Stochastic expansions and closed pricing formulas of European options

Mohammed Miri

INSTITUT POLYTECHNIQUE DE GRENOBLE

N^o attribué par la bibliothèque
|_|_|_|_|_|_|_|_|_|_|

THESE

pour obtenir le grade de

DOCTEUR DE L'institut polytechnique de Grenoble

***Spécialité*:** "Mathématiques appliquées"

préparée au laboratoire Jean Kuntzmann

dans le cade de **l'Ecole Doctorale:** "Mathématiques, Sciences et Technologies de l'Information, Informatique"

Présentée et soutenue publiquement

par

MOHAMMED MIRI

le 17 Décembre 2009

DEVELOPPEMENT STOCHASTIQUE ET FORMULES FERMEES DE PRIX POUR LES OPTIONS EUROPEENNES

EMMANUEL GOBET
ERIC BENHAMOU

JURY

Mme.	Nicole EL KAROUI	, Président
M.	Jean-Pierre FOUQUE	, Rapporteur
M.	Denis TALAY	, Rapporteur
M.	Emmanuel GOBET	, Directeur de thèse
M.	Eric BENHAMOU	, Co-Directeur en entreprise
M.	Philippe BRIAND	, Examinateur
M.	Etienne KOEHLER	, Examinateur

Remerciements

Je voudrais remercier Emmanuel car il a été un directeur de thèse exemplaire. J'ai beaucoup apprécié ses qualités humaines et professionnelles. Je le remercie donc pour la qualité et la rigueur de son suivi. Je voudrais remercier Eric pour son suivi efficace et régulier dans le cadre mes travaux ainsi que pour sa disponibilité. Je remercie Emmanuel et Eric pour avoir pu me consacrer du temps et avoir rendu ce contrat CIFRE réalisable malgré la distance qui sépare le laboratoire (Grenoble) et l'entreprise (Paris).

Cette thèse est le résultat d'une collaboration active entre le laboratoire Jean Kuntzmann et l'entreprise Pricing Partners. Je tiens donc à remercier tous les membres du laboratoire et l'ensemble de mes collègues à Pricing Partners pour leur soutien.

Pour finir, je tiens à remercier mes parents et ma famille qui m'ont soutenu et aidé à arriver jusque là. Je remercie tous mes amis pour leurs encouragements réguliers. Enfin, je voudrais dédier ma thèse à mes deux parents en signe de ma profonde reconnaissance.

Contents

List of Figures

Chapter 1

French introduction

Cette thèse développe une nouvelle méthodologie permettant d'établir des approximations analytiques pour les prix des options européennes. Notre approche combine astucieusement des expansions stochastiques et le calcul de Malliavin afin d'obtenir des formules explicites et des évaluations d'erreur précises. L'intérêt de ces formules réside dans leur temps de calcul qui est aussi rapide que celui de la formule de Black et Scholes. Notre motivation vient du besoin croissant de calculs et de procédures de calibration en temps réel, tout en contrôlant les erreurs numériques reliées aux paramètres du modèle. Il existe dans la littérature plusieurs manières d'établir des formules fermées soit par un calcul direct du problème d'évaluation d'option soit en utilisant des approximations. Pour des modèles ayant une densité explicite du sous-jacent comme le modèle de Black et Scholes, le modèle CEV ou tout autre modèle dont la fonction caractéristique de la distribution du sous-jacent est explicite (comme le modèle de Heston ou les modèles affines), on peut facilement trouver des formules analytiques (parfois aussi appelées formules fermées par anglicisme). Ceci est à l'origine de résultats bien connus pour les calls et les puts. En revanche, dans le cas de modèles n'ayant pas de densité explicite, on doit se tourner vers des méthodes numériques (techniques d'EDP, simulations de Monte Carlo, ...). Par contre, pour obtenir des formules explicites, on doit établir des approximations analytiques en utilisant des méthodes de perturbation ou l'analyse asymptotique. Les méthodes de perturbation sont très générales. Pour les appliquer, on utilise habituellement un modèle simplifié (dit modèle proxy), pour lequel les calculs sont plus simples et le plus souvent explicites. Dans la section 1.1, nous énumérons les modèles pour lesquels les prix des call-put sont explicites: ces modèles peuvent être utilisés par la suite comme des modèles proxy. Dans la section 1.2, nous présentons des approximations analytiques pour des équations différentielles ordinaires et stochastiques générales et leur domaine de validité. Le domaine de la validité spécifie les restrictions de la méthode d'approximation. Dans la section 1.3, nous exposons brièvement les approximations analytiques utilisées en finance. La vue d'ensemble des méthodes précédentes nous donne une idée claire de leurs limitations qui sont le point de départ de notre travail. Plus spécifiquement, si nous déployons une perturbation basée sur les propriétés ergodiques, l'approximation est seulement valide pour les longues maturités; si nous déployons une perturbation basée sur les propriétés géodésiques, l'approximation est limitée aux maturités courtes ; de la même manière, si nous déployons une perturbation utilisant l'opérateur d'EDP, l'approximation peut être explicitée seulement pour des coefficients homogènes du temps. Ces restrictions nous incitent à penser à une nouvelle méthodologie qui peut être appliquée dans un cadre plus large :maturité courte ou longue, paramètres non homogènes en temps, ... Cette nouvelle méthodologie présentée dans la section 1.4 constitue le sujet de ce travail de thèse. On présente aussi une comparaison détaillée de notre approche avec celle de Watanabe dans 1.5. La structure et les principaux résultats sont énoncés dans la section 1.6.

1.1 Formules analytiques en finance

Le modèle de Black et Scholes. Le modèle de Black et Scholes (voir [22]) suppose que le sous-jacent suit une loi lognormale avec une volatilité constante. Autrement dit, le sous-jacent (S_t) suit la diffusion suivante:

$$\frac{dS_t}{S_t} = (r-q)dt + \sigma dW_t,$$

où W est un mouvement brownien, σ est une volatilité constante du sous-jacent, r (resp. q) est un taux risque-neutre déterministe (resp. une taux de dividende continu déterministe). Cette dynamique est écrite sous la mesure risque-neutre utilisée pour l'évaluation d'options.
Le prix des options call et put dans ce modèle a une formule fermée due au calcul explicite de la fonction

de répartition de la loi normale. Le prix du call avec comme temps initial t, maturité T, cours S, et prix d'exercice K a une expression bien connue:

$$Call_{BS}(t,S;K,T) = Se^{-q(T-t)}\mathcal{N}(d_1(T-t,Se^{(r-q)(T-t)},K)) - Ke^{-r(T-t)}\mathcal{N}(d_0(T-t,Se^{(r-q)(T-t)},K)).$$

où $d_0(t,x,y) = \frac{1}{\sigma\sqrt{t}}\log(\frac{x}{y}) - \frac{\sigma\sqrt{t}}{2}$, $d_1(t,x,y) = \frac{1}{\sigma\sqrt{t}}\log(\frac{x}{y}) + \frac{\sigma\sqrt{t}}{2}$ et $\mathcal{N}$ est la fonction de répartition de la loi normale. Une expression similaire est obtenue lorsque les paramètres r, q and σ deviennent dépendants du temps (car le cours S reste une variable lognormle).

Remarque 1.1.1. *Quand σ dépend seulement du sous-jacent, il n'existe seulement que quelques formules fermées (voir [1]). De plus, si la volatilité σ est une fonction séparable du sous-jacent et du temps, on peut établir une formule approchée pour le prix des options vanilles (call, put) en utilisant des techniques de perturbation singulière comme expliqué par Hagan et al dans [62]. Mais, dans le cas de formes générales de fonctions de volatilités dépendants à la fois du temps et du sous-jacent, il n'y a plus de formule analytique pour les options call. Ces formes générales de volatilité sont incluses dans les modèles à volatilités locales ou modèles à la Dupire (voir [40]).*

Le modèle de Merton. Le modèle de Merton (voir [85]) peut être vu comme une extension du modèle de Black et Scholes en ajoutant des sauts de Poisson indépendants avec une amplitude de saut normalement distribuée:

$$\frac{dS_t}{S_t} = (r-q-\lambda(1-e^{\eta_J+\frac{\gamma_J^2}{2}}))dt + \sigma dW_t + (e^{J_t}-1)dN_t,$$

où

- le processus de Poisson composé (J_t) et le mouvement brownien (W_t) sont indépendant,
- $J_t = \sum_{i=1}^{N_t} Y_i$ où Y_i sont i.i.d. variables normales avec moyenne η_J et volatilité γ_J,
- N_t est un processus de poisson avec intensité λ.

En conditionnant par le nombre de saut N_T, on peut exprimer le prix du call dans le modèle de Merton comme une somme infinie de prix de type Black et Scholes:

$$Call_{Merton}(t,S;K,T) = \sum_{i=0}^{\infty} \frac{(\lambda(T-t))^i}{i!} e^{-(\lambda+r)T} BSCall\left(F_T e^{i(\eta_J+\frac{\gamma_J^2}{2})}, K, T-t, \sqrt{\sigma^2 + \frac{i\gamma_J^2}{T-t}}\right),$$

où

$$F_T = Se^{(r-q+\lambda(1-\exp(\eta_J+\gamma_J^2/2)))(T-t)},$$

et $BSCall(S,K,T,v)$ est le prix Black-Scholes pour un call ayant le sous-jacent S_t avec la condition initiale $S_0 = S$, la volatilité v, exercé à la maturité T et au strike K, où le taux risque-neutre et le taux de dividende sont nuls.

Remarque 1.1.2. *Le prix du call dans le modèle de Merton a encore une formule fermée quand les paramètres r, q and σ deviennent dépendants du temps. En outre, quand la volatilité σ devient une fonction à la fois du sous-jacent et du temps, on retrouve exactement la définition du modèle d'Andersen et Andreasen ([8]). Dans un tel modèle, il y a des méthodes numériques efficaces comme la méthode Forward PIDE utilisée pour le calcul des prix des calls (voir [8] et [33])). Mais, il n'y a plus de formule analytique des prix de call dans le modèle d'Andersen et Andreasen.*

Le modèle CEV. Dans le cas du modèle "Constant Elasticity of Variance" (connu comme le modèle CEV), les options call (put) ont des formules fermées. Dans ce cas, le sous-jacent (S_t) suit la dynamique suivante:

$$\frac{dS_t}{S_t} = (r-q)dt + \nu S_t^{\beta-1} dW_t, S_0 > 0.$$

Le modèle CEV a été étudié au début par Cox dans [34] pour le cas $\beta < 1$. Le cas $\beta > 1$ a été traité après par Emanuel et MacBeth dans [42]. Le prix du call dans ce modèle peut être calculé à l'aide de la fonction complémentaire non centrée khi-carré Q:

$$Call_{CEV}(t,S;K,T) = e^{-q(T-t)}Q(2x,n,2y) - e^{-r(T-t)}Q(2y,n-2,2x) \tag{1.1}$$

où

$$\begin{aligned} n &= 2 + \frac{1}{1-\beta}, \\ x &= \frac{(r-q)S^{-2(\beta-1)}}{\nu^2(\beta-1)(e^{2(r-q)(\beta-1)(T-t)}-1)}, \\ y &= \frac{(r-q)K^{-2(\beta-1)}}{\nu^2(\beta-1)(1-e^{-2(r-q)(\beta-1)(T-t)})}. \end{aligned}$$

Le calcul de la distribution non centrée Chi-carré Q peut être réalisé soit en utilisant un algorithme récursif (voir l'algorithme de Schroder dans [104]) soit en utilisant une intégration de fonctions de type Bessel.

Remarque 1.1.3. *Quand les paramètres r, q and ν sont dépendants du temps, le prix du call a encore une formule fermée en utilisant des techniques d'algèbre de Lie (voir [66]). Mais, quand β devient dépendent du temps, il n'y a plus de formule analytique à notre connaissance.*

Le modèle de Heston. Le modèle de Heston est une extension du modèle de Black et Scholes pour le sous-jacent (S_t) mais avec une volatilité stochastique:

$$\mathrm{d}X_t = \sqrt{v_t} dW_t - \frac{v_t}{2}\mathrm{d}t,\ X_0 = x_0, \tag{1.2}$$

$$\mathrm{d}v_t = \kappa(\theta - v_t)\mathrm{d}t + \xi\sqrt{v_t}\mathrm{d}B_t,\ v_0 > 0, \tag{1.3}$$

$$\mathrm{d}\langle W,B\rangle_t = \rho \mathrm{d}t,$$

où

- X_t est le logarithme du forward $e^{(q-r)t}S_t$, r and q sont respectivement le taux risque neutre et le taux de dividendes.
- v_0 est la valeur initiale de la volatilité,
- κ est le paramètre de retour à la moyenne,
- θ est le niveau long-terme,
- ξ est la volatilité de la volatilité,

- ρ est la corrélation.

Le calcul du prix du call-put dans le modèle de Heston peut être réalisé à l'aide de l'inversion de Fourier; ceci est du au fait que la fonction caractéristique du logarithme du sous-jacent est explicite dans ce cadre (les paramètres du modèles ne dépendent pas du temps). Le prix du call dans le modèle de Heston peut être obtenu grâce à la formule de Lewis [1] ([79]):

$$Call_{Heston}(t,S_t,v_t;T,K) = S_t e^{-q(T-t)} - \frac{Ke^{-(T-t)r}}{2\pi} \int_{\frac{i}{2}-\infty}^{\frac{i}{2}+\infty} e^{-izX} \phi_T(-z) \frac{dz}{z^2 - iz}$$

où $X = \log\left(\frac{S_t e^{-(T-t)q}}{Ke^{-(T-t)r}}\right)$ et $\phi_T(z) = \mathbb{E}(e^{z(X_T - X_t)} | \mathscr{F}_t)$.
La fonction caractéristique $\phi_T(z) = \mathbb{E}(e^{z(X_T - X_t)} | \mathscr{F}_t)$ est explicite quand les paramètres sont constants. Quand les paramètres θ, ξ et ρ deviennent constants par morceaux, la fonction caractéristique $\phi_T(z) = \mathbb{E}(e^{z(X_T - X_t)} | \mathscr{F}_t)$ peut être calculée récursivement en utilisant les méthodes d'EDP (voir [86]) ou un argument de Markov pour les modèles affines (voir [41]).

Remarque 1.1.4. *Il n'existe plus de formule fermée quand on a une dépendance en temps générale des paramètres du modèle de Heston. Le temps de calcul de la formule d'inversion de Fourier est loin d'être aussi rapide que celui de la formule de Black et Scholes ou Merton. En effet, l'inversion de Fourier est très coûteuse en temps et souffre d'instabilité pour des strikes grands et des maturités longues (voir [68]).*

1.2 Introduction générale sur les approximations analytiques

On a vu précédemment qu'on peut obtenir des formules fermées de l'option call ou put lorsque la densité du sous-jacent est explicite ou sa fonction caractéristique est explicite. En dehors de ces deux cas, il n'existe pas de formule fermée pour les options vanilles. Donc, on pourrait utiliser des méthodes de perturbation. L'objectif de cette section est de donner une introduction courte et générale sur ces méthodes de perturbation.

1.2.1 Les équations différentielles ordinaires

Dans cette sous-section, on introduit brièvement les dites "méthodes de perturbations" utilisées dans la littérature spécialement pour le problème de perturbations singulières.

Le développement asymptotique par recollement. Le principe de cette méthode consiste à diviser le domaine de la solution en une séquence de deux ou trois sous-intervalles. On distingue souvent deux types de solutions: la solution intérieure et la solution extérieure. Ces solutions sont nommés en raison de leurs relations à la couche limite; la couche limite est située souvent dans les bords du problème et est la source de termes de correction non négligeables de la perturbation. Dans chaque intervalle, la théorie de perturbation est appliquée afin d'obtenir une solution asymptotique valide dans cet intervalle. Le recollement est demandé afin de combiner les solutions intérieures et extérieures de telle façon que l'approximation a la même forme fonctionnelle dans chacun de ces intervalles. Finalement, on obtient

[1] On verra les détails de l'autre formule de Heston ([63]) dans la partie III sous-section 9.1.2

une solution approchée valide pour le domaine entier. Considérons l'exemple suivant extrait du chapitre 2 de [64]

$$\varepsilon y^{(2)}(x) + 2y^{(1)}(x) + 2y(x) = 0, \quad pour \quad 0 < x < 1,$$

où $y(0) = 0$, $y(1) = 1$ et $0 < \varepsilon << 1$. Ce type de problème n'est pas évident. Si on prend $\varepsilon = 0$, on retrouve le problème suivant:

$$2y^{(1)}(x) + 2y(x) = 0, \quad for \quad 0 < x < 1.$$

La solution de cette équation a la forme $y(x) = Ae^{-x}$ et les limites au bord à 0 et 1 donnent $A = 0$ et $A = e^1$. Comme cette équation n'admet pas de solution, le problème est un problème de perturbation singulière. On peut appliquer la technique de perturbation comme suit: on trouve d'abord la solution intérieure et extérieure du problème, après on suppose que 0 est une couche limite [2] et ensuite on combine les deux solutions afin d'obtenir une solution adéquate pour le domaine entier (pour plus de détails, voir [64]).

La méthode à plusieurs échelles. Cette méthode commence de la solution générale et diffère de la méthode de développement asymptotique par recollement. En effet, elle introduit des coordonnées pour chaque région; ces coordonnées sont mutuellement indépendantes. Donc, cette méthode rajoute de nouvelles variables à l'équation différentielle ordinaire et la transforme en une équation différentielle partielle. En physique, la méthode à échelles est utilisée souvent pour le temps tandis que la méthode de recollement s'applique aux variables d'espace. Cette méthode à plusieurs échelles intervient quand les termes de correction ne sont pas négligeables et peuvent êtres non bornés. Donc, en utilisant ces nouvelles variables, on permet à l'amplitude de varier lentement et on évite que les termes correctifs soient non bornés (pour plus de détails, voir Chapitre 6 dans [87]).

La méthode de WKB. Dans la méthode de recollement, on divise le domaine de la solution en sous-intervalles afin de rendre explicite la dépendance de la solution en la couche limite. Pour la méthode à plusieurs échelles, on utilisait la dépendance en une nouvelle échelle afin de résoudre une nouvelle équation aux dérivées partielles. La méthode de WKB est un cas spécial de méthode à plusieurs échelles et suppose une dépendance exponentielle de la solution. Cette méthode est utilisée pour résoudre des équations ordinaires linéaires où la dérivée du plus grand ordre est multipliée par un paramètre petit ε. WKB est un acronyme pour la méthode d'approximation de Wentzel-Kramers-Brillouin. Les physiciens Wentzel, Kramers et Brillouin ont développé cette théorie dans les années 1920. Après, cette méthode a été utilisée par Jeffreys afin de trouver une approximation adéquate pour l'équation de Schrödinger. A partir d'une équation différentielle du type:

$$\varepsilon y^{(n)}(x) + a_{n-1}y^{(n)}(x) + \cdots + a_1 y^{(1)}(x) + a_0 y(x) = 0,$$

on suppose une série asymptotique de la solution de la forme:

$$y(x) = e^{\frac{1}{\mu}\sum_n \mu^n z_n(x)}$$

pour μ proche de 0. En utilisant cette forme dans l'équation différentielle, on obtient un nombre arbitraire d'équations qui nous permettent de calculer d'une manière récursive les solutions z_n (pour plus de détails sur cette méthode, voir chapitre 10 dans [16]).

[2] Cet a priori hypothèse est la seul façon de trouver la vraie location de la couche limite (voir [111])

La méthode d'homogénéisation. Scientifiques et ingénieurs rencontrent des matériaux qui impliquent des constituants hétérogènes (par exemple, on cite les plaques laminées et les fluides avec bulles). Bien que hétérogène, pour les étudier on ne considère pas les propriétés de chaque sous-composant à part, mais on suppose que le matériel est continûment distribué et on moyenne les informations caractéristiques. Cette méthode peut être vue comme une extension de la méthode à plusieurs échelles pour deux espaces (pour plus de détails, voir [5] et [64]

Autres approximations. On cite parmi ces approximations:

- L'expansion à plusieurs échelles pour les équations aux dérivées partielles: Ce type de problèmes implique spécialement les équations aux dérivées partielles ayant pour variables, le temps et l'espace (pour plus de détails, voir chapitre 6 in [30]).
- Variation de paramètres et les méthodes de moyennisation: Cette méthode traite des coefficients dépendants du temps pour des équations différentielles ordinaires ou partielles. Cette approximation suppose que les paramètres bougent lentement dans le temps. Donc, on peut trouver des paramètres équivalents qui approchent la solution (voir chapitre 5 in [87]).
- Coordonnées tendues. Cette technique traite la non uniformité dans les développements asymptotiques. En effet, elle introduit des transformations proches de l'identité dans les variables de l'équation différentielle afin d'obtenir des solutions approchées et uniformes (pour plus de détails, voir chapitre 3 dans [87]).

1.2.2 Les équations différentielles stochastiques

Perturbation à petit bruit (trajectoire). Freidlin et Wentzell ([47]) ont considéré le système dynamique aléatoire:

$$\dot{X}_t^\varepsilon = b(X_t^\varepsilon, \varepsilon\xi_t),$$

où ε est un paramètre petit, b est une fonction continue, ξ_t est un processus aléatoire continu et la solution X_t^ε part de la condition initiale x_0. Les auteurs démontrent qu'on peut développer la solution $(X_t^\varepsilon)_{t\in[0,T]}$ uniformément sur $[0,T]$ en puissance de ε:

$$x_t + \varepsilon Y_t^1 + \cdots + \varepsilon^n Y_t^n + o(\varepsilon^n),$$

où x_t est la solution du système dynamique non perturbé et la fonction $b(x,y)$ a $n+1$ dérivées bornées par rapport à x et y. (Pour plus de détails, voir chapitre 2 dans [47]). On cite également des résultats analogues pour le flot des équations différentielles stochastiques (Kunita dans [72] pour les diffusions, Fujiwara et Kunita dans [48] pour les diffusions avec sauts).

Perturbation à paramètre petit (distribution). Watanabe dans [112] montre que pour toute fonction bornée f et pour toute variable aléatoire F^ε régulière en ε et régulière aussi au sens de Malliavin avec la matrice de covariance $\gamma(F^\varepsilon)$ qui est inversible et d'inverse intégrable autour de $\varepsilon = 0$ qu'il existe un développement faible de Taylor à tout ordre $n \geq 0$:

$$\mathbb{E}[f(F^\varepsilon)] = \mathbb{E}[f(F_0) + \varepsilon G^1 + \cdots + \varepsilon^n G^n] + o(\varepsilon^n),$$

où les corrections G_i vérifient l'égalité:

$$\mathbb{E}[G^i] = \mathbb{E}[f(F^0)\pi_i],$$

et π_i sont des poids explicites à condition qu'on connaisse la matrice de Malliavin $\gamma(F^0)$ et son inverse.

Mais, en pratique, la question d'identifier un petit paramètre pertinent est primordiale et très difficile. Il y a plusieurs façon de paramétrer un modèle et les approximations qui en découlent peuvent être très différentes. On donne deux exemples.

- petit bruit:

$$dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dW_t. \tag{1.4}$$

- temps petit:

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t. \tag{1.5}$$

  On est intéressé par la loi de X_t pour t petit. Par un changement en espace et temps, $(X_{t\varepsilon})_{t\geq 0}$ a la même loi que $(X_t^\varepsilon)_{t\geq 0}$ défini par

$$dX_t^\varepsilon = \varepsilon b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dW_t.$$

Pour plus de détails sur la différence de notre approche avec celle de Watanabe, on renvoie à la section 1.5.

Théorie des grandes déviations. La théorie des grandes déviations donne des estimations de la queue de probabilité de certaines distributions (voir [110] pour détails). Freidlin et Wentzell dans [47] établissent un Théorème de grandes déviations pour les équations différentielles ordinaires perturbées par un mouvement brownien (perturbation à petit bruit). Ils considèrent l'équation différentielle stochastique suivante:

$$dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}dW_t, X_0^\varepsilon = x_0,$$

où W est un mouvement brownien multidimensionnel de dimension $d \in \mathbb{N}^*$ et la fonction $b : \mathbb{R}^d \to \mathbb{R}^d$ est bornée et Lipschitz aussi. Si C and G sont respectivement des fermés et ouverts de l'ensemble $\{f : [0,T] \to \mathbb{R}^d, f \text{ est continue}, \dot{f} \text{ est une fonction de carré intégrable}, f(0) = x_0\}$, on obtient alors:

$$\limsup_{\varepsilon\downarrow 0} \varepsilon \log \mathbb{P}(X^\varepsilon \in C) \leq -\inf_{f\in C} I(f),$$
$$\liminf_{\varepsilon\downarrow 0} \varepsilon \log \mathbb{P}(X^\varepsilon \in G) \geq -\inf_{f\in G} I(f),$$

où $I(f) = \frac{1}{2}\int_0^T |\dot{f}_t - b(f_t)|^2 dt$. Les auteurs étendent aussi ce résultat pour l'équation stochastique différentielle:

$$dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dW_t, X_0^\varepsilon = x_0,$$

où $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ est une fonction bornée et Lipschitz (pour plus de détails, voir Section 6 dans [110]).

Remarque 1.2.1. *La théorie introduite par Freidlin et Wentzell est très intéressante car elle décrit le comportement asymptotique pour les distributions des équations stochastiques différentielles. Mais, elle ne fournit que des estimations logarithmiques.*

Sensibilité par rapport à un paramètre. Gobet et Munoz dans ([55]) considèrent une diffusion multidimensionnelle $(X_t^\alpha)_t$ dont les dynamiques sous-jacentes dépendent d'un paramètre α. Ils expriment la sensibilité $\partial_\alpha \mathbb{E}[f(X_t^\alpha)]$ pour une fonction bornée mesurable f comme une espérance impliquant seulement f et non ses dérivées. On rencontre ce type de problème dans la théorie du contrôle stochastique. Les auteurs utilisent trois approches: Le calcul de Malliavin, l'approche adjointe et l'approche martingale. Des résultats numériques avec les erreurs associées aux schémas de discrétisation sont également présentés.

La projection Markovienne. Gyöngi dans [59] considère un processus stochastique (ξ_t) partant de 0 de décomposition d'Itô :

$$d\xi_t = \delta(t,\omega)dW_t + \beta(t,\omega)dt,$$

où W est un $\mathscr{F}_t$ adapté mouvement brownien, δ et β sont bornées et sont des processus $\mathscr{F}_t$ adaptées avec $\delta\delta^*$ étant définie positive (uniformément). L'auteur montre qu'il existe un processus markovien (X_t) qui a la même marginale unidimensionnelle que (ξ_t) (i.e. $\mathscr{L}(X_t) = \mathscr{L}(\xi_t) \quad \forall t$), et qui est une solution faible de l'équation ordinaire stochastique différentielle suivante:

$$dX_t = \sigma(t,X_t)dW_t + \beta(t,X_t)dt, X_0 = 0,$$

où

$$\sigma\sigma^*(t,x) = \mathbb{E}[\delta\delta^*(t)|\xi_t = x], b(t,x) = \mathbb{E}[\beta(t)|\xi_t = x].$$

Brunick ([28]) relâche l'hypothèse d'ellipticité δ.

Remarque 1.2.2. *La projection Markovienne a été la clé de plusieurs approximations spécialement en finance comme le calcul de la fonction de la volatilité locale pour les modèles à volatilité stochastique (voir [94] et [76]). Mais, aucun contrôle d'erreur n'a été établi pour ces approximations.*

1.3 Les approximations analytiques appliquées à la finance

Après avoir donné une idée générale sur les méthodes de perturbation, on se focalise ici sur la littérature des méthodes de perturbation utilisées en finance. Cette section a pour objectif de présenter plusieurs idées originales sur les méthodes de perturbation et leur limitations.

1.3.1 Les méthodes de perturbation

Dans ce domaine, on cite

- Le développement du prix par rapport à la corrélation entre le sous-jacent et sa volatilité. Antonelli and Scarlatti considèrent dans [12] un modèle stochastique général et établissent un développement en série du prix du Call en fonction de la corrélation. Leurs approximations incluent le modèle de Heston comme cas particulier. Chaque terme de correction est aussi approché avec une estimation d'erreur. En plus, les bornes d'erreur de l'approximation sont estimées; il faut préciser que l'approximation est établie seulement pour des paramètres homogènes en temps.

- Développement en petite volatilité de volatilité. Lewis dans [79] établit un développement du prix du call en fonction de la volatilité de la volatilité. Son travail est basé sur un développement formel de l'opérateur d'EDP pour la transformée de Fourier du prix du Call. Cette approximation est valable pour des modèles à volatilité stochastique comme les modèles de Heston, Heston généralisé et GARCH. Par conséquence, l'auteur établit des formules approchées pour les volatilités implicites Black Scholes dans ces modèles à volatilité stochastique. L'approximation est obtenue seulement dans le cas de paramètres constants.

- Technique de moyennisation. Cette technique, très importante, a été introduite par Piterbarg [93]. Elle peut être vue comme une application en finance du Théorème de projection markovienne établie par Gyöngi. Piterbarg détermine des paramètres équivalents constants pour un modèle de Heston dépendant du temps pour approcher le prix du Call. En effet, il établit dans le domaine des dérivés des taux d'intérêts des formules pour les skew et la volatilité stochastique équivalents reliés à des paramètres dépendants du temps. La formule s'étend aussi à d'autres domaines comme les dérivés actions ou change. L'approximation est valable seulement dans le cas de corrélation nulle.

1.3.2 L'analyse asymptotique

Les strikes extrêmes. Lee a montré dans [78] que la variance implicite est bornée par une fonction linéaire à la moneyness $\ln(\frac{K}{F})$ pour des strikes grands. L'auteur donne des formules explicites qui relient les gradients des ailes de la borne supérieure de la variance implicite et les moments finis maximaux du sous-jacent. Par exemple, il a montré pour l'aile gauche que si $q^* := \sup\{q : \mathbb{E}[S_T^{-q}] < \infty\}$ et

$$\beta^* := \limsup_{K \to 0^+} \frac{\sigma_{imp}^2(T,K)T}{|\ln(\frac{K}{F})|},$$

alors $\beta^* \in [0,2]$ et

$$q^* = \frac{1}{2}(\frac{1}{\sqrt{\beta^*}} - \frac{\sqrt{\beta^*}}{2})^2.$$

En plus, Benaim et Friz ([15]) ont amélioré la formule de Lee. En effet, ils ont prouvé que la borne supérieure de Lee peut devenir une limite à condition que certaines hypothèses techniques soient satisfaites. Ces hypothèses sont réalisées pour une vaste classe de modèles.

Maturités longues. On cite le travail de Tehranchi ([108]) qui donne une formule asymptotique des volatilités implicites Black et Scholes pour des maturités longues avec un contrôle précis de l'erreur. En effet, il a montré sous l'hypothèse que ($S_t \to 0$ presque sûrement quand $t \uparrow \infty$) que

$$T\sigma_{imp}^2(T,K) = 8|\ln(\mathbb{E}(S_T \wedge \frac{K}{S_0}))| - 4\ln(|\ln(\mathbb{E}(S_T \wedge \frac{K}{S_0}))|) + 4\ln(\frac{K}{S_0}) - 4\ln(\pi) + \varepsilon(\ln(\frac{K}{S_0}),T),$$

où

$$\sup_{-M \le x \le M} |\varepsilon(x,T)| \underset{T \to \infty}{\to} 0$$

pour tout $M > 0$.
On cite aussi les travaux de Gatheral dans [49] qui donne des bornes d'arbitages pour le skew des volatilités Black et Scholes implicites à la monnaie. Ces bornes d'arbitrages sont de l'ordre de $O(T^{-\frac{1}{2}})$.
En plus, Rogers et Tehranchi ([99]) ont démontré le Théorème du smile conjecturé par S. Ross. En effet, ils ont prouvé que la forme du smile ne peut pas bouger par décalage parallèle.

1.3.3 Mélange de méthodes de perturbation et d'analyse asymptotique

L'approche ergodique. Fouque et al ([44]) considèrent un modèle à volatilité stochastique où la volatilité est une fonctionnelle de Y_t^ε qui est un processus d'Ornshtein Uhlenbeck (on le dénote OU) avec un temps de corrélation petit ε:

$$\begin{aligned} \frac{dS_t^\varepsilon}{S_t^\varepsilon} &= rdt + f(Y_t^\varepsilon)dW_t, \\ dY_t^\varepsilon &= \frac{1}{\varepsilon}(m - Y_t^\varepsilon)dt + \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}}dB_t, \\ \langle W, B \rangle &= \rho dt. \end{aligned}$$

La normalisation par rapport à ε est réalisée de telle façon que la distribution asymptotique de Y_t^ε quand t tend vers l'infini soit une loi gaussienne $\mathcal{N}(m, \nu^2)$.

Les auteurs écrivent le générateur infinitésimal $\mathcal{L}^\varepsilon$ comme une sommation de trois termes:

$$\mathcal{L}^\varepsilon = \frac{1}{\varepsilon}\mathcal{L}^{(0)} + \frac{1}{\sqrt{\varepsilon}}\mathcal{L}^{(1)} + \mathcal{L}^{(2)}, \tag{1.6}$$

où

- $\mathcal{L}^{(0)} = \nu^2 \frac{\partial^2}{\partial y^2} + (m - y)\frac{\partial}{\partial y}$ est le générateur infinitésimal du processus OU Y défini par:

$$dY_t = (m - Y_t)dt + \nu\sqrt{2}dB_t, \tag{1.7}$$

- $\mathcal{L}^{(1)} = \sqrt{2}\rho x f(y)\frac{\partial^2}{\partial x \partial y}$ contient les dérivées mixtes dues au terme de corrélation.

- $\mathcal{L}^{(2)} = \frac{\partial}{\partial t} + \frac{1}{2}f(y)^2 x^2 \frac{\partial^2}{\partial x^2} + r(x\frac{\partial}{\partial x} - .)$ est l'opérateur Black et Scholes avec volatilité $f(y)$.

Fouque et al ont supposé que le paramètre ε est petit ce qui fait du problème (1.6) un problème de perturbation singulière. Ils ont développé le prix du call coté en t, avec maturité T, cours S et strike K, en puissance de $\sqrt{\varepsilon}$:

$$\text{Call}^\varepsilon = \text{Call}_{BS}^{(0)} + \sqrt{\varepsilon}\text{Correction}^{(1)} + \cdots$$

où le terme de base est le prix Black-Scholes $\text{Call}_{BS}^{(0)} = Call_{BS}(t, S, K, T, \bar{\sigma})$ et la variance $\bar{\sigma}^2$ est la moyenne de la fonction f par rapport la distribution invariante $\mathcal{N}(m, \nu^2)$ de l'OU (Y) défini par l'équation (1.7):

$$\bar{\sigma}^2 = \frac{1}{\nu\sqrt{2\pi}} \int_{\mathbb{R}} f^2(y) e^{\frac{-(m-y)^2}{2\nu^2}} dy \equiv \langle f^2 \rangle, \tag{1.8}$$

et les termes correctifs sont des combinaisons linéaires de Grecques du terme principal $\text{Call}_{BS}^{(0)}$:

$$\sqrt{\varepsilon}\text{Correction}^{(1)} = -(T-t)(V_2 S^2 \frac{\partial^2 Call_{BS}(t, S, K, T, \bar{\sigma})}{\partial S^2} + V_3 S^3 \frac{\partial^3 Call_{BS}(t, S, K, T, \bar{\sigma})}{\partial S^3}),$$

où les coefficients V_2 et V_3 sont calculés comme la volatilité $\bar{\sigma}$ dans l'équation 1.8 utilisant l'opérateur $\langle . \rangle$:

$$V_2 = \sqrt{2}\rho\nu\langle f\phi' \rangle,$$
$$V_3 = \frac{\rho\nu}{\sqrt{2}}\langle f\phi' \rangle,$$

et ϕ est la solution de l'équation de Poisson:

$$\mathscr{L}^{(0)}\phi(y) = f(y)^2 - \langle f^2 \rangle.$$

En plus, les auteurs ont démontré dans [45] que l'erreur d'approximation pour les options call(put) se comporte comme:

$$\lim_{\varepsilon \downarrow 0} \frac{|(\text{Call}^{\varepsilon} - \text{Call}_{BS}^{(0)} - \sqrt{\varepsilon}\text{Correction}^{(1)})|}{\varepsilon|\ln(\varepsilon)|^{1+p}} = 0,$$

pour tout $p > 0$.

Remarque 1.3.1. *L'approximation est intéressante parce qu'elle donne une formule analytique rapide exprimée comme combinaison du terme principal de Black et Scholes et ses Grecques. En plus, les bornes d'erreurs sont données pour toutes les options call(put). Mais, cette approximation n'est pas valide pour les maturités courtes car elle est reliée à la propriété de retour à la moyenne de la volatilité, un comportement qui n'est pas instantané. En plus, elle est restreinte seulement à des coefficients homogènes en temps parce qu'elle utilise le générateur d'EDP.*

L'approche géodésique. Il existe plusieurs livres et papiers de recherche traitant l'approche géodésique et ses applications en finance:

- voir Chavel dans ([29]) pour une introduction à propos de la géométrie riemannienne et Varadhan ([109]) pour une asymptotique de la densité en temps petit,
- voir Berestycki et d'autres dans [21], Labordere dans [74] and [75], Lewis dans [80], Forde dans [43], Benhamou et autres dans [17] pour l'application de la géométrie Riemannienne en finance.

Ces travaux montrent que la fonction de Green π (la densité de valorisation) qui est solution de:

$$\frac{\partial \pi}{\partial t} = \sum_{i,j} g^{i,j} \frac{\partial^2 \pi}{\partial S^i \partial S^j} + \sum_i h_i \frac{\partial \pi}{\partial S^i},$$

avec la condition initiale $\pi(t_0, S_0, t_0, S) = \delta(S_0 = S)$, a un développement en temps petit:

$$\pi(t_0, S_0, t, S) = e^{\frac{-d^2(S_0,S)}{2(t-t_0)}} (G_0(S_0,S) + (t-t_0)G_1(S_0,S) + \cdots)$$

où $d(S_0, S)$ est la distance géodésique associée à l'espace riemannien défini par la distance métrique $ds^2 = \sum_{i,j} g_{i,j} dx^i dx^j$ où $g_{i,j}$ est l'inverse de la matrice de $g^{i,j}$. Le terme G_0 est relié à la distribution gaussienne. En effet la densité de valorisation se comporte comme une gaussienne quand on utilise ces nouvelles variables géodésiques. Le terme G_1 est relié à la courbure riemannienne.

De plus, Berestycki et al (voir Th 1.2 dans [21]) ont montré pour un modèle de volatilité stochastique générale que la volatilité implicite Black et Scholes près de la maturité est:

$$\lim_{T \downarrow t} \sigma_{imp}(T,K) = \frac{\ln(\frac{S}{K})}{d(\ln(\frac{S}{K}), y_0)} \tag{1.9}$$

où t est le temps de cotation, T la maturité, S le cours, K le strike, y_0 la valeur initiale de la volatilité stochastique et d est la distance géodésique signée qui résout l'équation (1.12) du Théorème 1.2 dans [21]. Il convient aussi de remarquer que Berestycki et al ([21]) ont établi dans certains cas des termes correctifs pour le développement en temps petit pour les volatilités implicites Black et Scholes.

Remarque 1.3.2. *Cette approche géodésique est très intéressante du fait qu'elle donne des formules analytiques pour la volatilité implicite Black et Scholes près de la maturité*[3]*. En outre, on peut dériver des développements asymptotiques pour les volatilités implicites Black et Scholes pour des maturités courtes. Mais, ces approximations sont restreintes seulement pour des maturités courtes et des paramètres homogènes.*

1.4 Les motivations et la méthodologie

A travers la section précédente, on a présenté plusieurs idées originales sur le calcul exact ou les méthodes analytiques. Remarquons aussi que les formules fermées couvrent une petite classe de modèles: Black et Scholes, Merton, CEV, ... Pour des modèles plus généraux, on peut établir des techniques d'approximation. Mais, ces formules analytiques ont été souvent obtenues sous des restrictions comme courte maturité (le développement géodésique), longue maturité (l'approche ergodique), les strikes extrêmes (l'approche de Lee), corrélation nulle (l'approche de Piterbarg) où des coefficients homogènes en temps (toutes les précédentes méthodes de perturbation à part celle de Piterbarg). Notre objectif dans cette thèse est d'élaborer une méthode d'approximation précise qui s'applique aux courtes et longues maturités, aux petits et grands strikes, à des paramètres non homogènes en temps ou à des corrélations non nulles. Pour assurer qu'une méthode de perturbation soit utilisée efficacement en pratique, la quantité d'intérêt (i.e. le prix) devrait être décomposée comme une sommation de termes explicites: un terme principal auquel on rajoute quelques termes correctifs. Dans notre approche, la partie principale est calculée par un modèle proxy et les termes correctifs quantifient la distance entre le vrai modèle et le modèle proxy; cette distance au modèle proxy est caractérisée par un petit paramètre Δ. Par exemple,

- dans le modèle de Heston, un modèle proxy possible peut être le modèle de Black et Scholes en prenant la volatilité de la volatilité égal à 0. Dans ce cas, on peut définir $\Delta = \xi$.
- dans les modèles à volatilité locale, le modèle de Black et Scholes peut être un modèle proxy en figeant la fonction de la volatilité locale σ. Dans ce cas, Δ peut être la norme supérieure des dérivées de σ.

En fait, cette analyse de la distance au modèle proxy est très grossière et dans cette thèse, on donnera des mesures précises pour cette distance par rapport au modèle proxy. Maintenant, si on note $(X_t)_t$ le vrai modèle pour le prix du sous-jacent (ou du log du sous-jacent) et $(X_t^P)_t$ le modèle proxy, on peut écrire,

[3]Remarquez aussi que ces calculs analytiques des volatilités implicites ne sont pas toujours explicites comme c'est le cas du modèle de Heston

au moins formellement, un développement de Taylor pour le prix des options vanilles h pour le cours X qui expirent à la date T:

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] + \mathbb{E}[h^{(1)}(X_T^P)(X_T - X_T^P)] + \cdots + \mathbb{E}[h^{(j)}(X_T^P)\frac{(X_T - X_T^P)^j}{j!}] + \mathscr{R}esid_j. \tag{1.10}$$

Le terme principal $\mathbb{E}[h(X_T^P)]$ est explicite, puisque souvent les prix dans le modèle proxy sont explicites (en effet, c'est une contrainte qu'on s'impose pour le choix du modèle proxy). Le deuxième terme $\mathbb{E}[h^{(1)}(X_T^P)(X_T - X_T^P)]$ jour le rôle du premier terme de correction, mais il n'est pas explicite pour des processus généraux et doit être approché. Pour mieux approcher ce terme, nous avons pensé à le décomposer comme une sommation de Grecques (dans le modèle proxy) plus un résidu:

$$\mathbb{E}[h^{(1)}(X_T^P)(X_T - X_T^P)] = \sum_{i=1}^{n_1} a_i^{(1)} \partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0} + O(\Delta^k). \tag{1.11}$$

Les Grecques dans le modèle proxy doivent être explicites aussi. La réalisation de la décomposition est fortement dépendante du modèle et on peut l'établir en utilisant le calcul de Malliavin combiné avec une paramétrisation fine du modèle de X. Il convient de préciser que cette étape est spécifique au modèle, on ne peut écrire une théorie générale et on se réfère à l'introduction de chaque partie de la thèse pour plus de détails. Pour aider le lecteur, on va expliquer pourquoi les Grecques apparaissent naturellement dans (1.11). L'identification en utilisant les Grecques peut être considérée comme une procédure inverse de celle utilisée dans la littérature pour la formule d'intégration par parties et le calcul de Malliavin ([46]). En effet, on sait que $\partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0} = \mathbb{E}[h^{(i)}(X_T^P)] = \mathbb{E}[h(X_T^P) H_T^i]$ pour certains "poids de Malliavin" H_T^i et pour établir (1.11), on devrait expliciter les $(a_i^{(1)})_i$ afin que $\sum_{i=1}^n a_i^{(1)} H_T^i \approx (X_T - X_T^P)$. Après, on répète la décomposition (1.11) pour chaque terme du type $\mathbb{E}[h^{(j)}(X_T^P)\frac{(X_T - X_T^P)^j}{j!}]$ et l'écrire aussi de la forme $\mathbb{E}[h^{(j)}(X_T^P)\frac{(X_T - X_T^P)^j}{j!}] = \sum_{i=1}^{n_j} a_i^{(j)} \partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0} + O(\Delta^k)$. Finalement, on ramène tous ces éléments dans l'équation (1.10) jusqu'à l'ordre minimal j_k qui vérifie $\mathscr{R}esid_{j_k} = O(\Delta^k)$. Ceci implique:

$$\begin{aligned}
\mathbb{E}[h(X_T)] &= \mathbb{E}[h(X_T^P)] + \sum_{j=1}^{j_k} [\sum_{i=1}^{n_j} a_i^{(j)} \partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0}] + O(\Delta^k) \\
&= \mathbb{E}[h(X_T^P)] + \sum_{i=1}^{max_{j \le j_k} n_j} (\sum_{j=1}^{j_k} a_i^{(j)} \mathbb{1}_{i \le n_j}) \partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0} + O(\Delta^k).
\end{aligned}$$

La forme générale est la suivante:

$$\boxed{\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] + \textit{ Somme pondérée de Grecques } \partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0} + \textit{ erreur.}} \tag{1.12}$$

Évidemment, ces arguments formels ont besoin d'être mathématiquement clarifiés en prenant en compte les considérations suivantes:

- le payoff h n'est pas régulier (la dérivée seconde du payoff du call n'existe pas dans le sens classique, ce qui évite d'écrire directement (1.10)).
- on devrait utiliser prudemment le calcul de Malliavin pour établir des développements explicites pour les coefficients $(a_j^{(i)})$.
- L'estimation d'erreur est une tache très difficile car elle dépend du modèle et du payoff.

1.5 La comparaison avec l'approche de Watanabe

Il est utile de signaler que le cadre mathématique de Watanabe (voir [112]) proposant les développements asymptotiques pour approcher des fonctionnelles de Wiener est une référence très importante, et qu'il existe des similitudes et des différences significatives entre l'approche de Watanabe et celle que nous introduisons dans ce travail de thèse. Dans ses travaux, Watanabe considère une famille $(F^\varepsilon)_{\varepsilon\geq 0}$ de variables aléatoires définies dans l'espace de Wiener et régulières au sens de Malliavin. Supposons qu'on puisse écrire un développement asymptotique de F^ε en puissance de ε à tout ordre k:

$$F^\varepsilon - (f_0 + \varepsilon f_1 + \cdots + \varepsilon^k f_k) = O(\varepsilon^{k+1}) \text{ quand } \varepsilon \to 0^+ \tag{1.13}$$

Les variables aléatoires $(f_i)_i$ sont régulières au sens de Malliavin et l'égalité ci-dessus avec le symbole de Landau est vraie dans les normes de Sobolev $\|\cdot\|_{\mathbb{D}^{l,p}}$. Supposons qu'on ait de plus une condition de non dégénérescence uniforme:

$$\limsup_{\varepsilon\to 0^+} \|\det(\gamma_{F^\varepsilon}^{-1})\|_p < \infty \text{ pour tout } p \geq 1 \tag{1.14}$$

où γ_{F^ε} est la matrice de covariance de Malliavin de F^ε. Watanabe montre que, pour toute fonction h avec croissance polynomiale, on obtient:

$$\mathbb{E}[h(F^\varepsilon)] = \mathbb{E}[h(F^0)] + \varepsilon\mathbb{E}[h(F^0)\pi_1] + \cdots + \varepsilon^k\mathbb{E}[h(F^0)\pi_k] + O(\varepsilon^{k+1}) \tag{1.15}$$

où $(\pi_i)_i$ est une suite de variables aléatoires. Cette égalité peut être étendue aux distributions h (Théorème 2.3 dans [112]).

A première vue, les formules (1.12) et (1.15) sont du même type en prenant $X_T = F^\varepsilon$ où $F^0 = X_T^P$. Avec cette analogie, on pourrait essayer de relier les Grecques dans (1.12) avec les termes $(\mathbb{E}[h(F^0)\pi_i])_i$ dans (1.15), mais en fait cette identification est loin d'être évidente.

Une caractéristique importante de l'approche de Watanabe réside dans le fait que la précision du développement est écrite en fonction du petit paramètre ε. Cette paramétrisation par rapport ε est cruciale comme le montre les exemples précédents dans les équations (1.4) and (1.5). Le dernier point à soulever, qui a toute son importance, est le fait que l'impact des paramètres des modèles n'est pas pris en compte dans les estimations. Ce point est un inconvénient significatif de tels calculs. Pour illustrer ceci, on considère le modèle jouet:

$$F^\varepsilon = \sigma W_1 + \sqrt{\varepsilon} B_1$$

où (W,B) est un mouvement brownien deux-dimensionnel, et σ est un nombre strictement positif. Développons $\mathbb{E}[h(F^\varepsilon)]$ en puissance de ε pour $h(x) = x^2$ et $h(x) = x^+$. On utilise le fait que F^ε est distribuée comme une $\mathcal{N}(0, \sigma^2 + \varepsilon)$. Clairement $F^0 = \sigma W_1$.

1. Cas $h(x) = x^2$. On a

$$\begin{aligned}\mathbb{E}[h(F^\varepsilon)] = \mathbb{E}[(F^\varepsilon)^2] &= \sigma^2 + \varepsilon \\ &= E[h(F^0)] + \varepsilon.\end{aligned}$$

2. Cas $h(x) = x^+$. En utilisant un changement d'échelle, on obtient:

$$\begin{aligned}\mathbb{E}[h(F^\varepsilon)] &= \sqrt{\sigma^2+\varepsilon}\,\mathbb{E}[(W_1)^+]\\ &= \sqrt{\frac{\sigma^2+\varepsilon}{\sigma^2}}\,\mathbb{E}[(\sigma W_1)^+]\\ &= E[h(F^0)] + \frac{1}{2}\frac{\varepsilon}{\sigma}\mathbb{E}[(W_1)^+] + O(\frac{\varepsilon^2}{\sigma^3}).\end{aligned}$$

Ces calculs sont cohérents avec les résultats du développement de Watanabe:

$$\mathbb{E}[h(F^\varepsilon)] = E[h(F^0)] + \varepsilon c_1 + O(\varepsilon^2).$$

On observe que les coefficients c_1 dépendent fortement du modèle et de la fonction h (dans le second cas, $\varepsilon c_1 = \frac{1}{2}\frac{\varepsilon}{\sigma}\mathbb{E}[(W_1)^+]$). Dans le cas où σ est petit aussi, on remarque que la précision du développement est fortement reliée au ratio $\frac{\varepsilon}{\sigma}$ et non seulement à ε. Ceci représente un autre argument fort contre l'approche directe du développement asymptotique de Watanabe. Dans nos travaux, on donne des estimations non-asymptotiques, qui nous permettent de déduire le domaine de validité de nos formules par rapport à tous les paramètres du modèle. Autrement dit, une partie significative de nos travaux met en relief l'impact des paramètres du modèle sur l'approximation. Ceci est confirmé dans ce travail de thèse par de nombreux résultats numériques (voir Chapitre 4, 7, 10, 12). On montre aussi que l'ordre de grandeur de l'erreur dépend de la régularité du payoff. L'exemple précédent est une illustration convaincante de ce phénomène, tandis que dans l'approche de Watanabe, la régularité du payoff ne joue aucun rôle dans les estimations.

Maintenant, on discute en détail les différences entre les démonstrations et aussi entre les méthodologies.

1. Dans notre approche, on quantifie les erreurs par rapport à la régularité du payoff: on réalise un développement de Taylor pour les options régulières comme on a montré auparavant. On utilise ensuite une méthode de régularisation et une intégration par parties (Calcul de Malliavin) afin de majorer les erreurs et exprimer les termes tronqués comme une combinaison de Grecques du terme principal.

2. Contrairement à notre approche, Watanabe donne un développement asymptotique pour la densité d'un processus en exprimant la densité comme une espérance d'une distribution de Dirac. Dans le théorème 2.3 de [112] pour l'approximation des distributions de fonctionnelles de Wiener: il utilise l'intégration par parties (calcul de Malliavin) appliquée aux distributions afin de retrouver des fonctions tests régulières. Ensuite, il utilise un développement de Taylor afin d'exprimer les corrections en fonction des dérivées de la distribution. On estime que le développement par la formule d'intégration par partie en utilisant le calcul de Malliavin est moins souple comparé à notre approche directe. Autrement dit, pour avoir des formules fermées, il est plus facile de "développer et ensuite intégrer par parties" que "d'intégrer par parties et ensuite développer".

3. Un facteur commun entre les deux approches est qu'on suppose que le proxy (X_T^P où F^0) est relié à un processus (variable aléatoire) gaussien. De plus, dans notre cas, on utilise des proxys log-normaux.

4. Une autre différence technique réside dans les hypothèses utilisées dans le développement. Dans les résultats de Watanabe, la régularité $\mathscr{C}^\infty$ pour les coefficients des modèles est requise. Dans

notre cadre, on suppose une régularité optimale pour les hypothèses (par exemple, voir hypothèse (R_4) dans la section 4.1). On traite aussi des cas non réguliers comme le modèle de Heston (du a la racine carré dans le coefficient de diffusion).

L'approche Watanabe en finance. Yoshida dans [113] [114], Kunitomo et Takahashi dans [73] appliquent l'approche de Watanabe en finance et utilisent le calcul de Malliavin pour le contrôle de l'erreur. De plus, Yoshida affaiblit l'hypothèse de non dégénérescence (1.14) dans une version localisée (permettant la dégénérescence sur un ensemble de mesure exponentiellement petite); voir Théorème 4.1 page 152 dans [113]. Teichmann et Siopacha dans [106] appliquent aussi les résultats de Watanabe pour le modèle du Libor en utilisant des simulations de Monte Carlo. Ils obtiennent des expressions facilement utilisables pour une valorisation précise.

5. En ce qui concerne les calculs des termes correctifs, on utilise directement la formule d'intégration par partie (comme on a montré auparavant) et on obtient que les termes correctifs sont une combinaison linéaire des Grecques dans le modèle proxy avec des coefficients dépendant seulement des paramètres du modèle. Yoshida (Théorème 2.1 in [113], Théorème 4.1 in [114]) and Takahashi (Théorème 3.3 in [73]) utilisent des calculs explicites reliés à des espérances conditionnelles pour des vecteurs gaussiens et expriment les termes correctifs comme une intégrale de produits d'une densité gaussienne avec des polynômes. En d'autre termes, la densité en question est approchée par une combinaison linéaire de dérivées de la densité gaussienne, montrant ainsi une forte similarité entre les différentes approches. Remarquons aussi que notre approche nous permette de traiter aussi le cas d'un processus de Poisson avec des sauts gaussiens, en utilisant une paramétrisation appropriée et dépendante du modèle (voir Partie I).

1.6 La structure de la thèse et les principaux résultats

La thèse est divisée en quatre parties. Chaque partie traite des modèles spécifiques et contient trois chapitres:

- Le premier chapitre donne une introduction pour le modèle étudié. En plus, il présente les méthodes existantes numériques et analytiques utilisés pour la valorisation de options européennes dans un tel modèle. Finalement, il présente la motivation pour un tel travail et les résultats de la partie.
- Le deuxième chapitre contient tous les résultats mathématiques et leur preuves.
- Le troisième chapitre traite les résultats numériques, comme la robustesse de la procédure calibration et la précision de la formule pour des strikes extrêmes et des maturités longues.

Partie I. Dans cette partie, on considère le modèle d'Andersen et Andreasen ([8]) qui est un modèle à volatilité locale plus des sauts gaussiens:

$$dX_t = \sigma(t,X_{t^-})dW_t + \mu(t,X_{t^-})dt + dJ_t, X_0 = x_0, \tag{1.16}$$

où

- X_t est le logarithme du forward $F_t = S_t e^{\int_0^t (q_s - r_s)ds}$, r est le taux déterministe risque neutre et q est le taux continu de dividende,

- le processus de Poisson composé (J_t) et le mouvement brownien (W_t) sont indépendants,
- $J_t = \sum_{i=1}^{N_t} Y_i$ et Y_i sont des variables normales i.i.d. avec moyenne η_J et volatilité γ_J,
- N_t est un processus de Poisson avec intensité λ,
- $\mu(t,x) = -\frac{\sigma^2(t,x)}{2} + \lambda(1 - e^{\eta_J + \frac{\gamma_J^2}{2}})$ afin de garantir la propriété de martingale pour (e^{X_t}).

Notre objectif dans cette partie est

- d'établir **une approximation analytique précise** pour l'option européenne:

$$\mathbb{E}[h(X_T)]$$

  où h est une **fonction non régulière** et T est la maturité de l'option. Pour le cas d'une option de call sans taux d'intérêts et sans versement de dividendes, on a $h(x) = (e^x - K)^+$.
- de calibrer dans **un temps de calcul** inférieur à une seconde (**1sec**) tous les paramètres du modèle: les paramètres de saut λ , η_J, γ_J et la fonction de volatilité $\sigma(t,x)$.

On peut montrer qu'un bon modèle proxy peut être le modèle de Merton:

$$dX_t^M = \sigma(t,x_0)dW_t + \mu(t,x_0)dt + dJ_t, X_0^M = x_0,$$

et l'option vanille dans un tel modèle est:

$$E[h(X_T)] = Merton\ price + corrections\ terms + errors, \tag{1.17}$$

où

- Les termes correctifs de l'équation (1.17) sont des combinaisons linéaires de Grecques (Delta, Gamma, Epsilon) du modèle de Merton (Il ne faut pas oublier que les Grecques ont des formules fermées dans le cas du modèle de Merton). Ces corrections sont explicitées dans le chapitre 4 du Theorème 4.2.1.
- Le terme d'erreur *errors* dans l'équation (3.8) est estimé pour les payoffs vanilles dans le chapitre 4 dans le Théorème 4.5.2. Une interprétation naive de l'erreur est la suivante: si les dérivées de volatilité σ et la taille des sauts impliqués dans l'équation (3.5) sont de l'ordre de Δ, alors le terme d'erreur est de l'ordre de $(\Delta\sqrt{T})^3$.

Remarquons aussi que l'approximation établie dans le chapitre 4 ne couvre pas seulement les options Call-Put mais aussi les payoffs vanilles dépendants de X_T (réguliers, vanilles, digitales). En effet, les erreurs sont analysées en fonction de la régularité du payoff (voir par exemple le Théorème 4.5.1 pour les payoffs réguliers en général, le Thèorème 4.5.2 pour les options vanilles, le Théorème 4.5.3 pour les options digitales). On montre aussi dans le chapitre 3 que la précision de notre formule s'avère excellente (les erreurs pour les volatilités implicites n'excèdent pas 2 points de base pour de nombreux strikes et maturités). En conséquence, la calibration d'un tel modèle devient très rapide.

Le chapitre 3 donne une idée générale sur les méthodes utilisés pour générer le smile et on introduit aussi le modèle d'Andersen et Andreasen. De plus, dans ce chapitre, on fournit des détails sur le modèle de Merton, la calibration utilisant la forward PIDE et on énonce finalement les résultats principaux de la partie. Le chapitre 4 dans la partie I est la reproduction exacte de l'article "smart expansion and fast calibration for jump diffusions" publié dans la revue "Finance and Stochastics". Le chapitre 5 dans la partie I donne des résultats numériques supplémentaires concernant la précision de la formule quand on stresse les paramètres, sans oublier la robustesse de la calibration.

Partie II. Dans cette partie, on s'intéresse à des modèles à volatilité locale([40]):

$$dX_t = \sigma(t,X_t)dW_t + \mu(t,X_t)dt, X_0 = x_0, \tag{1.18}$$

où X_t peut être le sous-jacent ou son logarithme.

L'objectif de cette partie est **d'établir des formules fermées pour les options vanilles dans un modèle à volatilité locale générale**. Cette formule fermée est en effet un développement de Taylor et peut être tronqué facilement à tout ordre. Par conséquent, le prix de l'option vanille peut être écrit à tout ordre comme la somme des termes suivants:

- le prix Black Scholes avec la volatilité à la monnaie. Comme dans la partie I, ce modèle peut être vu comme un modèle proxy pour le modèle à volatilité locale. L'avantage de ce modèle proxy réside dans le calcul explicite des prix et des Grecques de l'option vanille.

- Une combinaison linéaire des Grecques du prix Black Scholes avec des poids explicites dépendants de la volatilité, de la dérive et leurs dérivées.

- Une erreur résiduelle avec des bornes explicites.

Ceci est réalisé dans le chapitre 7. L'approximation pour les options vanilles au second ordre est calculée dans le Théorème 7.2.1 qui est un cas particulier du Théorème 4.2.1 du chapitre 4 lorsqu'il n'y a pas de sauts. De plus, le calcul explicite de l'approximation pour les options vanilles au troisième ordre est fournie dans le Théorème 7.2.2. En outre, ces termes correctifs et les erreurs sont estimés pour tout ordre dans les Théorèmes 7.4.1-7.4.2-7.4.3 en fonction de la régularité du payoff (réguliers, vanilles, digitales). La précision de l'approximation s'avère être excellente. En plus, on n'a besoin que de quelques termes pour donner des résultats précis pour les options vanilles. Cette méthodologie nous permet de calculer des paramètres équivalents pour un modèle CEV dépendant du temps.

Le chapitre 6 donne une introduction générale sur le modèle à volatilité locale, le modèle CEV et les approximations analytiques existantes utilisées pour la valorisation d'un tel modèle. En plus, on détaille dans ce chapitre les motivations et les résultats principaux de la partie. Le chapitre 7 est la reproduction exacte de l'article "Closed forms for European options in a local volatility model" accepté pour publication dans le journal "International Journal of Theoretical and Applied Finance". Le chapitre 8 détaille le comportement du smile pour le modèle CEV quand on varie ses paramètres au cours du temps. On fournit aussi des résultats numériques concernant la précision de l'approximation pour des strikes très grands et aussi le domaine d'arbitrage des formules d'approximations.

Part III. Cette partie traite le modèle de Heston dépendant du temps ([63]):

$$\mathrm{d}X_t = \sqrt{v_t}dW_t - \frac{v_t}{2}\mathrm{d}t,\ X_0 = x_0, \tag{1.19}$$

$$\mathrm{d}v_t = \kappa(\theta_t - v_t)\mathrm{d}t + \xi_t\sqrt{v_t}\mathrm{d}B_t,\ v_0, \tag{1.20}$$

$$\mathrm{d}\langle W,B\rangle_t = \rho_t\mathrm{d}t,$$

où

- X_t est le logarithme du forward $e^{(q-r)t}S_t$, r et q sont respectivement le taux risque neutre et le taux continu de dividende,

- v_0 est la valeur initiale du carré de la volatilité,

- κ est le paramètre de retour à la moyenne,
- θ est le niveau long-terme,
- ξ est la volatilité de la volatilité,
- ρ est la corrélation

Notre objectif dans cette partie est

- d'établir **une approximation analytique précise** pour le prix de l'option Call-Put
$$e^{-rT}\mathbb{E}[(K-e^{(r-q)T+X_T})_+] \tag{1.21}$$
- de travailler dans **un cadre général de modèle de Heston dépendent du temps** afin d'obtenir des approximations qui couvrent à la fois des maturités courtes et longues, des paramètres dépendant du temps et des corrélations non nulles également.
- de réaliser **un temps de calcul très petit par rapport à l'inversion de Fourier (un gain par un facteur de 100 ou plus)**.

Ceci est réalisé dans le chapitre 10 utilisant un développement précis en petite volatilité de volatilité avec des techniques de calcul de Malliavin et quelques lemmes techniques.

Le modèle proxy ici est le modèle de Heston sans volatilité de volatilité. Autrement dit, c'est le modèle Black et Scholes avec une volatilité dépendante du temps:

$$\begin{aligned} \mathrm{d}X_t^{BS} &= \sqrt{v_{0,t}}dW_t - \frac{v_{0,t}}{2}\mathrm{d}t,\ X_0^{BS} = x_0, \\ \mathrm{d}v_{0,t} &= \kappa(\theta_t - v_{0,t})\mathrm{d}t, v_0. \end{aligned}$$

On prouve que:

$$\begin{aligned} e^{-rT}\mathbb{E}[(K-e^{(r-q)T+X_T})_+] = &\ Put\ \ Price\ \ in\ \ BS\ \ model \\ &+ Correction\ \ terms + Errors. \end{aligned}$$

- Les termes correctifs sont des combinaisons linéaires de Grecques du prix Black Scholes avec des poids dépendant uniquement des paramètres du modèle. Ces calculs sont faits à l'aide du calcul de Malliavin dans le Théorème 10.2.1.
- Les erreurs sont estimées dans le Théorème 10.2.2 par:
$$Errors = O(|\xi|_\infty^3 T^2).$$

  Un choix possible de la mesure Δ peut être la norme supérieure de la fonction vol de vol ξ.

A partir de la formule approchée, on déduit des corollaires reliés premièrement à des paramètres équivalents pour le modèle de Heston (une extension du travail de Piterbarg pour les modèles à volatilité stochastique) et deuxièmement pour la procédure de calibration en terme de problèmes mal-posés (voir section 10.2.6).

Le chapitre 9 donne une introduction générale sur les formules d'inversion de Fourier et les approximations analytiques dans un tel modèle. De plus, on énonce les résultats principaux de la partie dans ce chapitre. Le chapitre 10 est exactement la reproduction de l'article "Time dependent Heston model" en révision pour la revue "SIAM Journal on Financial Mathematics". Le chapitre 11 donne des résultats numériques additionnels concernant le comportement du smile pour le modèle de Heston avec des paramètres constants et dépendant du temps aussi. De plus, dans ce chapitre, on détaille des résultats numériques concernant les moments négatifs de l'intégrale du processus de CIR.

Partie IV. Dans cette Partie on traite des modéles hybrides avec une composante action modélisée avec une volatilité locale et une autre composante taux d'intérêt stochastiques:

$$\frac{dS_t}{S_t} = r_t dt + \sigma_t dW_t,$$

où (S_t) est le spot, (σ_t) sa volatilité aléatoire et (r_t) est le taux stochastique de dynamique HJM.

L'objectif de cette Partie est de calculer de nouvelles approximations pour les options européennes dans un modèle à volatilité locale et des taux stochastiques. Nous introduisons un modèle de volatilité locale sur le sous-jacent actualisé. Cette modélisation nous permet d'obtenir des formules faciles à implémenter pour les options vanilles dans un tel modèle.

Dans le Théorème 12.2.1, on donne une formule d'approximation au second ordre pour les options vanilles dans un modèle à volatilité locale générale combiné avec des taux stochastiques Gaussiens qui suivent le cadre HJM. Cette formule est la somme:

- du prix dans le modèle proxy de Black et Scholes avec des taux stochastiques.
- et des termes correctifs qui sont une combinaison linéaire de Grecques du terme principal avec des poids explicites dépendant des paramètres de la diffusion et des taux stochastiques.

On donne aussi la formule d'approximation au troisième ordre (voir Théorème 12.2.2) dans un tel modèle. Dans le cas de volatilité locale homogène avec un modèle de Hull and White pour les taux, on donne des calculs explicites pour les poids des corrections dans la sous-section 12.2.1. On étend aussi nos résultats pour le cas de dividendes stochastiques et aussi pour le cas des taux de convenance stochastiques (voir Section 12.3). Par exemple, dans le domaine des matières premières, on peut voir notre travail comme une extension du modèle de Gibson Schwartz afin de supporter des fonctions de volatilité locale. Dans la Section 12.4, on donne des exemples numériques illustrant la précision de notre formule d'approximation. En effet, on compare notre formule avec les simulations Monte Carlo avec variable de contrôle. La précision de nos formules s'avère être excellente.

Chapter 2

Introduction

This thesis develops a new methodology for deriving analytical approximations of the prices of European options. Our approach smartly combines stochastic expansions and Malliavin calculus to obtain explicit formulas and tight error estimates. The striking feature of these formulas is their rapidity to be evaluated (as quick as Black and Scholes formula). Our motivation comes from the increasing need for real-time computations and calibration procedures, while controlling numerical errors with respect to the model parameters. There are many ways to derive closed-form formulas, either by a straight computation of the option pricing problem or by some approximations. For models with explicit density for the spot like the Black Scholes model, CEV model or any model with explicit characteristic function of the spot distribution (like the Heston or affine models), one can easily find closed-form formulas. It leads to well known results for call and put options. In the case of non explicit density, one must turn to numerical methods (PDE techniques, Monte Carlo simulations, ...). But to obtain explicit formulas, one has to derive analytical approximations (using perturbation methods or asymptotic analysis). Perturbation methods are very general. But to perform this method, one usually relies on a known proxy, for which the computations are simpler. In Section 2.1, we list the models where the prices of call-put are explicit and that may be used as proxy models. In Section 2.2, we present analytical approximations for general ordinary and stochastic differential equations and their domain of validity. The domain of validity specifies the restrictions of the approximation method. In Section 2.3, we briefly expose the analytical approximations used in finance. The overview of the previous methods gives us a clear idea about their limitations which are the starting point for our work. More specifically, if we use a perturbation based on ergodic properties, the approximation is only valid for long maturities; if we use one based on geodesic properties, the approximation is restricted to short maturities; similarly, if we use a perturbation using the PDE operator, the approximation can be explicitly worked out only for time homogeneous PDE coefficients. These restrictions motivate us to think about a new methodology which can be applied to a wider framework (short or long maturity, time inhomogeneous parameters,...). Our methodology is presented in Section 2.4 and is the subject of this thesis. A comparison of our approach with Watanabe's one is detailed in Section 2.5. The outline and the main results of the thesis are given in Section 2.6.

2.1 Closed-form formulas in finance

Black Scholes model. The Black Scholes paradigm (see [22]) assumes the spot to be a lognormal diffusion with constant volatility. In other words, the spot (S_t) follows the diffusion:

$$\frac{dS_t}{S_t} = (r-q)dt + \sigma dW_t,$$

where W is a Brownian motion, σ is the constant volatility of the spot, r (resp. q) is the deterministic risk free rate (resp. the deterministic dividend yield). This dynamics is written under the risk-neutral measure used for the option valuation.
The price of call and put options in this model has a closed-form formula due to the explicit computation of the cumulative function of the Gaussian variable. The price of the call with starting time t, expiry time T, spot S and strike K has the following well known expression:

$$Call_{BS}(t,S;K,T) = Se^{-q(T-t)}\mathcal{N}(d_1(T-t, Se^{(r-q)(T-t)}, K)) - Ke^{-r(T-t)}\mathcal{N}(d_0(T-t, Se^{(r-q)(T-t)}, K)).$$

where $d_0(t,x,y) = \frac{1}{\sigma\sqrt{t}}\log(\frac{x}{y}) - \frac{\sigma\sqrt{t}}{2}$, $d_1(t,x,y) = \frac{1}{\sigma\sqrt{t}}\log(\frac{x}{y}) + \frac{\sigma\sqrt{t}}{2}$ and $\mathcal{N}$ is the cumulative function of the standard normal distribution. This closed-form formula can be extended easily when the parameters r, q and σ are time dependent (since the spot S is still a lognormal variable).

Remark 2.1.1. *Note that when σ depends only on the spot, there are a few closed-form formulas (see [1]). Moreover, when the volatility σ becomes a separable function on the spot and the time, one can derive an asymptotic expansion for the price of vanilla options (call, put) using singular perturbation techniques as explained by Hagan et al in [62]. However, for general forms of a volatility functions σ depending on the spot and the time, there is no analytical formula for call options. These general forms of volatility are included in local volatility models or models à la Dupire (see [40]).*

Merton model. The Merton model (see [85]) can be seen as an extension of the Black Scholes model with the addition of independent Poisson jumps with the jump size normally distributed:

$$\frac{dS_t}{S_t} = (r - q - \lambda(1 - e^{\eta_J + \frac{\gamma_J^2}{2}}))dt + \sigma dW_t + (e^{J_t} - 1)dN_t,$$

where

- the compound Poisson process (J_t) and the Brownian motion (W_t) are independent,
- $J_t = \sum_{i=1}^{N_t} Y_i$ and Y_i are i.i.d. normal variables with mean η_J and volatility γ_J,
- N_t is a Poisson process with intensity λ.

Conditioning by the number of the jumps N_T, one can express the call price in the Merton model as an infinite sum of Black Scholes prices:

$$Call_{Merton}(t,S;K,T) = \sum_{i=0}^{\infty} \frac{(\lambda(T-t))^i}{i!} e^{-(\lambda+r)T} BSCall\left(F_T e^{i(\eta_J + \frac{\gamma_J^2}{2})}, K, T-t, \sqrt{\sigma^2 + \frac{i\gamma_J^2}{T-t}}\right),$$

where

$$F_T = Se^{(r-q+\lambda(1-\exp(\eta_J+\gamma_J^2/2)))(T-t)},$$

and $BSCall(S,K,T,v)$ is the Black-Scholes price for a call on an underlying S_t with initial condition $S_0 = S$, volatility v, exercised at maturity T and strike K, where the risk-free rate and the dividend yield are set at 0%.

Remark 2.1.2. *The call price in the Merton model has still closed-form formula when the parameters r, q and σ become time dependent. Moreover, when the volatility σ becomes a function of both the spot and the time, we retrieve exactly the definition of the Andersen Andreasen model ([8]). In such a model, there exist numerical methods like Forward PIDE method for the calculus of the call prices (see [8] and [33]). However, there is no analytical formula for the call price in the Andersen Andreasen model.*

CEV model. In the case of the Constant Elasticity of Variance model (known as CEV model), the call (put) price has a closed-form formula. In that case, the spot (S_t) has the following dynamics:

$$\frac{dS_t}{S_t} = (r-q)dt + \nu S_t^{\beta-1} dW_t, S_0 > 0.$$

The CEV model has been originally studied by Cox in [34] for the case $\beta < 1$. The case $\beta > 1$ has been treated after by Emanuel and MacBeth in [42]. The call price in this model can be computed using the complementary non central Chi-square distribution Q:

$$Call_{CEV}(t,S;K,T) = e^{-q(T-t)}Q(2x,n,2y) - e^{-r(T-t)}Q(2y,n-2,2x) \tag{2.1}$$

where

$$n = 2 + \frac{1}{1-\beta},$$
$$x = \frac{(r-q)S^{-2(\beta-1)}}{\nu^2(\beta-1)(e^{2(r-q)(\beta-1)(T-t)}-1)},$$
$$y = \frac{(r-q)K^{-2(\beta-1)}}{\nu^2(\beta-1)(1-e^{-2(r-q)(\beta-1)(T-t)})}.$$

The computation of the non central Chi-square distribution Q can be performed using a recursive algorithm (see Schroder algorithm in [104]) or an integration of Bessel functions.

Remark 2.1.3. *When the parameters r, q and ν are time dependent, the Call price still has closed formula using Lie-algebraic techniques (see [66]). However, when β becomes time dependent, there is no analytical formula to our knowledge.*

Heston model. The Heston model is an extension of the Black Scholes model for the underlying (S_t) but with stochastic volatility:

$$\mathrm{d}X_t = \sqrt{v_t}dW_t - \frac{v_t}{2}\mathrm{d}t,\ X_0 = x_0, \tag{2.2}$$
$$\mathrm{d}v_t = \kappa(\theta - v_t)\mathrm{d}t + \xi\sqrt{v_t}\mathrm{d}B_t,\ v_0 > 0, \tag{2.3}$$
$$\mathrm{d}\langle W,B\rangle_t = \rho\mathrm{d}t,$$

where

- X_t is the logarithm of the forward $e^{(q-r)t}S_t$, r and q are respectively the risk free rate and the dividend yield,
- v_0 is the initial square of volatility,
- κ is the mean reversion parameter,
- θ is the long-term level,
- ξ is the volatility of volatility,
- ρ is the correlation.

The computation of the call-put price in the Heston model can be done using Fourier inversion since the characteristic functions of the logarithm of the underlying is explicit in this framework (the model parameters do not depend on time).

The call price in Heston's model can be written using Lewis' formula[1] ([79]):

$$Call_{Heston}(t,S_t,v_t;T,K) = S_t e^{-q(T-t)} - \frac{Ke^{-(T-t)r}}{2\pi}\int_{\frac{i}{2}-\infty}^{\frac{i}{2}+\infty} e^{-izX}\phi_T(-z)\frac{dz}{z^2-iz}$$

where $X = \log\left(\frac{S_t e^{-(T-t)q}}{Ke^{-(T-t)r}}\right)$ and $\phi_T(z) = \mathbb{E}(e^{z(X_T-X_t)}|\mathscr{F}_t)$.

The characteristic function $\phi_T(z) = \mathbb{E}(e^{z(X_T-X_t)}|\mathscr{F}_t)$ is explicit when the parameters are constant. When the parameters θ, ξ and ρ are piecewise constant, the characteristic function $\phi_T(z) = \mathbb{E}(e^{z(X_T-X_t)}|\mathscr{F}_t)$ is computed recursively using PDE methods (see [86]) or a Markov argument for affine models (see [41]).

[1]For details on the other formula derived in [63], we refer to Part III Subsection 9.1.2

Remark 2.1.4. *When the time-dependency of the Heston parameters becomes more general, there is no analytical formula anymore. Remark also that the computation of the Fourier inversion formula is far to be as quick as the Black Scholes or the Merton formulas. Indeed, this Fourier inversion is time-consuming and suffers from instabilities for large strikes and long maturities (see [68]).*

2.2 General overview on analytical approximations

We have seen that we can derive closed-form formulas for call or put option when the density of the spot is explicit or its characteristic function is explicit. Otherwise, there is no closed formula. Then, one may use perturbations methods. The aim of this section is to give a brief overview of these perturbation methods in general.

2.2.1 Ordinary differential equations

In this subsection, we briefly introduce the so called "perturbations methods" used in the literature especially for the field of singular perturbation problem.

Matched asymptotic expansions. The principle of this method consists in splitting the domain of the boundary value problem into a sequence of two or more subintervals. Often, we distinguish two kinds of solutions: the inner solution and the outer solution. These solutions are named because of for their relationships to the boundary layer; the boundary layer occurs often in the domain boundary and it is the place of non negligible corrections terms for the perturbation method. In each interval, the perturbation theory is used to obtain an asymptotic solution valid only on this interval. The matching is then required in order to combine the outer and inner solutions in such a way that the approximation has the same functional form on each of these intervals. Finally, this gives an approximate solution valid for the entire domain.

Consider the following example borrowed to Chapter 2 in [64]:

$$\varepsilon y^{(2)}(x) + 2y^{(1)}(x) + 2y(x) = 0, \quad for \quad 0 < x < 1,$$

where $y(0) = 0$, $y(1) = 1$ and $0 < \varepsilon << 1$. This kind of problem is not straightforward. If we directly take $\varepsilon = 0$, we retrieve the following problem:

$$2y^{(1)}(x) + 2y(x) = 0, \quad for \quad 0 < x < 1.$$

The solution of this equation has the form $y(x) = Ae^{-x}$ and the boundary values at 0 and 1 gives $A = 0$ and $A = e^1$. Hence, the limit equation of that problem has no solution and makes this problem a singular perturbation problem. One may apply the perturbation technique as follows: find the outer and inner solution of that problem, assume that 0 is a boundary layer[2] and then match the two solutions in order to obtain an accurate solution for the whole domain (for more details see [64]).

The multiple scales method. This method starts from the general solution and differs from the matched expansion method. Indeed, it introduces coordinates for each region; these coordinates are mutually independent. Hence, this method consists in adding new variables to the initial ordinary differential equation problem and transforming it into a partial differential equation one. In physic systems,

[2] This a priori assumption is the only way to find the true location of the boundary layer (see [111])

this scaling method is often used for the time variable while the matching expansion method applies to space variables. The multiple scales method intervenes when the corrections terms of the regular perturbation terms are not negligible and may be unbounded. Hence, using extra variables, we allow the amplitude to vary slowly and we avoid unboundedness of the corrections terms (for details about this expansion, we refer to Chapter 6 in [87]).

