Stochastic expansion for the diffusion processes and applications to option pricing

Romain Bompis

École Polytechnique

ÉCOLE DOCTORALE: Mathématiques Appliquées

THÈSE

pour obtenir le titre de

Docteur de l'École Polytechnique

Specialité : Mathématiques Appliquées

présentée et soutenue publiquement par

Romain Bompis

Stochastic expansion for the diffusion processes and applications to option pricing

Thèse préparée au CMAP

soutenue le 11 décembre 2013

Jury :

Directeur:	Emmanuel Gobet	-	École Polytechnique
Rapporteurs:	Andrea Pascucci	-	Université de Bologne
	Peter Tankov	-	Université Paris 7 - Denis Diderot
Président du jury:	Nicole El Karoui	-	Université Paris 6 - Pierre et Marie-Curie
Examinateurs:	Stéphane Crépey	-	Université d'Évry
	Monique Jeanblanc	-	Université d'Évry
	Anis Matoussi	-	Université du Maine
	Nizar Touzi	-	École Polytechnique

Remerciements

Tout d'abord je voudrais remercier mon directeur de thèse Emmanuel Gobet. C'est lui qui m'a initié au monde de la recherche et je le remercie de m'avoir poussé à travailler sur ce passionnant sujet sur les développements stochastiques qui m'a enthousiasmé durant ces trois années. Je le remercie également d'avoir été présent dans les "coups durs".

Ensuite je remercie tous les membres du jury. Dans un premier temps Andrea Pascucci et Peter Tankov, pour avoir accepté de rapporter ce travail, ce dont je leur suis très reconnaissant. Puis Stéphane Crépey, Nicole El Karoui, Monique Jeanblanc, Anis Matoussi et Nizar Touzi qui me font l'honneur de faire partie de mon jury.

Puis j'adresse mes remerciements à tous les membres du CMAP qui m'ont permis de réaliser cette thèse dans les meilleurs conditions. J'en profite également pour remercier vivement Arturo Kohatsu-Higa pour son invitation au Japon à l'hiver 2011. Je n'oublierai jamais la qualité de son acceuil.

Pour finir je tiens à remercier chaleureusement mes amis et ma famille pour leur soutien et leurs encouragements réguliers. En particulier j'exprime toute ma gratitude à mon épouse Cécile pour sa patience et son réconfort dans les moments difficiles. Merci de m'avoir supporté pendant ces trois années, merci pour ton amour, je n'aurais pas réussi à finaliser ce travail sans toi.

Stochastic expansion for the diffusion processes and applications to option pricing

Abstract: This thesis deals with the approximation of the expectation of a functional (possibly depending on the whole path) applied to a diffusion process (possibly multidimensional).

The motivation for this work comes from financial mathematics where the pricing of options is reduced to the calculation of such expectations. The rapidity for price computations and calibration procedures is a very strong operational constraint and we provide real-time tools (or at least more competitive than Monte Carlo simulations in the case of multidimensional diffusions) to meet these needs.

In order to derive approximation formulas, we choose a proxy model in which analytical calculus are possible and then we use stochastic expansions around the proxy model and Malliavin calculus to approach the quantities of interest. In situation where Malliavin calculus can not be applied, we develop an alternative methodology combining Itô calculus and PDE arguments. All the approaches (from PDEs to stochastic analysis) allow to obtain explicit formulas and tight error estimates in terms of the model parameters. Although the final result is generally the same, the derivation can be quite different and we compare the approaches, first regarding the way in which the corrective terms are made explicit, second regarding the error estimates and the assumptions used for that.

We consider various classes of models and functionals throughout the four Parts of the thesis.

In the Part I, we focus on local volatility models and provide new price, sensitivity (delta) and implied volatility approximation formulas for vanilla products showing an improving accuracy in comparison to previous known formulas. We also introduce new results concerning the pricing of forward start options.

The Part II deals with the analytical approximation of vanilla prices in models combining both local and stochastic volatility (Heston type). This model is very difficult to analyze because its moments can explode and because it is not regular in the Malliavin sense. The error analysis is original and the idea is to work on an appropriate regularization of the payoff and a suitably perturbed model, regular in the Malliavin sense and from which the distance with the initial model can be controlled.

The Part III covers the pricing of regular barrier options in the framework of local volatility models. This is a difficult issue due to the indicator function on the exit times which is not considered in the literature. We use an approach mixing Itô calculus, PDE arguments, martingale properties and temporal convolutions of densities to decompose the approximation error and to compute correction terms. We obtain explicit and accurate approximation formulas under a martingale hypothesis.

The Part IV introduces a new methodology (denoted by SAFE) for the efficient weak analytical approximation of multidimensional diffusions in a quite general framework. We combine the use of a Gaussian proxy to approximate the law of the multidimensional diffusion and a local interpolation of the terminal function using Finite Elements. We give estimates of the complexity of our methodology. We show an improved efficiency in comparison to Monte Carlo simulations in small and medium dimensions (up to 10).

Keywords: Stochastic expansion, diffusion process, weak approximation, financial mathematics, option pricing, stochastic analysis, Malliavin calculus, partial differential equation, local volatility model, stochastic volatility, hitting time, barrier option, finite elements.

Développement stochastique pour les processus de diffusion et applications à la valorisation d'options

Résumé: Cette thèse est consacrée à l'approximation de l'espérance d'une fonctionnelle (pouvant dépendre de toute la trajectoire) appliquée à un processus de diffusion (pouvant être multidimensionnel).

La motivation de ce travail vient des mathématiques financières où la valorisation d'options se réduit au calcul de telles espérances. La rapidité des calculs de prix et des procédures de calibration est une contrainte opérationnelle très forte et nous apportons des outils temps-réel (ou du moins plus compétitifs que les simulations de Monte Carlo dans le cas multidimensionnel) afin de combler ces besoins.

Pour obtenir des formules d'approximation, on choisit un modèle proxy dans lequel les calculs analytiques sont possibles, puis nous utilisons des développements stochastiques autour de ce modèle proxy et le calcul de Malliavin afin d'approcher les quantités d'intérêt. Dans le cas où le calcul de Malliavin ne peut pas être appliqué, nous développons une méthodologie alternative combinant calcul d'Itô et arguments d'EDP. Toutes les approches (allant des EDPs à l'analyse stochastique) permettent d'obtenir des formules explicites et des estimations d'erreur précises en fonction des paramètres du modèle. Bien que le résultat final soit souvent le même, la dérivation explicite du développement peut être très différente et nous comparons les approches, tant du point de vue de la manière dont les termes correctifs sont rendus explicites que des hypothèses requises pour obtenir les estimées d'erreur.

Nous considérons différentes classes de modèles et fonctionnelles lors des quatre Parties de la thèse.

Dans la Partie I, nous nous concentrons sur les modèles à volatilité locale et nous obtenons des nouvelles formules d'approximation pour les prix, les sensibilités (delta) et les volatilités implicites des produits vanilles surpassant en précision les formules connues jusque-là. Nous présentons aussi des résultats nouveaux concernant la valorisation des options à départ différé.

La Partie II traite de l'approximation analytique des prix vanilles dans les modèles combinant volatilité locale et stochastique (type Heston). Ce modèle est très délicat à analyser car ses moments ne sont pas tous finis et qu'il n'est pas régulier au sens de Malliavin. L'analyse d'erreur est originale et l'idée est de travailler sur une régularisation appropriée du payoff et sur un modèle habilement modifié, régulier au sens de Malliavin et à partir duquel on peut contrôler la distance par rapport au modèle initial.

La Partie III porte sur la valorisation des options barrières régulières dans le cadre des modèles à volatilité locale. C'est un cas non considéré dans la littérature, difficile à cause de l'indicatrice des temps de sorties. Nous mélangeons calcul d'Itô, arguments d'EDP, propriétés de martingales et de convolutions temporelles de densités afin de décomposer l'erreur d'approximation et d'expliciter les termes correctifs. Nous obtenons des formules d'approximation explicites et très précises sous une hypothèse martingale.

La Partie IV présente une nouvelle méthodologie (dénotée SAFE) pour l'approximation en loi efficace des diffusions multidimensionnelles dans un cadre assez général. Nous combinons l'utilisation d'un proxy Gaussien pour approcher la loi de la diffusion multidimensionnelle et une interpolation locale de la fonction terminale par éléments finis. Nous donnons une estimation de la complexité de notre méthodologie. Nous montrons une efficacité améliorée par rapport aux simulations de Monte Carlo dans les dimensions petites et moyennes (jusqu'à 10).

Mots clés: Développement stochastique, processus de diffusion, approximation en loi, mathématiques financières, valorisation d'options, analyse stochastique, calcul de Malliavin, équation aux dérivées partielles, modèle à volatilité locale, volatilité stochastique, temps d'atteinte, option barrière, éléments finis.

Contents

Chapter 1

Introduction

The Sections 1.1, 1.2 and 1.3 of this introduction Chapter have been published in the Chapter "Asymptotic and non asymptotic approximations for option valuation" of the book "Recent Developments in Computational Finance Foundations, Algorithms and Applications", Thomas Gerstner and Peter Kloeden (Ed.) 2012, World Scientific Publishing Company.

Contents

The aim of this thesis is to provide analytical approximations of the law of $(X_t)_{t\in[0,T]}$ solution of the following stochastic differential equation (SDE):

$$X_t = x_0 + \int_0^t b(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}W_s, \tag{1.1}$$

with applications to financial mathematics, where W is a Brownian motion (possibly multidimensional) and b and σ are functions which the regularity is not specified for the moment. More precisely we want to obtain explicit approximation formulas of the expectation of a functional φ possibly involving the whole trajectory of X on $[0, T]$:

$$\mathbb{E}[\varphi(X_t : 0 \le t \le T)]. \tag{1.2}$$

These problems are connected to the pricing of financial derivative products which prices can be represented with expectations of the form (1.2). Apart from some particular cases of processes X or functionals φ, the explicit calculus of (1.2) is not possible. Among them we cite the case where φ is a function of the terminal value of the process X and where the density of X_T (like in the Gaussian or log-normal models) or its characteristic function (like in the affine models) are explicitly known. Otherwise one must turn to numerical methods like PDE techniques or Monte Carlo simulations. The principal drawback of these numerical methods is that there are time-costing whereas the rapidity in the different computations and calibration procedures is a very strong operational constraint in the financial industry. An other approach to obtain real-time and explicit formulas is to derive analytical approximations. The approximations can be of very different nature: asymptotic or non-asymptotic approach, large choice of parameters under consideration, methodology used (rather stochastic analysis or rather PDE)... In Section 1.3 we give a broad overview of approximation methods to derive analytical formulas for accurate and quick evaluation of option prices. We compare different approaches, from the theoretical point of view regarding the tools they require, and also from the numerical point of view regarding their performances. We nevertheless mention some limitations of certain methods quoted further that we would like to circumvent:

- The approximation is oftenly explicit under a condition of time-homogeneity of the coefficients b and σ of the diffusion (see (1.1)), because its determination involves time-change of variable-change techniques or necessitates the explicit resolution of elliptic PDEs.
- Some works perform an asymptotic expansion which the accuracy is given in terms of a small parameter ϵ. However this accuracy may depend on the regularity of the functional φ and on the magnitude of the diffusion coefficients b and σ. Thus there might be a competition between these parameters and ϵ, what rises some problems of interpretation of the obtained theoretical accuracy.
- Sometimes, only the existence of an expansion is proven without the explicit calculation of the coefficients. On the contrary an heuristic expansion can be performed using formal calculations but rigorous error estimates are not provided (especially in the difficult cases of irregular diffusion coefficients or non smooth terminal functions).

To overcome these drawbacks, Benhamou, Gobet and Miri [Benhamou 2009] provide a new non asymptotic methodology, the so called "proxy principle". The idea is to:

1. choose a proxy model in which analytical calculus are possible,
2. find a way to establish a connexion between the initial model and the proxy model,
3. perform a non asymptotic expansion to approach the quantities of interest. The expansions take the following generic form (written in the particular case where φ is a function of X_T):
$$\mathbb{E}[\varphi(X_T)] = \mathbb{E}[\varphi(X_T^P)] + \sum_{\alpha} w_\alpha \, \partial^{|\alpha|}_{\epsilon_{\alpha_1}\cdots\epsilon_{\alpha_{|\alpha|}}} \mathbb{E}[\varphi(X_T^P + \epsilon)]|_{\epsilon=0} + \text{Error},$$
where:
   - $\partial^{|\alpha|}_{\epsilon_{\alpha_1}\cdots\epsilon_{\alpha_{|\alpha|}}} \mathbb{E}[\varphi(X_T^P + \epsilon)]|_{\epsilon=0}$ are sensitivities in the Proxy model,
   - w_α are explicit and depend on b, σ and T.
4. Estimate Error according to the magnitude of b, σ and their derivatives as well T and the regularity of φ.

The choice of the proxy is left to the expertise of the user. As an evidence of the efficiency of this approach, we cite some cases where it has been possible to handle the pricing of European and exotic options with very accurate and real-time approximation formulas:

- in local volatility models [Benhamou 2010a], taking as proxy the Black-Scholes model obtained by freezing the local volatility at spot,
- possibly incorporating Gaussian jumps [Benhamou 2009] choosing as proxy the Merton model [Merton 1976] or incorporating Hull & White type interest rates [Hull 1990] handled with the Black model proxy [Black 1976],
- in time-dependent Heston model [Benhamou 2010b] (see [Heston 1993] for the Heston model) using a suitable Black-Scholes model proxy obtained by vanishing the volatility of volatility,

- in general average options like Asian or Basket options [Gobet 2012a] taking as proxy Gaussian averages.

This thesis proposes to continue these works and to deepen them in several directions:

- improvement of the approximations in the local volatility models: freezing the local volatility function at different points that the spot, implied volatility expansions, analytical formulas for the sensitivities, choice of more accurate proxy models. . . Forward implied volatility expansions for the pricing of forward start options,
- approximation formulas for models combining local and stochastic volatility,
- expansion formulas of some path-dependent products like Barrier options,
- weak approximations and efficient calculations in multidimensional diffusions,

and thus to contribute to the development of real-time stochastic tools allowing to provide risk management indicators on advanced models or sophisticated products.
This introduction Chapter is organized as follows. In Section 1.1 we develop the application context and we list the different computational approaches. We give in Section 1.2 some useful notations repeatedly used throughout this Chapter and sometimes utilized in the remainder of the thesis. Section 1.3 of this Chapter gives an overview of asymptotic and non-asymptotic results: wing formulas, long maturity behavior, large deviations type results, regular and singular perturbation for PDEs, asymptotic expansions of Wiener functionals and other stochastic analysis approaches. The choice of the *small/large* parameter is of course crucial and is usually left to the expertise of the user. In particular, we compare in Subection 1.3.5 the asymptotic and non-asymptotic approaches and we show that there might be a competition between different small/large parameters and the accuracy order might not be the natural one. This motivates for deriving non asymptotic results and this is our emphasize throughout the next Chapters of the thesis. Besides we already provide a brief presentation of the "proxy principle" in Subsection 1.3.4.4. The outline and the main results of the thesis are given in Section 1.4.

1.1 Application context and overview of different computational approaches

In the two last decades, numerous works have been devoted to designing efficient methods in order to give exact or approximative pricing formulas for many financial products in various models. This quest of efficiency originates in the need for more and more accurate methods, when one takes into account an increasing number of sources of risk, while maintaining a competitive computational time. The current interest in real-time tools (for pricing, hedging, calibration) is also very high.

Let us give a brief overview of different computational approaches. While explicit formulas are available in simple models (Black-Scholes model associated to log-normal distribution [Black 1973], or Bachelier model related to normal distributions [Bachelier 1900] for instance), in general no closed forms are known and numerical methods have to be used. As a numerical method, it is usual to perform PDE solvers for one or two-dimensional sources of risk (see [Achdou 2005] for instance) or Monte Carlo methods for higher dimensional problems [Glasserman 2004]: both approaches are popular, efficiently developed and many improvements have been proposed for years. However, these methods are not intrinsically real-time methods, due to the increasing number of points required in the PDE discretization grid or due to the increasing number of paths needed in the Monte Carlo procedure. Not being *real-time method* means, for example, that when used for calibration routine based on data consisting of (say) 30 vanilla options, it usually takes more than one minute (in the most favorable situations) to achieve the calibration parameters. The approaches presented below are aimed at reducing this computational time to less than one second.

The class of affine models (such as Heston model, exponential Levy model ...) offers an alternative approach related to Fourier computations: on the one hand, in such models the characteristic function of the marginal distribution of the log-asset is explicitly known; on the other hand, there are general relations between Call/Put prices and the characteristic function of the log-asset. These relations write as follows.

- Following Carr and Madan [Carr 1998b], consider the difference $z_T(k)$ between the Call price in a given model and that price in an arbitrary Black-Scholes model (with volatility σ), both with maturity T and log-strike k. For zero interest rate (to simplify), it is equal to $z_T(k) = \mathbb{E}(e^{X_T} - e^k)_+ - \mathrm{Call}^{\mathrm{BS}}(k)$, where X is the logarithm of the asset. A direct computation gives explicitly the Fourier transform $\widehat{z_T}(v)$ in the log-strike variable:

$$\widehat{z_T}(v) = \int_{\mathbb{R}} e^{ivk} z_T(k)\mathrm{d}k = \frac{\Phi_T^X(1+iv) - \Phi_T^{\mathrm{BS}}(1+iv)}{iv(1+iv)},$$

where $\Phi_T^{X,\mathrm{BS}}(u) := \mathbb{E}(e^{uX_T})$ is either computed in the X model or in the Black-Scholes model. Since $\Phi_T^X(\cdot)$ is required to be known, we get the X-model Call price $z_T(k) + \mathrm{Call}^{\mathrm{BS}}(k)$ simultaneously for any log-strike using a Fast Fourier Transform.

- Alternatively, following the Lewis approach [Lewis 2000, Chapter 2], let $\alpha > 0$ be a damping constant, set $h(y) = (e^y - K)_+ e^{-(1+\alpha)y}$ which belongs to $L^2(\mathbb{R}, \mathrm{Leb.})$ and assume that $\mathbb{E}(e^{(1+\alpha)X_T}) < +\infty$: from the Parceval-Plancherel identity, assuming that the density p_{X_T} of X_T w.r.t. the Lebesgue measure exists, we obtain

$$\begin{aligned}
\mathbb{E}(e^{X_T} - K)_+ &= \int_{\mathbb{R}} h(y) e^{(1+\alpha)y} p_{X_T}(y)\mathrm{d}y \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{h}(-\xi)[\widehat{e^{(1+\alpha)\cdot} p_{X_T}(\cdot)}](\xi)\mathrm{d}\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{-(\alpha+i\xi)\log(K)}}{(i\xi+\alpha)(i\xi+\alpha+1)}\Phi_T^X(1+\alpha+i\xi)\mathrm{d}\xi.
\end{aligned}$$

The final identity still holds without assuming the existence of density: this can be proved by adding a small Brownian perturbation (considering $X_T + \varepsilon W_T$ instead of X_T), and taking the limit as the perturbation ε goes to 0. From the above formula, using an extra numerical integration method (to compute the ξ-integral), we recover Call prices. For higher numerical performance, Lewis recommends a variant of the formula above, obtained through the decomposition $(e^{X_T} - K)_+ = e^{X_T} - \min(e^{X_T}, K)$: it finally writes

$$\mathbb{E}(e^{X_T} - K)_+ = S_0 - \frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{(\frac{1}{2}-i\xi)\log(K)}}{\frac{1}{4}+\xi^2}\Phi_T^X(\frac{1}{2}+i\xi)\mathrm{d}\xi. \tag{1.3}$$

Regarding computational time, both Fourier-based approaches perform well, since they are essentially reduced to a one-dimensional integration problem. But they can be applied only to specific models for which the characteristic function is given in an explicit and tractable form: in particular, it rules out the local volatility models, the local and stochastic volatility models.

The last approach consists of explicit analytical approximations and this is the main focus of this thesis: it is based on the general principle of expanding the quantity of interest (price, hedge, implied volatility...) with respect to some small/large parameters (possibly multidimensional). The parameters under consideration may be of very different nature: for instance in the case of Call/Put options of strike K and maturity T, it ranges from the asymptotic behavior as K is small or large, to the case of short or long maturity T, passing through coupled asymptotics, or small/fast volatility variations, and so on... A detailed description with references is presented in Section 1.3 of this introduction Chapter. Due to the plentiful and recent litterature on the subject, it is likely that we will not be exhaustive in the references. But we will do our best to give the main trends and to expose whenever possible what are the links between different viewpoints; we will compare the mathematical tools to achieve these approximations (rather PDE techniques or stochastic analysis ones), in order to provide to the reader a clarified presentation of this prolific topic.

1.2 Notations used throughout the introduction Chapter

▷ **Models.** In all this Chapter, financial products are written w.r.t. a single asset, which price at time t is denoted by S_t. The dynamics of S is modeled through a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ where $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration of a standard linear Brownian motion W, augmented by the $\mathbb{P}$-null sets. The risk-free rate is set[1] to 0; most of the time and unless stated otherwise, S follows a local volatility model, i.e. it is solution of the stochastic diffusion equation

$$\mathrm{d}S_t = S_t \sigma(t, S_t)\mathrm{d}W_t, \tag{1.4}$$

where the dynamics is directly under the pricing measure. Assumptions on the local volatility σ are given later. We assume that the complete market framework holds and that an option with payoff $h(S_T)$ paid at maturity T has a fair value at time 0 equal to $\mathbb{E}(h(S_T))$.

For positive S, we define the log-asset $X = \log(S)$ which satisfies

$$\mathrm{d}X_t = a(t, X_t)\mathrm{d}W_t - \frac{1}{2}a^2(t, X_t)\mathrm{d}t, \tag{1.5}$$

where $a(t,x) = \sigma(t, e^x)$.

