur à [1] pour un exposé clair sur ces notions.

Ce processus est défini par la distribution de ses accroissements indépendants sur les intervalles de la forme $[t, t+h]$, qui suivent tous une loi gamma. Plus exactement, si on désigne par $f_h(x)$ la densité de l'accroissement $\gamma_{t+h}^{\mu, \nu} - \gamma_t^{\mu, \nu}$, alors on a la formule suivante :

$$\forall x > 0, \quad f_h(x) = \left(\frac{\mu}{\nu}\right)^{\frac{\mu^2 h}{\nu}} \frac{x^{\frac{\mu^2 h}{\nu}} \exp(-\frac{\mu}{\nu} x)}{\Gamma(\frac{\mu^2 h}{\nu})}$$

où $\Gamma(x)$ désigne la fonction gamma usuelle. A partir de ce résultat, on peut grace à un calcul rapide obtenir une formule de la fonction caractéristique $\phi_{\gamma_t}(u) = E(e^{iu\gamma_t^{\mu, \nu}})$ du processus :

$$\phi_{\gamma_t}(u) = \left(\frac{1}{1 - iu\frac{\nu}{\mu}}\right)^{\frac{\mu^2 t}{\nu}}$$

et donc également une formule de l'exposant de Laplace défini dans la section précédente :

$$l_{\gamma_t}(u) = \left(\frac{1}{1 - u\frac{\nu}{\mu}}\right)^{\frac{\mu^2 t}{\nu}}$$

Il est également intéressant de noter que l'on dispose d'une formule de la densité de la mesure de Lévy, notée $\rho^{\mu, \nu}(x)$, associée à ce processus :

$$\rho^{\mu, \nu}(x) = \frac{\mu^2}{\nu} \frac{e^{-\frac{\mu}{\nu} x}}{x} 1_{x > 0}$$

FIG. 3.1 – Réalisations du processus à Variance Gamma pour différentes valeurs du paramètre ν (variance du processus gamma). $\nu_{bleu} = 0.2$, $\nu_{vert} = 0.05$, $\nu_{rouge} = 0.001$

On constate donc que pour sur tout intervalle $[t, t_1]$ le nombre de sauts du processus gamma est infini (la plupart de ces sauts étant bien évidemment petit sous peine de contredire l'hypothèses càdlàg). Le processus gamma est donc un processus de pure saut, et l'on peut l'approximer par un processus de Poisson composé en tronquant sa mesure de Lévy au voisinage de zero. Nous ne détaillerons pas ici cette procédure.

Formellement, on définit le processus à variance gamma $(X_t^{\sigma, \nu, \theta})_{t \geq 0}$ à partir d'un mouvement brownien avec drift $(B_t^\theta)_{t \geq 0}$ et d'un processus gamma de variance $\mu = 1$:

$$\forall t \geq 0, \quad X_t^{\sigma, \nu, \theta} = B_{\gamma_t^{1, \nu}}^\theta$$

On dispose donc de trois paramètres : σ la volatilité du mouvement brownien subordonné, θ le drift du brownien paramétrisé, et γ la variance du subordonateur. Nous verrons plus loin que l'intérêt majeur de cette paramétrisation est que les paramètres θ et σ sont liés à deux

FIG. 3.2 – Loi des incréments pour le processus à variance gamma. Valeurs des paramètres : $\sigma = 0.2$, $\theta = -0.16$, $\kappa = 0.01$, $h = \Delta t = 0.01$

grandeurs statistiques le skew et la kurtosis κ de la distribution des log-rendements.

De cette définition du processus à variance gamma en termes de subordination nous pouvons déduire plusieurs résultats. Tout d'abord, le processus $(X_t^{\sigma, \nu, \theta})_{t \geq 0}$ est bien un processus de Lévy : c'est une conséquence directe du théorème (4.2). Mais ce théorème nous dit plus, on peut également en déduire la forme de la fonction caractéristique, notée $\phi_{X_t} = E(e^{iuX_t^{\sigma, \nu, \theta}})$ du processus à partir de la formule de l'exposant de Laplace du subordonateur. Ceci se traduit par la formule suivante :

$$\phi_{X_t}(u) = \left(\frac{1}{1 - i\theta\nu u + (\sigma^2\nu)/2u^2} \right)^{\frac{t}{\nu}}$$