The WKB method. In the matching method, we divide the domain of the boundary value problems into subintervals in order to make explicit the dependence of the solution on the boundary layer variable. For the multiple scales method, we use the dependence of a new timescale in order to solve a new partial differential Equation. The WKB is a special case of the multiple scales method and assumes an exponential dependence for the solution of the boundary value problem. This method is used to solve linear differential equation where the higher derivative is multiplied by a small parameter ε. WKB is an acronym for Wentzel-Kramers-Brillouin approximation. The physicists Wentzel, Kramers, and Brillouin developed this theory in the 1920s. This method was used by Jeffreys in order to give accurate approximation for the Schrödinger equation. From a differential equation of the type:

$$\varepsilon y^{(n)}(x) + a_{n-1}y^{(n)}(x) + \cdots + a_1 y^{(1)}(x) + a_0 y(x) = 0,$$

we assume an asymptotic series of the solution in the form of:

$$y(x) = e^{\frac{1}{\mu}\sum_n \mu^n z_n(x)}$$

for μ close to 0. By using this form in the differential equation, we obtain an arbitrary number of differential equations which allow us to get in a recursive manner the solutions z_n (for more details about the method, we refer to Chapter 10 in [16]).

The Homogenization method. Scientists and engineers encounter real problem and deal with materials involving heterogeneous constituents (for instance, we cite laminated plates and bubble fluids). Hence, the scientists do not consider each component separately using its own information, own mass. But, they assume the material to be continuously distributed and use averaged information for material parameters like the mass density. This method can be viewed as an extension of the multiple scales method in two space scales method (for details on this method, we refer to [5] and [64]).

Alternative approaches. We cite among them:

- Multiple scale expansion for partial differential equations: This kind of problems involves especially partial differential problems with space and time variables (for more details see Chapter 6 in [30]).
- Parameters variations and averaging methods : This method deals with time dependent coefficients for ordinary or partial differential Equations. This method assumes that these parameters evolve slowly in time. Hence, one can find equivalent constant parameters which approximates the solution (see Chapter 5 in [87]).
- Strained coordinates. This technique deals with non uniformities on asymptotic expansions. Indeed, it introduces near-identity transformations on the variables of the differential Equation in order to obtain uniform approximate solutions (for more details about this technique we refer to Chapter 3 in [87]).

2.2.2 Stochastic differential equations

Small noise expansion (pathwise). Freidlin and Wentzell ([47]) consider the random dynamical system:

$$\dot{X}_t^\varepsilon = b(X_t^\varepsilon, \varepsilon \xi_t),$$

where ε is a small parameter, b is a continuous function, ξ_t is a given continuous random process and the solution X_t^ε starts from the initial condition x_0. The authors prove that we can expand the solution $(X_t^\varepsilon)_{t\in[0,T]}$ uniformly on $[0,T]$ in powers of ε:

$$x_t + \varepsilon Y_t^1 + \cdots + \varepsilon^n Y_t^n + o(\varepsilon^n),$$

where x_t is the solution of the non perturbed dynamical system and the function $b(x,y)$ has $n+1$ bounded derivatives with respect to x and y (for more details, see Chapter 2 in [47]). We also cite analogous results for the flow in stochastic differential equations (Kunita [72] for diffusions, Fujiwara and Kunita [48] for jump diffusions).

Small parameter expansion (distribution). Watanabe in [112] shows that, for every random variable F^ε smooth in ε and in the Malliavin sense with Malliavin covariance matrix $\gamma(F^\varepsilon)$ which is invertible with integrable inverse around $\varepsilon = 0$ and for every bounded function f, there exists a weak Taylor approximation of any order $n \geq 0$:

$$\mathbb{E}[f(F^\varepsilon)] = \mathbb{E}[f(F_0) + \varepsilon G^1 + \cdots + \varepsilon^n G^n] + o(\varepsilon^n),$$

where the corrections G_i verify the Equality

$$\mathbb{E}[G^i] = \mathbb{E}[f(F^0)\pi_i],$$

and π_i are explicit weights provided that we know the Malliavin matrix $\gamma(F^0)$ and its inverse.

However, in practice, the question of identifying a relevant small parameter is crucial and not simple. There are usually many ways to define a parametrisation for the model and the resulting approximations may be very different. We give two examples.

- small noise:

$$dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dW_t. \tag{2.4}$$

- small time

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t. \tag{2.5}$$

  We are interested in the law of X_t for t small. By a space-time scaling, $(X_{t\varepsilon})_{t\geq 0}$ has the same law than $(X_t^\varepsilon)_{t\geq 0}$ defined by

$$dX_t^\varepsilon = \varepsilon b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dW_t.$$

For more details on the difference of our approach with Watanabe's one, we refer to Section 2.5.

Large deviation theory. Large deviations theory provides tail estimates of certain distributions (see [110] for details). Freidlin and Wentzell in [47] derive a large deviations Theorem for ordinary differential equations perturbed by a Brownian Motion (small noise perturbation). They consider the following SDE:

$$dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}dW_t, X_0^\varepsilon = x_0,$$

where W is multidimensional Brownian motion with dimension $d \in \mathbb{N}^*$ and the function $b : \mathbb{R}^d \to \mathbb{R}^d$ is bounded and Lipschitz as well. If C and G are respectively closed and open subsets in the set $\{f : [0,T] \to \mathbb{R}^d, f \text{ is continuous}, \dot{f} \text{ is a square integrable function}, f(0) = x_0\}$, then one obtains:

$$\limsup_{\varepsilon\downarrow 0} \varepsilon \log \mathbb{P}(X^\varepsilon \in C) \leq -\inf_{f\in C} I(f),$$
$$\liminf_{\varepsilon\downarrow 0} \varepsilon \log \mathbb{P}(X^\varepsilon \in G) \geq -\inf_{f\in G} I(f),$$

where $I(f) = \frac{1}{2}\int_0^T |\dot{f}_t - b(f_t)|^2 dt$. They also extended the result to the stochastic differential equation:

$$dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dW_t, X_0^\varepsilon = x_0,$$

where $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ is a bounded and Lipschitz function (for more details see Section 6 in [110]).

Remark 2.2.1. *The theory introduced by Freidlin and Wentzell is interesting since it allows to emphasize the asymptotic behavior for the distributions of stochastic differential Equations. However, it gives only logarithmic estimates.*

Parameter sensitivity. Gobet and Munos in ([55]) consider a multidimensional diffusion process $(X_t^\alpha)_t$ where the dynamics depend on a parameter α. They express the sensitivity $\partial_\alpha \mathbb{E}[f(X_t^\alpha)]$ for a bounded measurable function f as an expectation involving the function f and not its derivatives. This kind of problem arises in stochastic control problems. The authors use three different methods: Malliavin calculus, adjoint approach and martingale approach. Numerical results with associated errors for the corresponding discretization schemes are also given.

Markov projection. Gyöngi in [59] considers a stochastic process (ξ_t) starting from 0 with the Itô form:

$$d\xi_t = \delta(t,\omega)dW_t + \beta(t,\omega)dt,$$

where W is an $\mathscr{F}_t$-adapted Brownian Motion, δ and β are bounded and $\mathscr{F}_t$-adapted process with $\delta\delta^*$ being uniform positive definite. He shows that there exists a Markovian process (X_t) which has the same one-dimensional marginal as (ξ_t) (i.e. $\mathscr{L}(X_t) = \mathscr{L}(\xi_t) \quad \forall t$), and which is a weak solution to the following stochastic differential equation:

$$dX_t = \sigma(t,X_t)dW_t + \beta(t,X_t)dt, X_0 = 0,$$

where

$$\sigma\sigma^*(t,x) = \mathbb{E}[\delta\delta^*(t)|\xi_t = x], b(t,x) = \mathbb{E}[\beta(t)|\xi_t = x].$$

Brunick ([28]) relaxes the assumption of ellipticity on δ.

Remark 2.2.2. *This Markov projection method was the key for many approximations especially in finance like computing equivalent local volatility functions for stochastic volatility models (see [94] and [76]). However, no error estimates for these approximations are available.*

2.3 Analytical approximations applied to finance

In the previous Section, we have given a general overview about the perturbation methods. Here, we focus on the literature related to the perturbation methods used in finance. This section aims at presenting many original trends of used perturbation methods with their limitations.

2.3.1 Perturbation methods

In this field, we cite:

- Price expansion w.r.t. correlation between the spot and the volatility. Antonelli and Scarlatti consider in [12] a general stochastic model and derive a series expansion for the Call price w.r.t. correlation; Their approximations includes Heston model as a particular case. Each correction term is also approximated with estimates of the error. Moreover, the error bounds of the series expansion w.r.t. correlation are available; the approximation is available only for time homogeneous parameters.
- Small volatility of volatility expansion. Lewis in [79] derives a call price expansion w.r.t. the volatility of volatility. His work is based on formal expansion of the PDE operator for Fourier transform of the Call price. This approximation handles general stochastic volatility models like Heston, generalized Heston and GARCH models. As a consequence, he derives accurate formulas for implied Black Scholes in these stochastic volatility models. The approximation is available only for constant parameters.
- Averaging technique. The averaging technique, introduced by Piterbarg [93] has emerged as an important technique. It can be viewed as an application in finance of the Markov projection theorem derived by Gyöngi. Piterbarg derives averaging constant Heston parameters for time dependent Heston model to approximate the call price. Indeed, he derives in the field of interest rate derivatives, formulas for "effective" skew and stochastic volatility which are related to the time dependent parameters. The application of the formula is also valid for Equity or FX derivatives. The approximation is derived only for zero correlation.

2.3.2 Asymptotic analysis

Extreme strikes. Lee shows in [78] that the implied variance is bounded from above by a function linear w.r.t. the log moneyness $\ln(\frac{K}{F})$ for large strikes. He gives explicit formulas which relate the gradients of the wings of the upper bound of the implied variance and the maximal finite moments of the spot. For instance, for the left wing, he shows that if $q^* := \sup\{q : \mathbb{E}[S_T^{-q}] < \infty\}$ and

$$\beta^* := \limsup_{K \to 0^+} \frac{\sigma_{imp}^2(T,K)T}{|\ln(\frac{K}{F})|},$$

then $\beta^* \in [0,2]$ and

$$q^* = \frac{1}{2}(\frac{1}{\sqrt{\beta^*}} - \frac{\sqrt{\beta^*}}{2})^2.$$

Moreover, Benaim and Friz in [15] sharpen Lee's formula. Indeed, they show that Lee's upper bound may become a limit under some technical conditions which are satisfied for a large class of models.

Long maturities. We cite the work of Tehranchi in [108] that gives asymptotic formula for implied Black Scholes volatilities far from maturity with precise control of the error. Assuming that, almost surely $S_t \to 0$ as $t \uparrow \infty$, Tehranchi shows the following formula holds:

$$T\sigma^2_{imp}(T,K) = 8|\ln(\mathbb{E}(S_T \wedge \frac{K}{S_0}))| - 4\ln(|\ln(\mathbb{E}(S_T \wedge \frac{K}{S_0}))|) + 4\ln(\frac{K}{S_0}) - 4\ln(\pi) + \varepsilon(\ln(\frac{K}{S_0}), T),$$

where

$$\sup_{-M \le x \le M} |\varepsilon(x,T)| \underset{T\to\infty}{\to} 0$$

for all $M > 0$.

We also mention Gatheral's work in [49] who derives arbitrage bounds on the skew of the implied B-S volatility at-the-money. These bounds are of the order of $O(T^{-\frac{1}{2}})$.

Moreover, Rogers and Tehranchi prove in [99] the smile Theorem conjectured by S. Ross. Indeed, they prove that the smile shape can not move by parallel shifts.

2.3.3 Combining perturbation methods and asymptotic analysis

Ergodic approach. Fouque et al in [44] consider a stochastic volatility model where the volatility is a functional of Y^ε_t which is an Ornstein Uhlenbeck process (we denote it OU) with small correlation time ε:

$$\begin{aligned} \frac{dS^\varepsilon_t}{S^\varepsilon_t} &= rdt + f(Y^\varepsilon_t)dW_t, \\ dY^\varepsilon_t &= \frac{1}{\varepsilon}(m - Y^\varepsilon_t)dt + \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}}dB_t, \\ d\langle W,B\rangle_t &= \rho dt. \end{aligned}$$

The normalisation w.r.t. ε is such that the asymptotic distribution of Y^ε_t as t goes to infinity is a Gaussian law $\mathcal{N}(m,\nu^2)$.

They write the infinitesimal generator $\mathcal{L}^\varepsilon$ as the summation of three terms

$$\mathcal{L}^\varepsilon = \frac{1}{\varepsilon}\mathcal{L}^{(0)} + \frac{1}{\sqrt{\varepsilon}}\mathcal{L}^{(1)} + \mathcal{L}^{(2)}, \tag{2.6}$$

where

- $\mathcal{L}^{(0)} = \nu^2 \frac{\partial^2}{\partial y^2} + (m-y)\frac{\partial}{\partial y}$ is the infinitesimal generator of the OU process Y defined by

$$dY_t = (m - Y_t)dt + \nu\sqrt{2}dB_t, \tag{2.7}$$

- $\mathcal{L}^{(1)} = \sqrt{2}\rho x f(y)\frac{\partial^2}{\partial x \partial y}$ contains the mixed derivatives due to the correlation term,
- $\mathcal{L}^{(2)} = \frac{\partial}{\partial t} + \frac{1}{2}f(y)^2x^2\frac{\partial^2}{\partial x^2} + r(x\frac{\partial}{\partial x} - .)$ is the Black-Scholes operator with volatility $f(y)$.

Fouque et al suppose that the parameter ε is small which makes the problem (2.6) to be a singular perturbation problem. They expand the call price with starting time t, maturity T, spot S and strike K, in power of $\sqrt{\varepsilon}$:

$$\text{Call}^{\varepsilon} = \text{Call}_{BS}^{(0)} + \sqrt{\varepsilon}\text{Correction}^{(1)} + \cdots$$

where the leading term is the Black Scholes price $\text{Call}_{BS}^{(0)} = Call_{BS}(t,S,K,T,\bar{\sigma})$ and the variance $\bar{\sigma}^2$ is the averaging of the function f with respect to the invariant distribution $\mathcal{N}(m,\nu^2)$ of the OU (Y) defined in Equation (2.7):

$$\bar{\sigma}^2 = \frac{1}{\nu\sqrt{2\pi}} \int_{\mathbb{R}} f^2(y) e^{\frac{-(m-y)^2}{2\nu^2}} dy \equiv \langle f^2 \rangle, \tag{2.8}$$

and the correction term is a combination of Greeks of the leading term $\text{Call}_{BS}^{(0)}$:

$$\sqrt{\varepsilon}\text{Correction}^{(1)} = -(T-t)(V_2 S^2 \frac{\partial^2 Call_{BS}(t,S,K,T,\bar{\sigma})}{\partial S^2} + V_3 S^3 \frac{\partial^3 Call_{BS}(t,S,K,T,\bar{\sigma})}{\partial S^3}),$$

where the coefficients V_2 and V_3 are computed like the volatility $\bar{\sigma}$ in Equation 2.8 using the operator $\langle . \rangle$:

$$V_2 = \sqrt{2}\rho\nu\langle f\phi' \rangle,$$
$$V_3 = \frac{\rho\nu}{\sqrt{2}}\langle f\phi' \rangle,$$

and ϕ is a solution of the Poisson equation:

$$\mathcal{L}^{(0)}\phi(y) = f(y)^2 - \langle f^2 \rangle.$$

Moreover, the authors show in [45] that the error of the approximation for call(put) option behaves like:

$$\lim_{\varepsilon \downarrow 0} \frac{|\text{Call}^{\varepsilon} - \text{Call}_{BS}^{(0)} - \sqrt{\varepsilon}\text{Correction}^{(1)}|}{\varepsilon |\ln(\varepsilon)|^{1+p}} = 0,$$

for any $p > 0$.

Remark 2.3.1. *This approximation is interesting since it gives fast analytical formula of the price as a combination of a leading Black Scholes price and some related Greeks. Moreover, the error bounds are given for call (put) options. But, this approximation is not valid for small maturities since it relies on the mean reverting property of the volatility, a behavior which is not instantaneous. Moreover, it is restricted to homogeneous parameters in order to get explicit solutions to some PDE.*

Geodesic approach. There are many interesting papers and books about the geodesic approach and its applications in finance:

- see Chavel ([29]) for an introduction about Riemannian geometry and Varadhan ([109]) for an asymptotic of the density for small time,
- see Berestycki et al [21], Labordere [74] and [75], Lewis [80], Forde [43], Benhamou et al [17] for the application of the Riemannian geometry in finance.

These works show that the Green function π (the pricing density) which solves:

$$\frac{\partial \pi}{\partial t} = \sum_{i,j} g^{i,j} \frac{\partial^2 \pi}{\partial S^i \partial S^j} + \sum_i h_i \frac{\partial \pi}{\partial S^i},$$

with initial condition $\pi(t_0, S_0, t_0, S) = \delta(S_0 = S)$, has the short maturity expansion:

$$\pi(t_0, S_0, t, S) = e^{\frac{-d^2(S_0,S)}{2(t-t_0)}} \left(G_0(S_0,S) + (t-t_0)G_1(S_0,S) + \cdots\right)$$

where $d(S_0,S)$ is the geodesic distance associated with the Riemannian space defined by the metric distance $ds^2 = \sum_{i,j} g_{i,j} dx^i dx^j$, where $g_{i,j}$ is the inverse of the matrix $g^{i,j}$. The term G_0 is related to the Gaussian distribution. Indeed, the pricing density behaves like a Gaussian one when we use these new geodesic variables. The term G_1 is related to the curvature of the Riemannian space.

Moreover, Beresticky et al (see Th 1.2 in [21]) show in a general stochastic volatility model that the implied Black Scholes volatility near the expiry is:

$$\lim_{T \downarrow t} \sigma_{imp}(T,K) = \frac{\ln(\frac{S}{K})}{d(\ln(\frac{S}{K}), y_0)} \tag{2.9}$$

where t the starting time, T the maturity, S the spot, K the strike, y_0 the value of the initial stochastic volatility and d is the signed geodesic distance that solves the Eikonal Equation (1.12) in Theorem 1.2 in [21]. Notice also that Beresticky et al in [21] derive corrections terms for the short time expansion of the implied Black Scholes for some cases.

Remark 2.3.2. *This geodesic approach is very interesting since it gives closed-form formulas for the implied Black Scholes volatility near the expiry*[3]*. Moreover, one can derive asymptotic expansion of the implied Black Scholes for short maturities. However, these kinds of approximation are restricted to short maturities and homogeneous parameters.*

2.4 Motivation and methodology

Through the previous Section, we have seen many trends of original ideas for exact computation or fast analytical methods. Note that the closed formulas cover a very small class of models: Black Scholes, Merton, CEV, Heston,... For general models, one can perform approximation formulas. However, these kinds of analytical formulas have been performed within some restriction like short maturity (Geodesic expansion), long maturity (Ergodic approach), extreme strikes (Lee's approach), zero correlation (Piterbarg's approach) or time homogeneous parameters (all the previous perturbation methods except Piterbarg's one). Our aim in this thesis is to design an accurate expansion method that applies to both short and long maturities, to both small and large strikes, to time inhomogeneous parameters and to non null correlations as well.

To ensure that a perturbation method can be efficiently used in practice, the quantity of interest (i.e. the price) should be decomposed as a summation of explicit terms, namely a principal part plus some correction terms. The main idea is that the principal part is given by the price in a proxy model. The distance to the proxy model is represented by a small parameter Δ. For instance,

[3]Remark also that the analytical computation of the implied Black Scholes volatility is not always explicit like in the case of Heston's model.

- in Heston model, a possible proxy could be a Black Scholes model by taking the vol of vol ξ equal to 0. In that case, we could define $\Delta = \xi$.
- in local volatility models, the Black Scholes model could be a proxy by freezing the local volatility function σ at its initial value. In that case, Δ could be the sup norm of the derivatives of σ.

Actually, this analysis of the distance to the proxy is too rough and in the thesis, we will provide much more tight measures of this distance to proxy.

Now, if we denote $(X_t)_t$ the true model for the underlying asset price (or log-asset price) and by $(X_t^P)_t$ the proxy model, at least formally we can write a Taylor expansion for the price of a vanilla option h on the asset X maturing at time T:

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] + \mathbb{E}[h^{(1)}(X_T^P)(X_T - X_T^P)] + \cdots + \mathbb{E}[h^{(j)}(X_T^P)\frac{(X_T - X_T^P)^j}{j!}] + \mathscr{R}esid_j. \quad (2.10)$$

The leading term $\mathbb{E}[h(X_T^P)]$ is explicit, because usually the prices in the proxy model are explicit (actually, it is somehow a constraint in the choice of our proxy model). The second term $\mathbb{E}[h^{(1)}(X_T^P)(X_T - X_T^P)]$ plays the role of the first correction term, but it is not explicit for general processes and has to be approximated. The trick is to decompose this term as a summation of Greeks (in the proxy model) plus a residual term:

$$\mathbb{E}[h^{(1)}(X_T^P)(X_T - X_T^P)] = \sum_{i=1}^{n_1} a_i^{(1)} \partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0} + O(\Delta^k). \quad (2.11)$$

The Greeks in the proxy model have to be explicit as well. The derivation of this decomposition is strongly model-dependent and we derive it using Malliavin calculus combined with a smart parameterization of the model for X. Since this step is very model-specific, it is not possible to write a general theory and we refer for the details to the introduction of each part of the thesis. For the convenience of the reader, we shall explain why Greeks appear naturally in (2.11). The identification using Greeks can be seen as an inverse procedure used in the literature about integration by parts formula and Malliavin calculus ([46]). Indeed, we know that $\partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0} = \mathbb{E}[h^{(i)}(X_T^P)] = \mathbb{E}[h(X_T^P)H_T^i]$ for some "Malliavin weights" H_T^i and to get (2.11), we should identify explicit $(a_1)_i$ such that $\sum_{i=1}^n a_i^{(1)} H_T^i \approx (X_T - X_T^P)$. Then, we repeat the decomposition (2.11) for each term of the type $\mathbb{E}[h^{(j)}(X_T^P)\frac{(X_T - X_T^P)^j}{j!}]$, by writing $\mathbb{E}[h^{(j)}(X_T^P)\frac{(X_T - X_T^P)^j}{j!}] = \sum_{i=1}^{n_j} a_i^{(j)} \partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0} + O(\Delta^k)$. Finally, we bring together all these contributions in (2.10) up to the minimal order j_k verifying $\mathscr{R}esid_{j_k} = O(\Delta^k)$. This gives

$$\begin{aligned}\mathbb{E}[h(X_T)] &= \mathbb{E}[h(X_T^P)] + \sum_{j=1}^{j_k}[\sum_{i=1}^{n_j} a_i^{(j)} \partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0}] + O(\Delta^k) \\ &= \mathbb{E}[h(X_T^P)] + \sum_{i=1}^{max_{j \le j_k} n_j} (\sum_{j=1}^{j_k} a_i^{(j)} \mathbb{1}_{i \le n_j}) \partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0} + O(\Delta^k).\end{aligned}$$

The general form is the following:

$$\boxed{\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] + \textit{ weighted sum of Greeks } \partial_x^i \mathbb{E}[h(X_T^P + x)]|_{x=0} + \textit{ error.}} \quad (2.12)$$

Of course, these formal arguments need to be mathematically clarified with many respects:

- the payoff h is not smooth (the second derivative of a call payoff does not exist in the classical sense, which avoids to write directly (2.10)).
- one has to carefully use Malliavin calculus to derive explicit expansions for the coefficients $(a_j^{(i)})$.
- the estimate of error terms is a very difficult task and actually, it depends on the model and the payoff.

2.5 Comparison with Watanabe's approach

It is worth emphasizing that the Watanabe approach ([112]) using an asymptotic expansion of Wiener functionals is an important reference and there exist similarities and significant differences between Watanabe's approach and the one we introduce in this thesis. In his works, Watanabe considers a family $(F^\varepsilon)_{\varepsilon\geq 0}$ of random variables defined on the Wiener space and smooth in the Malliavin sense. Below, we follow his notation. Suppose that we can write an asymptotic expansion of F^ε in powers of ε at any order k:

$$F^\varepsilon - (f_0 + \varepsilon f_1 + \cdots + \varepsilon^k f_k) = O(\varepsilon^{k+1}) \text{ as } \varepsilon \to 0^+ \tag{2.13}$$

The random variables $(f_i)_i$ are smooth in the Malliavin sense and the above equality with Landau symbol has to be understood in the Sobolev norms $\|\cdot\|_{\mathbb{D}^{l,p}}$. Assume additionally a uniform non degeneracy condition:

$$\limsup_{\varepsilon\to 0^+} \|\det(\gamma_{F^\varepsilon}^{-1})\|_p < \infty \text{ for all } p \geq 1 \tag{2.14}$$

where γ_{F^ε} is the Malliavin covariance matrix of F^ε. Then, Watanabe shows that, for any function h with polynomial growth, we have:

$$\mathbb{E}[h(F^\varepsilon)] = \mathbb{E}[h(F^0)] + \varepsilon\mathbb{E}[h(F^0)\pi_1] + \cdots + \varepsilon^k\mathbb{E}[h(F^0)\pi_k] + O(\varepsilon^{k+1}) \tag{2.15}$$

where $(\pi_i)_i$ is a sequence of random variables. This equality can be extended to distributions h (Theorem 2.3 in [112]).

At first sight, (2.12) and (2.15) are of the same type by taking $X_T = F^\varepsilon$ and $F^0 = X_T^P$. Within this analogy, we may try to relate the Greeks in (2.12) with the terms $(\mathbb{E}[h(F^0)\pi_i])_i$ in (2.15), but actually this identification is not straightforward at all.

An important feature of Watanabe's approach is that the expansion accuracy is written in terms of the small parameter ε. The parameterization w.r.t. ε is crucial (see previous examples in Equations 2.4 and 2.5). Last but not least, the impact of other model parameters do not enter in the estimates. And this point is a significant drawback of such computations. To illustrate this, consider the toy model

$$F^\varepsilon = \sigma W_1 + \sqrt{\varepsilon} B_1$$

where (W,B) is a two-dimensional Brownian motions, and σ is positive. Let us expand $\mathbb{E}[h(F^\varepsilon)]$ in powers of ε for $h(x) = x^2$ and $h(x) = x^+$. We use that F^ε is distributed as a $\mathcal{N}(0, \sigma^2 + \varepsilon)$. Clearly $F^0 = \sigma W_1$.

1. Case $h(x) = x^2$. We have

$$\begin{aligned}\mathbb{E}[h(F^\varepsilon)] = \mathbb{E}[(F^\varepsilon)^2] &= \sigma^2 + \varepsilon \\ &= E[h(F^0)] + \varepsilon.\end{aligned}$$

2. Case $h(x) = x^+$. By a scaling argument, we have:

$$\begin{aligned}
\mathbb{E}[h(F^\varepsilon)] &= \sqrt{\sigma^2+\varepsilon}\,\mathbb{E}[(W_1)^+] \\
&= \sqrt{\frac{\sigma^2+\varepsilon}{\sigma^2}}\,\mathbb{E}[(\sigma W_1)^+] \\
&= E[h(F^0)] + \frac{1}{2}\frac{\varepsilon}{\sigma}\mathbb{E}[(W_1)^+] + O(\frac{\varepsilon^2}{\sigma^3}).
\end{aligned}$$

These computations are coherent with the Watanabe expansion results:

$$\mathbb{E}[h(F^\varepsilon)] = E[h(F^0)] + \varepsilon c_1 + O(\varepsilon^2).$$

But we see that the coefficient c_1 depends strongly on the model and on the function h (in the second case, $\varepsilon c_1 = \frac{1}{2}\frac{\varepsilon}{\sigma}\mathbb{E}[(W_1)^+]$). In case where σ is small as well, we see that the expansion accuracy is heavily related to ratio of $\frac{\varepsilon}{\sigma}$ and thus not only to ε. This is another strong argument against a direct application of the Watanabe asymptotic expansion. In our works, we provide non-asymptotic estimates, which crucially enables us to deduce the domain of validity of our formulas regarding all the model parameters. In other words, a significant part of our work emphasises the impact of model parameters on the approximation. This is confirmed by numerous numerical experiments (see Chapters 4, 7, 10, 12). We also show that the magnitude of the error is impacted by the payoff smoothness. The previous toy example is a convincing illustration of this phenomenon, whereas within Watanabe's approach, the payoff regularity does not play any role in the estimates.

We now discuss in more details the differences in the proofs and methodologies.

1. In our approach, we quantify the error according to the payoff smoothness: We perform a Taylor expansion for smooth options as we show before. Then, we use a regularisation method and integration by parts (Malliavin calculus) in order to upper bound the errors and express the truncated terms as a combination of Greeks of the first term.

2. As a difference, Watanabe gives asymptotic expansions of the density of the process by expressing the density as an expectation of a Dirac distribution. In Theorem 2.3 in [112] for the pull-back of distributions by Wiener functionals, he uses the integration by parts formula (Malliavin calculus) applied to the distributions in order to retrieve smooth test functions. Then, he uses Taylor expansions in order to express the corrections as a function of the distribution derivatives. We guess that the expansion of Malliavin integration by parts formula is less tractable compared to our direct approach. In other words, in view of having closed formulas, it is easier to "expand and integrate by parts" than "integrate by parts and expand".

3. As a common fact within the two approaches, it is both assumed that the proxy model (X_T^P or F^0) is related to a Gaussian process/ random variable. Actually, in our case, we also use log-normal proxys.

4. Another technical difference lies in the assumptions used for the expansions. In Watanabe results, $\mathscr{C}^\infty$ smoothness of the model coefficients is required. In our framework, we assume optimal regularity assumptions (for instance, see Assumption (R_4) in Section 4.1). We also handle non smooth cases like Heston model (because of the square root in the diffusion coefficient).

Watanabe approach in finance. Yoshida in [113] [114] , Kunitomo and Takahashi in [73] applied Watanabe's approach in finance and used also Malliavin calculus for the control of the error. Actually, Yoshida weakens the non degeneracy Assumption (2.14) in a localised version (allowing the degeneracy on a set of exponentially small measure); see Theorem 4.1 page 152 in [113]. Teichmann and Siopacha in [106] applied also Watanabe's result for Libor market model using Monte Carlo simulations. They obtain easily tractable formulas for accurate pricing.

5. Regarding the computations of the correction terms, we directly use the inverse of the integration by parts formula (as we explained before) and obtain that the correction terms are a linear combination of Greeks in the proxy model with coefficients depending only on the model parameters. Yoshida (Theorem 2.1 in [113], Theorem 4.1 in [114]) and Takahashi (Theorem 3.3 in [73]) used explicit computations related to conditional expectations for Gaussian vectors and expressed the correction terms as an integral of a products of Gaussian density and polynomials. In other way, the density of interest is approximated by a linear combination of derivatives of the Gaussian density, showing strong similarities between different approaches. Remark also that within our approach, we can handle Poisson process with Gaussian jumps as well, using a suitable and model-adapted parameterization (see Part I).

2.6 Structure of thesis and main results

The thesis is divided into four parts. Each part deals with a specific model and contain three Chapters:

- the first Chapter gives an introduction about the model studied. Moreover, it presents the existing numerical and analytical methods used for the pricing of European options in such a model. Finally, it provides the motivation of such works and the main results of the related part.
- the second Chapter contains all the results of the current part and the technical proofs of these theorems.
- the third Chapter deals with numerical results like calibration, robustness procedure or accuracy of the formula for extreme strikes and long maturities.

Part I. In this part, we consider the Andersen Andreasen model ([8]) which is a local volatility model plus Gaussian jumps:

$$dX_t = \sigma(t,X_{t^-})dW_t + \mu(t,X_{t^-})dt + dJ_t, X_0 = x_0, \tag{2.16}$$

where

- X_t is the logarithm of the forward $F_t = S_t e^{\int_0^t (q_s - r_s)ds}$, r is the deterministic risk free rate and q is the deterministic dividend yield.
- the compound Poisson process (J_t) and the Brownian motion (W_t) are independent.
- $J_t = \sum_{i=1}^{N_t} Y_i$ and Y_i are i.i.d. normal variables with mean η_J and volatility γ_J.
- N_t is a Poisson process with intensity λ.
- $\mu(t,x) = -\frac{\sigma^2(t,x)}{2} + \lambda(1 - e^{\eta_J + \frac{\gamma_J^2}{2}})$ in order to ensure martingale properties for (e^{X_t}).

Our aim in this part is to

- give an **accurate analytic approximation** for the European option:

$$\mathbb{E}[h(X_T)]$$

  where h is a **non smooth function** and T is the maturity of the option. For the case of call option without risk free rate and dividend yield, one has $h(x) = (e^x - K)^+$.

- Calibrate within a **computational time** smaller than one second (**1sec**) all the model parameters: the jump parameters λ, η_J, γ_J and the volatility function $\sigma(t,x)$.

We show that a good proxy for the model will be the following Merton model:

$$dX_t^M = \sigma(t,x_0)dW_t + \mu(t,x_0)dt + dJ_t, X_0^M = x_0,$$

and that the vanilla price in such a model is:

$$E[h(X_T)] = Merton\ price + corrections\ terms + errors, \tag{2.17}$$

where

- The correction terms in Equation (2.17) are a combination of Greeks (Delta, Gamma, Epsilon) of the Merton Model (keep in mind that Greeks have closed formula in Merton's model). These corrections terms are given explicitly in Chapter 4 in Theorem 4.2.1.

- The error term *errors* in Equation (3.8) is estimated for vanilla payoff in Chapter 4 in Theorem 4.5.2. A naive interpretation of the error is the following: if the volatility derivatives σ and the jump size involved in the SDE (3.5) are of the order of Δ, then the error term is of the order of $(\Delta\sqrt{T})^3$.

Note also that the approximation derived in Chapter 4 covers not only call-put options but also all vanilla payoffs depending on X_T (smooth, vanilla, binary). Indeed, the errors are analyzed according to the payoff smoothness (see for instance Theorem 4.5.1 for smooth payoff in general, Theorem 4.5.2 for vanilla options, Theorem 4.5.3 for binary options). We show also in Chapter 3 that the accuracy of our formula turns out to be excellent (the errors for implied B-S volatilities are smaller than 2 bp for various strikes and maturities). As a consequence, the calibration of such model becomes very fast.

Chapter 3 gives an overview about the different models used in order to manage the smile and one introduces the Andersen Andreasen model. Moreover, in this Chapter, we gives details about the Merton model, the calibration using the forward PIDE and we state finally the main results of the Part. Chapter 4 in Part I is exactly the article "smart expansion and fast calibration for jump diffusions" published in the journal "Finance and Stochastics", volume 13(4), pages 563:589, 2009. Chapter 5 in Part I provides additional numerical results concerning the accuracy of the approximation formula for bumped parameters and calibration/optimization robustness as well.

Part II. Now, we look at local volatility models ([40]):

$$dX_t = \sigma(t,X_t)dW_t + \mu(t,X_t)dt, X_0 = x_0, \tag{2.18}$$

where X_t may be the asset of the log asset.

The aim of this part is to **derive explicit closed formula for vanilla options for general forms of local volatility functions**. This closed formula is a Taylor expansion and can be truncated easily at any order. Therefore, the expected price can be written at any order as a summation of:

- The Black Scholes price with at the money volatility. As in Part I, this model can be seen as the proxy of the local volatility model. The advantage of this proxy lies in the explicit calculus of the prices and the Greeks of vanilla options.
- A combination of the Greeks of the leading Black Scholes price with explicit weights depending on the volatility, the drift functions and their derivatives.
- A residual error with explicit upper bounds.

This is achieved in Chapter 7. The approximation for vanilla options at the second order is computed in Theorem 7.2.1 which is a particular case of Theorem 4.2.1 of Chapter 4 when there is no jump. Moreover, the explicit calculus of the approximation for vanilla options at the third order is derived in Theorem 7.2.2. In addition, the corrections and the error terms at any order of the closed formula are estimated in Theorems 7.4.1-7.4.2-7.4.3 according to the payoff smoothness (smooth, vanilla, digital). The accuracy of the expansion turns out to be excellent. Moreover, we need only few terms to give accurate results for vanilla options. As a consequence of our methodology, we derive averaging parameters for time dependent CEV models.

Chapter 6 gives an overview about local volatility model, Dupire's Formula, CEV model and existing analytical approximation methods used for the pricing in such model. Moreover, we detail also in this Chapter the motivations and main results of the Part II. Chapter 7 is exactly the article "Closed forms for European options in a local volatility model" accepted for publication in the journal "International Journal of Theoretical and Applied Finance". The Chapter 8 presents smile behaviours for the CEV model when varying its parameters through the time. We also provide numerical results concerning the accuracy of the approximation formula for large strikes and concerning the domain of arbitrage of the approximation formulas.

Part III. This parts deals with the time dependent Heston model ([63]):

$$\mathrm{d}X_t = \sqrt{v_t} dW_t - \frac{v_t}{2}\mathrm{d}t,\ X_0 = x_0, \tag{2.19}$$

$$\mathrm{d}v_t = \kappa(\theta_t - v_t)\mathrm{d}t + \xi_t\sqrt{v_t}\mathrm{d}B_t,\ v_0 > 0, \tag{2.20}$$

$$\mathrm{d}\langle W,B\rangle_t = \rho_t \mathrm{d}t,$$

where

- X_t is the logarithm of the forward $e^{(q-r)t}S_t$, r and q are respectively the risk free rate and the dividend yield,
- v_0 is the initial square of volatility,
- κ is the mean reversion parameter,
- θ is the long-term level,
- ξ is the vol of vol,
- ρ is the correlation.

Our aim in this part is to

- give an **accurate analytic approximation** for the price of call-put option

$$e^{-rT}\mathbb{E}[(K-e^{(r-q)T+X_T})_+], \tag{2.21}$$

- work in a **very general time dependent Heston framework** to obtain an approximation which can be valid for both short and long maturities and can handle time dependent parameters and non null correlation as well,

- achieve a **computational time cheaper than Fourier inversion (gain by a factor 100 or more)**.

This is achieved in the Chapter 10 using an accurate expansion on small volatility on volatility combined with Malliavin calculus techniques and some technical Lemmas.

The proxy of this model is the Heston model without volatility of volatility. In other words, it is the Black Scholes model with time dependent volatility:

$$\begin{aligned} \mathrm{d}X_t^{BS} &= \sqrt{v_{0,t}}dW_t - \frac{v_{0,t}}{2}\mathrm{d}t,\ X_0^{BS} = x_0, \\ \mathrm{d}v_{0,t} &= \kappa(\theta_t - v_{0,t})\mathrm{d}t, v_0. \end{aligned}$$

We prove that:

$$e^{-rT}\mathbb{E}[(K-e^{(r-q)T+X_T})_+] = Put\ \ Price\ \ in\ \ BS\ \ model + Correction\ \ terms + Errors.$$

- The corrections terms are a linear combination of Greeks of the leading Black Scholes price with explicit weights depending only on the model parameters. These computations are done using Malliavin calculus in Theorem 10.2.1.

- The errors are estimated in Theorem 10.2.2 by:

$$Errors = O(|\xi|_\infty^3 T^2).$$

  A possible choice of the measure Δ here will be the sup norm of the time dependent vol of vol function ξ.

From the approximative formula, we also derive some corollaries related first to equivalent Heston models (extending some work of Piterbarg on stochastic volatility models [93]) and second, to the calibration procedure in terms of ill-posed problems (see Subsection 10.2.6).

The Chapter 9 gives an overview about the Fourier inversion formulas, the analytical approximations in such a model. Moreover, we state the main results in this Chapter. The Chapter 10 is exactly the article "Time dependent Heston model" in revision for the journal "SIAM Journal on Financial Mathematics". The Chapter 11 provides numerical results concerning the smile behavior for Heston's model with constant and time dependent parameters as well. Moreover, in this Chapter, we review some results concerning the negative moments of the integrated CIR process.

Part IV. In this Part, we focus on hybrid models composed by the dynamics of the spot and the stochastic rate as well:

$$\frac{dS_t}{S_t} = r_t dt + \sigma_t dW_t,$$

where (S_t) is the spot, σ is the random volatility and (r_t) is the stochastic rate which follows the HJM framework.

The aim of this Part is to derive new approximation formulas for European options on a local volatility model with stochastic rates. Hence, we model the discounted spot. This modeling enable us to obtain user's friendly formulas for vanilla options in such a model.

In Theorem 12.2.1, we give a second approximation formula for vanilla options in a general local volatility model combined with Gaussian stochastic rates following the HJM framework. This formula is a sum of

- the price in the proxy Black-Scholes model with stochastic rates.
- the corrections terms which are a linear combination of Greeks of the leading term with explicit weights depending on the diffusion and the stochastic rate parameters.

We give also third order approximation (see Theorem 12.2.2) in such a model. In the case of time homogenous local volatility and Hull and White model, we give nice computations of the corrections weights in Subsection 12.2.1. We extend also our results for stochastic dividends and convenience yields (see Section 12.3). For example, in commodity field, our work can be seen as an extension to Gibson Schwartz model to handle local volatility functions. In Section 12.4, we give numerical examples for the accuracy of our approximation formula. Indeed, we benchmark our formula with Monte Carlo simulations and control variate. The accuracy of our approximation formulas turns out to be excellent.

Part I

Jump diffusion models

Chapter 3

Introduction

After the 1987 crash, practitioners realized that the standard Black Scholes model had serious flaws. Among these, the assumption of constant implied volatility for all strikes and maturities had become a great concern. In order to manage the so-called smile, academics and practitioners offered various extensions which are briefly introduced in the following.

Local volatility models. Introduced by Dupire [40], Rubinstein [101] and Derman Khani [36], these models do not assume anymore a constant volatility for all strikes but rather that the underlying (local) volatility is now a function of the spot and the time. This allows to price consistently options with different level of implied Black Scholes volatilities. Interestingly, one can derive an explicit formula for the local volatility σ in terms of the call prices $Call(T,K)$ for different strikes K and maturities T:

$$\sigma^2(T,K) = 2\frac{\frac{\partial Call(T,K)}{\partial T} + r_T K \frac{\partial Call(T,K)}{\partial K}}{K^2 \frac{\partial^2 Call(T,K)}{\partial K^2}}. \tag{3.1}$$

This equation is often referred as the Dupire's formula (for more details about this model see Chapter 6). Moreover, local volatility models are consistent with the market completeness assumption.

Stochastic volatility models. The key idea in these models is to assume that the volatility itself (like the stock price) is also stochastic. Theses types of models can generate easily different smiles and skews. The most common stochastic models are

- The SABR model introduced by Hagan in [61]:

$$dF_t = \sigma_t F_t^{\beta} dW_t, F_0 = f,$$
$$d\sigma_t = \nu \sigma_t dB_t, \sigma_0 = \alpha,$$
$$d\langle W,B\rangle_t = \rho dt,$$

  where (F_t) is the forward and

  - α is the initial volatility,
  - β is the skew parameter,
  - ν is the volatility of volatility,
  - ρ is the correlation.

  This model can be viewed as an extension of a CEV model but with stochastic volatility. The advantage of this model is that the implied Black Scholes volatility has an accurate analytic expression using geodesic transformations (see [61] and [74]).

- The Heston model (see [63]) is an extension of Black Scholes model for the underlying (S_t) with stochastic volatility:

$$\frac{dS_t}{S_t} = (r-q)\mathrm{d}t + \sqrt{v_t}dW_t,\ S_0 > 0, \tag{3.2}$$
$$\mathrm{d}v_t = \kappa(\theta - v_t)\mathrm{d}t + \xi\sqrt{v_t}\mathrm{d}B_t,\ v_0 > 0, \tag{3.3}$$
$$\mathrm{d}\langle W,B\rangle_t = \rho \mathrm{d}t, \tag{3.4}$$

  where

- r and q are respectively the risk free rate and the dividend yield,
- v_0 is the initial square of volatility,
- κ is the mean reversion parameter,
- θ is the long-term level,
- ξ is the volatility of volatility,
- ρ is the correlation.

Call (and put) can be easily priced in this model using the characteristic function and Fourier transform (for more details about Fourier transforms see Chapter 9).

In this class of models, as opposed to local volatility model, the assumption of market completeness does not hold anymore as there are two source of risk (namely two Brownian motions) for one single asset used to hedge the risk. To complete the market, one should use another option for the hedging (see [100]).

Exponential Levy models. In these models, the underlying is assumed to follow the exponential of a Levy process (for an introduction of Lévy process see [103]). In this model, one has

$$S_t = e^{rt+X_t},$$

where (X_t) is Levy process. We cite among these models:

- Merton model: $X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i$ where $(N_t)_{t\geq 0}$ is the Poisson process counting the jumps of X and the jump sizes Y_i are i.i.d. normal variables (for more details see [85]).
- Kou model: $X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i$ where $(N_t)_{t\geq 0}$ is the Poisson process counting the jumps of X and the jump sizes Y_i are distributed according to an asymmetric exponential law (for more details see [71]).
- Infinite intensity models: Variance Gamma and NIG model (for details about these models see [32]).

Note also that the market is still incomplete in these models (see [32]).

Local volatility plus Poisson jumps models. This kind of model has been introduced by Andersen and Andreasen in [8]. In this Part, we focus our work on this model.

The Andersen Andreasen model (we call it AA model) has been introduced after the 87 crash in order to handle a larger smile for short maturities. Their model can be seen as a perturbation of local volatility models by Poisson jumps. Then, the underlying in the AA model follows the stochastic differential equation (SDE) :

$$dX_t = \sigma(t, X_{t^-})dW_t + \mu(t, X_{t^-})dt + dJ_t, X_0 = x_0, \tag{3.5}$$

where

- $X_t = \log(S_t) + \int_0^t (q_s - r_s)ds$ is the logarithm of the forward, r is the deterministic risk free rate and q is the deterministic dividend yield,

- the compound Poisson process (J_t) and the Brownian motion (W_t) are independent,
- $J_t = \sum_{i=1}^{N_t} Y_i$ and Y_i are i.i.d. normal variables with mean η_J and volatility γ_J,
- N_t is a Poisson process with intensity λ,
- $\mu(t,x) = -\frac{\sigma^2(t,x)}{2} + \lambda(1 - e^{\mu_J + \frac{\gamma_J^2}{2}})$ in order to ensure martingale properties for (e^{X_t}).

The market is incomplete in such a model because of the presence of the jumps. Indeed, the class of equivalent martingale measures is infinite (see Chapters 9 and 10 of [32]). Moreover, one can find a martingale measure which is equivalent to a given prior and hence is able to reproduce quoted prices (see [31] and Chapter 13 of [32]).

Using standard arguments (see [8]), one can derive easily the backward partial integro differential equation (denoted by PIDE) for the call price $Call(t,S;T,K)$ which begins at time t with spot S, expires at maturity T and is exercised at strike K:

$$
\begin{aligned}
&Call_t(t,S) + (r_t - q_t - \lambda m) S Call_S(t,S) + \frac{1}{2}\sigma^2(t, \int_0^t (q_s - r_s) ds + \ln(S)) S^2 Call_{SS}(t,S) \\
&+ \lambda \mathbb{E}[Call(t, e^{J_t} S) - Call(t,S)] - r_t Call(t,S) = 0,
\end{aligned}
$$

where $m = e^{\mu_J + \frac{\gamma_J^2}{2}} - 1$.

Moreover, the European call satisfies the forward PIDE:

$$
\begin{aligned}
&-Call_T(T,K) + (q_T - r_T - \lambda m) K Call_K(T,K) + \frac{1}{2}\sigma^2(T, \int_0^T (r_s - q_s) ds + \ln(K)) K^2 Call_{KK}(T,K) \\
&+ \lambda \mathbb{E}[e^{J_T} Call(T, e^{-J_T} K) - (m+1) Call(T,K)] - q_T Call(T,K) = 0,
\end{aligned}
$$

The proof is based on an application of the Tanaka-Meyer formula (see [8]).

3.1 Merton model

The AA model has closed-form formulas only when the function σ do not depend on the forward. In other words, for every real number x, $\sigma(t,x) = \sigma(t,x_0)$. In this case, we retrieve exactly Merton's model:

$$
dX_t^M = \sigma(t,x_0) dW_t + \mu(t,x_0) dt + dJ_t, X_0^M = x_0. \tag{3.6}
$$

Then the call price in the Merton model is:

$$
Call_{Merton}(0,S;K,T) = \sum_{i=0}^{\infty} \frac{(\lambda T)^i}{i!} e^{-\lambda T - \int_0^T r(u) du} BSCall\left(F_T e^{i(\eta_J + \frac{\gamma_J^2}{2})}, K, T, \sqrt{\frac{\int_0^T \sigma^2(t,x_0) dt + i\gamma_J^2}{T}}\right),
$$

where

$$
F_T = e^{x_0 + \int_0^T (r(u) - q(u)) du + \lambda(1 - \exp(\eta_J + \gamma_J^2/2)) T},
$$

and $BSCall(S,K,T,v)$ is the Black-Scholes price for a call on an underlying S_t with initial condition $S_0 = S$, volatility v, exercised at maturity T and strike K, where the risk-free rate and the dividend yield are set at 0%.

Proof. For simplicity, we consider the functions $r, q, \sigma(., x_0)$ constant. In this case, the call price in Merton's model is:

$$
\begin{aligned}
Call_{Merton}(0,S;K,T) &= e^{-rT}\mathbb{E}[(e^{x_0+(r-q+\lambda(1-\exp(\eta_J+\gamma_J^2/2))-\frac{\sigma^2}{2})T+\sigma W_T+\sum_{j=1}^{N_T}Y_j}-K)^+] \\
&= e^{-rT}\mathbb{E}[\mathbb{E}[(e^{x_0+(r-q+\lambda(1-\exp(\eta_J+\gamma_J^2/2))-\frac{\sigma^2}{2})T+\sigma W_T+\sum_{j=1}^{N_T}Y_j}-K)^+|N_T]] \\
&= e^{-rT}\sum_{i=0}^{\infty}e^{-\lambda T}\frac{(\lambda T)^i}{i!}\mathbb{E}[(e^{x_0+(r-q+\lambda(1-\exp(\eta_J+\gamma_J^2/2))-\frac{\sigma^2}{2})T+\sigma W_T+\sum_{j=1}^{i}Y_j}-K)^+] \\
&= e^{-rT}\sum_{i=0}^{\infty}e^{-\lambda T}\frac{(\lambda T)^i}{i!}BSCall\left(e^{x_0+(r-q)T+\lambda(1-\exp(\eta_J+\gamma_J^2/2))T+i(\eta_J+\frac{\gamma_J^2}{2})},K,T,\sqrt{\frac{\sigma^2T+i\gamma_J^2}{T}}\right),
\end{aligned}
$$

where we have conditioned on the number of jumps N_T in the second Equality, we have used the independence of the number of jumps N_T and the random variables (W_T, $(Y_i)_{i\geq 1}$) in the third Equality, we have used the definition of the BSCall function and the independence of the Gaussian variables (W_T, $(Y_i)_{i\geq 1}$) for the last one. □

3.2 Forward PIDE and calibration

For general forms of the local volatility σ, there are neither closed-form formulas nor efficient approximation formulas. Otherwise, one can use numerical methods and especially numerical partial differential schemes applied to forward PIDE for calibration purpose. We recall that Forward PIDE means that PIDE equation depends on forward parameters (strike and maturity) while Backward PIDE depends on backward parameters (underlying and starting time). The advantage of such forward PIDE arises mainly when one needs to compute the model price for many strikes and maturities. Indeed, the forward PIDE gives the price function for any T and K (at fixed S and t), while there are as many backward PIDEs as the number of couples (T,K) (which define the boundary condition). This need to compute model prices for many T and K naturally arises in the calibration procedure, where we have to solve the following optimization problem:

$$
\begin{aligned}
\min_{Model\ parameters}\sum_{\text{market data}}&(\text{Model Call Price}(T_i,K_i)-\text{Market Call Price}(T_i,K_i))^2 \\
&+ eventually\ a\ volatility\ regularization\ term^1. \qquad (3.7)
\end{aligned}
$$

Each iteration in the optimization routine involves only the resolution of one PIDE, the forward one. However, although using this trick, the calibration is computationally expensive and takes about one minute ([8]).

3.3 Motivation and main results

Our aim in this Part is to

- give an **accurate analytic approximation** for the European option:

$$\mathbb{E}[h(X_T)]$$

[1] The calibration of jump diffusion is an ill posed problem. The regularization or the convex penalization term is chosen in order to obtain a unique solution in a stable manner (for more details about possible choices see [31]).

where h is a **non smooth function** and T is the maturity of the option. For the case of call option without risk free rate and dividend yield, one has $h(x) = (e^x - K)^+$.

- Calibrate within a **computational time** smaller than one second (**1sec**) all the model parameters: the jump parameters λ , η_J, γ_J and the volatility function $\sigma(t,x)$.

Intuition leading the choice of the proxy: In financial markets

- the volatility of the stock price is small,
- and the frequency (or the size) of the price jumps is small.

Therefore, the log price process $(X_t)_{0\leq t\leq T}$ is not far from x_0. Hence, by replacing X_{t^-} by x_0 in the SDE (3.5) of the AA model:

$$dX_t = \sigma(t,X_{t^-})dW_t + \mu(t,X_{t^-})dt + dJ_t, X_0 = x_0,$$

we retrieve exactly the Merton model (3.6):

$$dX_t^M = \sigma(t,x_0)dW_t + \mu(t,x_0)dt + dJ_t, X_0^M = x_0.$$

In other words, **the Merton's model** (3.6) **is a good proxy for the AA model** (3.5). However, it is not enough to improve the approximation. We need to add some corrections in order to better approximate the AA model. To achieve that, we use a suitable parameterization of this model as follows:

$$dX_t^\varepsilon = \varepsilon(\sigma(t,X_{t^-}^\varepsilon)dW_t + \mu(t,X_{t^-}^\varepsilon)dt + dJ_t), X_0^\varepsilon = x_0$$

so that $X_t^1 = X_t$ and $X_t^0 = X_t^M$. Therefore, we can perform a Taylor expansion w.r.t. ε for $\varepsilon = 1$ around 0:

$$X_T = \underbrace{x_0 + \frac{\partial^1 (X_T^\varepsilon)}{\partial \varepsilon}|_{\varepsilon=0}}_{:=X_T^M} + \frac{\partial^2 (X_T^\varepsilon)}{2\partial \varepsilon^2}|_{\varepsilon=0} + \cdots .$$

Suppose h is smooth enough, then using a Taylor expansion for the function h at the first order, we get:

$$\begin{aligned} E[h(X_T)] &= E[h(X_T^M + \frac{\partial^2 (X_T^\varepsilon)}{2\partial \varepsilon^2}|_{\varepsilon=0} + \cdots)] \\ &= E[h(X_T^M)] + \mathbb{E}[h^{(1)}(X_T^M)\frac{\partial^2 (X_T^\varepsilon)}{2\partial \varepsilon^2}|_{\varepsilon=0}] + \cdots \\ &= Merton\ price + corrections\ terms + errors, \end{aligned} \tag{3.8}$$

where

- The correction term $\mathbb{E}[h^{(1)}(X_T^M)\frac{\partial^2 (X_T^\varepsilon)}{2\partial \varepsilon^2}|_{\varepsilon=0}]$ in Equation (3.8) is a combination of Greeks[2] (Delta, Gamma, Epsilon) of the Merton Model:

$$\mathbb{E}[h^{(1)}(X_T^M)\frac{\partial^2 (X_T^\varepsilon)}{2\partial \varepsilon^2}|_{\varepsilon=0}] = \sum_{i=1}^{3} \alpha_{i,T} \mathrm{Greek}_i^h(X_T^M) + \sum_{i=1}^{3} \beta_{i,T} \mathrm{Greek}_i^h(X_T^M + Y'),$$

[2] Keep in mind that Greeks have closed formula in Merton's model

where

$$\begin{aligned}
\text{Greek}_i^h(Z) =&\partial_x^i \mathbb{E}[h(x+Z)]|_{x=0},\\
\alpha_{1,T} =&\frac{1}{2}C_1 + \lambda(e^{\eta_J+\frac{\gamma^2}{2}}-1)C_2, \alpha_{2,T} = -\frac{3}{2}C_1 - \lambda(e^{\eta_J+\frac{\gamma^2}{2}}-1)C_2,\\
\alpha_{3,T} =&-\alpha_{1,T}-\alpha_{2,T},\\
\beta_{1,T} =&-\lambda\eta_J C_2, \beta_{2,T} = \lambda(\eta_J - \gamma_J^2)C_2,\\
\beta_{3,T} =&-\beta_{1,T}-\beta_{2,T},
\end{aligned}$$

and $C_1 = \int_0^T \sigma^2(t,x_0)(\int_t^T \sigma(s,x_0)\partial_x\sigma(s,x_0)ds)$, $C_2 = \int_0^T t\sigma(t,x_0)\partial_x\sigma(t,x_0)dt$. These corrections terms are made explicit in Chapter 4 in Theorem 4.2.1.

- The error term *errors* in Equation (3.8) is estimated for vanilla payoff in Chapter 4 in Theorem 4.5.2. A simple interpretation of the error is the following: if the volatility σ and the jump size involved in the SDE (3.5) are of the order of Δ, then the error term is of the order of $(\Delta\sqrt{T})^3$.

Note also that the approximation derived in Chapter 4 covers not only call-put options but also all vanilla payoffs depending on X_T (smooth, vanilla, binary). Indeed, the errors are analyzed according to the payoff smoothness (see for instance Theorem 4.5.1 for smooth payoff in general, Theorem 4.5.2 for vanilla options, Theorem 4.5.3 for binary options).

The following chapter is exactly the article "smart expansion and fast calibration for jump diffusions" published in the journal "Finance and Stochastics", volume 13(4), pages 563:589, 2009. Chapter 5 provides additional numerical results concerning the accuracy of the approximation formula for bumped parameters and calibration/optimization robustness as well.

Chapter 4

Smart expansion and fast calibration for jump diffusion

Published in "Finance and Stochastics", volume 13(4), pages 563:589, 2009.

Using Malliavin calculus techniques, we derive in this chapter an analytical formula for the price of European options, for any model including local volatility and Poisson jump process. We show that the accuracy of the formula depends on the smoothness of the payoff function. This approach relies on an asymptotic expansion related to small diffusion and small jump frequency/size. The formula derived in this chapter has excellent accuracy (the error on implied Black-Scholes volatilities for call option is smaller than 2 bp for various strikes and maturities). As a result, model calibration becomes very rapid.

4.1 Introduction

The standard Black-Scholes formula (1973) was derived under the assumption of lognormal diffusion with constant volatility to price calls and puts. However, this hypothesis is unrealistic under real market conditions because we need to use different volatilities to equate different option strikes K and maturities T. Market data shows that the shape of the implied volatilities takes the form of a smile or a skew.

In order to fit the smile or the skew, Dupire (in [40]) and Rubinstein (in [101]) use a local volatility $\sigma_{loc}(t,f)$ depending on time t and state f to fit the market. This hypothesis of local volatility is interesting for hedging because it maintains the completeness of the market. However, in a few cases [1], one has closed formulas. In the case of homogeneous volatility, singular perturbation techniques in [62] have been used to obtain asymptotic expression for the price of vanilla options (call, put). Other cases have been derived using an asymptotic expansion of the heat kernel for short maturity (see [74]).

But Andersen and Andreasen in [8] show that this sole assumption of local volatility is not compatible with empirical evidence (for instance, the post-crash of implied volatility of the S&P500 index). Hence, they derived a model with local volatility plus a jump process to fit the smile (we write it AA model). Their model may be seen as a perturbation of pure local volatility models. Of course, this is not the only alternative modeling[1]. The AA model fits some market data well (see calibration results in [8] and those in this work), although we are aware that it does not work systematically nicely. In the following, we do not discuss the relevance of this model in specific situations. We simply focus on this model in order to illustrate our new approach for numerical pricing and fast calibration. For an analogous study on the time dependent Heston model, we refer to the chapter 10. Andersen and Andreasen [8] calibrate their model by solving the equivalent forward PIDE. This sort of problem could be handled numerically using: an ADI-FFT scheme in [8], a Finite Element Method in [83], an explicit implicit PIDE-FFT method for general Lévy processes in [33] or Predictor Corrector methods to improve the accuracy of the PIDE in [19]. In the best case, all of these methods lead to a time of calibration of roughly one minute (see [8]). Can we reduce this computational time? Is it possible to reach a time of calibration as short as the computational time of a closed formula such as Merton's [85]? Our present research responds positively to the above questions.

In order to handle even more general situations we consider, for the one dimensional underlying state process, the solution of the stochastic differential equation (SDE):

$$dX_t = \sigma(t,X_{t-})dW_t + \mu(t,X_{t-})dt + dJ_t, \quad X_0 = x_0. \tag{4.1}$$

For instance one may think of $(X_t)_t$ as the log asset price. Here $(W_t)_{0\le t\le T}$ is a standard real Brownian motion on a filtered probability space $(\Omega,\mathscr{F},(\mathscr{F}_t)_{0\le t\le T},\mathbb{P})$ with the usual assumption on the filtration $(\mathscr{F}_t)_{0\le t\le T}$ and $(J_t)_{0\le t\le T}$ is a compound Poisson process independent of $(W_t)_t$ defined by: $J_t = \sum_{i=1}^{N_t} Y_i$ where N_t is a counting Poisson process with constant jump intensity λ and $(Y_i)_{i\in\mathbb{N}^*}$ are i.i.d. normal

[1] for instance, see the book by Lewis [79] on stochastic volatility models or the one by Gatheral [50] on models explaining the volatility surface.

variables with mean η_J and volatility γ_J. Our main objective is to give an accurate analytic approximation of the expected payoff (or fair price of this option)

$$\mathbb{E}(h(X_T))$$

for a given terminal function h and for a fixed maturity T.

The approximation can be applied to the following models:

Example 4.1.1. *AA model on the log-asset.*
In this case, $(X_t)_t$ is the logarithm of the underlying asset, $\sigma(t,x)$ is its local volatility and $\mu(t,x) = \lambda(1-e^{\eta_J+\frac{\gamma_J^2}{2}}) - \frac{\sigma^2(t,x)}{2}$ in order to guarantee the martingale property for $(e^{X_t})_t$. For a call exercised at maturity T, with strike K, $h(x) = e^{-\int_0^T r(u)du}(e^{\int_0^T (r(u)-q(u))du}e^x - K)_+$ where r is the deterministic risk-free rate term and q is the deterministic dividend term. This model was derived in [8]. In this work we mainly focus our discussion on this model.

Example 4.1.2. *Jump diffusion model on the asset.*
$(X_t)_t$ is the forward contract with maturity T, $\sigma(t,x)$ is its volatility and $\mu(t,x) = -\lambda\eta_J$. For a call exercised at maturity T, $h(x) = e^{-\int_0^T r(u)du}(x-K)_+$. The primary focus of this model is the implied normal volatility instead of standard implied Black-Scholes volatility (Japanese markets in [61]) and it includes the presence of price jumps.

Heuristics of our approximation and model proxy. In practice, at first glance, it is reasonable to think that $(X_t)_t$ (in the AA model) is approximated by a Merton model, where the coefficients μ and σ only depend on time. We denote this proxy by $(X_t^M)_t$ and it is defined by

$$dX_t^M = \sigma(t,x_0)dW_t + \mu(t,x_0)dt + dJ_t, \quad X_0^M = x_0. \tag{4.2}$$

This approximation can be justified by one of the following situations.

i) The functions $\mu(\cdot)$ and $\sigma(\cdot)$ have small variations, which means that $\sigma(t,X_{t-}) \approx \sigma(t,x_0)$ and analogously for μ.

ii) The diffusion component is small (i.e. $|\mu|_\infty + |\sigma|_\infty$ is small) and the jump component as well (i.e. $\lambda(|\eta_J| + \gamma_J)$ is small, meaning that the jump frequency or the jump size is small), which results in $X_t \approx x_0$. This case is not equivalent to situation *i)* because the functions may be small and yet have large variations.

iii) Another obvious reason may be that the maturity T is small, leading to $X_T \approx x_0$.

The heuristics *i)* and *ii)* are coherent with the parameter values taken in [8]. When the three conditions are carried out at the same time, we expect our approximation to become even more accurate. Note also that no jump cases ($\lambda = 0$) are allowed. The above qualitative features *i)* and *ii)* are encoded into quantitative constants M_0, M_1 and M_J defined in (4.5) and will be discussed later in this work. The above heuristic rule implies that

$$\mathbb{E}(h(X_T)) = \mathbb{E}(h(X_T^M)) + \text{error}.$$

The term $\mathbb{E}(h(X_T^M))$ is the price in the Merton proxy which is explicit (see Remark 4.2.1). But this sole approximation is too rough to be sufficiently accurate. Our work consists of deriving correction terms for the above equality to attain a remarkably good approximation.

Smart expansion. To perform a rigorous analysis, we use a suitable parameterization w.r.t. $\varepsilon \in [0,1]$:

$$dX_t^\varepsilon = \varepsilon(\sigma(t,X_{t-}^\varepsilon)dW_t + \mu(t,X_{t-}^\varepsilon)dt + dJ_t), X_0^\varepsilon = x_0, \tag{4.3}$$

so that $X_t^1 = X_t$. We write

$$g(\varepsilon) = \mathbb{E}(h(X_T^\varepsilon)) \tag{4.4}$$

and our approach consists of expanding the price (4.4) with respect to ε. For related expansion results, see [48] and [106]. But the accuracy of the expansion is not related to ε because the value of interest $\varepsilon = 1$ is not small. This is a significant difference as compared with singular perturbation techniques. Parameterization is just a tool to derive convenient representations. By using an asymptotic expansion in the context of small diffusions and small jumps (relative to the frequency or to the size), we can establish estimates of the derivatives. This allows us to make an explicit contribution at given order and to control the error. This is achieved by using the infinite dimensional analysis of Malliavin calculus. Here, we focus our analysis on the first terms[2], for which we provide explicit formulas. We also give explicit upper bounds of the errors for general forms of $\mu(\cdot)$ and $\sigma(\cdot)$. However, the smaller the parameters $\mu(\cdot), \sigma(\cdot)$ and $\lambda(|\eta_J| + \gamma_J)$ are, the smaller the maturity T is, or the smaller the derivatives of the functions $\sigma(\cdot)$ and $\mu(\cdot)$ w.r.t. the second variable are, the more accurate the expansion is. Given realistic parameters, the accuracy is indeed very good (less than **2bp** in implied volatilities for various strikes and maturities). As a result of these expansions, we prove that the price (4.4) in our general model (4.1) equals the price in the Merton model plus a combination of Greeks (still in the Merton model). Hence, numerical evaluation of all these terms is straightforward, with a computational cost equivalent to the closed Merton formula. The residual terms (otherwise stated as error) are also estimated and their amplitudes depend on the smoothness of the payoff. We distinguish three cases: smooth, vanilla (call, put) and binary payoffs. In practice, the vanilla case is likely to be the most useful.

This is our main contribution. Furthermore, from the approximation price we observe that one may obtain a volatility smile for short maturities (since we use the Merton model as a proxy) and a volatility skew for long maturities (due to local volatility function).

Comparison with the literature. We refer in particular to Hagan et al. in [61] for the SABR model, to Fouque et al. in [44] for stochastic volatility models, or to Antonelli-Scarletti in [12]. In all these works, as opposed to our approach, a perturbation analysis w.r.t. the volatility, the mean reversion parameters, or the correlation, is performed and this leads to writing the price as a main term (essentially a Black-Scholes price) plus an integral of Greeks over maturities. In the time homogeneous case, the authors successfully compute or approximate this integral, which strongly relies on PDE arguments. In our case, we do not approximate the underlying PDE (or the related operator) but owing to Malliavin calculus, we directly focus on the law of the random variable X_T given $X_0 = x_0$ and not necessarily on the process for any initial condition. Thus, we are able to handle time inhomogeneous coefficients and jumps as well, without extra effort. This is a very significant difference from previous research.

Outline of the chapter. In the following, we present some notations and assumptions that will be used throughout the chapter. Section 4.2 is aimed at presenting our methodology in an heuristic way to approximate the expected cost. Rigorous results are proved in Section 4.5. In Section 4.3, we derive financial modeling consequences from these formulas. These observations lead to justifying simplified

[2] in the former version of this work, terms at any order have been analyzed.

choices of the local volatility (of the CEV type), to predict the form of all attainable smiles with their dynamics. In Section 4.4, we first give a methodology for implementing the approximation formula. Secondly, we show how to efficiently use our formula for calibrating the model using a relevant algorithm. Finally, we detail numerical applications in calibration for real market data using our simplified form of local volatility. In Section 4.5, we analyze the amplitude of the correction and error terms of the approximation formula; the analysis depends on the kinds of payoff (smooth payoff in Theorem 4.5.1, vanilla options in Theorem 4.5.2, binary options in Theorem 4.5.3).

Notations used throughout the chapter.
Differentiation. If these derivatives have a meaning, we write:

- $\psi_t^{(i)}(x) = \frac{\partial^i \psi}{\partial x^i}(t,x)$ for every function ψ of two variables.
- $X_{i,t} = \frac{\partial^i X_t^\varepsilon}{\partial \varepsilon^i}|_{\varepsilon=0}$. These processes play a crucial role in the work that follows.
- When there is no ambiguity, we simply write $\sigma_t = \sigma(t,x_0), \mu_t = \mu(t,x_0), \sigma_t^{(i)} = \frac{\partial^i \sigma}{\partial x^i}(t,x_0), \mu_t^{(i)} = \frac{\partial^i \mu}{\partial x^i}(t,x_0)$.

The following definition is used to distinguish the payoff functions h.

Definition 4.1.1. *As per usual, we define $\mathscr{C}_0^\infty(\mathbb{R})$ as the space of real infinitely differentiable functions h with compact support (smooth payoffs). The sup-norm of the function h is denoted by $|h|_\infty$. We define $\mathscr{H}$ as the space of functions with growth being at most exponential. In other words, a function h belongs to $\mathscr{H}$ if $|h(x)| \le c_1 e^{c_2|x|}$ for any x, for two constants c_1 and c_2.*

The following notation provides a convenient representation of the correction terms.

Definition 4.1.2. ***Greeks.*** *Let Z be a random variable. Given a payoff function h, we define the i^{th} Greek for the variable Z by the quantity (when it has a meaning) :*

$$\mathrm{Greek}_i^h(Z) = \frac{\partial^i \mathbb{E}[h(Z+x)]}{\partial x^i}|_{x=0}.$$

Given appropriate smoothness assumptions concerning h, one also has

$$\mathrm{Greek}_i^h(Z) = \mathbb{E}[h^{(i)}(Z)].$$

Assumptions. In order to get accurate approximations, we may assume that coefficients σ and μ are smooth enough.

- ***Assumption*** *(R_4). The functions $\sigma(\cdot)$ and $\mu(\cdot)$ are continuously differentiable w.r.t. x up to order 4. In addition, these functions and their derivatives are uniformly bounded.*

The functions and their derivatives could be piecewise continuous w.r.t. the time variable, without changing the following approximation formulas and the following error bounds.

The assumption (R_4) seems to be restrictive because one requires $\sigma(\cdot), \mu(\cdot)$ and their derivatives w.r.t. x to be bounded. On the one hand, this hypothesis is clearly too strong for us to use in the derivation of our smart expansion: indeed, the reader may check that polynomial growth conditions are sufficient for this purpose. On the other hand, assuming that the derivatives are bounded is much more convenient for explanation purposes. It enables us to state all our error estimates purely in terms of the following constants:

$$\begin{cases} M_1 = & \max_{1\leq i\leq 4}(|\sigma^{(i)}|_\infty + |\mu^{(i)}|_\infty), \\ M_0 = & \max_{0\leq i\leq 4}(|\sigma^{(i)}|_\infty + |\mu^{(i)}|_\infty), \\ M_J = & |\eta_J| + \gamma_J. \end{cases} \tag{4.5}$$

M_1, M_0 and M_J play complementary roles.

a) The constant M_1 is a measure of the norm of the derivatives (w.r.t. x) of the objective functions $\sigma(\cdot)$ and $\mu(\cdot)$. All our error estimates (see Theorems 4.5.1-4.5.2-4.5.3) are linear w.r.t. M_1, which corroborates the proxy intuition explained in item *i)*. The smaller the value of M_1 is, the closer X and X^M are, and as a result, approximation is increasingly accurate. At the limit $M_1 = 0$, the initial model and the proxy coincide ($X_t = X_t^M$) and our approximation formula becomes exact.

b) The constants M_0 and M_J also include estimates of the amplitudes of $\sigma(\cdot), \mu(\cdot)$ and of the jump components. All our error estimates also depend on powers of M_0 and M_J. This mathematically justifies proxy intuition *ii)*. The smaller M_0 and M_J are, the better the resulting accuracy.

In our next theorems, we also clarify the dependence of our estimates regarding jump frequency λ and maturity T, because as these parameters decrease, the approximation becomes increasingly accurate.

To perform the infinitesimal analysis, we rely on smoothness properties which are not provided by the payoff functions, but rather by the law of the underlying stochastic models (this is related to Malliavin calculus). The following ellipticity assumption on volatility combined with (R_4) guarantees these smoothness properties.

- ***Assumption** (E). σ does not vanish and for a positive constant C_E, one has*

$$1 \leq \frac{|\sigma|_\infty}{\sigma_{inf}} \leq C_E$$

where $\sigma_{inf} = \inf_{(t,x)\in\mathbb{R}^+\times\mathbb{R}} \sigma(t,x)$.

We also need to separate our analysis according to payoff smoothness. We thus divide our analysis into three cases.

- ***Assumption** (H_1). h belongs to $\mathscr{C}_0^\infty(\mathbb{R})$. This case corresponds to smooth payoffs.*

- ***Assumption** (H_2). h is almost everywhere differentiable. In addition, h and $h^{(1)}$ belong to $\mathscr{H}$. This case corresponds to vanilla options (call, put).*

- ***Assumption** (H_3). h belongs to $\mathscr{H}$. This case includes binary options (digital).*

4.2 Smart Taylor Development

In this section, we formally show how to replace the price $\mathbb{E}(h(X_T))$ by using that found in the Merton model $\mathbb{E}(h(X_T^M))$ with appropriate correction terms. Rigorous justification of the following expansions is postponed to Section 4.5.

The initial trick of our smart expansion lies in the use of the parameterized process $(X_t^\varepsilon)_t$ for $\varepsilon \in [0,1]$, defined in (4.3). Under assumption (R_4), almost surely for any t, X_t^ε is C^3 w.r.t ε (see Theorem 2.3 in [48]). If we put $X_{i,t}^\varepsilon = \frac{\partial^i X_t^\varepsilon}{\partial \varepsilon^i}$, we get

$$\begin{aligned} dX_{1,t}^\varepsilon =& \sigma_t(X_{t-}^\varepsilon)dW_t + \mu_t(X_{t-}^\varepsilon)dt + dJ_t \\ &+ \varepsilon X_{1,t-}^\varepsilon(\sigma_t^{(1)}(X_{t-}^\varepsilon)dW_t + \mu_t^{(1)}(X_{t-}^\varepsilon)dt), \quad X_{1,0}^\varepsilon = 0. \end{aligned}$$

From the definitions, $X_{i,t} \equiv \frac{\partial^i X_t^\varepsilon}{\partial \varepsilon^i}|_{\varepsilon=0}$, $\sigma_t^{(i)} \equiv \sigma^{(i)}(t,x_0)$ and $\mu_t^{(i)} \equiv \mu^{(i)}(t,x_0)$, we easily get

$$\begin{aligned} dX_{0,t} =& 0, \quad X_{0,0} = x_0, \\ dX_{1,t} =& \sigma_t dW_t + \mu_t dt + dJ_t, \quad X_{1,0} = 0, \\ dX_{2,t} =& 2X_{1,t-}(\sigma_t^{(1)} dW_t + \mu_t^{(1)} dt), \quad X_{2,0} = 0. \end{aligned}$$

Thus, the Merton model is obtained by the first order expansion of X^ε at $\varepsilon = 0$:

$$X_{0,T} + X_{1,T} = x_0 + X_{1,T} = X_T^M.$$

We now use the Taylor formula twice: first, for X_T^ε at the second order w.r.t ε around $\varepsilon = 0$, second for smooth function h at the first order w.r.t x around $x_0 + X_{1,T} = X_T^M$. One gets:

$$\mathbb{E}[h(X_T^1)] = \mathbb{E}[h(x_0 + X_{1,T} + \frac{X_{2,T}}{2} + \cdots)] = \mathbb{E}[h(X_T^M)] + \mathbb{E}[h^{(1)}(X_T^M)\frac{X_{2,T}}{2}] + \cdots.$$

Then, the price $\mathbb{E}[h(X_T)]$ can be approximated by a summation of two terms :

- $\mathbb{E}[h(X_T^M)]$: The leading order which corresponds to the Merton price (BS price when $\lambda = 0$) for the payoff h.
- $\mathbb{E}[h^{(1)}(X_T^M)\frac{X_{2,T}}{2}]$: The correction term which is made explicit in the next theorem.

Theorem 4.2.1. ***(Main approximation price formula).***
Suppose that the process data fulfills (R_4) and (E) and that the payoff function fulfills one of the assumptions (H_1), (H_2) or (H_3). Then

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^M)] + \sum_{i=1}^{3} \alpha_{i,T} \mathrm{Greek}_i^h(X_T^M) + \sum_{i=1}^{3} \beta_{i,T} \mathrm{Greek}_i^h(X_T^M + Y') + \mathrm{Error}, \tag{4.6}$$

where

$$\begin{aligned} \alpha_{1,T} =& \int_0^T \mu_t (\int_t^T \mu_s^{(1)} ds) dt, \\ \alpha_{2,T} =& \int_0^T (\sigma_t^2 (\int_t^T \mu_s^{(1)} ds) + \mu_t (\int_t^T \sigma_s \sigma_s^{(1)} ds)) dt, \\ \alpha_{3,T} =& \int_0^T \sigma_t^2 (\int_t^T \sigma_s \sigma_s^{(1)} ds) dt, \\ \beta_{1,T} =& \lambda \eta_J \int_0^T t \mu_t^{(1)} dt, \\ \beta_{2,T} =& \lambda \int_0^T t(\gamma_J^2 \mu_t^{(1)} + \eta_J \sigma_t \sigma_t^{(1)}) dt, \\ \beta_{3,T} =& \lambda \gamma_J^2 \int_0^T t \sigma_t \sigma_t^{(1)} dt, \end{aligned}$$

Y' is an independent copy of the variables $(Y_i)_{i\in\mathbb{N}^*}$.
In addition, estimates for the error term Error *in the cases (H_1), (H_2) and (H_3) are respectively given in Theorems 4.5.1, 4.5.2 and 4.5.3.*

To prove Theorem 4.2.1, it remains to show that $\mathbb{E}[h^{(1)}(X_T^M)\frac{X_{2,T}}{2}]$ is equal to the two summations of (4.6). The reader familiar with Malliavin calculus for the computations of Greeks (see [46], [54], ...) may recognize in the expansion of $\mathbb{E}[h^{(1)}(X_T^M)\frac{X_{2,T}}{2}]$ the generic form of some derivatives (or Greeks) of $\mathbb{E}[h^{(1)}(X_T^M)]$, derivatives which are written as the expectation of $h^{(1)}(X_T^M)$ multiplied by random weights. This is indeed our methodology to explicitly compute the correction terms in the formula (4.6).