▷ **Call options.** Let us denote by $\mathrm{Call}(S_0, T, K)$ the price at time 0 of a Call option with maturity T and strike K, written on the asset S. "Price" usually means the price given by a model on S, that is

$$\mathrm{Call}(S_0, T, K) = \mathbb{E}(S_T - K)_+. \tag{1.6}$$

This Model Price should equalize the Market Price taken from Market datas (calibration step). As usual, ATM (At The Money) Call refers to $S_0 \approx K$, ITM (In The Money) to $S_0 \gg K$, OTM (Out The Money) to $S_0 \ll K$.

▷ **Black-Scholes Call price function.** For convenience of the reader, we give the *Black-Scholes Call price* function depending on log-spot x, total variance y and log-strike z:

$$\mathrm{Call}^{\mathrm{BS}}(x, y, z) = e^x \mathcal{N}(d_1(x, y, z)) - e^z \mathcal{N}(d_2(x, y, z)) \tag{1.7}$$

where:

$$\mathcal{N}(x) = \int_{-\infty}^{x} \mathcal{N}'(u)\mathrm{d}u, \qquad \mathcal{N}'(u) = \frac{e^{-u^2/2}}{\sqrt{2\pi}},$$

[1] for non-zero but deterministic risk-free rate, we are reduced to the previous case by considering the discounted asset; see also the discussion in [Benhamou 2012] for stochastic interest rates.

$$d_1(x,y,z) = \frac{x-z}{\sqrt{y}} + \frac{1}{2}\sqrt{y}, \qquad d_2(x,y,z) = d_1(x,y,z) - \sqrt{y}.$$

This value $\mathrm{Call}^{\mathrm{BS}}(x,y,z)$ equals $\mathrm{Call}(e^x,T,e^z)$ in (1.6) when the volatility in (1.4) is only time-dependent and $y = \int_0^T \sigma^2(t)\mathrm{d}t$. Note that $\mathrm{Call}^{\mathrm{BS}}$ is a smooth function (for $y > 0$) and there is in addition a simple relation between its partial derivatives:

$$\partial_y \mathrm{Call}^{\mathrm{BS}}(x,y,z) = \frac{1}{2}(\partial^2_{x^2} - \partial_x)\mathrm{Call}^{\mathrm{BS}}(x,y,z) = \frac{1}{2}(\partial^2_{z^2} - \partial_z)\mathrm{Call}^{\mathrm{BS}}(x,y,z). \tag{1.8}$$

In the following, $x_0 = \log(S_0)$ (which is the initial value of the process X defined in (1.5)) will represent the log-spot, $k = \log(K)$ the log-strike, $x_{avg} = (x_0+k)/2 = \log(\sqrt{S_0K})$ the mid-point between the log-spot and the log-strike, $m = x_0 - k = \log(S_0/K)$ the log-moneyness.
The reader can find in Proposition 2.6.1.2 of Chapter 2 Section 2.6 the definition of $\mathrm{Vega}^{\mathrm{BS}}$, $\mathrm{Vomma}^{\mathrm{BS}}$ and $\mathrm{Ultima}^{\mathrm{BS}}$ which are the first three derivatives of $\mathrm{Call}^{\mathrm{BS}}$ w.r.t. a volatility parameter.

For (x,T,z) given, the *implied Black-Scholes volatility* of a price $\mathrm{Call}(e^x,T,e^z)$ is the unique non-negative parameter $\sigma_{\mathrm{I}}(x,T,z)$ such that

$$\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2_{\mathrm{I}}(x,T,z)T,z) = \mathrm{Call}(e^x,T,e^z). \tag{1.9}$$

▷ **Bachelier Call price function.** We now recall the *Bachelier Call price* as a function of spot S, total variance Y and strike Z:

$$\mathrm{Call}^{\mathrm{BA}}(S,Y,Z) = (S-Z)\mathcal{N}\Big(\frac{S-Z}{\sqrt{Y}}\Big) + \sqrt{Y}\mathcal{N}'\Big(\frac{S-Z}{\sqrt{Y}}\Big), \tag{1.10}$$

which coincides with $\mathrm{Call}(S,T,Z)$ when the volatility in (1.4) is such that $x\sigma(t,x) = \Sigma(t)$ and $Y = \int_0^T \Sigma^2(t)\mathrm{d}t$. The function $\mathrm{Call}^{\mathrm{BA}}$ is smooth (for $Y > 0$) and we have:

$$\partial_Y \mathrm{Call}^{\mathrm{BA}}(S,Y,Z) = \frac{1}{2}\partial^2_{S^2}\mathrm{Call}^{\mathrm{BA}}(S,Y,Z) = \frac{1}{2}\partial^2_{Z^2}\mathrm{Call}^{\mathrm{BA}}(S,Y,Z).$$

We frequently use the notation $S_{avg} = (S_0+K)/2$ and $M = S_0 - K$ for the *Bachelier moneyness*. Proposition 2.6.2.2 postponed to Chapter 2 Section 2.6 defines the sensitivities of $\mathrm{Call}^{\mathrm{BA}}$ w.r.t. the volatility parameter: $\mathrm{Vega}^{\mathrm{BA}}$, $\mathrm{Vomma}^{\mathrm{BA}}$ and $\mathrm{Ultima}^{\mathrm{BA}}$.

For (S,T,Z) given, the *implied Bachelier volatility* of a price $\mathrm{Call}(S,T,Z)$ is the unique non-negative parameter $\Sigma_{\mathrm{I}}(S,T,Z)$ such that

$$\mathrm{Call}^{\mathrm{BA}}(S,\Sigma^2_{\mathrm{I}}(S,T,Z)T,Z) = \mathrm{Call}(S,T,Z). \tag{1.11}$$

Black-Scholes and Bachelier implied volatilities are compared in [Schachermayer 2008].

1.3 An overview of approximation results

The increasing need in evaluating financial risks at a very global level and in a context of high-frequency market exchanges is a significant incentive for the computational methods to be efficient in evaluating the exposure of large portfolio to market fluctuations (VaR computations, sensitivity analysis), in quickly calibrating the models to the market data. Hence, in the two last decades, many numerical methods have been developed to meet these objectives: in particular, regarding the option pricing, several approximation results have been derived, following one or another asymptotic point of view. We give a summary of these different approaches, stressing the limits of applicability of the methods.

1.3.1 Large and small strikes, at fixed maturity

The Call price $\mathrm{Call}(S_0, T, K)$ as a function of strike is convex and its left/right derivatives are related to the distribution function of S_T [Musiela 2005, Chapter 7]: $\partial_K^- \mathrm{Call}(S_0, T, K) = -\mathbb{P}(S_T \geq K)$ and $\partial_K^+ \mathrm{Call}(S_0, T, K) = -\mathbb{P}(S_T > K)$. Beyond the important fact that the family of Call prices $\{\mathrm{Call}(S_0, T, K) : K \geq 0\}$ completely characterizes the marginal distribution of S_T, this relation also shows that the tails of the law of S_T are intrinsically related to the decay of $\mathrm{Call}(S_0, T, K)$ as $K \to +\infty$. In terms of implicit volatility, the heuristics is the following: the larger the implied volatility of OTM options, the larger the right tail of S_T. This is similar for small strikes K, using Put options. Lee [Lee 2005] has been the first one to quantify these features relating the behavior of implied volatility to the tails of S_T, with an encoding of the tails through the existence of positive/negative moments. These are model-free relations, that can be applied to any model with $\mathbb{E}(S_T) < +\infty$ and not only to local volatility ones like in (1.4). The well-known Lee moment formulas write as follows, using the log-variables $x_0 = \log(S_0)$ and $m = \log(S_0/K) = x_0 - k$.

Theorem 1.3.1.1. *Define*
• *the maximal finite positive moment order* $p_R := \sup\{p \geq 0 : \mathbb{E}(S_T^{1+p}) < +\infty\}$,
• *the maximal finite negative moment order*[2] $p_L := \sup\{p \geq 0 : \mathbb{E}(S_T^{-p}) < +\infty\}$. *Then, the right tail-wing of the Black-Scholes implied volatility defined in* (1.9) *is such that*

$$\limsup_{m \to -\infty} \frac{T\sigma_{\mathrm{I}}^2(x_0, T, x_0 - m)}{|m|} = \phi(p_R) := \beta_R,$$

while the left tail-wing is such that

$$\limsup_{m \to +\infty} \frac{T\sigma_{\mathrm{I}}^2(x_0, T, x_0 - m)}{m} = \phi(p_L) := \beta_L,$$

where $\phi(x) = 2 - 4(\sqrt{x^2 + x} - x) \in [0, 2]$.

Proof. We refer to [Lee 2005] for a detailed proof. We only give the two main arguments for proving the right tail-wing, the left one being similar.
• The *first argument* relies on a tight connection between moments and asymptotics of Call/Put as $K \to +\infty$. Indeed, on the one hand, convexity inequalities give $(s - K)_+ \leq \frac{s^{p+1}}{p+1}(\frac{p}{p+1})^p \frac{1}{K^p}$ (for $p \geq 0$), and taking the expectation yields

$$\mathrm{Call}(S_0, T, K) \leq \frac{\mathbb{E}(S_T^{p+1})}{p+1}\left(\frac{p}{p+1}\right)^p \frac{1}{K^p}. \tag{1.12}$$

In other words, the more integrability of S_T, the faster the decay of $\mathrm{Call}(S_0, T, K)$ as $K \to +\infty$. Conversely, the Carr formula states that the Call/Put prices form a pricing generating system for any payoff equal to a difference of convex functions: making this principle particular to the *power* payoff, we obtain

$$\mathbb{E}(S_T^{1+p}) = \int_0^\infty p(p+1)K^{p-1}\mathrm{Call}(S_0, T, K)\mathrm{d}K, \tag{1.13}$$

i.e. the faster the decreasing of $\mathrm{Call}(S_0, T, K)$ as $K \to +\infty$, the higher the integrability of S_T.
• The *second argument* is based on exponential decreasing behaviors of Call/Put in terms of Black-Scholes implied volatility, as the log-moneyness $m \to \pm\infty$. Reparameterizing the implied volatility

[2] $1 + p_R$ and p_L are respectively called right-tail and left-tail indices.

$\sigma_{\rm I}(x_0,T,x_0-m) = \sqrt{\beta|m|/T}$ with $\beta \in (0,2]$ (β is interpreted as a slope of the total variance per log-moneyness), we obtain

$$\mathrm{Call}^{\mathrm{BS}}(x_0,\beta|m|,x_0-m) = S_0\mathcal{N}(-\sqrt{f_-(\beta)|m|}) - S_0e^{-m}\mathcal{N}(-\sqrt{f_+(\beta)|m|})$$

where $f_\pm(\beta) = \frac{1}{\beta}+\frac{\beta}{4}\pm 1$. Then, a direct computation shows a dichotomic behavior related to β:

$$\lim_{m\to-\infty} e^{-cm}\mathrm{Call}^{\mathrm{BS}}(x_0,\beta|m|,x_0-m) = +\infty\mathbb{1}_{c>f_-(\beta)/2}. \tag{1.14}$$

Comparing (1.12-1.13-1.14) and setting $p_R := f_-(\beta_R)/2$ (or equivalently $\beta_R = \phi(p_R)$) yields the tail-wing formulas. □

Since the original contribution of Lee, several improvements to Theorem 1.3.1.1 have been established. For instance, the lim sup can be removed by a simple limit, under the additional assumptions that S_T has a regularly varying density, see [Benaim 2009]. More recently, Gulisashvili [Gulisashvili 2010] and his co-authors have proved refined expansions of the form

$$\sigma_{\rm I}(x_0,T,k) = \frac{\sqrt{2}}{\sqrt{T}}\Big[\sqrt{\log K + \log\frac{1}{\mathrm{Call}(S_0,T,K)} - \frac{1}{2}\log\log\frac{1}{\mathrm{Call}(S_0,T,K)}}$$
$$-\sqrt{\log\frac{1}{\mathrm{Call}(S_0,T,K)} - \frac{1}{2}\log\log\frac{1}{\mathrm{Call}(S_0,T,K)}}\Big] + O((\log\frac{1}{\mathrm{Call}(S_0,T,K)})^{-\frac{1}{2}})$$

as K becomes large, which allows precise asymptotics of $\sigma_{\rm I}(x_0,T,k)$ through those of $\mathrm{Call}(S_0,T,K)$.

These kinds of asymptotics are now well-known for most of the usual models, like CEV models (no right tail-wing), Heston model (tail-wing depending on the maturity)... see [Gulisashvili 2012] for more references. Different models may have the same strike asymptotics. We can use this information on extreme strikes in different manners: first, comparing with the asymptotic market implied volatility smile, it allows for selecting a coherent model. Second, it helps the calibration procedure by setting approximately some parameter values (those having an impact on the tails). Third, it can be used to appropriately extrapolate market data.

In practice, these asymptotic formulas refer to far OTM or ITM options, for which the accuracy of market data is really questionable (large bid-ask spread, low liquidity). Thus, a direct application is usually not straightforward.

1.3.2 Long maturities, at fixed strike

Another asymptotics is large maturity. It has been studied by Rogers and Tehranchi, see [Tehranchi 2009] and [Rogers 2010], proving the following.

Theorem 1.3.2.1. *Assume that S remains positive with probability 1. Then, for any $\lambda > 0$,* $\lim_{T\to+\infty}\sup_{|m|\le\lambda}\left|\sigma_{\rm I}(x_0,T,x_0-m) - \sqrt{\frac{8}{T}|\ln(\mathbb{E}(S_T\wedge S_0))|}\right| = 0.$

As before, the proof is based on the careful derivation of asymptotics of Black-Scholes formula (1.7). The above limit states that for strikes in a fixed neighborhood of the spot S_0, the implied volatility behaves like $\sqrt{\frac{8}{T}|\ln(\mathbb{E}(S_T\wedge S_0))|}$ for large maturity, and thus it does not depend on the strike. In other words, the implied volatility surface flattens as maturity becomes large, which is coherent with market data. There are also some refined and higher order asymptotics: assuming that the *a.s.* large-time limit of the martingale S is 0, then

$$T\sigma_{\rm I}^2(x_0,T,k) = 8|\ln(\mathbb{E}(S_T\wedge\frac{K}{S_0}))| - 4\ln(|\ln(\mathbb{E}(S_T\wedge\frac{K}{S_0}))|) + 4\ln(\frac{K}{\pi S_0}) + o(1),$$

where the reminder is locally uniform in the log-moneneyss $m = x_0 - k$.

1.3.3 Long maturities, with large/small strikes

In view of the preceeding results, the asymptotics of the smile for large maturity becomes very simple regarding the strike variable, unless one allows the strike to be large/small together with the maturity. Indeed to recover interesting information at the limit, we should consider strikes of the form $K = S_0e^{xT}$ with $x \neq 0$, or equivalently $k = x_0 + xT$. From the linearization of the payoff, one obtains

$$\begin{aligned}\text{Call}(S_0, T, S_0e^{xT}) &= \mathbb{E}(S_T \mathbb{1}_{S_T \geq S_0e^{xT}}) - S_0e^{xT}\mathbb{P}(S_T \geq S_0e^{xT})\\ &= S_0\mathbb{P}^S\Big(\frac{1}{T}\log(S_T/S_0) \geq x\Big) - S_0e^{xT}\mathbb{P}\Big(\frac{1}{T}\log(S_T/S_0) \geq x\Big)\end{aligned}$$

where the new measure $\mathbb{P}^S$ is the one associated to the numéraire S. Under this form, it appears clearly that for x large enough (say larger than the asymptotic $\mathbb{P}$-mean or $\mathbb{P}^S$-mean of $\frac{1}{T}\log(S_T/S_0)$ whenever it exists), both probabilities above correspond to the evaluation of large deviation events. The role of Large Deviation Principle satisfied by the sequence $(\frac{1}{T}\log(S_T/S_0))_{T\geq 0}$ as $T \to +\infty$ has been outlined in [Forde 2011] in the case of Heston model, and in [Jacquier 2011] for more general affine models. Saddle point arguments combined with Lewis formula (1.3) have been performed in [Gatheral 2011] for the Heston model, to recover the *Stochastic Volatility Inspired* parameterization of Gatheral [Gatheral 2004]: the squared implied volatility $\sigma_{\text{I}}^2(x_0, T, x_0 + xT)$ has the simple asymptotic shape

$$\sigma_\infty^2(x) = \frac{\omega_1}{2}(1 + \omega_2\rho x + \sqrt{(\omega_2 x + \rho)^2 + 1 - \rho^2}). \tag{1.15}$$

For more general affine models like Heston model, without or with jumps, or Bates model, or Barndorff-Nielsen-Shephard model (see [Duffie 2003] and [Jacquier 2011]), it is possible to derive similar limits. Let $\Lambda_t(u) = \log(\mathbb{E}(S_t^u))$ be the exponent of the moment generating function, which is convex in u: in the aforementioned model we can define and compute its asymptotic average $\Lambda(u) = \lim_{t\to\infty}\frac{1}{t}\Lambda_t(u)$, which still satisfies to the convexity feature. We associate its Fenchel-Legendre transform $\Lambda^*(x) = \sup_{u\in\mathbb{R}}(ux - \Lambda(u))$ and it turns out that $(\frac{1}{T}\log(S_T/S_0))_{T\geq 0}$ satisfies a LDP under $\mathbb{P}$ (resp. $\mathbb{P}^S$) with rate function $x \mapsto \Lambda^*(x)$ (resp. $x \mapsto \Lambda^*(x) - x$).

Theorem 1.3.3.1. *Under some assumptions (see [Jacquier 2011]), for any $x \in \mathbb{R}$, the asymptotic implied volatility $\sigma_\infty(x)$ is given by*

$$\lim_{T\to\infty}\sigma_{\text{I}}(x_0, T, x_0 + xT) = \sqrt{2}\Big[\text{sgn}(\Lambda'(1) - x)\sqrt{\Lambda^*(x) - x} + \text{sgn}(x - \Lambda'(0))\sqrt{\Lambda^*(x)}\Big].$$

In the Black-Scholes model with constant volatility σ, one has $\Lambda(u) = \frac{\sigma^2}{2}(u^2 - u)$, $\Lambda^*(x) = \frac{1}{2\sigma^2}(x + \frac{\sigma^2}{2})^2$, and we get obviously $\sigma_\infty(x) = \sigma$. For Heston model, Λ is explicit as well and we finally recover the SVI parsimonious parameterization (1.15). Here again, different models may have the same asymptotic smiles, see [Jacquier 2011].

1.3.4 Non large maturities and non extreme strikes

To obtain approximation formulas in that situation, the asymptotics should originate from different large/small parameters that are rather related to the model and not to the contract characteristics (maturity and strike). These different asymptotics are generally well intuitively interpreted. For the sake of clarity, we spend time to detail a bit the arguments, in order to make clearer the differences between the further expansion results and the tools to obtain them. To the best of our knowledge, such comparative presentation does not exist in the literature and the reader may find it interesting.

1.3.4.1 Small noise expansion

This is inspired by the Freidlin-Wentzell approach [Freidlin 1998] in which the noise in the Stochastic Differential Equation of interest is small. Denote by Y the scalar SDE under study (which can be X or S in our framework), solution of

$$\mathrm{d}Y_t = \mu(t, Y_t)\mathrm{d}t + \nu(t, Y_t)\mathrm{d}W_t, \qquad Y_0 \text{ given.} \tag{1.16}$$

Assume that ν is small, or equivalently that ν becomes $\varepsilon\nu$ with $\varepsilon \to 0$: after making this small noise parameterization, the model writes

$$\mathrm{d}Y_t^\varepsilon = \mu(t, Y_t^\varepsilon)\mathrm{d}t + \varepsilon\nu(t, Y_t^\varepsilon)\mathrm{d}W_t, \quad Y_0^\epsilon = Y_0.$$

For $\varepsilon = 0$, it reduces to an ODE

$$Y_t^0 = y_{0,t} = Y_0 + \int_0^t \mu(s, y_{0,s})\mathrm{d}s \tag{1.17}$$

and this deterministic model serves as a zero-order approximation for the further expansion. Under smooth coefficient assumptions [Freidlin 1998], we can derive a stochastic expansion of Y^ε in powers of ε:

$$Y_t^\varepsilon = y_{0,t} + \varepsilon Y_{1,t} + \frac{1}{2}\varepsilon^2 Y_{2,t} + o(\varepsilon^2). \tag{1.18}$$

For instance Y_1 solves a linear Gaussian SDE

$$Y_{1,t} = \int_0^t \partial_x\mu(s, y_{0,s})Y_{1,s}\mathrm{d}s + \int_0^t \nu(s, y_{0,s})\mathrm{d}W_s = \int_0^t e^{\int_s^t \partial_x\mu(r,y_{0,r})\mathrm{d}r}\nu(s, y_{0,s})\mathrm{d}W_s.$$

Similarly, Y_2 solves

$$\begin{aligned} Y_{2,t} &= \int_0^t [\partial_x\mu(s, y_{0,s})Y_{2,s} + \partial_x^2\mu(s, y_{0,s})Y_{1,s}^2]\mathrm{d}s + \int_0^t 2\partial_x\nu(s, y_{0,s})Y_{1,s}\mathrm{d}W_s \\ &= \int_0^t e^{\int_s^t \partial_x\mu(r,y_{0,r})dr}(\partial_x^2\mu(s, y_{0,s})Y_{1,s}^2\mathrm{d}s + 2\partial_x\nu(s, y_{0,s})Y_{1,s}\mathrm{d}W_s). \end{aligned}$$

Higher order expansions are available under higher smoothness assumptions. The notation $o(\varepsilon^2)$ in (1.18) means that the related error term has a L^p-norm (for any p) that can be neglected compared to ε^2 as $\varepsilon \to 0$. The stochastic expansion (1.18) becomes a weak expansion result when we compute $\mathbb{E}(h(Y_T))$ for a test function h.