De cette définition découle également une première expression de la densité du processus à variance gamma. En effet, conditionnellement à une réalisation du subordonateur, $X_t^{\sigma, \nu, \theta}$ suit une loi normale de variance σ^2 . Dès lors en intégrant la densité de la loi normale contre la densité du processus gamma déjà calculée, on obtient la formule suivante pour la densité de $X_t^{\sigma, \nu, \theta}$, notée $f_{X_t}(x)$:

$$f_{X_t}(x) = \int_0^{+\infty} \frac{1}{\sigma \sqrt{(2\pi y)}} e^{-\frac{(x-\theta y)^2}{2\sigma^2 y}} y^{\frac{t}{\nu}-1} e^{-\frac{y}{\nu}} \frac{1}{\nu \Gamma(\frac{t}{\nu})} dy$$

La mesure de Lévy du processus $(X_t^{\sigma, \nu, \theta})_{t \geq 0}$, peut se formuler de plusieurs manières. Dans cette optique, il est intéressant de remarquer que ce processus peut s'exprimer comme la différence de deux processus gamma :

$$\forall t \geq 0, \quad X_t^{\sigma, \nu, \theta} = \gamma_p^{\mu_p, \nu_p}(t) - \gamma_n^{\mu_n, \nu_n}(t)$$

dont les moyennes et variances sont liées aux paramètres du processus à variance gamma par les égalités :

$$\begin{aligned} \mu_p &= \frac{1}{2} \sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \frac{\theta}{2} \\ \mu_n &= \frac{1}{2} \sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \frac{\theta}{2} \\ \nu_p &= \left(\frac{1}{2} \sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \frac{\theta}{2} \right)^2 \nu \\ \nu_n &= \left(\frac{1}{2} \sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \frac{\theta}{2} \right)^2 \nu \end{aligned}$$

Si l'on privilégie cette représentation, alors la densité de la mesure de Lévy du processus $(X_t^{\sigma, \nu, \theta})_{t \geq 0}$ peut s'écrire de la manière suivante :

$$\xi_X(x) = \frac{\mu_n^2}{\nu_n} \frac{e^{-\frac{\mu_n}{\nu_n}|x|}}{|x|} 1_{x < 0} + \frac{\mu_p^2}{\nu_p} \frac{e^{-\frac{\mu_p}{\nu_p}x}}{x} 1_{x > 0}$$

Mais on peut également l'écrire à l'aide des seuls trois paramètres du processus gamma, sans recours à la précédente décomposition :

$$\xi_X(x) = \frac{\theta x}{\nu |x|} e^{\left(-\sqrt{\frac{2}{\nu} + \frac{\theta^2}{\sigma^2}} |x| \right)}$$

Les figures (1), (2) représentent ces diverses fonctions.

FIG. 3.3 – Mesure de Lévy correspondant au processus à variance gamma. Valeurs des paramètres : $\sigma = 0.2$, $\theta = -0.16$, $\kappa = 0.01$

3.4.2 Modélisation risque neutre

On se place ici sous la probabilité risque-neutre notée Q . Sous cette probabilité, on suppose que la valeur du sous jacent S_t vérifie

$$\forall t \geq 0, \quad S_t = S_0 e^{rt + X_t^{\sigma, \nu, \theta} + \omega t} \quad (3.4)$$

Sous Q , le prix de l'actif actualisé est une martingale, ce qui implique $E(e^{-rt} S_t) = S_0$. Cette propriété explique la nécessaire adjonction dans l'univers risque-neutre du terme ω , dont on déduit la formule à partir du théorème (4.3) :

$$\omega = \frac{1}{\nu} \ln \left(1 - \theta \nu - \frac{\sigma^2 \nu}{2} \right)$$

Grace aux différents résultats déjà recensés à l'occasion de la construction du processus à variance gamma, nous pouvons énoncer la proposition suivante :