Proof. Define the new function G by $G(x) = h(x + x_0 + \int_0^T \mu_t dt)$. One has:

$$\begin{aligned}\mathbb{E}[\frac{X_{2,T}}{2}h^{(1)}(X_T^M)] &= \mathbb{E}[\frac{X_{2,T}}{2}G^{(1)}(\int_0^T \sigma_t dW_t + J_T)]\\ &= \mathbb{E}[(\int_0^T X_{1,t-}(\sigma_t^{(1)}dW_t + \mu_t^{(1)}dt))G^{(1)}(\int_0^T \sigma_t dW_t + J_T)].\end{aligned}$$

Write $(X_{1,t}^c)_t$ for the continuous part of $(X_{1,t})_t$. Using Lemma 4.6.2 (since J_T is independent of $(W_t)_{t\in[0,T]}$) and $\text{Leb}\{t\in[0,T] : X_{1,t} = X_{1,t-}\} = 0$ a.s. (see Chapter 1, page 6, Equation (1.10) in [103]), one gets:

$$\begin{aligned}\mathbb{E}[\frac{X_{2,T}}{2}h^{(1)}(X_T^M)] &= \mathbb{E}[(\int_0^T \sigma_t\sigma_t^{(1)}X_{1,t}^c dt)G^{(2)}(\int_0^T \sigma_t dW_t + J_T)]\\ &+ \mathbb{E}[(\int_0^T \mu_t^{(1)}X_{1,t}^c dt)G^{(1)}(\int_0^T \sigma_t dW_t + J_T)]\\ &+ \mathbb{E}[(\int_0^T \sigma_t\sigma_t^{(1)}J_t dt)G^{(2)}(\int_0^T \sigma_t dW_t + J_T)]\\ &+ \mathbb{E}[(\int_0^T \mu_t^{(1)}J_t dt)G^{(1)}(\int_0^T \sigma_t dW_t + J_T)].\end{aligned}$$

Apply Lemmas 4.6.3 and 4.6.4 and use Definition 4.1.2 of Greeks to get the result. □

Remark 4.2.1. *The above price approximation is a summation of three terms:*

1. *$\mathbb{E}[h(X_T^M)]$: The leading order corresponding to the price when the functions $(\sigma_t)_t$ and $(\mu_t)_t$ are deterministic. We know that in this case, there is a closed formula : either the Merton closed formula for call (put), or FFT tools for any other payoff because the characteristic function of X_T^M is explicit. For instance, the formula for a call in the Merton model (see [85]) on the log asset is:*

$$\sum_{i=0}^{\infty}\frac{(\lambda T)^i}{i!}e^{-\lambda T-\int_0^T r(u)du}BSCall\left(F_T e^{i(\eta_J+\frac{\gamma_J^2}{2})}, K, T, \sqrt{\frac{\int_0^T \sigma_t^2 dt + i\gamma_J^2}{T}}\right),$$

where

$$F_T = e^{x_0+\int_0^T (r(u)-q(u))du+\lambda(1-\exp(\eta_J+\gamma_J^2/2))T},$$

and $BSCall(S,K,T,v)$ is the Black-Scholes price for a call on an underlying S_t with initial condition $S_0 = S$, volatility v, exercised at maturity T and strike K, where the risk-free rate and the dividend yield are set at 0%.

2. $\sum_{i=1}^{3} \alpha_{i,T} \mathrm{Greek}_i^h(X_T^M)$: *The volatility and drift correction term which depends on the first derivatives of μ and σ. This term can be computed as easily as the main term.*

3. $\sum_{i=1}^{3} \beta_{i,T} \mathrm{Greek}_i^h(X_T^M + Y')$: *The jump correction term which depends on the first derivatives of μ, σ and on the jump parameters. Since Y' is also Gaussian and independent of X_T^M, the computation of these Greeks are similar to the previous ones, by adding to the mean $\int_0^T \mu_t dt$ and variance $\int_0^T \sigma_t^2 dt$ the quantities η_J and γ_J^2.*

Remark 4.2.2. *In the AA model on the log-asset, one has:*

$$\begin{aligned}
\alpha_{1,T} =& \frac{1}{2}\int_0^T \sigma_t^2 \Big(\int_t^T \sigma_s \sigma_s^{(1)} ds\Big) dt + \lambda\big(e^{\eta_J + \frac{\gamma_J^2}{2}} - 1\big) \int_0^T t\sigma_t \sigma_t^{(1)} dt, \\
\alpha_{2,T} =& -\frac{3}{2}\int_0^T \sigma_t^2 \Big(\int_t^T \sigma_s \sigma_s^{(1)} ds\Big) dt - \lambda\big(e^{\eta_J + \frac{\gamma_J^2}{2}} - 1\big) \int_0^T t\sigma_t \sigma_t^{(1)} dt, \\
\alpha_{3,T} =& \int_0^T \sigma_t^2 \Big(\int_t^T \sigma_s \sigma_s^{(1)} ds\Big) dt, \\
\beta_{1,T} =& -\lambda \eta_J \int_0^T t\sigma \sigma_t^{(1)} dt, \\
\beta_{2,T} =& \lambda(\eta_J - \gamma_J^2) \int_0^T t\sigma_t \sigma_t^{(1)} dt, \\
\beta_{3,T} =& \lambda \gamma_J^2 \int_0^T t\sigma_t \sigma_t^{(1)} dt.
\end{aligned}$$

Thus, the computation of these constants is simply reduced to that of $\int_0^T t\sigma_t \sigma_t^{(1)} dt$ and $\int_0^T \sigma_t^2 (\int_t^T \sigma_s \sigma_s^{(1)} ds) dt$.

We note that we can perform higher order approximation formulas that remain explicit. The only difference is that the number of random variables used as arguments for the Greeks will increase with each order, and it is within the set $(X_{1,T} + Y_1' + \cdots + Y_i')_{i\in\mathbb{N}}$.

4.3 Financial Modeling Consequences

For simplicity, we consider the AA model on the log-asset (an analogous statement would be available for the jump diffusion model on the asset).

The standard Gaussian framework as developed by Black-Scholes (1973) and Merton (1976) is realized by choosing a constant volatility function $\sigma(\cdot)$ (the computation is still possible for a function dependent only on time). In order to arrive at a coherent, appropriate analysis and modeling for a fixed income market (without jump) Andersen and Andreasen [9] take a parametric form for σ:

$$\sigma(t,x) = \nu(t) e^{(\beta(t)-1)x}, \tag{4.7}$$

where $\nu(t)$ the relative volatility function, $\beta(t)$ is a time-dependent constant elasticity of variance (CEV). Piterbarg[3] [91] uses the same form but applies it to Power Reverse Dual Currency swaps in order to handle the skew for the FX.

Because of $\mu(t,x) = \lambda(1 - e^{\eta_J + \frac{\gamma_J^2}{2}}) - \frac{\sigma^2(t,x)}{2}$, the approximation formula (4.6) depends only on $\sigma(t,x_0), \sigma^{(1)}(t,x_0), \lambda, \eta_J$ and γ_J. The volatility given in equation (4.7) may generate all possible

[3] If σ_{Pit} is the local volatility used in [91] and $L(t) = e^{\int_0^t (r(u)-q(u))du}$, one has $\sigma(t,x) = \sigma_{Pit}(t, L_t e^x)$.

values of the following time-dependent functions $\sigma(t,x_0) = \nu(t)e^{(\beta(t)-1)x_0}$ and $\sigma^{(1)}(t,x_0) = (\beta(t)-1)\nu(t)e^{(\beta(t)-1)x_0}$, because it has two degrees of freedom $\nu(t)$ and $\beta(t)$. So this kind of volatility potentially creates all attainable prices in this class of models, and thus all attainable Black-Scholes smiles. This justifies interest in CEV-type volatility (4.7).

Attainable Black-Scholes smiles using the model. Can we predict the general form of the smiles generated by this model?

- For short maturity: using our approach, the model is close to the Merton model related to X_T^M. Therefore, the shape of implied volatilities forms a smile centered on a point close to the money, which is on the left when $\eta_J + \frac{\gamma_J^2}{2} > 0$ (on the right when $\eta_J + \frac{\gamma_J^2}{2} < 0$).
  Formal Proof: Using the approximation formula, the correction terms are $O(T)$. So when T decreases to zero, the price converges to the Merton price. The second statement is easy to check. One can follow the approach of ([84] or [50] in Chapter 5 page 62 Equation (5.10)) using characteristic functions, or can prove it directly using some derivations of the Merton formula [85].

- For long maturity: the smile becomes a skew which is due to the local volatility function (because the smile for the Merton model flattens for long maturity).

4.4 Numerical Experiments

In this section, we give details of the implementation for the approximation (4.6) and illustrate the accuracy of our formula. After that, a generic bootstrap algorithm for calibration purposes is derived. Finally, a numerical application of this algorithm is applied to market data (currency options).

4.4.1 Numerical Implementation

The case of time homogeneous parameters $\sigma_t, \sigma_t^{(1)}, \mu_t$ and $\mu_t^{(1)}$ gives us the coefficients α and β exactly (see their expressions in Theorem 4.2.1). Indeed,in the case of constant parameters (we denote $\sigma_t \equiv \sigma$, $\sigma_t^{(1)} \equiv \sigma^{(1)}$,the same notation holds for μ), one obtains that:

$$
\begin{aligned}
\alpha_{1,T} =& \frac{\mu\mu^{(1)}T^2}{2},\\
\alpha_{2,T} =& (\frac{\sigma^2\mu^{(1)}}{2} + \frac{\mu\sigma\sigma^{(1)}}{2})T^2,\\
\alpha_{3,T} =& \frac{\sigma^3\sigma^{(1)}T^2}{2},\\
\beta_{1,T} =& \frac{\lambda\eta_J\mu^{(1)}T}{2},\\
\beta_{2,T} =& \frac{\lambda(\gamma_J^2\mu^{(1)} + \eta_J\sigma_t\sigma^{(1)})T}{2},\\
\beta_{3,T} =& \frac{\lambda\gamma_J^2\sigma\sigma^{(1)}T}{2}.
\end{aligned}
$$

In addition, when these parameters are time-dependent, there are two cases.

- Either the data are smooth. In this case, we use a Gauss-Legendre quadrature formula (see Chapter 4 section 5 page 151 in [95]) for the calculation of the coefficients α and β.

- Or the data are piecewise constant. In this case, we can give explicit expressions of α and β in terms of the piecewise constant data. Let $T_0 = 0 \le T_1 \le \cdots \le T_n = T$ such that $\sigma_t, \sigma_t^{(1)}, \mu_t$ and $\mu_t^{(1)}$ are constant at each interval $]T_i, T_{i+1}]$ and are equal respectively to $\sigma_{T_{i+1}}, \sigma_{T_{i+1}}^{(1)}, \mu_{T_{i+1}}$ and $\mu_{T_{i+1}}^{(1)}$. Before giving the recursive formula, we need to introduce the following functions: $\omega_{1,t} = \int_0^t \sigma_s^2 ds, \omega_{2,t} = \int_0^t \mu_s ds$.

Proposition 4.4.1. ***Recursive formula.***
For piecewise constant coefficients, one has:

$$\begin{aligned}
\alpha_{1,T_{i+1}} =& \alpha_{1,T_i} + (T_{i+1} - T_i)\mu_{T_{i+1}}^{(1)}\omega_{2,T_i} + \frac{(T_{i+1} - T_i)^2}{2}\mu_{T_{i+1}}\mu_{T_{i+1}}^{(1)},\\
\alpha_{2,T_{i+1}} =& \alpha_{2,T_i} + (T_{i+1} - T_i)(\mu_{T_{i+1}}^{(1)}\omega_{1,T_i} + \sigma_{T_{i+1}}\sigma_{T_{i+1}}^{(1)}\omega_{2,T_i})\\
&+ \frac{(T_{i+1} - T_i)^2}{2}(\sigma_{T_{i+1}}^2\mu_{T_{i+1}}^{(1)} + \mu_{T_{i+1}}\sigma_{T_{i+1}}\sigma_{T_{i+1}}^{(1)}),\\
\alpha_{3,T_{i+1}} =& \alpha_{3,T_i} + (T_{i+1} - T_i)\sigma_{T_{i+1}}\sigma_{T_{i+1}}^{(1)}\omega_{1,T_i} + \frac{(T_{i+1} - T_i)^2}{2}\sigma_{T_{i+1}}^3\sigma_{T_{i+1}}^{(1)},\\
\beta_{1,T_{i+1}} =& \beta_{1,T_i} + \lambda\eta_J\frac{(T_{i+1}^2 - T_i^2)}{2}\mu_{T_{i+1}}^{(1)},\\
\beta_{2,T_{i+1}} =& \beta_{2,T_i} + \lambda\frac{(T_{i+1}^2 - T_i^2)}{2}(\gamma_J\mu_{T_{i+1}}^{(1)} + \eta_J\sigma_{T_{i+1}}\sigma_{T_{i+1}}^{(1)}),\\
\beta_{3,T_{i+1}} =& \beta_{3,T_i} + \lambda\gamma_J\frac{(T_{i+1}^2 - T_i^2)}{2}\sigma_{T_{i+1}}\sigma_{T_{i+1}}^{(1)},\\
\omega_{1,T_{i+1}} =& \omega_{1,T_i} + (T_{i+1} - T_i)\sigma_{T_{i+1}}^2,\\
\omega_{2,T_{i+1}} =& \omega_{2,T_i} + (T_{i+1} - T_i)\mu_{T_{i+1}}.
\end{aligned}$$

Proof. According to Theorem 4.2.1, one has:

$$\begin{aligned}
\alpha_{1,T_{i+1}} =& \int_0^{T_i} \mu_t(\int_t^{T_{i+1}} \mu_s^{(1)} ds)dt + \int_{T_i}^{T_{i+1}} \mu_t(\int_t^{T_{i+1}} \mu_s^{(1)} ds)dt\\
=& \alpha_{1,T_i} + \int_0^{T_i} \mu_t(\int_{T_i}^{T_{i+1}} \mu_s^{(1)} ds)dt + \int_{T_i}^{T_{i+1}} \mu_t(\int_t^{T_{i+1}} \mu_s^{(1)} ds)dt\\
=& \alpha_{1,T_i} + (\int_{T_i}^{T_{i+1}} \mu_s^{(1)} ds)\int_0^{T_i} \mu_t dt + \int_{T_i}^{T_{i+1}} \mu_t(\int_t^{T_{i+1}} \mu_s^{(1)} ds)dt\\
=& \alpha_{1,T_i} + (T_{i+1} - T_i)\mu_{T_{i+1}}^{(1)}\omega_{2,T_i} + \frac{(T_{i+1} - T_i)^2}{2}\mu_{T_{i+1}}\mu_{T_{i+1}}^{(1)}.
\end{aligned}$$

The other terms are calculated analogously. □

4.4.2 Accuracy of the approximation

Here, we give a short example of the performance of our method. The jump parameters have been set to: $\lambda = 30\%, \eta_J = -8\%, \gamma_J = 35\%$. These parameters are not small, especially for the jump intensity λ and the jump volatility γ_J. The piecewise constant functions ν and β defined in (4.7) are equal respectively at each interval of the form $[\frac{i}{20}, \frac{i+1}{20}]$ to $25\% - i \times 0.11\%$ and $100\% - i \times 0.75\%$. The spot, the risk-free rate and the dividend yield are set respectively to $100, 4\%$ and 0%.

We observe in the table below that the errors of implied Black-Scholes volatilities between our approximation and the price calculated using a PIDE method do not exceed **2 bp** for a large range of strikes and maturities. The computational time of our formula is less than four milliseconds on a 2.6 GHz Pentium PC. The accuracy of our formula turns out to be excellent.

Table 4.1: Error in implied Black-Scholes volatilities (in bp=0.01%) between the approximation formula and the PIDE method expressed as a function of maturities in fractions of years and relative strikes.

T/K	70%	85%	100%	120%	150%
3M	0.02	-0.03	-0.92	-0.07	-0.12
1Y	0.04	0.06	0.15	-0.11	0.01
3Y	0.22	-0.23	0.11	0.41	0.31
5Y	1.39	1.06	-0.01	1.85	1.76

4.4.3 Calibration issues

For this kind of model (AA model on the log asset or on the asset itself), calibration is still challenging as this model has no analytical formula. We can still perform a numerical calibration using the forward PIDE as explained in [8], but the time of calibration remains quite long (about one minute). With our approach, we can shorten the duration of calibration to less than one second, because our computation of the model price takes four milliseconds as previously mentioned. We achieve that by a simple bootstrapping algorithm using the path dependent formula.

Bootstrap algorithm for piecewise data . Suppose that we want to fit option prices for n maturities $T_0 = 0 \leq T_1 \leq \cdots \leq T_n$ and m strikes $K_1, \cdots, K_m$. First, we search the parameters λ, η_J and γ_J with best fit. At each interval $]T_{i-1}, T_i]$, the data σ, $\sigma^{(1)}$, μ and $\mu^{(1)}$ are constant, equal respectively to σ_{T_i}, $\sigma^{(1)}_{T_i}$, μ_{T_i} and $\mu^{(1)}_{T_i}$, and depending on the vector $\chi_i = (\nu(T_i), \beta(T_i))$ (see formula 4.7). Starting at $i = 1$, we express the coefficients α_{j,T_i} and β_{j,T_i} as a function of χ_i, recursively using Proposition (4.4.1). We apply a local minimization algorithm (for instance, the Levenberg-Marquardt as described in Chapter 15 section 5 page 683 in [95]) in order to fit the implied volatilities for all strikes $K_1, \cdots, K_m$ at maturity T_i using our approximation (4.6). Once the vector χ_i is found, we go to the next step $i+1$, update α and β and compute χ_{i+1}.

This calibration procedure is not completely safe. Sometimes we encounter instability problems. The final parameters depend on the initial guess. Moreover, there are many local minima. To avoid these problems, we could use a regularization method based on relative entropy (see [31]), but these issues are not in direct relation with the accuracy of our formula. We think that the set of calibrated options (call/put) does not contain enough information on the future volatility to ensure a good calibration. Therefore, it is presumably worth including volatility options in the set of calibrated instruments. This is a topic for further research.

Calibration results. Here, we calibrate the EUR/USD exchange rate. The surface of implied Black-Scholes volatility is given in table 5.2.

Table 4.2: Implied Black-Scholes volatilities for the EUR/USD rate expressed as a function of maturities in fractions of years and relative strikes. The spot is equal to 1.54.

T/K	92%	96%	100%	108%
6M	10.82%	10.65%	10.53%	10.56%
1Y	10.84%	10.70%	10.63%	10.66%
1.5Y	10.71%	10.60%	10.56%	10.58%
2Y	10.60%	10.48%	10.46%	10.47%

The jump parameters for the calibrated model are $\lambda = 1.21\%$, $\eta_J = -19.07\%$ and $\gamma_J = 40.30\%$. The diffusion parameters ν and β for the calibrated model are given in table 5.3. These values are realistic. The errors between the implied volatilities generated by the calibrated model and the market data are given in table 5.4.

Table 4.3: Calibrated values of the piecewise constant functions ν and β.

T	ν	β
6M	10.31%	98.81%
1Y	10.27%	100%
1.5Y	9.90%	100%
2Y	9.43%	100%

Table 4.4: Errors between implied Black-Scholes volatilities for the EUR/USD rate and those calculated within the calibrated model (in bp) expressed as a function of maturities in fractions of years and relative strikes. The spot is equal to 1.54.

T/K	92%	96%	100%	108%
6M	-4	3	-1	-3
1Y	2	1	0	2
1.5Y	-1	-3	-2	1
2Y	2	-1	1	4

The errors show that our model is a good model for the FX rate EUR/USD. Within our relevant algorithm, we are able to fit a 4×4 grid of quoted prices in less than **1 s**.

4.5 Error Analysis

This section is devoted to the mathematical justification of Theorem 4.2.1 and to the statement and proofs of upper bounds for the error term in (4.6). For this, the analysis differs according to the payoff smoothness (smooth, vanilla or binary). We start with the smooth case (subsection 4.5.1), which is less technical. Then, we handle the two other cases (call/put and binary options), which requires the use of Malliavin calculus.

Throughout these computations, we aim at emphasizing the dependence of error upper bounds in terms of: the constants M_0, M_1 and M_J defined in (4.5), the jump frequency λ and the maturity T, in order to support the heuristic choice of the model proxy (see the discussion in the introduction).

Additional notation.

- *About floating constants and upper bounds.* In the following statements and proofs, for the upper bounds we use numerous constants, that are not relabelled during the computations. We simply use the unique notation
$$A \le_c B$$
to assert that $A \le cB$, where c is a positive constant depending on the model parameters M_0, M_1, M_J, λ, T, C_E (defined in assumption (E)) and on other universal constants. The constant c remains bounded when the model parameters go to 0, and it is uniform w.r.t. the parameter $\varepsilon \in [0,1]$. When informative, we make clear the dependence of upper bounds w.r.t. M_0, M_1, M_J, λ and T.

- *Miscellaneous.* As usual, the $\mathbf{L}_p$-norm of a real random variable Z is denoted by $\|Z\|_p = [\mathbb{E}|Z|^p]^{1/p}$. In the proofs, the derivatives of the parameterized process X^ε are useful: they are defined by $X^\varepsilon_{i,t} = \frac{\partial^i X^\varepsilon_t}{\partial \varepsilon^i}$.

4.5.1 Error analysis for smooth payoff (under (H_1))

We begin our error analysis with the case of smooth payoff ($h \in \mathscr{C}_0^\infty(\mathbb{R})$).

Theorem 4.5.1. ***Error for smooth payoff.*** *Assume that (R_4) holds and that the payoff function h fulfills Assumption (H_1). Then the error term in Theorem 4.2.1 satisfies the following estimate:*

$$|\text{Error}| \le_c \sup_{j=1,2} |h^{(j)}|_\infty (M_1\sqrt{T})\big((M_0\sqrt{T})^2 + M_J^2\sqrt{\lambda T}\big). \tag{4.8}$$

Let us briefly comment on the upper bound, making reference to the introduction. If the functions $\sigma(\cdot)$ and $\mu(\cdot)$ are only time dependent ($M_1 = 0$), the approximation formula (4.6) is exact (the model and the proxy coincide). If they do not vary much w.r.t. x (M_1 is small), the accuracy is still good in view of (4.8). If the coefficients $\sigma(\cdot), \mu(\cdot)$ and their derivatives and the jump size parameters are all small, the formula becomes very accurate. For instance, in a multiplicative case where $\sigma(t,x) = \Delta s(t,x)$, $\mu(t,x) = \Delta m(t,x)$ and $|\eta_J| + \gamma_J \le \Delta$ for a small parameter Δ, it readily follows that $M_1, M_0, M_J = O(\Delta)$. Thus
$$|\text{Error}| = O\big(\Delta^3 T[\sqrt{T} + \sqrt{\lambda}]\big).$$
Consequently, we may refer to the formula (4.6) in Theorem 4.2.1 as an approximation of order 2 w.r.t. the amplitudes of the data (with error terms of order 3).

These features arise similarly for the other examples of payoff smoothness.

Proof. It is divided into several steps. First, we write the SDEs satisfied by the three first derivatives of X_t^ε w.r.t. ε. Second, we give tight $\mathbf{L}_p$ upper bounds on these derivatives. Finally, we combine these estimates with our smart expansion to complete the proof of Theorem 4.5.1.

Step 1. Differentiation of X^ε. Under (R_4), almost surely X_t^ε is C^3 w.r.t ε for any t (see Theorem 2.3 in [48]) and the derivatives are obtained by successive differentiations of the initial SDE (4.3). Thus, direct computations lead to

$$\begin{aligned} dX_{1,t}^\varepsilon =&\sigma_t(X_{t-}^\varepsilon)dW_t+\mu_t(X_{t-}^\varepsilon)dt+dJ_t+\varepsilon X_{1,t-}^\varepsilon(\sigma_t^{(1)}(X_{t-}^\varepsilon)dW_t+\mu_t^{(1)}(X_{t-}^\varepsilon)dt), \quad (4.9)\\ dX_{2,t}^\varepsilon =&[2X_{1,t-}^\varepsilon\sigma_t^{(1)}(X_{t-}^\varepsilon)+\varepsilon(X_{1,t-}^\varepsilon)^2\sigma_t^{(2)}(X_{t-}^\varepsilon)]dW_t\\ &+[2X_{1,t-}^\varepsilon\mu_t^{(1)}(X_{t-}^\varepsilon)+\varepsilon(X_{1,t-}^\varepsilon)^2\mu_t^{(2)}(X_{t-}^\varepsilon)]dt\\ &+\varepsilon X_{2,t-}^\varepsilon(\sigma_t^{(1)}(X_{t-}^\varepsilon)dW_t+\mu_t^{(1)}(X_{t-}^\varepsilon)dt), \quad (4.10)\\ dX_{3,t}^\varepsilon =&[3X_{2,t-}^\varepsilon\sigma_t^{(1)}(X_{t-}^\varepsilon)+3(X_{1,t-}^\varepsilon)^2\sigma_t^{(2)}(X_{t-}^\varepsilon)+3\varepsilon X_{1,t-}^\varepsilon X_{2,t-}^\varepsilon\sigma_t^{(2)}(X_{t-}^\varepsilon)\\ &+\varepsilon(X_{1,t-}^\varepsilon)^3\sigma_t^{(3)}(X_{t-}^\varepsilon)]dW_t+[3X_{2,t-}^\varepsilon\mu_t^{(1)}(X_{t-}^\varepsilon)+3(X_{1,t-}^\varepsilon)^2\mu_t^{(2)}(X_{t-}^\varepsilon)\\ &+3\varepsilon X_{1,t-}^\varepsilon X_{2,t-}^\varepsilon\mu_t^{(2)}(X_{t-}^\varepsilon)+\varepsilon(X_{1,t-}^\varepsilon)^3\mu_t^{(3)}(X_{t-}^\varepsilon)]dt\\ &+\varepsilon X_{3,t-}^\varepsilon(\sigma_t^{(1)}(X_{t-}^\varepsilon)dW_t+\mu_t^{(1)}(X_{t-}^\varepsilon)dt). \quad (4.11) \end{aligned}$$

Their initial conditions are all equal to 0. Notice that unlike X^ε and X_1^ε, the processes X_2^ε and X_3^ε are continuous.

Step 2. Tight upper bounds. We aim at proving the following estimates for any $p\geq 2$:

$$\mathbb{E}|X_{1,t}^\varepsilon|^p \leq_c (M_0\sqrt{T})^p+M_J^p\lambda T, \quad (4.12)$$
$$\mathbb{E}|X_{2,t}^\varepsilon|^p \leq_c (M_1\sqrt{T})^p\big((M_0\sqrt{T})^p+M_J^p\lambda T\big), \quad (4.13)$$
$$\mathbb{E}|X_{3,t}^\varepsilon|^p \leq_c (M_1\sqrt{T})^p\big((M_0\sqrt{T})^{2p}+M_J^{2p}\lambda T\big), \quad (4.14)$$

uniformly for $t\leq T$.
The existence of any moment is easy to establish, but here, we emphasize the dependence of the upper bounds w.r.t. the constants M_0, M_1, M_J, λ and T. Let us first prove the inequality (4.12). From (4.9), apply Lemma 4.6.5 to the jump component and Burkholder-Davis-Gundy inequalities to the Brownian part, to deduce

$$\begin{aligned} \mathbb{E}|X_{1,t}^\varepsilon|^p \leq_c& t^{p/2-1}\int_0^t\mathbb{E}|\sigma_s(X_s^\varepsilon)|^p ds+t^{p-1}\int_0^t\mathbb{E}|\mu_s(X_s^\varepsilon)|^p ds+M_J^p\lambda t\\ &+t^{p/2-1}\int_0^t\mathbb{E}|X_{1,s}^\varepsilon\sigma_s^{(1)}(X_s^\varepsilon)|^p ds+t^{p-1}\int_0^t\mathbb{E}|X_{1,s}^\varepsilon\mu_s^{(1)}(X_s^\varepsilon)|^p ds\\ \leq_c& T^{p/2}M_0^p+M_J^p\lambda T+T^{p/2-1}M_1^p\int_0^t\mathbb{E}|X_{1,s}^\varepsilon|^p ds. \end{aligned}$$

Using Gronwall's lemma, we easily complete the proof of (4.12). For the second inequality (4.13), we proceed analogously and we obtain:

$$\mathbb{E}|X_{2,t}^\varepsilon|^p \leq_c T^{p/2}M_1^p(\sup_{s\leq t}\mathbb{E}|X_{1,s}^\varepsilon|^p+\sup_{s\leq t}\mathbb{E}|X_{1,s}^\varepsilon|^{2p}).$$

Thus, plugging the estimate (4.12) into the previous inequality directly leads to (4.13). Now let us prove the inequality (4.14). As before, apply BDG inequalities combined with Gronwall's lemma to obtain that

$$\mathbb{E}|X^{\varepsilon}_{3,t}|^p \leq_c T^{p/2} M_1^p (\sup_{s\leq t}\mathbb{E}|X^{\varepsilon}_{2,s}|^p + \sup_{s\leq t}\mathbb{E}|X^{\varepsilon}_{1,s}|^{2p} + \sup_{s\leq t}\mathbb{E}|X^{\varepsilon}_{1,s}X^{\varepsilon}_{2,s}|^p + \sup_{s\leq t}\mathbb{E}|X^{\varepsilon}_{1,s}|^{3p}).$$

Use $\mathbb{E}|X^{\varepsilon}_{1,s}X^{\varepsilon}_{2,s}|^p \leq \frac{1}{2}(\mathbb{E}|X^{\varepsilon}_{1,s}|^{2p} + \mathbb{E}|X^{\varepsilon}_{2,s}|^{2p})$ and the previous inequalities (4.12-4.13). Then bringing together different contributions easily leads to the required estimate (4.14).

Step 3. Completion of the proof. We follow the formal computations done at the beginning of Section 4.2, but more carefully. Let us introduce

$$\overline{X}_{2,T} = \int_0^1 X^{\varepsilon}_{2,T}(1-\varepsilon)d\varepsilon, \quad \overline{X}_{3,T} = \int_0^1 X^{\varepsilon}_{3,T}\frac{(1-\varepsilon)^2}{2}d\varepsilon. \tag{4.15}$$

Then applications of Taylor expansions of X^{ε}_T at $\varepsilon = 0$ readily give these equalities:

$$X_T = X_T^M + \overline{X}_{2,T}, \quad X_T = X_T^M + \frac{1}{2}X_{2,T} + \overline{X}_{3,T}$$

where we have used $X_T^M = x_0 + X_{1,T}$. Thus a second order Taylor expansion of h at point X_T^M writes

$$\begin{aligned}\mathbb{E}[h(X_T^1)] =& \mathbb{E}[h(X_T^M + \frac{X_{2,T}}{2} + \overline{X}_{3,T})] \\ =& \mathbb{E}[h(X_T^M)] + \mathbb{E}[h^{(1)}(X_T^M)\frac{X_{2,T}}{2}] + \mathbb{E}[h^{(1)}(X_T^M)\overline{X}_{3,T}] \\ &+ \int_0^1 \mathbb{E}[h^{(2)}((1-v)X_T^M + vX_T)(\overline{X}_{2,T})^2](1-v)dv.\end{aligned}$$

This proves that the Error term in (4.6) for smooth payoff equals

$$\text{Error} = \mathbb{E}[h^{(1)}(X_T^M)\overline{X}_{3,T}] + \int_0^1 \mathbb{E}[h^{(2)}((1-v)X_T^M + vX_T)(\overline{X}_{2,T})^2](1-v)dv. \tag{4.16}$$

Then it readily follows that

$$|\text{Error}| \leq_c |h^{(1)}|_\infty \sup_{\varepsilon\in[0,1]} (\mathbb{E}|X^{\varepsilon}_{3,T}|^2)^{\frac{1}{2}} + |h^{(2)}|_\infty \sup_{\varepsilon\in[0,1]} \mathbb{E}|X^{\varepsilon}_{2,T}|^2.$$

It is now straightforward to obtain Theorem 4.5.1, by using estimates (4.13-4.14) with $p = 2$. □

A careful inspection of the previous proof shows that assumption (R_3) is sufficient to derive the error estimate (4.8).

4.5.2 Error analysis for vanilla payoff (under (H_2))

This case has practical importance, because it includes call/put options. Regarding the error estimates related to Theorem 4.2.1, we have paved the way with the case of smooth payoff. Nevertheless, there are some technical differences. The main one is that our previous proof represents the error in terms of the second derivative of the payoff, which is meaningless here. The additional ingredient is the Malliavin calculus integration by parts formula to avoid this second derivative appearing. We now state our main result when the payoff h is almost everywhere differentiable (with sub-exponential growth conditions).

Theorem 4.5.2. ***Error for vanilla payoff.*** *Assume that* (R_4) *and* (E) *hold, and that the payoff function h fulfills Assumption* (H_2)*. Then the error term in Theorem 4.2.1 satisfies to the following estimate:*

$$|\text{Error}| \leq_c \big(\|h^{(1)}(X_T^M)\|_2 + \int_0^1 \|h^{(1)}((1-v)X_T^M + vX_T)\|_3 dv\big)$$
$$\times \frac{M_0}{\sigma_{inf}}(M_1\sqrt{T})\big((M_0\sqrt{T})^2 + M_J^2\sqrt{\lambda T}\big). \tag{4.17}$$

The shape of the upper bound regarding h is used for convenience in the proof. In view of the growth condition on $h^{(1)}$, the two first terms depending on $h^{(1)}$ are finite and uniformly bounded as M_0, M_1, M_J, λ and$T$ go to 0.

Analogously to the smooth case (Theorem 4.5.1), the approximation error in (4.6) is of order 3 w.r.t. the amplitudes of the model data, meaning that (4.6) is a second order approximation formula.

Proof. We split the proof into several steps. First, we assume that the payoff is smooth and we establish estimates that depend only on $h^{(1)}$, the first derivative of h. For this, we need extra tools from Malliavin calculus, together with tight estimates on the Malliavin derivatives of the parameterized process. Then, we apply a density argument to approximate h under (H_2) by a sequence of smooth payoffs.

Step 1. Malliavin calculus. For the usual Malliavin calculus on the Wiener space, we refer to Nualart [88]. But our case is slighty different because of jumps. However, in the following, our Malliavin differentiation is w.r.t. the Brownian motion W and not w.r.t. the Poisson measure κ. Hence formally, it is performed by leaving the jump component fixed, computing the Malliavin derivatives or integration by parts w.r.t. W, and then integrating out w.r.t. the jumps. This principle has been formalized in several papers, for instance in [25] Section 3. We briefly recall a few facts using their notations.
The model jumps are associated with the Poisson measure κ, with intensity $g_{\eta_J,\gamma_J}(x)dx\,\lambda dt$, where g_{η_J,γ_J} is the Gaussian density on $\mathbb{R}$ with mean η_J and variance γ_J^2. The set of integer-valued measures on $[0,T]\times\mathbb{R}$ is denoted by Ω_κ. For $l(.) \in \mathscr{L} = \mathbf{L}_2([0,T],\mathbb{R})$, the Wiener stochastic integral $\int_0^T l(t)dW_t$ is denoted by $W(l)$. Let $\mathscr{S}$ denote the class of simple random variables of the form $F = f(W(l_1),\dots,W(l_N);\kappa)$ where $N \geq 1$, $(l_1,\dots,l_N) \in \mathscr{L}^N$, $f : \mathbb{R}^N \times \Omega_\kappa \mapsto \mathbb{R}$ is bounded and infinitely differentiable w.r.t. its N first components (with bounded derivatives). We denote by D the Malliavin derivative operator with respect to the Brownian motion. For $F \in \mathscr{S}$, it is defined as the $\mathscr{L}$-valued random variable given by

$$D_tF = \sum_{i=1}^N \partial_{x_i} f(W(l_1),\dots,W(l_N);\kappa)l_i(t).$$

The operator D is closable as an operator from $\mathbf{L}_p(\Omega)$ to $\mathbf{L}_p(\Omega,\mathscr{L})$, for any $p \geq 1$. Its domain is denoted by $\mathbb{D}^{1,p}$ with respect to the norm $\|\cdot\|_{1,p}$ given by $\|F\|_{1,p}^p = \mathbb{E}|F|^p + \mathbb{E}(\int_0^T |D_tF|^2 dt)^{p/2}$. We can define the iteration of the operator D in such a way that for a smooth random variable F, the derivative D^kF is a random variable with values in $\mathscr{L}^{\otimes k}$. As in the case $k=1$, the operator D^k is closable from $\mathscr{S} \subset \mathbf{L}_p(\Omega)$ into $\mathbf{L}_p(\Omega;\mathscr{L}^{\otimes k})$, $p \geq 1$. Its domain is denoted by $\mathbb{D}^{k,p}$ w.r.t. the norm $\|F\|_{k,p} = [\mathbb{E}|F|^p + \sum_{j=1}^k \mathbb{E}(\|D^jF\|_{\mathscr{L}^{\otimes j}}^p)]^{1/p}$. With this construction, the operator D enjoys the same properties as the usual operator on the Wiener space (see [25] for more details). This justifies, in the case under study, the application of the usual results established without jumps (in particular the integration by parts formula and the related general $\mathbf{L}_p$ estimates, see the proof of Lemma 4.5.1).

Step 2. Estimates of Malliavin derivatives. Under our regularity assumptions (R_4), we know that for any $t \le T$, any $\varepsilon \in [0,1]$ and any $p \ge 1$, we have $X_t^\varepsilon \in \mathbb{D}^{4,p}$, $X_{1,t}^\varepsilon \in \mathbb{D}^{3,p}$, $X_{2,t}^\varepsilon \in \mathbb{D}^{2,p}$, $X_{3,t}^\varepsilon \in \mathbb{D}^{1,p}$ (see the arguments in [25]). Actually, we aim at proving the following tight estimates for any $p \ge 2$:

$$\mathbb{E}|D_r X_t^\varepsilon|^p \le_c |\sigma|_\infty^p, \tag{4.18}$$
$$\mathbb{E}|D_r X_t^M|^p \le_c |\sigma|_\infty^p, \tag{4.19}$$
$$\mathbb{E}|D_{r,s}^2 X_t^\varepsilon|^p \le_c |\sigma|_\infty^p M_1^p, \tag{4.20}$$
$$\mathbb{E}|D_{r,s,u}^3 X_t^\varepsilon|^p \le_c |\sigma|_\infty^p M_1^{2p}, \tag{4.21}$$
$$\mathbb{E}|D_r X_{1,t}^\varepsilon|^p \le_c M_0^p, \tag{4.22}$$
$$\mathbb{E}|D_{r,s}^2 X_{1,t}^\varepsilon|^p \le_c M_0^p M_1^p, \tag{4.23}$$
$$\mathbb{E}|D_r X_{2,t}^\varepsilon|^p \le_c M_1^p\big((M_0\sqrt{T})^p + M_J^p \lambda T\big), \tag{4.24}$$
$$\mathbb{E}|D_{r,s}^2 X_{2,t}^\varepsilon|^p \le_c M_0^p M_1^p, \tag{4.25}$$
$$\mathbb{E}|D_r X_{3,t}^\varepsilon|^p \le_c M_1^p\big((M_0\sqrt{T})^{2p} + M_J^{2p} \lambda T\big), \tag{4.26}$$

uniformly in $(r,s,t,u) \in [0,T]^4$ and $\varepsilon \in [0,1]$. Here again, the existence of any moment is easy to establish and we will skip the details. We prefer to focus on the dependence of the upper bounds w.r.t. M_0, M_1, M_J, λ and T. The bounds (4.23-4.25-4.26) are not used for vanilla payoffs, but only for binary ones.

Proof of (4.18). For $r > t$, $D_r X_t^\varepsilon = 0$. Now take $r \le t$, in this case $(D_r X_t^\varepsilon)_{r \le t \le T}$ solves the following SDE (see [25]):

$$D_r X_t^\varepsilon = \varepsilon \sigma_r(X_{r-}^\varepsilon) + \int_r^t D_r X_{u-}^\varepsilon \varepsilon(\sigma_u^{(1)}(X_{u-}^\varepsilon) dW_u + \mu_u^{(1)}(X_{u-}^\varepsilon) du), \tag{4.27}$$

which defines a continuous process. Now, we proceed as in the proof of (4.12-4.13-4.14), combining BDG inequalities and Gronwall's lemma. This gives

$$\mathbb{E}|D_r X_t^\varepsilon|^p \le_c \mathbb{E}|\sigma_r(X_{r-}^\varepsilon)|^p + T^{p/2-1} M_1^p \int_0^T \mathbb{E}|D_r X_u^\varepsilon|^p du \le_c |\sigma|_\infty^p,$$

and proves the announced inequality. Besides, in light of (4.2) one has $D_r X_t^M = \mathbf{1}_{r \le t} \sigma_r$, which directly gives (4.19).

Proof of (4.20). Take for instance $r < s \le T$, the other cases are handled in the same way. We have

$$\begin{aligned} D_{r,s}^2 X_t^\varepsilon =& \varepsilon D_r X_{s-}^\varepsilon \sigma_s^{(1)}(X_{s-}^\varepsilon) \\ &+ \int_s^t D_s X_{u-}^\varepsilon \varepsilon(D_r X_{u-}^\varepsilon \sigma_u^{(2)}(X_{u-}^\varepsilon) dW_u + D_r X_{u-}^\varepsilon \mu_u^{(2)}(X_{u-}^\varepsilon) du) \\ &+ \int_s^t D_{r,s}^2 X_{u-}^\varepsilon \varepsilon(\sigma_u^{(1)}(X_{u-}^\varepsilon) dW_u + \mu_u^{(1)}(X_{u-}^\varepsilon) du), \end{aligned}$$

which implies, in particular, that $t \mapsto D_{r,s}^2 X_t^\varepsilon$ is continuous. It readily follows that

$$\begin{aligned} \mathbb{E}|D_{r,s}^2 X_t^\varepsilon|^p &\le_c M_1^p \mathbb{E}|D_r X_s^\varepsilon|^p + T^{p/2} M_1^p \sup_{r<s\le u\le T} \mathbb{E}|D_s X_u^\varepsilon D_r X_u^\varepsilon|^p \\ &\le_c M_1^p \mathbb{E}|D_r X_s^\varepsilon|^p + T^{p/2} M_1^p \sup_{r<s\le u\le T} (\mathbb{E}|D_s X_u^\varepsilon|^{2p} + \mathbb{E}|D_r X_u^\varepsilon|^{2p}) \le_c |\sigma|_\infty^p M_1^p \end{aligned}$$

where we have used the Young inequality $ab \leq \frac{1}{2}(a^2+b^2)$ in the second line and (4.18) in the last inequality. The estimate (4.21) can be established in the same way.
Proof of (4.22). We only consider $r \leq t$. Here one has

$$\begin{aligned} D_r X^{\varepsilon}_{1,t} =& \sigma_r(X^{\varepsilon}_{r-}) + \varepsilon X^{\varepsilon}_{1,r-}\sigma^{(1)}_r(X^{\varepsilon}_{r-}) \\ &+ \int_r^t D_r X^{\varepsilon}_{u-}(\sigma^{(1)}_u(X^{\varepsilon}_{u-}) + \varepsilon X^{\varepsilon}_{1,u-}\sigma^{(2)}_u(X^{\varepsilon}_{u-}))dW_u \\ &+ \int_r^t D_r X^{\varepsilon}_{u-}(\mu^{(1)}_u(X^{\varepsilon}_{u-}) + \varepsilon X^{\varepsilon}_{1,u-}\mu^{(2)}_u(X^{\varepsilon}_{u-}))du \\ &+ \int_r^t D_r X^{\varepsilon}_{1,u-}\varepsilon(\sigma^{(1)}_u(X^{\varepsilon}_{u-})dW_u + \mu^{(1)}_u(X^{\varepsilon}_{u-})du). \end{aligned}$$

It readily follows that

$$\mathbb{E}|D_r X^{\varepsilon}_{1,t}|^p \leq_c \mathbb{E}|\sigma_r(X^{\varepsilon}_{r-}) + \varepsilon X^{\varepsilon}_{1,r-}\sigma^{(1)}_r(X^{\varepsilon}_{r-})|^p + T^{p/2}M_1^p \sup_{r\leq u\leq T}(\mathbb{E}|D_r X^{\varepsilon}_u|^p + \mathbb{E}|D_r X^{\varepsilon}_u X^{\varepsilon}_{1,u}|^p).$$

Since a fixed time r is equal to a jump time with null probability and thanks to the Young inequality, we obtain

$$\mathbb{E}|D_r X^{\varepsilon}_{1,t}|^p \leq_c |\sigma|^p_\infty + M_1^p \mathbb{E}|X^{\varepsilon}_{1,r}|^p + T^{p/2}M_1^p \sup_{r\leq u\leq T}(\mathbb{E}|D_r X^{\varepsilon}_u|^p + \mathbb{E}|D_r X^{\varepsilon}_u|^{2p} + \mathbb{E}|X^{\varepsilon}_{1,u}|^{2p}).$$

It remains to take advantage of the inequalities (4.12) and (4.18), and to use $|\sigma|_\infty \leq M_0$ and $M_1 \leq M_0$ to complete the proof of (4.22).
Proof of (4.23-4.24-4.25-4.26). They can be proved similarly, with long and tedious computations. Since there is no extra difficulty, we will skip further details.

Step 3. Bounding the error using only $h^{(1)}$, when h is smooth. We come back to the representation (4.16) for the error. The first term can be estimated using a Cauchy-Schwartz inequality and (4.14):

$$\begin{aligned} |\mathbb{E}[h^{(1)}(X^M_T)\overline{X}_{3,T}]| &\leq_c \|h^{(1)}(X^M_T)\|_2 \sup_{\varepsilon\in[0,1]} \|X^{\varepsilon}_{3,T}\|_2 \\ &\leq_c \|h^{(1)}(X^M_T)\|_2 (M_1\sqrt{T})\big((M_0\sqrt{T})^2 + M_J^2\sqrt{\lambda T}\big). \end{aligned}$$

This fits the required upper bound (4.17) well, because $M_0 \geq \sigma_{inf}$.

The second term in (4.16) requires a little extra work because of $h^{(2)}$. For this, we state a lemma, proof of which is given at the end.

Lemma 4.5.1. *Assume (E) and (R_3). Let Z belong to $\cap_{p\geq 1}\mathbb{D}^{2,p}$. For any $v \in [0,1]$, for $k = 1,2$, there exists a random variable Z^v_k in any L_p ($p \geq 1$) such that for any function $l \in \mathscr{C}^\infty_0(\mathbb{R})$, one has*

$$\mathbb{E}[l^{(k)}(vX_T + (1-v)X^M_T)Z] = \mathbb{E}[l(vX_T + (1-v)X^M_T)Z^v_k].$$

Moreover, one has $\|Z^v_k\|_p \leq_c \frac{\|Z\|_{k,p+\frac{1}{2}}}{(\sigma_{inf}\sqrt{T})^k}$, *uniformly in v.*

Apply this Lemma with $k=1$ and $Z = (\overline{X}_{2,T})^2$ defined in (4.15). From the estimates (4.13-4.24), we readily obtain

$$\|Z^v_1\|_{\frac{3}{2}} \leq_c \frac{(M_1\sqrt{T})^2\big((M_0\sqrt{T})^2 + M_J^2\sqrt{\lambda T}\big)}{\sigma_{inf}\sqrt{T}}.$$

We have proved the upper bound (4.17).

Step 4. Bounding the error under the sole assumption (H_2). So far, our error estimates depend on $h^{(1)}$, but they have been established for smooth payoffs h. It remains to justify that the error upper bound still holds for payoffs that are only almost everywhere differentiable (assumption (H_2)). We argue by regularization, which is somewhat standard but a bit tricky here. We follow the proof of [55].

Denote by ρ the measure defined by $\int_{\mathbb{R}} g(x)\rho(dx) = \mathbb{E}(g(X_T)) + \mathbb{E}(g(X_T^M)) + \mathbb{E}(g(X_T^M + Y')) + \int_0^1 \mathbb{E}(g(vX_T + (1-v)X_T^M))dv$. It is well known (see Chapter 3, Theorem 3.14 in [102] for instance) that there exists a sequence $(h_n)_{n\in\mathbb{N}}$ of smooth functions converging to h in $\mathbf{L}_3(\rho)$ as well as its first derivative, as n goes to infinity. Thus, we can pass to the limit for $\mathbb{E}(h_n(X_T))$ and $\mathbb{E}(h_n(X_T^M))$. In view of (4.17), we can also pass to the limit for the error bound. It remains to pass to the limit for the corrections terms, i.e. for the greeks $\text{Greek}_i^{h_n}(X_T^M)$ and $\text{Greek}_i^{h_n}(X_T^M + Y')$. To accomplish this, we represent them as $\mathbb{E}(h_n(X_T^M)Z_i)$ and $\mathbb{E}(h_n(X_T^M + Y')Z_i)$ using Lemma 4.5.1 with $Z = 1$. Since Z_i is in $\mathbf{L}_{3/2}$, we can pass to the limit as $n \to \infty$ to get $\mathbb{E}(h(X_T^M)Z_i) = \text{Greek}_i^h(X_T^M)$ and $\mathbb{E}(h(X_T^M + Y')Z_i) = \text{Greek}_i^h(X_T^M + Y')$. □

Proof of Lemma 4.5.1. Take $k = 1$ or 2.

Step 1. $F_v = vX_T + (1-v)X_T^M$ is a non degenerate random variable (in the Malliavin sense). Under (R_4), we know that F_v is in $\cap_{p\geq 1}\mathbb{D}^{4,p}$. One has to prove that $\gamma_{F_v} = \int_0^T (D_sF_v)^2 ds$ is almost surely positive and its inverse is in any $\mathbf{L}_p$ ($p \geq 1$). From the linear SDE (4.27) satisfied by $(D_sX_t)_{s\leq t\leq T}$, we obtain

$$\gamma_{F_v} = \int_0^T (v\sigma_s(X_{s-})e^{\int_s^T \sigma_u^{(1)}(X_{u-})dW_u + (\mu_u^{(1)} - \frac{1}{2}(\sigma_u^{(1)})^2)(X_{u-})du} + (1-v)\sigma_s(x_0))^2 ds,$$

which clearly leads to our claim. Besides, for any $p \geq 1$, we derive

$$\|\gamma_{F_v}^{-1}\|_p \leq_c (\sigma_{inf}\sqrt{T})^{-2}.$$

Step 2. Integration by Parts formula. Using Proposition 2.1.4 and Proposition 1.5.6 in [88], one gets the existence of Z_k^v in $\mathbf{L}_p$ with

$$\|Z_k^v\|_p \leq_c \|\gamma_{F_v}^{-1}\|_{k,2^k p(2p+1)}^k \|DF_v\|_{k,2^k p(2p+1)}^k \|Z\|_{k,p+\frac{1}{2}}.$$

Step 3: Upper bound of $\|DF_v\|_{k,q}, \|\gamma_{F_v}^{-1}\|_{k,q}$ for $q \geq 2$. On the one hand, using the inequalities (4.18-4.19-4.20-4.21), we easily obtain

$$\|DF_v\|_{k,q} \leq_c |\sigma|_\infty\sqrt{T}. \tag{4.28}$$

On the other hand, with the same inequalities, we get $\sup_{r\leq T}\mathbb{E}|D_r\gamma_{F_v}|^p \leq_c T^p|\sigma|_\infty^{2p}M_1^p$ and $\sup_{r,s\leq T}\mathbb{E}|D_{r,s}^2\gamma_{F_v}|^p \leq_c T^p|\sigma|_\infty^{2p}M_1^{2p}$ for any $p \geq 2$. Then, after some computations, it follows that

$$\|\gamma_{F_v}^{-1}\|_{2,p} \leq_c (\sigma_{inf}\sqrt{T})^{-2}\big(1 + \frac{M_1|\sigma|_\infty^2 T^{1/2}}{\sigma_{inf}^2} + \frac{M_1^2|\sigma|_\infty^2 T}{\sigma_{inf}^2} + \frac{M_1^2|\sigma|_\infty^4 T}{\sigma_{inf}^4}\big) \tag{4.29}$$

for any $p \geq 2$. Finally using $|\sigma|_\infty \leq C_E\sigma_{inf}$ (assumption (E)) combined with (4.28) and (4.29), we get

$$\|\gamma_{F_v}^{-1}\|_{k,2^k p(2p+1)}^k \|DF_v\|_{k,2^k p(2p+1)}^k \leq_c (\sigma_{inf}\sqrt{T})^{-2k}(|\sigma|_\infty\sqrt{T})^k \leq_c (\sigma_{inf}\sqrt{T})^{-k}.$$

This completes our proof. □

4.5.3 Error analysis for binary payoff (under (H_3))

For this kind of option, the payoff h is not necessarily smooth. We only assume that h is in $\mathscr{H}$. The results below are easy extensions of the vanilla options case, we leave the proof to the reader.

Theorem 4.5.3. ***Error for binary payoff.*** *Assume that (R_4) and (E) hold, and that the payoff function h fulfills Assumption (H_3). Then the error term in Theorem 4.2.1 satisfies the following estimate:*

$$|\text{Error}| \leq_c \big(\|h(X_T^M)\|_3 + \int_0^1 \|h((1-v)X_T^M + vX_T)\|_3 dv\big) \times \big(\frac{M_1}{\sigma_{inf}} + \frac{M_1^2}{\sigma_{inf}^2}\big)\big((M_0\sqrt{T})^2 + M_J^2\sqrt{\lambda T}\big).$$

Unlike the cases of smooth and vanilla payoff, for binary payoffs the approximation formula (4.6) is of first order w.r.t. the amplitudes of the model data (with error terms of order 2). This is inherent to the lack of regularity of the payoff.

4.6 Appendix

4.6.1 Technical results related to explicit correction terms

In this subsection, we bring together the results (and their proofs) which allow us to derive the explicit terms in the formula (4.6).
In the following, (u_t) (resp. (v_t) and (ν_t)) are square integrable and predictable (resp. deterministic) process and l is a smooth function with compact support.

Lemma 4.6.1. *For any continuous (or piecewise continuous) function f, any continuous semimartingale Z vanishing at t=0, one has:*

$$\int_0^T f_t Z_t dt = \int_0^T (\int_t^T f_s ds) dZ_t.$$

Proof. This follows from the Itô formula applied to the product $(\int_t^T f_s ds)Z_t$. ☐

Lemma 4.6.2. *One has:*

$$\mathbb{E}[(\int_0^T u_t dW_t) l(\int_0^T v_t dW_t)] = \mathbb{E}[(\int_0^T v_t u_t dt) l^{(1)}(\int_0^T v_t dW_t)].$$

In the case of deterministic u, it is equal to $\int_0^T v_t u_t dt\, \text{Greek}_1^l(\int_0^T v_t dW_t)$.

Proof. We first give the proof in a particular case when u and v are equal to 1. By a usual integration by parts formula, one has:

$$\mathbb{E}[l(W_T)W_T] = \int_{-\infty}^{\infty} l(\sqrt{T}x)\sqrt{T}x \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx = \int_{-\infty}^{\infty} T l^{(1)}(\sqrt{T}x) \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx = T\mathbb{E}[l^{(1)}(W_T)].$$

For the general proof: apply the duality relationship of Malliavin calculus (see Lemma 1.2.1 in [88]), identifying Itô's integral and Skorohod operator for adapted integrands. ☐

Lemma 4.6.3. *Write $(X^c_{1,t})_t$ for the continuous part of $(X_{1,t})_t$. One has :*

$$\mathbb{E}[(\int_0^T \nu_t X^c_{1,t}dt)l(\int_0^T \nu_t dW_t)] = \int_0^T \nu_t \sigma_t(\int_t^T \nu_s ds)dt\,\mathrm{Greek}^l_1(\int_0^T \nu_t dW_t) + \int_0^T \mu_t(\int_t^T \nu_s ds)dt\,\mathrm{Greek}^l_0(\int_0^T \nu_t dW_t).$$

Proof. Applying first Lemma 4.6.1 to $f(t) = \nu_t$ and $Z_t = X^c_{1,t}$, one has:

$$\begin{aligned}\mathbb{E}[(\int_0^T \nu_t X^c_{1,t}dt)l(\int_0^T \nu_t dW_t)] &= \mathbb{E}[(\int_0^T (\int_t^T \nu_s ds)dX^c_{1,t})l(\int_0^T \nu_t dW_t)]\\ &= \mathbb{E}[(\int_0^T (\int_t^T \nu_s ds)(\sigma_t dW_t + \mu_t dt)l(\int_0^T \nu_t dW_t)]\\ &= (\int_0^T \nu_t \sigma_t(\int_t^T \nu_s ds)dt)\mathbb{E}[l^{(1)}(\int_0^T \nu_t dW_t)]\\ &+ (\int_0^T \mu_t(\int_t^T \nu_s ds)dt)\mathbb{E}[l(\int_0^T \nu_t dW_t)],\end{aligned}$$

and we have used Lemma 4.6.2 for the last equality. □

Lemma 4.6.4. *One has:*

$$\mathbb{E}[(\int_0^T \nu_t J_t dt)l(J_T)] = \lambda(\eta_J \int_0^T t\nu_t dt\,\mathrm{Greek}^l_0(J_T + Y') + \gamma_J^2 \int_0^T t\nu_t dt\,\mathrm{Greek}^l_1(J_T + Y')),$$

such that Y' is an independent copy of the variables $(Y_i)_{i\in\mathbb{N}^*}$.

Proof. Using the independence of increments for J, one has:

$$\mathbb{E}[(\int_0^T \nu_t J_t dt)l(J_T)] = \int_0^T \nu_t \mathbb{E}[J_t l(J_T - J_t + J_t)]dt = \int_0^T \nu_t \mathbb{E}[\iota(J_T - J_t)]dt.$$

Using a conditioning argument and since $\sum_{j=1}^k Y_j$ is a Gaussian random variable, one has:

$$\begin{aligned}\iota(x) = \mathbb{E}[J_t l(x + J_t)] &= \sum_{k\in\mathbb{N}^*} \mathbb{P}(N_t = k)\mathbb{E}[\sum_{j=1}^k Y_j l(x + \sum_{j=1}^k Y_j)]\\ &= \sum_{k\in\mathbb{N}^*} \mathbb{P}(N_t = k)k(\eta_J \mathbb{E}[l(x + \sum_{j=1}^k Y_j)] + \gamma_J^2 \mathbb{E}[l^{(1)}(x + \sum_{j=1}^k Y_j)])\\ &= \sum_{k\in\mathbb{N}} \lambda t \mathbb{P}(N_t = k)(\eta_J \mathbb{E}[l(x + \sum_{j=1}^{k+1} Y_j)] + \gamma_J^2 \mathbb{E}[l^{(1)}(x + \sum_{j=1}^{k+1} Y_j)])\\ &= \lambda t(\eta_J \mathbb{E}[l(x + J_t + Y')] + \gamma_J^2 \mathbb{E}[l^{(1)}(x + J_t + Y')]),\end{aligned}$$

with Y' as in the lemma statement. □

4.6.2 Upper bound for compound Poisson process

Lemma 4.6.5. *The* $\mathbf{L}_p$ *norm* ($p \geq 1$) *of the compound Poisson process at time* $t \leq T$ *can be estimated as follows:*

$$\mathbb{E}|J_t|^p \leq_c M_J^p \lambda t.$$

Proof. Set $Z_j = (Y_j - \eta_J)/\gamma_J$. The random variables $(Z_j)_j$ are i.i.d. Gaussian variables, with zero mean and unit variance. Then

$$|J_t| = |\sum_{j=1}^{N_t} \eta_J + \gamma_J Z_j| \leq |\eta_J| N_t + \gamma_J |\sum_{j=1}^{N_t} Z_j| \leq M_J(N_t + |\sum_{j=1}^{N_t} Z_j|).$$

Now it only remains to compute the p-th moment of N_t and $K_t = |\sum_{j=1}^{N_t} Z_j|$, which is considered a standard exercise. We give few details about the second term K_t. First compute the characteristic function $\varphi(u) = \mathbb{E}(e^{iu\sum_{j=1}^{N_t} Z_j}) = \exp(\lambda t(e^{-u^2/2} - 1))$. Then for an even integer p, one has $\mathbb{E}(\sum_{j=1}^{N_t} Z_j)^p = \mathbb{E}(K_t^p) = i^p \varphi^{(p)}(0) = O(\lambda t)$. For odd values of p of the form $p = 2k+1$, we apply the inequality $ab \leq \frac{1}{2}(a^2 + b^2)$ to write $K_t^p \leq \frac{1}{2}(K_t^{2k} + K_t^{2k+2})$. The result then follows by using the estimates from the previous case (p even). □

Chapter 5

Appendix

5.1 Bumped parameters

In this section, we give numerical results concerning the accuracy of the formula for the AA model when we bump one of the following CEV and Jump model parameters:

- the diffusion level ν,
- or the diffusion skew β,
- or the jump intensity λ.

The aim of this section is to observe if the errors are still acceptable when we bump these parameters and when they become large.

Here, we choose as initial parameters the following:

- $S_0 = 1$,
- $r = q = 0$,
- $T = 1Y$,
- $\sigma = 0.25$,
- $\beta = 0.95$,
- $\lambda = 0.3$,
- $\mu_J = -0.08$,
- $\sigma_J = 0.35$.

In the table 5.1, we give the implied Black Scholes volatilities using our approximation formula (4.6), the PDE method and give the related errors (in bp). Hence, we observe that the approximation formula is still very good when σ or λ increase. Moreover, the errors increase slightly when β goes away from 1 and become large when β is close to 0. This is quite expected because in this case the AA model is far from the proxy (Merton's model).

Table 5.1: Implied BS volatilities for the approximation formula, the PDE method and the errors (in bp).

App. form.	80%	90%	100%	110%	120%
initial parameters	30.98%	30.36%	30.09%	30.05%	30.21%
bumped parameters: $\beta = 0.5$	32.10%	30.86%	29.97%	29.36%	29.00%
β=0.1	33.08%	31.31%	29.86%	28.75%	27.91%
σ=0.5	53.54%	53.33%	53.16%	53.02%	52.91%
$\lambda = 1$	41.71%	40.88%	40.46%	40.38%	40.56%
$\lambda = 3$	63.77%	63.24%	62.91%	62.76%	62.75%
$\beta = 0.1, \lambda = 0.0$	27.42%	26.18%	25.00%	23.92%	22.90%
PDE Method	80%	90%	100%	110%	120%
initial parameters	30.96%	30.36%	30.09%	30.07%	30.24%
bumped parameters: β=0.5	32.04%	30.91%	30.12%	29.63%	29.42%
β=0.1	33.11%	31.45%	30.18%	29.28%	28.74%
$\sigma = 0.5$	53.54%	53.34%	53.19%	53.07%	52.98%
$\lambda = 1$	41.69%	40.89%	40.50%	40.45%	40.65%
$\lambda = 3$	63.81%	63.32%	63.03%	62.91%	62.93%
$\beta = 0.1, \lambda = 0.0$	27.66%	26.26%	25.05%	23.99%	23.04%
Errors	80%	90%	100%	110%	120%
initial parameters	1.97	0.54	-0.49	-1.75	-3.19
bumped parameters: β=0.5	5.88	-4.36	-14.94	-26.77	-41.94
β=0.1	-3.37	-14.25	-31.32	-52.86	-83.81
σ=0.5	0.25	-1.64	-3.24	-5.34	-6.74
λ=1	2.02	-1.17	-3.57	-6.68	-8.99
λ=3	-4.23	-8.47	-11.44	-15.08	-18.29
$\beta = 0.1, \lambda = 0.0$	-24.00	-8.25	-4.98	-6.52	-14.45

5.2 Calibration of Index option

Here, we calibrate the EURO STOXX 50 Index. The surface of implied Black Scholes volatility is given in the table 5.2 and plotted in the figure 5.1.

We use in this example the same calibration procedure described in Subsection 4.4.3. After calibration, the jump parameters are $\lambda = 28.52\%$, $\eta_J = -31.32\%$ and $\gamma_J = 5.11\%$ and the time dependent diffusion parameters ν and β are given in the table 5.3. The calibrated values are realistic.

The errors between the implied volatilities generated by the calibrated model and the market data are given in the table 5.4 and plotted in figure 5.2. The errors show that our model is a good model for the Index EURO STOXX 50. Within our relevant algorithm, we are able to fit a 6×5 grid of quoted prices of Index options in less than **400 milliseconds**.

5.3 Robustness of the parameters optimization/calibration

Through the following tests, we aim to check the robustness of our calibration. In other words, we want to verify that we are able to retrieve the true model from an initial guess. To this end, we apply 3 tests in a progressive way: In test 1, we treat the case of Merton's model ($\beta = 1$). In test 2, we allow β to be

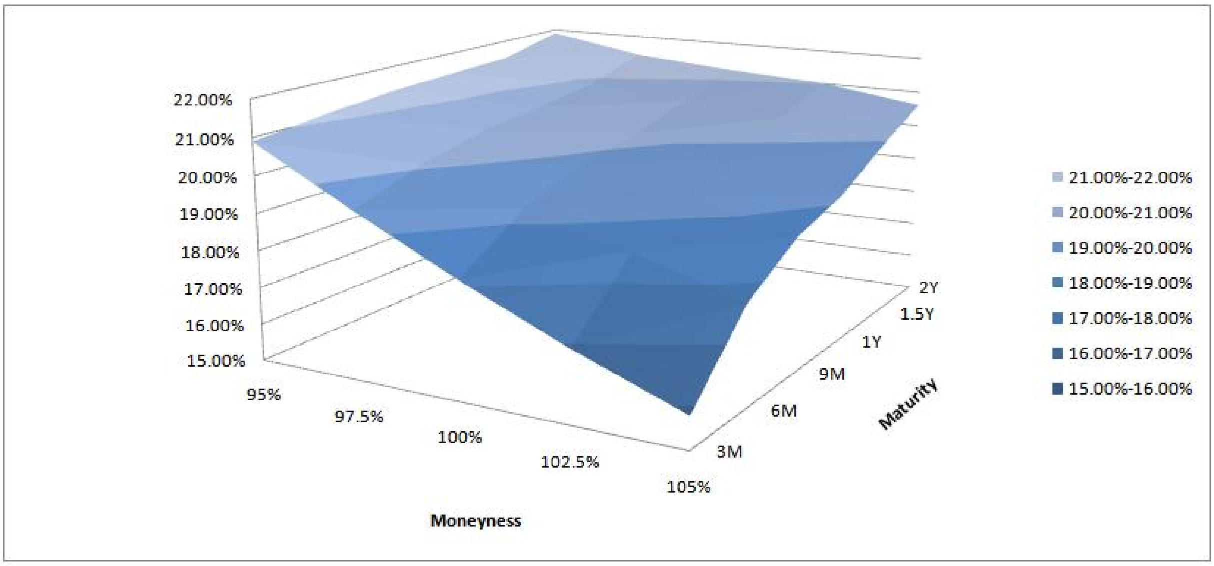


Figure 5.1: Market implied Black Scholes volatilities for Index option: EUROSTOXX 50.

Table 5.2: Implied B-S volatilities for the EURO STOXX Index expressed as a function of maturities in fractions of years (T) and relative strikes (K). The risk free rate is equal to 4.08%.

T/K	95%	97.5%	100%	102.5%	105%
3M	20.88%	19.47%	18.13%	16.91%	15.85%
6M	21.12%	20.07%	19.26%	18.55%	17.70%
9M	21.30%	20.47%	19.86%	19.33%	18.65%
1Y	21.39%	20.67%	20.16%	19.71%	19.11%
1.5Y	21.46%	20.90%	20.61%	20.40%	19.92%
2Y	21.89%	21.41%	21.18%	21.02%	20.61%

different than 1 which corresponds to CEV model with jump. In test 3, we treat the general case of CEV model with jump and time dependent parameters.

Test 1. The set of calibration is a grid with relative strikes $80\%, 90\%, 100\%, 110\%, 120\%$ and maturities: 3M and 6M. The spot is normalised and equal to 1. *The parameter β is fixed to 1 (Merton's model).* The results detailed in table 5.5 show that from an initial guess not far from the true model, we can retrieve the true solution for the Merton's model[1].

Test 2. CEV model with jumps. The set of calibration: relative strikes $80\%, 90\%, 100\%, 110\%, 120\%$; maturities: 3M and 6M. Spot=1. The results are shown in table 5.6 from which observe that the global calibration of the five parameters still has good results and gives the true solution with small errors.

Test 3. CEV model with jumps and time dependent parameters. The set of calibration: relative strikes $80\%, 90\%, 100\%, 110\%, 120\%$; maturities: 3M, 6M, 9M, 1Y, 18M, 2Y. Spot=1. Here, we use the algorithm developed in Subsection 4.4.3. First, we calibrate the jumps parameters for the first maturi-

[1]Notice for an initial guess far from the solution, a simple calibration is an ill posed problem and can have different minimums (see [31]). This justifies the need of a regularisation method or to take into the calibration set other financial instruments.

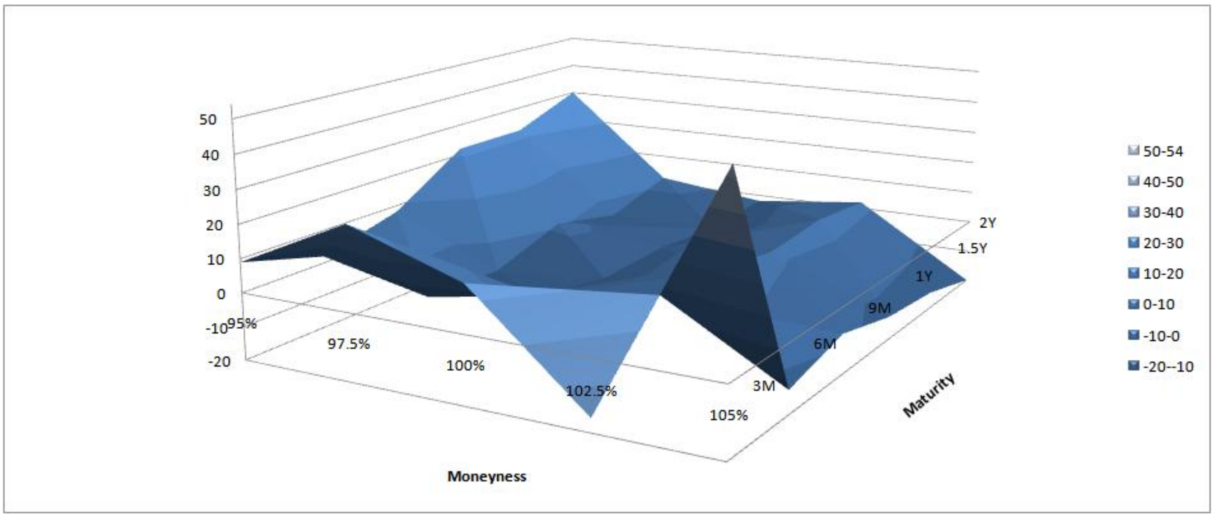


Figure 5.2: Implied Black Scholes volatilities errors (in bp) between our approximation formula and the PDE method.

Table 5.3: Calibrated values of the piecewise constant functions ν and β.

T	ν	β
3M	21.48%	94.36%
6M	18.73%	95.63%
9M	21.46%	93.81%
1Y	21.41%	93.39%
1.5Y	18.06%	96.60%
2Y	18.15%	98.38%

ties (which correspond to the two first maturities in the table 5.7). Second, we calibrate the diffusion parameters ν and β. The table 5.7 gives exhaustive results for the calibration of such a time dependent model.

Table 5.4: Errors (in bp) between implied B-S volatilities for the EURO STOXX 50 Index and those calculated within the calibrated model expressed as a function of maturities in fractions of years (T) and relative strikes (K). The risk free rate is equal to 4.08%.

T/K	95%	97.5%	100%	102.5%	105%
3M	9	25	14	-16	54
6M	2	-5	1	7	-13
9M	8	-6	-3	5	-8
1Y	22	1	-2	2	-13
1.5Y	22	-4	-4	4	-15
2Y	30	2	-2	2	-20

Table 5.5: Calibration for Merton's model.

Parameters (in %)	ν	β	λ	μ_J	γ_J
True model	25	(100)	10	-5	30
Initial guess	22	(100)	5	0	35
After optimization	25.02	(100)	9.32	-5.32	30.98

Table 5.6: Calibration for CEV model with jumps.

Parameters (in %)	ν	β	λ	μ_J	γ_J
True model	25	95	10	-5	30
Initial guess	22	100	5	0	35
After optimization	25.02	94.93	9.33	-5.28	30.96

Table 5.7: Calibrated for CEV model with jumps and time dependent parameters. True jump parameters: $\lambda = 10\%$, $\mu_J = -5\%$ $\gamma_J = 30\%$. Jump parameters before optimization: $\lambda = 5\%$, $\mu_J = 0\%$ $\gamma_J = 35\%$. Jump parameters after optimization: $\lambda = 9.32\%$, $\mu_J = -5.32\%$ $\gamma_J = 30.98\%$.

	True model		Init. guess		After optimiz.	
Maturity	ν in %	β in %	ν in %	β in %	ν in %	β in %
3M	25	(100)	22	(100)	25	(100)
6M	25	(100)	22	(100)	25	(100)
9M	23	98.30	22	100	24.12	97.24
1Y	22	95.95	22	100	22.04	95.79
18M	21	94.03	22	100	21.02	94
2Y	20	93.05	22	100	20.01	92.90

Part II

Local volatility models

Chapter 6

Introduction

The local volatility models have been introduced by Dupire (in [40]) and Rubinstein (in [101]), in order to quote different market prices for different strikes and maturities. Indeed, they derived an explicit formula for the local volatility σ in terms of the call prices for different strikes K and maturities T:

$$\sigma^2(T,K) = 2\frac{\frac{\partial Call}{\partial T} + r_T K \frac{\partial Call}{\partial K}}{K^2 \frac{\partial^2 Call}{\partial K^2}}. \tag{6.1}$$

This equation[1] is known as the Dupire's formula and is derived from the Fokker Plank equation (also known as the forward Kolmogorov Equation) and given by:

$$-\frac{\partial Call}{\partial T} + \frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2 Call}{\partial K^2} - r_T K \frac{\partial Call}{\partial K} = 0.$$

6.1 CEV model

In the case of Constant Elasticity Variance model (known as CEV model), call (and put) options have closed form-formulas. In this model, the spot (S_t) has a well known dynamics given by:

$$\frac{dS_t}{S_t} = (r-q)dt + \nu S_t^{\beta-1} dW_t, S_0 > 0.$$

The CEV model has been originally studied by Cox in [34] for the case of $\beta < 1$. The case $\beta > 1$ has been treated after by Emanuel and MacBeth in [42]. The Call price in this model can be computed using the complementary non central Chi-square distribution Q:

$$Call_{CEV}(S;K,T) = e^{-qT}Q(2x,n,2y) - e^{-rT}Q(2y,n-2,2x) \tag{6.2}$$

where

$$\begin{aligned} n &= 2 + \frac{1}{1-\beta}, \\ x &= \frac{(r-q)S^{-2(\beta-1)}}{\nu^2(\beta-1)(e^{2(r-q)(\beta-1)T}-1)}, \\ y &= \frac{(r-q)K^{-2(\beta-1)}}{\nu^2(\beta-1)(1-e^{-2(r-q)(\beta-1)T})}. \end{aligned}$$

The computation of the non central Chi-square distribution Q can be performed using a recursive algorithm (see Schroder algorithm in [104]) or an integration of Bessel functions.

Proof. We briefly indicate here a probabilistic proof (see [104]) having the following steps:

- In the case of no drift ($r-q=0$), remark that the variable ($z_t = \frac{\nu|\beta-1|}{S_t^{-(\beta-1)}}$) is a Bessel process of order $\frac{1}{2(\beta-1)}$ (see Revuz and Yor in [98] for an introduction to Bessel processes). Moreover, we exploit the fact that the Bessel process has an explicit density (see Borodin and Salminen in [23]).

[1]From this equation, one can derive an explicit equation that gives the local volatility in terms of the implied BS volatility $\sigma_{imp}(T,K)$ and choose a smooth parameterization for the implied BS volatility to derive exact and instantaneous solution for the local volatility (for more details, see Gatheral book [50]).

- After that, we express the drifted CEV process density in terms of the no drifted CEV process density (see Goldenberg's result in [57]).
- Then, we write the Call price in this model as a linear combination of two probabilities (like in the proof of Black Scholes model).
- Finally we express these probabilities in terms of the non central Chi-square distribution Q.

□

Remark 6.1.1. *Hsu et al in [65]) derive another formula for the call price as a series of the complementary Gamma distribution. Their proof uses the fact that the fundamental solution of the related PDE can be expressed as a series of modified Bessel functions.*

6.2 Review of Analytical approximations

Only in a few cases the local volatility model admits closed form-formulas, as explained in [1] (the CEV model belongs to these cases). Otherwise, there are analytical approximations in the following cases.

Separable local volatility. In this case, the local volatility function is written as the product of two independent functions of time and underlying, $\sigma(t,f) = \alpha(t)A(f)$. Hagan et al in [62] derive an asymptotic expansion for the implied Black Scholes volatility using singular perturbation techniques.

Short maturity. In this framework, Berestycki et al in [20] derived, from Dupire formula's, the local volatility function $\sigma(t,x)$ in terms of the implied Black volatility $\sigma_{imp}(T,K)$ in the form of a parabolic partial differential equation (the same relation has been derived also by Andersen and Brotherton-Ratcliffe in [10]). Then, they gave explicit formula of the implied volatility near the expiry under an ellipticity assumption:

$$\lim_{T\to 0} \frac{1}{\sigma_{imp}(T,K)} = \frac{1}{\ln(\frac{F}{K})} \int_K^F \frac{df'}{f'\sigma(0,f')} = \int_0^1 \frac{ds}{\sigma(0, Ke^{s\ln(\frac{F}{K})})}. \tag{6.3}$$

This formula means that the implied volatility near the expiry is approximately an harmonic mean of local volatility from moneyness 0 to moneyness $\ln(\frac{F}{K})$. If the local volatility σ has a limit for extreme strikes ($\lim_{\ln(K)\to\pm\infty} \sigma_{imp}(T,K) = \sigma_{\pm}(T)$), then the implied Black volatilities is given by:

$$\lim_{\ln(K)\to\pm\infty} \sigma_{imp}(T,K) = (\frac{1}{T}\int_0^T \sigma_{\pm}^2(s)ds)^{\frac{1}{2}}. \tag{6.4}$$

Another type of asymptotic expansion can be also derived from an expansion of the heat kernel (see the work of Labordere in [74]).

Long maturities. We cite the work of Tehranchi in [108] that gives asymptotic formula for the implied Black Scholes volatilities far from maturity with precise estimate of the error. Assuming that, almost surely $S_t \to 0$ as $t \uparrow \infty$, Tehranchi shows the following formula holds:

$$T\sigma_{imp}^2(T,K) = 8|\ln(\mathbb{E}(S_T \wedge \frac{K}{S_0}))| - 4\ln(|\ln(\mathbb{E}(S_T \wedge \frac{K}{S_0}))|) + 4\ln(\frac{K}{S_0}) - 4\ln(\pi) + \varepsilon(\ln(\frac{K}{S_0}), T),$$

where

$$\sup_{-M\le x\le M} |\varepsilon(x,T)| \underset{T\to\infty}{\to} 0$$

for all $M > 0$.

We also mention Gatheral's work in [49] who derives arbitrage bounds on the skew of the implied at-the-money Black Scholes volatility. These bounds are of the order of $O(T^{-\frac{1}{2}})$.

Moreover, Rogers and Tehranchi prove in [99] the smile Theorem conjectured by S. Ross. Indeed, they prove that the smile shape can not move by parallel shifts.

Extreme strikes. Lee shows in [78] that the implied variance $\sigma^2_{imp}(T,K)T$ is bounded from above by a function linear w.r.t. the log moneyness $\ln(\frac{K}{F})$ for large strikes. He gives explicit formulas which relate the gradients of the wings of the upper bound of the implied variance and the maximal finite moments of the spot. For instance, for the left wing, he shows that if $q^* := \sup\{q : \mathbb{E}[S_T^{-q}] < \infty\}$ and

$$\beta^* := \limsup_{K\to 0^+} \frac{\sigma^2_{imp}(T,K)T}{|\ln(\frac{K}{F})|},$$

then $\beta^* \in [0,2]$ and

$$q^* = \frac{1}{2}\Big(\frac{1}{\sqrt{\beta^*}} - \frac{\sqrt{\beta^*}}{2}\Big)^2.$$

Moreover, Benaim and Friz in [15] sharpen Lee's formula. Indeed, they show that Lee's upper bound may become a limit provided some technical assumptions which are satisfied for a large class of models.