▷ The case of smooth h. If h is smooth enough, we obviously obtain

$$\begin{aligned} \mathbb{E}(h(Y_T)) =& h(y_{0,T}) + \varepsilon h'(y_{0,T})\mathbb{E}(Y_{1,T}) \\ &+ \varepsilon^2(h'(y_{0,T})\mathbb{E}(\frac{Y_{2,T}}{2}) + \frac{1}{2}h''(y_{0,T})\mathbb{E}(Y_{1,T}^2)) + o(\varepsilon^2). \end{aligned}$$

Observe that $\mathbb{E}(Y_{1,T}) = 0$ since $Y_{1,T}$ is a Wiener integral. To make the above expansion fully effective in practice, it is necessary to make the coefficients $\mathbb{E}(Y_{2,T})$ and $\mathbb{E}(Y_{1,T}^2)$ explicit: this is quite straightforward thanks to the linear equations solved by $Y_{1,.}$ and $Y_{2,.}$. The L^2-isometry property of the Wiener integral yields $\mathbb{E}(Y_{1,t}^2) = \int_0^t e^{2\int_s^t \partial_x\mu(r,y_{0,r})\mathrm{d}r}\nu^2(s, y_{0,s})\mathrm{d}s$. In addition, we have $\mathbb{E}(Y_{2,t}) = \int_0^t e^{\int_s^t \partial_x\mu(r,y_{0,r})\mathrm{d}r}\partial_x^2\mu(s, y_{0,s})\mathbb{E}(Y_{1,s}^2)\mathrm{d}s$. The coefficients computation is reduced to the evaluation of nested time-integrals which are simple to compute using standard n-points integral discretization, with a computational complexity[3] of order n. The above expansion analysis is a *regular perturbation* analysis, using a *stochastic analysis* point of view.

[3] Observe that although the time integrals are multidimensional, we are reduced to iterative one-dimensional computations since the function to integrate is separable in all its variables.

To derive this expansion in powers of ε, we could alternatively use a *PDE point of view* based on Feynman-Kac representation, which states that $u^\varepsilon:(t,x)\mapsto u^\varepsilon(t,x)=\mathbb{E}(h(Y_T^\varepsilon)|Y_t^\varepsilon=x)$ solves the perturbed PDE

$$\begin{cases} \partial_t u^\varepsilon(t,x)+\mu(t,x)\partial_x u^\varepsilon(t,x)+\frac{1}{2}\varepsilon^2\nu^2(t,x)\partial_x^2 u^\varepsilon(t,x)=0 & \text{for } t<T,\\ u^\varepsilon(T,x)=h(x). \end{cases}$$

Setting $\mathcal{L}^\varepsilon=\partial_t+\mu\partial_x+\frac{1}{2}\varepsilon^2\nu^2\partial_x^2:=\mathcal{L}_0+\varepsilon^2\mathcal{L}_2$, the above PDE writes $\mathcal{L}^\varepsilon u^\varepsilon=0$ plus boundary conditions at time T. Seeking an expansion of the form $u^\varepsilon=u_0+\varepsilon u_1+\frac{1}{2}\varepsilon^2u_2+o(\varepsilon^2)$, we obtain

$$\mathcal{L}_0u_0+\varepsilon\mathcal{L}_0u_1+\varepsilon^2[\frac{1}{2}\mathcal{L}_0u_2+\mathcal{L}_2u_0]+o(\varepsilon^2)=0.$$

A formal identification of each coefficient of ε^i $(i=0,1\dots)$ to 0, we obtain a system of PDEs:

$$\mathcal{L}_0u_0=0,\qquad \mathcal{L}_0u_1=0,\qquad \frac{1}{2}\mathcal{L}_0u_2+\mathcal{L}_2u_0=0,$$

with the boundary conditions $u_0(T,.)=h(.),u_1(T,.)=u_2(T,.)=0$. The justification of this kind of expansion and its related error analysis can be made under appropriate smoothness assumptions on h, μ and ν; we refer to [Fleming 1986, Theorem 5.1], [Fournié 1997, Theorem 3.1] or [Fouque 2011, Chapter 4] where a similar error analysis is made. The PDE solutions are then given by

$$u_0(t,x)=h(y_T^{t,x}),\qquad u_1\equiv 0,\qquad u_2(t,x)=\int_t^T 2\mathcal{L}_2u_0(s,y_s^{t,x})\mathrm{d}s$$

where $(y_s^{t,x})_{s\geq t}$ stands for the solution of the ODE (1.17) with initial condition (t,x). Under this form of system of PDEs, the derivation of an explicit expression for u_2 is not as easy as within the stochastic analysis approach. However, we can obtain the same expansion (fortunately!), i.e. the same formula for u_2 at $(0,Y_0)$:

$$u_2(0,Y_0)=h'(y_{0,T})\mathbb{E}(Y_{2,T})+h''(y_{0,T})\mathbb{E}(Y_{1,T}^2) \tag{1.19}$$

with $\mathbb{E}(Y_{2,T})$ and $\mathbb{E}(Y_{1,T}^2)$ given as before. To see this, start from $\mathcal{L}_2$ and write $u_2(t,x)=\int_t^T\nu^2(s,y_s^{t,x})\partial_y^2u_0(s,y_s^{t,x})\mathrm{d}s$. We have $\partial_xu_0(t,x)=h'(t,y_T^{t,x})\partial_xy_T^{t,x}$ and $\partial_x^2u_0(t,x)=h'(t,y_T^{t,x})\partial_x^2y_T^{t,x}+h''(t,y_T^{t,x})(\partial_xy_T^{t,x})^2$. Then to recover (1.19), use the notation $y_{0,t}=y_t^{0,Y_0}$, the flow property $y_s^{t,y_{0,t}}=y_{0,s}$ for $s\geq t$, and the explicit expressions for $\partial_xy_s^{t,x}$ and $\partial_x^2y_s^{t,x}$: for instance $\partial_xy_s^{t,x}=1+\int_t^s\partial_x\mu(r,y_r^{t,x})\partial_xy_r^{t,x}\mathrm{d}r=e^{\int_t^s\partial_x\mu(r,y_r^{t,x})\mathrm{d}r}$. We skip further details. This completes the *PDE* approach to derive a *regular perturbation* analysis. Observe that the derivation of explicit formula is delicate because of the system of PDEs to solve (more complicate than solving the system of SDEs arising within the stochastic analysis approach).

▷ THE CASE OF NON-SMOOTH h. The previous derivation which involves h',h'' and possibly higher derivatives is mathematically incorrect if h is not smooth. This fact is clear using the stochastic analysis approach. It is also clear using PDE arguments: indeed, it would involve the perturbed PDE solution $(t,x)\mapsto\mathbb{E}(h(Y_T^\varepsilon)|Y_t^\varepsilon=x)$, that is not uniformly smooth (because the regularization parameter ε shrinks to 0). If $h(x)=\mathbb{1}_{x\geq K}$ (like digital payoff), i.e. we evaluate $p^\varepsilon=\mathbb{P}(Y_T^\varepsilon\geq K)$, and $y_{0,T}\neq K$, large deviation arguments [Azencott 83] show that the probability p^ε is exponentially close to 0 or 1 w.r.t. $1/\varepsilon^2$ (i.e. $\log(p^\varepsilon)\approx -c/\varepsilon^2$ if $y_{0,T}<K$), and thus an expansion in power of ε provides zero coefficients at any order. To get a non degenerate and interesting situation, we should consider the case K is close to $y_{0,T}$ in the sense $K=y_{0,T}+\lambda\,\varepsilon$, that is

$$p^\varepsilon=\mathbb{P}(Y_T^\varepsilon\geq y_{0,T}+\lambda\,\varepsilon)=\mathbb{P}(\frac{Y_T^\varepsilon-y_{0,T}}{\varepsilon}\geq\lambda).$$

In other words, to overcome the difficulty of the singularity of h, we have leveraged a homogenization argument (*singular perturbation*), by considering the rescaled variable (usually called fast variable) $Z_t^\varepsilon = \frac{Y_t^\varepsilon - y_{0,t}}{\varepsilon} = Y_{1,t} + \frac{1}{2}\varepsilon Y_{2,t} + o(\varepsilon)$. If the law of $Y_{1,T}$ is non degenerate (for instance Gaussian law with non-zero variance), the latter quantity can be expanded in powers of ε. Actually, for less specific functions h, Watanabe [Watanabe 1987] has developed a Malliavin calculus-based machinery to establish a general expansion result of $\mathbb{E}(h(Z_T^\varepsilon))$ in powers of ε, available even for Schwartz distributions h, assuming stochastic expansions in Malliavin sense of

$$Z_T^\varepsilon = Z_{0,T} + \varepsilon Z_{1,T} + \varepsilon^2 Z_{2,T} + \cdots + \varepsilon^n Z_{n,T} + O(\varepsilon^{n+1})$$

for any $n \geq 1$ and asymptotic (in ε) non-degeneracy in Malliavin sense of Z_T^ε:

$$\limsup_{\varepsilon \to 0} \|1/\det(\gamma^{Z_T^\varepsilon})\|_p < +\infty \tag{1.20}$$

for any $p \geq 1$, where γ^Z is the Malliavin covariance matrix of a random variable Z. The Watanabe result states the existence of random variables $(\pi_k)_{k\geq 1}$ such that for any polynomially bounded function h, we have

$$\mathbb{E}(h(Z_T^\varepsilon)) = \mathbb{E}(h(Z_{0,T})) + \sum_{k=1}^n \varepsilon^k \mathbb{E}(h(Z_{0,T})\pi_k) + O(\varepsilon^{n+1}), \qquad \forall n \geq 1. \tag{1.21}$$

Compared to the non-smooth case, the possibility to get an expansion result is due to the non-degeneracy condition which has a (asymptotic) regularization effect on the non-smooth function h. With our previous notation $Z_T^\varepsilon = \frac{Y_T^\varepsilon - y_{0,T}}{\varepsilon} = Y_{1,T} + \frac{1}{2}\varepsilon Y_{2,T} + o(\varepsilon)$, the asymptotic non-degeneracy (1.20) implies that the Gaussian random variable $Y_{1,T}$ has a non-zero variance, i.e. $\int_0^T e^{2\int_s^T \partial_x \mu(r, y_{0,r}) \mathrm{d}r} \nu^2(s, y_{0,s}) \mathrm{d}s > 0$: in the case of time-independent coefficient $\mu(s,y) = \mu(y), \nu(s,y) = \nu(y)$, it reads $\nu(y_{0,T}) \neq 0$. The converse result ($\nu(y_{0,T}) \neq 0$ implies (1.20)) holds true in the case of time-independent coefficient and in a multidimensional setting, see [Watanabe 1987, Theorem 3.4]. Yoshida [Yoshida 1992b, Theorem 2.2] has weakened the assumption (1.20) into a localized version allowing degeneracy on a set of polynomially small probability measure. This approach has been successfully applied to different pricing problems in finance, mainly by Yoshida, Takahashi and their co-authors: see [Yoshida 1992a, Uchida 2004, Kunitomo 2001] or the unpublished work [Osajima 2007]. Their methodology consists in expanding the density of the random variable $Z_T^\varepsilon = \frac{Y_T^\varepsilon - y_{0,T}}{\varepsilon}$ using the Gaussian density as the zero-order term, and then going back to $\mathbb{E}(h(Y_T^\varepsilon))$ by suitable integration computations. The advantage of this approach is that the expansion result (1.21) holds in a large generality, provided that we assume infinitely differentiable coefficients and uniform non degeneracy. However, observe two difficulties or restrictions:

- within usual financial models like Heston model, the required regularity assumption is not satisfied and we even know that the Malliavin differentiability of high order may fail, see [Alòs 2008].

- the existence of the Malliavin weights $(\pi_k)_k$ does not provide an explicit and numerically computable expansion: very involved additional computations are required to obtain explicit formulas. One might compare these tricky computations to those necessary to solve the aforementioned system of PDEs.

Last, this approach usually leads to normal approximations of financial models (Bachelier prices) whereas log-normal approximations (Black-Scholes prices) might be more accurate (numerical evidences are given in Chapter 2 Section 2.5).

After this presentation of *singular perturbation* using stochastic analysis, we now turn to the PDE approach. It has been developed in the financial context by Hagan and co-authors [Hagan 1999, Hagan 2002]. To be as close as possible to the quoted work, assume that the drift coefficient is $\mu \equiv 0$. In the case of Call payoff, the original valuation PDE u^ε writes

$$\begin{cases} \partial_t u^\varepsilon(t,x) + \frac{1}{2}\varepsilon^2 v^2(t,x)\partial_x^2 u^\varepsilon(t,x) = 0 \quad \text{for } t < T, \\ u^\varepsilon(T,x) = (x-K)_+; \end{cases}$$

now, if we consider ATM strikes ($K - Y_0 = O(\varepsilon)$ similarly to before), we should consider the fast variable $y = (x-K)/\varepsilon$ and the rescaled solution $v^\varepsilon(t,y) = \frac{1}{\varepsilon}u^\varepsilon(t, K+\varepsilon y)$ which solves

$$\begin{cases} \partial_t v^\varepsilon(t,y) + \frac{1}{2}v^2(t, K+\varepsilon y)\partial_y^2 v^\varepsilon(t,y) = 0 \quad \text{for } t < T, \\ v^\varepsilon(T,y) = y_+. \end{cases} \tag{1.22}$$

At this stage, the analysis follows the routine similar to before, by seeking a solution under the form

$$v^\varepsilon = v_0 + \varepsilon v_1 + o(\varepsilon) \tag{1.23}$$

solving $\mathcal{L}^\varepsilon v^\varepsilon = 0$ where $\mathcal{L}^\varepsilon = \partial_t + \frac{1}{2}v^2(t, K+\varepsilon y)\partial_y^2 = \mathcal{L}_0 + \varepsilon\mathcal{L}_1 + o(\varepsilon)$ with $\mathcal{L}_0 = \partial_t + \frac{1}{2}v^2(t,K)\partial_y^2$, $\mathcal{L}_1 = vv'(t,K)y\partial_y^2$. A formal identification leads to a system of PDEs:

$$\mathcal{L}_0 v_0 = 0,\ v_0(T,y) = y_+ \quad \text{and} \quad \mathcal{L}_0 v_1 + \mathcal{L}_1 v_0 = 0,\ v_1(T,y) = 0.$$

The solution v_0 is obviously given by the Call price in a Bachelier model (1.10) $\mathrm{d}X_t^{\mathrm{BA}} = v(t,K)\mathrm{d}W_t$ with time-dependent diffusion coefficient, and the first correction is given by $v_1(t,y) = \mathbb{E}(\int_t^T \mathcal{L}_1 v_0(s, X_s^{\mathrm{BA}})\mathrm{d}s | X_t^{\mathrm{BA}} = y)$. Although the new terminal function $h(y) = y_+$ is not infinitely smooth, non-zero function v induces a smoothing effect due to a non-degenerate heat kernel (this feature is analogous to the previous non-degeneracy in the Malliavin sense): hence, v_0 is smooth with derivatives possibly blowing up as time gets close to T and a careful analysis shows that v_1 is well defined too. Here again, the explicit computation of v_1 is not an easy exercise and it requires some tricks. Finally v_1 can be written as the weighted sum of derivatives of v_0 (interpreted as Greeks). In Chapter 2 Sections 2.1 and 2.3, we provide a more direct and generic way to compute this kind of correction terms using stochastic analysis instead of PDE arguments.

Regarding the careful justification of the above PDE regular expansion with error estimates, quite surprisingly we have not been able to find literature references when the terminal condition is non-smooth (like $y \mapsto y_+$). We nevertheless cite the works of Fouque et al. [Fouque 2003, Fouque 2004] in which error estimates are provided for Call options in the homogenization framework (see below). But these estimates are obtained at the cost of loss of accuracy during the regularization step of the payoff function.

Once obtained the expansion of v^ε for a given local volatility function $\sigma(.,.)$ (i.e. $v(t,y) = y\,\sigma(t,y)$), one can derive an expansion of the Black-Scholes implied volatility σ_{I} by identifying the previous expansion with that in the case ($v_{\mathrm{I}}(t,y) = y\,\sigma_{\mathrm{I}}$): see [Hagan 1999] where the analysis is successfully performed for time-independent volatility $\sigma(t,y) = \sigma(y)$ (or separable function $\sigma(t,y) = \sigma(y)\alpha(t)$ by a simple time-change). It is possible that the case of general time-dependent volatility has been considered out of reach by the authors of [Hagan 1999, Hagan 2002] using PDE arguments, whereas we will see later how much stochastic analysis tools are suitable even in the case time-dependent coefficients.

1.3.4.2 Short maturity

In this asymptotics, the terminal time T is small. When one has to evaluate $\mathbb{E}(h(Y_T))$ for Y solution of the SDE (1.16) and for h smooth (say infinitely differentiable with bounded derivatives), iterative applications of Itô's formula give

$$\mathbb{E}(h(Y_T)) = h(Y_0) + \int_0^T \mathbb{E}([\mathcal{L}h](t, Y_t))\mathrm{d}t$$

$$= h(Y_0) + T[\mathcal{L}h](0, Y_0) + \int_0^T \int_0^t \mathbb{E}([\mathcal{L}^2 h](s, Y_s))\mathrm{d}s\mathrm{d}t$$

where $\mathcal{L}$ is the infinitesimal generator associated to Y. Iterating the procedure, we obtain an expansion in powers of T:

$$\mathbb{E}(h(Y_T)) = \sum_{k=0}^{n} \frac{T^k}{k!}[\mathcal{L}^k h](0, Y_0) + O(T^{n+1}), \qquad n \geq 0.$$

The numerical evaluation of such formula is straightforward. We refer the reader to [Kloeden 1995, Chap. 5] for a more comprehensive exposure of related Itô-Taylor expansions.

As in the case of small noise expansion, the case of non-smooth h requires a different treatment because $\mathcal{L}^k h$ is not defined. For this, we transform the problem of small terminal time with fixed coefficients into a problem of fixed terminal time with small coefficients, by leveraging the scaling property of the Brownian motion. Actually, having T small is equivalent to replace T by $\varepsilon^2 T$ with $\varepsilon \to 0$: then, starting from the SDE (1.16), we consider the time-rescaled process $(Y_{\varepsilon^2 t})_{0\leq t\leq T}$ which has the same distribution as $(Y_t^\varepsilon)_{0\leq t\leq T}$ defined as the solution of

$$\mathrm{d}Y_t^\varepsilon = \varepsilon^2 \mu(\varepsilon^2 t, Y_t^\varepsilon)\mathrm{d}t + \varepsilon\nu(\varepsilon^2 t, Y_t^\varepsilon)\mathrm{d}W_t, \quad Y_0^\epsilon = Y_0, \tag{1.24}$$

see [Watanabe 1987, p.17]. Observe that this leads to a different parameterization compared to the small noise case (in particular, the drift coefficient is multiplied by ε^2). However the expansion methodology is similar: in the case of non-smooth function, it is more appropriate to rescale the process by setting $Z_t^\varepsilon = \frac{Y_t^\varepsilon - y_{0,t}}{\varepsilon} = Y_{1,t} + \frac{1}{2}\varepsilon Y_{2,t} + o(\varepsilon)$, where

$$\begin{aligned}
Y_{1,t} &= \partial_\varepsilon Y_t^\varepsilon|_{\varepsilon=0} = \nu(0, Y_0)W_t,\\
Y_{2,t} &= \partial_\varepsilon^2 Y_t^\varepsilon|_{\varepsilon=0} = 2\mu(0, Y_0)t + 2\partial_y\nu(0, Y_0)\int_0^t Y_{1,s}\mathrm{d}W_s\\
&= 2\mu(0, Y_0)t + \partial_y\nu(0, Y_0)\nu(0, Y_0)(W_t^2 - t).
\end{aligned}$$

Once the *fast variable* is selected, observe that we are reduced to a regular perturbation problem, that can be handled using stochastic analysis tools (namely Watanabe approach [Watanabe 1987]) or using PDE tools. We skip details since it is similar to what have been presented before. See also the book by Henry-Labordère [Henry-Labordère 2008], where short-time asymptotics of density functions are derived through geometry considerations or the work of Gatheral et al. [Gatheral 2012] where the authors obtain approximations of European Call option prices and of implied volatilities for small maturities in local volatility models.

Parametrix approach and Fourier methods. We also mention the so called parametrix approach of Pascucci et al. [Corielli 2010] based on PDE tools and heat kernel expansions. The method consists in decomposing the density of the underlying process as a series of time iterated integrals of heat kernels. Then the authors propose an approximation of the density by truncating the series and performing approximations in short maturity framework. That leads to explicit results which are accurate if the maturity T is small.

In [Pagliarani 2011], Pascucci and Pagliarani revisit the work of Hagan and Woodward [Hagan 1999] to derive an expansion of the transition density in local volatility models in the context of small maturity.

Pascucci, Pagliarani et al. extend the framework to local volatility-Lévy jumps models [Pagliarani 2013b] and local-stochastic volatility with Lévy jumps [Pagliarani 2013a] and provide expansions of the characteristic function. Pricing formulas are obtained using Fourier methods. We also mention the recent work of Lorig, Pascucci and Pagliarani [Lorig 2013b] in which the authors provide a general explicit approximation formula based on Taylor expansions at any order and in any dimension for local and stochastic volatility models.

1.3.4.3 Fast volatility

Since the end of the nineties, another popular approximation approach has been developed by Fouque, Papanicolaou and Sircar, see [Papanicolaou 1999, Fouque 2000]. It emphasizes that the asset volatility $(\sigma_t)_{t\geq 0}$ has usually slow variations compared to the variations of the asset itself (multiscale modeling). This is achieved in two different ways.

- Either the natural time scale of stochastic volatility is short, which leads to a model of the form (1.24) for $(\sigma_t)_t$, while the asset dynamics is unchanged. Thus, at order zero, we obtain a Black-Scholes model with a constant volatility equal to the initial stochastic volatility σ_0, see [Papanicolaou 1999, Section 2].

- Or the fluctuations of the stochastic volatility $(\sigma_t)_t$ are so fast that they give the appearance of a constant (in time) volatility, when considered at a longer time scale. This second point of view has been much developed by Fouque, Papanicolaou and Sircar and their co-authors, in many respects, and this is presented below.