Proposition 3.7 *On suppose que $(S_t)_{t \geq 0}$ suit le modèle (4.1), alors la variable aléatoire $\ln(S_t/S_0)$ admet la densité suivante*

$$h(z) = \frac{2e^{\theta x/\sigma^2}}{\nu^{\frac{t}{\nu}} \sqrt{2\pi} \sigma \Gamma(\frac{t}{\nu})} \left(\frac{x^2}{2\sigma^2/\nu + \theta^2} \right)^{\frac{t}{2\nu} - \frac{1}{4}} K_{\frac{t}{\nu} - \frac{1}{2}} \left(\frac{1}{\sigma^2} \sqrt{x^2(2\sigma^2/\nu + \theta^2)} \right)$$

où K est une fonction de Bessel de deuxième espèce, et

$$x = z - rt - \omega t$$

Remarque 3.7 *Cette formule dont nous donnons l'expression sous la probabilité risque neutre peut s'avérer également très utile sous la probabilité historique P , notamment en vue d'une estimation statistique des paramètres de la loi, à l'aide par exemple du maximum de vraisemblance. Il faut alors substituer à r un nouveau paramètre m .*

Dans Madan et Milne, les auteurs proposent de calculer le prix de l'option par intégration numérique à partir du résultat précédent. Une amélioration significative de ce résultat est proposée dans [12]. Ces derniers proposent une formule fermée pour le prix du call. Celle-ci est l'objet de la proposition suivante :

Proposition 3.8 *On suppose que $(S_t)_{t \geq 0}$ suit le modèle (4.1), alors le prix $P(0, S_0, K, T)$ d'une option d'achat de strike K et de maturité T est donné à l'instant 0 par :*

$$\begin{aligned} P(0, S_0, K, T) &= S_0 \Psi \left(d \sqrt{\frac{1-c_1}{\nu}}, (\alpha + s) \sqrt{\frac{\nu}{1-c_1}}, \frac{t}{\nu} \right) \\ &\quad - K e^{-rt} \Psi \left(d \sqrt{\frac{1-c_2}{\nu}}, \alpha \sqrt{\frac{\nu}{1-c_2}}, \frac{t}{\nu} \right) \end{aligned} \quad (3.5)$$

avec

$$d = \frac{1}{s} \left[\ln \left(\frac{S_0}{K} \right) + rt + \frac{t}{\nu} \ln \left(\frac{1 - c_1}{1 - c_2} \right) \right]$$

où $\alpha = \zeta s$, et ζ, s, c_1, c_2 vérifiant :

$$\begin{aligned} \zeta &= -\frac{\theta}{\sigma^2} \\ s &= \frac{\sigma}{\sqrt{1 + \left(\frac{\theta}{\sigma}\right)^2 \frac{\nu}{2}}} \\ c_1 &= \frac{\nu(\alpha + s)^2}{2} \\ c_2 &= \frac{\nu\alpha^2}{2} \end{aligned}$$

Remarque 3.8 *La formule précédente est séduisante et familière : elle est proche de celle du modèle de Black-Scholes. Mais la fonction Ψ , définie à l'aide des fonctions de Bessel modifiées de seconde espèce, et des fonctions hypergéométriques dégénérées de deux variables est très difficile à évaluer numériquement. Nous ne nous étendons pas plus sur ce problème.*

3.4.3 Pricing par Monte-Carlo

Les ouvrages et articles de référence pour cette section sont [1] et [15]. Dans le modèle (4.1), on dispose donc d'une formule fermée d'implémentation difficile qui n'est de plus valide que dans les cas particuliers des options d'achat et de vente. Nous nous intéressons ici au pricing d'options européennes dans ce même modèle par la méthode de Monte-Carlo, dont l'implémentation est plus simple.

On considère une option européenne de payoff $H_T \in L^2(\Omega)$, que l'on peut mettre sous la forme $H_T = H^1(S_T)$ (que nous étudierons plus précisément ici) ou sous la forme $H_T = H^2((S_t)_{t \in [0, T]})$. Sous la probabilité risque neutre Q , le prix d'un tel actif à l'instant $t < T$ est donné par la formule :

$$P_t = E^Q \left[e^{-r(T-t)} H_T \right]$$

On estime cette quantité par la méthode de Monte-Carlo, on obtient un prix \hat{P}_t approximé :

$$\hat{P}_t = e^{-r(T-t)} \frac{1}{M} \sum_{i=1}^M H_T(\omega^i)$$

où M représente le nombre de simulation indépendante du processus réalisé.