6.3 Motivation and main results

However in general cases of local volatility functions, there is no analytical formula. Therefore, the aim of this Part is to derive an explicit closed formula for European options for general forms of local volatility functions. This closed formula is a Taylor expansion and can be truncated easily at any order. Therefore, the expected price can be written at any order as a summation of:

- The Black Scholes price with at the money volatility. As in Part I, this model can be seen as the proxy of the local volatility model. The advantage of this proxy lies in the explicit calculus of the prices and the Greeks of vanilla options.
- A combination of the Greeks of the leading Black Scholes price with explicit weights depending on the volatility, the drift functions and their derivatives[2].
- A residual error with explicit upper bounds.

This is achieved in Chapter 7. The approximation for vanilla options at the second order is computed in Theorem 7.2.1 which is a particular case of Theorem 4.2.1 of Chapter 4 when there is no jump. Moreover, the explicit calculus of the approximation for vanilla options at the third order is derived in Theorem 7.2.2. In addition, the corrections and the error terms at any order of the closed formula are estimated in

[2] We do not give the expressions of the weight coefficients here because it does not enlighten more the result.

Theorems 7.4.1-7.4.2-7.4.3 according to the payoff smoothness (smooth, vanilla, digital). The accuracy of the expansion turns out to be excellent. Moreover, we need only few terms to give accurate results for vanilla options. As a consequence of these expansions, we derive averaging parameters for time dependent CEV models.

The following Chapter is exactly the article "Closed forms for European options in a local volatility model" accepted in the journal "International Journal of Theoretical and Applied Finance". The Chapter 8 details smile behaviors for the CEV model when varying its parameters through the time. We also provide numerical results concerning the accuracy of the approximation formula for large strikes and concerning the domain of "numerical" arbitrage of the approximation formulas.

Chapter 7

Closed forms for European options in a local volatility model

Accepted for Publication in "International Journal of Theoretical and Applied Finance".

Because of its very general formulation, the local volatility model does not have an analytical solution for European options. In this Chapter, we present a new methodology to derive closed form solutions for the price of any European options. The formula results from an asymptotic expansion, terms of which are Black-Scholes price and related Greeks. The accuracy of the formula depends on the payoff smoothness and it converges with very few terms.

7.1 Introduction

The local volatility model, introduced by Dupire [40], Rubinstein [101] and Derman Khani[36], has the main advantage of fitting all call and put option prices. However, in contrast to the seminal Black-Scholes model, this model has no more closed form solution for general European options. This comes from the very general form of the local volatility function. Only in a few cases this model admits closed formulas, as explained in [1]. In the special case of a separable local volatility function written as the product of two independent functions of time and underlying, $\sigma_{loc}(t,f) = \alpha(t)A(f)$, one can derive an asymptotic expansion for the price of vanilla options (call, put) using singular perturbation techniques as explained in [62]. Another type of asymptotic expansion can be also derived from an expansion of the heat kernel as shown in [74].However, for the general case, there is no methodology so far. This paper tackles precisely this challenge.

The overall idea is to do an asymptotic expansion directly on the diffusion using Malliavin calculus. We will consider a local volatility model, in which the underlying asset is classically related to the diffusion process

$$dX_t = \sigma(t,X_t)dW_t + \mu(t,X_t)dt, \quad X_0 = x_0. \tag{7.1}$$

Typically, in the following, X stands for the log-price of the underlying asset[1]. $\sigma(t,X_t)$ is the volatility term whereas $\mu(t,X_t)$ is the drift term. Our aim is to give an analytical accurate[2] approximation of any European option, written as the expected value under the risk neutral probability measure of a payoff function h evaluated at the maturity time T:

$$\mathbb{E}(h(X_T)) \tag{7.2}$$

where $\mathbb{E}$ stands for the standard expectation operator. To accomplish this, we introduce a parametrized process given by:

$$dX_t^\varepsilon = \varepsilon(\sigma(t,X_t^\varepsilon)dW_t + \mu(t,X_t^\varepsilon)dt), X_0^\varepsilon = x_0, \tag{7.3}$$

where the parameter ε lies in the range $[0,1]$. Obviously, this parametrized process is equal to the initial one for $\varepsilon = 1$. Remarkably, it is much easier to calculate the price (7.2) as an expansion formula with respect to ε. Once we have derived all the terms of the expansion, we see that the price of the European option is obtained by taking $\varepsilon = 1$.

Compared to standard expansion methods, the accuracy of this expansion is not related to the perturbation parameter ε. Indeed, the limit value $\varepsilon = 1$ is not small at all. This is a significant difference compared to singular perturbation techniques. Our expansion is just a way to derive convenient closed form solution. This asymptotic expansion is achieved using the infinite dimensional analysis of Malliavin calculus. A key feature of our approach is that we can provide explicit formulas for the terms at any order and explicit upper bounds of the errors, for general forms of the drift term μ and the volatility term σ. The derivation of expansion terms at any order completes for pure diffusion some earlier work done in Chapter 4.

[1] when explicitely stated, X may alternatively stand for the asset price.

[2] in some sense detailed later in this Chapter

In practice, we compute a limited number of terms. The main term is the price in a suitable Black-Scholes model, while the other terms are a weighted summation of sensitivities (Greeks). These terms are straightforward to evaluate numerically, with a computational cost equivalent to the Black-Scholes formula. The smaller the parameters μ and σ are, the smaller the maturity T is, or the smaller the derivatives of the functions μ and σ with respect to their second variable are, the faster the convergence of the expansion is. This means that in practice, we need to calculate the expansion up to the second order, or possibly to the third order, to achieve an excellent accuracy (smaller than 2 bp on implied volatilities for various strikes and maturities). In addition, as a consequence of our approximation formulas, we establish that, for any fixed maturity, a time dependent CEV model is equivalent to a CEV model with appropriate constant parameters (parameter averaging principle).

Outline of the Chapter. In the following, we give some notations and assumptions used throughout the Chapter. The next section presents in an heuristic way our methodology to approximate the expected cost. We provide approximation formulas at the second and third order, using a log-normal or a normal proxy. In Section 7.3, we detail the approximation formulas for the case of time dependent CEV volatility. In Section 7.4, we analyse the magnitude of the correction and error terms of the general approximation formula (and at any order). The analysis depends on the payoff smoothness. The proofs of the main theorems 7.4.1-7.4.2-7.4.3 are postponed to section 7.5. In appendix 7.6, we bring together useful results to make our expansion explicit.

Definitions

Definition 7.1.1. *As usual, we define $\mathscr{C}_0^\infty(\mathbb{R})$ as the space of real infinitely differentiable functions h with compact support. We also define $\mathscr{H}$ as the space of functions having at most an exponential growth. A function h belongs to $\mathscr{H}$ if $|h(x)| \leq c_1 e^{c_2|x|}$ for any x, for two constants c_1 and c_2.*

Notations

The following notation will be used extensively throughout the Chapter.

Notation 7.1.1. ***Differentiation.***
If these derivatives have a meaning, we write:

- $\psi_t^{(i)}(x) = \frac{\partial^i \psi}{\partial x^i}(t,x)$ *for any function ψ of two variables.*
- $X_{i,t}^\varepsilon = \frac{\partial^i X_t^\varepsilon}{\partial \varepsilon^i}$ *is the i^{th} derivative of the parametrized process with respect to ε.*
- $X_{i,t} = \frac{\partial^i X_t^\varepsilon}{\partial \varepsilon^i}|_{\varepsilon=0}$. *These processes play a crucial role in this work.*
- $\sigma_t = \sigma(t,x_0), \mu_t = \mu(t,x_0), \sigma_t^{(i)} = \sigma^{(i)}(t,x_0), \mu_t^{(i)} = \mu^{(i)}(t,x_0)$.

The following notation of Greeks will be useful for interpreting the expansion terms.

Notation 7.1.2. ***Greeks.***
Let Z be a random variable. Given a payoff function h, we define the i^{th} Greek for the variable Z by the quantity (if it has a meaning) :

$$\text{Greek}_i^h(Z) = \frac{\partial^i \mathbb{E}[h(Z+x)]}{\partial x^i}|_{x=0}.$$

Assumptions. In order to derive accurate approximations, we may assume that coefficients σ and μ are smooth enough. In what follows, N is an integer greater than 4.

- ***Assumption*** *(R_N). The functions σ and μ are bounded and of class C^N w.r.t x. Their derivatives up to order N are bounded.*

This assumption may be restrictive because σ and μ have to be bounded as well their derivatives. Actually, this statement is made only to simplify a bit our analysis, but we can prove that our approximation remains valid if some boundedness requirements are partially relaxed.

Notation 7.1.3. *Function amplitudes.*
Under (R_N), we set

$$M_0 = max(|\sigma|_\infty, \cdots, |\sigma^{(N)}|_\infty, |\mu|_\infty, \cdots, |\mu^{(N)}|_\infty), \tag{7.4}$$

$$M_1 = max(|\sigma^{(1)}|_\infty, \cdots, |\sigma^{(N)}|_\infty, |\mu^{(1)}|_\infty, \cdots, |\mu^{(N)}|_\infty). \tag{7.5}$$

Although M_0 and M_1 may depend on N, we remove this dependence in our notation, for the sake of simplicity.

Remark 7.1.1. *The constant M_0 measures the amplitude of the objective functions μ, σ and their derivatives w.r.t. the second variable, whereas M_1 measures only the amplitude of their derivatives. Notice that $M_1 \leq M_0$ and in case of deterministic functions σ and μ, one has $M_1 = 0$.*

To perform the infinitesimal analysis, we rely on smoothness properties not related to the payoff function itself but rather to the law of the underlying stochastic models.

- ***Assumption*** *(E). The function σ does not vanish and its oscillation is bounded, meaning $1 \leq \frac{|\sigma|_\infty}{\sigma_{inf}} \leq C_E$ where $\sigma_{inf} = \inf_{(t,x)\in\mathbb{R}^+\times\mathbb{R}} \sigma(t,x)$.*

The assumption (E) is commonly called an ellipticity assumption.

We also need to divide our analysis according to the payoff smoothness. We split our analysis into three cases.

- ***Assumption*** *(H_1). h belongs to $\mathscr{C}_0^\infty(\mathbb{R})$. This case corresponds to smooth payoffs.*
- ***Assumption*** *(H_2). h and $h^{(1)}$ belongs to $\mathscr{H}$. This case corresponds to vanilla options (call-put).*
- ***Assumption*** *(H_3). h belongs to $\mathscr{H}$. This is the case of binary options (digital).*

7.2 Smart Taylor Development

In the following, we provide several approximation formulas, at the second and third order. These formulas are different if X models the logarithm of the underlying asset price or if it models directly the asset price. In the first case, our approximation is equivalent to take a lognormal proxy (or Black-Scholes proxy) for the asset price; in the second case, it is equivalent to use a normal proxy.

7.2.1 Second order approximation

Here, we consider that the dynamics (7.1) for X models the logarithm of the underlying asset. In the case of call option, the payoff is then $h(x) = (e^x - K)_+$, where K is the strike price.

Our perturbation approach relies on the Taylor expansion of the parameterized process. We have paved the way in our previous work Chapter 4 (see Section 4.2). For the sake of completeness, we briefly recall the main steps to achieve a closed approximative formula.

From the definitions, $X_{i,t} \equiv \frac{\partial^i X_t^\varepsilon}{\partial \varepsilon^i}|_{\varepsilon=0}$, we can expand the perturbed process X_T^ε as follows:

$$X_T^\varepsilon = X_T^\varepsilon|_{\varepsilon=0} + \varepsilon X_{1,T} + \frac{\varepsilon^2}{2!} X_{2,T} + \dots \tag{7.6}$$

Indeed, under the assumption (R_5), almost surely for any t, X_t^ε is C^4 w.r.t ε (see Theorem 2.3 in [72]). The diffusion dynamics of $(X_{i,t}^\varepsilon \equiv \frac{\partial^i X_t^\varepsilon}{\partial \varepsilon^i})_{t\geq 0}$ is obtained by a straight differentiation of the parameters of the diffusion equation of X^ε. The first order term $X_{1,t}^\varepsilon$ is easily obtained as follows:

$$\begin{aligned} dX_{1,t}^\varepsilon =& \sigma_t(X_t^\varepsilon)dW_t + \mu_t(X_t^\varepsilon)dt \\ &+ \varepsilon X_{1,t}^\varepsilon(\sigma_t^{(1)}(X_t^\varepsilon)dW_t + \mu_t^{(1)}(X_t^\varepsilon)dt), X_{1,0}^\varepsilon = 0. \end{aligned} \tag{7.7}$$

From the definitions, we have $\sigma_t \equiv \sigma(t,x_0)$, $\mu_t \equiv \mu(t,x_0)$, $\sigma_t^{(i)} \equiv \sigma^{(i)}(t,x_0)$ and $\mu_t^{(i)} \equiv \mu^{(i)}(t,x_0)$. Then, we obtain

$$\begin{aligned} dX_{1,t} =& \sigma_t dW_t + \mu_t dt, X_{1,0} = 0, \\ dX_{2,t} =& 2X_{1,t}(\sigma_t^{(1)} dW_t + \mu_t^{(1)} dt), X_{2,0} = 0. \end{aligned}$$

Applying the expansion (7.6) at $\varepsilon = 1$, we conclude that $x_0 + X_{1,T}$ is a proxy for X_T. This is a Gaussian proxy for X, hence a lognormal proxy for the asset price (or Black-Scholes diffusion proxy). It justifies the notation

$$X_T^{BS} = x_0 + X_{1,T} = x_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s. \tag{7.8}$$

To obtain an approximation formula, we use the Taylor formula twice: first, for X_T^1 at the second order w.r.t ε around x_0, secondly for smooth function h at the first order w.r.t x around X_T^{BS}. This leads to:

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^{BS} + \frac{X_{2,T}}{2} + \dots)] = \mathbb{E}[h(X_T^{BS})] + \mathbb{E}[h^{(1)}(X_T^{BS})\frac{X_{2,T}}{2}] + \dots$$

To achieve an explicit formula, it remains to transform the correction term involving $X_{2,T}$ into a summation of greeks computed in the Black-Scholes proxy. This is performed using the Malliavin calculus. We refer to Chapter 4 (Proof of Theorem 4.2.1) where the computations are detailed, or to the proof of Theorem 7.2.2 in this Chapter. This leads to the following theorem, which is a particular case of Theorem 4.2.1 of Chapter 4 when there is no jump.

Theorem 7.2.1. ***(Second order approximation price formula using lognormal proxy).***
Assume that the process (X_t) fulfills (R_5) and (E), and that the payoff function fulfills one of the assumptions (H_1), (H_2) or (H_3). Then

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^{BS})] + \sum_{i=1}^{3} \alpha_{i,T} \mathrm{Greek}_i^h(X_T^{BS}) + \mathrm{Resid}_2, \tag{7.9}$$

where

$$\begin{aligned} \alpha_{1,T} =& \int_0^T \mu_t (\int_t^T \mu_s^{(1)} ds) dt, \\ \alpha_{2,T} =& \int_0^T (\sigma_t^2 (\int_t^T \mu_s^{(1)} ds) + \mu_t (\int_t^T \sigma_s \sigma_s^{(1)} ds)) dt, \\ \alpha_{3,T} =& \int_0^T \sigma_t^2 (\int_t^T \sigma_s \sigma_s^{(1)} ds) dt. \end{aligned}$$

Additionally, estimates of the error term Resid$_2$ *(otherwise stated as residual terms) are given in Theorems 7.4.1, 7.4.2 and 7.4.3, according to the cases* (H_1), (H_2) *or* (H_3).

Formula (7.9) is refereed as a second order approximation formula because we establish, in Theorem 7.4.2 for call/put option, that the error term Resid$_2$ is of order three with respect to the amplitudes of coefficients.
The above approximation of the price is a sum of two terms:

1. $\mathbb{E}[h(X_T^{BS})]$ is the leading order, corresponding to the price when the parameters σ and μ are deterministic. In the case of call/put option, it is given by the Black- Scholes formula. For other payoffs, we can use numerical integration because the density of the random variable X_T^{BS} is explicit.
2. $\sum_{i=1}^{3} \alpha_{i,T} \text{Greek}_i^h(X_T^{BS})$ are the volatility and drift correction terms, which depend on the first derivatives of μ and σ. These terms can be computed as easily as the main term.

The above formula may be simplified when the asset (i.e. $(e^{X_t})_{t\geq 0}$) is a martingale under the pricing measure[3] (also refered to Dupire model). Then, $\mu(t,x) = -\frac{1}{2}\sigma^2(t,x)$ and the formula writes

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^{BS})] + C_{1,T}(\frac{1}{2}\text{Greek}_1^h(X_T^{BS}) - \frac{3}{2}\text{Greek}_2^h(X_T^{BS}) + \text{Greek}_3^h(X_T^{BS})) + \text{Resid}_2$$

with

$$C_{1,T} = \int_0^T \sigma_t^2 (\int_t^T \sigma_s \sigma_s^{(1)} ds) dt. \tag{7.10}$$

7.2.2 Third order approximation using a lognormal proxy

If the original model is close to its lognormal proxy, the formula (7.9) is very accurate (see the numerical results in Section 7.3). Otherwise, we can obtain higher accuracy by adding third order correction terms. The following result provides explicit expressions for these terms in the Dupire model ($\mu(t,x) = -\frac{1}{2}\sigma^2(t,x)$) for vanilla payoffs. Before, we introduce an appropriate definition, which will enable us to represent the coefficients of the greeks as iterated time integrals.

Definition 7.2.1. ***Integral Operator.***
The integral operator ω^T *is defined as follows: for any integrable function l, we set*

$$\omega(l)_t^T = \int_t^T l_u du$$

for $t \in [0,T]$. *Its n-times iteration is defined analogously: for any integrable functions* $(l_1, \cdots, l_n)$, *we set*

$$\omega(l_1, \cdots, l_n)_t^T = \omega(l_1 \omega(l_2, \cdots, l_n)_.^T)_t^T$$

for $t \in [0,T]$.

Theorem 7.2.2. ***(Third order approximation price formula in the Dupire model using lognormal proxy).*** *Assume that the process* $(X_t)_{t\geq 0}$ *fulfills* (R_7) *and* (E), *and that the payoff function fulfills the assumption* (H_2). *Then*

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^{BS})] + \sum_{i=1}^{6} \eta_{i,T} \text{Greek}_i^h(X_T^{BS}) + Resid_3, \tag{7.11}$$

[3] for instance, when one models the evolution of the forward price.

where

$$
\begin{aligned}
\eta_{1,T} =& \frac{C_{1,T}}{2} - \frac{C_{2,T}}{2} - \frac{C_{3,T}}{2} - \frac{C_{4,T}}{4} - \frac{C_{5,T}}{4} - \frac{C_{6,T}}{2},\\
\eta_{2,T} =& -\frac{3C_{1,T}}{2} + \frac{C_{2,T}}{2} + \frac{C_{3,T}}{2} + \frac{5C_{4,T}}{4} + \frac{5C_{5,T}}{4} + \frac{7C_{6,T}}{2} + \frac{C_{7,T}}{2} + \frac{C_{8,T}}{4},\\
\eta_{3,T} =& C_{1,T} - 2C_{4,T} - 2C_{5,T} - 6C_{6,T} - 3C_{7,T} - \frac{3C_{8,T}}{2},\\
\eta_{4,T} =& C_{4,T} + C_{5,T} + 3C_{6,T} + \frac{13C_{7,T}}{2} + \frac{13C_{8,T}}{4},\\
\eta_{5,T} =& -6C_{7,T} - 3C_{8,T},\\
\eta_{6,T} =& 2C_{7,T} + C_{8,T},
\end{aligned}
$$

and

$$
\begin{aligned}
C_{1,T} &= \omega(\sigma^2, \sigma\sigma^{(1)})_0^T, & C_{2,T} &= \omega(\sigma^2, (\sigma^{(1)})^2)_0^T,\\
C_{3,T} &= \omega(\sigma^2, \sigma\sigma^{(2)})_0^T, & C_{4,T} &= \omega(\sigma^2, \sigma^2, (\sigma^{(1)})^2)_0^T,\\
C_{5,T} &= \omega(\sigma^2, \sigma^2, \sigma\sigma^{(2)})_0^T, & C_{6,T} &= \omega(\sigma^2, \sigma\sigma^{(1)}, \sigma\sigma^{(1)})_0^T,\\
C_{7,T} &= \omega(\sigma^2, \sigma^2, \sigma\sigma^{(1)}, \sigma\sigma^{(1)})_0^T, & C_{8,T} &= \omega(\sigma^2, \sigma\sigma^{(1)}, \sigma^2, \sigma\sigma^{(1)})_0^T.
\end{aligned}
$$

In addition, the estimate of the error term $Resid_3$ is given in Theorem 7.4.2.

An application of Theorem 7.4.2 yields that $Resid_3$ is of order four with respect to the volatility coefficient.

The proof of Theorem 7.2.2 is postponed to subsection 7.6.3.

7.2.3 Third order approximation using a normal proxy

In the previous third order approximation formula, numerous correction terms appear because the *smart expansion* involves simultaneously the volatility and the drift coefficients. If we consider directly a model on the asset price (and not on its logarithm), our expansion simplifies much because the drift in the Dupire model vanishes:

$$dX_t = \sigma(t, X_t)dW_t. \tag{7.12}$$

The above function σ for the asset price X and the volatility function σ in (7.1) for the log-asset are different, they are simply related by a change of variables of exponential type. Similarly, here the call payoff is equal to $h(x) = (x-K)_+$. Then, we can perform our expansion approach using the parametrized process X^ε that solves $dX_t^\varepsilon = \varepsilon\sigma(t, X_t^\varepsilon)dW_t$. We obtain that the model proxy for the asset price is defined by

$$X_t^N = x_0 + \int_0^t \sigma(s, x_0)dW_s, \tag{7.13}$$

which is a Gaussian process. We call it normal proxy. Formal computations of our *smart expansion* are analogous to those done for the lognormal proxy. We will skip details regarding the proof and the assumptions. We do not provide a rigorous estimation of the error term, we prefer to focus on the expressions of correction terms to achieve a third order approximation formula.

Theorem 7.2.3. ***(Third order approximation price formula in the Dupire model using normal proxy).*** *For a vanilla payoff h, we have*

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^N)] + \sum_{i=1}^{6} \eta_{i,T} \text{Greek}_i^h(X_T^N) + Error, \tag{7.14}$$

where

$$\eta_{1,T} = 0, \qquad \eta_{2,T} = \frac{C_{2,T}}{2} + \frac{C_{3,T}}{2}, \qquad \eta_{3,T} = C_{1,T},$$
$$\eta_{4,T} = C_{4,T} + C_{5,T} + 3C_{6,T}, \qquad \eta_{5,T} = 0, \qquad \eta_{6,T} = 2C_{7,T} + C_{8,T}.$$

The coefficients $(C_{j,T})_{1\leq j\leq 8}$ *are defined as in Theorem 7.2.2.*

In the case of call/put option, the computations of the main term $\mathbb{E}[h(X_T^N)]$ and of the related greeks $(\text{Greek}_i^h(X_T^N))_{1\leq i\leq 6}$ are straightforward because the proxy (7.13) is normal. Numerical results are reported in Section 7.3.

If one prefers to restrict to a second order approximation formula, it simpy writes

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^N)] + C_{1,T} \text{Greek}_3^h(X_T^N) + Error. \tag{7.15}$$

7.2.4 Parameter averaging in CEV model

The time dependent CEV model on the underlying asset is defined by

$$dX_t = \nu_t X_t^{\beta_t} dW_t.$$

We suppose that the risk-free rate $(r_t)_t$ and the dividend yield $(q_t)_t$ are both deterministic. For simplicity in the following discussion, we assume $X_0 = 1$ in order to have a normalized model.

As discussed in Chapter 4 (see Section 4.3), the time dependent CEV model is interesting because it generates all the possible values of $(\sigma_t)_t$ and $(\sigma_t^{(1)})_t$ by appropriate choices of $(\nu_t)_t$ and $(\beta_t)_t$. Thus, in view of (7.9) and (7.15), this model may potentially generate all the possible prices at the second order.

When the coefficients $(\nu_t)_t$ and $(\beta_t)_t$ are constant, there is a closed formula for the call price (see [104]). For general time dependent coefficients, we may use our approximation formulas based on log-normal or normal proxy. Alternatively, we may look for an equivalent CEV model with constant coefficients $\bar{\nu}$ and $\bar{\beta}$, with which the prices coincide at the second order. This is possible maturity by maturity. This principle has been studied for stochastic volatility models by Piterbarg [92]. Owing to our approximation formulas, we retrieve that

$$\bar{\nu} = \sqrt{\frac{\int_0^T \nu_t^2 dt}{T}}, \qquad \bar{\beta} = \int_0^T \beta_t \rho_t dt, \qquad \text{with } \rho_t = \frac{\nu_t^2 \int_0^t \nu_s^2 ds}{\int_0^T \nu_t^2 \int_0^t \nu_s^2 ds}. \tag{7.16}$$

Proof. In the context of lognormal proxy (β close to 1), we take

$$\sigma(t,x) = \nu_t e^{(\beta_t - 1)x}, \quad \mu(t,x) = -\frac{1}{2}\sigma^2(t,x), \quad h(x) = e^{-\int_0^T r_s ds}(e^{\int_0^T (r_s - q_s)ds} e^x - K)_+.$$

Then, our approximation formula (7.9) depends only on two constants $\int_0^T \nu_t^2 dt$ and $\int_0^T \nu_t^2 \int_t^T (\beta_s - 1)\nu_s^2 ds dt$. Consequently, two models must coincide with respect these two quantities in order to provide the same approximation formula (with lognormal proxy) up to second order. This easily leads to

the identification (7.16).
When the model is close to normal proxy (β close to 0), we take

$$\sigma(t,x) = \nu_t x^{\beta_t}, \qquad \mu(t,x) = 0, \qquad h(x) = e^{-\int_0^T r_s ds}(e^{\int_0^T (r_s-q_s)ds}x - K)_+.$$

Then, using a similar approach based on formula (7.15), one retrieves exactly the same averaged parameters (7.16).

We conjecture that the averaging rule (7.16) is true not only for β close to 0 or 1, but also for various values in between. A numerical result (see Table 7.4) illustrates this averaging property. □ □

7.3 Numerical Experiments

In this section, we compare approximation formulas given in Theorem 7.2.1, Theorem 7.2.2 and Theorem 7.2.3, applied to Dupire model for call option. We assume that the risk-free rate and the dividend yield are both set at 0. For the following numerical results, we choose a CEV-type function for the local volatility. When the model is applied directly to the asset price (see (7.12) and Theorem 7.2.3), we have

$$\sigma(t,x) = \nu_t x^{\beta_t}, \qquad \mu(t,x) = 0, \qquad h(x) = (x-K)_+.$$

When the model is used for the log-asset price (see (7.1), Theorems 7.2.1 and 7.2.2), we have

$$\sigma(t,x) = \nu_t e^{(\beta_t-1)x}, \qquad \mu(t,x) = -\frac{1}{2}\sigma^2(t,x), \qquad h(x) = (e^x - K)_+.$$

When the functions $(\nu_t)_t$ and $(\beta_t)_t$ do not depend on time (and thus are constant), we use the closed formula for call price [104] as a benchmark. Otherwise, for time dependent functions, we use PDE methods to obtain reference values.

7.3.1 Accuracy of the second order formula (7.9) (based on a log-normal proxy)

Constant parameters. In the case of time independent volatility, the coefficient $C_{1,T}$ becomes:

$$C_{1,T} = \sigma_0^3 \sigma_0^{(1)} \frac{T^2}{2}.$$

In Table 7.1, we report related numerical results, which show that our formula is very accurate (errors in implied volatilities are smaller[4] than 2 bp) for β close to 1. This is coherent with the estimate of the error term $Resid_2$, because this model is close to the lognormal one. In Table 7.2, analogous tests are reported with $\beta = 0.2$. Here, the errors are roughly equal to 20 bp, which is quite satisfactory. This case motivates the use of the third order approximation formula to obtain a better accuracy, this is discussed in the following subsection (see Table 7.6).

Piecewise constant parameters. Here, the functions ν and β are piecewise constant on each interval $[T_i, T_{i+1}[$ for each $i \le n$. Therefore, $C_{1,.}$ can be calculated recursively

$$C_{1,T_{i+1}} = C_{1,T_i} + (T_{i+1}-T_i)\sigma_{T_i}\sigma_{T_i}^{(1)} \sum_{j=1}^{i-1} \sigma_{T_j}^2 (T_{j+1}-T_j) + \frac{(T_{i+1}-T_i)^2}{2}\sigma_{T_i}^3 \sigma_{T_i}^{(1)},$$

[4] 1 bp on implied volatilities is equal to 0.01%.

Table 7.1: Errors on implied Black-Scholes volatilities (in bp) between the second order approximation formula (7.9) and the closed formula for CEV model (6.2), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\beta = 0.8$, $\nu = 0.2$ and $x_0 = 0$.

T/K	80%	90%	100%	110%	120%
6M	-1.63	-0.22	-0.08	-0.17	-0.86
1Y	-1.11	-0.26	-0.15	-0.22	-0.63
1.5Y	-0.98	-0.32	-0.21	-0.28	-0.60
2Y	-0.95	-0.38	-0.28	-0.34	-0.62
3Y	-0.98	-0.51	-0.41	-0.46	-0.69
5Y	-1.16	-0.77	-0.67	-0.70	-0.89
10Y	-1.70	-1.37	-1.26	-1.27	-1.40

Table 7.2: Errors on implied Black-Scholes volatilities (in bp) between the second order approximation formula (7.9) and the closed formula for CEV model (6.2), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\beta = 0.2$, $\nu = 0.2$ and $x_0 = 0$.

T/K	80%	90%	100%	110%	120%
6M	-22.85	-3.33	-1.07	-2.61	-14.87
1Y	-16.60	-4.07	-2.14	-3.21	-10.20
1.5Y	-15.21	-5.11	-3.21	-4.03	-9.31
2Y	-15.13	-6.23	-4.27	-4.92	-9.29
3Y	-16.36	-8.53	-6.39	-6.74	-10.12
5Y	-20.47	-13.19	-10.60	-10.42	-12.74
10Y	-32.01	-24.45	-20.77	-19.45	-20.26

with $C_{1,T_1} = \sigma_0^3 \sigma_0^{(1)} \frac{T_1{}^2}{2}$. In our tests, the piecewise constant functions ν and β are equal respectively on each interval of the form $[\frac{i}{20}, \frac{i+1}{20}[$ to $25\% - i \times 0.11\%$ and $100\% - i \times 0.75\%$. Results given in Table 7.3 show that our second order approximation formula is still very accurate for time dependent parameters ν and β. Using the same time dependent coefficients, we test the parameter averaging principle, that is described in paragraph 7.2.4. Results are reported in Table 7.4. The accuracy is still very good.

7.3.2 Accuracy of the third order formula (7.11)

Constant parameters. Tables 7.5 and 7.6 show that the third order approximation (7.11) is very good for various values of β. The use of this formula has much improved the accuracy in the case $\beta = 0.2$, for which the model is not close to the log-normal proxy.

7.3.3 Accuracy of the third order formula using normal approximation

Constant parameters. Tables 7.7 and 7.8 show that the third order approximation (7.14) is also very good for various values of β. The computation of this formula is slightly quicker than that with a log-normal proxy, because there are fewer terms.

Table 7.3: Errors on implied Black-Scholes volatilities (in bp) between the second order approximation formula (7.9) and the PDE method, expressed as a function of maturities in fractions of years and relative strikes. Parameters: time dependent ν and β, $x_0 = 0$.

T/K	80%	90%	100%	110%	120%
6M	-0.67	-0.09	0.03	-0.07	-0.35
1Y	-0.44	0.10	0.06	-0.09	-0.26
1.5Y	-0.38	-0.13	0.09	0.11	-0.25
2Y	0.37	0.15	-0.11	-0.14	-0.26

Table 7.4: Errors on implied Black-Scholes volatilities (in bp) between the closed CEV formula (6.2) applied to an equivalent CEV model (7.16) and the PDE method, expressed as a function of relative strikes. Parameters: time dependent ν and β, $x_0 = 0$ and $T = 1Y$.

T/K	80%	90%	100%	110%	120%
1Y	0,09	-0,27	-0,20	-0,07	0,00

7.4 General results about error analysis

In this section, we analyse the error terms according to the payoff smoothness (smooth, vanilla or binary). To accomplish this, we first give some notations that will be used throughout the theorems and the proofs. Then, we provide a general expansion formula of the price $\mathbb{E}[h(X_T)]$ at any order, making explicit the order of magnitude of each term. This expansion is different according to the payoff smoothness: smooth payoff in Theorem 7.4.1, vanilla payoff in Theorem 7.4.2 under an additional ellipticity condition on σ and binay payoff in Theorem 7.4.3.

For the three cases, we discuss the form of error estimates. We show that the second order approximation formula (7.9) (and those at any order) is accurate under one of the following conditions:

- the maturity of the option T is small.
- the derivatives of the volatility σ and the drift μ w.r.t. the second variables are small. This is measured by the constant M_1 defined in (7.5). In particular, the model and the proxy coincide ($X \equiv X^{BS}$) when these derivatives vanish ($M_1 = 0$, see remark 7.1.1). This is coherent with our estimates since the correction and the error terms are estimated as $O(M_1)$ where O is the Landau symbol.
- The volatility, the drift and their derivatives are small. This dependence is represented using the constant M_0 defined in (7.4).

Moreover, when the three conditions are all satisfied, the approximation formula becomes even more accurate.

All the proofs are given in Section 7.5.

Notations.

Table 7.5: Error in implied Black-Scholes volatilities (in bp) between the third order approximation formula (7.11) and the closed formula for CEV model (6.2), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\beta = 0.8$, $\nu = 0.2$ and $x_0 = 0$.

T/K	80%	90%	100%	110%	120%
6M	-0.08	-0.02	-0.01	0.00	0.00
1Y	-0.06	-0.03	-0.01	-0.01	0.00
1.5Y	-0.06	-0.03	-0.02	-0.01	0.00
2Y	-0.06	-0.04	-0.02	-0.01	0.00
3Y	-0.08	-0.05	-0.03	-0.01	0.00
5Y	-0.10	-0.06	-0.04	-0.01	0.01
10Y	-0.16	-0.10	-0.06	-0.02	0.01

Table 7.6: Error in implied Black-Scholes volatilities (in bp) between the third order approximation formula (7.11) and the closed formula for CEV model (6.2), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\beta = 0.2$, $\nu = 0.2$ and $x_0 = 0$.

T/K	80%	90%	100%	110%	120%
6M	-1.23	-0.18	-0.01	0.12	0.53
1Y	-0.93	-0.34	-0.03	0.22	0.52
1.5Y	-1.19	-0.51	-0.06	0.31	0.68
2Y	-1.51	-0.68	-0.09	0.39	0.85
3Y	-2.22	-1.05	-0.19	0.52	1.17
5Y	-3.71	-1.87	-0.47	0.67	1.69
10Y	-7.32	-4.13	-1.56	0.55	2.38

- *About floating constants and upper bounds.* In the following statements and proofs, for the upper bounds we use numerous constants, that are not relabelled during the computations. We simply use the unique notation
$$A \le_c B$$
to assert that $A \le cB$, where c is a positive constant depending on the model parameters M_0, M_1, T, C_E (defined in assumption (E)) and on other universal constants. The constant c remains bounded when the model parameters go to 0, and it is uniform w.r.t. the parameter $\varepsilon \in [0,1]$. When informative, we make clear the dependence of upper bounds w.r.t. M_0, M_1 and T.

- *Model differentiation.* In the proofs, the derivatives of the parameterized process X^ε are useful: they are defined by $X^\varepsilon_{i,t} = \frac{\partial^i X^\varepsilon_t}{\partial \varepsilon^i}$ when these derivatives have a meaning. Additionally, we write:

$$Y^\varepsilon_T = X^\varepsilon_T - (x_0 + \varepsilon X_{1,T}), \qquad Y^\varepsilon_{k,i,T} = \frac{\partial^i((Y^\varepsilon_T)^k)}{\partial \varepsilon^i}, \qquad Y_{k,i,T} = Y^0_{k,i,T},$$
$$R_{k,i,T} = \frac{\int_0^1 Y^{(1-\lambda)}_{k,i+1,T} \lambda^i d\lambda}{i!}.$$

Table 7.7: Error in implied Black-Scholes volatilities (in bp) between the third order approximation formula (7.14) and the closed formula for CEV model (6.2), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\beta = 0.8$, $\nu = 0.2$ and $x_0 = 0$.

T/K	80%	90%	100%	110%	120%
6M	-1.61	-0.07	-0.01	0.03	0.77
1Y	-0.88	-0.08	-0.02	0.03	0.45
1.5Y	-0.61	-0.11	-0.02	0.04	0.31
2Y	-0.51	-0.15	-0.03	0.06	0.25
3Y	-0.49	-0.23	-0.05	0.10	0.23
5Y	-0.71	-0.44	-0.11	0.16	0.30
10Y	-1.70	-1.09	-0.37	0.22	0.56

Table 7.8: Error in implied Black-Scholes volatilities (in bp) between the third order approximation formula (7.14) and the closed formula for CEV model (6.2), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\beta = 0.2$, $\nu = 0.2$ and $x_0 = 0$.

T/K	80%	90%	100%	110%	120%
6M	0.22	0.06	-0.01	-0.06	-0.16
1Y	0.41	0.11	0.00	-0.10	-0.26
1.5Y	0.56	0.17	0.00	-0.13	-0.34
2Y	0.71	0.24	0.02	-0.16	-0.41
3Y	1.02	0.39	0.06	-0.20	-0.53
5Y	1.75	0.79	0.21	-0.23	-0.71
10Y	4.71	2.55	1.15	0.10	-0.84

- *Miscellaneous.* As usual, the $\mathbf{L}_p$-norm of a real random variable Z is denoted by $\|Z\|_p = [\mathbb{E}|Z|^p]^{1/p}$.

7.4.1 Error analysis for smooth payoff

Theorem 7.4.1. ***Asymptotic expansion for the price of smooth payoff ($h \in \mathscr{C}_0^\infty(R)$).***
For $m \geq 2$ assume that (R_{m+2}) holds. If the payoff h fulfills Assumption (H_1), then one has

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^{BS})] + \sum_{i=2}^{m} Ord_i + Resid_m, \tag{7.17}$$

where different terms are estimated as follows.

- *The contribution for order $i \in \{2,\dots,m\}$: $Ord_i = \sum_{k=1}^{\lfloor \frac{i}{2} \rfloor} \mathbb{E}[h^{(k)}(X_T^{BS}) \frac{Y_{k,i,T}}{k!i!}]$ and it is estimated by*

$$|Ord_i| \leq_c \sup_{1\leq j \leq \lfloor \frac{i}{2} \rfloor -1} |h^{(j)}|_\infty M_1 M_0^{i-1} (\sqrt{T})^i. \tag{7.18}$$

- *The residual term for order m is :* $Resid_m = \mathbb{E}[\sum_{k=1}^{[\frac{m}{2}]} h^{(k)}(X_T^{BS})\frac{R_{k,m,T}}{k!} + \frac{(Y_T^1)^{\lfloor\frac{m}{2}\rfloor+1}}{\lfloor\frac{m}{2}\rfloor!}\int_0^1 h^{(\lfloor\frac{m}{2}\rfloor+1)}(vX_T+(1-v)(X_T^{BS}))(1-v)^{\lfloor\frac{m}{2}\rfloor}dv]$, *such that*

$$|Resid_m| \leq_c \sup_{1\leq j\leq\lfloor\frac{m}{2}\rfloor}|h^{(j)}|_\infty M_1 M_0^m(\sqrt{T})^{m+1}. \tag{7.19}$$

In the *multiplicative case* ($\sigma(t,x) = \Delta a(t,x)$ and $\mu(t,x) = \Delta b(t,x)$), we have $M_0 \leq_c \Delta$ and $M_1 \leq_c \Delta$. Thus, we obtain

$$Ord_i = O((\Delta\sqrt{T})^i) \quad \text{for } 2 \leq i \leq m, \qquad\qquad Resid_m = O((\Delta\sqrt{T})^{m+1}).$$

This justifies that Equation (7.17) should be viewed as an approximation formula of order m.

Notice that the above theorem provides which terms have to be computed to achieve a given accuracy. But to effectively compute these terms as a summation of Greeks (as in Theorems 7.2.1 and 7.2.2), we shall use results in Appendix 7.6.

7.4.2 Error analysis for vanilla payoff

The payoff h for this kind of option is not necessarily smooth, it is almost everywhere differentiable and belongs to the space $\mathscr{H}$. The previous expansion in the case of smooth payoff is no more valid. Indeed, the i-th order contribution Ord_i has been represented using the derivatives of $h^{(1)}$ that do not necessarily exist anymore. Therefore we introduce some new variables in order to represent higher contributions only using $h^{(1)}$ (and not higher order derivatives).

Lemma 7.4.1. *Given $m \geq 2$, assume (R_{3m-2}) and (E). Let $v \in [0,1]$. There exist random variables $(G_i)_{2\leq i\leq m}, S_m, I_{m,v} \in \cap_{p\geq 1}\mathbf{L}_p$ such that for any $l \in \mathscr{C}_0^\infty(\mathbb{R})$, one has*

$$\sum_{k=1}^{i-1}\frac{1}{k!}\mathbb{E}[l^{(k)}(X_T^{BS})\frac{Y_{k,k+i-1,T}}{(k+i-1)!}] = \mathbb{E}[l^{(1)}(X_T^{BS})G_i] \quad \textit{for } 2\leq i\leq m,$$

$$\sum_{k=1}^{m-1}\frac{1}{k!}\mathbb{E}[l^{(k)}(X_T^{BS})R_{k,k+m-1,T}] = \mathbb{E}[l^{(1)}(X_T^{BS})S_m],$$

$$\mathbb{E}[\frac{(Y_T^1)^m}{(m-1)!}l^{(m)}(vX_T+(1-v)X_T^{BS})] = \mathbb{E}[l^{(1)}(vX_T+(1-v)X_T^{BS})I_{m,v}].$$

Additionally, we have for any $p \geq 1$

$$\|G_i\|_p \leq_c (\frac{M_0}{\sigma_{inf}})^{i-2}M_1M_0^{i-1}(\sqrt{T})^i, \tag{7.20}$$

$$\|S_m\|_p + \sup_{v\in[0,1]}\|I_{m,v}\|_p \leq_c (\frac{M_0}{\sigma_{inf}})^{m-1}M_1M_0^m(\sqrt{T})^{m+1}. \tag{7.21}$$

The proof of this lemma is postponed to Subsection 7.5.2.

The random variables in the above lemma are now used to represent conveniently successive contributions in the general approximation formula for vanilla payoffs. This is the following statement.

Theorem 7.4.2. ***Asymptotic expansion for the price of vanilla payoff*** *($h \in \mathscr{H}$ and $h' \in \mathscr{H}$). Given $m \geq 2$, assume (R_{3m-2}) and (E). If the payoff h fulfills Assumption (H_2), then we have*

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^{BS})] + \sum_{i=2}^{m} Ord_i + Resid_m, \tag{7.22}$$

where different terms are estimated as follows.

- *The contribution of order $i \in \{2,\ldots,m\}$ is $Ord_i = \mathbb{E}[h^{(1)}(X_T^{BS})G_i]$ and it is estimated by:*

$$|Ord_i| \leq_c \|h^{(1)}(X_T^{BS})\|_2 (\frac{M_0}{\sigma_{inf}})^{i-2} M_1 M_0^{i-1} (\sqrt{T})^i. \tag{7.23}$$

- *The residual for order m is $Resid_m = \mathbb{E}[h^{(1)}(X_T^{BS})S_m] + \int_0^1 \mathbb{E}[h^{(1)}(vX_T + (1-v)X_T^{BS})I_{m,v}](1-v)^{m-1}dv$, such that*

$$\begin{aligned}|Resid_m| \leq_c & (\|h^{(1)}(X_T^{BS})\|_2 + \sup_{v\in[0,1]} \|h^{(1)}(vX_T + (1-v)X_T^{BS})\|_2) \\ & (\frac{M_0}{\sigma_{inf}})^{m-1} M_1 M_0^m (\sqrt{T})^{m+1}.\end{aligned} \tag{7.24}$$

Notice that the error term in Theorem 7.2.1 for vanilla payoff is $Resid_2$. For the third order approximation formula of Theorem 7.2.2, it is $Resid_3$. Let us comment on the above theorem.

- The label Ord_i is due to the fact that this term is bounded by $M_1 M_0^{i-1}(\sqrt{T})^i$ multiplied by an ellipticity factor of the form $(\frac{M_0}{\sigma_{inf}})^n$. This ellipticity factor is new compared to the case of smooth payoffs. To have a clear view on each contribution, one should have in mind the multiplicative case ($\sigma(t,x) = \Delta a(t,x)$ and $\mu(t,x) = \Delta b(t,x)$) which leads to $\max(M_1, M_0) \leq_c \Delta$ and

  $$Ord_i = O((\Delta\sqrt{T})^i) \quad \text{for } 2 \leq i \leq m, \qquad Resid_m = O((\Delta\sqrt{T})^{m+1}).$$

  That is why we refer to Equation (7.22) as an approximation formula of order m.

- Correction terms are brought together in a different way than in the case of smooth payoffs. Indeed, the hierarchy (in terms of amplitudes) is modified according to the payoff smoothness. However, it is easy to check that the second order approximation is the same for smooth payoffs and vanilla ones. For higher orders, there is no more coincidence with the smooth case.

- Similarly to the smooth case, the above formula provides the appropriate terms to compute to reach a given level of accuracy. It remains to explicitly compute these terms as a summation of Greeks, using results in Appendix 7.6. This is done in Theorems 7.2.1 and 7.2.2 for $m = 2$ and $m = 3$.

- Finally to accommodate irregular payoffs, we require extra smoothness properties on μ and σ.

7.4.3 Error analysis for binary payoff

For this kind of option, the payoff h is not necessarily smooth, it is at least in $\mathscr{H}$. The results below are easy extensions of the case of vanilla options, we leave the proof to the reader.

Lemma 7.4.2. *Given $m \geq 1$, assume (R_{3m+2}) and (E). Let $v \in [0,1]$. There exist random variables $(P_i)_{1\leq i\leq m}, Q_m, T_{m,v} \in \cap_{p\geq 1} \mathbf{L}_p$ such that, for any $l \in \mathscr{C}_0^\infty(\mathbb{R})$, one has:*

$$\sum_{k=1}^{i} \frac{1}{k!}\mathbb{E}[l^{(k)}(X_T^{BS})\frac{Y_{k,k+i,T}}{(k+i)!}] = \mathbb{E}[l(X_T^{BS})P_i] \quad \textit{for } 1 \leq i \leq m,$$
$$\sum_{k=1}^{m} \frac{1}{k!}\mathbb{E}[l^{(k)}(X_T^{BS})R_{k,k+m,T}] = \mathbb{E}[l(X_T^{BS})Q_m],$$
$$\mathbb{E}[\frac{(Y_T^1)^{m+1}}{m!} l^{(m+1)}(vX_T + (1-v)X_T^{BS})] = \mathbb{E}[l(vX_T + (1-v)X_T^{BS})T_{m,v}].$$

Moreover, they are estimated in the $\mathbf{L}_p$ norm as follows:

$$\|P_i\|_p \leq_c (\frac{M_0}{\sigma_{inf}})^i M_1 M_0^{i-1}(\sqrt{T})^i,$$
$$\|Q_m\|_p + \sup_{v\in[0,1]} \|T_{m,v}\|_p \leq_c (\frac{M_0}{\sigma_{inf}})^{m+1} M_1 M_0^m (\sqrt{T})^{m+1}.$$

We are now in a position to state an expansion formula of order m.

Theorem 7.4.3. ***Asymptotic expansion for the price of binary payoff** ($h \in \mathscr{H}$).*
Given $m \geq 1$, assume (R_{3m+2}) and (E). If the payoff h fulfills Assumption (H_3), then we have

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^{BS})] + \sum_{i=1}^{m} Ord_i + Resid_m, \tag{7.25}$$

where different terms are as follows.

- *The contribution for order $i \in \{1, \ldots, m\}$ is $Ord_i = \mathbb{E}[h(X_T^{BS})P_i]$ and it is estimated by:*

$$|Ord_i| \leq_c \|h(X_T^{BS})\|_2 (\frac{M_0}{\sigma_{inf}})^i M_1 M_0^{i-1}(\sqrt{T})^i. \tag{7.26}$$

- *The residual term for order m is $Resid_m = \mathbb{E}[h(X_T^{BS})Q_m] + \int_0^1 \mathbb{E}[h(vX_T + (1-v)X_T^{BS})T_{m,v}](1-v)^m dv$, such that*

$$|Resid_m| \leq_c (\|h(X_T^{BS})\|_2 + \sup_{v\in[0,1]} \|h(vX_T + (1-v)X_T^{BS})\|_2)$$
$$(\frac{M_0}{\sigma_{inf}})^{m+1} M_1 M_0^m (\sqrt{T})^{m+1}. \tag{7.27}$$

Notice that the second order approximation for smooth payoffs and vanilla options is only a first order approximation for binary options. This is due to the lack of regularity of the binary payoffs.

7.5 Proofs

For the following, we use the same definitions and notations as in Chapter 1 of [88]. Before giving the proofs for the main theorems, we need to upper bound the $\mathbf{L}_p$ norm of the derivatives $X_{i,t}^\varepsilon$ to state Theorem 7.4.1, to upper bound also the $\mathbf{L}_p$ norm of the Malliavin derivatives $D_{t_1,\cdots,t_j}^j X_{i,t}^\varepsilon$, and use the key lemma 7.5.3 in order to state Theorems 7.4.2 and 7.4.3.

7.5.1 Proof of Theorem 7.4.1 (Smooth payoff)

The proof of Theorem 7.4.1 is performed through two steps:

- **Step 1:** Upper bound the $\mathbf{L}_p$ norm of $X^{\varepsilon}_{i,t}$.
- **Step 2:** Completion of the proof of Theorem 7.4.1.

We first recall that $\varepsilon \to X^{\varepsilon}_t$ is almost surely C^{N-1} w.r.t. ε under assumption (R_N).

Step 1: Upper bounds for the $\mathbf{L}_p$ norm of $X^{\varepsilon}_{i,t}$

We aim at proving the following result, which may be useful, independently of our work.

Theorem 7.5.1. *Given $N \geq 2$, assume (R_N). For every $\varepsilon \in [0,1]$ and $p \geq 1$, we have*

$$\sup_{t\leq T} \|X^{\varepsilon}_{1,t}\|_p \leq_c M_0\sqrt{T}; \tag{7.28}$$

$$\sup_{t\leq T} \|X^{\varepsilon}_{i,t}\|_p \leq_c M_1 M_0^{i-1}(\sqrt{T})^i, \quad \forall i \in \{2,\ldots,N-1\}. \tag{7.29}$$

A meaning of the first inequality is that the first derivative has the same amplitude as the implicit total standard deviation $M_0\sqrt{T}$. The second inequality shows that the bounds of the derivative estimates decrease successively by the implicit total standard deviation $M_0\sqrt{T}$. Furthermore, the dependence w.r.t. the constant M_1 shows that the derivatives $(X_{i,t})_{i\geq 2}$ are null if the function σ and μ are deterministic (see Remark 7.1.1). In this case, X is the Black-Scholes model.

Proof. The existence of any moment is easy to establish, we will skip details. In the following, we rather focus on their dependence w.r.t. M_0, M_1 and $\sqrt{T}$.

Clearly, it is sufficient to prove estimates for $p \geq 2$. Take $p \geq 2$, note that $X^{\varepsilon}_{1,.}$ is the solution of the linear SDE:

$$\begin{aligned} dX^{\varepsilon}_{1,t} &= \sigma_t(X^{\varepsilon}_t)dW_t + \mu_t(X^{\varepsilon}_t)dt + \varepsilon X^{\varepsilon}_{1,t}(\sigma^{(1)}_t(X^{\varepsilon}_t)dW_t + \mu^{(1)}_t(X^{\varepsilon}_t)dt), \\ X^{\varepsilon}_{1,0} &= 0. \end{aligned}$$

To estimate the $\mathbf{L}_p$ norm of the solution of the above *linear* equation, we state a lemma, that will be repeatedly used in the following computations.

Lemma 7.5.1. *Assume that Z is an Itô process such that*

i) $\sup_{t\leq T}\|Z_t\|_p < +\infty$ *for some $p \geq 2$;*

ii) Z solves a linear equation

$$Z_t = \int_0^t Z_s(a_s dW_s + b_s ds) + \int_0^t \alpha_s dW_s + \beta_s ds,$$

where $\sup_{t\leq T}(\|\alpha_t\|_p + \|\beta_t\|_p) < +\infty$, a and b are bounded.

Then, for a constant c (depending only on p and T), we have

$$\sup_{t\leq T}\|Z_t\|_p \leq c \sup_{t\leq T}(\|\alpha_t\|_p + \|\beta_t\|_p)\sqrt{T} e^{c(|a|_\infty + |b|_\infty)^p T^{p/2}}. \tag{7.30}$$

The proof is quite standard: it results from easy calculations using BDG inequalities and Gronwall's lemma. We omit further details.

From this, it readily follows that

$$\sup_{t\le T}\|X^{\varepsilon}_{1,t}\|_p \le_c \max(|\sigma|_\infty,|\mu|_\infty)\sqrt{T} \le_c M_0\sqrt{T}.$$

This proves the first inequality (7.28).

We now prove the second inequality (7.29) which is not straightforward. To accomplish this, it is useful to scale the parameters. Let us define the new variables:

$$\tilde{X}^{\varepsilon}_t = X_t^{\frac{\varepsilon}{M_0\sqrt{T}}}, \tag{7.31}$$

$$\tilde{\sigma}(t,x) = \frac{\sigma(t,x)}{M_0}, \qquad \tilde{\mu}(t,x) = \frac{\mu(t,x)}{M_0}. \tag{7.32}$$

From Equation (7.3), one obtains the dynamics of the rescaled process $(\tilde{X}^{\varepsilon}_t)_t$:

$$d\tilde{X}^{\varepsilon}_t = \varepsilon(\tilde{\sigma}_t(\tilde{X}^{\varepsilon}_t)\frac{dW_t}{\sqrt{T}} + \tilde{\mu}_t(\tilde{X}^{\varepsilon}_t)\frac{dt}{\sqrt{T}}), \tilde{X}^{\varepsilon}_0 = x_0, \tag{7.33}$$

where $\varepsilon \in [0, M_0\sqrt{T}]$. The advantage of this change of parameters is that the constant M_0 associated to the new coefficients $\tilde{\sigma}$ and $\tilde{\mu}$ is bounded by 1 (thus, it is model-free):

$$\max(|\tilde{\sigma}|_\infty,\cdots,|\tilde{\sigma}^{(N)}|_\infty,|\tilde{\mu}|_\infty,\cdots,|\tilde{\mu}^{(N)}|_\infty) = 1.$$

Additionally, there is a simple relation between derivatives of X^{ε} and those of $\tilde{X}^{\varepsilon}$:

$$\tilde{X}^{\varepsilon}_{i,t} \equiv \frac{\partial^i(\tilde{X}^{\varepsilon}_t)}{\partial\varepsilon^i} = \frac{\partial^i(X_t^{\frac{\varepsilon}{M_0\sqrt{T}}})}{\partial\varepsilon^i} = \frac{1}{(M_0\sqrt{T})^i}X_{i,t}^{\frac{\varepsilon}{M_0\sqrt{T}}}.$$

Using this notation, the proof of Inequality (7.29) is reduced to prove that

$$\sup_{t\le T}\|\tilde{X}^{\varepsilon}_{i,t}\|_p \le_c \frac{M_1}{M_0}. \tag{7.34}$$

for every $\varepsilon \in [0, M_0\sqrt{T}]$ and $i \in \{2,\ldots,N-1\}$.

Proof of (7.34) . By successive differentiation of (7.33), it is not hard to prove

$$\sup_{t\le T}\|\tilde{X}^{\varepsilon}_{i,t}\|_p \le_c 1. \tag{7.35}$$

Indeed, we obtain linear SDEs[5] solved by $\tilde{X}^{\varepsilon}_{i,\cdot}$, to which we can apply Lemma 7.5.1. It gives uniform bounds because the arising processes (a,b,α,β) are proportional to $1/\sqrt{T}$ and then multiplied by $\sqrt{T}$ in Lemma 7.5.1. Another heuristic argument, to get that the bound (7.35) is indeed equal to 1, is the following: on the one hand, the integrands $\frac{W_t}{\sqrt{T}}$ and $\frac{t}{\sqrt{T}}$ in the SDE (7.33) are $O(1)$ over the maturity T. On the other hand, the uniform bounds for the derivatives of $\tilde{\sigma}$ and $\tilde{\mu}$ up to order N are smaller than 1. Consequently, the $\mathbf{L}_p$ estimates (7.35) remain uniformly bounded.

However, the inequality (7.35) is not equivalent to the inequality (7.34) because $\frac{M_1}{M_0} \le 1$. But this preliminary estimate is useful to establish the final one as follows. To prove the required inequality, we first show that $\tilde{X}^{\varepsilon}_{i,\cdot}$ solves a linear equation, this is stated in the following proposition.

[5] this is fully justified in Proposition 7.5.1.

Proposition 7.5.1. *Given $N \geq 2$, assume (R_N). For $2 \leq i \leq N-1$, $\tilde{X}_{i,.}$ is the solution of the linear SDE:*

$$d\tilde{X}^{\varepsilon}_{i,t} = dH^{\varepsilon}_{i,t} + \tilde{X}^{\varepsilon}_{i,t} dL^{\varepsilon}_t, \qquad \tilde{X}^{\varepsilon}_{i,0} = 0, \tag{7.36}$$
$$dL^{\varepsilon}_t = \varepsilon(\tilde{\sigma}^{(1)}_t(\tilde{X}_t)\frac{dW_t}{\sqrt{T}} + \tilde{\mu}^{(1)}_t(\tilde{X}_t)\frac{dt}{\sqrt{T}}),$$
$$dH^{\varepsilon}_{i,t} = P^{\varepsilon}_{\tilde{\sigma},i,t}\frac{dW_t}{\sqrt{T}} + P^{\varepsilon}_{\tilde{\mu},i,t}\frac{dt}{\sqrt{T}},$$

where the processes $(P^{\varepsilon}_{\tilde{\sigma},i,t})_{t\geq 0}$ and $(P^{\varepsilon}_{\tilde{\mu},i,t})_{t\geq 0}$ are defined in the proof.

Proof. Take $i \geq 2$, the SDE for the i^{th} derivative is obtained from Equation (7.33) using differentiation under the integral sign (see [72]):

$$d\tilde{X}^{\varepsilon}_{i,t} = \frac{\partial^i(\varepsilon\tilde{\sigma}_t(\tilde{X}^{\varepsilon}_t))}{\partial\varepsilon^i}\frac{dW_t}{\sqrt{T}} + \frac{\partial^i(\varepsilon\tilde{\mu}_t(\tilde{X}^{\varepsilon}_t))}{\partial\varepsilon^i}\frac{dt}{\sqrt{T}}, \qquad \tilde{X}^{\varepsilon}_{i,0} = 0. \tag{7.37}$$

The application of the Leibniz formula for the i^{th} derivative of the product (that is $(\varepsilon f(\varepsilon))^{(i)} = \varepsilon f^{(i)}(\varepsilon) + i f^{(i-1)}(\varepsilon))$ gives:

$$\frac{\partial^i(\varepsilon\tilde{\sigma}_t(\tilde{X}^{\varepsilon}_t))}{\partial\varepsilon^i} = \varepsilon\frac{\partial^i(\tilde{\sigma}_t(\tilde{X}^{\varepsilon}_t))}{\partial\varepsilon^i} + i\frac{\partial^{i-1}(\tilde{\sigma}_t(\tilde{X}^{\varepsilon}_t))}{\partial\varepsilon^{i-1}}.$$

Using the Faà di Bruno formula for derivative of composite function (apply Lemma 7.6.4 with $g(x) = \tilde{\sigma}_t(x)$ and $f(\varepsilon) = \tilde{X}^{\varepsilon}_t$), one obtains

$$\begin{aligned}\frac{\partial^i(\varepsilon\tilde{\sigma}(t,\tilde{X}^{\varepsilon}_t))}{\partial\varepsilon^i} =& \varepsilon \sum_{\substack{k=(k_1,\cdots,k_i)\in\mathbb{N}^i \\ \sum_{j=1}^i jk_j = i}} d_k \tilde{\sigma}_t^{(\sum_{j=1}^i k_j)}(\tilde{X}^{\varepsilon}_t)\prod_{j=1}^{i}(\tilde{X}^{\varepsilon}_{j,t})^{k_j} \\ &+ i \sum_{\substack{k=(k_1,\cdots,k_{i-1})\in\mathbb{N}^{i-1} \\ \sum_{j=1}^{i-1} jk_j = i-1}} d_k \tilde{\sigma}_t^{(\sum_{j=1}^{i-1} k_j)}(\tilde{X}^{\varepsilon}_t)\prod_{j=1}^{i-1}(\tilde{X}^{\varepsilon}_{j,t})^{k_j}.\end{aligned}$$

Notice that the i^{th} component k_i can take only two values 0 or 1 (because $ik_i \leq \sum_{j=1}^i jk_j = i$). When $k_i = 1$, one has $k_j = 0$ for $j < i$ and $d_k = 1$ (see Lemma 7.6.4). Thus, we obtain

$$\begin{aligned}\frac{\partial^i(\varepsilon\tilde{\sigma}(t,\tilde{X}^{\varepsilon}_t))}{\partial\varepsilon^i} =& \varepsilon\tilde{\sigma}^{(1)}_t(\tilde{X}^{\varepsilon}_t)\tilde{X}^{\varepsilon}_{i,t} \\ &+ \varepsilon \sum_{\substack{k=(k_1,\cdots,k_{i-1},0)\in\mathbb{N}^i \\ \sum_{j=1}^{i-1} jk_j = i}} d_k \tilde{\sigma}_t^{(\sum_{j=1}^{i-1} k_j)}(\tilde{X}^{\varepsilon}_t)\prod_{j=1}^{i-1}(\tilde{X}^{\varepsilon}_{j,t})^{k_j} \\ &+ i \sum_{\substack{k=(k_1,\cdots,k_{i-1})\in\mathbb{N}^{i-1} \\ \sum_{j=1}^{i-1} jk_j = i-1}} d_k \tilde{\sigma}_t^{(\sum_{j=1}^{i-1} k_j)}(\tilde{X}^{\varepsilon}_t)\prod_{j=1}^{i-1}(\tilde{X}^{\varepsilon}_{j,t})^{k_j} \\ :=& \varepsilon\tilde{\sigma}^{(1)}_t(\tilde{X}^{\varepsilon}_t)\tilde{X}^{\varepsilon}_{i,t} + P^{\varepsilon}_{\tilde{\sigma},i,t}.\end{aligned} \tag{7.38}$$

We define analogously $P^{\varepsilon}_{\tilde\mu,i,t}$ by replacing $\tilde\sigma$ by $\tilde\mu$ in the expression (7.38). It writes

$$\frac{\partial^i(\varepsilon\tilde\mu(t,\tilde X^{\varepsilon}_t))}{\partial\varepsilon^i} = \varepsilon\tilde\mu^{(1)}_t(\tilde X^{\varepsilon}_t)\tilde X^{\varepsilon}_{i,t} + P^{\varepsilon}_{\tilde\mu,i,t}. \tag{7.39}$$

The two equalities (7.38) and (7.39) plugged into the relation (7.37) give immediately the result. □ □

End of proof of (7.34). Owing to Equation (7.5.1), $\tilde X_{i,\cdot}$ is the solution of a linear SDE, to which we apply Lemma 7.5.1. We obtain

$$\sup_{t\le T}\|\tilde X^{\varepsilon}_{i,t}\|_p \le_c \sup_{t\le T}\|P^{\varepsilon}_{\tilde\sigma,i,t}\|_p + \sup_{t\le T}\|P^{\varepsilon}_{\tilde\mu,i,t}\|_p.$$

In view of the expression of $P^{\varepsilon}_{\tilde\sigma,i,t}$ in Equation (7.38), using the Hölder inequality and the preliminary estimates (7.35), we obtain

$$\sup_{t\le T}\|P^{\varepsilon}_{\tilde\sigma,i,t}\|_p \le_c \sum_{\substack{k=(k_1,\cdots,k_{i-1},0)\in\mathbb{N}^i\\ \sum_{j=1}^{i-1} jk_j=i}} |\tilde\sigma^{(\sum_{j=1}^{i-1}k_j)}|_\infty + \sum_{\substack{k=(k_1,\cdots,k_{i-1})\in\mathbb{N}^{i-1}\\ \sum_{j=1}^{i-1} jk_j=i-1}} |\tilde\sigma^{(\sum_{j=1}^{i-1}k_j)}|_\infty$$

Since $\sum_{j=1}^{i-1} jk_j \ge 1$ and k_j are integers, we have $\sum_{j=1}^{i-1} k_j \ge 1$. It readily follows

$$\begin{aligned}\sup_{t\le T}\|P^{\varepsilon}_{\tilde\sigma,i,t}\|_p &\le_c \max(|\tilde\sigma^{(1)}|_\infty,\cdots,|\tilde\sigma^{(N-1)}|_\infty) = c\frac{\max(|\sigma^{(1)}|_\infty,\cdots,|\sigma^{(N-1)}|_\infty)}{M_0}\\ &\le_c \frac{M_1}{M_0}.\end{aligned}$$

The same inequality holds for $P^{\varepsilon}_{\tilde\mu,i,t}$, which finishes the proof of (7.34). Consequently, Theorem 7.5.1 is proved. □ □

Step 2: Proof of Theorem 7.4.1 (Smooth payoff)

Before performing the Taylor expansion, we recall the notations:

$$Y^{\varepsilon}_T = X^{\varepsilon}_T - (x_0 + \varepsilon X_{1,T}), \qquad Y^{\varepsilon}_{k,i,T} = \frac{\partial^i((Y^{\varepsilon}_T)^k)}{\partial\varepsilon^i}, \qquad Y_{k,i,T} = Y^0_{k,i,T},$$

$$R_{k,i,T} = \frac{\int_0^1 Y^{(1-\lambda)}_{k,i+1,T}\lambda^i d\lambda}{i!}.$$

Clearly one has $X_T = X_T^{BS} + Y_T^1$. We write

$$\begin{aligned}
\mathbb{E}[h(X_T)] =& \mathbb{E}[h(X_T^{BS})] + \sum_{k=1}^{\lfloor \frac{m}{2} \rfloor} \frac{1}{k!} \mathbb{E}[h^{(k)}(X_T^{BS})(Y_T^1)^k] \\
&+ \int_0^1 \mathbb{E}[\frac{(Y_T^1)^{\lfloor \frac{m}{2} \rfloor + 1}(1-v)^{\lfloor \frac{m}{2} \rfloor}}{\lfloor \frac{m}{2} \rfloor !} h^{(\lfloor \frac{m}{2} \rfloor + 1)}(vX_T + (1-v)X_T^{BS})]dv \\
=& \mathbb{E}[h(X_T^{BS})] + \sum_{k=1}^{\lfloor \frac{m}{2} \rfloor} \frac{1}{k!} \mathbb{E}[h^{(k)}(X_T^{BS})(\sum_{i=2k}^{m} \frac{Y_{k,i,T}}{i!} + R_{k,m,T})] \\
&+ \int_0^1 \mathbb{E}[\frac{(Y_T^1)^{\lfloor \frac{m}{2} \rfloor + 1}(1-v)^{\lfloor \frac{m}{2} \rfloor}}{\lfloor \frac{m}{2} \rfloor !} h^{(\lfloor \frac{m}{2} \rfloor + 1)}(vX_T + (1-v)X_T^{BS})]dv \\
=& \mathbb{E}[h(X_T^{BS})] + \sum_{i=2}^{m} \sum_{k=1}^{[\frac{i}{2}]} \frac{1}{k!} \mathbb{E}[h^{(k)}(X_T^{BS}) \frac{Y_{k,i,T}}{i!}] \\
&+ \sum_{k=1}^{\lfloor \frac{m}{2} \rfloor} \frac{1}{k!} \mathbb{E}[h^{(k)}(X_T^{BS}) R_{k,m,T}] \\
&+ \int_0^1 \mathbb{E}[\frac{(Y_T^1)^{\lfloor \frac{m}{2} \rfloor + 1}(1-v)^{\lfloor \frac{m}{2} \rfloor}}{\lfloor \frac{m}{2} \rfloor !} h^{(\lfloor \frac{m}{2} \rfloor + 1)}(vX_T + (1-v)X_T^{BS})]dv \\
=& \mathbb{E}[h(X_T^{BS})] + \sum_{i=2}^{m} Ord_i + Resid_m,
\end{aligned}$$

where we have used a Taylor expansion twice for the two first identities (notice that $Y_{k,i,T} = 0$ for $i \leq 2k-1$), and we have interchanged the summations for the third one. The equation (7.17) is proved.

Now we establish estimates (7.18) and (7.19). By differentiation of composite function using the Faà di Bruno formula (see Lemma 7.6.4 with $g(x) = x^k$ and $f(\varepsilon) = Y_T^\varepsilon$), we obtain

$$Y_{k,i,T}^\varepsilon = \sum_{\substack{\alpha = (\alpha_1, \cdots, \alpha_i) \in \mathbb{N}^i \\ \sum_{j=1}^i j\alpha_j = i,\ \sum_{j=1}^i \alpha_j \leq k}} d_\alpha \frac{k!}{(\sum_{j=1}^i \alpha_j)!} (Y_T^\varepsilon)^{k - \sum_{j=1}^i \alpha_j} \prod_{j=1}^{i} (Y_{1,j,T}^\varepsilon)^{\alpha_j}. \tag{7.40}$$

Here, we restrict to the indices α such that $\sum_{j=1}^i \alpha_j \leq k$ because we have $g^{(\sum_{j=1}^i \alpha_j)}(x) = 0$ when $\sum_{j=1}^i \alpha_j > k$. Using Equation (7.29), one deduces for each $j \in \{2, \ldots, i\}$ that

$$\|Y_{1,j,T}^\varepsilon\|_p = \|X_{j,T}^\varepsilon\|_p \leq_c M_1 M_0^{j-1} (\sqrt{T})^j \tag{7.41}$$

for any $p \geq 1$. For $j = 1$, the inequality (7.41) is also available because we can write $Y_{1,1,T}^\varepsilon = \int_0^\varepsilon X_{2,T}^\lambda d\lambda$, which readily implies

$$\|Y_{1,1,T}^\varepsilon\|_p \leq_c M_1 M_0 (\sqrt{T})^2 \leq_c M_1 \sqrt{T}.$$

For any indices α, we have the rough estimate $\|(Y_T^\varepsilon)^{k - \sum_{j=1}^i \alpha_j}\|_p \leq_c 1$. Using the above estimate, (7.41) and the Hölder inequality, we finally get

$$\|Y^{\varepsilon}_{k,i,T}\|_p \leq_c \sum_{\substack{\alpha=(\alpha_1,\cdots,\alpha_i)\in\mathbb{N}^i\\ \sum_{j=1}^i j\alpha_j=i}} \prod_{j=1}^{i}(M_1 M_0^{j-1}(\sqrt{T})^j)^{\alpha_j}$$
$$\leq_c \sum_{\substack{\alpha=(\alpha_1,\cdots,\alpha_i)\in\mathbb{N}^i\\ \sum_{j=1}^i j\alpha_j=i}} (\frac{M_1}{M_0})^{\sum_{j=1}^i \alpha_j} M_0^{\sum_{j=1}^i j\alpha_j}(\sqrt{T})^{\sum_{j=1}^i j\alpha_j}$$
$$\leq_c \frac{M_1}{M_0} M_0^i(\sqrt{T})^i = cM_1 M_0^{i-1}(\sqrt{T})^i, \tag{7.42}$$

where we used $\frac{M_1}{M_0} \leq 1$ and $\sum_{j=1}^i \alpha_j \geq 1$ (since $(\alpha_j)_j$ are integers that satisfy $\sum_{j=1}^i j\alpha_j = i \geq 1$). The inequality (7.42) gives immediately the inequality (7.18). It also leads to

$$\|R_{k,m,T}\|_p \leq_c M_1 M_0^m(\sqrt{T})^{m+1}. \tag{7.43}$$

Since $Y^1_T = X_{2,T} + R_{1,2,T}$, one has

$$\|Y^1_T\|_p \leq \|X_{2,T}\|_p + \|R_{1,2,T}\|_p \leq_c M_1 M_0(\sqrt{T})^2 + M_1 M_0^2(\sqrt{T})^3 \leq_c M_1 M_0(\sqrt{T})^2$$

(recall our definition of generic constants). Therefore

$$\|(Y^1_T)^{\lfloor \frac{m}{2}\rfloor+1}\|_p \leq_c M_1^{\lfloor \frac{m}{2}\rfloor+1} M_0^{\lfloor \frac{m}{2}\rfloor+1}(\sqrt{T})^{2(\lfloor \frac{m}{2}\rfloor+1)} \leq_c M_1 M_0^m(\sqrt{T})^{m+1}, \tag{7.44}$$

where we have used $M_1 \leq M_0$ and $2\lfloor \frac{m}{2}\rfloor \geq m-1$. The inequalities (7.43) and (7.44) readily leads to the inequality (7.19). The proof is complete. □

7.5.2 Proof of Lemma 7.4.1

For Malliavin calculus, we use the notation of Nualart [88] for the Sobolev spaces $\mathbb{D}_{k,p}$ associated to the norm $\|.\|_{k,p}$. We divide the proof of Lemma 7.4.1 into three steps:

- **Step 1**: Upper bounds for the $\mathbb{D}^{k,p}$ norm of $X^{\varepsilon}_{i,t}$.
- **Step 2**: Statement of a suitable integration by parts formula (Lemma 7.5.3) in order to handle the irregularity of vanilla payoffs.
- **Step 3**: Completion of the proof of Lemma 7.4.1.

In all this subsection, we assume (R_{3m-2}) for a given $m \geq 2$.

Step 1: Upper Bounds for the $\mathbb{D}^{k,p}$ norm of $X^{\varepsilon}_{i,t}$

The aim of this paragraph is to show that, for every $\varepsilon \in [0,1]$, we have

- $X^{\varepsilon}_T \in \mathbb{D}^{3m-2,\infty}$ with

$$\|DX^{\varepsilon}_T\|_{3m-3,p} \leq_c |\sigma|_\infty \sqrt{T}, \tag{7.45}$$

- for each $i \in \{1, \ldots, 3m-3\}$, $X^{\varepsilon}_{i,T}$ belongs to $\mathbb{D}^{3m-3-i,\infty}$ with

$$\|X^{\varepsilon}_{1,T}\|_{3m-4,p} \leq_c M_0\sqrt{T}, \tag{7.46}$$

$$\|X^{\varepsilon}_{i,T}\|_{3m-3-i,p} \leq_c M_1 M_0^{i-1}(\sqrt{T})^i, \quad i \geq 2. \tag{7.47}$$

Only the proofs of upper bounds need few details. To prove the inequality (7.45), we use the following lemma.

Lemma 7.5.2. *For any $t \in [0,T]$ and $\varepsilon \in [0,1]$, X^{ε}_t belongs to $\mathbb{D}^{3m-2,\infty}$. Moreover, the j first Malliavin derivatives of X^{ε}_t satisfy the following estimates:*

$$\sup_{(t_1,\cdots,t_j)\in[0,T]^j, t\in[0,T]} \|D^j_{t_1,\cdots,t_j} X^{\varepsilon}_t\|_p \leq_c |\sigma|_\infty.$$

Proof. We first take $j=1$; for $t_1 \in [0,t]$, using formula (2.59) in [88] p.126, we have

$$D_{t_1} X^{\varepsilon}_t = \varepsilon\sigma(t_1, X^{\varepsilon}_{t_1}) e^{\int_{t_1}^t \varepsilon(\sigma^{(1)}_s(X^{\varepsilon}_s)dW_s + (\mu^{(1)}_s - \varepsilon\frac{(\sigma^{(1)}_s)^2}{2})(X^{\varepsilon}_s)ds)}.$$

This leads to the announced estimate when $j=1$. The result for $j \geq 2$ is easily obtained by induction. □ □

From the definition of the $\mathbb{D}^{k,p}$ norm, it follows that

$$\begin{aligned}
\|DX^{\varepsilon}_T\|_{3m-3,p} &= (\sum_{j=1}^{3m-3} \mathbb{E}[(\int_0^T \cdots \int_0^T (D_{t_1,\cdots,t_j} X^{\varepsilon}_T)^2 dt_1 \cdots dt_j)^{\frac{p}{2}}])^{\frac{1}{p}} \\
&\leq (\sum_{j=1}^{3m-3} T^{\frac{j}{2}} \sup_{(t_1,\cdots,t_j)\in[0,T]^j} \|D^j_{t_1,\cdots,t_j} X^{\varepsilon}_T\|_p^p)^{\frac{1}{p}} \leq_c |\sigma|_\infty \sqrt{T}
\end{aligned}$$

using Lemma 7.5.2 at the last inequality. This proves the first inequality (7.45).
Now, to establish the upper bounds (7.46) and (7.47), we note that it is equivalent to prove, for every $\varepsilon \in [0, M_0\sqrt{T}]$, that

$$\|\tilde{X}^{\varepsilon}_{1,T}\|_{3m-4,p} \leq_c 1, \tag{7.48}$$

$$\|\tilde{X}^{\varepsilon}_{i,T}\|_{3m-3-i,p} \leq_c \frac{M_1}{M_0}, \quad i \geq 2, \tag{7.49}$$

where $(\tilde{X}^{\varepsilon}_t)_t$ is the rescaled process introduced in (7.31). Using similar arguments as for (7.35), we obtain

$$\|\tilde{X}^{\varepsilon}_{i,T}\|_{N-1-i,p} \leq_c 1, \tag{7.50}$$

for any $\varepsilon \in [0, M_0\sqrt{T}]$. The inequality (7.48) is proved but not (7.49), because $\frac{M_1}{M_0} \leq 1$. To establish (7.49), we proceed as for the proof of Theorem 7.5.1. We will skip further details.

Step 2: Statement of the integration by part Lemma

To handle non-smooth payoffs, our computations rely on a non-degenerate condition on the volatility (stated in assumption (E)). This type of condition is essential to prove the following lemma.

Lemma 7.5.3. *Assume (E) and (R_{k+1}) for a given $k \geq 1$. Let Z belong to $\cap_{p\geq 1}\mathbb{D}^{k,p}$. For any $v \in [0,1]$, there exists a random variable Z_k^v in any $\mathbf{L}_p$ ($p \geq 1$) such that for any function $l \in \mathscr{C}_0^\infty(\mathbb{R})$, we have*

$$\mathbb{E}[l^{(k)}(vX_T + (1-v)X_T^{BS})Z] = \mathbb{E}[l(vX_T + (1-v)X_T^{BS})Z_k^v].$$

Moreover, we have $\|Z_k^v\|_p \leq_c \frac{\|Z\|_{k,2p}}{(\sigma_{inf}\sqrt{T})^k}$, uniformly in v.

This is a straightforward adaptation of Lemma 4.5.1 in Chapter 4, we omit the proof.