As an illustration of their methodology, we consider the asset model $\mathrm{d}S_t = S_t\sigma_t\mathrm{d}W_t$ and an Ornstein-Uhlenbeck process for modeling $(\sigma_t = f(\Sigma_t))_t$ with

$$\mathrm{d}\Sigma_t = \frac{1}{\varepsilon}(\overline{\Sigma}_\infty - \Sigma_t)\mathrm{d}t + \nu\sqrt{\frac{2}{\varepsilon}}\mathrm{d}B_t.$$

For instance in the Scott model [Scott 1987], $f(x) = e^x$ and (W, B) is a standard bi-dimensional correlated Brownian motion ($\mathrm{d}\langle W, B\rangle_t = \rho\mathrm{d}t$). As time goes to infinity, the random variable Σ_t weakly converges to the stationary Gaussian distribution with mean $\overline{\Sigma}_\infty$ and variance $(\nu\sqrt{\frac{2}{\varepsilon}})^2/(2\frac{1}{\varepsilon}) = \nu^2$. In other words, although the fluctuations are fast (the characteristic time being ε), the distribution remains the same (at least for time larger than ε). It allows the application of ergodic theorem to obtain large-time asymptotics of integrals of the realized volatility: for any polynomially bounded function Ψ, we have

$$\lim_{T\to+\infty}\frac{1}{T}\int_0^T\Psi(\sigma_s)\mathrm{d}s = \int_{\mathbb{R}}\Psi(e^y)\frac{1}{\sqrt{2\pi\nu^2}}e^{-(y-\overline{\Sigma}_\infty)^2/(2\nu^2)}\mathrm{d}y := \sigma_{BS}^2$$

in the almost sure sense and in the L^1-sense. For $\Psi(y) = y^2$, we obtain a constant large-time approximation of $\frac{1}{T}\int_0^T\sigma_s^2\mathrm{d}s$ to be used as a zero-order approximation in a Black-Scholes formula. To derive correction terms, the authors employ *singular pertubation* PDE techniques: indeed, the price function $u^\varepsilon(t, x, y) = \mathbb{E}(h(S_T)|S_t = x, \Sigma_t = y)$ solves $\mathcal{L}^\varepsilon u^\varepsilon = 0$ with

$$\begin{aligned}\mathcal{L}^\epsilon &= \partial_t + \frac{1}{2}x^2f^2(y)\partial^2_{xx} + \sqrt{\frac{2}{\varepsilon}}\rho\nu x f(y)\partial^2_{xy} + \frac{1}{\varepsilon}(\nu^2\partial^2_y + (\overline{\Sigma}_\infty - y)\partial_y)\\ &:= \mathcal{L}_0 + \frac{1}{\sqrt{\varepsilon}}\mathcal{L}_1 + \frac{1}{\varepsilon}\mathcal{L}_2.\end{aligned}$$

As in the previous approaches, by decomposing u^ε in powers of $\sqrt{\varepsilon}$ and by gathering the contributions of the same order, we obtain a system of PDEs characterizing the main order term and the correction terms. Actually the analysis is quite intricate because one has to take into account the ergodic property of σ (which leads to solving elliptic PDEs of the form of Poisson equation): see [Fouque 2000, Chapter 5] where the error analysis is made for smooth payoffs and [Fouque 2003] for the Call options case. The final approximation pricing formula writes

$$\mathbb{E}(S_T - K)_+ = \mathrm{Call}^{\mathrm{BS}}(\log(S_0), T\sigma^2_{BS}, \log(K))$$

$$+\ \sqrt{\varepsilon}\times \text{linear combination of } \partial^i_{S^i}\mathrm{Call}^{\mathrm{BS}}(\log(S_0), T\sigma^2_{BS}, \log(K)) \text{ for } i=2,3+\dots$$

with some explicit coefficients as the factors for the Greeks. Consequently, the approximation formula is straightforward to evaluate on a computer since Black-Scholes price and Greeks are known in closed form and available in any pricing software. In this analysis and similarly to any PDE approaches, assuming time homogeneous coefficients simplifies much the derivation of explicit formula. In the context of fast volatility, some extensions are possible, see [Fouque 2004].

1.3.4.4 Proxy expansion

We complete our overview section by presenting a different point of view, which is going to be developed further in the next Chapters of the thesis. As a difference with previous works, this is rather a non-asymptotic approach, relying on the a priori knowledge of proxy of the model to handle; for this reason, it may appear as more understandable and more intuitive for practioners. Consider the model (1.4) on S:

$$\mathrm{d}S_t = S_t\sigma(t,S_t)\mathrm{d}W_t,$$

and assume that by expertise, S behaves closely to a Gaussian model, i.e. the fluctuations of $S_t\sigma(t,S_t)$ are small. Then, it is reasonable to take the Bachelier model S^P with parameter $(\Sigma_t = S_0\sigma(t,S_0))_t$ as a proxy, that is

$$\mathrm{d}S^P_t = \Sigma_t\mathrm{d}W_t, \qquad S^P_0 = S_0. \qquad \text{(NORMAL PROXY)}$$

The Call price in S model should be close to that in the proxy; since this approximation may be too crude, it is recommended to add correction terms.
Alternatively, one could guess that S rather behaves as a log-normal model with parameter $(a_t)_t$, i.e. $X = \log(S)$ may be approximated by

$$\mathrm{d}X^P_t = -\frac{1}{2}a^2_t\mathrm{d}t + a_t\mathrm{d}W_t, \qquad X^P_0 = x_0. \qquad \text{(LOG-NORMAL PROXY)}$$

The proxy volatility may be taken to $a_t = a(t,x_0) = \sigma(t,S_0)$ for instance, but another point could be chosen (for instance, the strike K or the mid-point $(K+S_0)/2$). This description does not put an emphasize on a specific asymptotic, but one has to quantify how $S_t\sigma(t,X_t) \approx \Sigma_t$ or $a(t,X_t)\approx a_t$.

To compute correction terms to the relation $\mathbb{E}(S_T-K)_+ \approx \mathbb{E}(S^P_T-K)_+$ or $\mathbb{E}(e^{X^P_T}-K)_+$, it is necessary to derive a convenient representation of the distance to the proxy $S_T - S^P_T$ or $X_T - X^P_T$. The linear interpolation $X^\eta_T = X^P_T + \eta(X_T - X^P_T)$ does not lead to illuminating computations. It is much better to consider the following *interpolation*: for $\eta\in[0,1]$, define

$$\mathrm{d}X^\eta_t = \eta(-\frac{1}{2}a^2(t,X^\eta_t)\mathrm{d}t + a(t,X^\eta_t)\mathrm{d}W_t), \quad X^\eta_0 = x_0. \tag{1.25}$$

Note that η is not a small parameter but an *interpolation* parameter. Observe also that this parameterization is different from that in small time or small noise asymptotics.

A direct computation shows that $X^1_t = X_t$, $X^0_t = x_0$ and $\partial_\eta X^\eta_t|_{\eta=0} = \int_0^t a(s,x_0)[\mathrm{d}W_s - \frac{1}{2}a(s,x_0)\mathrm{d}s]$: this shows that $X_t - X^P_t = X^1_t - (X^0_t + \partial_\eta X^\eta_t|_{\eta=0})$ writes as a Taylor formula at order 1. Thus, the natural candidate for the first contribution in $X_t - X^P_t$ is $\frac{1}{2}\partial^2_\eta X^\eta_t|_{\eta=0}$. The above interpolation is equivalent to

$$\mathrm{d}\hat{X}^\eta_t = -\frac{1}{2}a^2(t,x_0+\eta(\hat{X}^\eta_t - x_0))\mathrm{d}t + a(t,x_0+\eta(\hat{X}^\eta_t - x_0))\mathrm{d}W_t, \quad \hat{X}^\eta_0 = x_0, \tag{1.26}$$

which is related to X^η_t by the relation $X^\eta_t = x_0 + \eta(\hat{X}^\eta_t - x_0)$. We refer the reader to Chapter 4 for a discussion on the choice of the proxy process and of the parameterization to link the initial process and its proxy.

1.3.5 Asymptotic expansion versus non-asymptotic expansion

Deriving an asymptotic expansion sheds the light on the crucial role of one model parameter compared to the other ones, to explain and approximate the option prices: for instance, in small noise expansion, we focus only on the volatility by putting an ε in front of the dW-term, and so on. It finally leads to a generic expansion of the form

$$u^{\varepsilon} = u_0 + \varepsilon u_1 + \frac{1}{2}\varepsilon^2 u_2 + \dots \tag{1.27}$$

or in powers of $\sqrt{\varepsilon}$ in the *fast volatility* framework. Our previous discussion has shown how this expansion is obtained in a Markovian framework using PDE (regular or singular pertubations), or more generally using Malliavin calculus (Watanabe approach).

Actually, it is important to observe that writing such an expansion implicitly means that apart of the parameter related to ε, the other parameters have no important asymptotics for the problem under consideration. Below we consider a simple toy example to show that there might be a competition between all the model parameters, and moreover there is a necessary trade-off with the payoff regularity. In other words, deriving (1.27) does not necessarily mean that the first order approximation $u_0 + \varepsilon u_1$ is really accurate and taking more terms do not necessarily improve the accuracy, because of the possible crucial influence of other large or small parameters. Our toy example is the following perturbed Brownian model:

$$X_1^{\varepsilon} = \sigma W_1 + \sqrt{\varepsilon} B_1$$

where (W, B) is a two-dimensional Brownian motion, and σ is positive. This toy model can be viewed as the simplest way to perturb a volatility model (we could have taken $B = W$ without changing the conclusion of the discussion below) and thus, it is quite realistic compared to the further situations to handle.

1. *Case* $h(x) = 1 + x^2$. We have $\mathbb{E}[h(X_1^{\varepsilon})] = 1 + \sigma^2 + \varepsilon = \mathbb{E}[h(X_1^0)] + \varepsilon$.

2. *Case* $h(x) = 1 + x_+$. By a scaling argument, we have:

$$\mathbb{E}[h(X_1^{\varepsilon})] = 1 + \sqrt{\sigma^2 + \varepsilon}\,\mathbb{E}[(W_1)^+] = \mathbb{E}[h(X_1^0)] + \frac{1}{2}\frac{\varepsilon}{\sigma}\mathbb{E}[(W_1)^+] + O(\frac{\varepsilon^2}{\sigma^3}).$$

3. *Case* $h(x) = 1_{x \le x_0}$. We have:

$$\mathbb{E}[h(X_1^{\varepsilon})] = \mathcal{N}\left(\frac{x_0}{\sqrt{\sigma^2 + \varepsilon}}\right) = \mathcal{N}\left(\frac{x_0}{\sigma}\right) - \mathcal{N}'\left(\frac{x_0}{\sigma}\right)\frac{x_0}{\sigma}\frac{\varepsilon}{2\sigma^2} + O(\frac{\varepsilon^2}{\sigma^4}).$$

These simple computations show that the expansion order depends on the relative magnitude of ε and σ, and also on the regularity of the function h. For instance, if σ is also small, say $\varepsilon = \sigma^3 \to 0$, then the expansion order w.r.t. σ in the case (1), (2), (3) is respectively equal to 3, 2 and 1. These subtleties do not appear in the expansions (1.21) of Watanabe type or (1.23) of PDE type, because the focus is made only on a single small parameter ε.

It means that in some situations, asymptotic expansions may be misleading or may not give the best possible approximations; then, we should take into account the influence of all (or many) model parameters. In the context of fast volatility, multi-scale modeling and its related asymptotic analysis are very recently developed in [Fouque 2011, Chapter 4]; see also [Kevorkian 1985, Chapter 3 and Section 4.4].

In the sequel of this thesis, we consider non asymptotic expansions, mainly for local volatility models (except in Part II of the thesis where the case of model combining local and stochastic volatility is investiguated), and analyse the approximation error taking into account several parameters at the same time, in order to determine in which extent they play complementary or opposite roles. For instance, it is informative to see the simultaneous influences of maturity, of volatility amplitude or of derivatives of volatility function on the option prices. Their impacts depend on the payoff smoothness: the accuracy is expected to be improved for smooth payoff compared to non-smooth payoffs.

Final considerations. After this (hopefully complete) overview, the reader may wonder what is the best approximation method among those presented. Of course, it depends on the required accuracy and the computational time allowed for the numerical evaluation. From this point of view, all methods are not equivalent. The choice of relevant asymptotics/approximations guarantees to catch the main features of the pricing problem, and as a consequence, it will likely lead to an expansion of low order to achieve a good accuracy (with low computational time or complexity). In these respects, the proxy expansion has immediate advantages: the better or the more intuitive the proxy, the smaller the number of correction terms.

One should also take care of the preservation of some model properties in the approximation.

- One of them is the martingale property of $S = e^X$ (serving as a base for Call/Put parity relation). For instance, a small noise approximation of X defined in (1.5) does not maintain the martingale property since the volatility coefficient is scaled by ε while the drift remains unchanged: as a result, the final approximation may suffer from numerical arbitrage. On the contrary the transition density approximations provided in [Pagliarani 2011] integrate to one, thus avoiding the introduction of arbitrage opportunities.

- Another property is positivity of S. Taking a Normal Proxy for S may give wrong results if the values of S close to 0 have a prominent role in the computation of $\mathbb{E}(h(S_T))$ (for instance, Call/Put with small strikes). The inadequacy of the Normal Proxy is also widely examined in [Pagliarani 2011].

These kinds of consideration may help to choose between different methods, with the additional help of comparative numerical tests.

1.4 Structure of thesis and main results

The thesis contains four Parts which deal with different problematics and models:

- The Part I provides new expansion formulas for the local volatility models for vanilla products and introduces new results concerning the pricing of forward start options.

- The Part II deals with the analytical approximation of vanilla prices in models combining both local and stochastic volatility (Heston type).

- The Part III covers the pricing of regular barrier options.

- The Part IV gives a new methodology for the efficient weak analytical approximation of multidimensional diffusions.

▷ **Part I.** This Part is divided into three chapters. Chapter 2 deals with local volatility models (see

Equations (1.4)-(1.5)) and the aim is to revisit the Proxy principle and to provide **new analytical approximations** for Call options and the delta of the options which is defined as the first sensitivity w.r.t. the spot S_0:

$$\delta = \partial_{S_0}\mathrm{Call}(S_0, T, K) = \partial_{S_0}\mathbb{E}(S_T - K)_+.$$

As seen in Subsection 1.3 of this Chapter, the literature on this subject is very profuse. Using the Proxy principle presented in Subsection 1.3.4.4, which seems to us to be an intuitive methodology and providing very accurate results, we develop new approximation formulas. We show that for the pricing of Call options in local volatility models, a good proxy is to consider Gaussian models with the local volatility frozen at mid-point between the strike and the spot:

$$\begin{aligned}
\mathrm{d}X_t^P &= -\frac{1}{2}a^2(t, \log(\sqrt{S_0K}))\mathrm{d}t + a(t, \log(\sqrt{S_0K}))\mathrm{d}W_t, \qquad X_0^P = x_0 \qquad \text{(Log-Normal Proxy)},\\
\mathrm{d}S_t^P &= \Sigma(t, \frac{S_0+K}{2})\mathrm{d}W_t, \qquad S_0^P = S_0. \qquad \text{(Normal Proxy)},
\end{aligned}$$

and that efficient approximations of the Call option are given by:

$$\mathbb{E}(S_T - K)_+ = \begin{cases} \mathbb{E}(e^{X_T^P} - K)_+ + \sum_{i=1}^{6} \eta_i(a; \log(\sqrt{S_0K}))_0^T \partial_{x^i}^i \mathbb{E}[(e^{X_T^P + x} - K)_+]|_{x=0} + \mathrm{Error}_{\mathrm{LN}},\\ \mathbb{E}(S_T^P - K)_+ + \sum_{i=1}^{6} \zeta_i(\Sigma; \frac{S_0+K}{2})_0^T \partial_{s^i}^i \mathbb{E}[(S_T^P + s - K)_+]|_{s=0} + \mathrm{Error}_{\mathrm{N}}, \end{cases}$$

where:

- the weights η (respectively ζ) in front of the Greeks w.r.t. the log-spot (respectively the spot) are multiple time-integrals of $a(., \log(\sqrt{S_0K}))$ and its spatial derivatives (respectively $\Sigma(., \frac{S_0+K}{2})$).
- $\mathrm{Error}_{\mathrm{LN}}$ (respectively $\mathrm{Error}_{\mathrm{N}}$) is an error term estimated w.r.t. the magnitude of a (respectively Σ) and its derivatives as well the maturity T. Basically if we denote by $\mathcal{M}(a)$ a control on a and its derivatives, $\mathrm{Error}_{\mathrm{LN}}$ is of order $(\mathcal{M}(a)\sqrt{T})^m$ with $m = 3$ (order 2 expansion) or 4 (order 3 expansion).

These new results are resumed in Theorems 2.1.3.2 and 2.3.2.1. To obtain more tractable (and hopefully more accurate) formulas, we also provide expansions of the implied volatility (of type Black-Scholes or Bachelier, see (1.9)-(1.11)) to say that the price approximation is simply a Black-Scholes (or Bachelier) Call price with a suitable volatility parameter. For example with the Log-Normal Proxy, our implied volatility expansions read as polynomials of order one or two w.r.t. the log-moneyness $\log(\frac{S_0}{K})$ with an error of order 3 or 4 (see Theorems 2.1.4.1-2.3.3.1). In addition we give approximations of the option delta (see Theorems 2.4.0.2-2.4.0.3). We prove throughout many numerical experiments the extreme accuracy of our formulas, greatly improving the previous expansions (developed in [Benhamou 2010a]) with the local volatility at spot and outperforming standard approximations given by Hagan [Hagan 1999] and Henry-Labordère [Henry-Labordère 2008]. That confirms the duality played by the variables spot and strike and that there is a theoretical (symmetric formulas with sometimes some reductions) as well a practical (extreme accuracy) interest to use the local volatility at mid-point.
We also propose an alternative technique of proof to derive the formulas based on arguments mixing Itô calculus and PDE tools instead of a pure stochastic analysis point of view. We compare the different approaches and give early stages for a third method involving only PDE arguments in Section 2.2.

In Chapter 3, we focus on the pricing of a forward start option, which can be view as a forward on an option. More precisely, if $t_i > 0$ denotes the forward date and $t_i + T$ with $T > 0$ the expiration date, we are interesting by the **accurate numerical evaluation** of:

$$\begin{cases} \mathrm{Call}^{\mathrm{FS},A}(S_0, t_i, T, K) = & \mathbb{E}[(\frac{S_{t_i+T}}{S_{t_i}} - K)_+] = \mathbb{E}[(e^{X_{t_i+T}-X_{t_i}} - e^k)_+], \\ \mathrm{Call}^{\mathrm{FS},B}(S_0, t_i, T, K) = & \mathbb{E}[(S_{t_i+T} - S_{t_i}K)_+] = \mathbb{E}[(e^{X_{t_i+T}} - e^{k+X_{t_i}})_+], \end{cases}$$

where S is the solution assumed to be positive of the stochastic diffusion equation (1.4) and $X = \log(S)$ (see (1.5)). For (x_0, t_i, T, k) given, the forward implied Black-Scholes volatilities of type A and B are the unique non-negative parameters $\sigma_{\mathrm{I,F,A}}(x_0, t_i, T, k)$ and $\sigma_{\mathrm{I,F,B}}(x_0, t_i, T, k)$ such that:

$$\begin{cases} \mathrm{Call}^{\mathrm{FS},A}(e^{x_0}, t_i, T, e^k) = & \mathrm{Call}^{\mathrm{BS}}(0, \sigma^2_{\mathrm{I,F,A}}(x_0, t_i, T, k)T, k), \\ \mathrm{Call}^{\mathrm{FS},B}(e^{x_0}, t_i, T, e^k) = & \mathrm{Call}^{\mathrm{BS}}(x_0, \sigma^2_{\mathrm{I,F,B}}(x_0, t_i, T, k)T, x_0 + k). \end{cases}$$

Using a conditioning argument and the results of vanilla implied volatility expansions developed in Chapter 2, we express the price of the forward start option of type A as an expectation of the Black-Scholes price function with a volatility depending on the local volatility function frozen at some stochastic point involving X_{t_i}, plus an error. Then we provide forward implied volatility of type A approximations by performing a volatility expansion to freeze the local volatility function at some deterministic point. The results are obtained under a uniform (w.r.t. time and space variables) ellipticity condition: $\inf_{(t,x)\in[0,T]\times\mathbb{R}} a(t,x) > 0$. Forward implied volatility of type B approximations are obtained employing a change of probability measure argument. Numerical results confirm the very good accuracy of the developed approximation formulas.

In Chapter 4 which concludes this Part, we discuss and give insight about two essential facts in the Proxy Principle: the choice of the parameterization to link the initial and the Proxy processes, and the choice of the Proxy process. All parameterizations and approaches are not equivalent and even if they can lead to the same result, the difficulty in the intermediate calculations and the required assumptions can be quite different. We illustrate this fact with some examples and we notably show how to consider a Log-Normal proxy with a direct suitable parameterization of the initial process S itself rather than its logarithm X (see Theorem 4.2.0.2).
The choice of the Proxy process is also crucial and relies both on the initial process and on the practitioner intuition. So far, only zero order Proxy processes have been studied (Normal or Log-normal), i.e. capturing the behaviour of the local volatility only throughout its value at a specific spatial point. Order one Proxy processes can be considered by incorporating the knowledge of the skew (i.e. the first derivative of the local volatility frozen at some suitable point) in their dynamic. The most natural surrogate is probably the following Displaced Log-Normal process in the time-homogeneous framework:

$$\mathrm{d}S^P_t = [\Sigma(S_0) + \Sigma^{(1)}(S_0)(S^P_t - S_0)]\mathrm{d}W_t, \quad S^P_0 = S_0. \tag{1.28}$$

We give means to link S and S^P using an appropriate transformation of the model and we provide in Theorem 4.3.0.3 2nd and 3rd Call options approximations using the Displaced Log-Normal Proxy S^P. Interestingly, the obtained expansions contain less terms than those using zero order Proxy processes and consequently seem more tractable. We also show with some numerical tests how this approach looks promising.