Il nous suffit donc de préciser la manière dont on simule une trajectoire approximée sur une grille temporelle discrète $0 = t_0 < t_1 < \dots < t_N = T$. On va bien évidemment utiliser le concept de subordination. Les étapes sont les suivantes :

- on commence par construire un échantillon discret de $N+1$ variables aléatoires $\{h_n\}_{n=0\dots N}$ vérifiant :

$$h_0 = 0 \text{ et } \Delta h_n = h_{n+1} - h_n \sim G\left(\frac{\Delta t_n}{\nu}, \nu\right)$$

où $G(a, b)$ représente la loi de la distribution gamma de paramètres a et b (à ne pas confondre avec la moyenne et la variance). Elles représentent le nouveau référentiel discret en temps. On simule en fait sur une grille temporelle discrète les valeurs prises par le subordonateur (processus gamma)

- Conditionnellement à cette réalisation on calcule les valeurs $\{\omega_i\}_{0 \leq i \leq N}$ de manière itérative en remarquant que

$$\Delta \omega_n = \omega_{n+1} - \omega_n \sim N(\theta \Delta h_n, \sigma^2 \Delta h_n)$$

- La trajectoire $\{\omega_i\}_{0 \leq i \leq N}$ est une approximation discrète d'une trajectoire (continue) du processus à variance gamma.

On calcule finalement les prix des options en moyennant les valeurs du payoff pour M réalisations approximées, construites en utilisant la méthode précédente.

Remarque 3.9 *La trajectoire discrète n'est pas simulée suivant la loi exacte du processus à variance gamma. En conséquence, il faut théoriquement, pour cette méthode de simulation, conserver à l'esprit qu'à l'erreur statistique classique et inhérente à la méthode de Monte-Carlo s'ajoute une erreur dite de discrétisation*

Remarque 3.10 *Pour l'évaluation des payoffs "path-dependant", il est recommandé d'avoir recours comme dans le cas des diffusions à des méthodes de ponts dont la mise en oeuvre est abordée dans Ribeiro.*

3.4.4 Pricing par EID

Comme nous l'avons remarqué plus haut dans un cadre général, le prix d'une option européenne dans le modèle à Variance Gamma s'exprime comme la solution d'une équation intégral-différentielle aux dérivées partielles. Nous avons déjà évoqué cette classe d'équation fonctionnelle dans le cours de notre développement. Nous appliquons donc au processus VG les principes généraux déjà développés, et détaillons ici l'implémentation d'une méthode de résolution numérique.

Dans le modèle à variance gamma, le prix d'un put vérifie l'EID, suivante

$$\frac{\partial P}{\partial t} + (r + \omega)S \frac{\partial P}{\partial S} + \int_{-\infty}^{+\infty} \{P(t, Se^y) - P(t, S)\} k(dy) - rP = 0 \quad (3.6)$$

avec les conditions aux bords suivantes :

$$\begin{aligned}\forall t \in]0, T], \quad P(t, 0) &= Ke^{-r(T-t)} \\ \forall t \in]0, T], \quad P(t, +\infty) &= 0 \\ \forall S \in [0, \infty[, \quad P(T, S) &= (K - S)^+\end{aligned}$$

Bien sur comme dans le cas du modèle à volatilité locale, on doit restreindre le domaine de résolution, afin de mettre en place une méthode de résolution numérique. On se placera à présent sur le domaine $[0, T] \times [S_{min}, S_{max}]$, qui servira de base à la construction de la grille de discrétisation dans la suite. On effectue en suite le changement de variable suivant ;

$$\begin{aligned}x &= \ln(S) \\ U(t, x) &= P(t, S)\end{aligned}$$

On en déduit immédiatement :

$$U(t, x + y) = P(t, Se^y) \quad \frac{\partial U}{\partial x} = S \frac{\partial P}{\partial S}$$