Step 3: Proof of Lemma 7.4.1

Starting from Equation (7.40) with $i+k-1$ instead of i, we write

$$Y^{\varepsilon}_{k,i+k-1,T} = \sum_{\substack{\alpha=(\alpha_1,\cdots,\alpha_i)\in\mathbb{N}^i \\ \sum_{j=1}^{i+k-1} j\alpha_j = i+k-1, \sum_{j=1}^{i+k-1}\alpha_j \leq k}} d_\alpha \frac{k!}{(\sum_{j=1}^{i+k-1}\alpha_j)!}(Y_T^{\varepsilon})^{k-\sum_{j=1}^{i+k-1}\alpha_j}\prod_{j=1}^{i+k-1}(Y^{\varepsilon}_{1,j,T})^{\alpha_j}.$$

Using Equation (7.47) one deduces, for $2 \leq j \leq i+k-1$, that

$$\|Y^{\varepsilon}_{1,j,T}\|_{k-1,p} = \|X^{\varepsilon}_{j,T}\|_{k-1,p} \leq_c M_1 M_0^{j-1}(\sqrt{T})^j.$$

This inequality is also available for $j=1$, since

$$\|Y^{\varepsilon}_{1,1,T}\|_{k-1,p} = \|\int_0^{\varepsilon} X^{\lambda}_{2,T} d\lambda\|_{k-1,p} \leq_c M_1 M_0 (\sqrt{T})^2 \leq_c M_1\sqrt{T}.$$

Additionally, we note that $Y^{\varepsilon}_{k,i+k-1,T} \in \mathbb{D}^{k-1,\infty}$. Furthermore, using the Hölder inequality for the spaces $\mathbb{D}^{k-1,\infty}$ (see Proposition 1.5.6 in [88]), we obtain

$$\|Y^{\varepsilon}_{k,i+k-1,T}\|_{k-1,p} \leq_c M_1 M_0^{i+k-2}(\sqrt{T})^{i+k-1}. \tag{7.51}$$

We omit the details of the above computations because they are very similar to those used for (7.42). Then, Lemma 7.5.3 ensures the existence of a random variable G_i in $\mathbf{L}_p$. Its $\mathbf{L}_p$ norm is estimated using Lemma 7.5.3 and Inequality (7.51):

$$\|G_i\|_p \leq_c \sum_{k=1}^{i-1} \frac{M_1 M_0^{i+k-2}(\sqrt{T})^{i+k-1}}{(\sigma_{inf}\sqrt{T})^{k-1}} \leq_c \frac{M_1 M_0^{2(i-1)-1}(\sqrt{T})^i}{\sigma_{inf}^{i-2}},$$

using $\frac{M_0}{\sigma_{inf}} \geq 1$. For S_m and $I_{m,v}$, we proceed analogously.

7.5.3 Statement of Theorem 7.4.2 (Vanilla options)

We first assume that h is a smooth function. We have

$$\begin{aligned}
\mathbb{E}[h(X_T)] =&\mathbb{E}[h(X_T^{BS})]\\
&+\sum_{k=1}^{m-1}\frac{1}{k!}\mathbb{E}[h^{(k)}(X_T^{BS})(\sum_{i=k+1}^{m}\frac{Y_{k,k+i-1,T}}{(k+i-1)!}+R_{k,k+m-1,T})]\\
&+\int_0^1\mathbb{E}[\frac{(Y_T^1)^m(1-v)^{m-1}}{(m-1)!}h^{(m)}(vX_T+(1-v)X_T^{BS})]dv\\
=&\mathbb{E}[h(X_T^{BS})]+\sum_{i=2}^{m}\frac{1}{k!}\sum_{k=1}^{i-1}\mathbb{E}[h^{(k)}(X_T^{BS})\frac{Y_{k,k+i-1,T}}{(k+i-1)!}]\\
&+\sum_{k=1}^{m-1}\frac{1}{k!}\mathbb{E}[h^{(k)}(X_T^{BS})R_{k,k+m-1,T}]\\
&+\int_0^1\mathbb{E}[\frac{(Y_T^1)^m(1-v)^{m-1}}{(m-1)!}h^{(m)}(vX_T+(1-v)X_T^{BS})]dv\\
=&\mathbb{E}[h(X_T^{BS})]+\sum_{i=2}^{m}\mathbb{E}[h^{(1)}(X_T^{BS})G_i]\\
&+\mathbb{E}[h^{(1)}(X_T^{BS})S_m]\\
&+\int_0^1\mathbb{E}[h^{(1)}(vX_T+(1-v)X_T^{BS})I_{m,v}](1-v)^{m-1}dv,
\end{aligned}$$

where we have used a Taylor expansion in the first identity, interchanged the summations in the second equality, and used the Lemma 7.4.1 in the last one. So yields the identity (7.22) for smooth payoff. Additionally, using estimates (7.20) and (7.21) from Lemma 7.4.1, it is straightforward to deduce the inequalities (7.23) and (7.24).
It remains to extend the result to vanilla options (instead of smooth function h). Since all the estimates depend only on $h^{(1)}$, it can be achieved by a standard density argument. We refer to Chapter 4 (see Subsection 4.5.2) for details.

7.6 Appendix

Here, we bring together the results (and their proofs) which allow us to derive the explicit terms in the formulas (7.9), (7.11), (7.14) and (7.15).
In the following, (u_t) (resp. (v_t)) is a square integrable and predictable (resp. deterministic) process and l is a smooth function with compact support.

7.6.1 Technical results related to explicit correction terms

The two first lemmas are proved in Chapter 4 (see Section 4.6).

Lemma 7.6.1. *For any continuous (or piecewise continuous) function f, any continuous semimartingale Z vanishing at t=0, one has:*

$$\int_0^T f_tZ_tdt=\int_0^T(\int_t^T f_sds)dZ_t.$$

Lemma 7.6.2. *One has:*

$$\mathbb{E}[(\int_0^T u_t dW_t) l(\int_0^T v_t dW_t)] = \mathbb{E}[(\int_0^T v_t u_t dt) l^{(1)}(\int_0^T v_t dW_t)].$$

In the case of deterministic u, it is equal to $\int_0^T v_t u_t dt\, \mathrm{Greek}_1^l(\int_0^T v_t dW_t)$.

7.6.2 Explicit correction in the case of Dupire model

In this case ($\mu \equiv -\frac{1}{2}\sigma^2$), the SDEs solved by the derivatives $X_{i,.}$ become:

$$\begin{aligned}
dX_{1,t} =& \sigma_t dW_t - \frac{\sigma_t^2}{2} dt, X_{1,0} = 0,\\
dX_{2,t} =& 2X_{1,t}(\sigma_t^{(1)} dW_t - \sigma_t \sigma_t^{(1)} dt), X_{2,0} = 0,\\
dX_{3,t} =& 3(X_{2,t}(\sigma_t^{(1)} dW_t - \sigma\sigma_t^{(1)} dt) + (X_{1,t})^2(\sigma_t^{(2)} dW_t - (\sigma_t\sigma_t^{(2)} + (\sigma_t^{(1)})^2)dt)), X_{3,0} = 0.
\end{aligned}$$

Lemma 7.6.3. *We have*

$$\mathbb{E}[(\int_0^T v_t X_{1,t} dt) l(\int_0^T \sigma_t dW_t)] = \omega(\sigma^2, v)_0^T \big(\mathbb{E}[l^{(1)}(\int_0^T \sigma_t dW_t)] - \frac{1}{2}\mathbb{E}[l(\int_0^T \sigma_t dW_t)]\big), \tag{7.52}$$

$$\begin{aligned}
\mathbb{E}[(\int_0^T v_t \frac{X_{2,t}}{2} dt) l(\int_0^T \sigma_t dW_t)] = & \ \omega(\sigma^2, \sigma\sigma^{(1)}, v)_0^T \big(\mathbb{E}[l^{(2)}(\int_0^T \sigma_t dW_t)]\\
& - \frac{3}{2}\mathbb{E}[l^{(1)}(\int_0^T \sigma_t dW_t)] + \frac{1}{2}\mathbb{E}[l(\int_0^T \sigma_t dW_t)]\big),
\end{aligned} \tag{7.53}$$

$$\begin{aligned}
\mathbb{E}[(\int_0^T v_t \frac{(X_{1,t})^2}{2} dt) l(\int_0^T \sigma_t dW_t)] = & \ \omega(\sigma^2, \sigma^2, v)_0^T \big(\mathbb{E}[l^{(2)}(\int_0^T \sigma_t dW_t)] - \mathbb{E}[l^{(1)}(\int_0^T \sigma_t dW_t)]\big)\\
& + (\frac{1}{4}\omega(\sigma^2, \sigma^2, v)_0^T + \frac{1}{2}\omega(\sigma^2, v)_0^T)\mathbb{E}[l(\int_0^T \sigma_t dW_t)],
\end{aligned} \tag{7.54}$$

$$\begin{aligned}
\mathbb{E}[(\int_0^T \sigma_t \sigma_t^{(1)} \frac{X_{2,t} X_{1,t}}{2} dt) l(\int_0^T \sigma_t dW_t)] = & \ (-\frac{3}{2}C_{6,T} - \frac{1}{2}C_{7,T} - \frac{1}{4}C_{8,T})\mathbb{E}[l(\int_0^T \sigma_t dW_t)]\\
& + (2C_{6,T} + \frac{5}{2}C_{7,T} + \frac{5}{4}C_{8,T})\mathbb{E}[l^{(1)}(\int_0^T \sigma_t dW_t)]\\
& + (-4C_{7,T} - 2C_{8,T})\mathbb{E}[l^{(2)}(\int_0^T \sigma_t dW_t)]\\
& + (2C_{7,T} + C_{8,T})\mathbb{E}[l^{(3)}(\int_0^T \sigma_t dW_t)],
\end{aligned} \tag{7.55}$$

where

$$C_{6,T} = \omega(\sigma^2, \sigma\sigma^{(1)}, \sigma\sigma^{(1)})_0^T, \qquad C_{7,T} = \omega(\sigma^2, \sigma^2, \sigma\sigma^{(1)}, \sigma\sigma^{(1)})_0^T,$$
$$C_{8,T} = \omega(\sigma^2, \sigma\sigma^{(1)}, \sigma^2, \sigma\sigma^{(1)})_0^T.$$

Proof. Applying first Lemma 7.6.1 to $f(t) = v_t$ and $Z_t = X_{1,t}$, we obtain:

$$\begin{aligned}\mathbb{E}[(\int_0^T v_t X_{1,t} dt) l(\int_0^T \sigma_t dW_t)] =& \mathbb{E}[(\int_0^T (\int_t^T v_s ds) dX_{1,t}) l(\int_0^T \sigma_t dW_t)] \\ =& \mathbb{E}[(\int_0^T (\int_t^T v_s ds)(\sigma_t dW_t - \frac{\sigma_t^2}{2} dt) l(\int_0^T \sigma_t dW_t)] \\ =& (\int_0^T \sigma_t^2 (\int_t^T v_s ds) dt) \mathbb{E}[l^{(1)}(\int_0^T \sigma_t dW_t)] \\ & - (\int_0^T \frac{\sigma_t^2}{2} (\int_t^T v_s ds) dt) \mathbb{E}[l(\int_0^T \sigma_t dW_t)],\end{aligned}$$

and we have used Lemma 7.6.2 for the last equality. This gives (7.52).

To establish the equalities (7.53), (7.54) and (7.55), we proceed analogously. We only detail the computations for (7.55). Using Lemma 7.6.1 ($f(t) = \sigma_t \sigma_t^{(1)}$, $Z_t = \frac{X_{2,t} X_{1,t}}{2}$) to justify the first following identity and Lemma 7.6.2 for the second one, we can write

$$\begin{aligned}& \mathbb{E}[(\int_0^T \sigma_t \sigma_t^{(1)} \frac{X_{2,t} X_{1,t}}{2} dt) l(\int_0^T \sigma_t dW_t)] \\ =& \mathbb{E}[(\int_0^T \omega(\sigma\sigma^{(1)})_t^T (X_{1,t}^2 (\sigma_t^{(1)} dW_t - \sigma_t \sigma_t^{(1)} dt) \\ & + \frac{X_{2,t}}{2} (\sigma_t dW_t - \frac{\sigma_t^2}{2} dt) + \sigma_t \sigma_t^{(1)} X_{1,t} dt)) l(\int_0^T \sigma_t dW_t)] \\ =& \mathbb{E}[(\int_0^T \omega(\sigma\sigma^{(1)})_t^T (-\sigma_t \sigma_t^{(1)} X_{1,t}^2 - \sigma_t^2 \frac{X_{2,t}}{4} + \sigma_t \sigma_t^{(1)} X_{1,t}) dt) l(\int_0^T \sigma_t dW_t)] \\ & + \mathbb{E}[(\int_0^T \omega(\sigma\sigma^{(1)})_t^T (\sigma_t \sigma_t^{(1)} X_{1,t}^2 + \sigma_t^2 \frac{X_{2,t}}{2}) dt) l^{(1)}(\int_0^T \sigma_t dW_t)].\end{aligned}$$

Then, we obtain the announced identity by an application of the three first identities (7.52), (7.53) and (7.54). □ □

7.6.3 Proof of Theorem 7.2.2

Proof. Using Theorem 7.4.2 and Lemma 7.4.1, the price is approximated at the third order by

$$E[h(X_T^{BS})] + E[h^{(1)}(X_T^{BS}) \frac{X_{2,T}}{2}] + \mathbb{E}[h^{(1)}(X_T^{BS}) \frac{X_{3,T}}{3!}] + \mathbb{E}[h^{(2)}(X_T^{BS}) \frac{(\frac{X_{2,T}}{2})^2}{2}].$$

We compute each correction term separately.

Step 1: term with $X_{2,T}$. Owing to Lemma 7.6.2, we have

$$\begin{aligned}\mathbb{E}[h^{(1)}(X_T^{BS}) \frac{X_{2,T}}{2}] =& \mathbb{E}[h^{(1)}(X_T^{BS})(\int_0^T X_{1,t}(\sigma_t^{(1)} dW_t - \sigma_t \sigma_t^{(1)} dt))] \\ =& \mathbb{E}[h^{(2)}(X_T^{BS})(\int_0^T \sigma_t \sigma_t^{(1)} X_{1,t} dt)] \\ & - \mathbb{E}[h^{(1)}(X_T^{BS})(\int_0^T \sigma_t \sigma_t^{(1)} X_{1,t} dt)].\end{aligned}$$

Apply Lemma 7.6.3 (equality (7.52)) to obtain

$$\mathbb{E}[h^{(1)}(X_T^{BS})\frac{X_{2,T}}{2}] = C_{1,T}(\mathbb{E}[h^{(3)}(X_T^{BS})] - \frac{3}{2}\mathbb{E}[h^{(2)}(X_T^{BS})] + \frac{1}{2}\mathbb{E}[h^{(1)}(X_T^{BS})]),$$

where $C_{1,T} = \omega(\sigma^2, \sigma\sigma^{(1)})_0^T$.

Step 2: term with $X_{3,T}$. From Lemma 7.6.1 and 7.6.2, we obtain

$$\begin{aligned}\mathbb{E}[h^{(1)}(X_T^{BS})\frac{X_{3,T}}{3!}] =& \mathbb{E}[h^{(2)}(X_T^{BS})\frac{\int_0^T \sigma_t\sigma_t^{(1)}X_{2,t}dt}{2}] \\ &- \mathbb{E}[h^{(1)}(X_T^{BS})\frac{\int_0^T \sigma_t\sigma_t^{(1)}X_{2,t}dt}{2}] \\ &+ \mathbb{E}[h^{(2)}(X_T^{BS})\frac{\int_0^T \sigma_t\sigma_t^{(2)}(X_{1,t})^2dt}{2}] \\ &- \mathbb{E}[h^{(1)}(X_T^{BS})\frac{\int_0^T (\sigma_t\sigma_t^{(2)} + (\sigma_t^{(1)})^2)(X_{1,t})^2dt}{2}].\end{aligned}$$

An application of Lemma 7.6.3 (equalities (7.53) and (7.54)) gives:

$$\begin{aligned}\mathbb{E}[h^{(1)}(X_T^{BS})\frac{X_{3,T}}{3!}] =& (-\frac{1}{2}C_{2,T} - \frac{1}{2}C_{3,T} - \frac{1}{4}C_{4,T} - \frac{1}{4}C_{5,T} - \frac{1}{2}C_{6,T})\mathbb{E}[h^{(1)}(X_T^{BS})] \\ &+ (\frac{1}{2}C_{3,T} + C_{4,T} + \frac{5}{4}C_{5,T} + 2C_{6,T})\mathbb{E}[h^{(2)}(X_T^{BS})] \\ &+ (-C_{4,T} - 2C_{5,T} - \frac{5}{2}C_{6,T})\mathbb{E}[h^{(3)}(X_T^{BS})] \\ &+ (C_{5,T} + C_{6,T})\mathbb{E}[h^{(4)}(X_T^{BS})],\end{aligned}$$

where

$$\begin{aligned}&C_{2,T} = \omega(\sigma^2, (\sigma^{(1)})^2)_0^T, \quad C_{3,T} = \omega(\sigma^2, \sigma\sigma^{(2)})_0^T, \quad C_{4,T} = \omega(\sigma^2, \sigma^2, (\sigma^{(1)})^2)_0^T, \\ &C_{5,T} = \omega(\sigma^2, \sigma^2, \sigma\sigma^{(2)})_0^T, \quad C_{6,T} = \omega(\sigma^2, \sigma\sigma^{(1)}, \sigma\sigma^{(1)})_0^T.\end{aligned}$$

Step 3: term with $(X_{2,T})^2$. Similarly, we have

$$\begin{aligned}\mathbb{E}[h^{(2)}(X_T^{BS})\frac{(\frac{X_{2,T}}{2})^2}{2}] =& \mathbb{E}[h^{(2)}(X_T^{BS})\int_0^T (\sigma_t^{(1)}\frac{X_{1,t}X_{2,t}}{2}dW_t - \sigma_t\sigma_t^{(1)}\frac{X_{1,t}X_{2,t}}{2}dt + (\sigma_t^{(1)})^2\frac{X_{1,t}^2}{2}dt)] \\ =& \mathbb{E}[h^{(3)}(X_T^{BS})\int_0^T (\sigma_t\sigma_t^{(1)}\frac{X_{1,t}X_{2,t}}{2}dt] \\ &- \mathbb{E}[h^{(2)}(X_T^{BS})\int_0^T (\sigma_t\sigma_t^{(1)}\frac{X_{1,t}X_{2,t}}{2}dt] \\ &+ \mathbb{E}[h^{(2)}(X_T^{BS})\int_0^T (\sigma_t^{(1)})^2\frac{X_{1,t}^2}{2}dt].\end{aligned}$$

Using Lemma 7.6.3 (third and fourth equalities), it follows

$$\begin{aligned}\mathbb{E}[h^{(2)}(X_T^{BS})\frac{(\frac{X_{2,T}}{2})^2}{2}] =&(\frac{1}{2}C_{2,T}+\frac{1}{4}C_{4,T}+\frac{3}{2}C_{6,T}+\frac{1}{2}C_{7,T}+\frac{1}{4}C_{8,T})\mathbb{E}[h^{(2)}(X_T^{BS})]\\ &+(-C_{4,T}-\frac{7}{2}C_{6,T}-3C_{7,T}-\frac{3}{2}C_{8,T})\mathbb{E}[h^{(3)}(X_T^{BS})]\\ &+(C_{4,T}+2C_{6,T}+\frac{13}{2}C_{7,T}+\frac{13}{4}C_{8,T})\mathbb{E}[h^{(4)}(X_T^{BS})]\\ &+(-6C_{7,T}-3C_{8,T})\mathbb{E}[h^{(5)}(X_T^{BS})]\\ &+(2C_{7,T}+C_{8,T})\mathbb{E}[h^{(6)}(X_T^{BS})],\end{aligned}$$

where

$$\begin{aligned}C_{1,T}&=\omega(\sigma^2,\sigma\sigma^{(1)})_0^T, & C_{2,T}&=\omega(\sigma^2,(\sigma^{(1)})^2)_0^T,\\ C_{4,T}&=\omega(\sigma^2,\sigma^2,(\sigma^{(1)})^2)_0^T, & C_{6,T}&=\omega(\sigma^2,\sigma\sigma^{(1)},\sigma\sigma^{(1)})_0^T,\\ C_{7,T}&=\omega(\sigma^2,\sigma^2,\sigma\sigma^{(1)},\sigma\sigma^{(1)})_0^T, & C_{8,T}&=\omega(\sigma^2,\sigma\sigma^{(1)},\sigma^2,\sigma\sigma^{(1)})_0^T.\end{aligned}$$

Final step. To get the announced formula, we bring together all the previous contributions and use that $\mathbb{E}(h^{(i)}(X_T^{BS}))=\text{Greek}_i^h(X_T^{BS})$. □

7.6.4 Faà di Bruno's formula [37]

Lemma 7.6.4. *If g and f are functions that are sufficiently differentiable, then*

$$(g(f(\varepsilon)))^{(n)}=\sum_{\substack{k=(k_1,\cdots,k_n)\in\mathbb{N}^n\\ \sum_{j=1}^n jk_j=n}} d_k g^{(\sum_{j=1}^n k_j)}(f(\varepsilon))\prod_{j=1}^n(f^{(j)}(\varepsilon))^{k_j},$$

where d_k are integer numbers depending only on k. Notice that when $k_n=1$ one has $k_1=\cdots=k_{n-1}=0$ and $d_k=1$.

Chapter 8

Appendix

8.1 Smile behaviors for CEV model

The aim of this section is to study the smile's behavior induced by the CEV model. In all this section, we take $x_0 = 1$ ($S_0 = 1$), $r = q = 0$. The Black Scholes implied volatilities are computed in the following cases (using one of the previous numerical methods):

- *Constant parameters*. In this case, we consider the constant CEV model with level $\nu = 25\%$ and shift $\beta = 30\%$. We plot the implied Black Scholes volatilities, in the figure 8.1, for different maturities; for each maturity, the implied Black Scholes volatilities are plotted as a function of the log moneyness $\ln(\frac{K}{S})$. From the figure 8.1, we observe that the smile of implied Black volatilities doesn't change through the time. Moreover, in Figure 8.2, we plot both local volatility at the money and implied Black Scholes volatility for short maturity 3M. Then, we remark that the slope of the local volatility is approximatively twice the slope of the implied Black Scholes, as it can be derived from Equality (6.3) (see also page 88 in [50]).

- *Time varying parameter ν*. In this case, we consider a CEV model with constant shift $\beta = 30\%$ and piecewise constant level ν which is equal to $25\% - i \times 0.11\%$ on each interval of the form $[\frac{i}{20}, \frac{i+1}{20}[$. We remark from the figure 8.3 that the smile is shifted from maturity to maturity but keeps the same shape for all the maturities. In other words, the smile is bumped from maturity to maturity.

- *Time varying parameter β*. In this case, we consider a CEV model with constant level $\nu = 25\%$ and piecewise constant shift β which is equal to $25\% - i \times 0.11\%$ on each interval of the form $[\frac{i}{20}, \frac{i+1}{20}[$. We remark from the figure 8.4 that the smile's shape change from maturity to maturity because of the changes of the shift β.

- *Time varying parameters ν and β*. In this case, we consider a CEV model with piecewise constant parameters ν and β which are equal respectively to $25\% - i \times 0.11\%$ and $30\% + i \times 0.35\%$ on each interval of the form $[\frac{i}{20}, \frac{i+1}{20}[$. We observe from the figure 8.5, that the smile is bumped and that its shape changes from maturity to maturity. We plot also the surface of implied Black Scholes volatilities as a function of the maturities and strikes[1].

[1] In this example, the smile is emphasized for small maturities and becomes flat for large ones. This example coincides with real market data.

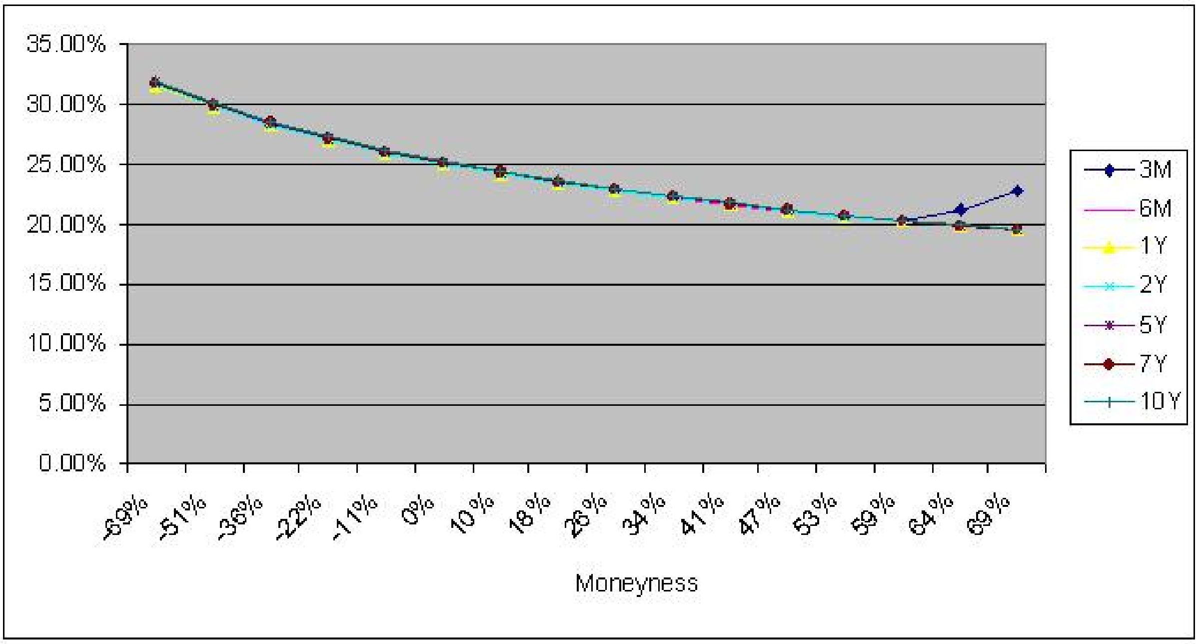


Figure 8.1: Implied Black Scholes volatilities for the constant CEV model.

To sum up, we deduce that the smile changes are due to the change of the shift β through the time and the smile bumps are due to the change of the level ν through the time. As a consequence, we remark that if we want to reproduce real market smiles using local volatilities, we need to use time inhomogeneous parameters in order to manage smile changes.

8.2 Large strikes

In this section, we detail results concerning the accuracy of the formulas (7.9), (7.11) for large strikes. This is tested for the case of constant CEV model in the tables 8.1 for $\beta = 0.8$ and 8.2 for $\beta = 0.2$.
For instance, the cell of the table 8.1 at the maturity $4Y$ and for the relative strike 40% contains three informations:

- 27.02% is the implied BS volatility using the second order lognormal approximation formula (7.9).
- 27.35% is the implied BS volatility using the third order lognormal approximation formula (7.11).
- 27.37% is the implied BS volatility using the CEV closed formula (6.2).

Therefore, we remark that our second order formula is still accurate for small skews ($\beta = 0.8$) and less accurate for large skews ($\beta = 0.2$). Moreover, our third order approximation formula is still very accurate for both small and large skews.

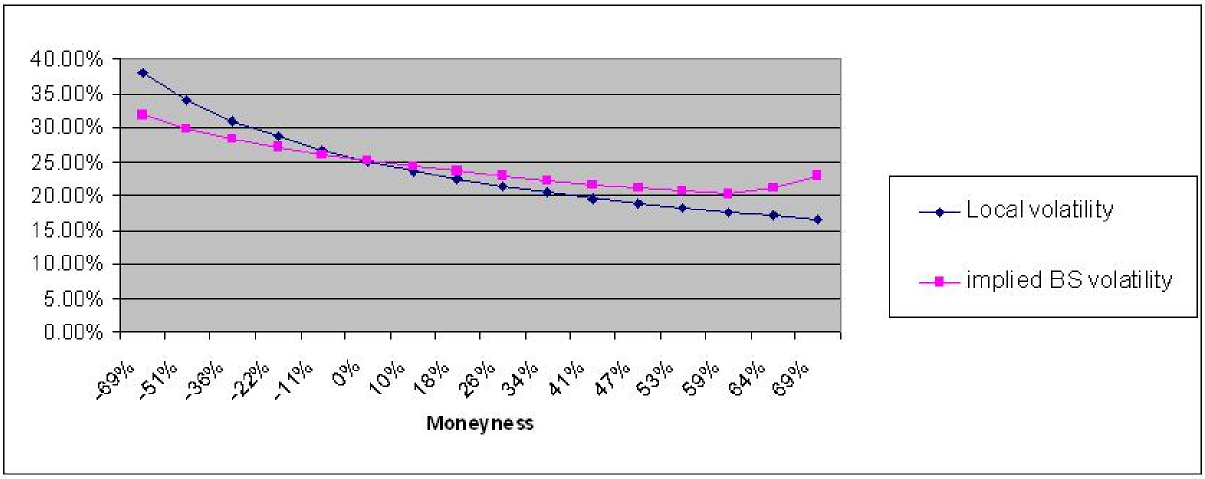


Figure 8.2: Local and implied Black Scholes volatilities for the constant CEV model.

8.3 Arbitrage interval

In this section, we aim at exhibiting the domain where the call price obtained using our approximation formulas (7.9) belongs to the no-arbitrage interval $](e^{x_0}-K)^+, e^{x_0}[$ (when $r=q=0$). Indeed, for some numerical methods, the errors are sometimes so large that the resulting numerical price is out of the range of the no-arbitrage bounds. This is potentially dramatic because one can make an arbitrage on model prices.
In the following tables, we write:

- T (abbreviation of True) if the call price obtained using our second order approximation (7.9) belongs to the no-arbitrage interval.
- F (abbreviation of False) if the call price obtained using our second order approximation (7.9) doesn't belong to the no-arbitrage interval.

We give the domain arbitrage in a progressive way for

- the second order approximation formula using a lognormal proxy (7.9). This is reported in tables 8.3 and 8.4.
- the third order approximation formula using a lognormal proxy (7.11). This is reported in tables 8.5 and 8.6.
- the third order approximation formula using a normal proxy (7.14). This is reported in tables 8.7 and 8.8.

We observe that our formula covers a very large domain of strikes which includes real market priced options. Moreover, we remark that the domain of no arbitrage widens with the order of the approximation formula. This shows that our formula is safe and trustful for large strikes and various maturities.

Table 8.1: Implied Black-Scholes volatilities of the second, third order approximation formula ((7.9), (7.11)) and the CEV closed formula (6.2), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\nu = 25\%$, $\beta = 80\%$.

T/K	40 %	70 %	100 %	150 %	180 %
4Y	27.02%	27.02%	27.02%	27.02%	27.02%
	27.35%	27.35%	27.35%	27.35%	27.35%
	27.37%	27.37%	27.37%	27.37%	27.37%
5Y	27.07%	27.07%	27.07%	27.07%	27.07%
	27.36%	27.36%	27.36%	27.36%	27.36%
	27.38%	27.38%	27.38%	27.38%	27.38%
6Y	27.11%	27.11%	27.11%	27.11%	27.11%
	27.37%	27.37%	27.37%	27.37%	27.37%
	27.38%	27.38%	27.38%	27.38%	27.38%
7Y	27.13%	27.13%	27.13%	27.13%	27.13%
	27.37%	27.37%	27.37%	27.37%	27.37%
	27.38%	27.38%	27.38%	27.38%	27.38%
8Y	27.15%	27.15%	27.15%	27.15%	27.15%
	27.38%	27.38%	27.38%	27.38%	27.38%
	27.38%	27.38%	27.38%	27.38%	27.38%
9Y	27.17%	27.17%	27.17%	27.17%	27.17%
	27.38%	27.38%	27.38%	27.38%	27.38%
	27.39%	27.39%	27.39%	27.39%	27.39%
10Y	27.18%	27.18%	27.18%	27.18%	27.18%
	27.38%	27.38%	27.38%	27.38%	27.38%
	27.39%	27.39%	27.39%	27.39%	27.39%

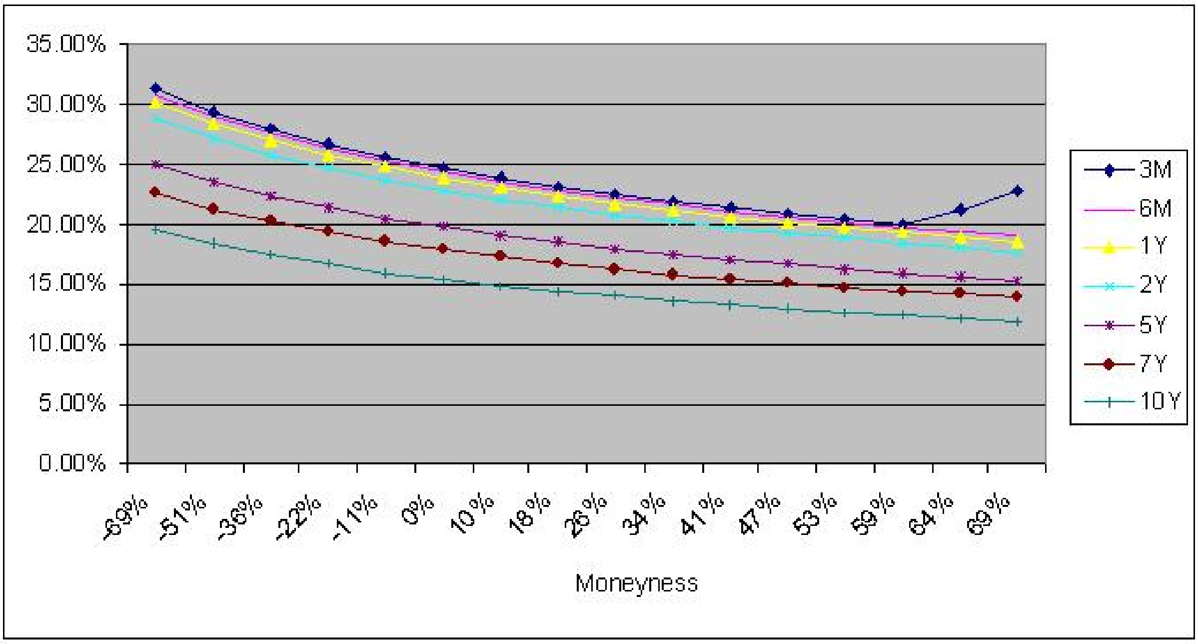


Figure 8.3: Implied Black Scholes volatilities for the CEV model with time varying parameter ν.

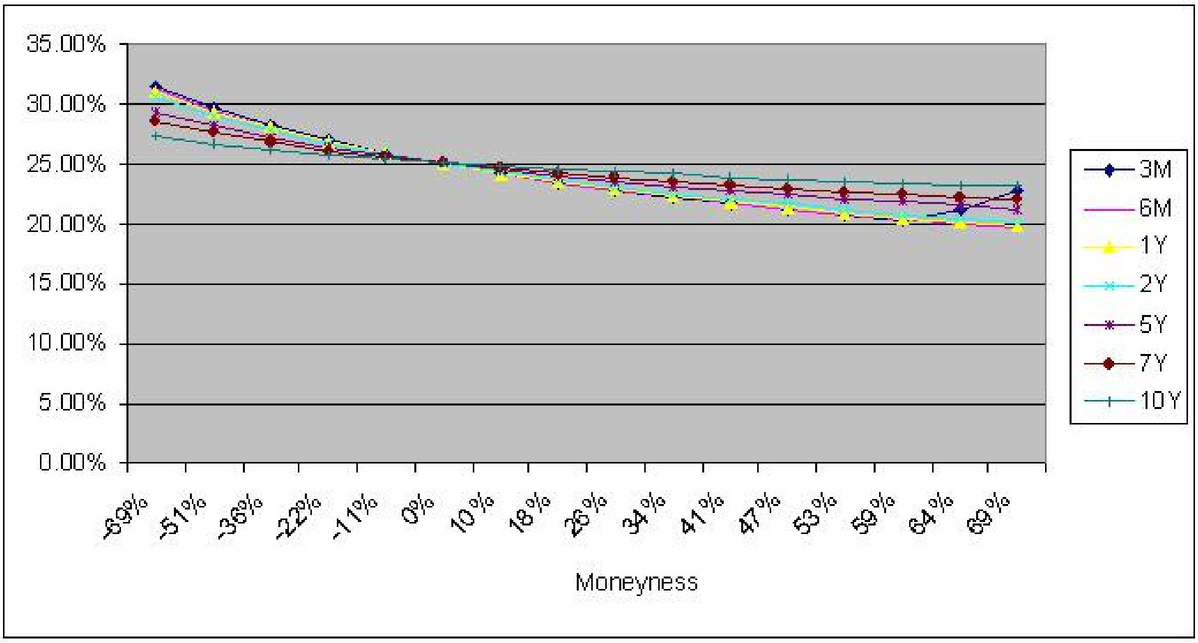
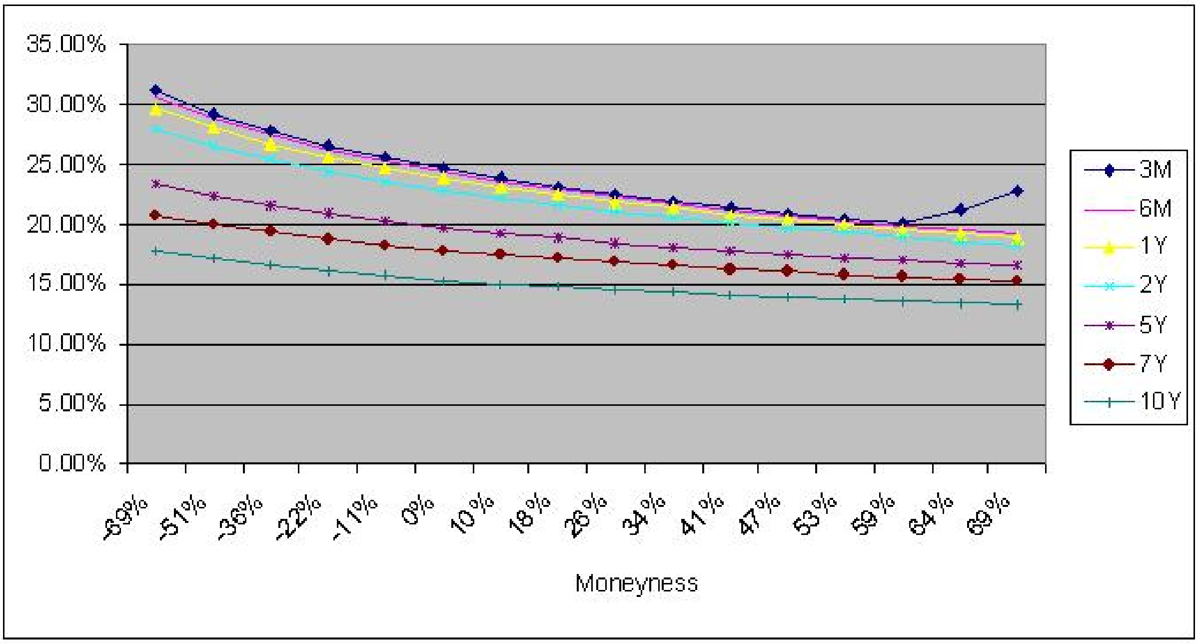


Figure 8.4: Implied Black Scholes volatilities for the CEV model with time varying parameter β.



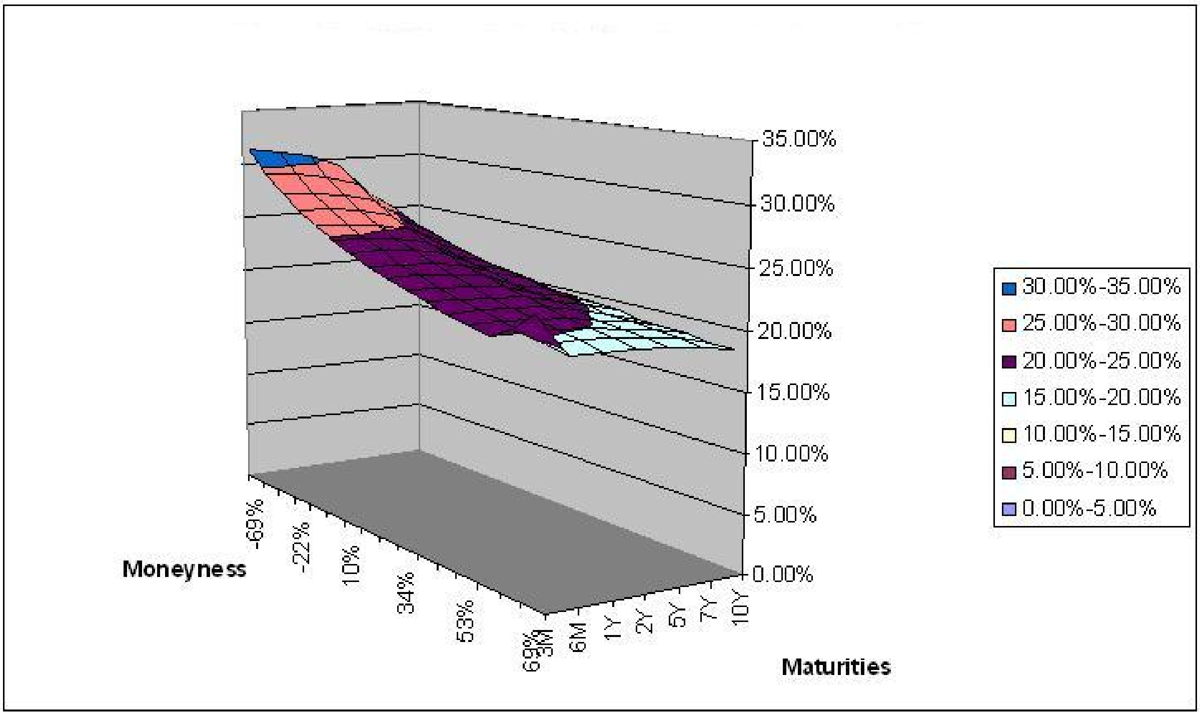


Figure 8.5: Implied Black Scholes volatilities for the CEV model with time varying parameters ν and β.

Table 8.2: Implied Black-Scholes volatilities of the second , third order approximation formula ((7.9), (7.11)) and the CEV closed formula (6.2), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\nu = 25\%$, $\beta = 20\%$.

T/K	40 %	70 %	100 %	150 %	180 %
4Y	31.45%	28.47%	25.00%	20.68%	17.20%
	34.84%	28.89%	25.16%	21.33%	19.86%
	35.71%	28.98%	25.16%	21.26%	19.66%
5Y	31.87%	28.50%	25.00%	20.75%	17.80%
	35.11%	28.93%	25.19%	21.36%	19.85%
	35.80%	29.04%	25.20%	21.29%	19.68%
6Y	32.19%	28.52%	25.00%	20.79%	18.12%
	35.30%	28.96%	25.23%	21.40%	19.86%
	35.86%	29.10%	25.24%	21.31%	19.70%
7Y	32.45%	28.53%	25.00%	20.83%	18.33%
	35.43%	29.00%	25.26%	21.43%	19.87%
	35.89%	29.15%	25.28%	21.34%	19.72%
8Y	32.66%	28.54%	25.00%	20.85%	18.48%
	35.54%	29.03%	25.29%	21.46%	19.89%
	35.89%	29.19%	25.32%	21.36%	19.74%
9Y	32.84%	28.55%	25.00%	20.87%	18.58%
	35.62%	29.07%	25.32%	21.49%	19.92%
	35.87%	29.22%	25.35%	21.38%	19.76%
10Y	32.99%	28.56%	25.00%	20.89%	18.67%
	35.70%	29.10%	25.35%	21.51%	19.94%
	35.83%	29.24%	25.38%	21.41%	19.78%

Table 8.3: No-arbitrage domain for the second order formula (7.9) expressed in booleans (T for True and F for False) as a function of maturities and strikes. Parameters: $\nu = 25\%$, $\beta = 80\%$.

T/K	0.1	0.30	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10
3M	T	T	T	T	T	T	T	T	F	F	F	F	F	F	F	F
6M	T	T	T	T	T	T	T	T	T	F	F	F	F	F	T	F
9M	T	T	T	T	T	T	T	T	T	T	F	F	F	F	F	F
1Y	T	T	T	T	T	T	T	T	T	T	T	F	F	F	F	F
2Y	T	T	T	T	T	T	T	T	T	T	T	T	T	F	F	F
3Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	F
5Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Table 8.4: No-arbitrage domain for the second order formula (7.9) expressed in booleans (T for True and F for False) as a function of maturities and strikes. Parameters: $\nu = 25\%$, $\beta = 20\%$.

T/K	0.1	0.30	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10
3M	T	T	T	T	T	T	T	F	F	F	F	F	F	F	F	F
6M	T	T	T	T	T	T	T	F	F	F	F	F	F	F	T	F
9M	T	T	T	T	T	T	T	T	F	F	F	F	F	F	F	F
1Y	T	T	T	T	T	T	T	T	F	F	F	F	F	F	F	F
2Y	T	T	T	T	T	T	T	T	T	F	F	F	F	F	F	F
3Y	T	T	T	T	T	T	T	T	T	F	F	F	F	F	F	F
5Y	T	T	T	T	T	T	T	T	T	T	F	F	F	F	F	F

Table 8.5: No-arbitrage domain for the third order approximation formula (7.11) expressed in booleans (T for True and F for False) as a function of maturities and strikes. Parameters: $\nu = 25\%$, $\beta = 80\%$.

T/K	0.1	0.30	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10
3M	F	F	T	T	T	T	T	T	T	T	T	T	T	F	F	F
6M	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
9M	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
1Y	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
2Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
3Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
5Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Table 8.6: No-arbitrage domain for the third approximation order formula (7.11) expressed in booleans (T for True and F for False) as a function of maturities and strikes. Parameters: $\nu = 25\%$, $\beta = 20\%$.

T/K	0.1	0.30	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10
3M	F	F	T	T	T	T	T	T	T	T	T	T	T	T	F	F
6M	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
9M	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
1Y	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
2Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
3Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
5Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Table 8.7: No-arbitrage domain for the third order approximation formula (7.14) expressed in booleans (T for True and F for False) as a function of maturities and strikes. Parameters: $\nu = 25\%$, $\beta = 80\%$.

T/K	0.1	0.30	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10
3M	T	T	T	T	T	T	T	T	T	T	F	F	F	F	F	F
6M	T	T	T	T	T	T	T	T	T	T	T	T	F	F	F	F
9M	T	T	T	T	T	T	T	T	T	T	T	T	T	T	F	F
1Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	F
2Y	T	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T
3Y	F	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T
5Y	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Table 8.8: No-arbitrage domain for the third approximation order formula (7.14) expressed in booleans (T for True and F for False) as a function of maturities and strikes. Parameters: $\nu = 25\%$, $\beta = 20\%$.

T/K	0.1	0.30	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30	2.50	2.70	2.90	3.10
3M	T	T	T	T	T	T	T	T	T	T	F	F	F	F	F	F
6M	T	T	T	T	T	T	T	T	T	T	T	T	F	F	F	F
9M	T	T	T	T	T	T	T	T	T	T	T	T	T	T	F	F
1Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	F
2Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
3Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
5Y	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T

Part III

Stochastic volatility models

Chapter 9

Introduction

Stochastic volatility models assumes the volatility of the spot itself to be stochastic, and then is governed by a stochastic differential Equation. These kind of models manage better the smile than local volatility models but they have the drawback of market incompleteness since they use two factors of uncertainty in order to hedge one asset. We cite among these models: The SABR model (see [61]) and the Heston one (see [63]). In this part, we focus our work on the Heston model.

The Heston model is an extension of the Black Scholes model for the underlying (S_t) with stochastic volatility:

$$\mathrm{d}X_t = \sqrt{v_t} dW_t - \frac{v_t}{2}\mathrm{d}t,\ X_0 = x_0, \tag{9.1}$$

$$\mathrm{d}v_t = \kappa(\theta_t - v_t)\mathrm{d}t + \xi_t\sqrt{v_t}\mathrm{d}B_t,\ v_0 > 0, \tag{9.2}$$

$$\mathrm{d}\langle W,B\rangle_t = \rho_t \mathrm{d}t,$$

where

- X_t is the logarithm of the forward $e^{(q-r)t}S_t$, r and q are respectively the risk free rate and the dividend yield,
- v_0 is the initial square of volatility,
- κ is the mean reversion parameter,
- θ is the long-term level,
- ξ is the volatility of volatility (vol of vol),
- ρ is the correlation.

We assume that the above dynamics is directly given under the pricing measure.

9.1 Heston model

The computation of the call-put price in the Heston model can be done using Fourier inversion since the characteristic function of the logarithm of the underlying is explicit when we have:

- either constant parameters θ, ξ and ρ,
- or piecewise constant parameters θ, ξ and ρ.

9.1.1 Lewis' formula

The call price in Heston's model can be written using the Lewis' formula ([79]):

$$\begin{aligned} Call_{Heston}(t,S_t,v_t;T,K) =& S_t e^{-q(T-t)} \\ & - \frac{Ke^{-(T-t)r}}{2\pi}\int_{\frac{i}{2}-\infty}^{\frac{i}{2}+\infty} e^{-izX}\phi_T(-z)\frac{dz}{z^2-iz} \end{aligned}$$

where $X = \log\left(\frac{S_t e^{-(T-t)q}}{Ke^{-(T-t)r}}\right)$ and $\phi_T(z) = \mathbb{E}(e^{z(X_T-X_t)}|\mathscr{F}_t)$.

The characteristic function $\phi_T(z) = \mathbb{E}(e^{z(X_T-X_t)}|\mathscr{F}_t)$ is explicit when the parameters θ, ξ and ρ are constant.

Proof references:

- *PDE approach: The Heston model belongs to affine models and has an explicit characteristic function, that is solution of Riccatti equations (see [63]). This is derived using PDE arguments.*
- *Probabilist approach: The CIR process v is a time-space changed Bessel process (see [56]).*

Moreover, the characteristic function $\phi_T(z) = \mathbb{E}(e^{z(X_T - X_t)}|\mathscr{F}_t)$ is computed recursively when the parameters θ, ξ and ρ are piecewise constant.
Proof references:

- *PDE approach: The characteristic function $\phi_T(z)$ can be computed recursively using nested Riccatti equations with constant coefficients (see [86]).*
- *Probabilist approach: We use the Markov property in order to compute recursively the characteristic function (see [41]).*

9.1.2 Heston's formula

Heston's formula for call option is an extension of Black Scholes formula for call option with spot S, strike K and maturity T:

$$Call_{Heston}(t, S_t, v_t, T, K) = S_t e^{-q(T-t)} P_1 - K e^{-r(T-t)} P_2, \tag{9.3}$$

Using a change of numéraire, the probabilities P_j can be considered as the conditional probability of $x_j(T)$:

$$P_j(\log S_t, v_t, t, T; ln(K)) = Prob[x_j(T) \geq \log K | x_j(t) = \log S_t, v_j(t) = v_t]$$

where x_j follows:

$$\begin{aligned} dx_j(t) &= ((r-q) + u_j v_j(t))dt + \sqrt{v_j(t)}dW_t^1 \\ dv_j(t) &= (a - b_j v_j(t))dt + \xi\sqrt{v_j(t)}dW_t^2, \\ d\langle W^1, W^2\rangle_t &= \rho dt, \end{aligned}$$

and $u_1 = \frac{1}{2}$, $u_2 = -\frac{1}{2}$, $a = \kappa\theta$, $b_1 = \kappa - \rho\xi$, $b_2 = \kappa$. We introduce the characteristic function $f_j(x, V, t, T; \phi)$ related to the complementary cumulative distribution $P_j(x, V, t, T; ln(K))$. Hence, using the Fourier's inversion formula:

$$P_j(x, V, t, T; ln(K)) = 1/2 + \frac{1}{\pi}\int_0^\infty Re[\frac{e^{-i\phi ln(K)} f_j(x, V, t, T; \phi)}{i\phi}]d\phi \tag{9.4}$$

where $f_j(x, V, t; T, \phi)$ satisfies:

$$\begin{cases} \frac{1}{2}V\frac{\partial^2 f_j}{\partial x^2} + \rho\xi V\frac{\partial^2 f_j}{\partial x \partial V} + \frac{1}{2}\xi^2 V\frac{\partial^2 f_j}{\partial V^2} + (r - q + u_j V)\frac{\partial f_j}{\partial x} + (a_j - b_j V)\frac{\partial f_j}{\partial V} + \frac{\partial f_j}{\partial t} = 0 & \text{if } 0 \leq t < T \\ f_j(x, V, T; \log K) = e^{i\phi x} & \text{if } t = T \end{cases}$$

For the proof of Fourier inversion, we refer to [58] and [53] and for financial application to [13] and [105].

Using a PDE verification argument ([79], [63], [77]) or the additivity property of the Bessel process (see [98]), the solutions $f_j(x, V, t = 0, T; \phi)$ have the form:

$$f_j(x, V, t, T; \phi) = e^{C(T-t;\phi) + D(T-t;\phi)V + i\phi x}, \tag{9.5}$$

where $C_j(\tau;\phi)$ and $D_j(\tau;\phi)$, with $\tau = T - t$, satisfy the following ordinary differential equations:

$$\begin{cases} \frac{dC_j(\tau;\phi)}{d\tau} - aD_j(\tau,\phi) - (r-q)\phi i = 0, \\ \frac{dD_j(\tau;\phi)}{d\tau} - \frac{\xi^2 D_j^2(\tau;\phi)}{2} + (b_j - \rho\xi\phi i)D_j(\tau,\phi) - u_j\phi i + \frac{1}{2}\phi^2 = 0 \end{cases}$$

with $C_j(0;\phi) = 0$ and $D_j(0;\phi) = 0$. The solutions for $C_j(\tau;\phi)$ and $D_j(\tau;\phi)$ are:

$$\begin{aligned} C_j(\tau;\phi) &= (r-q)\phi i\tau + \frac{a}{\xi^2}\{(b_j - \rho\xi\phi i + d_j)\tau - 2\log[\frac{1-g_j e^{d_j\tau}}{1-g_j}]\}, \\ D_j(\tau;\phi) &= \frac{b_j - \rho\xi\phi i + d_j}{\xi^2}[\frac{1-e^{d_j\tau}}{1-g_j e^{d_j\tau}}], \\ g_j &= \frac{b_j - \rho\xi\phi i + d_j}{b_j - \rho\xi\phi i - d_j}, \\ d_j &= -\sqrt{(\rho\xi\phi i - b_j)^2 - \xi^2(2u_j\phi i - \phi^2)}. \end{aligned}$$

Here we notice that the square root complex $d_j = -\sqrt{(\rho\xi\phi i - b_j)^2 - \xi^2(2u_j\phi i - \phi^2)}$ is used to avoid discontinuity (see [2]) instead of the original equation $d_j = \sqrt{(\rho\xi\phi i - b_j)^2 - \xi^2(2u_j\phi i - \phi^2)}$ derived in [63]. Moreover, we refer to [68] to handle the discontinuity of $\log[\frac{1-g_j e^{d_j\tau}}{1-g_j}]$. Heston's formula can be improved using a control variate method which is related to the closest Black Scholes price (Heston's price with same parameters without volatility of volatility). Notice that there are other methods of Fourier inversion formula and control variate methods in the excellent book of Lewis [79].

When Heston's parameters are piecewise constant, the characteristic function $f_j(x,V,t,T;\phi)$ still have closed formula. It is calculated recursively using PDE methods and nested Riccatti Equations (see [86]) or a Markov argument for affine models (see [41]).

9.2 Review of Analytical approximations

In this section, we give the approaches existing in the literature.

Ergodic approach. Fouque et al in [44] consider a stochastic volatility model where the volatility is a functional of Y_t^ε which is an Ornstein Uhlenbeck process (we denote it OU) with small correlation time ε:

$$\begin{aligned} \frac{dS_t^\varepsilon}{S_t^\varepsilon} &= rdt + f(Y_t^\varepsilon)dW_t, \\ dY_t^\varepsilon &= \frac{1}{\varepsilon}(m - Y_t^\varepsilon)dt + \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}}dB_t, \\ d\langle W,B\rangle_t &= \rho dt. \end{aligned}$$

They write the infinitesimal generator $\mathscr{L}^\varepsilon$ as the summation of three terms

$$\mathscr{L}^\varepsilon = \frac{1}{\varepsilon}\mathscr{L}^{(0)} + \frac{1}{\sqrt{\varepsilon}}\mathscr{L}^{(1)} + \mathscr{L}^{(2)}, \tag{9.6}$$

where

- $\mathscr{L}^{(0)} = \nu^2 \frac{\partial^2}{\partial y^2} + (m-y)\frac{\partial}{\partial y}$ is the infinitesimal generator of the OU process Y defined by

$$dY_t = (m - Y_t)dt + \nu\sqrt{2}dB_t, \tag{9.7}$$

- $\mathscr{L}^{(1)} = \sqrt{2}\rho x f(y)\frac{\partial^2}{\partial x \partial y}$ contains the mixed derivatives due to the correlation term,
- $\mathscr{L}^{(2)} = \frac{\partial}{\partial t} + \frac{1}{2}f(y)^2 x^2 \frac{\partial^2}{\partial x^2} + r(x\frac{\partial}{\partial x} - .)$ is the Black-Scholes operator with volatility $f(y)$.

Fouque et al suppose that the parameter ε is small which makes the problem (9.6) to be a singular perturbation problem. They expand the call price with starting time t, maturity T, spot S and strike K, in power of $\sqrt{\varepsilon}$:

$$\text{Call}^{\varepsilon} = \text{Call}_{BS}^{(0)} + \sqrt{\varepsilon}\text{Correction}^{(1)} + \cdots$$

where the leading term is the Black Scholes price $\text{Call}_{BS}^{(0)} = Call_{BS}(t,S,K,T,\bar{\sigma})$ and the variance $\bar{\sigma}^2$ is the averaging of the function f with respect to the invariant distribution $\mathscr{N}(m,\nu^2)$ of the OU (Y) defined in Equation (9.7):

$$\bar{\sigma}^2 = \frac{1}{\nu\sqrt{2\pi}}\int_{\mathbb{R}} f^2(y) e^{\frac{-(m-y)^2}{2\nu^2}} dy \equiv \langle f^2 \rangle, \tag{9.8}$$

and the correction term is a combination of Greeks of the leading term $\text{Call}_{BS}^{(0)}$:

$$\sqrt{\varepsilon}\text{Correction}^{(1)} = -(T-t)(V_2 S^2 \frac{\partial^2 Call_{BS}(t,S,K,T,\bar{\sigma})}{\partial S^2} + V_3 S^3 \frac{\partial^3 Call_{BS}(t,S,K,T,\bar{\sigma})}{\partial S^3}),$$

where the coefficients V_2 and V_3 are computed like the volatility $\bar{\sigma}$ in Equation 9.8 using the operator $\langle . \rangle$:

$$V_2 = \sqrt{2}\rho\nu\langle f\phi' \rangle,$$
$$V_3 = \frac{\rho\nu}{\sqrt{2}}\langle f\phi' \rangle,$$

and ϕ is a solution of the Poisson equation:

$$\mathscr{L}^{(0)}\phi(y) = f(y)^2 - \langle f^2 \rangle.$$

Moreover, the authors show in [45] that the error of the approximation for call(put) option behaves like:

$$\lim_{\varepsilon \downarrow 0} \frac{|\text{Call}^{\varepsilon} - \text{Call}_{BS}^{(0)} - \sqrt{\varepsilon}\text{Correction}^{(1)}|}{\varepsilon |\ln(\varepsilon)|^{1+p}} = 0,$$

for any $p > 0$. The drawback of such method is that the approximation is not valid for small maturities since it uses the asymptotic regime as a model proxy.

Geodesic approach. There are many interesting papers and books about the geodesic approach and its applications in finance:

- see Chavel [29] for an introduction about Riemannian geometry and Varadhan ([109]) for an asymptotic of the density for small time,

- see Berestycki et al [21], Labordere [74] and [75], Lewis [80], Forde [43], Benhamou et al [17] for the application of the Riemannian geometry in finance.

These works show that the Green function π (the pricing density) which solves:

$$\frac{\partial \pi}{\partial t} = \sum_{i,j} g^{i,j} \frac{\partial^2 \pi}{\partial S^i \partial S^j} + \sum_i h_i \frac{\partial \pi}{\partial S^i},$$

with initial condition $\pi(t_0, S_0, t_0, S) = \delta(S_0 = S)$, has the short maturity expansion:

$$\pi(t_0, S_0, t, S) = e^{\frac{-d^2(S_0,S)}{2(t-t_0)}} \left(G_0(S_0,S) + (t-t_0)G_1(S_0,S) + \cdots\right)$$

where $d(S_0,S)$ is the geodesic distance associated with the Riemannian space defined by the metric distance $ds^2 = \sum_{i,j} g_{i,j} dx^i dx^j$, where $g_{i,j}$ is the inverse of the matrix $g^{i,j}$. The term G_0 is related to the Gaussian distribution. Indeed, the pricing density behaves like a Gaussian one when we use these new geodesic variables. The term G_1 is related to the curvature of the Riemannian space.

Moreover, Beresticky et al (see Th 1.2 in [21]) show in a general stochastic volatility model that the implied Black Scholes volatility near the expiry is:

$$\lim_{T \downarrow t} \sigma_{imp}(T,K) = \frac{\ln(\frac{S}{K})}{d(\ln(\frac{S}{K}), y_0)} \tag{9.9}$$

where t the starting time, T the maturity, S the spot, K the strike, y_0 the value of the initial stochastic volatility and d is the signed geodesic distance that solves the Eikonal Equation (1.12) in Theorem 1.12 in [21]. Notice also that Beresticky et al in [21] derive corrections terms for the short time expansion of the implied Black Scholes for some cases.

Hence, in general stochastic models, the explicit computation of the implied Black Scholes volatility near the maturity requires the explicit computation of the geodesic distance d. Therefore, we give some applications of these results in the case of SABR and Heston models.

SABR model. We recall the SABR model introduced by Hagan in [61]. In this model, the spot (S_t) follows the dynamics:

$$\begin{aligned} dS_t &= rS_t dt + \sigma S_t^{\beta} dW_t, S_0 > 0, \\ d\sigma_t &= \nu \sigma_t dB_t, \sigma_0 = \alpha, \\ d\langle W, B \rangle_t &= \rho dt, \end{aligned}$$

where α is the initial volatility, β is the skew parameter, ν is the volatility of volatility and ρ is the correlation. The geodesic distance in this model can be computed explicitly (see paragraph 6.1 in [21]):

$$d(x,y) = \frac{1}{\nu} \ln\left(\frac{\nu z - \rho + \sqrt{1 - 2\rho\nu z + \nu^2 z^2}}{1-\rho}\right)$$

where $z = \frac{\hat{x}}{y}$, $\hat{x} = \int_0^x \frac{d\xi}{\sigma \xi^{\beta-1}}$. This formula combined with Equation 9.9 gives the implied Black Scholes volatility near the maturity. Analogously, Hagan et al in [61] give an accurate analytic formula for the implied Black Scholes using a short maturity expansion and related geodesic transformation[1].

[1] However, Obloj shows in [89] that there is a difference between Hagan's formula and the formula derived from Berestycki et al in [21] and advices the formula derived by Berestycki et al since it gives more accurate results.

Heston model. We recall that the Heston model (see [63]) is an extension of Black Scholes model for the underlying (S_t) with stochastic volatility:

$$\frac{dS_t}{S_t} = r\mathrm{d}t + \sqrt{v_t}dW_t,\ S_0 > 0, \tag{9.10}$$

$$\mathrm{d}v_t = \kappa(\theta - v_t)\mathrm{d}t + \xi\sqrt{v_t}\mathrm{d}B_t,\ v_0 > 0, \tag{9.11}$$

$$\mathrm{d}\langle W, B\rangle_t = \rho\mathrm{d}t, \tag{9.12}$$

where r is the risk free rate, v_0 is the initial square of volatility, κ is the mean reversion parameter, θ is the long-term level, ξ is the volatility of volatility and ρ is the correlation. In the case of Heston's model, the geodesic distance (see [43]) can be related to the Legendre transform Λ^* of the function $\Lambda : \mathbb{R} \to \mathbb{R}$ by:

$$d(x, \sqrt{v_0}) = \sqrt{2\Lambda^*(x)}, \quad \forall x \in R,$$

where

$$\begin{aligned}\Lambda(p) &= \frac{v_0 p}{\xi(\sqrt{1-\rho^2}\cot(\frac{1}{2}\xi p\sqrt{1-\rho^2}) - \rho)} \quad for \quad p \in]p_-, p_+[,\\ &= \infty \quad for \quad p \notin]p_-, p_+[,\end{aligned}$$

and the values p_-, p_+ are computed in Theorem 1.1 in [43]. Notice also the interesting series expansion for the geodesic distance in Heston's model developed in [80].

However, these kind of expansions using geodesic approach are accurate only for short maturities and valid for time homogeneous parameters.

Alternative approaches. There are various alternative approaches among which :

- Averaging technique. The averaging technique, introduced by Piterbarg [93] has emerged as an important technique. Piterbarg derives averaging constant Heston parameters for the Call price in a time dependent Heston model; the approximation is derived only under zero correlation assumption.

- Price expansion w.r.t. correlation. Antonelli and Scarlatti consider in [12] a general stochastic model and derive a series expansion for the Call price w.r.t. correlation; the error bounds of the series expansion w.r.t. correlation are available; the computations are available only for time homogeneous parameters.

- Small volatility of volatility expansion. Lewis in [79] derives a call price expansion w.r.t. the volatility of volatility. His work is based on formal expansion of the PDE operator for Fourier transform of the Call price and restricted to constant parameters. His results are probably the closest to ours.

9.3 Motivation and main results

Our aim in this Part is to

- give an **accurate analytic approximation** for the price of call-put option

$$e^{-rT}\mathbb{E}[(K-e^{(r-q)T+X_T})_+], \tag{9.13}$$

- work in a **very general time dependent Heston framework** to obtain an approximation which can be valid for both short and long maturities and can handle time dependent parameters and non null correlation as well,

- achieve a **computational time cheaper than Fourier inversion (gain by a factor 100 or more)**.

This is done in the chapter 10 using an accurate expansion w.r.t. small volatility of volatility combined with Malliavin calculus techniques and other stochastic analysis results.

Indeed, when the volatility of volatility equals to zero, the Heston model reduces to Black Scholes model with time dependent volatility:

$$\begin{aligned}
\mathrm{d}X_t^{BS} &= \sqrt{v_{0,t}}dW_t - \frac{v_{0,t}}{2}\mathrm{d}t,\ X_0^{BS} = x_0,\\
\mathrm{d}v_{0,t} &= \kappa(\theta_t - v_{0,t})\mathrm{d}t, v_0 > 0.
\end{aligned}$$

In this case, the put price is $P_{BS}(x_0, \int_0^T v_{0,t}\mathrm{d}t)$, where the function $(x,y) \to P_{BS}(x,y)$ is the put function price in a BS model with spot e^x, strike K, total variance y, risk-free rate r, dividend yield q and maturity T. We recall that $P_{BS}(x,y)$ has the following explicit expression:

$$\begin{aligned}
P_{BS}(x,y) =& Ke^{-rT}\mathcal{N}\left(\frac{1}{\sqrt{y}}\log(\frac{Ke^{-rT}}{e^x e^{-qT}}) + \frac{1}{2}\sqrt{y}\right)\\
&- e^x e^{-qT}\mathcal{N}\left(\frac{1}{\sqrt{y}}\log(\frac{Ke^{-rT}}{e^x e^{-qT}}) - \frac{1}{2}\sqrt{y}\right).
\end{aligned}$$

Using a suitable parameterization

$$\mathrm{d}v_t^{\varepsilon} = \kappa(\theta_t - v_t^{\varepsilon})\mathrm{d}t + \varepsilon\xi_t\sqrt{v_t^{\varepsilon}}\mathrm{d}B_t,$$

we have $v_t^1 = v_t$ and $v_t^0 = v_{0,t}$. Moreover, $X_T = x_0 + \int_0^T(\sqrt{v_t}dW_t - \frac{v_t}{2}\mathrm{d}t)$ conditioned on the sigma-field $\mathscr{F}_T^B$ generated by the Brownian motion B is a Gaussian distribution with mean $x_0 + \int_0^T \rho_t\sqrt{v_t}\mathrm{d}B_t - \frac{1}{2}\int_0^T v_t\mathrm{d}t$ and variance $\int_0^T(1-\rho_t^2)v_t\mathrm{d}t$ (see Renault and Touzi in [97]). Then, the put price $e^{-rT}\mathbb{E}[(K - e^{(r-q)T+X_T})_+]$ is an expectation of a stochastic Black Scholes put price:

$$\mathbb{E}[P_{BS}(x_0 + \int_0^T \rho_t\sqrt{v_t}\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2}v_t\mathrm{d}t, \int_0^T (1-\rho_t^2)v_t\mathrm{d}t)].$$

Therefore, using a Taylor expansion for the P_{BS} function at the second order combined with Malliavin calculus techniques and some technical Lemmas, we show that

$$\begin{aligned}
e^{-rT}\mathbb{E}[(K-e^{(r-q)T+X_T})_+] = P_{BS}(x_0, \int_0^T v_{0,t}dt)&\\
+ Correction\ \ terms + Errors.&
\end{aligned}$$

where

- The corrections terms are a linear combination of Greeks of the leading Black Scholes price $P_{BS}(x_0, \int_0^T v_{0,t}dt)$ with weights depending only on the model parameters:

$$Correction\ \ terms = \sum_{i=1}^{2} a_{i,T} \frac{\partial^{i+1} P_{BS}}{\partial x^i y}(x_0, \int_0^T v_{0,t}dt) + \sum_{i=0}^{1} b_{2i,T} \frac{\partial^{2i+2} P_{BS}}{\partial x^{2i} y^2}(x_0, \int_0^T v_{0,t}dt),$$

with

$$a_{1,T} = \int_0^T \int_{t_1}^T e^{\kappa t_1} \rho_{t_1} \xi_{t_1} v_{0,t_1} e^{-\kappa t_2} dt_2 dt_1,$$
$$a_{2,T} = \int_0^T \int_{t_1}^T \int_{t_2}^T e^{\kappa t_1} \rho_{t_1} \xi_{t_1} v_{0,t_1} \rho_{t_2} \xi_{t_2} e^{-\kappa t_3} dt_3 dt_2 dt_1,$$
$$b_{0,T} = \int_0^T \int_{t_1}^T \int_{t_2}^T e^{2\kappa t_1} \xi_{t_1}^2 v_{0,t_1} e^{-\kappa t_2} e^{-\kappa t_3} dt_3 dt_2 dt_1,$$
$$b_{2,T} = \frac{a_{1,T}^2}{2}.$$

These computation are done using Malliavin calculus in Theorem 10.2.1.

- The errors are estimated in Theorem 10.2.2 by:

$$Errors = O(|\xi|_\infty^3 T^2).$$

The proof is based on some technical Lemmas.

Remark 9.3.1. • *In case of the Heston model with constant parameters, one has:*

$$a_{1,T} = \rho\xi(p_0 v_0 + p_1\theta), a_{2,T} = (\rho\xi)^2(q_0 v_0 + q_1\theta), b_{0,T} = \xi^2(r_0 v_0 + r_1\theta).$$

where

$$p_0 = \frac{e^{-\kappa T}\left(-\kappa T + e^{\kappa T} - 1\right)}{\kappa^2}, p_1 = \frac{e^{-\kappa T}\left(\kappa T + e^{\kappa T}(\kappa T - 2) + 2\right)}{\kappa^2},$$
$$q_0 = \frac{e^{-\kappa T}\left(-\kappa T(\kappa T + 2) + 2e^{\kappa T} - 2\right)}{2\kappa^3},$$
$$q_1 = \frac{e^{-\kappa T}\left(2e^{\kappa T}(\kappa T - 3) + \kappa T(\kappa T + 4) + 6\right)}{2\kappa^3},$$
$$r_0 = \frac{e^{-2\kappa T}\left(-4e^{\kappa T}\kappa T + 2e^{2\kappa T} - 2\right)}{4\kappa^3},$$
$$r_1 = \frac{e^{-2\kappa T}\left(4e^{\kappa T}(\kappa T + 1) + e^{2\kappa T}(2\kappa T - 5) + 1\right)}{4\kappa^3}.$$

- *In case of Heston model with piecewise constant parameters, we use a recursive calculus of the coefficients $a_{1,T}, a_{2,T}, b_{0,T}$ (see Subsection 10.2.5).*

- *From the approximation formula, we also derive some corollaries related first to equivalent Heston models (extending some work of Piterbarg on stochastic volatility models [93]) and second, to the calibration procedure in terms of ill-posed problems (see Subsection 10.2.6).*

The following chapter is exactly the article "Time dependent Heston model" is in revision for the journal "SIAM on Financial Mathematics". The chapter 11 presents numerical results concerning the smile behavior for Heston's model with constant and time dependent parameters as well. Moreover, in this chapter, we review some results concerning the negative moments of the integrated CIR process.

Chapter 10

Time dependent Heston model

In revision for "SIAM Journal on Financial Mathematics".

The use of the Heston model is still challenging because it has a closed formula only when the parameters are constant [63] or piecewise constant [86]. Hence, using a small volatility of volatility expansion and Malliavin calculus techniques, we derive an accurate analytical formula for the price of vanilla options for any time dependent Heston model (the accuracy is less than a few bps for various strikes and maturities). In addition, we establish tight error estimates. The advantage of this approach over Fourier based methods is its rapidity (gain by a factor 100 or more), while maintaining a competitive accuracy. From the approximative formula, we also derive some corollaries related first to equivalent Heston models (extending some work of Piterbarg on stochastic volatility models [93]) and second, to the calibration procedure in terms of ill-posed problems.

10.1 Introduction

Stochastic volatility modeling has emerged in the late nineties as a way to manage the smile. In this work, we focus on the Heston model which is a lognormal model where the square of volatility follows a CIR[1] process. The call (and put) price has a closed formula in this model thanks to a Fourier inversion of the characteristic function (see Heston [63], Lewis [79] and Lipton [81]). When the parameters are piecewise constant, one can still derive a recursive closed formula using a PDE method (see Mikhailov and Nogel [86]) or a Markov argument in combination with affine models (see Elices [41]), but formula evaluation becomes increasingly time consuming. However, for general time dependent parameters there is no analytical formula and one usually has to perform Monte Carlo simulations. This explains the interest of recent works for designing more efficient Monte Carlo simulations: see Broadie and Kaya [27] for an exact simulation and bias-free scheme based on Fourier integral inversion; see Andersen [7] based on a Gaussian moment matching method and a user friendly algorithm; see Smith [107] relying on an almost exact scheme; see Alfonsi [4] using higher order schemes and a recursive method for the CIR process. For numerical partial differential equations, we refer the reader to Kluge's doctoral dissertation [70].

Comparison with the literature. A more recent trend in the quantitative literature has been the use of the so called approximation method to derive analytical formulae. This has led to an impressive number of papers, with many original ideas. For instance, Alòs et al. [6] have been studying the short time behavior of implied volatility for stochastic volatility using an extension of Itô's formula. Another trend has focused on analytical techniques to derive the asymptotic expansion of the implied volatility near expiry (see for instance Berestycki et al. [21], Labordere [74], Hagan et al. [61], Lewis [80], Osajima [90] or Forde [43]). But in these works the implied volatility near expiry does not have a closed formula because the related geodesic distance is not explicit. It can, however, be approximated by a series expansion [80]. The drawback to these methods is their inability to handle non-homogeneous (that is to say time dependent) parameters. For long maturities, another approach has been the asymptotic expansion w.r.t. the mean reversion parameter of the volatility as shown in [44]. In the case of zero correlation, averaging techniques as exposed in [93] and [92] can be used. Antonelli and Scarlatti take another view in [12] and have suggested price expansion w.r.t. correlation. For all of these techniques, the domain of availability of the expansion is restricted to either short or long maturities, to zero correlation, or to homogeneous parameters. In our work, we aim to give an analytical formula which covers both short and long maturities, that also handles time inhomogeneous parameters as well as non-null correlations. As a difference with several previously quoted papers, our purpose consists also of justifying mathematically our approximation.

The results closest to ours are probably those based on an expansion w.r.t. the volatility of volatility

[1] Nice properties for the CIR process are derived by Dufresne [39], Göing-Jaeschke and Yor [56], Diop [38] and Alfonsi [3].

by Lewis [79]: it is based on formal analytical arguments and is restricted to constant parameters. Our formula can be viewed as an extension of Lewis' formula in order to address a time dependent Heston model, using a direct probabilistic approach. In addition, we prove an error estimate which shows that our approximation formula for call/put is of order 2 w.r.t. the volatility of volatility. The advantage of this current approximation is that the evaluation is about 100 to 1000 times quicker than a Fourier based method (see our numerical tests).

Comparison with our previous works (see Chapters 4 and 7). Our approach here consists of expanding the price w.r.t. the volatility of volatility, and of computing the correction terms using Malliavin calculus. In these respects, the current approach is similar to our previous works (see Chapters 4 and 7), however, the techniques for estimating error are different. Indeed, we use the fact that the price of vanilla options can be expressed as an expectation of a smooth price function for stochastic volatility models. This is based on a conditioning argument as in [97]. Consequently, the smoothness hypotheses (H_1, H_2, H_3) of our previous Chapters are no longer required. Note also that the square root function arising in the martingale part of the CIR process is not Lipschitz continuous. Hence, the Heston model does not fit the smoothness framework previously used. Therefore, to overcome this difficulty, we derive new technical results in order to prove the accuracy of the formula.

Contribution of the chapter. We give an explicit analytical formula for the price of vanilla options in a time dependent Heston model. Our approach is based on an expansion w.r.t. a small volatility of volatility. This is practically justified by the fact that this parameter is usually quite small (of order 1 or less, see [79] or [27] for instance). The resulting formula is the sum of two terms: the leading term is the Black-Scholes price for the model without volatility of volatility while the correction term is a combination of Greeks of the leading term with explicit weights depending only on the model parameters. Proving the accuracy of the expansion is far from straightforward, but with some technicalities and a relevant analysis of error, we succeed in giving tight error estimates. Our expansion enables us to obtain averaged parameters for the dynamic Heston model.

Formulation of the problem. We consider the solution of the stochastic differential equation (SDE):

$$
\begin{aligned}
\mathrm{d}X_t &= \sqrt{v_t}dW_t - \frac{v_t}{2}\mathrm{d}t,\ X_0 = x_0, &(10.1)\\
\mathrm{d}v_t &= \kappa(\theta_t - v_t)\mathrm{d}t + \xi_t\sqrt{v_t}\mathrm{d}B_t,\ v_0, &(10.2)\\
\mathrm{d}\langle W, B\rangle_t &= \rho_t \mathrm{d}t,
\end{aligned}
$$

where $(B_t, W_t)_{0\le t\le T}$ is a two-dimensional correlated Brownian motion on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\le t\le T}, \mathbb{P})$ with the usual assumptions on filtration $(\mathscr{F}_t)_{0\le t\le T}$. In our setting, $(X_t)_t$ is the log of the forward price and $(v_t)_t$ is the square of the volatility which follows a CIR process with an initial value $v_0 > 0$, a positive mean reversion κ, a positive long-term level $(\theta_t)_t$, a positive volatility of volatility $(\xi_t)_t$ and a correlation $(\rho_t)_t$. These time dependent parameters are assumed to be measurable and bounded on $[0, T]$.