▷ **Part II.** This Part contains two chapters. In Chapter 5 we focus on models combining both local and stochastic volatility:

$$\mathrm{d}X_t = \sigma(t, X_t)\sqrt{V_t}\mathrm{d}W_t - \frac{1}{2}\sigma^2(t, X_t)V_t\mathrm{d}t, \; X_0 = x_0, \tag{1.29}$$

$$\begin{aligned} dV_t &= \alpha_t dt + \xi_t \sqrt{V_t} dB_t,\ V_0 = v_0, \\ d\langle W, B \rangle_t &= \rho_t dt. \end{aligned} \tag{1.30}$$

where:

- X_t is the logarithm of the asset price,
- V_t is the stochastic variance modelled with a square-root process,
- v_0 is the initial square of volatility,
- α is the drift parameter of the stochastic variance,
- ξ is the volatility of volatility,
- ρ is the correlation.

Notice that the general form of the local volatility allows (owing to a suitable space-change) to cover the case of a CIR process for the stochastic variance $dV_t = \kappa_t(\theta_t - V_t)dt + \xi_t \sqrt{V_t} dB_t$ where κ is the mean reversion parameter and θ the long-term variance level. In this Chapter we aim at:

- providing an **accurate analytical approximation** of non only Call-Put options but also all vanilla options (depending on X_T) writing as $\mathbb{E}[h(X_T)]$ with h a Lipschitz bounded payoff function,
- valid for both short and long maturities and covering general local volatility functions, non-null correlation as well time-dependent parameters,
- achieving a computational time close to zero on the contrary to very time consuming Monte Carlo simulations,
- providing a complete mathematical justification.

To achieve this we use as Proxy the time-dependent Black-Scholes model obtained by vanishing the volatility of volatility and by freezing spatially the local volatility at x_0:

$$\begin{aligned} dX_t^P &= \sigma(t, x_0) \sqrt{v_t} dW_t - \frac{1}{2}\sigma^2(t, x_0) v_t dt, \\ v_t &= v_0 + \int_0^t \alpha_s ds. \end{aligned}$$

To link the initial process (1.29)-(1.30) with the above proxy process, we introduce a two-dimensional parameterized process given by:

$$\begin{aligned} dX_t^\eta &= \sigma(t, \eta X_t^\eta + (1-\eta)x_0) \sqrt{V_t^\eta} dW_t - \frac{1}{2}\sigma^2(t, \eta X_t^\eta + (1-\eta)x_0) V_t^\eta dt,\ \ X_0^\eta = x_0, \\ dV_t^\eta &= \alpha_t dt + \eta \xi_t \sqrt{V_t^\eta} dB_t,\ v_0, \end{aligned}$$

where η is an interpolation parameter lying in the range $[0, 1]$, such that for $\eta = 1$, $X_t^1 = X_t$ and $V_t^1 = V_t$, and for $\eta = 0$, $X_t^0 = X_t^P$ and $V_t^0 = v_t$. The derivation of an expansion and its justification is far from straightforward due to the irregularity of the square root coefficient in the stochastic variance (1.30) and due to the function h supposed to be only Lipschitz. In spite of the inspiration by [Benhamou 2010a], the mathematical analysis must be quite different since the model (1.29)-(1.30) is not smooth in the Malliavin sense (because of the square-root process, see [Alòs 2008, De Marco 2011]). We briefly explain how we overcome this major problem: we use Malliavin calculus for a suitably perturbed payoff h_δ and a suitably perturbed random variable $\bar{X}_T$. We consider the Gaussian regularization $h_\delta(x) := \mathbb{E}[h(x+\delta\overline{W}_T)]$ for an independent Brownian motion $\overline{W}$ and a small parameter $\delta > 0$ suitably fixed. Then write the decomposition:

$$\mathbb{E}[h(X_T)] \approx \mathbb{E}[h_\delta(\bar{X}_T)] \approx \mathbb{E}[h_\delta(X_T^P)] + \mathbb{E}[h'_\delta(X_T^P)(\bar{X}_T - X_T^P)] + \dots$$

i) Since h is Lipschitz, the first approximation is easily justified if δ is small enough and if X_T and $\bar{X}_T$ are close enough in L^p-sense.

ii) The last expectation is computed (eventually up to an error) using the Malliavin calculus and can be expressed as a linear combination of sensitivities in the Proxy model:

$$\mathbb{E}[h'_\delta(X_T^P)(\bar{X}_T - X_T^P)] = \sum_i \eta_{i,T}\partial^i_{x^i}\mathbb{E}[h_\delta(X_T^P + x)]|_{x=0} + \text{Error} = \sum_i \eta_{i,T}\partial^i_{x^i}\mathbb{E}[h(X_T^P + x)]|_{x=0} + \text{Error}.$$

iii) The hard part of the analysis is related to the global error control, which enlightens the right choice of δ and $\bar{X}_T$. To account for non-smooth payoffs we use an integration by parts formula, which relies on the non-degeneracy of the interpolated random variable $X^\lambda := \lambda\bar{X}_T + (1-\lambda)X_T^P$ (for any fixed $\lambda \in]0,1[$). This excludes to take $\bar{X}_T = X_T$ that is not sufficiently Malliavin differentiable. Alternatively, we select a $\bar{X}_T \in \mathbb{D}^\infty$ which on the one hand, is close enough to X_T in L^p, and which on the other hand, is such that X^λ is uniformly non-degenerate or at least non-degenerate with high probability under the sole assumption $\int_0^T \sigma^2(t,x_0)v_t\mathrm{d}t > 0$ (which reads as a pointwise ellipticity assumption). By construction, $\bar{X}_T$ is potentially not uniformly non-degenerate what does not allow the use of an integration by parts formula. To handle this last problem we use a splitting noise property by writing $h_\delta(x) = \mathbb{E}[h_{\delta/\sqrt{2}}(x+\delta\overline{W}_{\frac{T}{2}})]$. Thus $X^\lambda + \delta\overline{W}_{\frac{T}{2}}$ is uniformly non-degenerate and we can obtain nice estimates provided that δ is not too small.

A precise tuning of δ and a construction of $\bar{X}_T$ are possible without loss of accuracy in the regularization step and the final approximation formula (see Theorem 5.2.2.1) takes the form of an explicit Gaussian representation under a local non-degeneracy condition:

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] + \sum_{i=1}^{6} \eta_{i,T}\partial^i_{x^i}\mathbb{E}[h(X_T^P + x)]|_{x=0} + \text{Error}_{3,h},$$

where:

- $\eta_{i,T}$ are integral operators depending on $(\partial^j_{x^j}\sigma(.,x_0))_{j\in\{0,1,2\}}, \rho_., \xi_., v_.$ and T,
- the error is estimated as $\text{Error}_{3,h} = O(L_h\mathcal{M}(\sigma)[\mathcal{M}(\sigma)+\xi_{\sup}]^3T^2)$ where L_h is the Lipschitz constant associated to h and $\mathcal{M}(\sigma)$ denotes a control on σ and its derivatives. This reads as a multi-parameter estimate.

From the price expansion formula, we also derive for the particular case of Call options (for which the results are valid owing to the Call/Put parity) an expansion of the implied volatility with the local volatility frozen at mid-point (see Theorems 5.4.2.1):

$$\mathbb{E}[(e^{X_T} - e^k)_+] = \text{Call}^{\text{BS}}(x_0, \tilde{\sigma}^2_{3,I}T, k) + \text{Error}^{\text{I}}_3,$$

where:

- $\tilde{\sigma}_{3,I} = \pi_{0,T} + \pi_{1,T}\ (x_0 - k) + \pi_{2,T}\ (x_0-k)^2$,
- $\pi_{i,T}$ are integral operators depending on $(\partial^j_{x^j}\sigma(., \frac{x_0+k}{2}))_{j\in\{0,1,2\}}, \rho_., \xi_., v_.$ and T,

- $\mathrm{Error}^{\mathrm{I}}_3 = O(e^k \mathcal{M}(\sigma)[\mathcal{M}(\sigma)+\xi_{\sup}]^3 T^2)$.

Numerical experiments show that our formulas, in addition to be computable in quasi real-time, present an accuracy turning out to be excellent.

In Chapter 6, we provide numerical results concerning the smile and the skew behaviors in the CEV-Heston model. We notably study the impact of the volatility of volatility, the correlation and the skew parameter of the CEV part on the implied volatility represented as a function of the log-moneyness and the maturity.

▷ **Part III.** In Chapter 7, we look at **the efficient price approximation of regular down Barrier options**, price which can be written as:

$$\mathbb{E}[h(X_T)\mathbb{1}_{\inf_{t\in[0,T]} X_t > b}] = \mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}],$$

where:

- X is the log-asset process solution to $X_t = x_0 + \int_0^t \sigma(s,X_s)(\mathrm{d}W_s - \frac{\sigma(s,X_s)}{2}\mathrm{d}s)$,
- $b < x_0$ is the level of the barrier,
- h is a locally Lipschitz payoff function such that $h(x) = 0$, $\forall x \le b$,
- $\tau_b = \inf\{t > 0 :\ X_t = b\}$ is the first hitting time of the level b for the process X.

As usual we use a Gaussian proxy process $(X^P_t)_{t\in[0,T]}$ obtained by freezing the space variable in the function σ:

$$X^P_t = x_0 + \int_0^t \sigma(s,x_0)(\mathrm{d}W_s - \frac{\sigma(s,x_0)}{2}\mathrm{d}s),$$

and we introduce $\tau^P_b = \inf\{t > 0 :\ X^P_t = b\}$ the first hitting time of the level b for X^P. At first glance we replace the unknown joint law of $(X_T, \inf_{t\in[0,T]} X_t)$ by that of $(X^P_T, \inf_{t\in[0,T]} X^P_t)$ fully specified and to derive corrective terms, we represent the error $\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}] - \mathbb{E}[h(X^P_T)\mathbb{1}_{\tau^P_b>T}]$ using the PDE associated to the Proxy:

$$v^{P,h}_{\underline{\varrho},T}(t,x) = \mathbb{E}[h(X^P_T)\mathbb{1}_{\inf_{s\in[t,T]} X^P_s > b}|X^P_t = x].$$

Under an intermediate ellipticity condition (uniform in the time variable for $x = x_0$): $\inf_{t\in[0,T]}\sigma(t,x_0) > 0$, $v^{P,h}_{\underline{\varrho},T} \in C^\infty([0,T[\times[b,+\infty[)$ and is the explicit solution of the Cauchy-Dirichlet problem:

$$\begin{cases} \partial_t v^{P,h}_{\underline{\varrho},T}(t,x) + \dfrac{1}{2}\sigma^2(t,x_0)(\partial^2_{x^2} - \partial_x)v^{P,h}_{\underline{\varrho},T}(t,x) = 0, \quad (t,x) \in [0,T[\times[b,+\infty[, \\ v^{P,h}_{\underline{\varrho},T}(t,b) = 0, \quad t \in [0,T], \\ v^{P,h}_{\underline{\varrho},T}(T,x) = h(x), \quad x \in]b,+\infty[. \end{cases}$$

Then using the regularity of the payoff, the ellipticity assumption, the Itô's formula for $v^{P,h}_{\underline{\varrho},T}(T\wedge\tau_b, X_{T\wedge\tau_b})$ and simplifications coming from the above PDE, we establish a Robustness-type formula:

$$\mathbb{E}[\underbrace{h(X_T)\mathbb{1}_{\tau_b>T}}_{=v^{P,h}_{\underline{\varrho},T}(T\wedge\tau_b, X_{T\wedge\tau_b})}] = \underbrace{\mathbb{E}[h(X^P_T)\mathbb{1}_{\tau^P_b>T}]}_{=v^{P,h}_{\underline{\varrho},T}(0,x_0)}$$

$$+\frac{1}{2}\mathbb{E}[\int_0^{T\wedge\tau_b}[\sigma^2(t,X_t)-\sigma^2(t,x_0)](\partial^2_{x^2}-\partial_x)v^{P,h}_{\underline{\varrho},T}(t,X_t)\mathrm{d}t].$$

Then combining expansions of σ, nested Robustness fomulas, Itô calculus, PDE arguments, martingale properties ($(\partial^n_{x^n}v^{P,h}_{\underline{\varrho},T}(t\wedge\tau^P_b,X^P_{t\wedge\tau^P_b}))_{t\in[0,T[}$ is a martingale for any integer n) and convolution of densities, we finally derive analytical price approximations writing as:

$$\begin{aligned}\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}] =&\mathbb{E}[h(X^P_T)\mathbb{1}_{\tau^P_b>T}]+\sum_i\eta_i(\sigma;x_0)^T_0\partial^i_{x^i}v^{P,h}_{\underline{\varrho},T}(0,x)|_{x=x_0}\\&-\sum_i\mathbb{E}[\mathbb{1}_{\tau^P_b\leq T}\eta_i(\sigma;x_0)^T_{\tau^P_b}\partial^i_{x^i}v^{P,h}_{\underline{\varrho},T}(\tau^P_b,x)|_{x=b}]+\mathrm{Error}_{h,b},\end{aligned}$$

where:

- $\eta_i(\sigma;x_0)^t_s$ are integral operators between the dates $s\leq t$ belonging to $[0,T]$ and depending on $(\partial^j_{x^j}\sigma(.,x_0))_{j\in\{0,1,2\}}$. Notice that if $x_0=b$, all the corrective terms vanish what is coherent with the fact that $\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}]=0$.
- $\mathrm{Error}_{h,b}=O([\mathcal{M}(\sigma)\sqrt{T}]^j)$ with $j=3$ (order 2 expansion, see Theorem 7.2.2.1) or $j=4$ (order 3 expansion, see Theorem 7.2.3.1).

Besides, we provide some corollaries related to the particular case of Barrier Call options pricing. First we derive expansion formulas with the local volatility at mid-point between the strike and the spot (see Theorem 7.5.2.1). Second if all the expansion coefficients remain explicit up to a numerical integration of the terms written as an expectation involving τ^P_b, we show that if σ is time-independent, we get very simple formulas involving only Gaussian cumulative and density functions owing to nice convolution properties of densities (see Proposition 7.5.3.1). Last but not least we give numerical examples of the excellent accuracy of our approximation formulas.

▷ **Part IV.** In this Part which contains two Chapters, we are interesting by the weak approximations and the efficient calculations in multidimensional diffusions in a quite general framework. Our goal is to provide a generic methodology applicable to a large class of terminal functions in order to broaden the scope of application beyond financial mathematics. We introduce in Chapter 8 a new method called **Stochastic Approximation Finite Elements (SAFE for short)** which combines the use of a Gaussian proxy to approximate the law of a multidimensional diffusion and a local interpolation of the terminal function applied to the diffusion using Finite Elements. More formally, considering for $d\geq 1$ a d-dimensional stochastic differential equation (SDE) defined by:

$$X_t=x_0+\sum_{j=1}^q\int_0^t\sigma_j(s,X_s)\mathrm{d}W^j_s+\int_0^t b(s,X_s)\mathrm{d}s,$$

where $(W_t)_{t\geq 0}$ is a standard Brownian motion in $\mathbb{R}^q$, σ is a $d\times q$ matrix and b is a d-dimensional vector, we aim at providing an analytical approximation of

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

for a given function h, at least Lipschitz continuous. As usual we begin with an approximation of the law of X by considering the Gaussian proxy process obtained by freezing at $x=x_0$ the diffusion coefficients:

$$X^P_t=x_0+\sum_{j=1}^q\int_0^t\sigma_j(s,x_0)\mathrm{d}W^j_s+\int_0^t b(s,x_0)\mathrm{d}s.$$

Adapting the stochastic tools developed in the previous Parts of the thesis to the multidimensional case, we derive a weak approximation in the form (see Theorem 8.2.1.1):

$$\mathbb{E}[h(X_T)] \approx \mathbb{E}[h(X_T^P)] + \sum_{|\alpha|\leq 3} w_{\alpha,T} \partial^{|\alpha|}_{\epsilon_{\alpha_1}\dots\epsilon_{\alpha_{|\alpha|}}} (\mathbb{E}[h(X_T^P+\epsilon)])\big|_{\epsilon=0},$$

where $\alpha \in \{1,\dots,d\}^{|\alpha|}$ is a multi-index and $w_{\alpha,T}$ are weights depending explicitly on the SDE coefficients. Most of the time, although the law of X_T^P is known, the explicit calculus of the main term $\mathbb{E}[h(X_T^P)]$ and of the sensitivities $\partial^{|\alpha|}_{\epsilon_{\alpha_1}\dots\epsilon_{\alpha_{|\alpha|}}} (\mathbb{E}[h(X_T^P+\epsilon)])\big|_{\epsilon=0}$ is not possible due to the general form of h. However a simple expectation form can be derived for straightforward Monte Carlo simulations (see Theorem 8.2.1.2) always faster than Monte Carlo simulations on the initial diffusion owing to the lack of discretization for the proxy simulation. The idea is to use the Malliavin calculus to express the sensitivities as an expectation of $h(X_T^P)$ weighted with an explicit random variable. One has for an explicit Malliavin weight $\mathcal{W}$:

$$\mathbb{E}[h(X_T^P)] + \sum_{|\alpha|\leq 3} w_{\alpha,T} \partial^{|\alpha|}_{\epsilon_{\alpha_1}\dots\epsilon_{\alpha_{|\alpha|}}} (\mathbb{E}[h(X_T^P+\epsilon)])\big|_{\epsilon=0} = \mathbb{E}[h(X_T^P)(1+\mathcal{W})].$$

To obtain fully analytical formulas, an other ingredient is needed: we suitably approximate the function h. Using interpolations based on Finite Elements, we get the final structure of approximation:

$$\mathbb{E}[h(X_T)] \approx \mathbb{E}[\hat{h}(X_T^P)] + \sum_{|\alpha|\leq 3} w_{\alpha,T} \partial^{|\alpha|}_{\epsilon_{\alpha_1}\dots\epsilon_{\alpha_{|\alpha|}}} (\mathbb{E}[\hat{h}(X_T^P+\epsilon)])\big|_{\epsilon=0},$$

where $\hat{h}$ denotes the resulting interpolation of h (based on multilinear or multi-quadratic Finite Elements of type Lagrange, see Theorems 8.2.2.1 and 8.2.4.1). The interpolation procedure is done in such a way ($\hat{h}$ is build depending on the law of X_T^P) that the calculus of the above expectations becomes fully explicit with very simple formulas involving the one-dimensional Gaussian cumulative function and its derivatives. We also provide an accuracy analysis (see Theorems 8.2.2.2-8.2.3.1) according to the h-regularity and show how tune the interpolation parameters (size of the grid, grid mesh) to obtain a global error of order at most equal to:

$$\mathcal{E} = [\mathcal{M}(\sigma,b)\sqrt{T}]^3,$$

where $\mathcal{M}(\sigma,b)$ denotes a control on the diffusion coefficients. We finally give estimates of the complexity of our methodology (see Corollaries 8.2.3.1-8.2.4.1) and show an improved efficiency in comparison to Monte Carlo simulations in small and medium dimensions (up to **10**). The theoretical performance of our algorithm is confirmed throughout various numerical experiments. For high dimensions (from 20 to 100) we also show the efficiency of the Monte Carlo simulations based on the proxy (exact and without discretization) which present a speed gain by a factor **100** (whatever is the dimension) in comparison to Monte Carlo simulations of the initial diffusion.

In Chapter 9, we provide additional numerical results concerning the pricing of multi-asset products (Basket, Geometrical mean, Worst of and Best of Put options) in multidimensional CEV models using the SAFE method.

Publications and submissions. Every Part of this thesis has been or should be published. More precisely:

- Sections 1.1, 1.2 and 1.3 of this introduction Chapter and Chapter 2 have been published in the Chapter "Asymptotic and non asymptotic approximations for option valuation" of the book "Recent Developments in Computational Finance Foundations, Algorithms and Applications", Thomas Gerstner and Peter Kloeden (Ed.) 2012, World Scientific Publishing Company. This is a joint work with my supervisor Emmanuel Gobet.

- Chapter 3 "Forward implied volatility expansions in local volatility models" is in preparation to be submitted soon. This is a common work with Julien Hok.

- Chapter 5 "Price expansion formulas for model combining local and stochastic volatility" will be the subject of a forthcoming publication.

- Chapter 7 "Price expansions for regular down barrier options" is too in preparation for an upcoming submission.

- Chapter 8 "Stochastic Approximation Finite Element method: analytical formulas for multidimensional diffusion process" have been submitted in *SIAM Journal of Numerical Analysis*.

Part I

New expansion formulas in local volatility models

Chapter 2

Revisiting the Proxy principle in local volatility models

Published in the Chapter "Asymptotic and non asymptotic approximations for option valuation" of the book "Recent Developments in Computational Finance Foundations, Algorithms and Applications", Thomas Gerstner and Peter Kloeden (Ed.) 2012, World Scientific Publishing Company.

Contents

We develop in this Chapter the principle of high order approximations related to an intuitive proxy and in the case of local volatility models with general time-dependency, we derive new formulas in terms of both prices and implied volatilities using the local volatility function at the mid-point between strike

and spot: in general, our approximations outperform previous ones by Hagan and Henry-Labordère. We also provide approximations of the option delta.
Here is the outline of the Chapter. In Section 2.1, we consider the simplest case of second order approximation in local volatility models, using log-normal or normal proxys. We give pedagogic proofs. Section 2.2 is devoted to a more detailed analysis: we first give arguments based on stochastic analysis (martingales, Malliavin calculus). We compare this derivation with a method mixing stochastic analysis and PDE, and with a pure PDE approach: we show in which respect our methodology is different. In Section 2.3, we provide various high-order approximation using proxys. In Section 2.4, approximations of the option delta are provided. Section 2.5 is gathering numerical experiments illustrating the performance of our formulas compared to those of Hagan et al. [Hagan 1999] and of Henry-Labordère [Henry-Labordère 2008]. Some intermediate and complementary results are postponed to Section 2.6.
In all this Chapter, we keep the notations introduced in the previous Chapter 1 Section 1.2.