On peut donc réécrire l'EID précédente en coordonnées logarithmiques :

$$\frac{\partial U}{\partial t} + (r + \omega) \frac{\partial U}{\partial x} + \int_{-\infty}^{+\infty} \{U(t, x + y) - U(t, x)\} k(dy) - rU = 0 \quad (3.7)$$

avec les conditions aux bords suivantes :

$$\begin{aligned}\forall t \in]0, T], \quad U(t, -\infty) &= Ke^{-r(T-t)} \\ \forall t \in]0, T], \quad U(t, +\infty) &= 0 \\ \forall x \in]-\infty, \infty[, \quad U(T, x) &= (K - e^x)^+\end{aligned}$$

Afin de résoudre numériquement, on est contraint de restreindre cette équation au domaine borné $[0, T] \times [\ln S_{min}, \ln S_{max}]$. Une difficulté apparaît alors qui consiste à donner un sens aux intégrales intervenant dans l'équation et définie sur \mathfrak{R} tout entier. Nous allons voir que ce problème peut être solutionné dans le cadre de la discrétisation. Il conduit cependant à l'introduction d'une nouvelle source d'erreur dite de troncature. On introduit tout d'abord quelques notations.

Le domaine $[0, T] \times [\ln S_{min}, \ln S_{max}]$ est discrétisé à l'aide d'une grille de $(M + 1)(N + 1)$ points, et on pose :

$$\begin{aligned}\Delta x &= \frac{x_{max} - x_{min}}{N} \\ \Delta t &= \frac{T}{M} \\ x_i &= x_{min} + i\Delta x \\ t_j &= t_{min} + j\Delta t\end{aligned}$$

On note de plus $(U_{i,j})$ le vecteur de $\mathfrak{R}^{(N+1) \times (M+1)}$ des valeurs de la solution au problème discrétisé aux points de la grille de discrétisation. Pour l'approximation des différentes dérivées on choisit :

$$\begin{aligned}\frac{\partial U}{\partial t}(t_j, x_i) &\approx \frac{U_{i,j+1} - U_{i,j}}{\Delta t} \\ \frac{\partial U}{\partial x}(t_j, x_i) &\approx \frac{U_{i+1,j} - U_{i-1,j}}{2\Delta x}\end{aligned}$$

Enfin, afin d'alléger les notations, on définit $a = r + \omega$. On peut alors définir un schéma de discrétisation pour l'EID en posant pour tout $i \in 1, \dots, N$ et $j \in 0, \dots, M$

$$\frac{U_{i,j+1} - U_{i,j}}{\Delta t} + a \frac{U_{i+1,j} - U_{i-1,j}}{2\Delta x} + \int_{-\infty}^{+\infty} \{U(t_j, x_i + y) - U_{i,j}\} k(dy) = rU_{i,j}$$

On définit une nouvelle constante :

$$h = \frac{a\Delta t}{2\Delta x}$$

Ce qui nous permet de reformuler le schéma numérique précédent :

$$hU_{i-1,j} + (1 + r\Delta t)U_{i,j} - hU_{i+1,j} = U_{i,j+1} + \int_{-\infty}^{+\infty} \{U(t_j, x_i + y) - U_{i,j}\} k(dy)$$

Afin d'obtenir un résultat implémentable, nous avons besoin d'une approximation du terme intégral présent dans l'égalité ci-dessus. Un bon moyen d'obtenir une telle approximation consiste dans un premier temps à découper cette intégrale de la manière suivante :

$$\begin{aligned}\int_{-\infty}^{+\infty} \{U(t_j, x_i + y) - U_{i,j}\} k(dy) &= \int_{-\infty}^{x_0 - x_i} \{U(t_j, x_i + y) - U_{i,j}\} k(dy) \\ &+ \int_{x_0 - x_i}^{-\Delta x} \{U(t_j, x_i + y) - U_{i,j}\} k(dy) \\ &+ \int_{-\Delta x}^0 \{U(t_j, x_i + y) - U_{i,j}\} k(dy) \\ &+ \int_0^{\Delta x} \{U(t_j, x_i + y) - U_{i,j}\} k(dy) \\ &+ \int_{\Delta x}^{(N-i)\Delta x} \{U(t_j, x_i + y) - U_{i,j}\} k(dy) \\ &+ \int_{(N-i)\Delta x}^{+\infty} \{U(t_j, x_i + y) - U_{i,j}\} k(dy)\end{aligned}$$