To develop our approximation method, we will examine the following perturbed process w.r.t. $\varepsilon \in [0, 1]$:

$$
\begin{aligned}
\mathrm{d}X_t^\varepsilon &= \sqrt{v_t^\varepsilon}dW_t - \frac{v_t^\varepsilon}{2}\mathrm{d}t, \quad X_0^\varepsilon = x_0,\\
\mathrm{d}v_t^\varepsilon &= \kappa(\theta_t - v_t^\varepsilon)\mathrm{d}t + \varepsilon\xi_t\sqrt{v_t^\varepsilon}\mathrm{d}B_t, \quad v_0^\varepsilon = v_0, &(10.3)
\end{aligned}
$$

so that our perturbed process coincides with the initial one for $\varepsilon = 1 : X_t^1 = X_t, v_t^1 = v_t$. For the existence of the solution v^ε, we refer to Chapter IX in [98] (moreover, the process is non-negative for $k\theta_t \ge 0$, see also the proof of Lemma 10.4.2). Our main purpose is to give an accurate analytic approximation, in a certain sense, of the expected payoff of a put option :

$$g(\varepsilon) = e^{-\int_0^T r_t \mathrm{d}t}\mathbb{E}[(K - e^{\int_0^T (r_t - q_t)\mathrm{d}t + X_T^\varepsilon})_+] \tag{10.4}$$

where r (resp. q) is the risk-free rate (resp. the dividend yield), T is the maturity and $\varepsilon = 1$. Extensions to call options and other payoffs are discussed later.

Outline of the chapter. In Section 10.2, we explain the methodology of the small volatility of volatility expansion. An approximation formula is then derived in Theorem 10.2.1 and its accuracy stated in Theorem 10.2.2. This section ends by explicitly expressing the formula's coefficients for general time dependent parameters (constant, smooth and piecewise constant). Our expansion allows us to give equivalent constant parameters for the time dependent Heston model (see Subsection 10.2.6). As a second corollary, the options calibration for Heston's model using only one maturity becomes an ill-posed problem; we give numerical results to confirm this situation. In section 10.3, we provide numerical tests to benchmark our formula with the closed formula in the case of constant and piecewise constant parameters. In Section 10.4, we prove the accuracy of the approximation stated in Theorem 10.2.2: this section is the technical core of the chapter. In Section 10.5, we establish lemmas used to make the calculation of the correction terms explicit (those derived in Theorem 10.2.1). In Section 10.6, we conclude this work and give a few extensions. In the appendix, we recall details about the closed formula (of Heston [63] and Lewis [79]) in the case of constant (and piecewise constant) parameters.

10.2 Smart Taylor expansion

10.2.1 Notations

Notation 10.2.1. ***Extremes of deterministic functions.***
For a càdlàg function $l : [0,T] \to \mathbb{R}$, we denote $l_{Inf} = \inf_{t\in[0,T]} l_t$ and $l_{Sup} = \sup_{t\in[0,T]} l_t$.

Notation 10.2.2. ***Differentiation.***

- *For a smooth function $x \mapsto l(x)$, we denote by $l^{(i)}(x)$ its i-th derivative.*
- *Given a fixed time t and for a function $\varepsilon \to f_t^\varepsilon$, we denote (if it has a meaning) the i^{th} derivative at $\varepsilon = 0$ by $f_{i,t} = \frac{\partial^i f_t^\varepsilon}{\partial \varepsilon^i}|_{\varepsilon=0}$.*

10.2.2 Definitions

In order to make the approximation explicit, we introduce the following family of operators indexed by maturity T.

Definition 10.2.1. ***Integral Operator.*** *We define the integral operator $\omega_{.,T}^{(.,.)}$ as follows:*

- *For any real number k and any integrable function l, we set*

$$\omega_{t,T}^{(k,l)} = \int_t^T e^{ku} l_u \mathrm{d}u, \ \forall t \in [0,T].$$

- *For any real numbers $(k_1,\cdots,k_n)$ and for any integrable functions $(l_1,\cdots,l_n)$, the n-times iteration is given by*
$$\omega_{t,T}^{(k_1,l_1),\cdots,(k_n,l_n)} = \omega_{t,T}^{(k_1,l_1\omega_{.,T}^{(k_2,l_2),\cdots,(k_n,l_n)})},\ \forall t\in[0,T].$$
- *When the functions $(l_1,\cdots,l_n)$ are equal to the unity constant function* 1*, we simply write*
$$\tilde{\omega}_{t,T}^{k_1,\cdots,k_n} = \omega_{t,T}^{(k_1,1),\cdots,(k_n,1)},\ \forall t\in[0,T].$$

10.2.3 About the CIR process

Assumptions. In order to bound the approximation errors, we need a positivity assumption for the CIR process.

***Assumption** (P). The parameters of the CIR process (10.2) verify the following conditions:*

$$\xi_{Inf} > 0, \quad (\frac{2\kappa\theta}{\xi^2})_{Inf} \geq 1.$$

This assumption is crucial to ensure the positivity of the process on $[0,T]$, which is stated in detail in Lemma 10.4.2 (remember that $v_0 > 0$). We have

$$\mathbb{P}(\forall t\in[0,T]:\ v_t > 0) = 1.$$

When the functions θ and ξ are constant, Assumption (P) coincides with the usual Feller test condition $\frac{2\kappa\theta}{\xi^2} \geq 1$ (see [69]).

Note that the above assumption ensures that the positivity property also holds for the perturbed CIR process (10.3): for any $\varepsilon\in[0,1]$, we have

$$\mathbb{P}(\forall t\in[0,T]:\ v_t^\varepsilon > 0) = 1$$

(see Lemma 10.4.2). We also need a uniform bound of the correlation in order to preserve the non degeneracy of the SDE (10.1) conditionally on $(B_t)_{0\leq t\leq T}$.

***Assumption** (R). The correlation is bounded away from -1 and +1:*

$$|\rho|_{Sup} < 1.$$

10.2.4 Taylor Development

In this paragraph, we present the main steps leading to our results. Complete proofs are given later.

If $(\mathscr{F}_t^B)_t$ denotes the filtration generated by the Brownian motion B, the distribution of X_T^ε conditionally to $\mathscr{F}_T^B$ is a Gaussian distribution with mean $x_0 + \int_0^T \rho_t\sqrt{v_t^\varepsilon}\mathrm{d}B_t - \frac{1}{2}\int_0^T v_t^\varepsilon\mathrm{d}t$ and variance $\int_0^T(1-\rho_t^2)v_t^\varepsilon\mathrm{d}t$ ($\varepsilon\in[0,1]$). Therefore, the function (10.4) can be expressed as follows:

$$g(\varepsilon) = \mathbb{E}[P_{BS}(x_0 + \int_0^T \rho_t\sqrt{v_t^\varepsilon}\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2}v_t^\varepsilon\mathrm{d}t, \int_0^T (1-\rho_t^2)v_t^\varepsilon\mathrm{d}t)], \tag{10.5}$$

where the function $(x,y) \to P_{BS}(x,y)$ is the put function price in a Black-Scholes model with spot e^x, strike K, total variance y, risk-free rate $r_{eq} = \frac{\int_0^T r(t)\mathrm{d}t}{T}$, dividend yield $q_{eq} = \frac{\int_0^T q(t)\mathrm{d}t}{T}$ and maturity T. For the sake of completeness, we recall that $P_{BS}(x,y)$ has the following explicit expression

$$Ke^{-r_{eq}T}\mathcal{N}\left(\frac{1}{\sqrt{y}}\log(\frac{Ke^{-r_{eq}T}}{e^x e^{-q_{eq}T}})+\frac{1}{2}\sqrt{y}\right) - e^x e^{-q_{eq}T}\mathcal{N}\left(\frac{1}{\sqrt{y}}\log(\frac{Ke^{-r_{eq}T}}{e^x e^{-q_{eq}T}})-\frac{1}{2}\sqrt{y}\right).$$

In the following, we expand $P_{BS}(.,.)$ with respect to its two arguments. For this, we note that P_{BS} is a smooth function (for $y>0$). In addition, there is a simple relation between its partial derivatives:

$$\frac{\partial P_{BS}}{\partial y}(x,y) = \frac{1}{2}(\frac{\partial^2 P_{BS}}{\partial x^2}(x,y) - \frac{\partial P_{BS}}{\partial x}(x,y)), \quad \forall x \in \mathbb{R}, \forall y > 0, \tag{10.6}$$

which can be proved easily by a standard calculation left to the reader.
Under assumption (P), for any t, v_t^ε is C^2 w.r.t ε at $\varepsilon = 0$ (differentiation in L_p-sense). This result will be shown later. In addition, v^ε does not vanish (for any $\varepsilon \in [0,1]$). Hence, by putting $v_{i,t}^\varepsilon = \frac{\partial^i v_t^\varepsilon}{\partial \varepsilon^i}$, we get

$$\mathrm{d}v_{1,t}^\varepsilon = -\kappa v_{1,t}^\varepsilon \mathrm{d}t + \xi_t\sqrt{v_t^\varepsilon}\mathrm{d}B_t + \varepsilon\xi_t\frac{v_{1,t}^\varepsilon}{2\sqrt{v_t^\varepsilon}}\mathrm{d}B_t, \qquad v_{1,0}^\varepsilon = 0,$$

$$\mathrm{d}v_{2,t}^\varepsilon = -\kappa v_{2,t}^\varepsilon \mathrm{d}t + \xi_t\frac{v_{1,t}^\varepsilon}{\sqrt{v_t^\varepsilon}}\mathrm{d}B_t + \varepsilon\xi_t\frac{v_{2,t}^\varepsilon}{2\sqrt{v_t^\varepsilon}}\mathrm{d}B_t - \varepsilon\xi_t\frac{[v_{1,t}^\varepsilon]^2}{4[v_t^\varepsilon]^{3/2}}\mathrm{d}B_t, \qquad v_{2,0}^\varepsilon = 0.$$

From the definitions $v_{i,t} \equiv \frac{\partial^i v_t^\varepsilon}{\partial \varepsilon^i}|_{\varepsilon=0}$, we easily deduce

$$v_{0,t} = e^{-\kappa t}(v_0 + \int_0^t \kappa e^{\kappa s}\theta_s \mathrm{d}s),$$

$$v_{1,t} = e^{-\kappa t}\int_0^t e^{\kappa s}\xi_s\sqrt{v_{0,s}}\mathrm{d}B_s, \tag{10.7}$$

$$v_{2,t} = e^{-\kappa t}\int_0^t e^{\kappa s}\xi_s\frac{v_{1,s}}{(v_{0,s})^{\frac{1}{2}}}\mathrm{d}B_s. \tag{10.8}$$

Note that $v_{0,t}$ coincides also with the expected variance $\mathbb{E}(v_t)$ because of the linearity of the drift coefficient of $(v_t)_t$. Now, to expand $g(\varepsilon)$, we use the Taylor formula twice, first applied to $\varepsilon \to v_t^\varepsilon$ and $\sqrt{v_t^\varepsilon}$ at $\varepsilon = 1$ using derivatives computed at $\varepsilon = 0$:

$$v_t^1 = v_{0,t} + v_{1,t} + \frac{v_{2,t}}{2} + \cdots,$$

$$\sqrt{v_t^1} = \sqrt{v_{0,t}} + \frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}} + \frac{v_{2,t}}{4(v_{0,t})^{\frac{1}{2}}} - \frac{v_{1,t}^2}{8(v_{0,t})^{\frac{3}{2}}} + \cdots,$$

secondly for the smooth function P_{BS} at the second order w.r.t. the first and second variable around $(x_0 + \int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2}v_{0,t}\mathrm{d}t, \int_0^T(1-\rho_t^2)v_{0,t}\mathrm{d}t)$. For convenience, we simply write

$$\tilde{P}_{BS} = P_{BS}(x_0 + \int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2}v_{0,t}\mathrm{d}t, \int_0^T(1-\rho_t^2)v_{0,t}\mathrm{d}t), \tag{10.9}$$

$$\frac{\partial^{i+j}\tilde{P}_{BS}}{\partial x^i y^j} = \frac{\partial^{i+j}P_{BS}}{\partial x^i y^j}(x_0 + \int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2}v_{0,t}\mathrm{d}t, \int_0^T(1-\rho_t^2)v_{0,t}\mathrm{d}t).$$

Then, one gets

$$
\begin{aligned}
g(1) =& \mathbb{E}[\tilde{P}_{BS}] && (10.10)\\
&+\mathbb{E}[\frac{\partial \tilde{P}_{BS}}{\partial x}(\int_0^T \rho_t(\frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}+\frac{v_{2,t}}{4(v_{0,t})^{\frac{1}{2}}}-\frac{v_{1,t}^2}{8(v_{0,t})^{\frac{3}{2}}})\mathrm{d}B_t-\int_0^T\frac{\rho_t^2}{2}(v_{1,t}+\frac{v_{2,t}}{2})\mathrm{d}t)] && (10.11)\\
&+\mathbb{E}[\frac{\partial \tilde{P}_{BS}}{\partial y}\int_0^T(1-\rho_t^2)(v_{1,t}+\frac{v_{2,t}}{2})\mathrm{d}t] && (10.12)\\
&+\frac{1}{2}\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial x^2}(\int_0^T\rho_t\frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}\mathrm{d}B_t-\int_0^T\frac{\rho_t^2}{2}v_{1,t}\mathrm{d}t)^2] && (10.13)\\
&+\frac{1}{2}\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}(\int_0^T(1-\rho_t^2)v_{1,t}\mathrm{d}t)^2] && (10.14)\\
&+\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial xy}(\int_0^T(1-\rho_t^2)v_{1,t}\mathrm{d}t)(\int_0^T\rho_t\frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}\mathrm{d}B_t-\int_0^T\frac{\rho_t^2}{2}v_{1,t}\mathrm{d}t)] && (10.15)\\
&+\mathscr{E} && (10.16)
\end{aligned}
$$

where $\mathscr{E}$ is the error in our Taylor expansion. In fact, we notice that:

$$
\begin{aligned}
\mathbb{E}[\tilde{P}_{BS}] &= \mathbb{E}[\mathbb{E}[e^{-\int_0^T r_t\mathrm{d}t}(K-e^{x_0+\int_0^T(r_t-q_t-\frac{v_{0,t}}{2})\mathrm{d}t+\int_0^T\sqrt{v_{0,t}}(\rho_t\mathrm{d}B_t+\sqrt{1-\rho_t^2}\mathrm{d}B_t^{\perp})})_+|\mathscr{F}_T^B]]\\
&= P_{BS}(x_0,\int_0^T v_{0,t}\mathrm{d}t),
\end{aligned}
$$

where $B^{\perp}$ is a Brownian motion independent on $\mathscr{F}_T^B$. Furthermore, the relation (10.6) remains the same for $\tilde{P}_{BS}$ and this enables us to simplify the expansion above. This gives:

Proposition 10.2.1. *The approximation* (10.16) *is equivalent to*

$$
g(1) = P_{BS}(x_0,\int_0^T v_{0,t}\mathrm{d}t)+\mathbb{E}[\frac{\partial \tilde{P}_{BS}}{\partial y}\int_0^T(v_{1,t}+v_{2,t})\mathrm{d}t]+\frac{1}{2}\mathbb{E}[\frac{\partial^2\tilde{P}_{BS}}{\partial y^2}(\int_0^T v_{1,t}\mathrm{d}t)^2]+\mathscr{E}.
$$

The details of the proof are given in Subsection 10.5.2. At first sight, the above formula looks like a Taylor formula of P_{BS} w.r.t. the cumulated variance. In fact, it is different, note that the coefficient of $v_{2,t}$ is not $1/2$ but 1. We do not have any direct interpretation of this formula.

The next step consists of making explicit the correction terms as a combination of Greeks of the BS price.

Theorem 10.2.1. *Under assumptions (P) and (R), the put*[2] *price is approximated by*

$$
\begin{aligned}
e^{-\int_0^T r_t\mathrm{d}t}\mathbb{E}[(K-e^{\int_0^T(r_t-q_t)\mathrm{d}t+X_T^1})_+] =& P_{BS}(x_0,var_T)+\sum_{i=1}^{2}a_{i,T}\frac{\partial^{i+1}P_{BS}}{\partial x^i y}(x_0,var_T)\\
&+\sum_{i=0}^{1}b_{2i,T}\frac{\partial^{2i+2}P_{BS}}{\partial x^{2i}y^2}(x_0,var_T)+\mathscr{E}, \qquad (10.17)
\end{aligned}
$$

[2] The approximation formula for the call price is obtained using the call/put parity relation: in (10.17), it consists of replacing on the l.h.s. the put payoff by the call one, and on the r.h.s., the put price function P_{BS} by the similar call price function, while coefficients remain the same.

where

$$var_T = \int_0^T v_{0,t}\mathrm{d}t, \qquad a_{1,T} = \omega_{0,T}^{(\kappa,\rho\xi v_{0,.}),(-\kappa,1)}, \qquad a_{2,T} = \omega_{0,T}^{(\kappa,\rho\xi v_{0,.}),(0,\rho\xi),(-\kappa,1)},$$

$$b_{0,T} = \omega_{0,T}^{(2\kappa,\xi^2 v_{0,.}),(-\kappa,1),(-\kappa,1)}, \qquad b_{2,T} = \frac{a_{1,T}^2}{2}.$$

The proof is postponed to Subsection 10.5.3. Finally, we give an estimate regarding the error $\mathscr{E}$ arising in the above theorem.

Theorem 10.2.2. *Under assumptions (P) and (R), the error in the approximation* (10.17) *is estimated as follows:*

$$\mathscr{E} = O\left([\xi_{Sup}\sqrt{T}]^3\sqrt{T}\right).$$

In view of Theorem 10.2.2, we may refer to the formula (10.17) as a second order approximation formula w.r.t. the volatility of volatility.

10.2.5 Computation of coefficients

Constant parameters The case of constant parameters (θ,ξ,ρ) gives us the coefficients a and b explicitly. Indeed in this case, the operator ω is a simple iterated integration of exponential functions. Using Mathematica, we derive the following explicit expressions.

Proposition 10.2.2. ***Explicit computations.*** *For constant parameters, one has:*

$$var_T = m_0 v_0 + m_1\theta, \qquad a_{1,T} = \rho\xi(p_0 v_0 + p_1\theta),$$
$$a_{2,T} = (\rho\xi)^2(q_0 v_0 + q_1\theta), \qquad b_{0,T} = \xi^2(r_0 v_0 + r_1\theta).$$

where

$$m_0 = \frac{e^{-\kappa T}\left(-1+e^{\kappa T}\right)}{\kappa}, \qquad m_1 = T - \frac{e^{-\kappa T}\left(-1+e^{\kappa T}\right)}{\kappa},$$
$$p_0 = \frac{e^{-\kappa T}\left(-\kappa T+e^{\kappa T}-1\right)}{\kappa^2}, \qquad p_1 = \frac{e^{-\kappa T}\left(\kappa T+e^{\kappa T}(\kappa T-2)+2\right)}{\kappa^2},$$
$$q_0 = \frac{e^{-\kappa T}\left(-\kappa T(\kappa T+2)+2e^{\kappa T}-2\right)}{2\kappa^3}, \qquad q_1 = \frac{e^{-\kappa T}\left(2e^{\kappa T}(\kappa T-3)+\kappa T(\kappa T+4)+6\right)}{2\kappa^3},$$
$$r_0 = \frac{e^{-2\kappa T}\left(-4e^{\kappa T}\kappa T+2e^{2\kappa T}-2\right)}{4\kappa^3}, \qquad r_1 = \frac{e^{-2\kappa T}\left(4e^{\kappa T}(\kappa T+1)+e^{2\kappa T}(2\kappa T-5)+1\right)}{4\kappa^3}.$$

Remark 10.2.1. *In the case of constant parameters (θ,ξ,ρ), we retrieve the usual Heston model. In this particular case, our expansion coincides exactly with Lewis' volatility of volatility series expansion (see Equation (3.4), page 84 in [79] for Lewis' expansion formula and page 93 in [79] for the explicit calculation of the coefficients $J^{(i)}$ with $\varphi = \frac{1}{2}$). Using his notation, we have $a_{1,T} = J^{(1)}$, $a_{2,T} = J^{(4)}$ and $b_{0,T} = J^{(3)}$.*

Smooth parameters In this case, we may use a Gauss-Legendre quadrature formula for the computation of the terms a and b.

Piecewise constant parameters The computation of the variance var_T is straightforward. Thus, it remains to provide explicit expressions of a and b as a function of the piecewise constant data. Let $T_0 = 0 \le T_1 \le \cdots \le T_n = T$ such that θ, ρ, ξ are constant on each interval $]T_i, T_{i+1}[$ and are equal respectively to $\theta_{T_{i+1}}, \rho_{T_{i+1}}, \xi_{T_{i+1}}$. Before giving the recursive relation, we need to introduce the following functions: $\tilde{\omega}_{1,t} = \omega_{0,t}^{(\kappa,\rho\xi v_{0,.})}$, $\tilde{\omega}_{2,t} = \omega_{0,t}^{(2\kappa,\xi^2 v_{0,.})}$, $\alpha_t = \omega_{0,t}^{(k,\rho\xi v_{0,.}),(0,\rho\xi)}$, $\beta_t = \omega_{0,t}^{(2k,\xi^2 v_{0,.}),(-\kappa,1)}$.

Proposition 10.2.3. ***Recursive calculations.*** *For piecewise constant coefficients, one has:*

$$
\begin{aligned}
a_{1,T_{i+1}} &= a_{1,T_i} + \tilde{\omega}_{T_i,T_{i+1}}^{-\kappa} \tilde{\omega}_{1,T_i} + \rho_{T_{i+1}} \xi_{T_{i+1}} f_{\kappa,v_{0,T_i}}^{1}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
a_{2,T_{i+1}} &= a_{2,T_i} + \tilde{\omega}_{T_i,T_{i+1}}^{-\kappa} \alpha_{T_i} + \rho_{T_{i+1}} \xi_{T_{i+1}} \tilde{\omega}_{T_i,T_{i+1}}^{0,-\kappa} \tilde{\omega}_{1,T_i} + (\rho_{T_{i+1}} \xi_{T_{i+1}})^2 f_{\kappa,v_{0,T_i}}^{2}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
b_{0,T_{i+1}} &= b_{0,T_i} + \tilde{\omega}_{T_i,T_{i+1}}^{-\kappa} \beta_{T_i} + \tilde{\omega}_{T_i,T_{i+1}}^{-\kappa,-\kappa} \tilde{\omega}_{2,T_i} + \xi_{T_{i+1}}^2 f_{\kappa,v_{0,T_i}}^{0}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
\alpha_{T_{i+1}} &= \alpha_{T_i} + \rho_{T_{i+1}} \xi_{T_{i+1}} (T_{i+1} - T_i) \tilde{\omega}_{1,T_i} + \rho_{T_{i+1}}^2 \xi_{T_{i+1}}^2 g_{\kappa,v_{0,T_i}}^{1}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
\beta_{T_{i+1}} &= \beta_{T_i} + \tilde{\omega}_{T_i,T_{i+1}}^{-\kappa} \tilde{\omega}_{2,T_i} + \xi_{T_{i+1}}^2 g_{\kappa,v_{0,T_i}}^{2}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
\tilde{\omega}_{1,T_{i+1}} &= \tilde{\omega}_{1,T_i} + \rho_{T_{i+1}} \xi_{T_{i+1}} h_{\kappa,v_{0,T_i}}^{1}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
\tilde{\omega}_{2,T_{i+1}} &= \tilde{\omega}_{2,T_i} + \xi_{T_{i+1}}^2 h_{\kappa,v_{0,T_i}}^{2}(\theta_{T_{i+1}}, T_i, T_{i+1}),\\
v_{0,T_{i+1}} &= e^{-\kappa(T_{i+1}-T_i)}(v_{0,T_i} - \theta_{T_{i+1}}) + \theta_{T_{i+1}},
\end{aligned}
$$

where

$$
\begin{aligned}
f_{\kappa,v_0}^{0}(\theta,t,T) &= \frac{e^{-2\kappa T}\left(e^{2\kappa t}(\theta-2v_0)+e^{2\kappa T}((-2\kappa t+2\kappa T-5)\theta+2v_0)+4e^{\kappa(t+T)}((-\kappa t+\kappa T+1)\theta+\kappa(t-T)v_0)\right)}{4\kappa^3},\\
f_{\kappa,v_0}^{1}(\theta,t,T) &= \frac{e^{-\kappa T}\left(e^{\kappa T}((-\kappa t+\kappa T-2)\theta+v_0)-e^{\kappa t}((\kappa t-\kappa T-2)\theta-\kappa t v_0+\kappa T v_0+v_0)\right)}{\kappa^2},\\
f_{\kappa,v_0}^{2}(\theta,t,T) &= \frac{e^{-\kappa(t+3T)}\left(2e^{\kappa(t+3T)}((\kappa(T-t)-3)\theta+v_0)+e^{2\kappa(t+T)}((\kappa(\kappa(t-T)-4)(t-T)+6)\theta-(\kappa(\kappa(t-T)-2)(t-T)+2)v_0)\right)}{2\kappa^3},\\
g_{\kappa,v_0}^{1}(\theta,t,T) &= \frac{2e^{\kappa T}\theta+e^{\kappa t}\left(\kappa^2(t-T)^2 v_0-(\kappa(\kappa(t-T)-2)(t-T)+2)\theta\right)}{2\kappa^2},\\
g_{\kappa,v_0}^{2}(\theta,t,T) &= \frac{e^{-\kappa T}\left(e^{2\kappa T}\theta-e^{2\kappa t}(\theta-2v_0)+2e^{\kappa(t+T)}(\kappa(t-T)(\theta-v_0)-v_0)\right)}{2\kappa^2},\\
h_{\kappa,v_0}^{1}(\theta,t,T) &= \frac{e^{\kappa T}\theta+e^{\kappa t}((\kappa t-\kappa T-1)\theta+\kappa(T-t)v_0)}{\kappa},\\
h_{\kappa,v_0}^{2}(\theta,t,T) &= \frac{\left(e^{\kappa t}-e^{\kappa T}\right)\left(e^{\kappa t}(\theta-2v_0)-e^{\kappa T}\theta\right)}{2\kappa},
\end{aligned}
$$

and $\tilde{\omega}_t^{u}(T) = \frac{-e^{tu}+e^{Tu}}{u}$, $\tilde{\omega}_t^{0,u}(T) = \frac{e^{Tu}(-tu+Tu-1)+e^{tu}}{u^2}$, $\tilde{\omega}_t^{u,u}(T) = \frac{\left(e^{tu}-e^{Tu}\right)^2}{2u^2}$.

Proof. According to Theorem 10.2.1, one has :

$$
\begin{aligned}
a_{1,T_{i+1}} &= \int_0^{T_i} e^{\kappa t} \rho_t \xi_t v_{0,t} \omega_{t,T_{i+1}}^{(-\kappa,1)} \mathrm{d}t + \int_{T_i}^{T_{i+1}} e^{\kappa t} \rho_t \xi_t v_{0,t} \omega_{t,T_{i+1}}^{(-\kappa,1)} \mathrm{d}t\\
&= a_{1,T_i} + \int_0^{T_i} e^{\kappa t} \rho_t \xi_t v_{0,t} \omega_{T_i,T_{i+1}}^{(-\kappa,1)} \mathrm{d}t + \int_{T_i}^{T_{i+1}} e^{\kappa t} \rho_t \xi_t v_{0,t} \omega_{t,T_{i+1}}^{(-\kappa,1)} \mathrm{d}t\\
&= a_{1,T_i} + \omega_{T_i,T_{i+1}}^{(-\kappa,1)} \int_0^{T_i} e^{\kappa t} \rho_t \xi_t v_{0,t} \mathrm{d}t + \int_{T_i}^{T_{i+1}} e^{\kappa t} \rho_t \xi_t v_{0,t} \omega_{t,T_{i+1}}^{(-\kappa,1)} \mathrm{d}t\\
&= a_{1,T_i} + \tilde{\omega}_{T_i,T_{i+1}}^{-\kappa} \tilde{\omega}_{1,T_i} + \rho_{T_{i+1}} \xi_{T_{i+1}} f_{\kappa,v_{0,T_i}}^{1}(\theta_{i+1}, T_i, T_{i+1}),
\end{aligned}
$$

where the functions f_{κ,v_0}^{1} and $\tilde{\omega}^{-\kappa}$ are calculated analytically using Mathematica. The other terms are calculated analogously. □

10.2.6 Corollaries of the approximation formula (10.17)

Averaging Heston's model parameters We derive a first corollary of the approximation formula in terms of equivalent Heston models. As explained in [93], this averaging principle may facilitate efficient calibration. Namely, we search for equivalent constant parameters $\bar{\kappa}, \bar{\theta}, \bar{\xi}, \bar{\rho}$ for the Heston model[3]

$$\begin{aligned} d\bar{X}_t &= \sqrt{\bar{v}_t} dW_t - \frac{\bar{v}_t}{2} \mathrm{d}t, \bar{X}_0 = x_0, \\ d\bar{v}_t &= \bar{\kappa}(\bar{\theta}_t - \bar{v}_t)\mathrm{d}t + \bar{\xi}\sqrt{\bar{v}_t}\mathrm{d}B_t, \bar{v}_0 = v_0, \\ \mathrm{d}\langle W, B\rangle_t &= \bar{\rho}\mathrm{d}t, \end{aligned}$$

that equalize the price of call/put options maturing at T in the time dependent model (equality up to the approximation error $\mathscr{E}$). The following rules give the equivalent parameters as a function of the variance var_T and the coefficients $a_{1,T}$, $a_{2,T}$, $b_{0,T}$ that are computed in the time dependent model. Results are expressed using

$$a = \frac{a_{2,T} m_1}{m_1 q_0 - m_0 q_1}, \quad b = -\frac{a_{1,T} m_1}{m_1 p_0 - m_0 p_1}, \quad c = var_T\left(\frac{p_1}{m_1 p_0 - m_0 p_1} - \frac{q_1}{m_1 q_0 - m_0 q_1}\right),$$

where $m_0, m_1, p_0, p_1, q_0, q_1, r_0$ and r_1 are given in Proposition 10.2.2.

Averaging rule in the case of zero correlation. If $\rho_t \equiv 0$, the equivalent constant parameters (for maturity T) are

$$\bar{\kappa} = \kappa, \quad \bar{\theta} = \frac{var_T - m_0 v_0}{m_1}, \quad \bar{\xi} = \sqrt{\frac{b_{0,T}}{r_0 v_0 + r_1 \bar{\theta}}}, \quad \bar{\rho} = 0.$$

Proof. Two sets of prices coincide at maturity T if they have the same approximation formula (10.17). In this case $a_{1,T} = a_{2,T} = b_{2,T} = 0$, thus the approximation formula depends only on two quantities var_T and $b_{0,T}$. It is quite clear that there is not a single choice of parameters to fit these two quantities. A simple solution results from the choice of $\bar{\kappa} = \kappa$ and $\bar{\rho} = 0$: then, using Proposition 10.2.2, we obtain the announced parameters $\bar{\theta}$ and $\bar{\xi}$. □

Remark 10.2.2. *In this case of zero correlation and $\theta = v_0 = \bar{\theta}$, we exactly retrieve Piterbarg's results for the averaged volatility of volatility $\bar{\xi}$ (see [93]).*

Averaging rule in the case of non zero correlation. We follow the same arguments as before. Now the approximation formula also depends on the four quantities var_T, $a_{1,T}$, $a_{2,T}$ and $b_{2,T}$. Thus, equalizing call/put prices at maturity T is equivalent to equalizing these four quantities in both models, by adjusting $\bar{\kappa}, \bar{\theta}, \bar{\xi}$ and $\bar{\rho}$. Unfortunately, we have not found a closed expression for these equivalent parameters. An alternative and simpler way of proceeding consists of modifying the unobserved initial value $\bar{v}_0$ of the variance process while keeping $\bar{\kappa} = \kappa$. For non-vanishing correlation $(\rho_t)_t$, it leads to two possibilities

$$\begin{aligned} \bar{v_0} &= b\frac{(b \pm \sqrt{b^2 - 4ac})}{2a} - \frac{p_1 var_T}{m_1 p_0 - m_0 p_1}, & \bar{\theta} &= \frac{var_T - m_0 \bar{v}_0}{m_1}, \\ \bar{\xi} &= \sqrt{\frac{b_{0,T}}{r_0 \bar{v}_0 + r_1 \bar{\theta}}}, & \bar{\rho} &= -\frac{2a}{\bar{\xi}(b \pm \sqrt{b^2 - 4ac})}. \end{aligned}$$

[3] In this approach, we leave the initial value $\bar{v}_0$ equal to v_0. Indeed, it is not natural to modify its value since it is not a parameter, but rather an unobserved factor.

In practice, only one solution gives realistic parameters. However, this rule is heuristic since there is a priori no guarantee that these averaged parameters satisfy the assumption (P), which is the basis for the arguments correctness.

Proof. Using Proposition 10.2.2, one has to solve the following system of equations

$$var_T = m_0\bar{v}_0 + m_1\bar{\theta}, \qquad a_{1,T} = \bar{\rho}\bar{\xi}(p_0\bar{v}_0 + p_1\bar{\theta}),$$
$$a_{2,T} = (\bar{\rho}\bar{\xi})^2(q_0\bar{v}_0 + q_1\bar{\theta}), \qquad b_{0,T} = \bar{\xi}^2(r_0\bar{v}_0 + r_1\bar{\theta}).$$

The first equation gives $\bar{\theta} = \frac{var_T - m_0\bar{v}_0}{m_1}$. Replacing this identity in $a_{1,T}$ and $a_{2,T}$ gives

$$\bar{v}_0 = \left(\frac{a_{1,T}}{(\bar{\rho}\bar{\xi})} - \frac{p_1 var_T}{m_1}\right)\frac{m_1}{p_0 m_1 - p_1 m_0}, \qquad \bar{v}_0 = \left(\frac{a_{2,T}}{(\bar{\rho}\bar{\xi})^2} - \frac{q_1 var_T}{m_1}\right)\frac{m_1}{q_0 m_1 - q_1 m_0}.$$

It readily leads to a quadratic equation $ax^2 + bx + c = 0$ with $x = \frac{1}{\bar{\rho}\bar{\xi}}$. By solving this equation, we easily complete the proof of the result. □

Collinearity effect in the Heston model Another corollary of the approximation formula (10.17) is that we can obtain the same vanilla prices at time T with different sets of parameters. For instance, take

Table 10.1: Error in implied Black-Scholes volatilities (in bp) between the closed formulas (see appendix) of the two models M_1 and M_2 expressed as relative strikes. Maturity is equal to one year.

strikes K	80%	90%	100%	110%	120%
model M_1	20.12%	19.64%	19.50%	19.62%	19.92%
model M_2	20.11%	19.65%	19.51%	19.62%	19.92%
errors (bp)	0.69	-0.35	-0.81	-0.42	0.34

on the one hand $v_0 = \theta = 4\%$, $\kappa_1 = 2$ and $\xi_1 = 30\%$ (model M_1) and on the other hand $v_0 = \theta = 4\%$, $\kappa_2 = 3$ and $\xi_2 = 38.042\%$ (model M_2), both models having zero correlation. The resulting error between implied volatilities within the two models are presented in Table 10.1: they are so small that prices can be considered as equal. Actually, this kind of example is easy to create even with non-null correlation: as before, in view of the approximation formula (10.17), it is sufficient to equalize the four quantities var_T, $a_{1,T}$, $a_{2,T}$ and $b_{2,T}$.

As a consequence, calibrating a Heston model using options with a single maturity is an ill-posed problem, which is not a surprising fact.

10.3 Numerical accuracy of the approximation

We give numerical results of the performance of our method. In what follows, the spot S_0, the risk-free rate r and the dividend yield q are set respectively to 100, 0% and 0%. The initial value of the variance process is set to $v_0 = 4\%$ (initial volatility equal to 20%). Then we study the numerical accuracy w.r.t. K, T, κ, θ, ξ and ρ by testing different values for these parameters.

In order to present more interesting results for various relevant maturities and strikes, we allow the range of strikes to vary over the maturities. The strike values evolve approximately as $S_0\exp(c\sqrt{\theta T})$

Table 10.2: Set of maturities and strikes used for the numerical tests.

T/K								
3M	70	80	90	100	110	120	125	130
6M	60	70	80	100	110	130	140	150
1Y	50	60	80	100	120	150	170	180
2Y	40	50	70	100	130	180	210	240
3Y	30	40	60	100	140	200	250	290
5Y	20	30	60	100	150	250	320	400
7Y	10	30	50	100	170	300	410	520
10Y	10	20	50	100	190	370	550	730

for some real numbers c and $\theta = 6\%$. The extreme values of c are chosen to be equal to ± 2.57, which represents the 1%-99% quantile of the standard normal distribution. This corresponds to very out-of-the-money options or very deep-in-the-money options. The set of pairs (maturity, strike) chosen for the tests is given in Table 10.2.

Constant parameters In Table 10.3, we report the numerical results when $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$ and $\rho = 0\%$, giving the errors of implied Black-Scholes volatilities between our approximation formula (see Equation (10.17)) and the price calculated using the closed formula (see appendix), for the maturities and strikes of Table 10.2. The table should be read as follows: for example, for one year maturity and strike equal to 170, the implied volatility is equal to 24.14% using the closed formula and 24.20% with the approximation formula, giving an error of -6.33 bps. In Table 10.3, we observe that the errors do not exceed **7 bps** for a large range of strikes and maturities. We notice that the errors are surprisingly higher for short maturities. At first sight, it is counterintuitive as one would expect our perturbation method to work better for short maturities and worse for long maturities, since the difference between our proxy model (BS with volatility $(v_{0,t})_t$) and the original one is increasing w.r.t. time. In fact, this intuition is true for prices but not for implied volatilities. When we compare the price errors (in Price bp[4]) for the same data, we observe in Table 10.4 that the error terms are not any bigger for short maturities but vary slightly over time with two observed effects. The error term first increases over time as the error between the proxy and the original model increases over time, as forecasted. But for long maturities, presumably because the volatility converges to its stationary regime, errors decrease. When we convert these prices to implied Black-Scholes volatilities, these error terms are dramatically amplified for short maturities due to very small vega. Finally, note that for fixed maturity, price errors are quite uniform w.r.t. strike K.

Impact of the correlation Analogous results for correlations equal to -20%, 20% and -50% are reported in Tables 10.5-10.6, 10.7-10.8 and 10.9-10.10. We notice that the errors are smaller for a correlation close to zero and become larger when the absolute value of the correlation increases. However, for realistic correlation values (-50% for instance), the accuracy for the usual maturities and strikes remains excellent (error smaller than 20 bps), except for very extreme strikes.

Impact of the volatility of volatility In view of Theorem 10.2.2, the smaller the volatility of volatility, the more accurate the approximation. In the following numerical tests, we increase ξ, while maintaining

[4]Error price bp= $\frac{\text{Price Approximation}-\text{True Price}}{\text{Spot}} \times 10000$

Table 10.3: Implied Black-Scholes volatilities of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$ and $\rho = 0\%$.

3M	23.24%	22.14%	21.43%	21.19%	21.39%	21.86%	22.14%	22.44%
	23.06%	22.19%	21.42%	21.19%	21.38%	21.88%	22.19%	22.49%
	18.01	**-4.86**	**0.53**	**0.38**	**0.65**	**-2.68**	**-4.86**	**-4.71**
6M	24.32%	23.29%	22.55%	21.99%	22.10%	22.75%	23.17%	23.60%
	24.12%	23.36%	22.57%	21.98%	22.09%	22.79%	23.24%	23.65%
	19.69	**-7.17**	**-1.89**	**0.93**	**1.05**	**-3.97**	**-7.12**	**-4.57**
1Y	24.85%	24.06%	23.14%	22.90%	23.06%	23.66%	24.14%	24.38%
	24.78%	24.12%	23.14%	22.89%	23.06%	23.71%	24.20%	24.42%
	7.72	**-6.49**	**0.26**	**1.12**	**0.72**	**-4.54**	**-6.33**	**-4.27**
2Y	24.86%	24.36%	23.82%	23.61%	23.73%	24.16%	24.46%	24.76%
	24.86%	24.40%	23.82%	23.61%	23.72%	24.19%	24.50%	24.78%
	-0.21	**-3.51**	**-0.12**	**0.68**	**0.37**	**-2.54**	**-3.62**	**-1.71**
3Y	24.95%	24.53%	24.10%	23.89%	23.98%	24.27%	24.53%	24.74%
	24.94%	24.55%	24.10%	23.89%	23.98%	24.28%	24.55%	24.75%
	1.80	**-2.12**	**-0.33**	**0.39**	**0.19**	**-1.27**	**-2.12**	**-1.26**
5Y	24.88%	24.56%	24.20%	24.12%	24.17%	24.38%	24.53%	24.69%
	24.86%	24.57%	24.20%	24.12%	24.17%	24.39%	24.54%	24.70%
	1.38	**-0.96**	**0.03**	**0.17**	**0.10**	**-0.58**	**-0.95**	**-0.59**
7Y	25.03%	24.46%	24.30%	24.23%	24.27%	24.42%	24.54%	24.65%
	24.97%	24.46%	24.30%	24.22%	24.27%	24.42%	24.55%	24.66%
	5.72	**-0.43**	**-0.02**	**0.09**	**0.04**	**-0.33**	**-0.54**	**-0.35**
10Y	24.72%	24.51%	24.34%	24.30%	24.34%	24.44%	24.54%	24.62%
	24.71%	24.51%	24.34%	24.30%	24.34%	24.44%	24.54%	24.62%
	0.42	**-0.28**	**0.02**	**0.05**	**0.02**	**-0.17**	**-0.29**	**-0.19**

Assumption (P). Thus, the new Heston's parameters are $\kappa = 10$, $\xi = 1$ and $\rho = -50\%$, the other parameters remaining unchanged. The comparative results on implied volatilties and prices are presented in Table 10.11 and 10.12. As expected, the approximation is less accurate than for $\xi = 30\%$, but still accurate enough to be efficiently used for fast calibration. The results for prices are more satisfactory than for implied volatilities. Once again, for short maturities, the errors in implied volatilities may be quite significant, except for options not-far-from-the-money.

Impact of the assumption (P) The assumption (P) is a technical assumption that we use to establish error estimates for the approximation formula (10.17). In the test that follows, we relax this assumption by taking new parameters $\theta = 3\%$, $\kappa = 2$, $\xi = 40\%$ and $\rho = 0\%$ for which the ratio $2\kappa\theta/\xi^2 = 0.75 < 1$. Results are reported in Tables 10.13 and 10.14. We observe that the approximation formula still works (errors are smaller than 20 bps) but it is less accurate (compare with Table 10.3 for which the ratio $2\kappa\theta/\xi^2$ is equal to $4 > 1$). An extension of the validity of our formula by relaxing Assumption (P) is presumably relevant. This investigation is left for further research.

Piecewise constant parameters Heston's constant parameters have been set to: $v_0 = 4\%, \kappa = 3$. In addition, the piecewise constant functions θ, ξ and ρ are equal respectively at each interval of the form $]\frac{i}{4}, \frac{i+1}{4}[$ to $4\% + i \times 0.05\%$, $30\% + i \times 0.5\%$ and $-20\% + i \times 0.35\%$.
In the same Tables 10.16 and 10.17, we report values using three different formulas. For a given maturity,

Table 10.4: Put prices of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$ and $\rho = 0\%$.

3M	30.00	20.08	10.87	4.22	1.14	0.24	0.10	0.04
	30.00	20.08	10.87	4.22	1.14	0.24	0.10	0.04
	0.03	**-0.11**	**0.06**	**0.08**	**0.09**	**-0.15**	**-0.14**	**-0.07**
6M	40.01	30.07	20.52	6.20	2.72	0.40	0.14	0.05
	40.01	30.08	20.52	6.19	2.71	0.40	0.14	0.05
	0.05	**-0.16**	**-0.18**	**0.26**	**0.26**	**-0.34**	**-0.29**	**-0.08**
1Y	50.01	40.11	21.84	9.12	3.08	0.51	0.15	0.09
	50.01	40.11	21.84	9.11	3.07	0.52	0.16	0.09
	0.04	**-0.21**	**0.06**	**0.44**	**0.23**	**-0.51**	**-0.29**	**-0.12**
2Y	60.03	50.20	32.08	13.26	4.71	0.79	0.28	0.11
	60.03	50.20	32.08	13.26	4.71	0.79	0.29	0.11
	0.00	**-0.18**	**-0.03**	**0.38**	**0.17**	**-0.43**	**-0.29**	**-0.06**
3Y	70.02	60.15	41.70	16.39	5.73	1.21	0.36	0.15
	70.02	60.15	41.70	16.39	5.73	1.21	0.37	0.15
	0.01	**-0.09**	**-0.08**	**0.27**	**0.11**	**-0.31**	**-0.22**	**-0.07**
5Y	80.01	70.15	43.80	21.26	8.50	1.61	0.58	0.21
	80.01	70.15	43.80	21.26	8.50	1.61	0.58	0.21
	0.01	**-0.04**	**0.01**	**0.15**	**0.08**	**-0.19**	**-0.15**	**-0.04**
7Y	90.00	70.42	53.15	25.14	9.32	1.97	0.66	0.26
	90.00	70.42	53.15	25.14	9.32	1.97	0.67	0.26
	0.00	**-0.04**	**-0.01**	**0.09**	**0.04**	**-0.14**	**-0.10**	**-0.03**
10Y	90.01	80.23	55.22	29.92	11.49	2.62	0.84	0.33
	90.01	80.23	55.22	29.92	11.49	2.62	0.84	0.33
	0.00	**-0.02**	**0.01**	**0.06**	**0.03**	**-0.09**	**-0.07**	**-0.02**

the first row is obtained using the closed formula with piecewise constant parameters (see appendix), the second row uses our approximation formula (10.17) and the third row uses the closed formula with constant parameters computed by averaging (see Section 10.2.6). In order to give complete information on our tests, we also report in Table 10.15 the values used for the averaging parameters (following Section 10.2.6).

Of course, the quickest approach is the use of the approximation formula (10.17). As before, its accuracy is very good, except for very extreme strikes. It is quite interesting is to observe that the averaging rules that we propose are extremely accurate.

Computational time Regarding the computational time, the approximation formula (10.17) yields essentially the same computational cost as the Black-Scholes formula, while the closed formula requires an additional space integration involving many exponential and trigonometric functions for which evaluation costs are higher. For instance, using a 2,6 GHz Pentium PC, the computations of the 64 numerical values in Table 10.3 (10.5, 10.7 or 10.9) take 4.71 ms using the approximation formula and 301ms using the closed formula. For the example with time dependent coefficients (reported in Table 10.16), the computational time for the 64 prices is about 40.2 ms using the approximation formula and 2574 ms using the closed formula. Roughly speaking, the use of the approximation formula enables us to speed up the valuation (and thus the calibration) by a factor 100 to 600.

Table 10.5: Implied Black-Scholes volatilities of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$ and $\rho = -20\%$.

3M	24.50%	23.07%	21.92%	21.16%	20.84%	20.91%	21.04%	21.21%
	24.04%	23.14%	21.93%	21.15%	20.82%	20.87%	21.06%	21.37%
	45.76	**-7.65**	**-1.25**	**0.38**	**2.35**	**3.68**	**-2.73**	**-16.51**
6M	25.68%	24.38%	23.31%	21.94%	21.65%	21.68%	21.88%	22.15%
	25.19%	24.45%	23.38%	21.93%	21.63%	21.64%	21.96%	22.47%
	49.49	**-7.75**	**-7.32**	**0.99**	**2.22**	**4.10**	**-8.10**	**-32.52**
1Y	26.20%	25.14%	23.65%	22.82%	22.47%	22.51%	22.72%	22.86%
	25.92%	25.23%	23.68%	22.81%	22.44%	22.49%	22.89%	23.17%
	28.04	**-8.22**	**-2.65**	**1.32**	**3.45**	**2.08**	**-16.41**	**-31.56**
2Y	26.03%	25.28%	24.29%	23.51%	23.18%	23.09%	23.17%	23.29%
	25.95%	25.35%	24.32%	23.50%	23.16%	23.08%	23.25%	23.56%
	7.83	**-6.41**	**-2.54**	**0.93**	**2.37**	**1.57**	**-8.04**	**-26.37**
3Y	26.06%	25.40%	24.57%	23.78%	23.47%	23.34%	23.36%	23.42%
	25.95%	25.44%	24.60%	23.78%	23.45%	23.32%	23.41%	23.58%
	11.21	**-3.39**	**-2.44**	**0.61**	**1.65**	**1.71**	**-5.11**	**-16.68**
5Y	25.83%	25.28%	24.47%	24.01%	23.75%	23.57%	23.55%	23.55%
	25.75%	25.30%	24.47%	24.01%	23.74%	23.56%	23.56%	23.65%
	8.29	**-1.76**	**-0.65**	**0.32**	**0.84**	**1.01**	**-1.92**	**-9.38**
7Y	26.02%	24.97%	24.56%	24.11%	23.86%	23.70%	23.65%	23.64%
	25.82%	24.99%	24.57%	24.11%	23.85%	23.69%	23.67%	23.70%
	20.23	**-1.59**	**-0.59**	**0.21**	**0.60**	**0.69**	**-1.50**	**-6.16**
10Y	25.43%	24.99%	24.49%	24.19%	23.97%	23.81%	23.75%	23.72%
	25.40%	25.00%	24.49%	24.18%	23.96%	23.80%	23.76%	23.76%
	3.46	**-0.94**	**-0.20**	**0.14**	**0.38**	**0.48**	**-0.95**	**-3.98**

10.4 Proof of Theorem 10.2.2

The proof is divided into several steps. In Subsection 10.4.1 we give the upper bounds for derivatives of the put function P_{BS}, in Subsection 10.4.2 the conditions for positivity of the squared volatility process v, in Subsection 10.4.3 the upper bounds for the negative moments of the integrated squared volatility $\int_0^T v_t \mathrm{d}t$, in Subsection 10.4.4 the upper bounds for derivatives of functionals of the squared volatility process v. Finally, in Subsection 10.4.5, we complete the proof of Theorem 10.2.2 using the previous Subsections.

Notations In order to alleviate the proofs, we introduce some notations specific to this section.
Differentiation. For every process Z^ε, we write (if these derivatives have a meaning):

- $Z_{i,t} = \frac{\partial^i Z_t^\varepsilon}{\partial \varepsilon^i}|_{\varepsilon=0}$,
- the i^{th} Taylor residual by $R_{i,t}^{Z^\varepsilon} = Z_t^\varepsilon - \sum_{j=0}^{i} \frac{\varepsilon^j}{j!} Z_{j,t}$.

Generic constants. We keep the same notation C for all non-negative constants

- depending on universal constants, on a number $p \geq 1$ arising in L_p estimates, on θ_{Inf}, v_0 and K;
- depending in a non decreasing way on κ, $\frac{1}{\sqrt{1-|\rho|_{Sup}^2}}$, θ_{Sup}, ξ_{Sup}, $\frac{\xi_{Sup}}{\xi_{Inf}}$ and T.

Table 10.6: Put prices of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$ and $\rho = -20\%$.

3M	30.01	20.10	10.93	4.22	1.07	0.19	0.07	0.03
	30.00	20.11	10.93	4.22	1.06	0.19	0.07	0.03
	0.10	**-0.21**	**-0.15**	**0.08**	**0.32**	**0.18**	**-0.06**	**-0.18**
6M	40.01	30.10	20.60	6.18	2.61	0.31	0.10	0.03
	40.01	30.10	20.60	6.18	2.60	0.31	0.10	0.03
	0.19	**-0.22**	**-0.74**	**0.28**	**0.54**	**0.30**	**-0.26**	**-0.41**
1Y	50.02	40.15	21.95	9.08	2.89	0.39	0.10	0.05
	50.02	40.15	21.96	9.08	2.88	0.39	0.10	0.06
	0.23	**-0.32**	**-0.60**	**0.52**	**1.08**	**0.20**	**-0.57**	**-0.64**
2Y	60.05	50.25	32.21	13.21	4.46	0.62	0.19	0.06
	60.05	50.25	32.21	13.20	4.44	0.62	0.20	0.07
	0.12	**-0.39**	**-0.69**	**0.52**	**1.09**	**0.23**	**-0.50**	**-0.69**
3Y	70.03	60.19	41.82	16.32	5.44	0.99	0.26	0.09
	70.03	60.19	41.83	16.31	5.43	0.99	0.26	0.10
	0.12	**-0.17**	**-0.62**	**0.41**	**0.94**	**0.38**	**-0.42**	**-0.63**
5Y	80.02	70.18	43.91	21.16	8.17	1.35	0.44	0.14
	80.02	70.18	43.91	21.16	8.16	1.35	0.44	0.14
	0.06	**-0.09**	**-0.28**	**0.28**	**0.66**	**0.30**	**-0.26**	**-0.51**
7Y	90.00	70.47	53.25	25.02	8.94	1.68	0.51	0.18
	90.00	70.47	53.26	25.02	8.93	1.68	0.51	0.18
	0.02	**-0.17**	**-0.24**	**0.21**	**0.55**	**0.26**	**-0.24**	**-0.44**
10Y	90.01	80.26	55.30	29.78	11.07	2.29	0.66	0.24
	90.01	80.26	55.30	29.78	11.06	2.29	0.67	0.24
	0.02	**-0.06**	**-0.11**	**0.16**	**0.43**	**0.24**	**-0.20**	**-0.38**

We write $A = O(B)$ when $|A| \leq CB$ for a generic constant.
Miscellaneous.

- We write $\sigma_t^\varepsilon = \sqrt{v_t^\varepsilon}$ for the volatility for the perturbed process.
- if $(Z)_{t\in[0,T]}$ is a càdlàg process, we denote by Z^* its running extremum: $Z_t^* = \sup_{s\leq t} |Z_s|, \forall t \in [0,T]$.
- The L_p norm of a random variable is denoted, as usual, by $\|Z\|_p = \mathbb{E}[|Z|^p]^{1/p}$.

10.4.1 Upper bounds for put derivatives

Lemma 10.4.1. *For every $(i,j) \in \mathbb{N}^2$, there exists a polynomial P with positive coefficients such that:*

$$\sup_{x\in\mathbb{R}} \left| \frac{\partial^{i+j} P_{BS}}{\partial x^i y^j}(x,y) \right| \leq \frac{P(\sqrt{y})}{y^{\frac{(2j+i-1)_+}{2}}}.$$

Proof. Note that it is enough to prove the estimates for $j = 0$, owing to the relation (10.6). We now take $j = 0$. For $i = 0$, the inequality holds because P_{BS} is bounded. Thus consider $i \geq 1$. Then by

Table 10.7: Implied Black-Scholes volatilities of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$ and $\rho = 20\%$.

3M	21.81%	21.10%	20.89%	21.22%	21.89%	22.71%	23.13%	23.54%
	22.41%	21.11%	20.87%	21.22%	21.90%	22.78%	23.20%	23.55%
	-59.86	**-1.80**	**2.68**	**0.27**	**-0.82**	**-7.12**	**-7.19**	**-1.20**
6M	22.75%	22.05%	21.72%	22.04%	22.53%	23.71%	24.31%	24.47%
	23.41%	22.16%	21.66%	22.03%	22.53%	23.81%	24.40%	24.45%
	-66.39	**-10.95**	**5.61**	**0.72**	**-0.08**	**-9.75**	**-8.77**	**2.21**
1Y	23.31%	22.83%	22.59%	22.97%	23.62%	24.72%	25.41%	24.80%
	23.83%	22.91%	22.55%	22.96%	23.64%	24.82%	25.46%	24.81%
	-52.67	**-8.05**	**3.84**	**0.88**	**-1.65**	**-9.85**	**-4.37**	**-1.19**
2Y	23.53%	23.33%	23.31%	23.70%	24.25%	25.16%	25.65%	24.93%
	23.77%	23.34%	23.28%	23.70%	24.27%	25.22%	25.68%	24.93%
	-23.90	**-1.04**	**2.80**	**0.47**	**-1.42**	**-6.19**	**-3.19**	**-0.67**
3Y	23.70%	23.56%	23.58%	23.99%	24.48%	25.15%	25.63%	24.83%
	23.93%	23.58%	23.56%	23.99%	24.49%	25.19%	25.64%	24.84%
	-23.06	**-1.95**	**2.19**	**0.22**	**-1.15**	**-3.93**	**-1.70**	**-0.92**
5Y	23.81%	23.76%	23.92%	24.23%	24.59%	25.15%	25.46%	24.74%
	23.96%	23.77%	23.91%	24.23%	24.59%	25.17%	25.47%	24.74%
	-14.87	**-0.62**	**0.82**	**0.04**	**-0.59**	**-2.06**	**-0.94**	**-0.49**
7Y	23.92%	23.90%	24.03%	24.34%	24.68%	25.12%	25.39%	24.68%
	24.21%	23.89%	24.02%	24.34%	24.68%	25.13%	25.39%	24.68%
	-28.79	**0.90**	**0.63**	**-0.01**	**-0.48**	**-1.30**	**-0.30**	**-0.30**
10Y	23.94%	23.99%	24.18%	24.42%	24.70%	25.06%	25.29%	24.63%
	23.99%	23.99%	24.18%	24.42%	24.71%	25.07%	25.29%	24.64%
	-5.60	**0.42**	**0.26**	**-0.03**	**-0.32**	**-0.79**	**-0.07**	**-0.17**

differentiating the payoff, one gets:

$$\begin{aligned}\frac{\partial^i P_{BS}}{\partial x^i}(x,y) &= \partial_x^i \mathbb{E}[e^{-\int_0^T r_t \mathrm{d}t}(K - e^{x+\int_0^T (r_t-q_t)\mathrm{d}t - \frac{y}{2} + \sqrt{\frac{y}{T}}W_T})_+] \\ &= -\partial_x^{i-1}\mathbb{E}[\mathbb{1}_{(e^{x+\int_0^T (r_t-q_t)\mathrm{d}t - \frac{y}{2} + \sqrt{\frac{y}{T}}W_T} \le K)} e^{x-\int_0^T q_t \mathrm{d}t - \frac{y}{2} + \sqrt{\frac{y}{T}}W_T}] \\ &= -\partial_x^{i-1}\mathbb{E}[\Psi(x+G)]\end{aligned}$$

where Ψ is a bounded function (by K) and G is a Gaussian variable with zero mean and variance equal to y. For such a function, we write $\mathbb{E}[\Psi(x+G)] = \int_{\mathbb{R}} \Psi(z)\frac{e^{-(z-x)^2/(2y)}}{\sqrt{2\pi y}}\mathrm{d}z$ and from this, it follows by a direct computation that

$$\left|\partial_x^{i-1}\mathbb{E}[\Psi(x+G)]\right| \le \frac{C}{y^{\frac{i-1}{2}}}$$

for any x and y. We have proved the estimate for $j = 0$ and $i \ge 1$. □

10.4.2 Positivity of the squared volatility process v

For a complete review related to time homogeneous CIR processes, we refer the reader to [56]. For time dependent CIR process, see [82] where the existence and representation using squared Bessel processes are provided.

Table 10.8: Put prices of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$ and $\rho = 20\%$.

3M	30.00	20.06	10.81	4.23	1.22	0.28	0.13	0.06
	30.00	20.06	10.81	4.23	1.22	0.29	0.13	0.06
	-0.05	**-0.03**	**0.30**	**0.05**	**-0.12**	**-0.43**	**-0.25**	**-0.02**
6M	40.00	30.05	20.45	6.21	2.82	0.48	0.19	0.07
	40.00	30.05	20.44	6.21	2.82	0.49	0.20	0.07
	-0.11	**-0.19**	**0.49**	**0.20**	**-0.02**	**-0.92**	**-0.43**	**0.05**
1Y	50.01	40.08	21.72	9.14	3.26	0.64	0.22	0.10
	50.01	40.08	21.71	9.14	3.26	0.65	0.22	0.10
	-0.20	**-0.20**	**0.83**	**0.35**	**-0.53**	**-1.25**	**-0.26**	**-0.04**
2Y	60.02	50.15	31.94	13.31	4.96	0.97	0.39	0.11
	60.02	50.15	31.94	13.31	4.96	0.98	0.39	0.11
	-0.20	**-0.04**	**0.73**	**0.26**	**-0.67**	**-1.18**	**-0.32**	**-0.03**
3Y	70.01	60.11	41.58	16.46	6.02	1.44	0.49	0.16
	70.01	60.11	41.57	16.46	6.03	1.45	0.49	0.16
	-0.12	**-0.07**	**0.53**	**0.15**	**-0.67**	**-1.06**	**-0.21**	**-0.05**
5Y	80.01	70.11	43.68	21.36	8.83	1.87	0.75	0.21
	80.01	70.11	43.67	21.35	8.84	1.88	0.75	0.21
	-0.06	**-0.02**	**0.35**	**0.03**	**-0.48**	**-0.74**	**-0.18**	**-0.04**
7Y	90.00	70.36	53.04	25.25	9.69	2.26	0.84	0.26
	90.00	70.36	53.04	25.25	9.70	2.27	0.84	0.26
	-0.01	**0.08**	**0.25**	**-0.01**	**-0.45**	**-0.57**	**-0.07**	**-0.03**
10Y	90.01	80.20	55.13	30.06	11.91	2.96	1.04	0.34
	90.01	80.19	55.13	30.06	11.91	2.96	1.04	0.34
	-0.02	**0.02**	**0.15**	**-0.04**	**-0.37**	**-0.45**	**-0.02**	**-0.02**

To prove the positivity of the process v, we show that it can be bounded from below by a suitable time homogeneous CIR process, time scale being the only difference (see definition 5.1.2. in [98]). The arguments are quite standard, but since we need a specific statement that is not available in the literature, we detail the result and its proof. The time change $t \mapsto A_t$ is defined by

$$t = \int_0^{A_t} \xi_s^2 \mathrm{d}s.$$

Because $\xi_{Inf} > 0$, A is a continuous, strictly increasing time change and its inverse A^{-1} enjoys the same properties.

Lemma 10.4.2. *Assume (P) and $v_0 > 0$. Denote by $(y_s)_{0 \le s \le A_T^{-1}}$ the CIR process defined by*

$$\mathrm{d}y_t = (\frac{1}{2} - \frac{\kappa}{\xi_{Inf}^2} y_t)\mathrm{d}t + \sqrt{y_t} d\tilde{B}_t,\ y_0 = v_0,$$

where $\tilde{B}$ is the Brownian motion given by

$$\tilde{B}_t = \int_0^{A_T} \xi_s \mathrm{d}B_s. \tag{10.18}$$

Then, a.s. one has $v_t \ge y_{A_t^{-1}}$ for any $t \in [0,T]$. In particular, $(v_t)_{0 \le t \le T}$ is a.s. positive.

Table 10.9: Implied Black-Scholes volatilities of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$ and $\rho = -50\%$.

3M	26.13%	24.29%	22.60%	21.11%	19.95%	19.22%	19.03%	18.92%
	25.57%	24.43%	22.63%	21.11%	19.90%	18.99%	18.91%	19.57%
	56.55	**-14.06**	**-2.51**	**0.19**	**4.35**	**23.24**	**11.67**	**-64.22**
6M	27.47%	25.81%	24.31%	21.85%	20.92%	19.80%	19.55%	19.47%
	26.89%	25.97%	24.44%	21.84%	20.89%	19.50%	19.61%	21.11%
	58.13	**-16.68**	**-12.19**	**0.82**	**3.38**	**29.46**	**-5.28**	**-164.16**
1Y	27.96%	26.57%	24.34%	22.68%	21.51%	20.49%	20.19%	20.11%
	27.67%	26.75%	24.39%	22.66%	21.43%	20.24%	20.77%	21.73%
	29.08	**-18.08**	**-5.01**	**1.53**	**7.49**	**24.84**	**-58.18**	**-162.76**
2Y	27.56%	26.51%	24.93%	23.34%	22.31%	21.30%	20.95%	20.73%
	27.52%	26.65%	24.98%	23.33%	22.25%	21.15%	21.19%	22.20%
	4.11	**-14.03**	**-4.75**	**1.43**	**5.50**	**14.43**	**-23.17**	**-146.81**
3Y	27.53%	26.56%	25.22%	23.61%	22.66%	21.81%	21.39%	21.16%
	27.42%	26.66%	25.26%	23.60%	22.62%	21.71%	21.53%	22.04%
	11.28	**-9.11**	**-4.59**	**1.06**	**3.97**	**9.79**	**-14.43**	**-88.86**
5Y	27.11%	26.25%	24.83%	23.83%	23.10%	22.28%	21.94%	21.66%
	27.01%	26.31%	24.84%	23.82%	23.08%	22.23%	21.98%	22.14%
	9.64	**-5.22**	**-1.23**	**0.62**	**1.98**	**5.14**	**-4.04**	**-47.56**
7Y	27.35%	25.67%	24.92%	23.93%	23.23%	22.55%	22.22%	21.98%
	27.03%	25.71%	24.93%	23.93%	23.21%	22.52%	22.25%	22.28%
	31.65	**-3.57**	**-1.09**	**0.43**	**1.46**	**3.26**	**-3.91**	**-30.07**
10Y	26.40%	25.66%	24.70%	24.01%	23.40%	22.82%	22.50%	22.29%
	26.36%	25.68%	24.70%	24.00%	23.39%	22.80%	22.53%	22.48%
	4.15	**-2.43**	**-0.35**	**0.29**	**0.93**	**2.02**	**-2.65**	**-18.89**

Proof. Note that $(\tilde{B}_t)_{0 \le t \le A_T^{-1}}$ is really a Brownian motion because by Lévy's Characterization Theorem, it is a continuous local martingale with $\langle \tilde{B}, \tilde{B} \rangle_t = t$ (see Proposition 5.1.5 [98] for the computation of the bracket). Now that we have set $\tilde{v}_t = v_{A_t}$, our aim is to prove that $\tilde{v}_t \ge y_t$ for $t \in [0, A_T^{-1}]$. Using Propositions 5.1.4 and 5.1.5 [98], we write

$$\tilde{v}_t = v_0 + \int_0^{A_t} (\kappa(\theta_s - v_s)\mathrm{d}s + \xi_s \sqrt{v_s}\mathrm{d}B_s) = v_0 + \int_0^t \left(\frac{\kappa}{\xi_{A_s}^2}(\theta_{A_s} - \tilde{v}_s)\mathrm{d}s + \sqrt{\tilde{v}_s}\mathrm{d}\tilde{B}_s\right).$$

Now we apply a comparison result for SDEs twice (see Proposition 5.2.18 in [69]).

1. First, one gets $\tilde{v}_t \ge n_t$, where $(n_s)_s$ is the (unique) solution of

   $$n_t = 0 + \int_0^t -\frac{\kappa}{\xi_{A_s}^2} n_s \mathrm{d}s + \sqrt{n_s}\mathrm{d}\tilde{B}_s,$$

   because $v_0 \ge 0$ and $\frac{\kappa}{\xi_{A_s}^2}(\theta_{A_s} - x) \ge -\frac{\kappa}{\xi_{A_s}^2} x$, for all $x \in \mathbb{R}$ and $s \in [0, A_T^{-1}]$. Of course $n_t = 0$, thus $\tilde{v}_t$ is non-negative.

Table 10.10: Put prices of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$ and $\rho = -50\%$.

3M	30.01	20.14	11.01	4.21	0.95	0.12	0.03	0.01
	30.01	20.15	11.02	4.21	0.94	0.11	0.03	0.01
	0.21	**-0.47**	**-0.31**	**0.04**	**0.57**	**0.82**	**0.16**	**-0.36**
6M	40.02	30.15	20.70	6.16	2.43	0.19	0.04	0.01
	40.02	30.15	20.71	6.15	2.42	0.17	0.04	0.02
	0.37	**-0.59**	**-1.33**	**0.23**	**0.81**	**1.59**	**-0.09**	**-1.05**
1Y	50.04	40.21	22.11	9.03	2.59	0.22	0.03	0.01
	50.04	40.22	22.12	9.02	2.57	0.21	0.05	0.03
	0.36	**-0.88**	**-1.17**	**0.61**	**2.27**	**1.67**	**-1.05**	**-1.69**
2Y	60.08	50.33	32.38	13.11	4.06	0.39	0.08	0.02
	60.08	50.34	32.39	13.10	4.03	0.37	0.09	0.04
	0.09	**-1.00**	**-1.32**	**0.80**	**2.47**	**1.59**	**-0.84**	**-2.00**
3Y	70.05	60.25	41.99	16.20	4.98	0.69	0.13	0.03
	70.05	60.25	42.00	16.19	4.96	0.67	0.13	0.05
	0.17	**-0.54**	**-1.21**	**0.72**	**2.20**	**1.73**	**-0.74**	**-1.80**
5Y	80.03	70.23	44.06	21.01	7.65	0.99	0.25	0.06
	80.03	70.23	44.07	21.00	7.64	0.98	0.26	0.07
	0.11	**-0.30**	**-0.53**	**0.54**	**1.54**	**1.29**	**-0.38**	**-1.50**
7Y	90.00	70.54	53.40	24.84	8.36	1.28	0.31	0.09
	90.00	70.55	53.40	24.84	8.35	1.27	0.32	0.10
	0.06	**-0.41**	**-0.44**	**0.43**	**1.32**	**1.04**	**-0.45**	**-1.32**
10Y	90.02	80.30	55.42	29.57	10.43	1.82	0.44	0.13
	90.02	80.30	55.42	29.57	10.42	1.81	0.44	0.14
	0.03	**-0.18**	**-0.20**	**0.34**	**1.04**	**0.89**	**-0.42**	**-1.17**

2. Secondly, using the non-negativity of $\tilde{v}$, we only need to compare drift coefficients for the non-negative variable x. Under (P), since

$$\frac{\kappa}{\xi_{A_s}^2}(\theta_{A_s} - x) \geq \frac{1}{2} - \frac{\kappa}{\xi_{Inf}^2}x \quad \forall x \geq 0, \forall s \in [0, A_T^{-1}],$$

we obtain $\tilde{v}_t \geq y_t$ for $t \in [0, A_T^{-1}]$ a.s.

Moreover, the positivity of y (and consequently that of v) is standard: indeed, y is a 2-dimensional squared Bessel process with a time/space scale change (see [56], or the proof of Lemma 10.4.3 below).
□

10.4.3 Upper bound for negative moments of the integrated squared volatility process $\int_0^T v_t \mathrm{d}t$

Lemma 10.4.3. *Assume (P). Then for every $p > 0$, one has:*

$$\sup_{0 \leq \varepsilon \leq 1} \mathbb{E}[(\int_0^T v_t^\varepsilon \mathrm{d}t)^{-p}] \leq \frac{C}{T^p}.$$

Table 10.11: Implied Black-Scholes volatilities of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 10$, $\xi = 1$ and $\rho = -50\%$.

3M	31.51%	28.04%	24.74%	21.83%	19.94%	19.45%	19.58%	19.85%
	30.68%	28.99%	24.95%	21.71%	19.38%	18.05%	19.76%	22.93%
	82.46	**-94.66**	**-21.22**	**12.10**	**56.44**	**140.23**	**-18.10**	**-308.17**
6M	31.45%	28.86%	26.52%	22.69%	21.36%	20.11%	20.05%	20.20%
	30.83%	29.59%	26.98%	22.58%	21.09%	19.14%	20.64%	24.03%
	62.40	**-73.58**	**-46.52**	**11.30**	**26.99**	**97.22**	**-59.12**	**-383.12**
1Y	30.09%	28.30%	25.44%	23.34%	21.89%	20.76%	20.49%	20.45%
	29.87%	28.72%	25.54%	23.28%	21.70%	20.30%	21.65%	23.17%
	21.52	**-42.32**	**-10.69**	**6.02**	**19.45**	**46.13**	**-115.72**	**-271.22**
2Y	28.45%	27.27%	25.51%	23.73%	22.58%	21.48%	21.12%	20.90%
	28.46%	27.47%	25.57%	23.71%	22.50%	21.28%	21.42%	22.75%
	-0.53	**-20.08**	**-6.39**	**2.42**	**8.11**	**19.97**	**-30.34**	**-184.76**
3Y	28.08%	27.05%	25.61%	23.88%	22.86%	21.96%	21.51%	21.27%
	27.98%	27.16%	25.66%	23.86%	22.81%	21.83%	21.67%	22.30%
	9.78	**-11.59**	**-5.41**	**1.39**	**4.91**	**12.13**	**-16.04**	**-102.46**
5Y	27.40%	26.52%	25.04%	24.00%	23.23%	22.38%	22.03%	21.75%
	27.31%	26.58%	25.05%	23.99%	23.21%	22.33%	22.07%	22.26%
	9.15	**-5.98**	**-1.31**	**0.71**	**2.20**	**5.85**	**-3.93**	**-51.20**
7Y	27.56%	25.84%	25.06%	24.05%	23.33%	22.63%	22.29%	22.05%
	27.24%	25.88%	25.08%	24.05%	23.31%	22.59%	22.33%	22.36%
	32.00	**-3.83**	**-1.14**	**0.47**	**1.57**	**3.57**	**-3.88**	**-31.56**
10Y	26.53%	25.77%	24.80%	24.09%	23.47%	22.88%	22.55%	22.34%
	26.49%	25.80%	24.80%	24.09%	23.46%	22.86%	22.58%	22.53%
	4.02	**-2.57**	**-0.36**	**0.31**	**0.97**	**2.15**	**-2.64**	**-19.49**

Before proving the result, we mention that analogous estimates appear in [**?**] (Lemmas A.1 and A.2): some exponential moments are stated under stronger conditions than those in assumption (P). In addition, the uniformity of the estimates w.r.t. ξ (or equivalently w.r.t. ε) is not emphasized. In our study, it is crucial to get uniform estimates w.r.t. ε.

Proof. Fix $p \geq \frac{1}{2}$ (for $0 < p < \frac{1}{2}$, we derive the result from the case $p = \frac{1}{2}$ using the Hölder inequality). The proof is divided into two steps. We first prove the estimates in the case of constant coefficients κ, θ, ξ with $\kappa\theta = \frac{1}{2}$, $\varepsilon = 1$ and $\xi = 1$. Then, using the time change of Lemma 10.4.2, we derive the result for $(v_t^\varepsilon)_t$. The critical point is to get estimates that are uniform w.r.t. ε.
Step 1. Take $\theta_t \equiv \theta$, $\xi_t \equiv 1$, $\kappa\theta = \frac{1}{2}$, $\varepsilon = 1$ and consider

$$\mathrm{d}y_t = (\frac{1}{2} - \kappa y_t)\mathrm{d}t + \sqrt{y_t}\mathrm{d}B_t,\ y_0 = v_0,$$

for a standard Brownian motion B. We represent y as a time space transformed squared Bessel process

Table 10.12: Put prices of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 6\%$, $\kappa = 10$, $\xi = 1$ and $\rho = -50\%$.

3M	30.05	20.30	11.28	4.35	0.95	0.13	0.04	0.01
	30.04	20.35	11.31	4.33	0.87	0.08	0.05	0.05
	0.99	**-4.95**	**-2.80**	**2.41**	**7.37**	**4.62**	**-0.31**	**-3.51**
6M	40.06	30.28	20.96	6.40	2.54	0.21	0.05	0.01
	40.05	30.32	21.02	6.36	2.47	0.15	0.06	0.06
	0.92	**-3.90**	**-5.83**	**3.18**	**6.51**	**5.23**	**-1.28**	**-4.72**
1Y	50.08	40.31	22.37	9.29	2.71	0.24	0.04	0.02
	50.07	40.33	22.40	9.26	2.65	0.21	0.07	0.06
	0.41	**-2.60**	**-2.58**	**2.39**	**5.95**	**3.19**	**-2.52**	**-3.89**
2Y	60.10	50.39	32.54	13.33	4.18	0.41	0.09	0.02
	60.10	50.40	32.56	13.31	4.14	0.38	0.10	0.05
	-0.01	**-1.58**	**-1.82**	**1.35**	**3.67**	**2.26**	**-1.17**	**-2.89**
3Y	70.06	60.28	42.09	16.38	5.09	0.71	0.13	0.03
	70.05	60.28	42.11	16.37	5.06	0.69	0.14	0.06
	0.17	**-0.73**	**-1.46**	**0.94**	**2.74**	**2.19**	**-0.86**	**-2.22**
5Y	80.04	70.25	44.15	21.15	7.76	1.02	0.26	0.06
	80.03	70.25	44.16	21.15	7.74	1.01	0.27	0.08
	0.11	**-0.36**	**-0.57**	**0.61**	**1.72**	**1.49**	**-0.38**	**-1.68**
7Y	90.00	70.56	53.46	24.96	8.45	1.31	0.32	0.09
	90.00	70.57	53.46	24.96	8.44	1.30	0.32	0.10
	0.06	**-0.44**	**-0.47**	**0.47**	**1.42**	**1.16**	**-0.46**	**-1.42**
10Y	90.02	80.31	55.47	29.67	10.51	1.85	0.44	0.13
	90.02	80.31	55.48	29.67	10.50	1.84	0.45	0.14
	0.03	**-0.19**	**-0.20**	**0.36**	**1.09**	**0.95**	**-0.42**	**-1.23**

(see [56])

$$y_t = e^{-\kappa t} z_{\frac{(e^{\kappa t}-1)}{4\kappa}}$$

where z is a 2-dimensional squared Bessel process. Therefore, using a change of variable and the explicit expression of Laplace transform for the integral of z (see [23] p.377), one obtains for any $u \geq 0$

$$\mathbb{E}[\exp(-u\int_0^T y_t \mathrm{d}t)] \leq \mathbb{E}[\exp(-4ue^{-2\kappa T}\int_0^{\frac{(e^{\kappa T}-1)}{4\kappa}} z_s \mathrm{d}s)]$$
$$\leq \cosh(\frac{\sqrt{2u}(1-e^{-\kappa T})}{2\kappa})^{-1}\exp(-\sqrt{2u}e^{-\kappa T}v_0\tanh(\frac{\sqrt{2u}(1-e^{-\kappa T})}{2\kappa})).$$

Combining this with the identity $x^{-p} = \frac{1}{\Gamma(p)}\int_0^\infty u^{p-1}e^{-ux}\mathrm{d}u$ for $x = \int_0^T y_t \mathrm{d}t$, one gets:

$$\mathbb{E}[(\int_0^T y_t \mathrm{d}t)^{-p}] \leq \frac{1}{\Gamma(p)}\int_0^\infty u^{p-1}\cosh(\frac{\sqrt{2u}(1-e^{-\kappa T})}{2\kappa})^{-1}\exp(-\sqrt{2u}e^{-\kappa T}v_0\tanh(\frac{\sqrt{2u}(1-e^{-\kappa T})}{2\kappa}))\mathrm{d}u.$$

Define the parameter $\lambda^2 = \frac{(e^{\kappa T}-1)}{2\kappa v_0}$ and the new variable $n = \frac{\sqrt{2u}(1-e^{-\kappa T})}{2\kappa} = v_0 e^{-\kappa T}\lambda^2\sqrt{2u}$. It readily follows that

$$\mathbb{E}[(\int_0^T y_t \mathrm{d}t)^{-p}] \leq C(\frac{e^{\kappa T}}{\lambda^2})^{2p}\int_0^\infty n^{2p-1}\cosh(n)^{-1}\exp(-\frac{\tanh(n)n}{\lambda^2})\mathrm{d}n,$$

Table 10.13: Implied Black-Scholes volatilities of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 3\%$, $\kappa = 2$, $\xi = 40\%$ and $\rho = 0\%$.