2.1 Approximation based on proxy

2.1.1 Notations and definitions

The following notations and definitions are repeatedly used in this Chapter and the next Chapter 4.

▷ **Differentiation.** If these derivatives have a meaning, we write $l_t^{(i)}(x) = \partial^i_{x^i} l(t,x)$ for any function l of two variables.

▷ **Integral Operator.** The integral operator ω^T is defined as follows: for any measurable and bounded 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 measurable and bounded 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 $t \in [0,T]$.

▷ **Time reversal.** For any measurable and bounded function l, we denote by $\widetilde{l}$ the function $\widetilde{l_t} = l_{T-t}$ for any $t \in [0,T]$. Notice the relation

$$\omega(\widetilde{l_1},\widetilde{l_2},..,\widetilde{l_n})_0^T = \omega(l_n,l_{n-1},..,l_1)_0^T \tag{2.1}$$

available for any measurable and bounded functions $(l_1,\cdots,l_n)$: in other words, reversing the time of integrands is equivalent to change the order of integration.

▷ **Quadratic mean on** $[0,T]$**.** For any measurable function $(l(t,x))_{(t,x)\in[0,T]\times\mathbb{R}}$ of two variables, bounded w.r.t. the time variable for any $x \in \mathbb{R}$, we denote by $\bar{l}_z$ its quadratic mean on $[0,T]$ at the spatial point z defined by:

$$\bar{l}_z = \sqrt{\frac{1}{T}\int_0^T l_t^2(z)\mathrm{d}t}.$$

This notation is frequently used for the function a at the points $z = x_0, k, x_{avg}$ and for the function Σ at $z = S_0, K, S_{avg}$.

▷ **Assumptions on** a **and** Σ**.**

- $(\mathcal{H}^a)$: a is a bounded measurable function of $(t,x)\in[0,T]\times\mathbb{R}$, and five times continuously differentiable in x with bounded[1] derivatives. Set
$$\mathcal{M}_1(a)=\max_{1\le i\le 5}\sup_{(t,x)\in[0,T]\times\mathbb{R}}|\partial^i_{x^i}a(t,x)| \text{ and } \mathcal{M}_0(a)=\max_{0\le i\le 5}\sup_{(t,x)\in[0,T]\times\mathbb{R}}|\partial^i_{x^i}a(t,x)|.$$
In addition, there exists a constant $c_a>0$ such that $|a(t,x)|\ge c_a$ for any $(t,x)\in[0,T]\times\mathbb{R}$.

- $(\mathcal{H}^a_z)$: assume $(\mathcal{H}^a)$ by replacing the last uniform ellipticity by the single condition $\int_0^T|a(t,z)|^2\mathrm{d}t>0$.

The above hypothesis will be considered at $z=x_0, z=k$ or $z=x_{avg}$.
We define similarly $(\mathcal{H}^\Sigma)$ or $(\mathcal{H}^\Sigma_z)$ by replacing a by Σ in $(\mathcal{H}^a)$ and $(\mathcal{H}^a_z)$. Then the hypothesis will be considered at $z=S_0, z=K$ or $z=S_{avg}$.

▷ **Constants.** Our next error estimates are stated following the notation below.

- "$A=O(B)$" means that $|A|\le CB$: here, C stands for a generic constant that is a non-negative increasing function of T, $\mathcal{M}_1(a)$, $\mathcal{M}_0(a)$ and of the oscillation ratio $\frac{\mathcal{M}_0(a)}{c_a}$ (if $(\mathcal{H}^a)$ is fulfilled) or $\frac{[\mathcal{M}_0(a)]^2T}{\int_0^T|a(t,z)|^2\mathrm{d}t}$ (if $(\mathcal{H}^a_z)$ is fulfilled).

  If $(\mathcal{H}^\Sigma)$ or $(\mathcal{H}^\Sigma_z)$ is satisfied, in the above dependence a has to replaced by Σ.

  Usually, a generic constant may depend on S_0, x_0, K and k; nevertheless, it remains uniformly bounded in these variables: it is possible to derive exact dependency but we skip it to keep the analysis short.

- Similarly, if A is positive, $A\le_c B$ means that $A\le CB$ for a generic constant C.

2.1.2 Proxy approximation: a primer using the local volatility at spot

▷ **Log-normal proxy.** Assume by expertise that the model (1.4) introduced in the Chapter 1 Section 1.2 behaves closely to a log-normal model, in the sense that a log-normal approximation seems to be reasonable. For instance, in the case of CEV type model (see [Cox 1975] and [Emanuel 1982])

$$S\sigma(t,S)=\nu_t S^{\beta_t}, \tag{2.2}$$

a log-normal heuristics is associated to β close to 1. Some numerical illustrations are given later.

As a first log-normal approximation, we freeze the volatility in space to the initial spot value: regarding the log-asset X defined in (1.5) Chapter 1 Section 1.2, it writes

$$\mathrm{d}X^P_t=-\frac{1}{2}a^2(t,x_0)\mathrm{d}t+a(t,x_0)\mathrm{d}W_t,\qquad X^P_0=x_0.$$

We refer to this proxy model as *log-normal proxy with volatility at spot*. The evaluation of the next correction terms requires a suitable representation of the distance between the model and the proxy: for this, we use the interpolated process (1.25) given by

$$\mathrm{d}X^\eta_t=\eta(-\frac{1}{2}a^2(t,X^\eta_t)\mathrm{d}t+a(t,X^\eta_t)\mathrm{d}W_t),\quad X^\eta_0=x_0.$$

for an interpolation parameter $\eta\in[0,1]$. Under $(\mathcal{H}^a_{x_0})$, the three first derivatives of $\eta\mapsto X^\eta_t$ are well defined (a.s. simultaneously for any t, see [Kunita 1984, Theorem 2.3]). Denote by $X^\eta_{i,t}$ and $X_{i,t}$ the i-th derivative respectively at η and $\eta=0$. Direct computations yield

$$\mathrm{d}X^\eta_{1,t}=-\frac{1}{2}a^2(t,X^\eta_t)dt+a(t,X^\eta_t)dW_t+\eta X^\eta_{1,t}(-[a\partial_xa](t,X^\eta_t)\mathrm{d}t+\partial_xa(t,X^\eta_t)\mathrm{d}W_t),\quad X^\eta_{1,0}=0. \tag{2.3}$$

[1] the boundedness assumption of a and its derivatives could be weakened to L^p-integrability conditions, up to extra works.

$$\begin{aligned}
\mathrm{d}X^{\eta}_{2,t} =&2X^{\eta}_{1,t}(-[a\partial_x a](t,X^{\eta}_t)\mathrm{d}t+\partial_x a(t,X^{\eta}_t)\mathrm{d}W_t)+\eta X^{\eta}_{2,t}(-[a\partial_x a](t,X^{\eta}_t)\mathrm{d}t+\partial_x a(t,X^{\eta}_t)\mathrm{d}W_t)\\
&+\eta[X^{\eta}_{1,t}]^2(-\partial_x[a\partial_x a](t,X^{\eta}_t)\mathrm{d}t+\partial^2_x a(t,X^{\eta}_t)\mathrm{d}W_t),\quad X^{\eta}_{2,0}=0. \qquad (2.4)\\
\mathrm{d}X^{\eta}_{3,t} =&3X^{\eta}_{2,t}(-[a\partial_x a](t,X^{\eta}_t)\mathrm{d}t+\partial_x a(t,X^{\eta}_t)\mathrm{d}W_t)\ +3[X^{\eta}_{1,t}]^2(-\partial_x[a\partial_x a](t,X^{\eta}_t)\mathrm{d}t+\partial^2_x a(t,X^{\eta}_t)\mathrm{d}W_t),\\
&+\eta X^{\eta}_{3,t}(-[a\partial_x a](t,X^{\eta}_t)\mathrm{d}t+\partial_x a(t,X^{\eta}_t)\mathrm{d}W_t)+3\eta[X^{\eta}_{1,t}][X^{\eta}_{2,t}](-\partial_x[a\partial_x a](t,X^{\eta}_t)\mathrm{d}t+\partial^2_x a(t,X^{\eta}_t)\mathrm{d}W_t)\\
&+\eta[X^{\eta}_{1,t}]^3(-\partial^2_x[a\partial_x a](t,X^{\eta}_t)\mathrm{d}t+\partial^3_x a(t,X^{\eta}_t)\mathrm{d}W_t),\quad X^{\eta}_{3,0}=0. \qquad (2.5)
\end{aligned}$$

Observe that $X^{\eta}_t|_{\eta=0}=x_0$, thus the derivatives at $\eta=0$ have simpler expressions:

$$\begin{aligned}
\mathrm{d}X_{1,t} =&-\frac{1}{2}a^2(t,x_0)\mathrm{d}t+a(t,x_0)\mathrm{d}W_t=\mathrm{d}X^P_t,\\
\mathrm{d}X_{2,t} =&2X_{1,t}(-[a\partial_x a](t,x_0)\mathrm{d}t+\partial_x a(t,x_0)\mathrm{d}W_t),
\end{aligned}$$

with $X_{i,0}=0$ for $i\geq 1$. Then notice that $X^P_t=x_0+X_{1,t}$: hence

$$X_T-X^P_T=X^1_T-(x_0+X_{1,T})=\int_0^1(1-\lambda)X^{\lambda}_{2,T}\mathrm{d}\lambda \qquad (2.6)$$

$$=\frac{1}{2}X_{2,T}+\int_0^1\frac{(1-\lambda)^2}{2}X^{\lambda}_{3,T}\mathrm{d}\lambda \qquad (2.7)$$

using the Taylor expansion formula. As a consequence of the above representation, we obtain an approximation of $\mathbb{E}(h(X_T))$ for a smooth function h:

$$\mathbb{E}[h(X_T)]=\mathbb{E}[h(X^P_T+\frac{X_{2,T}}{2}+...)]=\mathbb{E}[h(X^P_T)]+\mathbb{E}[h^{(1)}(X^P_T)\frac{X_{2,T}}{2}]+... \qquad (2.8)$$

The first term is related to a log-normal model and thus, it is expected to be easily computable numerically. The second term is more delicate: actually, we transform it into a weighed sum of sensitivities of $\mathbb{E}[h(X^P_T+\varepsilon)]$ w.r.t. $\varepsilon=0$. To achieve this transformation, we use a key lemma which proof is given in Subsection 2.6.4

Lemma 2.1.2.1. *Let φ be a C^{∞}_b function and $(\lambda_t)_t$ be a measurable and bounded deterministic function. Let $N\geq 1$ be fixed, and consider measurable and bounded deterministic functions $t\mapsto l_{i,t}$ for $i=1,\dots,N$. Then, using the convention $\mathrm{d}W^1_t=\mathrm{d}W_t$ and $\mathrm{d}W^0_t=\mathrm{d}t$, for any $(I_1,\dots,I_N)\in\{0,1\}^N$ we have:*

$$\begin{aligned}
&\mathbb{E}\Big(\varphi(\int_0^T\lambda_t\mathrm{d}W_t)\int_0^T l_{N,t_N}\int_0^{t_N}l_{N-1,t_{N-1}}\dots\int_0^{t_2}l_{1,t_1}\mathrm{d}W^{I_1}_{t_1}\dots\mathrm{d}W^{I_{N-1}}_{t_{N-1}}\mathrm{d}W^{I_N}_{t_N}\Big)\\
&=\omega(\widehat{l_1},\dots,\widehat{l_N})\partial^{I_1+\dots+I_N}_{\varepsilon^{I_1+\dots+I_N}}\mathbb{E}\Big(\varphi(\int_0^T\lambda_t\mathrm{d}W_t+\varepsilon)\Big)|_{\varepsilon=0}, \qquad (2.9)
\end{aligned}$$

where $\widehat{l_{k,t}}:=l_{k,t}$ if $I_k=0$ and $\widehat{l_{k,t}}:=\lambda_t l_{k,t}$ if $I_k=1$.

Now, apply the above identity to $\varphi(\cdot)=h^{(1)}(x_0-\frac{1}{2}\int_0^T a^2(t,x_0)\mathrm{d}t+\cdot)$, $\lambda_t=a(t,x_0)$ and

$$\frac{X_{2,T}}{2}=\int_0^T\Bigg(\int_0^{t_2}(-\frac{1}{2}a^2(t_1,x_0)\mathrm{d}t_1+a(t_1,x_0)\mathrm{d}W_{t_1})\Bigg)(-[a\partial_x a](t_2,x_0)\mathrm{d}t_2+\partial_x a(t_2,x_0)\mathrm{d}W_{t_2}),$$

to get

$$\mathbb{E}[h^{(1)}(X^P_T)\frac{X_{2,T}}{2}]=C_1(a;x_0)^T_0(\partial^3_{\varepsilon^3}-\frac{3}{2}\partial^2_{\varepsilon^2}+\frac{1}{2}\partial_{\varepsilon})\mathbb{E}(h(X^P_T+\varepsilon))|_{\varepsilon=0}$$

where the operator C_1 is defined by:

$$C_1(l;z)_0^T = \omega(l^2(z), l(z)l^{(1)}(z))_0^T = \int_0^T l_t^2(z) \int_t^T l_s(z) l_s^{(1)}(z) \mathrm{d}s\mathrm{d}t. \quad (2.10)$$

Combine this with (2.8) to obtain that $\mathbb{E}(h(X_T))$ can be approximated by

$$\mathbb{E}[h(X_T^P)] + C_1(a;x_0)_0^T(\partial^3_{\varepsilon^3} - \frac{3}{2}\partial^2_{\varepsilon^2} + \frac{1}{2}\partial_\varepsilon)\mathbb{E}(h(X_T^P + \varepsilon))|_{\varepsilon=0}.$$

So far, the payoff function h is smooth and this does not fit the Call/Put setting; actually, an extra regularization argument and a careful passing to the limit enables to extend the previous formula to any locally Lipschitz h. Additionally, some error estimates are available (see [Gobet 2012a, Theorem 2.2]). All the results are gathered in the following theorem.

Theorem 2.1.2.1. ***(2nd order log-normal approximation with local volatility at spot)****. Assume $(\mathcal{H}^a_{x_0})$. Assume that h is locally Lipschitz in the following sense: for some constant $C_h \geq 0$,*

$$|h(x)| \leq C_h e^{C_h|x|}, \qquad |\frac{h(y)-h(x)}{y-x}| \leq C_h e^{C_h(|x|+|y|)} \quad (\forall y \neq x).$$

Then

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] + C_1(a;x_0)_0^T(\partial^3_{\varepsilon^3} - \frac{3}{2}\partial^2_{\varepsilon^2} + \frac{1}{2}\partial_\varepsilon)\mathbb{E}[h(X_T^P + \varepsilon)]|_{\varepsilon=0} + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}}).$$

where the operator C_1 is defined in (2.10) and O depends notably of the constant C_h.

This formula is referred to as a second order approximation because the residual term is of order three with respect to the amplitude of the volatility coefficient.

Remark 2.1.2.1. *The reader should notice that the expansion formulas are exact for the particular payoff function $h(x) = e^x$ (indeed $\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] = \partial^i_{\varepsilon^i}\mathbb{E}[h(X_T^P + \varepsilon)]|_{\varepsilon=0} = e^{x_0}$ and the sum of the corrective terms is equal to zero). This notably implies that the Call/Put parity relationship is preserved within these approximations, which is an essential property. The reader can verify in Section 2.3 that this martingale property is preserved for higher order approximation formulas.*

Under the current assumptions ($\int_0^T a^2(t, x_0)\mathrm{d}t > 0$), the law of X_T^P is a non-degenerate Gaussian r.v. and thus, the above derivatives are meaningful even for non-smooth h. Following [Gobet 2012a], the Lipschitz regularity can be weakened to Hölder regularity but error estimates in the case of discontinuous function h are not available so far under the current set of assumptions.

▷ **Normal proxy.** Alternatively to a log-normal proxy, we could prefer the use of normal proxy on the asset S: for CEV-type model described in (2.2), it can be justified for β close to 0. The same analysis can be done by considering the *normal proxy* with diffusion coefficient computed at spot: it writes

$$\mathrm{d}S_t^P = \Sigma(t, S_0)\mathrm{d}W_t, \qquad S_0^P = S_0.$$

Then, the distance to the proxy is represented through the interpolation process

$$dS_t^\eta = \eta\Sigma(t, S_t^\eta)dW_t, \quad S_0^\eta = S_0.$$

All the previous computations are very similar, and even simpler because there is no dt-term. We skip details and state directly the result (see [Gobet 2012a, Theorem 2.1]).

Theorem 2.1.2.2. *(2nd order normal approximation with local volatility at spot). Assume $(\mathcal{H}^{\Sigma}_{S_0})$. Assume that h is locally Lipschitz in the following sense: for some constant $C_h \geq 0$,*

$$|h(x)| \leq C_h(1+|x|^{C_h}), \qquad |\frac{h(y)-h(x)}{y-x}| \leq C_h(1+|x|^{C_h}+|y|^{C_h}) \quad (\forall y \neq x).$$

Then

$$\mathbb{E}[h(S_T)] = \mathbb{E}[h(S^P_T)] + C_1(\Sigma; S_0)^T_0 \partial^3_{\varepsilon^3}\mathbb{E}[h(S^P_T+\varepsilon)]|_{\varepsilon=0} + O(\mathcal{M}_1(\Sigma)[\mathcal{M}_0(\Sigma)]^2 T^{\frac{3}{2}}).$$

Remark 2.1.2.2. *As for the log-normal proxy (see Remark 2.1.2.1), the approximation formulas involving the normal proxy do not suffer from numerical arbitrage when using Call/Put payoffs: indeed they are exact for the particular payoff function $h(x) = x$ (indeed $\mathbb{E}[h(S_T)] = \mathbb{E}[h(S^P_T)] = S_0$ and $\partial^i_{\varepsilon^i}\mathbb{E}[h(S^P_T+\varepsilon)]|_{\varepsilon=0} = 0$, $\forall i \geq 2$). This property holds again when considering higher order expansions (see Section 2.3).*

Applying two previous results to the pricing of Call option (i.e. $h(x) = (e^x - K)_+$ in the case of log-normal proxy, and $h(x) = (x-K)+$ in the case of normal proxy), we obtain two different expansions using respectively Black-Scholes formula and Bachelier formula.

Theorem 2.1.2.3. *(2nd order approximations for Call options with local volatility at spot). Assuming $(\mathcal{H}^{a}_{x_0})$ and using the log-normal proxy, one has*

$$\begin{aligned}\mathrm{Call}(e^{x_0}, T, e^k) =& \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_0}T, k) + C_1(a; x_0)^T_0(\partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_0}T, k)\\ &+ O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}}).\end{aligned}$$

Assuming $(\mathcal{H}^{\Sigma}_{S_0})$ and using the normal proxy, one has

$$\begin{aligned}\mathrm{Call}(S_0, T, K) =& \mathrm{Call}^{\mathrm{BA}}(S_0, \bar{\Sigma}^2_{S_0}T, K) + C_1(\Sigma; S_0)^T_0 \partial^3_{S^3}\mathrm{Call}^{\mathrm{BA}}(S_0, \bar{\Sigma}^2_{S_0}T, K)\\ &+ O(\mathcal{M}_1(\Sigma)[\mathcal{M}_0(\Sigma)]^2 T^{\frac{3}{2}}).\end{aligned}$$

2.1.3 Towards Call option approximations with the local volatility at strike and at mid-point

For general payoff functions, the most natural choice seems to choose a proxy with the local volatility frozen at spot. When we are dealing with Call or Put payoffs, the spot and strike variables play a symmetrical role [Dupire 1994], and there is a priori no reason to advantage one or the other one. A first attempt to exploit this duality in proxy expansion is analysed in [Gobet 2012b]. In this subsection, we briefly recall the expansion formulas with a local volatility at strike and then we present new expansion formulas with a local volatility at mid-point $x_{avg} = (x_0+k)/2 = \log\sqrt{S_0 K}$ or $S_{avg} = (S_0+K)/2$. We detail the analysis only for the log-normal proxy. The proofs for the normal proxy are very similar and are left as an exercise to the reader.

To directly obtain expansions formulas with local volatility frozen at strike, the idea is to follow the Dupire approach [Dupire 1994], using explicitly the PDE satisfied by the Call price function $(T, K) \mapsto \mathrm{Call}(S_0, T, K) = \mathbb{E}[(S_T - K)_+]$. Indeed we have that:

$$\begin{cases}\partial_T \mathrm{Call}(S_0, T, K) = \frac{1}{2}\sigma^2(T, K)K^2\partial^2_{K^2}\mathrm{Call}(S_0, T, K),\\ \mathrm{Call}(S_0, 0, K) = (S_0 - K)_+.\end{cases}$$

Thus we do not consider anymore a PDE in the backward variables (t, S) with a Call payoff as a terminal condition, but we now handle a PDE in the forward variables (T, K), with a put payoff as initial condition. This dual PDE has a probabilistic Feynman-Kac representation:

$$\mathrm{Call}(S_0, T, K) = \mathbb{E}[(S_0 - e^{k_T})_+], \tag{2.11}$$

where $(k_t)_{t\in[0,T]}$ is the diffusion process defined by:

$$\mathrm{d}k_t = a(T-t, k_t)\mathrm{d}W_t - \frac{1}{2}a^2(T-t, k_t)\mathrm{d}t,\ k_0 = k = \log(K),$$

where we recall that $a(t,z) = \sigma(t, e^z)$. Thus we are in a position to apply Theorem 2.1.2.1 for the Put payoff function $h(z) = (e^{x_0} - e^z)_+$ with log-strike $x_0 = \log(S_0)$, with a log-normal proxy starting from $K = e^k$ and with the local volatility $\widetilde{a}(t,z) = a(T-t, z)$. In the same way, we can apply Theorem 2.1.2.2 with a normal proxy. As a result, we obtain a variant of Theorem 2.1.2.3 where the Greeks w.r.t. the k_T-variable are naturally transformed into Greeks w.r.t. the strike variable. The final statement is the following result.