On approxime ensuite chacun de ces termes de manière indépendante. Pour simplifier les calculs, on introduit deux constantes λ_p et λ_n dont l'usage allège l'expression de la densité de Lévy $k(dx)$. Examinons tout d'abord le terme

$$\int_0^{\Delta x} \{U(t_j, x_i + y) - U_{i,j}\} k(y) dy = \int_0^{\Delta x} \{U(t_j, x_i + y) - U_{i,j}\} \frac{e^{-\lambda_p y}}{\nu y} dy$$

en remarquant que pour Δx petit on a :

$$U(t_j, x_i + y) - U_{i,j} \approx \frac{U(x_{i+1}, t_j) - U_{i,j}}{\Delta x} y$$

on en déduit

$$\begin{aligned} \int_0^{\Delta x} \{U(t_j, x_i + y) - U_{i,j}\} k(y) dy &= \int_0^{\Delta x} \frac{U(x_{i+1}, t_j) - U_{i,j}}{\Delta x} y \frac{e^{-\lambda_p y}}{\nu y} k(y) dy \\ &= \frac{[U(x_{i+1}, t_j) - U_{i,j+1}](1 - e^{-\lambda_p \Delta x})}{\nu \Delta x \lambda_p} \end{aligned}$$

Par un raisonnement analogue, on a immédiatement :

$$\int_{-\Delta x}^0 \{U(t_j, x_i + y) - U_{i,j}\} k(dy) = \frac{[U(x_{i-1}, t_j) - U_{i,j+1}](1 - e^{-\lambda_n \Delta x})}{\nu \Delta x \lambda_n}$$

Le quatrième terme de la décomposition est plus délicat à traiter. On commence par remarquer que :

$$\int_{\Delta x}^{(N-i)\Delta x} \{U(t_j, x_i + y) - U_{i,j}\} k(dy) = \int_{\Delta x}^{(N-i)\Delta x} \{U(t_j, x_i + y) - U_{i,j}\} \frac{e^{-\lambda_p y}}{\nu y} k(y) dy$$

on remarque ensuite astucieusement que cette intégrale peut à nouveau être scindée

$$\sum_{k=1}^{N-i-1} \int_{k\Delta x}^{(k+1)\Delta x} \{U(t_j, x_i + y) - U_{i,j}\} \frac{e^{-\lambda_p y}}{\nu y} k(y) dy$$

et que pour Δx petit on peut réaliser l'approximation :

$$U(t_{j+1}, x_i + y) - U(t_{j+1}, x_{i+k}) \approx \frac{U(x_{i+k+1}, t_{j+1}) - U(t_{j+1}, x_{i+k})}{\Delta x} (y - k\Delta x)$$

En réinjectant ce dernier résultat dans la somme on dispose d'une nouvelle expression :

$$\begin{aligned} \sum_{k=1}^{N-i-1} \int_{k\Delta x}^{(k+1)\Delta x} \left\{ \frac{U(x_{i+k+1}, t_{j+1}) - U(t_{j+1}, x_{i+k})}{\Delta x} (y - k\Delta x) \right. \\ \left. + U(t_{j+1}, x_{i+k}) - U_{i,j+1} \frac{e^{-\lambda_p y}}{\nu y} k(y) dy \right\} \end{aligned}$$

Et on obtient finalement l'expression suivante :

$$\begin{aligned} \sum_{k=1}^{N-i-1} \frac{1}{\nu \Delta x} [U(x_{i+k+1}, t_{j+1}) - U_{i+k,j+1}] \frac{1}{\lambda_p} \left\{ e^{-\lambda_p k \Delta x} - e^{-\lambda_p (k+1) \Delta x} \right\} \\ + \sum_{k=1}^{N-i-1} \frac{1}{\nu} [U(x_{i+k}, t_{j+1}) - U_{i,j+1} - k(U(x_{i+k+1}, t_{j+1}) - U_{i+k,j+1})] \\ \times \{ \text{expint}(\lambda_p k \Delta x) - \text{expint}(\lambda_p (k+1) \Delta x) \} \end{aligned}$$