3M	23.27%	21.25%	19.59%	18.86%	19.47%	20.64%	21.25%	21.85%
	22.35%	21.48%	19.56%	18.85%	19.43%	20.83%	21.48%	21.94%
	92.35	**-22.93**	**2.23**	**1.62**	**3.51**	**-18.90**	**-22.93**	**-9.79**
6M	24.10%	22.05%	20.22%	18.21%	18.68%	20.75%	21.78%	22.72%
	22.52%	22.26%	20.50%	18.14%	18.59%	21.16%	22.10%	22.51%
	158.79	**-20.74**	**-28.41**	**7.08**	**9.08**	**-40.69**	**-32.03**	**20.96**
1Y	23.96%	22.01%	18.89%	17.60%	18.51%	20.84%	22.23%	22.85%
	22.20%	22.14%	18.99%	17.45%	18.48%	21.42%	22.20%	22.30%
	175.41	**-12.81**	**-10.17**	**14.90**	**2.60**	**-57.41**	**2.27**	**54.89**
2Y	22.72%	21.05%	18.61%	17.24%	18.04%	20.26%	21.42%	22.42%
	21.40%	21.20%	18.83%	17.10%	18.06%	20.72%	21.32%	21.42%
	132.35	**-14.49**	**-22.09**	**14.34**	**-1.35**	**-46.40**	**9.96**	**100.04**
3Y	22.44%	20.84%	18.66%	17.16%	17.88%	19.60%	20.84%	21.67%
	20.74%	20.67%	18.93%	17.06%	17.91%	19.96%	20.67%	20.79%
	170.16	**16.92**	**-27.04**	**10.16**	**-3.17**	**-36.03**	**16.92**	**87.99**
5Y	21.56%	20.09%	17.86%	17.16%	17.61%	19.08%	19.94%	20.75%
	20.03%	19.88%	17.92%	17.10%	17.62%	19.28%	19.83%	20.03%
	153.81	**20.49**	**-5.89**	**5.27**	**-0.54**	**-19.86**	**11.43**	**72.25**
7Y	21.93%	19.01%	17.88%	17.17%	17.60%	18.76%	19.54%	20.16%
	19.51%	19.09%	17.95%	17.14%	17.62%	18.88%	19.39%	19.58%
	241.42	**-7.47**	**-6.53**	**3.16**	**-1.53**	**-12.47**	**14.41**	**58.41**
10Y	20.21%	18.92%	17.58%	17.20%	17.53%	18.42%	19.09%	19.61%
	19.24%	18.88%	17.60%	17.18%	17.54%	18.49%	18.97%	19.16%
	96.63	**4.46**	**-1.61**	**1.76**	**-0.93**	**-7.64**	**11.97**	**44.80**

where C is a constant depending only on v_0 and p. We upper bound the above integral differently according to the value of λ.

- If $\lambda \geq 1$, then

$$\mathbb{E}[(\int_0^T y_t \mathrm{d}t)^{-p}] \leq C(\frac{e^{\kappa T}}{\lambda^2})^{2p} \int_0^\infty n^{2p-1}\cosh(n)^{-1}\mathrm{d}n \leq Ce^{2p\kappa T}. \tag{10.19}$$

- If $\lambda \leq 1$, split the integral into two parts, $n \leq \operatorname{arctanh}(\lambda)$ and $n \geq \operatorname{arctanh}(\lambda)$. For the first part, simply use $n \geq \tanh(n)$ for any n. For the second part, use $\tanh(n) \geq \lambda$ and $\cosh(n)^{-1} \leq 1$. This gives

$$\begin{aligned}\mathbb{E}[(\int_0^T y_t \mathrm{d}t)^{-p}] \leq & C\Big[(\frac{e^{\kappa T}}{\lambda^2})^{2p}\int_0^{\operatorname{arctanh}(\lambda)} n^{2p-1}\cosh(n)^{-1}\exp(-\frac{\tanh^2(n)}{\lambda^2})\mathrm{d}n \\ & + (\frac{e^{\kappa T}}{\lambda^2})^{2p}\int_{\operatorname{arctanh}(\lambda)}^\infty n^{2p-1}\exp(-\frac{n}{\lambda})\mathrm{d}n\Big] := C[\mathcal{T}_1 + \mathcal{T}_2]. \end{aligned} \tag{10.20}$$

We upper bound the two terms separately.

1. First term $\mathcal{T}_1$. Using the change of variable $m = \frac{\tanh(n)}{\lambda}$, one has:

$$\mathcal{T}_1 \leq e^{2p\kappa T}\lambda^{-4p+1}\int_0^1 \operatorname{arctanh}(\lambda m)^{2p-1}\cosh(\operatorname{arctanh}(\lambda m))\exp(-m^2)\mathrm{d}m.$$

Table 10.14: Put prices of the closed formula, of the approximation formula and related errors (in bp), expressed as a function of maturities in fractions of years and relative strikes. Parameters: $\theta = 3\%$, $\kappa = 2$, $\xi = 40\%$ and $\rho = 0\%$.

3M	30.00	20.06	10.67	3.76	0.88	0.18	0.08	0.03
	30.00	20.07	10.67	3.76	0.88	0.18	0.08	0.03
	0.12	**-0.46**	**0.24**	**0.32**	**0.45**	**-0.88**	**-0.58**	**-0.13**
6M	40.01	30.05	20.32	5.13	1.91	0.25	0.09	0.04
	40.00	30.05	20.35	5.11	1.89	0.27	0.10	0.03
	0.29	**-0.37**	**-2.15**	**1.99**	**2.06**	**-2.72**	**-1.01**	**0.29**
1Y	50.01	40.06	20.98	7.01	1.73	0.25	0.08	0.05
	50.00	40.06	21.00	6.95	1.73	0.29	0.08	0.04
	0.55	**-0.27**	**-1.80**	**5.92**	**0.70**	**-4.43**	**0.07**	**0.98**
2Y	60.01	50.07	30.89	9.70	2.28	0.28	0.10	0.04
	60.01	50.07	30.93	9.62	2.29	0.33	0.10	0.03
	0.63	**-0.39**	**-4.17**	**8.03**	**-0.51**	**-4.44**	**0.40**	**1.61**
3Y	70.01	60.04	40.60	11.82	2.56	0.36	0.10	0.04
	70.00	60.04	40.64	11.75	2.58	0.41	0.09	0.03
	0.36	**0.28**	**-4.17**	**6.94**	**-1.42**	**-4.48**	**0.71**	**1.54**
5Y	80.00	70.03	41.45	15.21	3.75	0.38	0.11	0.04
	80.00	70.02	41.47	15.16	3.75	0.40	0.11	0.02
	0.15	**0.26**	**-1.76**	**4.61**	**-0.34**	**-2.79**	**0.58**	**1.24**
7Y	90.00	70.07	51.04	17.97	3.78	0.39	0.10	0.03
	90.00	70.08	51.06	17.94	3.80	0.41	0.09	0.02
	0.01	**-0.24**	**-1.63**	**3.25**	**-1.07**	**-1.94**	**0.70**	**1.00**
10Y	90.00	80.03	51.95	21.43	4.55	0.46	0.10	0.03
	90.00	80.03	51.96	21.41	4.56	0.48	0.09	0.02
	0.04	**0.06**	**-0.64**	**2.14**	**-0.80**	**-1.44**	**0.61**	**0.76**

Because of $\lambda \leq 1$, we have the following inequalities for $m \in [0,1[$:

$$\operatorname{arctanh}(\lambda m) \leq \lambda \operatorname{arctanh}(m), \quad \cosh(\operatorname{arctanh}(\lambda m)) \leq \cosh(\operatorname{arctanh}(m)).$$

Using $2p-1 \geq 0$, it readily follows that

$$\mathscr{T}_1 \leq (\frac{e^{2\kappa T}}{\lambda^2})^p \int_0^1 \operatorname{arctanh}(m)^{2p-1} \cosh(\operatorname{arctanh}(m)) \exp(-m^2) \mathrm{d}m. \tag{10.21}$$

2. Second term $\mathscr{T}_2$. Clearly, we have

$$\mathscr{T}_2 \leq (\frac{e^{\kappa T}}{\lambda^2})^{2p} \int_0^\infty n^{2p-1} \exp(-\frac{n}{\lambda}) \mathrm{d}n = (\frac{e^{2\kappa T}}{\lambda^2})^p \int_0^\infty v^{2p-1} e^{-v} \mathrm{d}v. \tag{10.22}$$

Combining (10.20), (10.21) and (10.22), we obtain $\mathbb{E}[(\int_0^T y_t \mathrm{d}t)^{-p}] \leq C(\frac{e^{2\kappa T}}{\lambda^2})^p$. In view of the inequality ($e^x - 1 \geq x, x \geq 0$), we have $\lambda^2 = \frac{(e^{\kappa T}-1)}{2\kappa v_0} \geq \frac{T}{2v_0}$, which gives

$$\mathbb{E}[(\int_0^T y_t \mathrm{d}t)^{-p}] \leq C \frac{e^{2p\kappa T}}{T^p}, \tag{10.23}$$

available when $\lambda \leq 1$.

Table 10.15: Equivalent averaged parameters.

T	$\bar{v}_0$	$\bar{\theta}$	$\bar{\xi}$	$\bar{\rho}$
3M	4 %	4 %	30 %	-20 %
6M	3.97 %	4.04 %	30.12 %	-19.93 %
1Y	3.28 %	4.38 %	30.89 %	-19.72 %
2Y	4.64 %	4.02 %	31.12 %	-18.95 %
3Y	56.24 %	4.04 %	32.10 %	-18.20 %
5Y	28.58 %	2.68 %	33.63 %	-16.52 %
7Y	84.92 %	0.59 %	35.41 %	-14.80 %
10Y	14.54 %	4.57 %	39.98 %	-12.32 %

To sum up (10.19) and (10.23), we have proved that

$$\mathbb{E}[(\int_0^T y_t \mathrm{d}t)^{-p}] \leq Ce^{2p\kappa T}(1+\frac{1}{T^p}), \tag{10.24}$$

for a constant C depending only on p and v_0.

Step 2. Take $\varepsilon \in]0,1]$. We apply Lemma 10.4.2 to v^ε, in order to write $v_t^\varepsilon \geq y^\varepsilon_{A^{-1}_{\varepsilon,t}}$ where $t = \int_0^{A_{\varepsilon,t}} (\varepsilon\xi_s)^2 \mathrm{d}s$ and $\mathrm{d}y_t^\varepsilon = (\frac{1}{2} - \frac{\kappa}{(\varepsilon\xi_{Inf})^2} y_t^\varepsilon)\mathrm{d}t + \sqrt{y_t^\varepsilon} d\tilde{B}_t^\varepsilon, y_0^\varepsilon = y_0$. Thus, we get $\int_0^T v_t^\varepsilon \mathrm{d}t \geq (\int_0^{A^{-1}_{\varepsilon,t}} y_s^\varepsilon \mathrm{d}s)/(\varepsilon\xi_{Sup})^2$ and in view of (10.24), it follows that

$$\begin{aligned}\mathbb{E}[(\int_0^T v_t^\varepsilon \mathrm{d}t)^{-p}] &\leq (\varepsilon\xi_{Sup})^{2p}\mathbb{E}(\int_0^{A^{-1}_{\varepsilon,T}} y_s^\varepsilon \mathrm{d}s)^{-p} \\ &\leq C(\varepsilon\xi_{Sup})^{2p} e^{2p\frac{\kappa}{(\varepsilon\xi_{Inf})^2}A^{-1}_{\varepsilon,T}}(1+\frac{1}{[A^{-1}_{\varepsilon,T}]^p}) \\ &\leq Ce^{2p\kappa\frac{\xi^2_{Sup}}{\xi^2_{Inf}}T}(\xi^{2p}_{Sup} + \frac{\xi^{2p}_{Sup}}{\xi^{2p}_{Inf}}\frac{1}{T^p})\end{aligned}$$

where we have used $\varepsilon^2\xi^2_{Inf}T \leq A^{-1}_{\varepsilon,T} \leq \varepsilon^2\xi^2_{Sup}T$.

Note that the upper bound does not depend on $\varepsilon \in]0,1]$. For $\varepsilon = 0$, the upper bound in Lemma 10.4.3 is also true because $(v_t^0)_t$ is deterministic and

$$\max(v_0, \theta_{Sup}) \geq v_t^0 \geq \min(v_0, \theta_{Inf}) > 0. \tag{10.25}$$

□

10.4.4 Upper bound for residuals of the Taylor development of $g(\varepsilon)$ defined in (10.4)

Throughout the following paragraph, we assume that (P) is in force. We define the variables:

$$P_T^\varepsilon = \int_0^T \rho_t(\sigma_t^\varepsilon - \sigma_{0,t})\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2}(v_t^\varepsilon - v_{0,t})\mathrm{d}t, \qquad Q_T^\varepsilon = \int_0^T (1-\rho_t^2)(v_t^\varepsilon - v_{0,t})\mathrm{d}t.$$

Notice that $(x_0 + \int_0^T \rho_t \sqrt{v_{0,t}}\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2} v_{0,t}\mathrm{d}t, \int_0^T (1-\rho_t^2)v_{0,t}\mathrm{d}t) + (P_T^1, Q_T^1) = (x_0 + \int_0^T \rho_t \sqrt{v_t^1}\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2} v_t^1 \mathrm{d}t, \int_0^T (1-\rho_t^2)v_t^1 \mathrm{d}t)$.

Table 10.16: Implied Black-Scholes volatilities of the closed formula, of the approximation formula and of the averaging formula, expressed as a function of maturities in fractions of years and relative strikes. Piecewise constant parameters.

3M	23.45%	21.88%	20.58%	19.70%	19.39%	19.55%	19.74%	19.97%
	22.73%	21.96%	20.60%	19.69%	19.35%	19.53%	19.84%	20.28%
	23.45%	21.88%	20.58%	19.70%	19.39%	19.55%	19.74%	19.97%
6M	24.09%	22.59%	21.30%	19.63%	19.33%	19.58%	19.92%	20.31%
	23.09%	22.60%	21.43%	19.61%	19.30%	19.58%	20.19%	20.93%
	24.09%	22.59%	21.30%	19.63%	19.33%	19.58%	19.92%	20.31%
1Y	23.95%	22.66%	20.76%	19.70%	19.37%	19.69%	20.12%	20.36%
	23.12%	22.66%	20.81%	19.68%	19.32%	19.78%	20.62%	21.05%
	23.95%	22.66%	20.76%	19.70%	19.37%	19.69%	20.12%	20.35%
2Y	23.26%	22.30%	21.01%	19.99%	19.66%	19.83%	20.09%	20.37%
	22.84%	22.33%	21.04%	19.96%	19.62%	19.90%	20.43%	21.02%
	23.26%	22.30%	21.01%	19.98%	19.66%	19.83%	20.09%	20.37%
3Y	23.28%	22.40%	21.27%	20.26%	19.96%	20.02%	20.23%	20.43%
	22.81%	22.38%	21.33%	20.24%	19.93%	20.04%	20.47%	20.90%
	23.28%	22.40%	21.27%	20.26%	19.96%	20.02%	20.23%	20.42%
5Y	23.22%	22.46%	21.34%	20.77%	20.54%	20.54%	20.65%	20.80%
	22.88%	22.44%	21.35%	20.77%	20.52%	20.55%	20.76%	21.09%
	23.22%	22.46%	21.34%	20.77%	20.54%	20.54%	20.64%	20.79%
7Y	23.86%	22.36%	21.81%	21.26%	21.06%	21.06%	21.16%	21.27%
	23.25%	22.39%	21.82%	21.26%	21.05%	21.07%	21.23%	21.45%
	23.86%	22.37%	21.81%	21.26%	21.06%	21.06%	21.15%	21.26%
10Y	23.59%	22.96%	22.30%	21.97%	21.82%	21.83%	21.92%	22.02%
	23.46%	22.98%	22.30%	21.97%	21.81%	21.84%	21.96%	22.12%
	23.59%	22.96%	22.30%	21.97%	21.82%	21.83%	21.92%	22.01%

The main result of this subsection is the following proposition, the statement of which uses the notation introduced at the beginning of Section 10.4.

Proposition 10.4.1. *One has the following estimates for every $p \geq 1$*

$$\begin{aligned}
\|P^1_T\|_p &\leq C(\xi_{Sup}\sqrt{T})\sqrt{T},\\
\|R^{P^1}_{2,T}\|_p &\leq C(\xi_{Sup}\sqrt{T})^3\sqrt{T},\\
\|R^{(P^1)^2}_{2,T}\|_p &\leq C(\xi_{Sup}\sqrt{T})^3 T,\\
\|Q^1_T\|_p &\leq C(\xi_{Sup}\sqrt{T})T,\\
\|R^{Q^1}_{2,T}\|_p &\leq C(\xi_{Sup}\sqrt{T})^3 T,\\
\|R^{(Q^1)^2}_{2,T}\|_p &\leq C(\xi_{Sup}\sqrt{T})^3 T^2,\\
\|R^{P^1Q^1}_{2,T}\|_p &\leq C(\xi_{Sup}\sqrt{T})^3 T^{\frac{3}{2}}.
\end{aligned}$$

To estimate the derivatives and the residuals for the variables P^{ε}_T and Q^{ε}_T, we need first to prove the existence of the derivatives and the residuals of the volatility process $\sigma^{\varepsilon}_t = \sqrt{v^{\varepsilon}_t}$ and its square v^{ε}. Finally we prove Proposition 10.4.1.

Table 10.17: Put prices of the closed formula, of the approximation formula and of the averaging formula, expressed as a function of maturities in fractions of years and relative strikes. Piecewise constant parameters.

3M	30.00	20.07	10.78	3.93	0.87	0.13	0.05	0.02
	30.00	20.08	10.78	3.93	0.87	0.13	0.05	0.02
	30.00	20.07	10.78	3.93	0.87	0.13	0.05	0.02
6M	40.01	30.06	20.41	5.53	2.06	0.18	0.05	0.01
	40.00	30.06	20.42	5.53	2.05	0.18	0.05	0.02
	40.01	30.06	20.41	5.53	2.06	0.18	0.05	0.01
1Y	50.01	40.07	21.33	7.85	1.97	0.17	0.03	0.02
	50.01	40.07	21.35	7.84	1.95	0.18	0.04	0.02
	50.01	40.07	21.33	7.85	1.97	0.17	0.03	0.02
2Y	60.02	50.11	31.38	11.23	2.92	0.24	0.06	0.01
	60.01	50.11	31.39	11.23	2.90	0.25	0.07	0.02
	60.02	50.11	31.38	11.23	2.92	0.24	0.06	0.01
3Y	70.01	60.07	41.07	13.92	3.55	0.41	0.08	0.02
	70.01	60.07	41.08	13.92	3.54	0.42	0.09	0.03
	70.01	60.07	41.07	13.92	3.55	0.41	0.08	0.02
5Y	80.01	70.07	42.64	18.37	5.74	0.61	0.15	0.04
	80.01	70.07	42.64	18.36	5.72	0.61	0.16	0.04
	80.01	70.07	42.64	18.37	5.74	0.61	0.15	0.04
7Y	90.00	70.24	52.22	22.15	6.46	0.86	0.21	0.06
	90.00	70.24	52.22	22.15	6.45	0.86	0.21	0.07
	90.00	70.24	52.22	22.15	6.46	0.86	0.21	0.06
10Y	90.01	80.14	54.13	27.17	8.71	1.42	0.35	0.11
	90.01	80.14	54.13	27.16	8.70	1.42	0.36	0.12
	90.01	80.14	54.13	27.17	8.71	1.42	0.35	0.11

Upper bounds for derivatives of σ^ε and v^ε

Under assumption (P), the volatility process σ_t^ε is governed by the SDE:

$$d\sigma_t^\varepsilon = ((\frac{\kappa\theta_t}{2} - \frac{\varepsilon^2\xi_t^2}{8})\frac{1}{\sigma_t^\varepsilon} - \frac{\kappa}{2}\sigma_t^\varepsilon)\mathrm{d}t + \frac{\varepsilon\xi_t}{2}\mathrm{d}B_t,\ \sigma_0^\varepsilon = \sqrt{v_0}, \tag{10.26}$$

where we have used Ito's Lemma and positivity of v_t^ε (see Lemma 10.4.2).
In order to estimate $R_{0,t}^{\sigma^\varepsilon}$, we are going to prove that it verifies a linear equation (Lemma 10.4.4) from which we deduce an a priori upper bound (Proposition 10.4.2). We iterate the same analysis for the residuals $R_{1,t}^{\sigma^\varepsilon}$ (Proposition 10.4.3) and $R_{2,t}^{\sigma^\varepsilon}$ (Proposition 10.4.4). Analogously, we give upper bounds for the residuals of v_t^ε (Proposition 10.4.1).

Lemma 10.4.4. *Under (P), the process $(R_{0,t}^{\sigma^\varepsilon} = \sigma_t^\varepsilon - \sigma_t^0)_{0\le t\le T}$ is given by*

$$R_{0,t}^{\sigma^\varepsilon} = U_t^\varepsilon \int_0^t (U_s^\varepsilon)^{-1}(-\frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}}\mathrm{d}s + \frac{\varepsilon\xi_s}{2}\mathrm{d}B_s),$$

where

$$\mathrm{d}U_t^\varepsilon = -\alpha_t^\varepsilon U_t^\varepsilon \mathrm{d}t,\ U_t^\varepsilon = 1,$$
$$\alpha_t^\varepsilon = (\frac{\kappa\theta_t}{2} - \frac{\varepsilon^2\xi_t^2}{8})\frac{1}{\sigma_t^\varepsilon\sigma_{0,t}} + \frac{\kappa}{2}.$$

Proof. From the definition $(\sigma_{0,t})_t = (\sigma_t^0)_t$ and the equation (10.26), one obtains the SDE

$$d\sigma_{0,t} = (\frac{\kappa\theta_t}{2\sigma_{0,t}} - \frac{\kappa}{2}\sigma_{0,t})\mathrm{d}t,\ \sigma_{0,0} = \sqrt{v_0}.$$

Substitute this equation in (10.26) to obtain

$$\mathrm{d}R_{0,t}^{\sigma^\varepsilon} = -\alpha_t^\varepsilon R_{0,t}^{\sigma^\varepsilon}\mathrm{d}t - \frac{\varepsilon^2\xi_t^2}{8\sigma_{0,t}}\mathrm{d}t + \frac{\varepsilon\xi_t}{2}\mathrm{d}B_t,\ R_{0,0}^{\sigma^\varepsilon} = 0. \tag{10.27}$$

Note that $R_{0,.}^{\sigma^\varepsilon}$ is the solution of a linear SDE. Hence, it can be explicitly represented using the process U^ε (see Th. 52 in [96]):

$$R_{0,t}^{\sigma^\varepsilon} = U_t^\varepsilon \int_0^t (U_s^\varepsilon)^{-1}(-\frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}}\mathrm{d}s + \frac{\varepsilon\xi_s}{2}\mathrm{d}B_s).$$

□

Proposition 10.4.2. *Under* (P)*, for every* $p \geq 1$ *one has*

$$\|(R_{0,.}^{\sigma^\varepsilon})_t^*\|_p \leq C\varepsilon\xi_{Sup}\sqrt{t}.$$

In particular, the application $\varepsilon \to \sigma_t^\varepsilon$ *is continuous*[5] *at* $\varepsilon = 0$ *in* L_p.

Proof. At first sight, the proof seems to be straightforward from Lemma 10.4.4. But actually, the difficulty lies in the fact that one can not uniformly in ε upper bound U_t^ε in L_p (because of the term with $1/\sigma_t^\varepsilon$ in α_t^ε).

Using Lemma 10.4.4 and Ito's formula for the product $(U_t^\varepsilon)^{-1}(\int_0^t \frac{\varepsilon\xi_s}{2}\mathrm{d}B_s)$, one has

$$R_{0,t}^{\sigma^\varepsilon} = U_t^\varepsilon \int_0^t (U_s^\varepsilon)^{-1}(-\frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}}\mathrm{d}s) + \int_0^t \frac{\varepsilon\xi_s}{2}\mathrm{d}B_s - U_t^\varepsilon \int_0^t (\int_0^s \frac{\varepsilon\xi_u}{2}\mathrm{d}B_u) d(U_s^\varepsilon)^{-1}.$$

Under (P), one has $\alpha_t^\varepsilon \geq \kappa/2 > 0$, which implies that $t \mapsto U_t^\varepsilon$ is decreasing and $t \mapsto (U_t^\varepsilon)^{-1}$ is increasing. Thus, $0 \leq U_t^\varepsilon (U_s^\varepsilon)^{-1} \leq 1$ for $s \in [0,t]$. Consequently, we deduce

$$\begin{aligned} |R_{0,t}^{\sigma^\varepsilon}| &\leq \int_0^t \frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}}\mathrm{d}s + (\int_0^\cdot \frac{\varepsilon\xi_s}{2}\mathrm{d}B_s)_t^* + (\int_0^\cdot \frac{\varepsilon\xi_s}{2}\mathrm{d}B_s)_t^*(1 - U_t^\varepsilon) \\ &\leq \int_0^t \frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}}\mathrm{d}s + (\int_0^\cdot \varepsilon\xi_s\mathrm{d}B_s)_t^*. \end{aligned} \tag{10.28}$$

Now we easily complete the proof by observing that $\sigma_{0,s} \geq \min(\sqrt{\theta_{Inf}}, \sqrt{v_0})$ and $\|(\int_0^\cdot \xi_s\mathrm{d}B_s)_t^*\|_p \leq C\xi_{Sup}\sqrt{t}$. □

We define

$$\sigma_{1,t} = U_t^0 \int_0^t (U_s^0)^{-1}\frac{\xi_s}{2}\mathrm{d}B_s.$$

[5]Note that from the upper bound (10.28) in the proof, we easily obtain that the continuity also holds a.s., and not only in L_p. Since only the latter is needed in what follows, we do not go into detail.

Therefore, $(\sigma_{1,t})_{0\le t\le T}$ solves the following SDE:

$$d\sigma_{1,t} = -(\frac{\kappa\theta_t}{2(\sigma_{0,t})^2} + \frac{\kappa}{2})\sigma_{1,t}\mathrm{d}t + \frac{\xi_t}{2}\mathrm{d}B_t,\ \sigma_{1,0} = 0, \tag{10.29}$$

and for every $p \ge 1$

$$\|(\sigma_{1,.})^*_t\|_p \le C\xi_{Sup}\sqrt{t}. \tag{10.30}$$

Proposition 10.4.3. *Under (P), the process $(R^{\sigma^\varepsilon}_{1,t} = \sigma^\varepsilon_t - \sigma^0_t - \varepsilon\sigma_{1,t})_{0\le t\le T}$ fulfills the equality:*

$$R^{\sigma^\varepsilon}_{1,t} = U^\varepsilon_t \int_0^t (U^\varepsilon_s)^{-1}(-\frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}} + \varepsilon\sigma_{1,s}((\frac{\alpha^\varepsilon_s}{\sigma_{0,s}} - \frac{\kappa}{2\sigma_{0,s}})R^{\sigma^\varepsilon}_{0,s} + \frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}}))\mathrm{d}s.$$

Moreover, for every $p \ge 1$, one has

$$\|(R^{\sigma^\varepsilon}_{1,.})^*_t\|_p \le C(\varepsilon\xi_{Sup}\sqrt{t})^2.$$

In particular, the application $\varepsilon \to \sigma^\varepsilon_t$ is $\mathscr{C}^1$ at $\varepsilon = 0$ in L_p sense with the first derivative at $\varepsilon = 0$ equal to $\sigma_{1,t}$ (justifying a posteriori the definition $R^{\sigma^\varepsilon}_{1,.}$).

Proof. From Equations (10.27) and (10.29), it readily follows that

$$\mathrm{d}R^{\sigma^\varepsilon}_{1,t} = -\alpha^\varepsilon_t R^{\sigma^\varepsilon}_{1,t}\mathrm{d}t - \varepsilon\sigma_{1,t}(\alpha^\varepsilon_t - \frac{\kappa\theta_t}{2(\sigma_{0,t})^2} - \frac{\kappa}{2})\mathrm{d}t - \frac{\varepsilon^2\xi_t^2}{8\sigma_{0,t}}\mathrm{d}t, R^{\sigma^\varepsilon}_{1,0} = 0.$$

Because of the identity

$$-(\alpha^\varepsilon_t - \frac{\kappa\theta_t}{2(\sigma_{0,t})^2} - \frac{\kappa}{2}) = ((\frac{\alpha^\varepsilon_t}{\sigma_{0,t}} - \frac{\kappa}{2\sigma_{0,t}})R^{\sigma^\varepsilon}_{0,t} + \frac{\varepsilon^2\xi_t^2}{8(\sigma_{0,t})^2}),$$

one deduces the equality

$$R^{\sigma^\varepsilon}_{1,t} = U^\varepsilon_t \int_0^t (U^\varepsilon_s)^{-1}(-\frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}} + \varepsilon\sigma_{1,s}((\frac{\alpha^\varepsilon_s}{\sigma_{0,s}} - \frac{\kappa}{2\sigma_{0,s}})R^{\sigma^\varepsilon}_{0,s} + \frac{\varepsilon^2\xi_s^2}{8(\sigma_{0,s})^2}))\mathrm{d}s.$$

Then

$$\begin{aligned}
|R^{\sigma^\varepsilon}_{1,t}| &\le \int_0^t U^\varepsilon_t(U^\varepsilon_s)^{-1}(\frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}} + \varepsilon|\sigma_{1,s}|((\frac{\alpha^\varepsilon_s}{\sigma_{0,s}} + \frac{\kappa}{2\sigma_{0,s}})|R^{\sigma^\varepsilon}_{0,s}| + \frac{\varepsilon^2\xi_s^2}{8(\sigma_{0,s})^2}))\mathrm{d}s \\
&\le \int_0^t U^\varepsilon_t(U^\varepsilon_s)^{-1}(\frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}} + \varepsilon|\sigma_{1,s}|(\frac{\kappa}{2\sigma_{0,s}}|R^{\sigma^\varepsilon}_{0,s}| + \frac{\varepsilon^2\xi_s^2}{8(\sigma_{0,s})^2}))\mathrm{d}s + \varepsilon\int_0^t U^\varepsilon_t(U^\varepsilon_s)^{-1}\frac{\alpha^\varepsilon_s}{\sigma_{0,s}}|\sigma_{1,s}||R^{\sigma^\varepsilon}_{0,s}|\mathrm{d}s \\
&\le \int_0^t (\frac{\varepsilon^2\xi_s^2}{8\sigma_{0,s}} + \varepsilon|\sigma_{1,s}|(\frac{\kappa}{2\sigma_{0,s}}|R^{\sigma^\varepsilon}_{0,s}| + \frac{\varepsilon^2\xi_s^2}{8(\sigma_{0,s})^2}))\mathrm{d}s + \varepsilon(\frac{\sigma_{1,.}R^{\sigma^\varepsilon}_{0,.}}{\sigma_{0,.}})^*_t,
\end{aligned}$$

where we have used $U^\varepsilon_t(U^\varepsilon_s)^{-1} \le 1$ for every $s \in [0,t]$ and $U^\varepsilon_t \int_0^t \alpha^\varepsilon_s(U^\varepsilon_s)^{-1}\mathrm{d}s = 1 - U^\varepsilon_t \le 1$ for the third inequality. Apply Proposition 10.4.2 and Inequality (10.30) to complete the proof of the estimate of $\|(R^{\sigma^\varepsilon}_{1,.})^*_t\|_p$. □

We define $(\sigma_{2,t})_{0\leq t\leq T}$ as the solution of the linear equation

$$d\sigma_{2,t} = (-(\frac{\kappa\theta_t}{2(\sigma_{0,t})^2} + \frac{\kappa}{2})\sigma_{2,t} + \kappa\theta_t\frac{(\sigma_{1,t})^2}{(\sigma_{0,t})^3} - \frac{\xi_t^2}{4\sigma_{0,t}})\mathrm{d}t,\ \sigma_{2,0} = 0. \tag{10.31}$$

Clearly, for $p \geq 1$, we have

$$\|(\sigma_{2,.})_t^*\|_p \leq C(\xi_{Sup}\sqrt{t})^2. \tag{10.32}$$

Proposition 10.4.4. *Under (P), the process $(R_{2,t}^{\sigma^\varepsilon} = \sigma_t^\varepsilon - \sigma_t^0 - \varepsilon\sigma_{1,t} - \frac{\varepsilon^2}{2}\sigma_{2,t})_{0\leq t\leq T}$ fulfills the equality:*

$$R_{2,t}^{\sigma^\varepsilon} = U_t^\varepsilon \int_0^t (U_s^\varepsilon)^{-1}[\varepsilon^2((\frac{\alpha_s^\varepsilon}{\sigma_{0,s}} - \frac{\kappa}{2\sigma_{0,s}})R_{0,s}^{\sigma^\varepsilon} + \frac{\varepsilon^2\xi_s^2}{8(\sigma_{0,s})^2})(\frac{\sigma_{2,s}}{2} - \frac{(\sigma_{1,s})^2}{\sigma_{0,s}})$$
$$+\varepsilon((\frac{\alpha_s^\varepsilon}{\sigma_{0,s}} - \frac{\kappa}{2\sigma_{0,s}})R_{1,s}^{\sigma^\varepsilon} + \frac{\varepsilon^2\xi_s^2}{8(\sigma_{0,s})^2})\sigma_{1,s}]\mathrm{d}s.$$

Moreover, for every $p \geq 1$, one has

$$\|(R_{2,.}^{\sigma^\varepsilon})_t^*\|_p \leq C(\varepsilon\xi_{Sup}\sqrt{t})^3.$$

In particular, the application $\varepsilon \to \sigma_t^\varepsilon$ is $\mathscr{C}^2$ at $\varepsilon = 0$ in L_p sense with the second derivative at $\varepsilon = 0$ equal to $\sigma_{2,t}$.

Proof. The equality is easy to check. The estimate is proved in the same way as in the proof of Proposition 10.4.3, we therefore skip the details. □

Corollary 10.4.1. *The application $\varepsilon \to v_t^\varepsilon$ is $\mathscr{C}^2$ at $\varepsilon = 0$ in L_p sense. The residuals for the squared volatility satisfy the following inequalities: for every $p \geq 1$, one has*

$$\|(R_{0,.}^{v^\varepsilon})_t^*\|_p \leq C\varepsilon\xi_{Sup}\sqrt{t},$$
$$\|(R_{1,.}^{v^\varepsilon})_t^*\|_p \leq C(\varepsilon\xi_{Sup}\sqrt{t})^2,$$
$$\|(R_{2,.}^{v^\varepsilon})_t^*\|_p \leq C(\varepsilon\xi_{Sup}\sqrt{t})^3.$$

Proof. Note that $v_t^\varepsilon = (\sigma_t^\varepsilon)^2 = (\sigma_{0,t} + R_{0,t}^{\sigma^\varepsilon})^2 = v_{0,t} + 2\sigma_{0,t}R_{0,t}^{\sigma^\varepsilon} + (R_{0,t}^{\sigma^\varepsilon})^2$. Thus, we have $R_{0,t}^{v^\varepsilon} = 2\sigma_{0,t}R_{0,t}^{\sigma^\varepsilon} + (R_{0,t}^{\sigma^\varepsilon})^2$, which leads to the required estimate using $\sigma_{0,t} \leq \max(\sqrt{v_0}, \sqrt{\theta_{Sup}})$ and Proposition 10.4.2. The other estimates are proved analogously using Propositions 10.4.3 and 10.4.4 and Inequalities (10.30) and (10.32). □

Proof of Proposition 10.4.1

We can write

$$P_T^1 = \int_0^T \rho_t R_{0,t}^{\sigma^1}\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2}R_{0,t}^{v^1}\mathrm{d}t, \quad R_{2,T}^{P^1} = \int_0^T \rho_t R_{2,t}^{\sigma^1}\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2}R_{2,t}^{v^1}\mathrm{d}t.$$

Then, using Propositions 10.4.2, 10.4.4 and Corollary 10.4.1, we prove the two first estimates of Proposition 10.4.1. The others inequalities are proved in the same way.

10.4.5 Proof of Theorem 10.2.2

For convenience, we introduce the following notation for $\lambda \in [0,1]$:

$$\begin{aligned}\bar{P}_{BS}(\lambda) =& P_{BS}\Big(x_0 + \int_0^T \rho_t((1-\lambda)\sqrt{v_{0,t}} + \lambda\sqrt{v_t^1})dB_t - \int_0^T \frac{\rho_t^2}{2}((1-\lambda)v_{0,t} + \lambda v_t^1)dt \\ &, \int_0^T (1-\rho_t^2)((1-\lambda)v_{0,t} + \lambda v_t^1)dt\Big), \\ \frac{\partial^{i+j}\bar{P}_{BS}}{\partial x^i y^j}(\lambda) =& \frac{\partial^{i+j}P_{BS}}{\partial x^i y^j}\Big(x_0 + \int_0^T \rho_t((1-\lambda)\sqrt{v_{0,t}} + \lambda\sqrt{v_t^1})dB_t - \int_0^T \frac{\rho_t^2}{2}((1-\lambda)v_{0,t} + \lambda v_t^1)dt \\ &, \int_0^T (1-\rho_t^2)((1-\lambda)v_{0,t} + \lambda v_t^1)dt\Big).\end{aligned}$$

Notice that $\tilde{P}_{BS}$ (see (10.9)) is a particular case of $\bar{P}_{BS}$ for $\lambda = 0$:

$$\tilde{P}_{BS} = \bar{P}_{BS}(0), \quad \frac{\partial^{i+j}\tilde{P}_{BS}}{\partial x^i y^j} = \frac{\partial^{i+j}\bar{P}_{BS}}{\partial x^i y^j}(0).$$

Now, we represent the error $\mathscr{E}$ in (10.16) using the previous notations. A second order Taylor expansion leads to

$$g(1) = \mathbb{E}(\bar{P}_{BS}(1)) = \mathbb{E}(\bar{P}_{BS}(0) + \partial_\lambda \bar{P}_{BS}(0) + \frac{1}{2}\partial_\lambda^2 \bar{P}_{BS}(0) + \int_0^1 d\lambda \frac{(1-\lambda)^2}{2}\partial_\lambda^3 \bar{P}_{BS}(\lambda)).$$

The first term $\mathbb{E}(\bar{P}_{BS}(0))$ is equal to (10.10). Approximations of the three above derivatives contribute to the error $\mathscr{E}$.

1. We have $\mathbb{E}(\partial_\lambda \bar{P}_{BS}(0)) = \mathbb{E}(\frac{\partial \tilde{P}_{BS}}{\partial x} P_T^1 + \frac{\partial \tilde{P}_{BS}}{\partial y} Q_T^1)$. These two terms are equal to (10.11) and (10.12) plus an error equal to
$$\mathbb{E}(\frac{\partial \tilde{P}_{BS}}{\partial x} R_{2,T}^{P^1} + \frac{\partial \tilde{P}_{BS}}{\partial y} R_{2,T}^{Q^1}).$$

2. Regarding the second derivatives, we have $\mathbb{E}(\frac{1}{2}\partial_\lambda^2 \bar{P}_{BS}(0)) = \mathbb{E}(\frac{1}{2}\frac{\partial^2 \tilde{P}_{BS}}{\partial x^2}(P_T^1)^2 + \frac{1}{2}\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}(Q_T^1)^2 + \frac{\partial^2 \tilde{P}_{BS}}{\partial xy} P_T^1 Q_T^1)$. These terms are equal to (10.13), (10.14) and (10.15), plus an error equal to
$$\mathbb{E}(\frac{1}{2}\frac{\partial^2 \tilde{P}_{BS}}{\partial x^2} R_{2,T}^{(P^1)^2} + \frac{1}{2}\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2} R_{2,T}^{(Q^1)^2} + \frac{\partial^2 \tilde{P}_{BS}}{\partial xy} R_{2,T}^{P^1 Q^1}).$$

3. The last term with $\partial_\lambda^3 \bar{P}_{BS}$ is neglected and thus is considered as an error.

To sum up, we have shown that

$$\begin{aligned}\mathscr{E} =& \sum_{i=0}^{1} \mathbb{E}[\frac{\partial^1 \bar{P}_{BS}}{\partial x^i y^{1-i}}(0) R_{2,T}^{(P^1)^i (Q^1)^{1-i}}] + \sum_{i=0}^{2} \frac{C_2^i}{2}\mathbb{E}[\frac{\partial^2 \bar{P}_{BS}}{\partial x^i y^{2-i}}(0) R_{2,T}^{(P^1)^i (Q^1)^{2-i}}] \\ &+ \int_0^1 \frac{(1-\lambda)^2}{2} \sum_{i=0}^{3} C_3^i \mathbb{E}[\frac{\partial^3 \bar{P}_{BS}}{\partial x^i y^{3-i}}(\lambda)(P_T^1)^i (Q_T^1)^{3-i}] d\lambda.\end{aligned}$$

Using Lemma 10.4.1 and Assumption (R), one obtains for all $\lambda \in [0,1]$

$$\begin{aligned}\|\frac{\partial^{i+j}\bar{P}_{BS}}{\partial x^i y^j}(\lambda)\|_2 &\leq C\|(\int_0^T((1-\lambda)v_{0,t}+\lambda v_t^1)\mathrm{d}t)^{\frac{-(2j+i-1)_+}{2}}\|_4\\ &\leq C((1-\lambda)\|(\int_0^T v_{0,t}\mathrm{d}t)^{\frac{-(2j+i-1)_+}{2}}\|_4+\lambda\|(\int_0^T v_t^1\mathrm{d}t)^{\frac{-(2j+i-1)_+}{2}}\|_4)\end{aligned}$$

where we have applied a convexity argument. Finally, apply Lemma 10.4.3 with $\varepsilon=0$ and $\varepsilon=1$ to conclude that

$$\|\frac{\partial^{i+j}\bar{P}_{BS}}{\partial x^i y^j}(\lambda)\|_2 \leq \frac{C}{(\sqrt{T})^{(2j+i-1)_+}},$$

uniformly w.r.t. $\lambda \in [0,1]$. Combining this with Proposition 10.4.1 yields that

$$\begin{aligned}|\mathscr{E}| &\leq C\left(\sum_{i=0}^{1}(\xi_{Sup}\sqrt{T})^3\frac{T^{1-i/2}}{(\sqrt{T})^{1-i}}+\sum_{i=0}^{2}(\xi_{Sup}\sqrt{T})^3\frac{T^{2-i/2}}{(\sqrt{T})^{3-i}}+\sum_{i=0}^{3}(\xi_{Sup}\sqrt{T})^3\frac{T^{3-i/2}}{(\sqrt{T})^{5-i}}\right)\\ &\leq C(\xi_{Sup}\sqrt{T})^3\sqrt{T}.\end{aligned}$$

Theorem 10.2.2 is proved.

10.5 Proof of Proposition 10.2.1 and Theorem 10.2.1

10.5.1 Preliminary results

In this section, we bring together the results (and their proofs) which allow us to derive the explicit terms in the formula (10.17).
In the following, α_t (resp. β_t) is a square integrable and predictable process (resp. deterministic) and l is a smooth function with derivatives having, at most, exponential growth.
For the next Malliavin calculus computations, we freely use standard notations from [88].

Lemma 10.5.1. *(Lemma 1.2.1 in [88]) Let $G \in \mathbb{D}^{1,\infty}(\Omega)$. One has*

$$\mathbb{E}[G\int_0^t \alpha_s \mathrm{d}B_s] = \mathbb{E}[\int_0^t \alpha_s D_s^B(G)\mathrm{d}s],$$

where $D^B(G) = (D_s^B(G))_{s\geq 0}$ is the first Malliavin derivative of G w.r.t. B.

Taking $G = l(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)$ gives the following result.

Lemma 10.5.2. *One has:*

$$\mathbb{E}[(\int_0^T \alpha_t \mathrm{d}B_t)l(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)] = \mathbb{E}[(\int_0^T \rho_t\sqrt{v_{0,t}}\alpha_t\mathrm{d}t)l^{(1)}(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)].$$

Lemma 10.5.3. *For any deterministic integrable function f and any continuous semimartingale Z vanishing at t=0, one has:*

$$\int_0^T f(t)Z_t\mathrm{d}t = \int_0^T \omega_{t,T}^{(0,f)}\mathrm{d}Z_t.$$

Proof. This is an application of the Itô formula to the product $\omega_{t,T}^{(0,f)}Z_t$. □

Lemma 10.5.4. *One has:*

$$\mathbb{E}[l(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)\int_0^T \beta_t v_{1,t}\mathrm{d}t] = \omega_{0,T}^{(\kappa,\rho\xi v_{0,\cdot}),(-\kappa,\beta)}\mathbb{E}[l^{(1)}(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)],$$
$$\mathbb{E}[l(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)\int_0^T \beta_t v_{1,t}^2\mathrm{d}t] = \omega_{0,T}^{(2\kappa,\xi^2 v_{0,\cdot}),(-2\kappa,\beta)}\mathbb{E}[l(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)]$$
$$+2\omega_{0,T}^{(\kappa,\rho\xi v_{0,\cdot}),(\kappa,\rho\xi v_{0,\cdot}),(-2\kappa,\beta)}\mathbb{E}[l^{(2)}(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)],$$
$$\mathbb{E}[l(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)\int_0^T \beta_t v_{2,t}\mathrm{d}t] = \omega_{0,T}^{(\kappa,\rho\xi v_{0,\cdot}),(0,\rho\xi),(-\kappa,\beta)}\mathbb{E}[l^{(2)}(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)].$$

Proof. Using Lemmas 10.5.2 ($f(t)=e^{-\kappa t}\beta_t$, $Z_t=\int_0^t e^{\kappa s}\xi_s\sqrt{v_{0,s}}\mathrm{d}B_s$) and 10.5.3, one has:

$$\mathbb{E}[l(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)\int_0^T \beta_t v_{1,t}\mathrm{d}t] = \mathbb{E}[l(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)\int_0^T e^{-\kappa t}\beta_t\int_0^t e^{\kappa s}\xi_s\sqrt{v_{0,s}}\mathrm{d}B_s\mathrm{d}t]$$
$$= \mathbb{E}[l(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)\int_0^T \omega_{t,T}^{(-\kappa,\beta)}e^{\kappa t}\xi_t\sqrt{v_{0,t}}\mathrm{d}B_t]$$
$$= \mathbb{E}[l^{(1)}(\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t)]\int_0^T \omega_{t,T}^{(-\kappa,\beta)}e^{\kappa t}\rho_t\xi_t v_{0,t}\mathrm{d}t,$$

which gives the first equality. The second and the third are proved in the same way. □

Lemma 10.5.5. *One has*

$$\mathbb{E}[\frac{\partial^{i+j}\tilde{P}_{BS}}{\partial x^i y^j}] = \frac{\partial^{i+j}P_{BS}}{\partial x^i y^j}(x_0,\int_0^T v_{0,t}\mathrm{d}t).$$

Proof. One has

$$\mathbb{E}[\frac{\partial^i\tilde{P}_{BS}}{\partial x^i}] = \partial^i_{x=x_0}\mathbb{E}[P_{BS}(x_0+\int_0^T \rho_t\sqrt{v_{0,t}}\mathrm{d}B_t-\int_0^T\frac{\rho_t^2}{2}v_{0,t}\mathrm{d}t,\int_0^T(1-\rho_t^2)v_{0,t}\mathrm{d}t)]$$
$$= \frac{\partial^i P_{BS}}{\partial x^i}(x_0,\int_0^T v_{0,t}\mathrm{d}t).$$

Since $\tilde{P}_{BS}$ verifies the following relation

$$\frac{\partial\tilde{P}_{BS}}{\partial y} = \frac{1}{2}(\frac{\partial^2\tilde{P}_{BS}}{\partial x^2}-\frac{\partial\tilde{P}_{BS}}{\partial x}), \tag{10.33}$$

we immediately obtain the result. □

10.5.2 Proof of Proposition 10.2.1

One has

$$\mathbb{E}[\frac{\partial\tilde{P}_{BS}}{\partial x}(\int_0^T \rho_t(\frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}+\frac{v_{2,t}}{4(v_{0,t})^{\frac{1}{2}}}-\frac{v_{1,t}^2}{8(v_{0,t})^{\frac{3}{2}}})\mathrm{d}B_t-\int_0^T\frac{\rho_t^2}{2}(v_{1,t}+\frac{v_{2,t}}{2})\mathrm{d}t)]$$
$$= \mathbb{E}[\frac{1}{2}(\frac{\partial^2\tilde{P}_{BS}}{\partial x^2}-\frac{\partial\tilde{P}_{BS}}{\partial x})\int_0^T\rho_t^2(v_{1,t}+\frac{v_{2,t}}{2})\mathrm{d}t]-\mathbb{E}[\frac{\partial^2\tilde{P}_{BS}}{\partial x^2}\int_0^T\frac{\rho_t^2 v_{1,t}^2}{8v_{0,t}}\mathrm{d}t]$$
$$= \mathbb{E}[\frac{\partial\tilde{P}_{BS}}{\partial y}\int_0^T\rho_t^2(v_{1,t}+\frac{v_{2,t}}{2})\mathrm{d}t]-\mathbb{E}[\frac{\partial^2\tilde{P}_{BS}}{\partial x^2}\int_0^T\frac{\rho_t^2 v_{1,t}^2}{8v_{0,t}}\mathrm{d}t],$$

where we have used Lemma 10.5.2 at the first equality and identity (10.33) at the second one. Plugging this relation into the approximation (10.16) and summing the second and third line, one has

$$\begin{aligned}
g(1) =&\mathbb{E}[\tilde{P}_{BS}]+\mathbb{E}[\frac{\partial \tilde{P}_{BS}}{\partial y}\int_0^T (v_{1,t}+\frac{v_{2,t}}{2})\mathrm{d}t] \\
&-\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial x^2}\int_0^T \frac{\rho_t^2 v_{1,t}^2}{8v_{0,t}}\mathrm{d}t]+\frac{1}{2}\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial x^2}(\int_0^T \rho_t\frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}\mathrm{d}B_t-\int_0^T \frac{\rho_t^2}{2}v_{1,t}\mathrm{d}t)^2] \\
&+\frac{1}{2}\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}(\int_0^T (1-\rho_t^2)v_{1,t}\mathrm{d}t)^2] \\
&+\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial xy}(\int_0^T (1-\rho_t^2)v_{1,t}\mathrm{d}t)(\int_0^T \rho_t\frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}\mathrm{d}B_t-\int_0^T \frac{\rho_t^2}{2}v_{1,t}\mathrm{d}t)]+\mathscr{E}.
\end{aligned}\tag{10.34}$$

In addition, one has

$$\begin{aligned}
&-\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial x^2}\int_0^T \frac{\rho_t^2 v_{1,t}^2}{8v_{0,t}}\mathrm{d}t]+\frac{1}{2}\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial x^2}(\int_0^T \rho_t\frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}\mathrm{d}B_t-\int_0^T \frac{\rho_t^2}{2}v_{1,t}\mathrm{d}t)^2] \\
&=\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial x^2}\int_0^T (\int_0^t \rho_s\frac{v_{1,s}}{2(v_{0,s})^{\frac{1}{2}}}\mathrm{d}B_s-\int_0^t \frac{\rho_s^2}{2}v_{1,s}\mathrm{d}s)(\rho_t\frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}\mathrm{d}B_t-\frac{\rho_t^2}{2}v_{1,t}\mathrm{d}t)] \\
&=\mathbb{E}[\frac{1}{2}(\frac{\partial^3 \tilde{P}_{BS}}{\partial x^3}-\frac{\partial^2 \tilde{P}_{BS}}{\partial x^2})\int_0^T (\int_0^t \rho_s\frac{v_{1,s}}{2(v_{0,s})^{\frac{1}{2}}}\mathrm{d}B_s-\int_0^t \frac{\rho_s^2}{2}v_{1,s}\mathrm{d}s)\rho_t^2 v_{1,t}\mathrm{d}t] \\
&=\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial xy}\int_0^T (\int_0^t \rho_s\frac{v_{1,s}}{2(v_{0,s})^{\frac{1}{2}}}\mathrm{d}B_s-\int_0^t \frac{\rho_s^2}{2}v_{1,s}\mathrm{d}s)\rho_t^2 v_{1,t}\mathrm{d}t],
\end{aligned}$$

where we have used Ito's Lemma for the square at the first equality, Lemma 10.5.2 at the second and Identity (10.33) at the third one. Substituting this relation in the approximation (10.34) and summing the second and fourth line, one gets

$$\begin{aligned}
g(1) =&\mathbb{E}[\tilde{P}_{BS}]+\mathbb{E}[\frac{\partial \tilde{P}_{BS}}{\partial y}\int_0^T (v_{1,t}+\frac{v_{2,t}}{2})\mathrm{d}t] \\
&+\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial xy}(\int_0^T (\int_0^t \rho_s\frac{v_{1,s}}{2(v_{0,s})^{\frac{1}{2}}}\mathrm{d}B_s-\int_0^t \frac{\rho_s^2}{2}v_{1,s}\mathrm{d}s)\rho_t^2 v_{1,t}\mathrm{d}t \\
&\quad+(\int_0^T (1-\rho_t^2)v_{1,t}\mathrm{d}t)(\int_0^T \rho_t\frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}\mathrm{d}B_t-\int_0^T \frac{\rho_t^2}{2}v_{1,t}\mathrm{d}t))] \\
&+\frac{1}{2}\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}(\int_0^T (1-\rho_t^2)v_{1,t}\mathrm{d}t)^2]+\mathscr{E}.
\end{aligned}\tag{10.35}$$

We now study the second term of (10.35). In the computations below, we use Ito's Lemma for the second equality, Lemma 10.5.2 and Identity (10.33) for the third equality and Lemma 10.5.1 ($G=\frac{\partial^2 \tilde{P}_{BS}}{\partial xy}v_{1,t}$) for the fourth one; it gives

$$
\begin{aligned}
A =&\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial xy}(\int_0^T(\int_0^t \rho_s \frac{v_{1,s}}{2(v_{0,s})^{\frac{1}{2}}}\mathrm{d}B_s - \int_0^t \frac{\rho_s^2}{2} v_{1,s}\mathrm{d}s)\rho_t^2 v_{1,t}\mathrm{d}t \\
&+(\int_0^T (1-\rho_t^2)v_{1,t}\mathrm{d}t)(\int_0^T \rho_t \frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}\mathrm{d}B_t - \int_0^T \frac{\rho_t^2}{2} v_{1,t}\mathrm{d}t))] \\
=&\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial xy}(\int_0^T(\int_0^t \rho_s \frac{v_{1,s}}{2(v_{0,s})^{\frac{1}{2}}}\mathrm{d}B_s - \int_0^t \frac{\rho_s^2}{2} v_{1,s}\mathrm{d}s)(\rho_t^2+1-\rho_t^2) v_{1,t}\mathrm{d}t \\
&+\int_0^T (\int_0^t (1-\rho_s^2)v_{1,s}\mathrm{d}s)(\rho_t \frac{v_{1,t}}{2(v_{0,t})^{\frac{1}{2}}}\mathrm{d}B_t - \frac{\rho_t^2}{2} v_{1,t}\mathrm{d}t))] \\
=&\int_0^T \mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial xy} v_{1,t}(\int_0^t \rho_s \frac{v_{1,s}}{2(v_{0,s})^{\frac{1}{2}}}\mathrm{d}B_s - \int_0^t \frac{\rho_s^2}{2} v_{1,s}\mathrm{d}s)]\mathrm{d}t \\
&+\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}\int_0^T (\int_0^t (1-\rho_s^2)v_{1,s}\mathrm{d}s)\rho_t^2 v_{1,t}\mathrm{d}t] \\
=&\int_0^T \mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial xy}(v_{1,t}(-\int_0^t \frac{\rho_s^2}{2} v_{1,s}\mathrm{d}s) + \int_0^t \rho_s \frac{v_{1,s}}{2\sqrt{v_{0,s}}} D_s^B v_{1,t}\mathrm{d}s) \\
&+\frac{\partial^3 \tilde{P}_{BS}}{\partial x^2 y} v_{1,t}\int_0^t \frac{\rho_s^2}{2} v_{1,s}\mathrm{d}s]\mathrm{d}t + \mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}\int_0^T (\int_0^t (1-\rho_s^2)v_{1,s}\mathrm{d}s)\rho_t^2 v_{1,t}\mathrm{d}t].
\end{aligned}
$$

From Equation (10.7), one has $D_s^B v_{1,t} = e^{-kt}e^{ks}\xi_s\sqrt{v_{0,s}}$. Hence it is deterministic. Thus, using Identity (10.33) and Lemma 10.5.2 for the first equality and Equation (10.8) for the second equality, one has:

$$
\begin{aligned}
A =&\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}\int_0^T ((\int_0^t \rho_s^2 v_{1,s}\mathrm{d}s)v_{1,t}\mathrm{d}t + (\int_0^t (1-\rho_s^2)v_{1,s}\mathrm{d}s)\rho_t^2 v_{1,t}\mathrm{d}t)] \\
&+\mathbb{E}[\frac{\partial \tilde{P}_{BS}}{\partial y}\int_0^T (\int_0^t \frac{v_{1,s}}{2v_{0,s}} e^{-kt}e^{ks}\xi_s\sqrt{v_{0,s}}\mathrm{d}B_s)\mathrm{d}t] \\
=&\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}\int_0^T ((\int_0^t \rho_s^2 v_{1,s}\mathrm{d}s)v_{1,t}\mathrm{d}t + (\int_0^t (1-\rho_s^2)v_{1,s}\mathrm{d}s)\rho_t^2 v_{1,t}\mathrm{d}t)] + \mathbb{E}[\frac{\partial \tilde{P}_{BS}}{\partial y}\int_0^T \frac{v_{2,t}}{2}\mathrm{d}t].
\end{aligned}
$$

Now, plug this last equality into (10.35) and use the identity

$$
\begin{aligned}
&\int_0^T ((\int_0^t \rho_s^2 v_{1,s}\mathrm{d}s)v_{1,t}\mathrm{d}t + (\int_0^t (1-\rho_s^2)v_{1,s}\mathrm{d}s)\rho_t^2 v_{1,t}\mathrm{d}t) + \frac{1}{2}(\int_0^T (1-\rho_t^2)v_{1,t}\mathrm{d}t)^2 = \\
&\int_0^T ((\int_0^t \rho_s^2 v_{1,s}\mathrm{d}s)v_{1,t}\mathrm{d}t + (\int_0^t (1-\rho_s^2)v_{1,s}\mathrm{d}s)(\rho_t^2+1-\rho_t^2)v_{1,t}\mathrm{d}t) = \\
&\int_0^T ((\int_0^t (\rho_s^2+1-\rho_s^2))v_{1,s}\mathrm{d}s)v_{1,t}\mathrm{d}t = \frac{1}{2}(\int_0^T v_{1,t}\mathrm{d}t)^2;
\end{aligned}
$$

it immediately gives the result.

10.5.3 Proof of Theorem 10.2.1

Proof. **Step 1:** We show the equality

$$
\mathbb{E}[\frac{\partial \tilde{P}_{BS}}{\partial y}\int_0^T (v_{1,t}+v_{2,t})\mathrm{d}t] = \sum_{i=1}^{2} a_{i,T}\frac{\partial^{i+1}P_{BS}(x_0,\int_0^T v_{0,t}\mathrm{d}t)}{\partial x^i y},
$$

where

$$a_{1,T} = \omega_{0,T}^{(\kappa,\rho\xi v_{0,.}),(-\kappa,1)}, \quad a_{2,T} = \omega_{0,T}^{(\kappa,\rho\xi v_{0,.}),(0,\rho\xi),(-\kappa,1)}.$$

Actually, the result is an immediate application of Lemma 10.5.4 and Lemma 10.5.5.
Step 2: We show the equality

$$\frac{1}{2}\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}(\int_0^T v_{1,t}\mathrm{d}t)^2] = \sum_{i=0}^{1} b_{2i,T}\frac{\partial^{2i+2}P_{BS}(x_0,\int_0^T v_{0,t}\mathrm{d}t)}{\partial x^{2i}y^2},$$

where

$$b_{0,T} = \omega_{0,T}^{(2\kappa,\xi^2 v_{0,.}),(-\kappa,1),(-\kappa,1)},$$
$$b_{2,T} = \omega_{0,T}^{(\kappa,\rho\xi v_{0,.}),(-\kappa,1),(\kappa,\rho\xi v_{0,.}),(-\kappa,1)} + 2\omega_{0,T}^{(\kappa,\rho\xi v_{0,.}),(\kappa,\rho\xi v_{0,.}),(-\kappa,1),(-\kappa,1)} = \frac{a_{1,T}^2}{2}.$$

Indeed, one has

$$\begin{aligned}
\frac{1}{2}\mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}(\int_0^T v_{1,t}\mathrm{d}t)^2] &= \mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}\int_0^T(\int_0^t v_{1,s}\mathrm{d}s)v_{1,t}\mathrm{d}t] \\
&= \mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}\int_0^T(\int_t^T e^{-\kappa s}\mathrm{d}s)(e^{\kappa t}v_{1,t}^2\mathrm{d}t + \xi_t\sqrt{v_{0,t}}e^{\kappa t}(\int_0^t v_{1,s}\mathrm{d}s)\mathrm{d}B_t)] \\
&= \mathbb{E}[\frac{\partial^2 \tilde{P}_{BS}}{\partial y^2}\int_0^T(\int_t^T e^{-\kappa s}\mathrm{d}s)e^{\kappa t}v_{1,t}^2\mathrm{d}t] + \mathbb{E}[\frac{\partial^3 \tilde{P}_{BS}}{\partial xy^2}\int_0^T \omega_{t,T}^{(\kappa,\rho\xi v_{0,.}),(-\kappa,1)}v_{1,t}\mathrm{d}t],
\end{aligned}$$

where we have used Lemma 10.5.3 ($f(t) = e^{-\kappa t}$, $Z_t = (\int_0^t v_{1,s}\mathrm{d}s)(e^{\kappa t}v_{1,t})$) for the second equality and Lemmas 10.5.2 and 10.5.3 ($f(t) = (\int_t^T e^{-\kappa s}\mathrm{d}s)\rho_t\xi_t v_{0,t}e^{\kappa t}$, $Z_t = \int_0^t v_{1,s}\mathrm{d}s$) for the last one.
An application of the first and second equality in Lemma 10.5.4 gives the announced result. Actually, it remains to show that $b_{2,T} = a_{1,T}^2/2$. Indeed, consider two càdlàg functions f and $g : [0,T] \to \mathbb{R}$. Then

$$\begin{aligned}
\frac{(\int_0^T f_t(\int_t^T g_s\mathrm{d}s)\mathrm{d}t)^2}{2} &= \frac{\int_0^T\int_0^T f_{t_1}(\int_{t_1}^T g_{t_3}\mathrm{d}t_3)f_{t_2}(\int_{t_2}^T g_{t_4}\mathrm{d}t_4)\mathrm{d}t_2\mathrm{d}t_1}{2} \\
&= \int_0^T f_{t_1}(\int_{t_1}^T\int_{t_1}^T g_{t_3}f_{t_2}(\int_{t_2}^T g_{t_4}\mathrm{d}t_4)\mathrm{d}t_3\mathrm{d}t_2)\mathrm{d}t_1 \\
&= \int_0^T f_{t_1}(\int_{t_1}^T f_{t_2}\int_{t_2}^T\int_{t_2}^T g_{t_3}g_{t_4}\mathrm{d}t_3\mathrm{d}t_4\mathrm{d}t_2 \\
&\qquad + \int_{t_1}^T g_{t_3}\int_{t_3}^T f_{t_2}\int_{t_2}^T g_{t_4}\mathrm{d}t_4\mathrm{d}t_2\mathrm{d}t_3)\mathrm{d}t_1 \\
&= 2\int_0^T f_{t_1}\int_{t_1}^T f_{t_2}\int_{t_2}^T g_{t_3}\int_{t_3}^T g_{t_4}\mathrm{d}t_3\mathrm{d}t_4\mathrm{d}t_2\mathrm{d}t_1 \\
&\qquad + \int_0^T f_{t_1}\int_{t_1}^T g_{t_3}\int_{t_3}^T f_{t_2}\int_{t_2}^T g_{t_4}\mathrm{d}t_4\mathrm{d}t_2\mathrm{d}t_3\mathrm{d}t_1.
\end{aligned}$$

Putting $f(t) = \rho_t\xi_t v_{0,t}e^{kt}$ and $g(t) = e^{-kt}$ in the previous equality readily gives $b_{2,T} = \frac{a_{1,T}^2}{2}$, which finishes the proof. □

10.6 Conclusion

We have established an approximation pricing formula for call/put options in the time dependent Heston models. We prove that the error is of order 3 w.r.t. the volatility of volatility and 2 w.r.t. the maturity. In practice, taking the Fourier method as a benchmark, the accuracy is excellent for a large range of strikes and maturities. In addition, the computational time is about 100 to 1000 times smaller than using an efficient Fourier method.

Following the arguments in Chapter 4, our formula extends immediately to other payoffs depending on S_T (note that the identities (10.6) and (10.33) are valid for any payoff of this type). As explained in Chapter 4, the smoother the payoff, the higher the error order w.r.t. T; the less smooth the payoff, the lower the error order w.r.t. T. For digital options, the error order w.r.t. T becomes $3/2$ instead of 2.

Extensions to exotic options and to the third order expansion formula w.r.t. the volatility of volatility are left for further research.

10.7 Appendix: closed formulas in Heston model

There are few closed representations for the call/put prices written on the asset $S_t = e^{\int_0^t (r_s - q_s)ds} e^{X_t}$ in the Heston model (defined in (10.1) and (10.2)). We focus on the Heston formula [63] and on the Lewis formula [79]. Both of them rely on the knowledge of the characteristic function of the log-asset price $(X_t)_t$ and on Fourier transform-based approaches.

- In [63], Heston obtains a representation in a *Black-Scholes* form:

$$Call_{Heston}(t, S_t, v_t; T, K) = S_t e^{-\int_t^T q_s ds} P_1 - K e^{-\int_t^T r_s ds} P_2,$$

  where both probabilities P_1 and P_2 are equal to a one-dimensional integral of characteristic functions.

- In [79], Lewis takes advantage of the generalized Fourier transform, by using an integration along a straight line in the complex plane parallel to the real axis. It is important to detect the strip where the integration is safe. Lewis suggests the use of complex numbers z such that $\mathscr{I}\mathrm{m}(z) = \frac{1}{2}$. His formula writes

$$Call_{Heston}(t, S_t, v_t; T, K) = S_t e^{-\int_t^T q_s ds} - \frac{K e^{-\int_t^T r_s ds}}{2\pi} \int_{\frac{i}{2}-\infty}^{\frac{i}{2}+\infty} e^{-izX} \phi_T(-z) \frac{dz}{z^2 - iz}$$

  where $X = \log\left(\frac{S_t e^{-\int_t^T q_s ds}}{K e^{-\int_t^T r_s ds}}\right)$ and $\phi_T(z) = \mathbb{E}(e^{z(X_T - X_t)} | \mathscr{F}_t)$. Then, the above integral is evaluated by numerical integration.

Using PDE arguments in combination with affine models, we can obtain an explicit formula for $\phi_T(z)$ in the case of constant Heston parameters. In addition, it can be computed without discontinuities in z, following the arguments in [68]. For piecewise constant parameters, the characterictic function $\phi_T(z)$ can be computed recursively using nested Riccatti equations with constant coefficients: we refer to the work by Mikhailov and Nogel [86].

In our numerical tests, we prefer the Lewis formula which gives better numerical results, in particular for very small or very large strikes, compared to the Heston formula.

Chapter 11

Appendix

11.1 Smile behaviors for Heston model

The aim of this section is to study the impact of Heston model parameters on the smile behaviors through time and for different strikes. This is done first for the constant Heston model and second for the time dependent Heston model.

11.1.1 Constant Heston model

Impact of the correlation. We remark from the figure 11.1 that

- for negative correlation, the center of the short maturity smile is shifted to the right. As the correlation becomes negative, the smile shape changes from a symmetric shape to a negative skew.
- for zero correlation, the smile is symmetric w.r.t. the moneyness. This confirms the property proved in [97].
- for positive correlation, the center of the short maturity smile is shifted to the left. As the correlation becomes positive, the smile changes from a symmetric shape to a positive skew.
- for all the correlations, the smile flattens for long maturity and converges to the value [1] $\sqrt{\theta}$ as it is proved in Chapter 6 page 182 Equation (2.3) of [79].

Impact of the volatility of volatility. We remark from the figure 11.2 that

- for small volatility of volatility, the smile for short maturity is less emphasized and not far from a flat surface.
- when the volatility of volatility increases, the smile for short maturity becomes noticeable and looks like a U shape.

Impact of the long term variance. We remark from the figure 11.3 that

- the value of the long term of variance doesn't impact much the short-term smile.

[1] In this case, one can not observe the convergence towards $\sqrt{\theta}$ since the initial value is already equal to $\sqrt{v_0} = \sqrt{\theta}$. We will be able to observe the convergence phenomenon in the paragraph about the impact of log term variance.

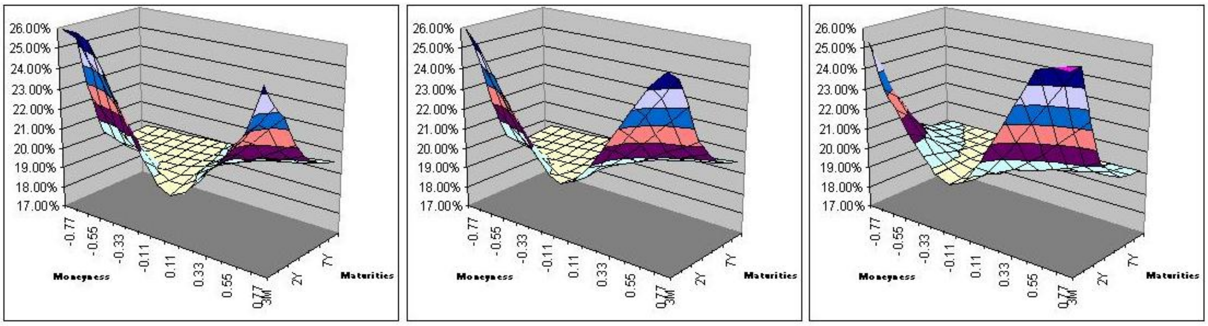


Figure 11.1: Implied Black Scholes volatilities for different correlations (on the left for $\rho = -20\%$, at the center for $\rho = 0\%$, on the right for $\rho = 20\%$) written as a function of log-moneyness and maturities (in fractions of years). Parameters: $x_0 = 0$, $\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$.

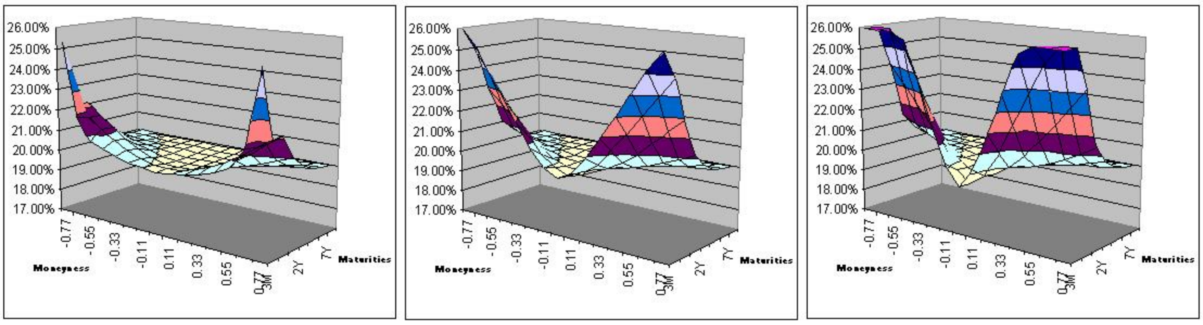


Figure 11.2: Implied Black Scholes volatilities for different volatility of volatilities (on the left for $\xi = 10\%$, at the center for $\xi = 30\%$, on the right for $\xi = 50\%$) written as a function of log-moneyness and maturities (in fractions of years). Parameters: $x_0 = 0$, $\theta = 6\%$, $\kappa = 3$, $\rho = 0\%$.

- this parameter impacts the implied Black Scholes volatility for long maturities. Indeed, as explained before, the limit of implied Black Scholes volatilities is the square root of the long term variance.

Impact of the mean reversion. We remark from the figure 11.4 that

- for higher mean reversion the smile for short maturity is less emphasized and not far from a Flat surface.
- when the mean reversion parameter decreases, the smile for short maturity becomes more important and takes the shape of U.

Indeed, this parameter plays a similar role than the volatility of volatility but in the inverse way. Then we can find different Heston models which reproduce the same smile for one maturity and doesn't have the same parameters. This property has been proved in Chapter 10 Subsection 10.2.6.

11.1.2 Time dependent Heston model

Impact of the time dependent correlation. In the figure 11.5, we plot two surfaces of implied Black Scholes volatilities: the first represents the implied Black Scholes volatilities with constant parameters

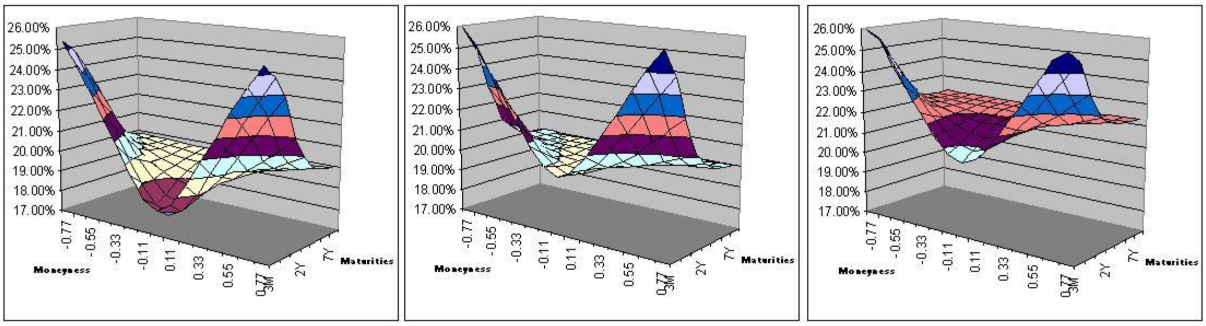


Figure 11.3: Implied Black Scholes volatilities for different long term variances (on the left for $\theta = 2\%$, at the center for $\theta = 4\%$, on the right for $\theta = 5\%$) written as a function of log-moneyness and maturities (in fractions of years). Parameters: $x_0 = 0$, $\rho = 0\%$, $\kappa = 3$, $\xi = 30\%$.

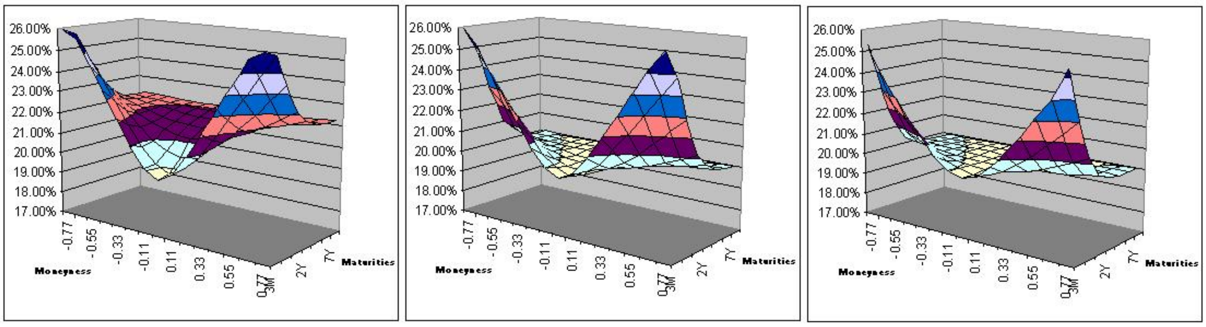


Figure 11.4: Implied Black Scholes volatilities for different mean reversions (on the left for $\kappa = 200\%$, at the center for $\kappa = 300\%$, on the right for $\kappa = 600\%$) written as a function of log-moneyness and maturities (in fractions of years). Parameters: $x_0 = 0$, $\theta = 6\%$, $\rho = 0$, $\xi = 30\%$.

($\theta = 4\%$, $\kappa = 3$, $\xi = 30\%$) and correlation $\rho = 0$. The second one is obtained with the same parameters ($\theta = 6\%$, $\kappa = 3$, $\xi = 30\%$) and piecewise constant correlation ρ equal to $0\% + i \times -2.5\%$ at each interval of the form $]\frac{i}{4}, \frac{i+1}{4}[$. Hence, we remark from the figure 11.5 that:

- the short maturity smile is not much impacted by the correlation since for this example the correlation for short maturity is not far from zero.
- for long maturities, there is an emphasized skew due to negative value of correlation for long maturities.

Impact of the volatility of volatility. In the figure 11.6, we plot two surfaces of implied Black Scholes volatilities: The first represents the implied Black Scholes volatilities with constant parameters ($\theta = 4\%$, $\kappa = 3$, $\rho = 0\%$) and volatility of volatility $\xi = 30\%$. The second one is obtained with the same parameters ($\theta = 4\%$, $\kappa = 3$, $\rho = 0\%$) and piecewise constant volatility of volatility ξ equal to $30\% + i \times 2.5\%$ at each interval of the form $]\frac{i}{4}, \frac{i+1}{4}[$. Hence, we remark from the figure 11.6 that:

- the short maturity smile is not really impacted by the piecewise constant volatility of volatility since it is not far from the initial volatility of volatility.

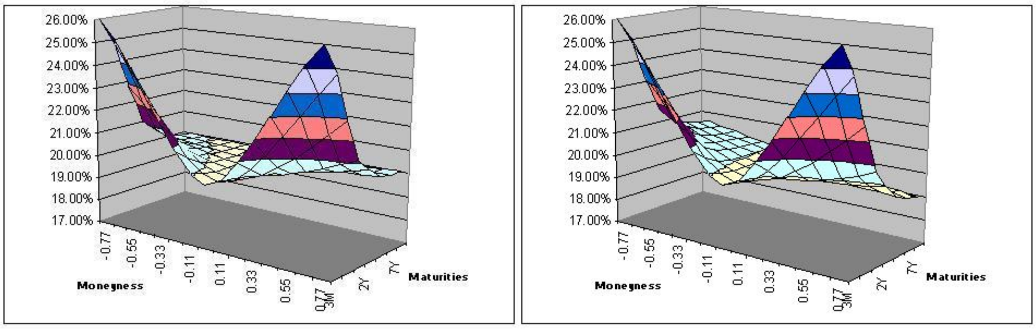


Figure 11.5: Implied Black Scholes volatilities for different correlations (on the left for constant parameters, on the right for piecewise constant correlation) written as a function of log-moneyness and maturities (in fractions of years). Parameters: $x_0 = 0$.

- for long maturities, the flat curve is replaced by a small smile. In this example, the change of the volatility of volatility makes the convergence of the implied Black Scholes volatilities at long maturities slows down.

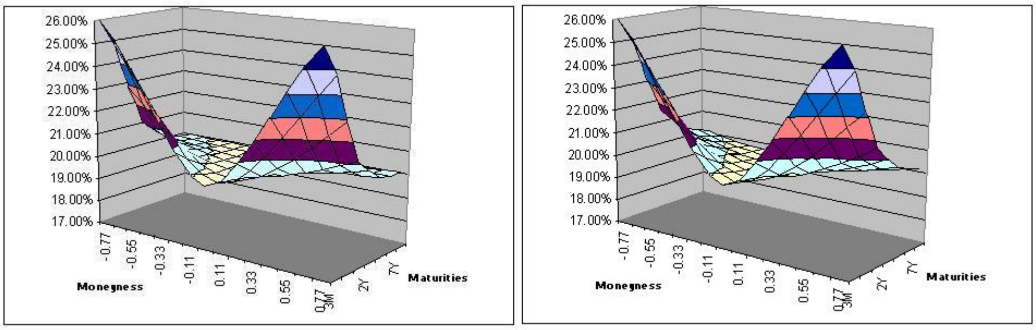


Figure 11.6: Implied Black Scholes volatilities for different volatility of volatilities (on the left for constant parameters, on the right for piecewise constant volatility of volatility) written as a function of log-moneyness and maturities (in fractions of years). Parameters: $x_0 = 0$.

Impact of the long term variance. In the figure 11.7, we plot two surfaces of implied Black Scholes volatilities: The first represents the implied Black Scholes volatilities with constant parameters ($\xi = 30\%$, $\kappa = 3$, $\rho = 0\%$) and long term variance $\theta = 4\%$. The second one is obtained with the same parameters ($\xi = 30\%$, $\kappa = 3$, $\rho = 0\%$) and piecewise constant long term variance θ equal to $4\% + i \times 0.05\%$ at each interval of the form $]\frac{i}{4}, \frac{i+1}{4}[$. Hence, we remark from the figure 11.7 that:

- as before, the short maturity smile is not impacted.