Theorem 2.1.3.1. ***(2nd order approximations for Call options with local volatility at strike)****. Assuming* $(\mathcal{H}_k^a)$ *and using the log-normal proxy, one has*

$$\begin{aligned}\mathrm{Call}(e^{x_0}, T, e^k) =& \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_k^2 T, k) + C_1(\widetilde{a}; k)_0^T(\partial^3_{z^3} - \frac{3}{2}\partial^2_{z^2} + \frac{1}{2}\partial_z)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_k^2 T, k)\\ &+ O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}}).\end{aligned}$$

Assuming $(\mathcal{H}_K^\Sigma)$ *and using the normal proxy, one has*

$$\begin{aligned}\mathrm{Call}(S_0, T, K) =& \mathrm{Call}^{\mathrm{BA}}(S_0, \bar{\Sigma}_K^2 T, K) + C_1(\widetilde{\Sigma}; K)_0^T \partial^3_{Z^3}\mathrm{Call}^{\mathrm{BA}}(S_0, \bar{\Sigma}_K^2 T, K)\\ &+ O(\mathcal{M}_1(\Sigma)[\mathcal{M}_0(\Sigma)]^2 T^{\frac{3}{2}}).\end{aligned}$$

Now, in order to obtain approximation formulas for the mid-points x_{avg} or S_{avg}, we perform a Taylor expansion of the local volatility function around these mid-points. We start from the expansions at spot and strike given in Theorems 2.1.2.3 and 2.1.3.1, we consider the average of these expansions and we transform each term to freeze the local volatility function at x_{avg} or S_{avg}. We only give details for the log-normal proxy. We first analyze the corrective terms.

Lemma 2.1.3.1. *Assume* $(\mathcal{H}_{x_0}^a)$*-*$(\mathcal{H}_k^a)$*-*$(\mathcal{H}_{x_{avg}}^a)$*. We have:*

$$\begin{aligned}&\frac{1}{2}C_1(a; x_0)_0^T(\partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_0}^2 T, k) + \frac{1}{2}C_1(\widetilde{a}; k)_0^T(\partial^3_{z^3} - \frac{3}{2}\partial^2_{z^2} + \frac{1}{2}\partial_z)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_k^2 T, k)\\ =&\frac{1}{2}\Big[C_1(a; x_{avg})_0^T - C_1(\widetilde{a}; x_{avg})_0^T\Big](\partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}}).\end{aligned}$$

Proof. We begin with the x_0-Greeks. Perform a zero order Taylor formula for the function $y \mapsto (\partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, y, k)$ at $y = \bar{a}_{x_0}^2 T = \omega(a^2(x_0))_0^T$ around $y = \bar{a}_{x_{avg}}^2 T = \omega(a^2(x_{avg}))_0^T$ and $\forall t \in [0, T]$, for the function $x \mapsto a_t^2(x)$ at $x = x_0$ around $x = x_{avg}$ to obtain:

$$\begin{aligned}&C_1(a; x_0)_0^T(\partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_0}^2 T, k)\\ =&[C_1(a; x_{avg})_0^T + R_1][(\partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) + R_2]\\ =&C_1(a; x_{avg})_0^T(\partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) + (\partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k)R_1 + C_1(a; x_0)_0^T R_2,\end{aligned}$$

where:

$$
\begin{aligned}
R_1 &= \frac{(x_0-k)}{2}\int_0^1 (\partial_x C_1(a;x)_0^T)|_{x=\lambda x_0+(1-\lambda)x_{avg}}\mathrm{d}\lambda,\\
R_2 &= T(\bar{a}_{x_0}^2-\bar{a}_{x_{avg}}^2)\int_0^1 (\partial^4_{yx^3}-\frac{3}{2}\partial^3_{yx^2}+\frac{1}{2}\partial^2_{yx})\mathrm{Call}^{\mathrm{BS}}(x_0,y,k)|_{y=T(\lambda\bar{a}^2_{x_0}+(1-\lambda)\bar{a}^2_{x_{avg}})}\mathrm{d}\lambda,\\
T(\bar{a}_{x_0}^2-\bar{a}_{x_{avg}}^2) &= \frac{(x_0-k)}{2}\int_0^1 (\partial_x\omega(a^2(x))_0^T)|_{x=\lambda x_0+(1-\lambda)x_{avg}}\mathrm{d}\lambda.
\end{aligned}
$$

In view of the definition (2.10) of C_1, the identity (1.8) in Chapter 1 Section 1.2, Corollary 2.6.1.1 and $(\mathcal{H}^a_{x_0})$-$(\mathcal{H}^a_{x_{avg}})$, we readily obtain

$$
\begin{aligned}
&\Big|(\partial^3_{x^3}-\frac{3}{2}\partial^2_{x^2}+\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)R_1\Big|\\
\le&\frac{1}{2}\Big|(\partial^3_{x^3}-\frac{3}{2}\partial^2_{x^2}+\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)(x_0-k)\Big|\Big|\int_0^1 (\partial_x C_1(a;x)_0^T)|_{x=\lambda x_0+(1-\lambda)x_{avg}}\mathrm{d}\lambda\Big|\\
\le_c&[\bar{a}^2_{x_{avg}}T]^{-\frac{1}{2}}\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2\le_c \mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}},
\end{aligned}
$$

$$
\begin{aligned}
\Big|C_1(a;x_0)_0^T R_2\Big| \le_c&[\mathcal{M}_0(a)]^3\mathcal{M}_1(a)T^2\mathcal{M}_0(a)\mathcal{M}_1(a)T\int_0^1[T(\lambda\bar{a}^2_{x_0}+(1-\lambda)\bar{a}^2_{x_{avg}})]^{-\frac{3}{2}}\mathrm{d}\lambda\\
\le_c&\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}.
\end{aligned}
$$

Similarly, using in addition $(\mathcal{H}^a_k)$ we show that:

$$
\begin{aligned}
&C_1(\overline{a};k)_0^T(\partial^3_{z^3}-\frac{3}{2}\partial^2_{z^2}+\frac{1}{2}\partial_z)\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_kT,k)\\
=&\ C_1(\overline{a};x_{avg})_0^T(\partial^3_{z^3}-\frac{3}{2}\partial^2_{z^2}+\frac{1}{2}\partial_z)\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}),\\
=&-C_1(\overline{a};x_{avg})_0^T(\partial^3_{x^3}-\frac{3}{2}\partial^2_{x^2}+\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}),
\end{aligned}
$$

where we have used at the last equality the relation (2.49) in Proposition 2.6.1.3. That completes the proof. □

Second, we analyze the leading order of the formula given in Theorems 2.1.2.3 and 2.1.3.1:

Lemma 2.1.3.2. *Assume $(\mathcal{H}^a_{x_0})$-$(\mathcal{H}^a_k)$-$(\mathcal{H}^a_{x_{avg}})$. We have:*

$$
\frac{1}{2}[\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k)+\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_kT,k)]=\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}).
$$

Proof. Apply a first order Taylor formula twice; firstly for the function $y\mapsto\mathrm{Call}^{\mathrm{BS}}(x_0,y,k)$ at $y=\bar{a}^2_{x_0}T$ around $y=\bar{a}^2_{x_{avg}}T$ and secondly, for the function $x\mapsto a_t^2(x)$ at $x=x_0$ around $x=x_{avg}$, $\forall t\in[0,T]$. It gives

$$
\begin{aligned}
\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k)=&\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)+\partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)T(\bar{a}^2_{x_0}-\bar{a}^2_{x_{avg}})+R_1,\\
=&\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)+\partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\omega(a(x_{avg})a^{(1)}(x_{avg}))_0^T(x_0-k)+R_2+R_1.
\end{aligned}
$$

where:

$$
R_1=T^2(\bar{a}^2_{x_0}-\bar{a}^2_{x_{avg}})^2\int_0^1(\partial^2_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0,y,k))|_{y=T(\lambda\bar{a}^2_{x_0}+(1-\lambda)\bar{a}^2_{x_{avg}})}(1-\lambda)\mathrm{d}\lambda,
$$

$$R_2 = \partial_y \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_{avg}} T, k)\frac{(x_0-k)^2}{4}\int_0^1 (\partial^2_{x^2}\omega(a^2(x))_0^T)|_{x=\lambda x_0+(1-\lambda)x_{avg}}(1-\lambda)\mathrm{d}\lambda.$$

Similar arguments previously employed in the proof of Lemma 2.1.3.1 easily lead to:

$$\begin{aligned}|R_1| \leq_c &[\mathcal{M}_1(a)]^2[\mathcal{M}_0(a)]^2T^2\int_0^1 [T(\lambda\bar{a}^2_{x_0}+(1-\lambda)\bar{a}^2_{x_{avg}})]^{-\frac{1}{2}}\mathrm{d}\lambda \leq_c \mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}},\\ |R_2| \leq_c &\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}.\end{aligned}$$

Similarly we have:

$$\begin{aligned}\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_k T,k) =& \mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) - \partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\omega(a(x_{avg})a^{(1)}(x_{avg}))_0^T(x_0-k)\\ &+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}).\end{aligned}$$

We are finished. □

Lemmas 2.1.3.1 and 2.1.3.2 lead to the following Theorem for the log-normal proxy, while similar arguments apply for the normal proxy.

Theorem 2.1.3.2. ***(2nd order approximations for Call options with local volatility at mid-point).*** *Under* $(\mathcal{H}^a_{x_0})$-$(\mathcal{H}^a_k)$-$(\mathcal{H}^a_{x_{avg}})$*, we have*

$$\begin{aligned}\mathrm{Call}(e^{x_0},T,e^k) =& \mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) + \frac{C_1(a;x_{avg})_0^T - C_1(\widetilde{a};x_{avg})_0^T}{2}(\partial^3_{x^3}-\frac{3}{2}\partial^2_{x^2}+\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\\ &+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}),\end{aligned} \tag{2.12}$$

Under $(\mathcal{H}^\Sigma_{S_0})$-$(\mathcal{H}^\Sigma_K)$-$(\mathcal{H}^\Sigma_{S_{avg}})$*, we have*

$$\begin{aligned}\mathrm{Call}(S_0,T,K) =& \mathrm{Call}^{\mathrm{BA}}(S_0,\bar{\Sigma}^2_{S_{avg}}T,K) + \frac{C_1(\Sigma;S_{avg})_0^T - C_1(\widetilde{\Sigma};S_{avg})_0^T}{2}\partial^3_{S^3}\mathrm{Call}^{\mathrm{BA}}(S_0,\bar{\Sigma}^2_{S_{avg}}T,K)\\ &+O(\mathcal{M}_1(\Sigma)[\mathcal{M}_0(\Sigma)]^2T^{\frac{3}{2}}).\end{aligned} \tag{2.13}$$

Remark 2.1.3.1. *If a (and consequently* Σ*) is time-independent or has separable variables, observe that the corrective terms vanish and we obtain remarkably simple formulas: the expansion formulas* (2.12) *and* (2.13) *reduce to only a Black-Scholes price and a Bachelier price, with the local volatility function frozen at the mid-point.*

2.1.4 Second order expansion of the implied volatility

Interestingly, the previous expansions of Call price (Theorems 2.1.2.3, 2.1.3.1 and 2.1.3.2) can be turned into expansions of Black-Scholes and Bachelier implied volatility defined respectively in (1.9) and (1.11) (see Chapter 1 Section 1.2). To achieve this, we use the relations between Greeks postponed in Propositions 2.6.1.3 and 2.6.2.3 in order to write the different approximation formulas in terms of the Vega. For example consider the second order log-normal expansion formula based on the ATM local volatility (Theorem 2.1.2.3): thanks to (2.49) in Proposition 2.6.1.3, it becomes:

$$\begin{aligned}\mathrm{Call}(e^{x_0},T,e^k) =& \mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k) - \mathrm{Vega}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k)\frac{C_1(a;x_0)_0^T m}{\bar{a}^3_{x_0}T^2} + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}),\\ \approx& \mathrm{Call}^{\mathrm{BS}}\Big(x_0,(\bar{a}_{x_0}-\frac{C_1(a;x_0)_0^T}{\bar{a}^3_{x_0}T^2}m)^2T,k\Big),\end{aligned}$$

where m is the log-moneyness $m = x_0 - k = \log(S_0/K)$. We have paved the way for the following result:

Theorem 2.1.4.1. *(**2nd order expansions of the implied volatility**). Assuming ($\mathcal{H}^a_{x_0}$)-($\mathcal{H}^a_k$)-($\mathcal{H}^a_{x_{avg}}$) and using the log-normal proxy, we have*

$$\sigma_{\mathrm{I}}(x_0,T,k) = \bar{a}_{x_0} - \frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3 T^2} m + \mathrm{Error}^{\mathrm{I}}_{2,\mathrm{x_0}}, \tag{2.14}$$

$$\sigma_{\mathrm{I}}(x_0,T,k) = \bar{a}_k + \frac{C_1(\widetilde{a};k)_0^T}{\bar{a}_k^3 T^2} m + \mathrm{Error}^{\mathrm{I}}_{2,\mathrm{k}}, \tag{2.15}$$

$$\sigma_{\mathrm{I}}(x_0,T,k) = \bar{a}_{x_{avg}} + \frac{(C_1(\widetilde{a};x_{avg})_0^T - C_1(a;x_{avg})_0^T)}{2\bar{a}_{x_{avg}}^3 T^2} m + \mathrm{Error}^{\mathrm{I}}_{2,\mathrm{x_{avg}}}. \tag{2.16}$$

Assuming ($\mathcal{H}^\Sigma_{S_0}$)-($\mathcal{H}^\Sigma_K$)-($\mathcal{H}^\Sigma_{S_{avg}}$) and using the normal proxy, we have

$$\Sigma_{\mathrm{I}}(S_0,T,K) = \bar{\Sigma}_{S_0} - \frac{C_1(\Sigma;S_0)_0^T}{\bar{\Sigma}_{S_0}^3 T^2} M + \mathrm{Error}^{\mathrm{I}}_{2,\mathrm{S_0}},$$

$$\Sigma_{\mathrm{I}}(S_0,T,K) = \bar{\Sigma}_K + \frac{C_1(\widetilde{\Sigma};K)_0^T}{\bar{\Sigma}_K^3 T^2} M + \mathrm{Error}^{\mathrm{I}}_{2,\mathrm{K}},$$

$$\Sigma_{\mathrm{I}}(S_0,T,K) = \bar{\Sigma}_{S_{avg}} + \frac{(C_1(\widetilde{\Sigma};S_{avg})_0^T - C_1(\Sigma;S_{avg})_0^T)}{2\bar{\Sigma}_{S_{avg}}^3 T^2} M + \mathrm{Error}^{\mathrm{I}}_{2,\mathrm{S_{avg}}},$$

where $S_{avg} = \frac{S_0+K}{2}$ and $M = S_0 - K$.

Remark 2.1.4.1. *We retrieve in our implied volatility approximation formulas the well-known properties that at the money (i.e. $m = 0$) and for short maturity, the value of the implied volatility is equal to the value of the local volatility function and the slope of the local volatility function is twice the slope of the implied volatility. We justify this assertion for the Black-Scholes implied volatility, the work being similar for the Bachelier one. If $T \ll 1$, in view of (2.14) and the definition (2.10) of C_1, assuming that $a(t,x_0)$ and $a^{(1)}(t,x_0)$ are continuous at $t = 0$, we obtain:*

$$[\sigma_{\mathrm{I}}(x_0,T,k)]|_{k=x_0} \approx a(0,x_0),$$

$$\partial_k[\sigma_{\mathrm{I}}(x_0,T,k)]|_{k=x_0} \approx \partial_k[\bar{a}_{x_0}]|_{k=x_0} - \frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3 T^2}\partial_k[(x_0-k)]|_{k=x_0} \approx 0 + \frac{a^3(0,x_0)a^{(1)}(0,x_0)\frac{T^2}{2}}{a^3(0,x_0)T^2} = \frac{a^{(1)}(0,x_0)}{2}.$$

We obtain the same estimates starting from (2.15) and (2.16), we skip details.

To conclude this paragraph, we estimate the residual terms of the above implied volatility expansions, in terms of $\mathcal{M}_0(a)$, $\mathcal{M}_1(a)$ and so on. Since the Vega is very small for far OTM/ITM Call options, deriving error bounds on implied volatility from Theorems 2.1.2.3, 2.1.3.1 and 2.1.3.2 gives poor estimates for extreme strikes. Actually, in the further numerical experiments, we also observe inaccuracies for extreme strikes. To obtain accurate theoretical error bounds, we restrict to log-moneyness m (resp moneyness M) belonging to a small ball by assuming that $|m| \le \xi\mathcal{M}_0(a)\sqrt{T}$ (resp. $|M| \le \xi\mathcal{M}_0(\Sigma)\sqrt{T}$) for a given $\xi > 0$.
For the sake of brevity, we only analyze the expansion (2.14), the other approximations being treated similarly. We assume in addition that $\mathcal{M}_0(a)$, $\mathcal{M}_1(a)$ and T are globally small enough to ensure that $\bar{a}_{x_0} - \frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3 T^2} m > 0$. Note that at the money (i.e. $m = 0$), this condition is automatically satisfied. A first order expansion readily gives

$$\mathrm{Call}^{\mathrm{BS}}(x_0,(\bar{a}_{x_0} - \frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3 T^2} m)^2 T,k) = \mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}_{x_0}^2 T,k) - \frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3 T^2} m \mathrm{Vega}^{\mathrm{BS}}(x_0,\bar{a}_{x_0}^2 T,k)$$

$$
\begin{aligned}
&+(\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3 T^2}m)^2\int_0^1 \mathrm{Vomma}^{\mathrm{BS}}(x_0,a^2T,k)|_{a=\bar{a}_{x_0}-\lambda\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3T^2}m}(1-\lambda)\mathrm{d}\lambda\\
=&\mathrm{Call}^{\mathrm{BS}}(x_0,\sigma_{\mathrm{I}}^2(x_0,T,k)T,k)+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}})\\
&+(\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3 T^2}m)^2\int_0^1 \mathrm{Vomma}^{\mathrm{BS}}(x_0,a^2T,k)|_{a=\bar{a}_{x_0}-\lambda\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3T^2}m}(1-\lambda)\mathrm{d}\lambda,
\end{aligned}
$$

applying Theorem 2.1.2.3 and using the definition of the Black-Scholes implied volatility. Expanding $a\mapsto \mathrm{Call}^{\mathrm{BS}}(x_0,a^2T,k)$ at $a=\sigma_{\mathrm{I}}(x_0,T,k)$ around $a=\bar{a}_{x_0}-\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3T^2}m$ gives:

$$
\begin{aligned}
&\mathrm{Error}_{2,x_0}^{\mathrm{I}}\int_0^1 \mathrm{Vega}^{\mathrm{BS}}(x_0,a^2T,k)|_{a=\bar{a}_{x_0}-\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3T^2}m+\lambda\,\mathrm{Error}_{2,x_0}^{\mathrm{I}}}\mathrm{d}\lambda\\
=&O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}})-(\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3 T^2}m)^2\int_0^1 \mathrm{Vomma}^{\mathrm{BS}}(x_0,a^2T,k)|_{a=\bar{a}_{x_0}-\lambda\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3T^2}m}(1-\lambda)\mathrm{d}\lambda.
\end{aligned}
$$

In view of the expression of $\mathrm{Vega}^{\mathrm{BS}}$ (see (2.42) in Proposition 2.6.1.2) and (2.45) in Corollary 2.6.1.2, the hypotheses made on m, $\mathcal{M}_0(a)$, $\mathcal{M}_1(a)$ and T guarantee the existence of a constant $C>0$ (depending on S_0) such that:

$$
\int_0^1 \mathrm{Vega}^{\mathrm{BS}}(x_0,a^2T,k)|_{a=\bar{a}_{x_0}-\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3T^2}m+\lambda\,\mathrm{Error}_{2,x_0}^{\mathrm{I}}}\mathrm{d}\lambda\geq C\sqrt{T}>0.
$$

In addition (2.46) and $(\mathcal{H}_{x_0}^a)$ readily yield

$$
\begin{aligned}
&\Big|(\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3 T^2}m)^2\int_0^1 \mathrm{Vomma}^{\mathrm{BS}}(x_0,a^2T,k)|_{a=\bar{a}_{x_0}-\lambda\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3T^2}m}(1-\lambda)\mathrm{d}\lambda\Big|\\
\leq_c&(\mathcal{M}_1(a)\mathcal{M}_0(a)\sqrt{T})^2\frac{\sqrt{T}}{\inf_{\lambda\in[0,1]}\bar{a}_{x_0}-\lambda\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3T^2}m}\leq_c \mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}},
\end{aligned}
$$

where the generic constant depends in an increasing way on ξ (and on the oscillation ratio $\frac{\mathcal{M}_0(a)}{\inf_{\lambda\in[0,1]}\bar{a}_{x_0}-\lambda\frac{C_1(a;x_0)_0^T}{\bar{a}_{x_0}^3T^2}m}$). That finally implies:

$$
\mathrm{Error}_{2,x_0}^{\mathrm{I}}=O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T).
$$

In view of the above upper bound, we interpret our implied volatility formulas as second order expansion ones.

2.2 Proofs: a comparative discussion between stochastic analysis and PDE techniques

In this section, our aim is to show how three different techniques ranging from stochastic analysis to PDE may lead to the same formulas given in Theorem 2.1.2.1. Although the final result is the same, the derivation is quite different, first regarding the way in which the expansion coefficients are made explicit, second regarding the error estimates and the assumptions used for that.