De manière symétrique pour l'évaluation du terme :

$$\int_{x_0-x_i}^{-\Delta x} \{U(t_j, x_i + y) - U_{i,j}\} k(dy)$$

on obtient le résultat suivant :

$$\begin{aligned} & \sum_{k=1}^i \frac{1}{\nu \Delta x} [U(x_{i-k-1}, t_{j+1}) - U_{i-k,j+1}] \frac{1}{\lambda_n} \left\{ e^{-\lambda_n k \Delta x} - e^{-\lambda_n (k+1) \Delta x} \right\} \\ & + \sum_{k=1}^i \frac{1}{\nu} [U(x_{i-k}, t_{j+1}) - U_{i,j+1} - k(U(x_{i-k-1}, t_{j+1}) - U_{i-k,j+1})] \\ & \quad \times \{ \expint(\lambda_n k \Delta x) - \expint(\lambda_n (k+1) \Delta x) \} \end{aligned}$$

Le calcul des termes restants dépend des conditions limites.

Remarque 3.11 *Cette méthode permet de retrouver des prix d'options similaires à ceux obtenus à l'aide des méthodes précédemment évoquées. Nous ne donnons pas ici de preuve de convergence pour ce schéma, car une telle preuve souvent éludée dans la littérature requiert des hypothèses et outils compliqués, dont l'énoncé n'a pas sa place ici. Mais une telle preuve existe dans le cas particulier de payoffs bornés (put par exemple), voir par exemple la thèse d'Ekatarina Votchkova [10].*

Conclusion

Ce rapport s'achève ici. J'espère y avoir exposé avec clarté les principales problématiques et méthodes traitées au cours de mon stage au sein d'Arbitragis. L'étude des deux modèles présentés, et le développement de pricer dans chacun de ces cadres constituèrent pour moi une expérience très enrichissante. J'aimerais en guise de conclusion évoquer quelques pistes de travail connexes que je n'ai malheureusement pas eu le temps d'approfondir lors de mon passage au sein d'Arbitragis.

Dans le cadre des modèles Lévy-exponentiels notamment, je pense qu'une extension des différentes méthodes de pricing, présentées ci-dessus dans le seul cadre du modèle VG (à l'exception du pricing par transformée de Fourier), à une vaste gamme de modèles est intéressante, possible et "peu coûteuse". Intéressante car le modèle à variance gamma, s'il offre certes un potentiel descriptif supérieur à celui du modèle de Black-Scholes, n'est qu'un modèle à sauts parmi d'autres. Certains modèles Lévy-exp alternatifs (loi hypergéométrique) apparaissent, à la lumière de l'analyse statistique, plus proches du comportement historique. Possible et "peu coûteuse", car les grandes lignes en sont tracées dans la littérature et sont dans l'esprit des méthodes déjà traitées ici.

Pour les méthodes Monte-Carlo, on pourra par exemple consulter dans [1] le chapitre traitant de la simulation des processus de Lévy, qui constitue ici le point crucial (payoffs européens). Pour les méthodes reposant sur les EDP des prix, la généralisation semble plus délicate car la méthode de différences finies ici présentée repose sur une troncature de l'intégrale propre au problème traité. Cependant, on trouve dans [10] une étude complète de ces questions. L'auteur insiste sur les problèmes liés à l'implémentation et elle traite des modèles Lévy-exp en toute généralité.

Le second point que j'aurais désiré approfondir bien que ce rapport ne fasse qu'effleurer la question est la calibration des modèles Lévy-exponentiels dans un cadre général. En effet, lors de ce stage et afin de tester certains pricers j'ai utilisé des procédés de calibration ad hoc qui me permettaient de retrouver les paramètres d'un processus Lévy-exp (basé bien entendu sur le modèle VG) à partir des skew et kurtosis exhibés par la courbe des prix. Cette méthode possède deux inconvénients majeurs : elle n'est pas directe (on cherche en fait skew et kurtosis à partir d'un développement limité des prix), et elle postule une forme a priori et paramétrisée pour la loi du processus (on restreint trop fortement l'ensemble des distributions susceptibles d'avoir produit les prix).