- for long maturities, the surface of the implied Black Scholes volatility is not constant but is shifted and changes with the maturity. Indeed, the implied Black Scholes variance for long maturities changes since the long term variance parameter changes through maturities under our tests assumptions.

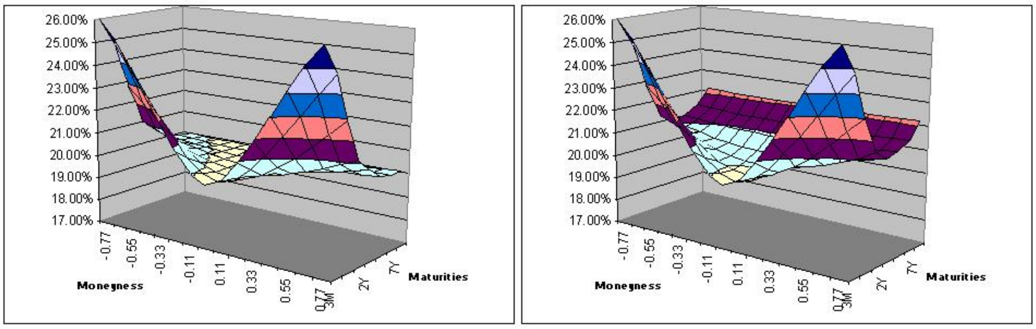


Figure 11.7: Implied Black Scholes volatilities for different long term variances (on the left for constant parameters, on the right for piecewise constant long term variance) written as a function of log-moneyness and maturities (in fractions of years). Parameters: $x_0 = 0$.

11.2 Literature on the negative moments of the integrated CIR process

In this section, we give results about the negative moments for the CIR process and the integrated CIR process. Most of the results are stated in [39] and [24]. Then, we focus our review on these articles and some related results. We extract also some technical Lemmas from these results. These Lemmas are interesting, but their domain of validity is restricted and the estimates are not uniform w.r.t. the volatility of volatility. In order to obtain uniform estimates under not restricted assumptions, we refer to Chapter 10 Subsection 10.4.3.

11.2.1 The integrated square-root process by Dufresne [39]

Dufresene in [39] derives interesting properties for the CIR process:

$$dv_t = \kappa(\theta - v_t)dt + \xi\sqrt{v_t}dB_t, v_0 > 0. \tag{11.1}$$

He calculates the CIR moments $\mathbb{E}[v_t^r]$ recursively using linear ordinary differential equations. By analyzing the Laplace transform of the CIR process he shows that the CIR variable v_t is a summation of two independent variables: a Gamma one plus a compound Poisson variable. He shows that there is a critical lower value of r for which the moment $\mathbb{E}[v_t^r]$ becomes infinite. This critical value is $-\frac{2\kappa\theta}{\xi^2}$. He also consider the integrated CIR process $Y_t = \int_0^t v_s ds$. Analogously, he derives its moments and gives explicit formula for the Laplace transform of the integrated CIR process. Hence, the process has finite moments $\mathbb{E}[Y_t^r]$ for any $r \in \mathbb{R}$. Because of the critical value for the negative moments of the process (v_t), the author conjectures[2] that the Laplace transform $\mathbb{E}[exp(p\int_0^t \frac{1}{v_t}dt)]$ is infinite for every positive index p. The author gives also interesting results which relates the Laplace transform of a random variable $U > 0$ to the Laplace transform of its inverse $\frac{1}{U}$.

[2]The answer to this conjecture is negative, see Lemma 11.2.1 in the next Subsection

11.2.2 An efficient discretization scheme for one dimensional SDEs by Bossy and Diop [24]

In the paper [24], the authors give nice properties about the CIR process and generally every process of the kind [3]

$$dv_t = \kappa(\theta - v_t)dt + \xi v_t^{\alpha} dB_t, v_0. \tag{11.2}$$

This SDE has been proposed by Cox ($\alpha = \frac{1}{2}$) in [35] and by Hull and White in [67] to model the short term rate. They give estimates of the moments of this process for any $p > 0$

$$\mathbb{E}[\sup_{t\in[0,T]} v_t^p] \le C(p)(1 + v_0^p).$$

Without extra effort, we can show that the constant $C(p)$ derived in the proof of Lemma 2.1 [24] is non decreasing on T, κ and ξ. They give under the condition $\kappa\theta > \xi^2$, estimates for the p moment (with $p \in [1, \frac{2\kappa\theta}{\xi^2} - 1[$) of the inverse of the CIR process ($\alpha = \frac{1}{2}$)

$$\mathbb{E}[\frac{1}{v_t^p}] \le \frac{1}{\Gamma(p)}(\frac{2e^{2\kappa t}}{\xi^2 t})^p. \tag{11.3}$$

This is an interesting inequality but it said that these moments may explode when ξ or t goes to 0. The authors show the following Lemma for CIR process ($\alpha = \frac{1}{2}$).

Lemma 11.2.1. *(Lemma A.2 in [24]) If $\kappa\theta \ge \xi^2$, then there exists a constant C depending on κ, θ, ξ and T such that*

$$\mathbb{E}[exp(\frac{\nu^2\xi^2}{8}\int_0^T \frac{dt}{v_t})] \le C(T)(1 + v_0^{-\frac{\nu}{2}}), \tag{11.4}$$

where $\nu = \frac{2\kappa\theta}{\xi^2} - 1$ and the positive constant $C(T)$ is non decreasing function w.r.t. T.

A nice consequence of this inequality is the following Proposition:

Proposition 11.2.1. *If $\kappa\theta \ge \xi^2$, then for every $p \ge 1$ there exists a constant C depending on κ, θ, ξ, v_0, p and T such that*

$$\mathbb{E}[(\int_0^T \frac{dt}{v_t})^p] \le C(T), \tag{11.5}$$

$$\mathbb{E}[(\int_0^T v_t dt)^{-p}] \le \frac{C(T)}{T^{2p}} \tag{11.6}$$

where $C(T)$ is a non decreasing function w.r.t. T.

Proof. Using the inequality $\frac{x^p}{p!} \le e^x$ for the value of $x = \frac{\nu^2\xi^2}{8}\int_0^T \frac{dt}{v_t}$ and the inequality 11.4, one deduces immediately the first inequality. For the second one, apply Cauchy-Schwartz inequality:

$$(\int_0^T dt)^2 \le (\int_0^T v_t dt)(\int_0^T \frac{1}{v_t} dt)$$

and thus

$$(\int_0^T v_t dt)^{-p} \le \frac{1}{T^{2p}}(\int_0^T \frac{1}{v_t} dt)^p,$$

which combined with the first inequality finishes the proof. □

[3] here we chose $b(x) = \kappa(\theta - x)$.

Notice that they give upper bound with precise information on the parameter T, but the dependence on κ, ξ are not monotone. In the following Lemma, we give explicit upper bounds w.r.t. κ, ξ and T.

Lemma 11.2.2. *If $2\kappa\theta \geq \xi^2$, for every $p \in]1, \frac{2\kappa\theta}{\xi^2}[$ there exists a constant C depending only on p and v_0 such that*

$$\mathbb{E}[(\int_0^T \frac{dt}{v_t})^p] \leq C\frac{e^{2p\kappa T}}{\xi^2}T^{p-1}, \tag{11.7}$$

$$\mathbb{E}[(\int_0^T v_t dt)^{-p}] \leq C\frac{e^{2p\kappa T}}{\xi^2}T^{-1-p}. \tag{11.8}$$

Proof. Besides the CIR process v_t is a time transformed Bessel square process:

$$v_t = e^{-\kappa t} z_{\frac{\xi^2(e^{\kappa t}-1)}{4\kappa}} \tag{11.9}$$

One has from Lemma ([38], page 83)

$$\mathbb{E}[\int_0^T \frac{dt}{z_t^p}] \leq \frac{1}{2\Gamma(p)\Gamma(p-1)v_0^{\frac{p-2}{2}}}$$

Using Holder inequality, one has

$$\begin{aligned}(\int_0^T \frac{dt}{v_t})^p &\leq (\int_0^T \frac{dt}{v_t^p})T^{p-1} \\ &\leq C\frac{4e^{2p\kappa T}T^{p-1}}{\xi^2}\int_0^{\frac{\xi^2(e^{\kappa T}-1)}{4\kappa}} \frac{du}{z_u^p},\end{aligned} \tag{11.10}$$

which gives the first inequality. The second inequality is straightforward using Cauchy-Schwartz inequality. □

Part IV

Hybrid models

Chapter 12

Analytical formulas for local volatility model with stochastic rates

This Chapter presents new approximation formulae of European options in a local volatility model with stochastic interest rates. This is a companion Chapter to our work on perturbation methods for local volatility models in Chapter 7 for the case of stochastic interest rates. The originality of this approach is to model the local volatility of the discounted spot and to obtain accurate approximations with tight estimates of the error terms. This approach can also be used in the case of stochastic dividends or stochastic convenience yields. We finally provide numerical results to illustrate the accuracy with real market data.

12.1 Introduction

Long term callable path dependent equity options have generated new modeling challenges as the path dependency requires consistency in the asset diffusion while the early exercise on long period suggests interest rates risk. To appropriately account both for the asset diffusion consistency and the interest rates risk, we consider in this paper a local volatility model with stochastic interest rates. Recent works have mostly focused on extending stochastic volatility models to stochastic interest rates, as described in [91], [14], [11] or [60]. However, very few works have been done on extending local volatility models to stochastic interest rates, except some work on the explicit bias between the local volatility in a stochastic and deterministic interest models [18].

Local volatility enables to infer a diffusion process, which is consistent with the whole volatility surface as explained in [40]. But the introduction of stochastic interest rates makes the calibration process much harder: indeed, the forward PDE approach is now much more computationally expensive as the forward PDE to solve includes an additional stochastic factor due to interest rates. In order to achieve real-time pricing computations, we revisit our perturbation approach (see Chapter 7) to derive approximation formulae in the case of stochastic interest rates.

As a preliminary to our computations, we briefly discuss the choice of the model for the volatility of the spot process. First, owing to the absence of arbitrage, we know that if the spot process $(S_t)_t$ and the interest rate instruments follow Itô-type dynamics, then necessarily

$$\frac{dS_t}{S_t} = r_t dt + \sigma_t dW_t^1,$$
$$r_t = f(0,t) - \int_0^t \gamma(s,t).\Gamma(s,t)ds + \int_0^t \gamma(s,t)dB_s,$$

where $(r_t)_t$ is the short term interest rate, $(f(0,t))_t$ is the forward rate curve at time 0 (deduced from the initial yield curve), $(\sigma_t)_t$ is the instantaneous volatility process, $(\Gamma(t,T))_t$ is the volatility of the zero coupon bond $(B(t,T))_t$ paying 1€ at time T, $\gamma(t,T) = -\partial_T \Gamma(t,T)$ is the volatility of the forward rates (this is the HJM framework). The above general decomposition is written under the risk-neutral measure $\mathbb{Q}$, under which W^1 and $B = (B^1, \cdots, B^n)$ are respectively a linear and a n-dimensional standard Brownian motions. So far, we have not defined the model for $(\sigma_t)_t$, $(\Gamma(t,T)_t)_t$ and the correlation between W^1 and B: this is the topic of the following discussion. If our pricing problem were only for European options, a standard market practice is to model the forward process and to perform a change of measure choosing the forward measure as numéraire. Namely, consider the forward process F_t^T for the maturity date T given by $F_t^T = \frac{S_t}{B(t,T)}$. As shown in [51], the pricing of a European option with final payoff $\varphi(S_T)$ can be reformulated in the forward measures as follows

$$\mathbb{E}[e^{-\int_0^T r_s ds}\varphi(S_T)] = B(0,T)\mathbb{E}_T[\varphi(F_T^T)] \tag{12.1}$$

where $\mathbb{E}_T$ is the expectation under the forward measure $\mathbb{Q}^T$. Interestingly, the forward process $(F_t^T)_{0\le t\le T}$ is a martingale under the forward measure $\mathbb{Q}^T$, meaning that only its volatility needs to be specified (it is known that it is equal to the difference of the volatilities of $(S_t)_t$ and $(B(t,T))_t$ up to correlation factors between W^1 and B). Thus, the equation (12.1) is illuminating as it shows that the stochastic interest rates risk seems to be eliminated from the pricing/calibration problem. However, this approach that models each forward process F^T under $\mathbb{Q}^T$, leads to as many volatility models as the number of maturities (in addition, each volatility model is written under a specific forward measure). Furthermore, in the case of path dependent options, it is not enough to model each F^T under $\mathbb{Q}^T$, since we can not extend the representation (12.1) using only (F_t^T). Thus, we are forced to model the full dynamics of S under $\mathbb{Q}$. This supports the choice of a local volatility model for S with stochastic interest rates. Being inspired by the previous arguments on modeling the forward that is martingale under the suitable forward probability, we choose to define a model on the discounted price process:

$$S_t^d = e^{-\int_0^t r_s ds} S_t$$

which is also a martingale (under $\mathbb{Q}$). We assume that

$$\frac{dS_t^d}{S_t^d} = \sigma^d(t,S_t^d)dW_t^1.$$

Equivalently, we study the log discounted process $X_t = \log(S_t^d) = \log(S_t) - \int_0^t r_s ds$, whose dynamics is

$$dX_t = \sigma(t,X_t)dW_t^1 - \frac{\sigma^2}{2}(t,X_t)dt, \quad X_0 = x_0, \tag{12.2}$$

where $\sigma(t,X_t)$ is the volatility term that can be related to the local volatility of the discounted process $\sigma(t,X_t) = \sigma^d(t,S_t^d)$.

Taking the log discounted process as a local volatility model is not very conventional as the local volatility function is now a function of the log discounted process and not of the log process itself. However, this new approach has the great advantage to remove the influence of stochastic interest rates in the local volatility function and presumably to lead to intuitive approximations. In addition, this approach leads to similar types of local volatility as the one for the forward process. Both approaches model the local volatility functions of martingale processes. For a trader accustomed to quote local volatility for the forward process, it becomes easy to shift to our approach.

To complete our preliminary discussion related to the choice of the model, it remains to specify the assumptions of the volatility of interest rates and the correlation. We consider Gaussian model for interest rates, by assuming that $\Gamma,\gamma : \mathbb{R}^+\times\mathbb{R}^+ \to \mathbb{R}^n$ are deterministic functions (n is the number of Gaussian factors). The Brownian motions W^1 and $B = (B^1,\cdots,B^n)$ are correlated using deterministic functions $(\rho_{i,t}^{S,r})_{i,t}$:

$$d\langle W^1,B^i\rangle_t = \rho_{i,t}^{S,r}dt \quad 1\le i\le n.$$

Now, our aim is to give an analytical accurate approximation of any European option price, written as the expected value under the risk neutral probability measure of a payoff function h evaluated at the maturity time T:

$$A = \mathbb{E}[e^{-\int_0^T r_s ds} h(\int_0^T r_s ds + X_T)], \tag{12.3}$$

where $h(x) = \varphi(e^x)$. The important cases are related to call/put for which $h(x) = (e^x - K)^+$ and $h(x) = (K - e^x)^+$.

Using the zero coupon $B((t,T))_t$ as a numéraire, one has

$$A = B(0,T)\mathbb{E}_T[h(\int_0^T r_s ds + X_T)], \tag{12.4}$$

where $\mathbb{E}_T$ is the expectation under the forward neutral probability $\mathbb{Q}^T$. The process $(X_t)_t$ has the following dynamics under the probability $\mathbb{Q}^T$:

$$dX_t = \sigma(t,X_t)dW_t^{1,T} + (\rho_t^{S,r}.\Gamma(t,T)\sigma(t,X_t) - \frac{\sigma^2}{2}(t,X_t))dt, \quad X_0 = x_0, \tag{12.5}$$

where $(W^{1,T})_t$ is a Brownian motion under $\mathbb{Q}^T$.

Black formula. An important case in our study is associated to time dependent volatility $\sigma(t,x) = \sigma_t$ for which the price is given explicitly by the Black formula for call/put payoffs. This feature relies on the Gaussian property of $\int_0^t r_s ds + X_T$ under $\mathbb{Q}^T$. One has

$$\begin{aligned}\int_0^T r_s ds + X_T = {} & \frac{1}{2}\int_0^T |\Gamma(t,T)|^2 dt - \int_0^T \Gamma(t,T)dB_t + \int_0^T f(0,t)dt \\ & + \log(S_0) + \int_0^T \sigma_t dW_t^1 - \frac{1}{2}\int_0^T \sigma_t^2 dt \\ = {} & \log(\frac{S_0}{B(0,T)}) + \int_0^T \sigma_t dW_t^{1,T} - \int_0^T \Gamma(t,T)dB_t^T \\ & - \frac{1}{2}\int_0^T |\Gamma(t,T)|^2 dt - \frac{1}{2}\int_0^T \sigma_t^2 dt + \int_0^T \sigma_t \rho_t^{S,r}.\Gamma(t,T)dt, \end{aligned} \tag{12.6}$$

where $(W^{1,T}, B^T)$ is a $\mathbb{Q}^T$-Brownian motion (with the same correlation than under $\mathbb{Q}$). Denote by σ^{Black} the equivalent Black volatility defined by

$$(\sigma^{Black})^2 T = \int_0^T [\sigma_t^2 + |\Gamma(t,T)|^2 - 2\sigma_t \rho_t^{S,r}.\Gamma(t,T)]dt.$$

Thus, $\int_0^T r_s ds + X_T$ is $\mathbb{Q}^T$-distributed as a Gaussian r.v. with mean $\log(\frac{S_0}{B(0,T)}) - \frac{1}{2}(\sigma^{Black})^2 T$ and variance $(\sigma^{Black})^2 T$. In particular for call options ($h(x) = (e^x - K)^+$), it follows that

$$A = B(0,T)[\frac{S_0}{B(0,T)}\mathcal{N}(\frac{1}{\sigma^{Black}\sqrt{T}}\log(\frac{S_0}{B(0,T)K}) + \frac{1}{2}\sigma^{Black}\sqrt{T}) - K\mathcal{N}(\frac{1}{\sigma^{Black}\sqrt{T}}\log(\frac{S_0}{B(0,T)K}) - \frac{1}{2}\sigma^{Black}\sqrt{T})].$$

General local volatility models. Then, to obtain the analytical approximation for general local volatility models, we follow the ideas from Chapter 7 and we introduce a parameterized process given by:

$$dX_t^\varepsilon = \varepsilon(\sigma(t,X_t^\varepsilon)dW_t^{1,T} + (\rho_t^{S,r}.\Gamma(t,T)\sigma(t,X_t^\varepsilon) - \frac{\sigma^2}{2}(t,X_t^\varepsilon))dt), X_0^\varepsilon = x_0, \tag{12.7}$$

where the parameter ε lies in the range $[0,1]$. Obviously, this parametrized process is equal to the initial one for $\varepsilon = 1$. Remarkably, it is much easier to calculate the price (12.4) as an expansion formula with respect to ε (this is related to Black formula). Once we have derived all the terms of the expansion, we see that the price of the European option is obtained by taking $\varepsilon = 1$ in the expansion.

Notations. The following notation will be used extensively throughout the paper.

Notation 12.1.1. ***Differentiation.***
If these derivatives have a meaning, we write:

- $\psi_t^{(i)}(x) = \frac{\partial^i \psi}{\partial x^i}(t,x)$ *for any function* ψ *of two variables.*
- $\sigma_t = \sigma(t,x_0), \sigma_t^{(i)} = \sigma^{(i)}(t,x_0)$.
- $X_{i,t}^{\varepsilon} = \frac{\partial^i X_t^{\varepsilon}}{\partial \varepsilon^i}$ *is the* i^{th} *derivative of the parameterized process with respect to* ε.
- $X_{i,t} = \frac{\partial^i X_t^{\varepsilon}}{\partial \varepsilon^i}|_{\varepsilon=0}$. *These processes play a crucial role in this work.*

The following notation of Greeks will be useful for interpreting the expansion terms.

Notation 12.1.2. ***Greeks.***
Let Z be a random variable. Given a payoff function h, we define the i^{th} *Greek for the variable Z by the quantity (if it has a meaning) :*

$$\text{Greek}_i^h(Z) = \frac{\partial^i \mathbb{E}_T[h(\int_0^T r_s ds + Z + x)]}{\partial x^i}|_{x=0}.$$

Assumptions. In order to derive tight upper bounds for our expansions, we assume that the coefficient σ is smooth enough. In what follows, N is an integer greater than 4.

- ***Assumption*** *(R_N). The function* σ *is bounded and of class* C^N *w.r.t x. Its derivatives up to order N are bounded.*

This assumption may be restrictive because σ has to be bounded as well its derivatives. Actually, this statement is made only to simplify a bit our analysis, but we can prove that our approximation remains valid if some boundedness requirements are partially relaxed.

Notation 12.1.3. ***Function amplitudes.***
Under (R_N)*, we set*

$$M_0 = max(|\sigma|_\infty, \cdots, |\sigma^{(N)}|_\infty), \tag{12.8}$$

$$M_1 = max(|\sigma^{(1)}|_\infty, \cdots, |\sigma^{(N)}|_\infty). \tag{12.9}$$

Although M_0 *and* M_1 *may depend on N, we remove this dependence in our notation, for the sake of simplicity. In our expansion, we expect these quantities to be small.*

Remark 12.1.1. *The constant* M_0 *measures the amplitude of the objective function* σ *and its derivatives w.r.t. the second variable, whereas* M_1 *measures only the amplitude of its derivatives. Notice that* $M_1 \leq M_0$ *and in case of deterministic function* σ*, one has* $M_1 = 0$.

In real market, the correlation between the asset and the short rate is close to zero (see table (12.1)). Therefore, the following assumption is consistent with real market data:

- ***Assumption*** *(Rho). The asset is not perfectly correlated (positively or negatively) to the interest rate:*

$$|\rho^{S,r}|_\infty = \sup_{t\in[0,T]} |\rho_t^{S,r}| < 1.$$

To perform the infinitesimal analysis, we rely on smoothness properties not related to the payoff function itself but rather to the law of the underlying stochastic models.

Table 12.1: Historical correlation between assets and short term interest rate EUR. Period: 23-Sep-2007 to 22-Sep-09.

Asset	Historical correlation
ADIDAS	18.32%
BELGACOM	4.09%
CARREFOUR	7.08%
DAIMLER	-0.94%
DANONE	7.23%
LVMH	4.53%
NOKIA	4.37%
PHILIPS	5.23%

- ***Assumption*** *(E). The function σ does not vanish and its oscillation is bounded, meaning* $1 \leq \frac{|\sigma|_\infty}{\sigma_{inf}} \leq C_E$ *where* $\sigma_{inf} = \inf_{(t,x)\in\mathbb{R}^+\times\mathbb{R}} \sigma(t,x)$.

The assumption (E) is commonly called an ellipticity assumption.

Definitions.

Definition 12.1.1. *As usual, we define $\mathscr{C}_0^\infty(\mathbb{R})$ as the space of real infinitely differentiable functions h with compact support. We also define $\mathscr{H}$ as the space of functions with exponential growth.*

As in Chapter 7, our analysis depend on the payoff smoothness. We split our analysis into three cases.

- ***Assumption*** *(H_1). h belongs to $\mathscr{C}_0^\infty(\mathbb{R})$. This case corresponds to smooth payoffs.*
- ***Assumption*** *(H_2). h and $h^{(1)}$ belongs to $\mathscr{H}$. This case corresponds to vanilla options (call/put).*
- ***Assumption*** *(H_3). h belongs to $\mathscr{H}$. This is the case of binary options (digital).*

12.2 Smart Taylor Development

Our perturbation approach relies on the Taylor expansion of the parameterized process. We have paved the way in our previous work (see Chapter 7). In the quoted reference, the parameterized process has the form

$$dX_t^\varepsilon = \varepsilon(\mu(t,X_t^\varepsilon)dt + \sigma(t,X_t^\varepsilon)dW_t)$$

and the aim was to approximate $\mathbb{E}[h(X_T^1)]$. Hence, compared to the current study, we take a specific form for μ, namely $\mu(t,x) = \rho_t^{S,t}.\Gamma(t,T)\sigma(t,x) - \frac{\sigma^2(t,x)}{2}$: with this respect, the expansion on the process (X_t^ε) is very similar to that of Chapter 7. On the other hand, in our case, the quantity of interest is $\mathbb{E}_T[h(\int_0^T r_s ds + X_T^1)]$ and the extra term $\int_0^T r_s ds$ induces significant differences when the correction terms are computed. For the convenience of the reader, we briefly expose the computations when similar to Chapter 7, and we detail the arguments when new compared to Chapter 7.

From the definitions, $X_{i,t} \equiv \frac{\partial^i X_t^\varepsilon}{\partial \varepsilon^i}|_{\varepsilon=0}$, we can expand the perturbed process X_T^ε as follows:

$$X_T^\varepsilon = X_T^\varepsilon|_{\varepsilon=0} + \varepsilon X_{1,T} + \frac{\varepsilon^2}{2!} X_{2,T} + \dots \tag{12.10}$$

Indeed, under the assumption (R_5), almost surely for any t, X_t^ε is C^4 w.r.t ε (see Theorem 2.3 in [72]). The diffusion dynamics of $(X_{i,t}^\varepsilon \equiv \frac{\partial^i X_t^\varepsilon}{\partial \varepsilon^i})_{t\geq 0}$ is obtained by a straight differentiation of the parameters of the diffusion equation of X^ε. The first order term $X_{1,t}^\varepsilon$ is easily obtained as follows:

$$\begin{aligned} dX_{1,t}^\varepsilon =& \sigma_t(X_t^\varepsilon)dW_t^{1,T} + (\rho_t^{S,r}.\Gamma(t,T)\sigma_t(X_t^\varepsilon)) - \frac{\sigma_t^2(X_t^\varepsilon)}{2})dt \\ &+ \varepsilon X_{1,t}^\varepsilon(\sigma_t^{(1)}(X_t^\varepsilon)dW_t^{1,T} + (\rho_t^{S,r}.\Gamma(t,T)\sigma_t^{(1)}(X_t^\varepsilon) - \sigma_t(X_t^\varepsilon)\sigma_t^{(1)}(X_t^\varepsilon))dt), X_{1,0}^\varepsilon = 0. \end{aligned} \tag{12.11}$$

From the definitions, we have $\sigma_t \equiv \sigma(t,x_0)$ and $\sigma_t^{(i)} \equiv \sigma^{(i)}(t,x_0)$. Then, we obtain

$$dX_{1,t} = \sigma_t dW_t^{1,T} + (\rho_t^{S,r}.\Gamma(t,T)\sigma_t - \frac{\sigma_t^2}{2})dt, X_{1,0} = 0, \tag{12.12}$$

$$dX_{2,t} = 2X_{1,t}(\sigma_t^{(1)}dW_t^{1,T} + (\rho_t^{S,r}.\Gamma(t,T)\sigma_t^{(1)} - \sigma_t\sigma_t^{(1)})dt), X_{2,0} = 0. \tag{12.13}$$

Applying the expansion (12.10) at $\varepsilon = 1$, we conclude that $x_0 + X_{1,T}$ is a proxy for X_T. It follows the notation:

$$X_T^B = x_0 + X_{1,T} = x_0 + \int_0^T (\rho_s^{S,r}.\Gamma(s,T)\sigma_s - \frac{\sigma_s^2}{2})ds + \int_0^T \sigma_s dW_s^{1,T}, \tag{12.14}$$

where the exponent B stands for Black, which is the proxy. To obtain an approximation formula as in Chapter 7, we assume that h is smooth and then, we obtain approximations which are valid even if h is not smooth, which allows us to handle finally the case of arbitrary payoffs. Use the Taylor formula twice: first, for X_T^1 at the second order w.r.t ε around x_0, secondly for smooth function h at the first order w.r.t x around X_T^B. This leads to:

$$\begin{aligned} A = B(0,T)\mathbb{E}_T[h(\int_0^T r_s ds + X_T)] &= B(0,T)\mathbb{E}_T[h(\int_0^T r_s ds + X_T^B + \frac{X_{2,T}}{2} + \dots)] \\ &= B(0,T)(\mathbb{E}_T[h(\int_0^T r_s ds + X_T^B)] + \mathbb{E}_T[h^{(1)}(\int_0^T r_s ds + X_T^B)\frac{X_{2,T}}{2}] + \dots). \end{aligned}$$

Note that the first term is explicit for call/put options since it is given by the Black formula previously mentioned. To achieve a fully explicit formula, it remains to transform the correction term involving $X_{2,T}$ into a summation of Greeks computed in the Black proxy. This is performed using the Malliavin calculus.

Theorem 12.2.1. ***(Second order approximation price formula).***
Assume that the model fulfills (R_5), (E) and (Rho), and that the payoff function fulfills one of the assumptions (H_1), (H_2) or (H_3). Then

$$\mathbb{E}[e^{-\int_0^T r_s ds}h(\int_0^T r_s ds + X_T)] = B(0,T)(\mathbb{E}_T[h(\int_0^T r_s ds + X_T^B)] + \sum_{i=1}^3 \alpha_{i,T}\text{Greek}_i^h(\int_0^T r_s ds + X_T^B) + \text{Resid}_2), \tag{12.15}$$

where

$$\begin{aligned}
\alpha_{1,T} =& -\int_0^T (\rho_t^{S,r}.\Gamma(t,T)\sigma_t - \frac{\sigma_t^2}{2})(\int_t^T a_s \sigma_s^{(1)} ds)dt,\\
\alpha_{2,T} =& -\alpha_{1,T} - \alpha_{3,T},\\
\alpha_{3,T} =& \int_0^T a_t \sigma_t (\int_t^T a_s \sigma_s^{(1)} ds)dt,\\
a_t =& \sigma_t - \rho_t^{S,r}.\Gamma(t,T).
\end{aligned}$$

Additionally, estimates of the error term Resid_2 *is analysed according to the payoff smoothness.*

- *For smooth payoff (assumption (H_1)), one has:*

$$|\mathrm{Resid}_2| \le C \sup_{1\le j\le \lfloor \frac{m}{2}\rfloor} |h^{(j)}|_\infty M_1 M_0^2 (\sqrt{T})^3.$$

- *For vanilla payoff (assumption (H_2)), one has:*

$$\begin{aligned}
|\mathrm{Resid}_2| \le& C(\|h^{(1)}(\int_0^T r_s ds + X_T^B)\|_2 + \sup_{v\in[0,1]} \|h^{(1)}(\int_0^T r_s ds + vX_T + (1-v)X_T^B)\|_2)\\
&\frac{M_0}{\sigma_{inf}\sqrt{1-|\rho^{S,r}|_\infty^2}} M_1 M_0^2 (\sqrt{T})^3.
\end{aligned}$$

- *For binary payoff (assumption (H_3)), one has:*

$$\begin{aligned}
|\mathrm{Resid}_2| \le& C(\|h(\int_0^T r_s ds + X_T^B)\|_2 + \sup_{v\in[0,1]} \|h(\int_0^T r_s ds + vX_T + (1-v)X_T^B)\|_2)\\
&(\frac{M_0}{\sigma_{inf}\sqrt{1-|\rho^{S,r}|_\infty^2}})^2 M_1 M_0 (\sqrt{T})^2.
\end{aligned}$$

In the above estimates, the constant C depends (in an increasing way) on the bounds of the model parameters and the maturity, and the norm $\|.\|_2$ is the L_2 norm under the probability measure $\mathbb{Q}^T$.

The proof of the above Theorem is postponed to Subsection 12.5.2.

Remark 12.2.1. *The above approximation is a summation of the leading term and a combination of some Greeks of the leading term:*

1. $B(0,T)\mathbb{E}_T[h(\int_0^T r_s ds + X_T^B)]$ *is the leading order, corresponding to the price when the parameters σ is deterministic. In the case of call/put option, it is given by the Black formula previously mentioned. For other payoffs, we can use numerical integration because the law of the random variable $\int_0^T r_s ds + X_T^B$ is Gaussian with known parameters.*

2. $B(0,T)\mathrm{Greek}_i^h(\int_0^T r_s ds + X_T^B)$ *is the i^{th} derivative of the leading term w.r.t. the initial value* $x_0 = \log(S_0)$.

3. *The coefficients $\alpha_{i,T}$ are explicit and depend on the function σ, its derivative at the point x_0, the zero coupon volatility Γ, its derivative γ and the correlation $\rho^{S,r}$. In the next subsection, these constants will be made specified for some important case to enlighten more their simple expressions w.r.t. the model parameters.*

As in Chapter 7, the accuracy of the approximation has different justifications which can be related from the error estimates on Resid_2. Either σ is only time dependent ($M_1 = 0$) and the formula is exact (Black formula). Or the shorter the maturity T or the smaller the volatility (measured by M_0), the more the accurate the approximation.

12.2.1 The case of the one factor Hull and white model plus time homogeneous diffusion

Here, we consider the case of $\gamma_1(t,T) = \xi e^{-\kappa(T-t)}$, a constant correlation ρ and a homogeneous volatility $\sigma(t,x) = \sigma(x)$. Then here $\sigma(t,x_0) = \sigma(x_0) \equiv \sigma$ and $\sigma^{(0,1)}(t,x_0) = \sigma^{(1)}(x_0) \equiv \sigma^{(1)}$. Using Mathematica, we can compute exactly the correction coefficients. Their expressions are

$$\begin{aligned}
\alpha_{1,T} =& \frac{e^{-2\kappa T}\sigma\sigma^{(1)}}{4\kappa^4}(2\rho^2\xi^2 + 2e^{\kappa T}\rho(\kappa\sigma(2\kappa T+1) + 2\rho(\kappa T-1)\xi)\xi \\
&+ e^{2\kappa T}\left(\sigma^2 T^2\kappa^4 + \rho\sigma(\kappa T(3\kappa T-2)-2)\xi\kappa + 2\rho^2(\kappa T-1)^2\xi^2\right), \\
\alpha_{2,T} =& -\alpha_{1,T} - \alpha_{3,T}, \\
\alpha_{3,T} =& \frac{e^{-2\kappa T}\sigma\sigma^{(1)}\left(\rho\xi + e^{\kappa T}\left(\sigma T\kappa^2 + \rho T\xi\kappa - \rho\xi\right)\right)^2}{2\kappa^4}.
\end{aligned}$$

12.2.2 Third order approximation formula

Notice also that in real market the amplitude of volatility of the Hull and White model is $\xi \approx 1\%$ while for the asset $\sigma \approx 20\%$. Therefore, one has presumably $|\Gamma(.)|_\infty = O(M_0^2)$. Thus, we expect that the third order approximation formula w.r.t. M_0 does not yield additional interest rate corrections. The proof follows the arguments of Chapter 7 and we skip the details. Therefore, in the following higher order, we neglect terms related to additional stochastic rate corrections.

Theorem 12.2.2. ***(Third order approximation price formula).***
Assume that the model fulfills (R_7), (E) and (Rho) with $|\Gamma(.)|_\infty = O(M_0^2)$, and that the payoff function is a vanilla payoff (assumption (H_2)). Then

$$\mathbb{E}[e^{-\int_0^T r_s ds}h(\int_0^T r_s ds + X_T)] = B(0,T)(\mathbb{E}_T[h(\int_0^T r_s ds + X_T^B)] + \sum_{i=1}^{6}\beta_{i,T}\mathrm{Greek}_i^h(\int_0^T r_s ds + X_T^B) + \mathrm{Resid}_3'), \tag{12.16}$$

where

$$\begin{aligned}
\beta_{1,T} =& \alpha_{1,T} - \frac{C_{2,T}}{2} - \frac{C_{3,T}}{2} - \frac{C_{4,T}}{4} - \frac{C_{5,T}}{4} - \frac{C_{6,T}}{2}, \\
\beta_{2,T} =& \alpha_{2,T} + \frac{C_{2,T}}{2} + \frac{C_{3,T}}{2} + \frac{5C_{4,T}}{4} + \frac{5C_{5,T}}{4} + \frac{7C_{6,T}}{2} + \frac{C_{7,T}}{2} + \frac{C_{8,T}}{4}, \\
\beta_{3,T} =& \alpha_{3,T} - 2C_{4,T} - 2C_{5,T} - 6C_{6,T} - 3C_{7,T} - \frac{3C_{8,T}}{2}, \\
\beta_{4,T} =& C_{4,T} + C_{5,T} + 3C_{6,T} + \frac{13C_{7,T}}{2} + \frac{13C_{8,T}}{4}, \\
\beta_{5,T} =& -6C_{7,T} - 3C_{8,T}, \\
\beta_{6,T} =& 2C_{7,T} + C_{8,T},
\end{aligned}$$

and

$$C_{2,T}=\omega(\sigma^2,(\sigma^{(1)})^2)_0^T,\qquad C_{3,T}=\omega(\sigma^2,\sigma\sigma^{(2)})_0^T,\qquad C_{4,T}=\omega(\sigma^2,\sigma^2,(\sigma^{(1)})^2)_0^T,$$
$$C_{5,T}=\omega(\sigma^2,\sigma^2,\sigma\sigma^{(2)})_0^T,\qquad C_{6,T}=\omega(\sigma^2,\sigma\sigma^{(1)},\sigma\sigma^{(1)})_0^T,$$
$$C_{7,T}=\omega(\sigma^2,\sigma^2,\sigma\sigma^{(1)},\sigma\sigma^{(1)})_0^T,\qquad C_{8,T}=\omega(\sigma^2,\sigma\sigma^{(1)},\sigma^2,\sigma\sigma^{(1)})_0^T.$$

In the above definition of the constants, the notation $\omega()$ is defined by

$$\omega(f_1,\cdots,f_k)=\int_0^T f_1(r_1)\int_{r_1}^T f_2(r_2)\cdots\int_{r_{k-1}}^T f_k(r_k)dr_1\cdots dr_k,\quad k\geq 0.$$

In addition, the error term $\text{Resid}_3^{'}$ *is estimated as follows*

$$\begin{aligned}|\text{Resid}_3^{'}|\leq & C(\|h^{(1)}(\int_0^T r_sds+X_T^B)\|_2+\sup_{v\in[0,1]}\|h^{(1)}(\int_0^T r_sds+vX_T+(1-v)X_T^B)\|_2)\\ & (\frac{M_0}{\sigma_{inf}\sqrt{1-|\rho^{S,r}|_\infty^2}})^2M_1M_0^3(\sqrt{T})^4\\ & +C\|h^{(1)}(\int_0^T r_sds+X_T^B)\|_2(\frac{M_0}{\sigma_{inf}\sqrt{1-|\rho^{S,r}|_\infty^2}})^5M_1M_0^3(\sqrt{T})^3,\end{aligned}$$

where the constant C depends (in an increasing way) on the bounds of the model parameters and the maturity.

Proof. Using an adaptation of Theorem 7.4.2 in Chapter 7, one has

$$\begin{aligned}\mathbb{E}[e^{-\int_0^T r_sds}h(\int_0^T r_sds+X_T)]=&B(0,T)(E[h(\int_0^T r_sds+X_T^B)]+E[h^{(1)}(\int_0^T r_sds+X_T^B)\frac{X_{2,T}}{2}]\\&+\mathbb{E}[h^{(1)}(\int_0^T r_sds+X_T^B)\frac{X_{3,T}}{3!}]+\mathbb{E}[h^{(2)}(\int_0^T r_sds+X_T^B)\frac{(\frac{X_{2,T}}{2})^2}{2}]+\text{Resid}_3).\end{aligned}$$

The error Resid_3 is estimated using an adaptation of Theorem 7.4.2 by :

$$\begin{aligned}|Resid_3|\leq & C(\|h^{(1)}(\int_0^T r_sds+X_T^B)\|_2+\sup_{v\in[0,1]}\|h^{(1)}(\int_0^T r_sds+vX_T+(1-v)X_T^B)\|_2)\\ & (\frac{M_0}{\sigma_{inf}\sqrt{1-|\rho^{S,r}|_\infty^2}})^2M_1M_0^3(\sqrt{T})^4.\end{aligned}$$

The first correction term $E[h^{(1)}(\int_0^T r_sds+X_T^B)\frac{X_{2,T}}{2}]$ is made explicit in Theorem 12.2.1. The other correction terms $\mathbb{E}[h^{(1)}(\int_0^T r_sds+X_T^B)\frac{X_{3,T}}{3!}]+\mathbb{E}[h^{(2)}(X_T^B)\frac{(\frac{X_{2,T}}{2})^2}{2}]$ are computed using integration by parts (Malliavin calculus) and a truncation argument of the weights. The truncation argument consists in neglecting the additional weights which are related to Γ since $|\Gamma(.)|_\infty=O(M_0^2)$. Hence, the computation of these corrections terms is reduced to the computation of the additional weights of the third order correction by taking $\Gamma\equiv 0$ (this is done in Theorem 7.2.2 of Chapter 7). By easy but tedious computations that are not detailed, we can prove that the related truncation error is estimated by:

$$|\text{TruncationError}|\leq C\|h^{(1)}(\int_0^T r_sds+X_T^B)\|_2(\frac{M_0}{\sigma_{inf}\sqrt{1-|\rho^{S,r}|_\infty^2}})^5M_1M_0^3(\sqrt{T})^3.$$

Note that the above error does not decrease to 0 as quickly as before w.r.t. the maturity T (the power is equal to 3 instead of 4).

□

Remark 12.2.2. *In the case of homogeneous volatility $\sigma(t,x) = \sigma(x)$. We write $\sigma(t,x_0) \equiv \sigma$, $\sigma^{(0,i)}(t,x_0) \equiv \sigma^{(i)}$. Then*

$$C_{2,T} = \sigma^2(\sigma^{(1)})^2\frac{T^2}{2}, \qquad C_{3,T} = \sigma^3\sigma^{(2)}\frac{T^2}{2},$$
$$C_{4,T} = \sigma^4(\sigma^{(1)})^2\frac{T^3}{6} = C_{6,T}, \qquad C_{5,T} = \sigma^5\sigma^{(2)}\frac{T^3}{6},$$
$$C_{7,T} = \sigma^6(\sigma^{(1)})^2\frac{T^4}{24}, \qquad C_{8,T} = C_{7,T}.$$

12.3 Extension to stochastic dividend and convenience yield

The current framework can be easily adapted to deal with stochastic dividends or stochastic convenience yield in a local volatility model applied to commodity models. This can be seen as an extension to Gibson Schwartz model to handle local volatility functions for example. We recall the SDE of the underlying spot in the Gibson Schwartz model [52]:

$$\begin{aligned}\frac{dS_t}{S_t} &= (r_t - y_t)dt + \sigma dW_t^1,\\ dy_t &= \kappa(\alpha_t - y_t)dt + \xi_t dW_t^2,\\ d\langle W^1, W^2\rangle_t &= \rho_t dt.\end{aligned}$$

Here, the interest rate $(r_t)_t$ is deterministic. $(\alpha_t)_t$ and $(\xi_t)_t$ are time dependent functions. Therefore, using similar modeling like for stochastic rates, we have the following framework:

$$\begin{aligned}dX_t &= \sigma(t,X_t)dW_t^1 - \frac{\sigma^2(t,X_t)}{2}dt,\\ dy_t &= \kappa(\alpha_t - y_t)dt + \xi_t dW_t^2,\\ d\langle W^1, W^2\rangle_t &= \rho_t dt,\end{aligned}$$

where $S_t = e^{X_t}e^{\int_0^t(r_s - y_s)ds}$ and $\sigma(t,x)$ is the local volatility function. Hence, our aim is to estimate:

$$e^{-\int_0^T r_s ds}\mathbb{E}[h(\int_0^T (r_s - y_s)ds + X_T)] \tag{12.17}$$

Analogously, the proxy X_t^B is:

$$dX_t^B = \sigma_t dW_t^1 - \frac{\sigma_t^2}{2}dt, X_0^B = x_0.$$

Hence, we obtain analogous corrections results:

Theorem 12.3.1. ***(Second order approximation price formula)***.
Assume that the model fulfills (R_5), (E) *and* (Rho), *and that the payoff function fulfills one of the assumptions* (H_1), (H_2) *or* (H_3). *Then*

$$e^{-\int_0^T r_s ds}\mathbb{E}[h(\int_0^T (r_s-y_s)ds+X_T)] = e^{-\int_0^T r_s ds}(\mathbb{E}[h(\int_0^T (r_s-y_s)ds+X_T^B)] + \sum_{i=1}^{3}\lambda_{i,T}\mathrm{Greek}_i^h(\int_0^T (r_s-y_s)ds+X_T^B)+\mathrm{Resid}_2), \tag{12.18}$$

where

$$\begin{aligned}
\lambda_{1,T} =& \int_0^T \frac{\sigma_t^2}{2}(\int_t^T \sigma_s\sigma_s^{(1)}ds)dt,\\
\lambda_{2,T} =& -\int_0^T b_t\sigma_t(\int_t^T \sigma_s\sigma_s^{(1)}ds)dt - \int_0^T \frac{\sigma_t^2}{2}(\int_t^T b_s\sigma_s^{(1)}ds)dt,\\
\lambda_{3,T} =& \int_0^T b_t\sigma_t(\int_t^T b_s\sigma_s^{(1)}ds)dt,\\
b_t =& \sigma_t - \rho_t\xi_t\int_t^T e^{-\kappa(s-t)}ds.
\end{aligned}$$

The error term Resid_2 *is estimated as in Theorem 12.2.1.*

12.4 Numerical Experiments

Here we give numerical examples for the accuracy of our approximation formula. As a benchmark, we use Monte Carlo methods with a variance reduction technique. We consider the one factor Hull and White model for interest rates, the CEV diffusion for the spot and constant correlation ρ. Then,

$$\gamma(t,T) = \xi e^{-\kappa(T-t)}, \sigma(t,x) = \nu e^{(\beta-1)x},$$

In this case the correction coefficients are computed in Paragraph 12.2.2. We consider the call $h(x) = (e^x - K)^+$, ensuring that the price and the Greeks in the Black proxy are explicit.

12.4.1 Monte Carlo with Control variate

Using the HJM framework for the Hull and White short rate (r_t), the integrated $\int_0^T r_s ds$ is a Gaussian variable with mean m and variance v (see [26])

$$\begin{aligned}
m =& \int_0^T f(0,t)dt + \frac{\xi^2}{2\kappa^2}(T+\frac{2}{\kappa}(e^{-\kappa T}-1)-\frac{1}{2\kappa}(e^{-2\kappa T}-1)),\\
v =& \frac{\xi^2}{\kappa^2}(T+\frac{2}{\kappa}e^{-\kappa T}-\frac{1}{2\kappa}e^{-2\kappa T}-\frac{3}{2\kappa}).
\end{aligned}$$

The simulated random variable is $e^{-\int_0^T r_t dt}(e^{\int_0^T r_t dt+X_T}-K)^+$. In order to reduce the statistical error, we use a control variate method. Namely, the control variate is $e^{-\int_0^T f(0,t)dt}(e^{\int_0^T f(0,t)dt+X_T}-K)^+$-$\mathbb{E}[e^{-\int_0^T f(0,t)dt}(e^{\int_0^T f(0,t)dt+X_T}-K)^+]$. The latter expectation is approximated analytically using the third order accurate formula using lognormal proxy approximation derived in Chapter 7.

We take null forward rates ($f(0,t)=0$). Indeed, this choice is arbitrary, and it does not influence the accuracy or the correction terms calculus.

12.4.2 Accuracy of the second and third order approximation formulas (12.15), (12.16)

In Tables 12.2, 12.3, 12.4, 12.5, 12.6, 12.7, 12.8, 12.9 (corresponding to the maturities $6M$, $1Y$, $5Y$, $10Y$ for a small and a large skew), we give detailed numerical results about the accuracy of the second order formula (12.15) and the third order formula (12.16). MC- and MC+ are the bounds of the 95%-confidence interval of the Monte Carlo estimator. Remark also that we increase the range of strike according to maturity in order to test our approximation formula for real quoted strikes. Therefore, we see that our formula (12.15) is very accurate (errors in implied volatilities are smaller[1] than few bps) for β close to 1. For various values of β, we remark that our third order formula (12.16) is extremely accurate.

Table 12.2: Implied Black-Scholes volatilities for the second order formula (12.15), the third order formula (12.16) and the Monte Carlo simulations (3×10^6 simulations using Euler scheme with 20 time steps) expressed as a function of strikes at the expiry $T = 6M$. Parameters: $\beta = 0.8$, $\nu = 0.2$, $\xi = 0.7\%$, $\kappa = 1\%$, $\rho = 15\%$ and $x_0 = 0$.

Strikes	80%	90%	100%	110%	120%
Second Order formula	20.46%	20.24%	20.03%	19.84%	19.66%
Third Order formula	20.48%	20.24%	20.03%	19.84%	19.67%
MC with control variate	20.48%	20.24%	20.03%	19.84%	19.67%
MC-	20.28%	20.18%	20.00%	19.82%	19.64%
MC+	20.68%	20.31%	20.07%	19.87%	19.70%

Table 12.3: Implied Black-Scholes volatilities for the second order formula (12.15), the third order formula (12.16) and the Monte Carlo simulations (3×10^6 simulations using Euler scheme with 20 time steps) expressed as a function of strikes at the expiry $T = 6M$. Parameters: $\beta = 0.2$, $\nu = 0.2$, $\xi = 0.7\%$, $\kappa = 1\%$, $\rho = 15\%$ and $x_0 = 0$.

Strikes	80%	90%	100%	110%	120%
Second Order formula	21.65%	20.86%	20.03%	19.26%	18.47%
Third Order formula	21.87%	20.89%	20.04%	19.29%	18.62%
MC with control variate	21.87%	20.90%	20.04%	19.29%	18.62%
MC-	21.70%	20.83%	20.01%	19.27%	18.60%
MC+	22.04%	20.96%	20.08%	19.32%	18.65%

12.5 Appendix

Here, we bring together the results (and their proofs) which allow us to derive the explicit terms in the formula (12.15). In the following, (u_t) (resp. (v_t)) is a square integrable and predictable (resp. deterministic) process and l is a smooth function with compact support. We recall that $a_t = \sigma_t - \rho_t^{S,r}.\Gamma(t,T)$.

[1] 1 bp on implied volatilities is equal to 0.01%.

Table 12.4: Implied Black-Scholes volatilities for the second order formula (12.15), the third order formula (12.16) and the Monte Carlo simulations (3×10^6 simulations using Euler scheme with 20 time steps) expressed as a function of strikes at the expiry $T = 1Y$. Parameters: $\beta = 0.8$, $\nu = 0.2$, $\xi = 0.7\%$, $\kappa = 1\%$, $\rho = 15\%$ and $x_0 = 0$.

Strikes	60%	80%	100%	120%	140%
Second Order formula	20.94%	20.50%	20.06%	19.69%	19.35%
Third Order formula	21.08%	20.51%	20.06%	19.70%	19.40%
MC with control variate	21.09%	20.51%	20.07%	19.70%	19.40%
MC-	19.29%	20.41%	20.03%	19.68%	19.37%
MC+	22.27%	20.62%	20.10%	19.73%	19.43%

Table 12.5: Implied Black-Scholes volatilities for the second order formula (12.15), the third order formula (12.16) and the Monte Carlo simulations (3×10^6 simulations using Euler scheme with 20 time steps) expressed as a function of strikes at the expiry $T = 1Y$. Parameters: $\beta = 0.2$, $\nu = 0.2$, $\xi = 0.7\%$, $\kappa = 1\%$, $\rho = 15\%$ and $x_0 = 0$.

Strikes	60%	80%	100%	120%	140%
Second Order formula	22.77%	21.75%	20.06%	18.55%	16.49%
Third Order formula	24.02%	21.91%	20.08%	18.66%	17.59%
MC with control variate	24.03%	21.92%	20.09%	18.67%	17.61%
MC-	23.29%	21.82%	20.05%	18.64%	17.58%
MC+	24.69%	22.01%	20.12%	18.69%	17.63%

12.5.1 Technical results related to explicit correction terms

The two first lemmas are proved in Section 4.6 of Chapter 4.

Lemma 12.5.1. *For any continuous (or piecewise continuous) function f, any continuous semimartingale Z vanishing at t=0, one has:*

$$\int_0^T f_t Z_t dt = \int_0^T (\int_t^T f_s ds) dZ_t.$$

Lemma 12.5.2. *One has:*

$$\mathbb{E}_T[(\int_0^T u_t dW_t^{1,T}) l(\int_0^T a_t dW_t^{1,T})] = \mathbb{E}_T[(\int_0^T a_t u_t dt) l^{(1)}(\int_0^T a_t dW_t^{1,T})].$$

If u is deterministic then $\mathbb{E}_T[(\int_0^T u_t dW_t^{1,T}) l(\int_0^T a_t dW_t^{1,T})] = (\int_0^T a_t u_t dt) \partial_x \mathbb{E}_T[l(\int_0^T a_t dW_t^{1,T} + x)]|_{x=0}$.

Lemma 12.5.3.

$$\begin{aligned}\mathbb{E}_T[l(\int_0^T a_t dW_t^{1,T})(\int_0^T v_t X_{1,t} dt)] =& (\int_0^T a_t \sigma_t (\int_t^T v_s ds) dt) \mathbb{E}_T[l^{(1)}(\int_0^T a_t dW_t^{1,T})] \\ &+ (\int_0^T (\rho_t^{S,r} . \Gamma(t,T) \sigma_t - \frac{\sigma_t^2}{2})(\int_t^T v_s ds) dt) \mathbb{E}_T[l(\int_0^T a_t dW_t^{1,T})]\end{aligned}$$

Table 12.6: Implied Black-Scholes volatilities for the second order formula (12.15), the third order formula (12.16) and the Monte Carlo simulations (3×10^6 simulations using Euler scheme with 20 time steps) expressed as a function of strikes at the expiry $T=5Y$. Parameters: $\beta=0.8$, $\nu=0.2$, $\xi=0.7\%$, $\kappa=1\%$, $\rho=15\%$ and $x_0=0$.

Strikes	40%	70%	100%	140%	180%
Second Order formula	21.90%	21.04%	20.36%	19.70%	19.15%
Third Order formula	22.18%	21.07%	20.36%	19.72%	19.25%
MC with control variate	22.22%	21.11%	20.40%	19.75%	19.28%
MC-	21.68%	21.02%	20.36%	19.72%	19.25%
MC+	22.72%	21.19%	20.44%	19.78%	19.31%

Table 12.7: Implied Black-Scholes volatilities for the second order formula (12.15), the third order formula (12.16) and the Monte Carlo simulations (3×10^6 simulations using Euler scheme with 20 time steps) expressed as a function of strikes at the expiry $T=5Y$. Parameters: $\beta=0.2$, $\nu=0.2$, $\xi=0.7\%$, $\kappa=1\%$, $\rho=15\%$ and $x_0=0$.

Strikes	40%	70%	100%	140%	180%
Second Order formula	25.15%	23.06%	20.37%	17.64%	13.91%
Third Order formula	27.79%	23.40%	20.47%	18.00%	16.48%
MC with control variate	27.86%	23.42%	20.50%	18.02%	16.49%
MC-	27.58%	23.34%	20.46%	17.98%	16.44%
MC+	28.14%	23.50%	20.54%	18.05%	16.54%

Proof. Applying first Lemma 12.5.1 to $f(t)=v_t$ and $Z_t=X_{1,t}$, we obtain:

$$\begin{aligned}
\mathbb{E}_T[(\int_0^T v_t X_{1,t}dt)l(\int_0^T a_t dW_t^{1,T})] =&\mathbb{E}_T[(\int_0^T(\int_t^T v_s ds)dX_{1,t})l(\int_0^T a_t dW_t^{1,T})]\\
=&\mathbb{E}_T[(\int_0^T(\int_t^T v_s ds)(\sigma_t dW_t^{1,T}+(\rho_t\Gamma_1(t,T)-\frac{\sigma_t^2}{2})dt)l(\int_0^T a_t dW_t^{1,T})]\\
=&(\int_0^T a_t\sigma_t(\int_t^T v_s ds)dt)\mathbb{E}_T[l^{(1)}(\int_0^T a_t dW_t^{1,T})]\\
&+(\int_0^T(\rho_t\Gamma_1(t,T)-\frac{\sigma_t^2}{2})(\int_t^T v_s ds)dt)\mathbb{E}_T[l(\int_0^T a_t dW_t^{1,T})],
\end{aligned}$$

where we have used Lemma 12.5.2 for the last equality. □

12.5.2 Proof of Theorem 12.2.1

Using Equation (12.6), the r.v. $\int_0^T r_s ds+X_T^B$ can be projected on the $\mathbb{Q}^T$-Brownian motion $W^{1,T}$ as follows:

$$\int_0^T r_s ds+X_T^B=\int_0^T(\sigma_t-\rho_t^{S,r}.\Gamma(t,T))dW_t^{1,T}+D_T,$$

Table 12.8: Implied Black-Scholes volatilities for the second order formula (12.15), the third order formula (12.16) and the Monte Carlo simulations (3×10^6 simulations using Euler scheme with 20 time steps) expressed as a function of strikes at the expiry $T = 10Y$. Parameters: $\beta = 0.8$, $\nu = 0.2$, $\xi = 0.7\%$, $\kappa = 1\%$, $\rho = 15\%$ and $x_0 = 0$.

Strikes	30%	60%	100%	160%	220%
Second Order formula	22.82%	21.82%	20.88%	19.89%	19.34%
Third Order formula	23.19%	21.86%	20.89%	19.94%	19.49%
MC with control variate	23.27%	21.93%	20.96%	19.99%	19.55%
MC-	22.82%	21.83%	20.92%	19.96%	19.52%
MC+	23.70%	22.03%	21.01%	20.03%	19.58%

Table 12.9: Implied Black-Scholes volatilities for the second order formula (12.15), the third order formula (12.16) and the Monte Carlo simulations (3×10^6 simulations using Euler scheme with 20 time steps) expressed as a function of strikes at the expiry $T = 10Y$. Parameters: $\beta = 0.2$, $\nu = 0.2$, $\xi = 0.7\%$, $\kappa = 1\%$, $\rho = 15\%$ and $x_0 = 0$.

Strikes	30%	60%	100%	160%	220%
Second Order formula	27.02%	24.57%	20.92%	16.71%	11.74%
Third Order formula	30.79%	25.29%	21.15%	17.66%	16.53%
MC with control variate	30.98%	25.29%	21.17%	17.61%	16.43%
MC-	30.74%	25.20%	21.13%	17.59%	16.40%
MC+	31.22%	25.37%	21.21%	17.64%	16.47%

where D_T is a Gaussian random variable independent on $(W_t^{1,T})_t$. Then using notation $a_t = \sigma_t - \rho_t^{S,r}.\Gamma(t,T)$, one gets

$$\begin{aligned}
\mathbb{E}_T[h(\int_0^T r_s ds + X_T^B)\frac{X_{2,T}}{2}] =& \mathbb{E}_T[h(\int_0^T a_t dW_t^{1,T} + D_T)\frac{X_{2,T}}{2}] \\
=& \mathbb{E}_T[h(\int_0^T a_t dW_t^{1,T} + D_T)\int_0^T X_{1,t}(\sigma_t^{(1)} dW_t^{1,T} + (\rho_t^{S,r}.\Gamma(t,T)\sigma_t^{(1)} - \sigma_t\sigma_t^{(1)})dt)], \\
=& \mathbb{E}_T[h^{(1)}(\int_0^T a_t dW_t^{1,T} + D_T)\int_0^T a_t\sigma_t^{(1)} X_{1,t} dt)] \\
&+ \mathbb{E}_T[h(\int_0^T a_t dW_t^{1,T} + D_T)\int_0^T (\rho_t^{S,r}.\Gamma(t,T)\sigma_t^{(1)} - \sigma_t\sigma_t^{(1)})X_{1,t} dt],
\end{aligned}$$

where we have used Equation (12.12) for second Equality and Lemma 12.5.2 for last one.
An application of Lemma 12.5.3 gives immediately the Equality (12.15).

Error analysis. For the smooth case, we only need estimates on $\varepsilon \to X_t^\varepsilon$ and its derivatives, in terms of M_0 and M_1. This is very similar to our previous work (see Chapter 7) and we skip the details (use our discussion in the beginning of Section 12.2). For the call/put case or digital case, once again we follow the lines of the proof of Chapters 4 and 7. The computations and estimates remain the same, except for Lemma 7.5.3 in Chapter 7 and Lemma 4.5.1 in Chapter 4 which now writes as follows. Here, appears the correlation assumption (*Rho*).

Lemma 12.5.4. *Assume (E), (Rho) and (R_{k+1}) for a given $k \geq 1$. Let Z belong to $\cap_{p\geq 1}\mathbb{D}^{k,p}$. For any $v \in [0,1]$, there exists a random variable Z_k^v in any $\mathbf{L}_p$ ($p \geq 1$) such that for any function $l \in \mathscr{C}_0^\infty(\mathbb{R})$, we have*

$$\mathbb{E}_T[l^{(k)}(\int_0^T r_s ds + vX_T + (1-v)X_T^B)Z] = \mathbb{E}_T[l(v\int_0^T r_s ds + X_T + (1-v)X_T^B)Z_k^v].$$

Moreover, we have $\|Z_k^v\|_p \leq C\frac{\|Z\|_{k,2p}}{(\sqrt{1-|\rho_\infty^{S,r}|^2}\sigma_{inf}\sqrt{T})^k}$, uniformly in v, where C depends (in an increasing way) on the bounds of the model coefficients and the maturity.

Proof. We follow the same approach from the quoted references. First, we prove that a suitable Malliavin covariance matrix is invertible and we estimate the L_p-norm of its inverse. Second, we apply the integration by parts from Malliavin calculus to get the existence of Z_k^v, and finally, we provide estimates for its L_p-norm. Only the first step is a bit different and is worth being detailed. For the other steps, we refer to the proof of Lemma 4.5.1 in Chapter 4. Let us denote $F_\nu = \int_0^T r_s ds + \nu X_T + (1-\nu)X_T^B$. All the calculus of stochastic variations will be performed relatively to the $(n+1)$-dimensional Brownian motion $(\tilde{W}^{1,T}, B^{1,T}, \cdots, B^{n,T})$. We define $\tilde{W}^{1,T}$ by the relation

$$dW_t^{1,T} = \sqrt{1-|\rho_t^{S,r}|^2} d\tilde{W}_t^{1,T} + \rho_t^{S,r}.dB_t^T,$$

from which we deduce that $(\tilde{W}^{1,T}, B^{1,T}, \cdots, B^{n,T})$ is indeed a standard $\mathbb{Q}^T$-Brownian motion. The key feature in this choice is that the first component of the Malliavin derivative of $\int_0^T r_s ds$ is zero: $D^{\tilde{W}^{1,T}}(\int_0^T r_s ds) \equiv 0$. Hence, we have:

$$D_t^{\tilde{W}^{1,T}} F_\nu = \nu\sigma(t,X_t)\sqrt{1-|\rho_t^{S,r}|^2} e^{\int_t^T \sigma_u^{(1)}(X_u)dW_u + (\rho_u^{S,r}\Gamma(u,T)\sigma_u^{(1)} - \sigma_u^{(1)}\sigma_u - \frac{1}{2}(\sigma_u^{(1)})^2)(X_u)du} + (1-\nu)\sigma_t\sqrt{1-|\rho_t^{S,r}|^2},$$

and thus

$$D_t^{\tilde{W}^{1,T}} F_\nu \geq \sigma_{inf}\sqrt{1-|\rho^{S,r}|_\infty^2} \inf_{0\leq t\leq T} e^{\int_t^T \sigma_u^{(1)}(X_u)dW_u + (\rho_u^{S,r}\Gamma(u,T)\sigma_u^{(1)} - \sigma_u^{(1)}\sigma_u - \frac{1}{2}(\sigma_u^{(1)})^2)(X_u)du}.$$

It is easy to deduce that the Malliavin covariance matrix is bounded from below by:

$$\begin{aligned}\gamma_{F_\nu} \geq &\int_0^T |D_t^{\tilde{W}^{1,T}} F_\nu|^2 dt \\ \geq &T\sigma_{inf}^2(1-|\rho^{S,r}|_\infty^2) \inf_{0\leq t\leq T} e^{2\int_t^T \sigma_u^{(1)}(X_u)dW_u + (\rho_u^{S,r}\Gamma(u,T)\sigma_u^{(1)} - \sigma_u^{(1)}\sigma_u - \frac{1}{2}(\sigma_u^{(1)})^2)(X_u)du},\end{aligned}$$

from which it readily follows (for $p \geq 1$)

$$\|\gamma_{F_\nu}^{-1}\|_p \leq C(\sigma_{inf}\sqrt{1-|\rho^{S,r}|_\infty^2}\sqrt{T})^{-2}.$$

□

Bibliography

[1] C. Albanese, G. Campolieti, P. Carr, and A. Lipton. Black-Scholes goes hypergeometric. *Risk magazine*, 14(12):99–103, 2001.

[2] H. Albrecher, P. Mayer, W. Schoutens, and J. Tistaert. The little Heston trap. *Wilmott Magazine*, (1):83–92, 2007.

[3] A. Alfonsi. On the discretization schemes for the cir (and bessel squared) processes. *Monte Carlo Methods and Applications*, 11(4):355–384, 2005.

[4] A. Alfonsi. High order discretization schemes for the cir process: application to affine term structure and Heston models. *To appear in Mathematics of Computation*, 2009.

[5] G. Allaire. *Shape optimization by the homogenization method.* Applied Mathematical Sciences 146, Springer Verlag, 2002.

[6] E. Alòs, J. A. León, and J. Vives. On the short-time behaviour of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics*, 11(4):571–589, 2007.

[7] L. Andersen. Efficient simulation of the heston stochastic volatility model. *The Journal of computational Finance*, 11(3), 2008.

[8] L. Andersen and J. Andreasen. Jump diffusion process: volatility smile fitting and numerical methods for pricing. *Review of Operations Research*, 4:231–262, 2000.

[9] L. Andersen and J. Andreasen. Volatile volatilities. *Risk magazine*, 15(12):163–168, 2002.

[10] L. Andersen and R. Brotherton-Ratcliffe. The equity option volatility smile: a finite difference approach. *Journal of Computational Finance*, 1(2):5–38, 1998.

[11] J. Andreasen. Closed form pricing of FX options under stochastic interest rates and volatility. *Global Derivatives Conference,Paris*, 2006.

[12] F. Antonelli and S. Scarlatti. Pricing options under stochastic volatility: a power series approach. *Finance and Stochastics*, 13(2):269–303, 2009.

[13] G. Bakshi and D. Madan. Spanning and derivative security valuation. *J. Finan. Econom*, 55:205–238, 2000.

[14] P. Balland. Stoch-vol model with interest rate volatility. *ICBI Global Derivatives preceedings*, 2005.

[15] S. Benaim and P. Friz. Regular variation and smile asymptotics. *Discussion paper Cambridge University, Department of Pure Mathematics and Mathematical Statistics*, 2006.

[16] C. Bender and S. Orszag. *Advanced Mathematical Methods for Scientists and Engineers*. McGraw-Hill, 1978.

[17] E. Benhamou and al. *Global Derivatives: Products, Theory and Practice*. Hardcover, 2007.

[18] E. Benhamou, A. Gruz, and A. Rivoira. Stochastic interest rates for local volatility hybrid models. *Wilmot Magazine*, 2008.

[19] E. Benhamou and M. Miri. Predictor corrector methods applied to PIDE and Monte Carlo simulations. *working paper, Pricing Partners*, 2006.

[20] H. Berestycki, J. Busca, and I. Florent. Asymptotics and calibration in local volatility models. *Quantitative finance*, 2:61–69, 2002.

[21] H. Berestycki, J. Busca, and I. Florent. Computing the implied volatility in stochastic volatility models. *Comm. Pure Appl. Math.*, 10:1352–1373, 2004.

[22] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–659, 1973.

[23] A. N. Borodin and P. Salminen. *Handbook of Brownian motion-facts and formulae*. Probability and its Applications, Birkhauser, second edition, 2002.

[24] M. Bossy and A. Diop. An efficient discretization scheme for one dimensional sdes with a diffusion coefficient function of the form $|x|^a$, a in [1/2,1). *Technical report, INRIA,preprint RR-5396-version 4*, 2007.

[25] B. Bouchard and R. Elie. Discrete-time approximation of decoupled Forward-Backward SDE with jumps. *Stochastic Processes and their Applications*, 118:53–75, 2008.

[26] D. Brigo and F. Mercurio. *Interest Rate Models: Theory and Practice*. Springer Finance, 2nd edition, 2006.

[27] M. Broadie and O. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 54(2), 2006.

[28] G. Brunick. *A weak existence result with application to the financial engineerŠs calibration problem*. PhD thesis, Carnegie Mellon University, 2008.

[29] I. Chavel. *Eigenvalues in Riemannian geometry*. Academic Press, 1984.

[30] J. D. Cole and J. Kevorkian. *Multiple scale and singular perturbation methods*. Springer, 1996.

[31] R. Cont and P. Tankov. Non-parametric calibration of jump diffusion option pricing models. *Journal of Computational Finance*, 7(3):1–49, 2003.

[32] R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman and Hall / CRC Press, 2004.

[33] R. Cont and E. Voltchkova. A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. *SIAM Journal on Numerical Analysis*, 43(4):1596–1626, 2005.

[34] J. Cox. Notes on option pricing I: Constant elasticity of variance diffusions. *Working paper, Stanford University (reprinted in Journal of Portfolio Management, 1996, 22, 15-17)*, 1975.

[35] J. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of the interest rates. *Econometrica*, 53, 1985.

[36] E. Derman and I. Kani. Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility. *International Journal of Theoretical and Applied Finance*, 1(1):61–110, 1998.

[37] C. F. Faá di Bruno. Note sur une nouvelle formule de calcul différentiel. *Quarterly J. Pure Appl. Math*, 1:359–360, 1857.

[38] A. Diop. *Sur la discrètisation et le comportement à petit bruit dÉDS multidimensionnelles dont les coefficients sont á dérivt'ees singulières*. PhD thesis, INRIA,(available at http://www.inria.fr/rrrt/0785.html), 2003.

[39] D. Dufresne. The integrated square-root process. *Working paper, University of Montreal*, 2001.

[40] B. Dupire. Pricing with a smile. *RISK magazine*, 7(1):18–20, 1994.

[41] A. Elices. Models with time-dependent parameters using transform methods: application to Heston's model. *arXiv:0708.2020v2*, 2008.

[42] D. Emanuel and J. MacBeth. Further results on the constant elasticity of variance call option pricing model. *J. Financial and Quant. Anal.*, (17):533–554, 1982.

[43] M. Forde. The small-time behavior of diffusion and time-changed diffusion processes on the line. *arXiv:math.PR/0609117*, 2008.

[44] J. P. Fouque, G. Papanicalaou, and R. Sircar. *Derivatives in financial Markets with stochastic volatility*. Cambridge University Press, 2000.

[45] J. P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna. Singular perturbations in option pricing. *SIAM J. Applied Math*, 63:1648–1665, 2002.

[46] E. Fournié, J. M. Lasry, J. Lebuchoux, P. L. Lions, and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3:391–412, 1999.

[47] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*. Second ed., Springer-Verlag, New York, 1998.

[48] T. Fujiwara and H. Kunita. Stochastic differential equations of jump type and Levy processes in diffeomorphisms group. *J. Math. Kyoto Univ.*, 25:71–106, 1985.

[49] J. Gatheral. The volatility skew: Arbitrage constraints and asymptotic behaviour. *Merrill Lynch*, 1999.

[50] J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. Wiley Finance, 2006.

[51] H. Geman, N. El Karoui, and J.C. Rochet. Changes of numeraire, changes of probability measure and option pricing. *Journal of Applied Probability*, 32(2):443–458, 1995.

[52] R. Gibson and E. S. Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, XLV(3):959–976, 1990.

[53] J. Gil-Pelaez. Note on the inverse theorem. *Biometrika*, 37:481–482, 1951.

[54] E. Gobet. Revisiting the Greeks for European and American options. In J. Akahori, S. Ogawa, and S. Watanabe, editors, *Proceedings of the International Symposium on Stochastic Processes and Mathematical Finance, Ritsumeikan University, Kusatsu, Japan*, pages 53–71. World Scientific Publishing, 2004.

[55] E. Gobet and R. Munos. Sensitivity analysis using Itô-Malliavin calculus and martingales. Application to stochastic control problem. *SIAM Journal of Control and Optimization*, 43(5):1676–1713, 2005.

[56] A. Göing-Jaeschke and M. Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[57] D. Goldenberg. A unified method for pricing options on diffusion processes. *J. Financial Econom.*, 29:3–34, 1991.

[58] J. Gurland. Inversion formulae for the distribution of ratios. *Ann. of Math. Statistics*, 19:228–237, 1948.

[59] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an ito differential. *Probability Theory and Related Fields*, 71(4):501–516, 1986.

[60] A. Haastrecht, R. Lord, A. Pelsser, and D. Schrager. Pricing long-maturity equity and fx derivatives with stochastic interest rates and stochastic volatility,. *ssrn working paper, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1131137*, 2008.

[61] P. S. Hagan, D. Kumar, A. S. Lesniewski, and D. E. Woodward. Managing smile risk. *Willmott Magazine*, pages 84–108, 2002.

[62] P. S. Hagan and D. E. Woodward. Equivalent Black volatilities. *Applied mathematical Finance*, 6:147–157, 1999.

[63] S. Heston. A closed-form solutions for options with stochastic volatility with applications to bond and currency options,. *Review of Financial Studies*, 6:327–343, 1993.

[64] M. H. Holmes. *Introduction to Perturbation Methods*. Hardcover, 1995.

[65] Y. L. Hsu., T. I. Lin, and C. F. Lee. Constant elasticity of variance (CEV) option pricing model: Integration and detailed derivation. *Math. Comput. Simul*, 1, 2007.

[66] C. H. Hui, C. F. Lo, and P. H. Yuen. Constant elasticity of variance option pricing model with time-dependent parameters. *International Journal of Theoretical and Applied Finance*, 3(4):661–674, 2000.

[67] J. C. Hull and A. White. Pricing interest-rate derivative securities. *Rev. Finan. Stud*, 3:573–592, 1990.

[68] P. Jackel and C. Kahl. Not-so-complex logarithms in the heston model. *Wilmott Magazine*, (19):94–103, 2005.

[69] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*. Springer-Verlag, New York, 1988.

[70] T. Kluge. *Pricing Derivatives in Stochastic Volatility Models using the Finite Difference Method*. PhD thesis, Technical University, Chemnitz, 2002.

[71] S. G. Kou. A jump-diffusion model for option pricing. *Management Science*, 48:1086–1101, 2002.

[72] H. Kunita. Stochastic flows of diffeomorphisms. *École d'été de Probabilités de St-Flour XII. Lecture notes.*, 1097, 1984.

[73] N. Kunitomo and A. Takahashi. The asymptotic expansion approach to the valuation of interest rate contingent claims. *Mathematical Finance*, (117-151), 2001.

[74] P. Henry Labordère. A general asymptotic implied volatility for stochastic volatility models. *arXiv:condmat/04317*, 2005.

[75] P. Henry Labordère. *Analysis, Geometry, and Modeling in Finance: Advanced Methods in Option Pricing*. Chapman and Hall, 2008.

[76] P. Henry Labordère. Calibration of local stochastic volatility models to market smiles. *preprint (submitted)*, 2008.

[77] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance*. Chapman and Hall, 1996.

[78] R. W. Lee. The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, 14(3):469–480, 2004.

[79] A. Lewis. *Option Valuation under Stochastic Volatility with Mathematica Code*. Finance Press, Newport Beach, California, 2000.

[80] A. Lewis. Geometries and smile asymptotics for a class of stochastic volatility models. *www.optioncity.net*, 2007.

[81] A. Lipton. The vol smile problem. *Risk Magazine*, pages 61–65, 2002.

[82] Y. Maghsoodi. Solution of the extended CIR term structure and bond option valuation. *Mathematical finance*, 6(1):89–109, 1996.

[83] A. M. Matache, T. von Petersdorff, and C. Schwab. Fast deterministic pricing of options on Lévy driven assets. *Mathematical Modelling and Numerical Analysis*, 38(1):37–71, 2004.

[84] A. Matytsin. Perturbative analysis of volatility smiles. *working paper, http://www.math.columbia.edu/ smirnov/matytsin2000.pdf*, 2000.

[85] R. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144, 1976.

[86] S. Mikhailov and U. Nogel. Heston's stochastic volatility model implementation, calibration and some extensions. *Wilmott magazine*, pages 74–79, 2003.

[87] A. H. Nayfeh. *Perturbation Methods*. Wiley Classics Library. Wiley-Interscience, 2000.

[88] D. Nualart. *Malliavin calculus and related topics*. Springer-Verlag, Berlin, second edition, 2006.

[89] J. Obloj. Fine-tune your smile: Correction to Hagan et al. *Wilmott Magazine*, 2008.

[90] Y. Osajima. The asymptotic expansion formula of implied volatility for dynamic SABR model and FX hybrid model. *Available at http://ssrn.com/abstract=965265*, 2007.

[91] V. V. Piterbarg. A multi-currency model with FX volatility skew. *ssrn working paper, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=685084*, 2005.

[92] V. V. Piterbarg. Stochastic volatility model with time-dependent skew. *Applied Mathematical Finance*, 12(2):147–185, 2005.

[93] V. V. Piterbarg. Time to smile. *Risk Magazine*, 18(5):71–75, 2005.

[94] V. V. Piterbarg. Markovian projection for volatility calibration. *Risk*, 20(4):84–89, 2007.

[95] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1992.

[96] P. Protter. *Stochastic integration and differential equations*. Springer, Heidelberg, 1990.

[97] E. Renault and N. Touzi. Option hedging and implied volatilities in a stochastic volatility model. *Mathematical Finance*, 6:279–302, 1996.

[98] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag, Hardcover, 3rd edition, 1999.

[99] C. Rogers and M. Tehranchi. Can the implied volatility surface move by parallel shifts? *To appear in Finance and Stochastics*, 2009.

[100] M. Romano and N. Touzi. Contingent claims and market completeness in a stochastic volatility model. *Mathematical Finance*, 7(4):399–410, 1997.

[101] M. Rubinstein. Implied binomial trees. *Journal of Finance*, 49:771–818, 1994.

[102] W. Rudin. *Real and complex analysis*. McGraw-Hill, Toronto, 1966.

[103] K. Sato. *Lévy processes and infinitely divisible distributions*. Cambridge University Press, 1999.

[104] M. Schroder. Computing the constant elasticity of variance option pricing formula. *Journal of Finance*, 44:211–219, 1989.

[105] L. Scott. Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: application of Fourier inversion methods. *Mathematical Finance*, 7:345–358, 1997.

[106] M. Siopacha and J. Teichmann. Weak and strong Taylor methods for numerical solutions of stochastic differential equations. *Forthcoming in Quantitative Finance*, 2007.

[107] R. D. Smith. An almost exact simulation method for the heston model. *Journal of Computational Finance*, 11(1):115–125, 2008.

[108] M. Tehranchi. Asymptotics of implied volatility far from maturity. *To appear in the Journal of Applied Probability*, 2009.

[109] S. R. S. Varadhan. Diffusion processes in small time intervals. *Communication Pure Appl. Math.*, 20:659–685, 1967.

[110] S. R. S. Varadhan. *Large deviations and applications*. SIAM, Philadelphia, 1984.

[111] F. Verhulst. *Methods and Applications of Singular Perturbations: Boundary Layers and Multiple Timescale Dynamics*. Springer, 2005.

[112] S. Watanabe. Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. *Annals of Probability*, 15(1):1–39, 1987.

[113] N. Yoshida. Asymptotic expansion for statistics related to small diffusions. *J. Japan Statistical Society*, 22(2):139–159, 1992.

[114] N. Yoshida. Asymptotic expansions of maximum likelihood estimator for small diffusion via the theory of malliavin-watanabe. *Probability Theory and Related Fields*, 92:275–311, 1992.