Dans [3], Tankov propose une méthode bien plus générale de calibration pour les modèles Lévy-exponentiels. Celle-ci est cependant relativement complexe et son implémentation difficile, bien qu'elle soit évoquée en longueur. Sans entrer dans les détails, elle repose sur une méthode de valorisation rapide des options européennes dans les modèles Lévy-exp par transformée de Fourier (très proche des techniques déjà évoquées et implémentées), une formulation précise et robuste du problème de calibration faisant intervenir la notion d'entropie par rapport à une mesure de référence et l'utilisation d'un algorithme de minimisation efficace. La technique

proposée possède en outre de bonnes propriétés dont la continuité du résultat par rapport aux données. La seule indétermination réside dans le choix d'une probabilité de référence : probabilité historique si on la connaît ou hypothèse sur la forme du processus.

L'élargissement des travaux sur le processus à variance gamma dans ces deux directions : généralisation des routines de valorisation et des méthodes de calibration offre donc des perspectives très intéressantes. Elle permettrait de saisir de manière fine l'intégration par les marchés du caractère non gaussien des queues de distributions notamment par le biais de la connaissance de la mesure de probabilité qui sert de fondement à l'évaluation des options. Couplé à une analyse statistique précise des cours, qui permette d'identifier les caractéristiques d'un titre sous la probabilité historique, ce savoir pourrait sans doute permettre d'exploiter des "différences de vue" se manifestant sous forme de différences de prix, entre une probabilité risque-neutre dérivant d'un modèle historique et d'une méthode de valorisation (+couverture), et une probabilité risque-neutre constatée via cette même méthode de calibration...

Bibliographie

- [1] RAMA CONT & PETER TANKOV *Financial Modelling With Jump Processes*, Chapman & Hall/CRC, 2003 .
- [2] RAMA CONT, *Empirical properties of asset returns : Stylized facts and statistical issues*, Quant. Finance, I (2001) pp. 1-14.
- [3] PETER TANKOV, *Processus de Lévy en Finance : Problèmes Inverses et Modélisation de Dépendance*, Thèse de l'école polytechnique (Septembre 2004).
- [4] PETER TANKOV, *Calibration de modèles et couvertures de produits dérivés* , Cours de DEA, université Paris VII (2003).
- [5] FREDDY DELBAEN, WALTER SCHACHERMAYER, *A general version of the fundamental theorem of asset pricing*, Mathematische Annalen.
- [6] BRUNO DUPIRE, *Pricing and Hedging with Smiles*, Proceedings of AFFI Conference, La Baule, June 1993
- [7] S. CRÉPEY, *Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization*, SIAM J.Math. Anal. Vol 34. No. 5.
- [8] HERBERT EGGER & HEINZ W.ENGL, *Tikhonov Regularization Applied to the Inverse Problem of Option Pricing : Convergence Analysis and rates*,
- [9] MARCO AVELLANEDA, CRAIG FRIEDMAN, RICHARD HOLMES & DOMINICK SAMPERI, *Calibrating Volatility Surfaces via Relatives Entropy Minimization*,
- [10] EKATERINA VOLTCHKOVA, *Equations intégrro-différentielles d'évolution : méthodes numériques et applications en finance*, Octobre 2005.
- [11] DILIP B. MADAN & PETER CARR, *Option valuation using the fast Fourier Transform*, J. Comput. Finance, (2) (1998) pp.61-73
- [12] CARL DE BOOR, *Bicubic Spline Interpolation*, J. Math. Phy, (41) (1963) pp.212-218
- [13] THOMAS F. COLEMAN, YUYING LEE & ARUN VERMA, *Reconstructing the unknown local volatility function*, J. Comput. Finance, (2) pp.77-102
- [14] J. F. BONNANS, J-M. COGNET & S. VOLLE, *Estimation de la volatilité locale d'actifs financiers par une méthode d'inversion numérique*, Rapport de recherche INRIA, (Novembre 2002)
- [15] FILIPPO FIORANI, *The Variance Gamma Process for Option Pricing*, (septembre 2001)
- [16] C. RIBEIRO & NICK WEBBER, *Valuing Path Dependant Options in the Variance Gamma Process by Monte-Carlo*, (2000)