We shall admit that our preference is for the stochastic analysis approach, because it is flexible regarding the model and the functionals under consideration, and it is slightly less demanding regarding the assumptions (pointwise ellipticity versus uniform ellipticity for instance). But the reader may argue differently, depending on its own fields of expertise.
As an illustration of flexibility of the stochastic analysis approach, it has been possible to handle Call/Put/digital options in local volatility models with Gaussian jumps [Benhamou 2009], Call/Put options in local volatility models with stochastic Gaussian interest rates [Benhamou 2012], Call/Put options in time-dependent Heston model [Benhamou 2010b], general average options (including Asian and Basket options) in local volatility models [Benhamou 2010a], and more recently local stochastic volatility models (we refer to Chapter 5 of the thesis).

2.2.1 A pure stochastic analysis approach

This is basically the derivation that we have performed in Subsection 2.1.2.

Smooth payoff h. We first deal with the case of infinitely differentiable function h with exponentially bounded derivatives. Resuming from (2.6-2.7-2.8) and using Taylor's formula, write

$$\begin{aligned}
\mathbb{E}[h(X_T)] =& \mathbb{E}[h(X_T^P)] + \mathbb{E}[h^{(1)}(X_T^P)(X_T - X_T^P)] + \int_0^1 \mathbb{E}[h^{(2)}(X_T^P + \lambda(X_T - X_T^P))(X_T - X_T^P)^2](1-\lambda)\mathrm{d}\lambda \\
=& \mathbb{E}[h(X_T^P)] + \mathbb{E}[h^{(1)}(X_T^P)\frac{X_{2,T}}{2}] + \int_0^1 \mathbb{E}[h^{(1)}(X_T^P)\frac{(1-\lambda)^2}{2}X_{3,T}^\lambda]\mathrm{d}\lambda \\
&+ \int_0^1 \mathbb{E}[h^{(2)}(X_T^P + \lambda(X_T - X_T^P))(\int_0^1 X_{2,T}^\eta (1-\eta)\mathrm{d}\eta)^2](1-\lambda)\mathrm{d}\lambda \\
:=& \mathbb{E}[h(X_T^P)] + \mathbb{E}[h^{(1)}(X_T^P)\frac{X_{2,T}}{2}] + \mathrm{Error}_2(h).
\end{aligned} \tag{2.17}$$

The first correction term $\mathbb{E}[h^{(1)}(X_T^P)\frac{X_{2,T}}{2}]$ is made explicit using the key Lemma 2.1.2.1, and it is equal to a weighted summation of sensitivities $\partial_\varepsilon^i \mathbb{E}[h(X_T^P + \varepsilon)]|_{\varepsilon=0}$ for $i = 1, 2, 3$ (see the statement of Theorem 2.1.2.1).

The evaluation of $\mathrm{Error}_2(h)$ requires to estimate the L^p-norms of $X_{2,T}^\lambda$ and $X_{3,T}^\lambda$ (uniformly in $\lambda \in [0, 1]$). Direct and standard stochastic calculus inequalities from (2.3-2.4-2.5) yield

$$|X_{2,T}^\lambda|_p \le_c \mathcal{M}_1(a)\mathcal{M}_0(a)T, \qquad |X_{3,T}^\lambda|_p \le_c \mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{3/2} \tag{2.18}$$

for any $p \ge 1$ and any $\lambda \in [0, 1]$. Combining these estimates with Hölder and Minkowski inequalities readily gives $\mathrm{Error}_2(h) = O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}})$, which completes the proof if h is smooth as above. Observe that we have only required the coefficients to be smooth enough, and nothing has been imposed on the non-degeneracy of a.

Locally Lipschitz function h. We now extend the analysis to functions satisfying conditions of Theorem 2.1.2.1 (thus almost everywhere differentiable), assuming additionally $(\mathcal{H}_{x_0}^a)$: observe that the pointwise ellipticity condition $\int_0^T a^2(t, x_0)\mathrm{d}t > 0$ is necessary to ensure that X^P-Greeks are well defined. The analysis below shows that the condition is also sufficient to obtain the expansion.
The new ingredient consists in *appropriately* smoothing h and in using integration-by-parts formula from Malliavin calculus to get rid of the derivatives of h; this follows the arguments of [Gobet 2012a]. Let B be another scalar Brownian motion independent of W and for $\delta > 0$, set

$$h_\delta(x) := \mathbb{E}(h(x + \delta B_{2T})) = \mathbb{E}(h_{\delta/\sqrt{2}}(x + \delta B_T)).$$

For any $\delta > 0$, the function h_δ is smooth and its derivatives are exponentially bounded, so that we can apply the previous expansion to h_δ instead of h in order to obtain:

$$\mathbb{E}[h_\delta(X_T)] = \mathbb{E}[h_\delta(X_T^P)] + C_1(a;x_0)_0^T(\partial^3_{\varepsilon^3} - \frac{3}{2}\partial^2_{\varepsilon^2} + \frac{1}{2}\partial_\varepsilon)\mathbb{E}[h_\delta(X_T^P+\varepsilon)]|_{\varepsilon=0} + \mathrm{Error}_2(h_\delta).$$

Take $\delta = \mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T$: then replacing $\mathbb{E}(h_\delta(X_T))$ and $\mathbb{E}(h_\delta(X_T^P))$ by $\mathbb{E}(h(X_T))$ and $\mathbb{E}(h(X_T^P))$ readily yields an extra error $O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}})$ which has the right magnitude regarding the expected global error. Moreover using $(\mathcal{H}^a_{x_0})$, we can also prove that computing the sensitivities with respect to h or to h_δ does not deteriorate the global accuracy (see [Gobet 2012a, Lemma 4.2]). It remains to prove that $\mathrm{Error}_2(h_\delta) = O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}})$. An inspection of the representation (2.17) of $\mathrm{Error}_2(h_\delta)$ shows immediately that the first contribution with $h^{(1)}_\delta$ is a $O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}})$, by simply using the exponential growth condition on $h^{(1)}$ and the finiteness of exponential moments of X_T^η .The second contribution with $h^{(2)}_\delta$ is the integral over $(\eta_1,\eta_2,\lambda) \in [0,1]^3$ of $(1-\eta_1)(1-\eta_2)(1-\lambda)$ times

$$\begin{aligned}\mathbb{E}[h^{(2)}_\delta(X_T^P + \lambda(X_T - X_T^P))X^{\eta_1}_{2,T}X^{\eta_2}_{2,T}] =& \mathbb{E}[h^{(2)}_{\delta/\sqrt{2}}(X_T^P + \lambda(X_T - X_T^P) + \delta B_T)X^{\eta_1}_{2,T}X^{\eta_2}_{2,T}]\\ =& \mathbb{E}[h^{(1)}_{\delta/\sqrt{2}}(X_T^P + \lambda(X_T - X_T^P) + \delta B_T)H_1^{\delta,\eta_1,\eta_2,\lambda}].\end{aligned}$$

The first equality follows from the definition of h_δ, whereas the second one is an integration by parts formula from Malliavin calculus [Nualart 2006, Proposition 2.1.4]. We do not enter into the derivation details, we only emphasize two points: first, it is allowed since $X_T^P + \lambda(X_T - X_T^P) + \delta B_T$ is a non-degenerate random variable (in Malliavin sense) thanks to the additional perturbation δB_T, and its Malliavin matrix has an inverse of order $(\int_0^T a^2(t,x_0)\mathrm{d}t)^{-1}$ in L^p-norms, owing to the ellipticity assumption in $(\mathcal{H}^a_{x_0})$. Second, the Malliavin norms of $X^\eta_{2,T}$ can be estimated similarly to (2.18) and it finally gives that $(\mathbb{E}|H_1^{\delta,\eta_1,\eta_2,\lambda}|^2)^{1/2} = O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}})$. This finishes the proof. Slight modifications in the above arguments would enable to handle functions with local Hölder smoothness.

Arbitrary function h. Here, we do not assume any regularity on h, only exponential growth. The analysis is similar but the regularization step for h is more complex, see [Benhamou 2009]: the expansion analysis has been done under the uniform ellipticity condition on $(\mathcal{H}^a)$, and not only under the pointwise ellipticity in $(\mathcal{H}^a_{x_0})$.

As a conclusion to this *stochastic analysis approach*:

- the derivation of expansion coefficients is direct and easy;
- the error analysis relies on delicate Malliavin calculus estimates;
- it applies to general function h under mild non-degeneracy condition.

2.2.2 Mixing stochastic analysis and PDE

Here, we directly prove the expansion result for locally Lipschitz function h. We represent the error $\mathbb{E}[h(X_T)] - \mathbb{E}[h(X_T^P)]$ using the PDE associated to the proxy:

$$u^{P,h}(t,x) = \mathbb{E}[h(X_T^P)|X_t^P = x]$$

Our methodology presents analogies with the decomposition approach of Alòs [Alòs 2012] who combines PDE arguments and the Itô calculus to obtain approximation formulas in the Heston model.
To get a smooth solution u^P, assume that $a(t,x_0) \neq 0$ for any $t \in [0,T]$ $(\widetilde{\mathcal{H}}^a_{x_0})$, which is stronger that

$\int_0^T a^2(t,x_0)\mathrm{d}t > 0$ considered in $(\mathcal{H}^a_{x_0})$. The generic constants appearing in our next error estimates depend in an increasing way on the oscillation ratio $\frac{\mathcal{M}_0(a)}{\inf_{t\in[0,T]} a(t,x_0)}$. Then,

$$\begin{cases}\partial_t u^{P,h}(t,x) + \frac{1}{2}a^2(t,x_0)(\partial^2_{x^2} - \partial_x)u^{P,h}(t,x) = 0, & \text{for } t < T,\\ u^{P,h}(T,x) = h(x),\end{cases} \tag{2.19}$$

$$|\partial^n_{x^n} u^{P,h}(t,x)| \leq_c e^{c|x|}(\int_t^T a^2(s,x_0)\mathrm{d}s)^{-\frac{n-1}{2}}. \tag{2.20}$$

The estimates (2.20) directly follow from the differentiation of the Gaussian density of X^P_T conditionally to $X^p_t = x$, taking into account the exponential growth of h. Then, apply Itô's formula to $u^{P,h}(t,X_t)$ between $t=0$ and $t=T$, combine this with simplifications coming from the PDE solved by $u^{P,h}$; it gives

$$\begin{aligned}\mathbb{E}[h(X_T)] =&\mathbb{E}[h(X^P_T)] + \frac{1}{2}\mathbb{E}[\int_0^T (a^2(t,X_t) - a^2(t,x_0))(\partial^2_{x^2} - \partial_x)u^{P,h}(t,X_t)\mathrm{d}t] &(2.21)\\ =&\mathbb{E}[h(X^P_T)] + \frac{1}{2}\int_0^T \partial_x[a^2](t,x_0)\mathbb{E}\Big[(X_t - x_0)(\partial^2_{x^2} - \partial_x)u^{P,h}(t,X_t)\Big]\mathrm{d}t\\ &+ \frac{1}{2}\mathbb{E}\Big[\int_0^T [a^2(t,X_t) - a^2(t,x_0) - \partial_x[a^2](t,x_0)(X_t - x_0)](\partial^2_{xx} - \partial_x)u^{P,h}(t,X_t)\mathrm{d}t\Big]. &(2.22)\end{aligned}$$

Remark 2.2.2.1. *The equality (2.21) says that if h is smooth (infinitely differentiable with compact support), then the approximation $\mathbb{E}[h(X_T)] \approx \mathbb{E}[h(X^P_T)]$ is already of order 2 and thus there is no need to additional corrective terms . We namely have in this case:*

$$\Big|\mathbb{E}[h(X_T)] - \mathbb{E}[h(X^P_T)]\Big| \leq_c \int_0^T \mathcal{M}_0(a)\mathcal{M}_1(a)|X_t - x_0|_4\mathrm{d}t \leq_c \mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}. \tag{2.23}$$

In light with Chapter 1 Subsection 1.3.5, we retrieve the fact that to achieve a target order error, the expansion strongly depends on the regularity of the function h. With the stochastic analysis approach, a straightforward and naive estimate using a Taylor expansion gives:

$$\Big|\mathbb{E}[h(X_T)] - \mathbb{E}[h(X^P_T)]\Big| \leq_c \Big|\int_0^1 \mathbb{E}[h^{(1)}(X^P_T + \lambda(X_T - X^P_T))(X_T - X^P_T)]\mathrm{d}\lambda\Big| = O(\mathcal{M}_1(a)\mathcal{M}_0(a)T).$$

However, we can refine the above estimate using the h-smoothness by writing $h^{(1)}(X^P_T + \lambda(X_T - X^P_T)) = h^{(1)}(x_0) + h^{(1)}(X^P_T + \lambda(X_T - X^P_T)) - h^{(1)}(x_0)$ to get:

$$\begin{aligned}\Big|\mathbb{E}[h(X_T)] - \mathbb{E}[h(X^P_T)]\Big| =&\Big|h^{(1)}(x_0)\mathbb{E}[X_T - X^P_T] + \int_0^1 \mathbb{E}[(h^{(1)}(X^P_T + \lambda(X_T - X^P_T)) - h^{(1)}(x_0))(X_T - X^P_T)]\mathrm{d}\lambda\Big|\\ \leq_c&(|h^{(1)}|_\infty + |h^{(2)}|_\infty)\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}},\end{aligned}$$

using the fact that $\mathbb{E}[X_T - X^P_T] = -\frac{1}{2}\mathbb{E}[\int_0^T (a_t^2(X_t) - a_t^2(x_0))\mathrm{d}t] = O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}})$ and that $\Big|h^{(1)}(X^P_T + \lambda(X_T - X^P_T)) - h^{(1)}(x_0)\Big| \leq |h^{(2)}|_\infty(|X^P_T - x_0| + |X_T - X^P_T|)$. We finally retrieve the estimate (2.23).

Taking advantage of (2.20), we can easily bound the last term in (2.22) by

$$C\int_0^T \mathcal{M}_0(a)\mathcal{M}_1(a)|X_t - x_0|_4^2(\int_t^T a^2(s,x_0)\mathrm{d}s)^{-\frac{1}{2}}\mathrm{d}t = O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}})$$

using standard increment estimates and uniform lower and upper bounds on a^2. Observe that the Lipschitz regularity of h gives rise to singular terms of the form $(T-t)^{-\frac{1}{2}}$, which are fortunately integrable at T.

Regarding the second term in (2.22), we have to approximate $\mathbb{E}\Big[(X_t-x_0)(\partial^2_{x^2}-\partial_x)u^{P,h}(t,X_t)\Big]$ for any $t\in[0,T[$: we apply again the previous decomposition by replacing T by t and $h(x)$ by $\phi_t(x)=(x-x_0)(\partial^2_{xx}-\partial_x)u^{P,h}(t,x)$. We denote by $v^{P,\phi}_t(s,x)=\mathbb{E}[\phi_t(X^P_t)|X^P_s=x]$ the solution of the system (2.19) on $[0,t[\times\mathbb{R}$ but with terminal condition ϕ_t. The term under study is thus equal to

$$\mathbb{E}[\phi_t(X^P_t)]+\frac{1}{2}\mathbb{E}[\int_0^t(a^2(s,X_s)-a^2(s,x_0))(\partial^2_{x^2}-\partial_x)v^{P,\phi}_t(s,X_s)\mathrm{d}s].$$

It remains to make explicit $v^{P,\phi}_t(s,x)$ in order to compute the first term and to estimate the second. For this, the trick lies in the observation that for any $k\geq 0$, $M_{k,t}=\partial^k_{x^k}u^{P,h}(t,X^P_t)$ is a martingale for $t<T$: this directly follows from the application of Itô's formula, combined with (2.19) and (2.20). Hence, successive applications of the equalities $\mathbb{E}[M_{k,t}|X^P_s=x]=\partial^k_{x^k}u^{P,h}(s,x)$ for $s\leq t$ and of the Lemma 2.1.2.1 give:

$$\begin{aligned}
v^{P,\phi}_t(s,x)=&\mathbb{E}\Big[(X^P_t-x_0)(\partial^2_{x^2}-\partial_x)u^{P,h}(t,X^P_t)|X^P_s=x\Big]\\
=&(x-x_0)\mathbb{E}\Big[(\partial^2_{x^2}-\partial_x)u^{P,h}(t,X^P_t)|X^P_s=x\Big]+\mathbb{E}\Big[(X^P_t-X^P_s)(\partial^2_{x^2}-\partial_x)u^{P,h}(t,X^P_t)|X^P_s=x\Big]\\
=&(x-x_0)(\partial^2_{x^2}-\partial_x)u^{P,h}(s,x)-\frac{1}{2}\int_s^t a^2(\xi,x_0)\mathrm{d}\xi\mathbb{E}\Big[(\partial^2_{x^2}-\partial_x)u^{P,h}(t,X^P_t)|X^P_s=x\Big]\\
&+\mathbb{E}\Big[(\int_s^t a(\xi,x_0)\mathrm{d}W_\xi)(\partial^2_{x^2}-\partial_x)u^{P,h}(t,x-\frac{1}{2}\int_s^t a^2(\xi,x_0)\mathrm{d}\xi+\int_s^t a(\xi,x_0)\mathrm{d}W_\xi)|X^P_s=x\Big]\\
=&(x-x_0)(\partial^2_{x^2}-\partial_x)u^{P,h}(s,x)-\frac{1}{2}\int_s^t a^2(\xi,x_0)\mathrm{d}\xi(\partial^2_{x^2}-\partial_x)u^{P,h}(s,x)\\
&+\int_s^t a^2(\xi,x_0)\mathrm{d}\xi\mathbb{E}\Big[(\partial^3_{x^3}-\partial^2_{x^2})u^{P,h}(t,X^P_t)|X^P_s=x\Big]\\
=&(x-x_0)(\partial^2_{x^2}-\partial_x)u^{P,h}(s,x)+\int_s^t a^2(\xi,x_0)\mathrm{d}\xi(\partial^3_{x^3}-\frac{3}{2}\partial^2_{x^2}+\frac{1}{2}\partial_x)u^{P,h}(s,x).
\end{aligned}$$

In particular the above calculus yields $\mathbb{E}[\phi_t(X^P_t)]=v^{P,\phi}_t(0,x_0)=\int_0^t a^2(s,x_0)\mathrm{d}s(\partial^3_{x^3}-\frac{3}{2}\partial^2_{x^2}+\frac{1}{2}\partial_x)u^{P,h}(0,x_0)$ and by multiplying by $\frac{1}{2}\partial_x[a^2](t,x_0)$ and integrating over $t\in[0,T]$ in (2.22), we recover the correction terms from Theorem 2.1.2.1.

On the other hand, combining (2.20) and the ellipticity assumption $(\widetilde{\mathcal{H}}^a_{x_0})$, we easily obtain

$$\begin{aligned}
\Big|(\partial^2_{x^2}-\partial_x)v^{P,\phi}_t(s,X_s)\Big|_p\leq_c&|X_s-x_0|_{2p}\Big|(\partial^4_{x^4}-2\partial^3_{x^3}+\partial^2_{x^2})u^{P,h}(s,X_s)\Big|_{2p}+\Big|(2\partial^3_{x^3}-3\partial^2_{x^2}+\partial_x)u^{P,h}(s,X_s)\Big|_p\\
&+\int_s^t a^2(\xi,x_0)\mathrm{d}\xi\Big|(\partial^5_{x^5}-\frac{5}{2}\partial^4_{x^4}+2\partial^3_{x^3}-\frac{1}{2}\partial^2_{x^2})u^{P,h}(s,X_s)\Big|_p\\
\leq_c&\frac{\sqrt{s}}{\inf\limits_{s\in[0,T]}a^2(s,x_0)(T-s)^{\frac{3}{2}}}+\frac{1}{\inf\limits_{s\in[0,T]}a^2(s,x_0)(T-s)},
\end{aligned}$$

for any $p\geq 1$, $t\in[0,T[$, $s\in[0,t[$. Consequently we obtain for the final error:

$$\begin{aligned}
&\Big|\int_0^T\partial_x[a^2](t,x_0)\mathbb{E}[\int_0^t(a^2(s,X_s)-a^2(s,x_0))(\partial^2_{x^2}-\partial_x)v^{P,\phi}_t(s,X_s)\mathrm{d}s\Big]\mathrm{d}t\Big|\\
\leq_c&\mathcal{M}_1(a)\mathcal{M}_0(a)\int_0^T\int_0^t|X_s-x_0|_2[\frac{\sqrt{s}}{(T-s)^{\frac{3}{2}}}+\frac{1}{(T-s)}]\mathrm{d}s\mathrm{d}t
\end{aligned}$$

$$\leq_c \mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}}.$$

We have retrieved the error estimate provided in Theorem 2.1.2.1. Once again, we would like to point out that the singular terms $(T-s)^{-3/2}$ and $(T-s)^{-1}$ appearing in the above time iterated integral remain integrable.

As a conclusion to this approach *mixing stochastic analysis and PDE*:

- the error analysis relies on usual estimates of derivatives of heat equations (PDE satisfied by the proxy) and it may be considered easier; however for digital options, the singularities arising in iterated time integrals are not integrable and the current approach seems to be inappropriate.
- this approach requires stronger non-degeneracy assumptions compared to the previous stochastic analysis approach;
- the explicit derivation of expansion coefficients is tricky and relies on appropriate combination of martingale properties and Itô calculus;
- we nevertheless mention that this approach could be potentially used in a framework where the Malliavin calculus fails, e.g. for barrier options: see Chapter 7 of the thesis.

Actually for higher order expansion, the latter explicit martingale computation is harder to write down, whereas a direct application of Lemma 2.1.2.1 remains direct.

2.2.3 A pure PDE approach

Alternatively, inspired by the interpolation (1.25-1.26) (see Chapter 1 Subsection 1.3.4.4), consider the solution of the PDE

$$\begin{cases} \partial_t u^\eta(t,x) + \frac{1}{2}a^2(t, x_0 + \eta(x-x_0))(\partial^2_{x^2} - \partial_x)u^\eta(t,x) = 0, & \text{for } t < T, \\ u^\eta(T,x) = h(x). \end{cases}$$

Observe that $u^1(0,x_0)$ coincides with $\mathbb{E}(h(X_T))$ whereas $u^0(0,x_0)$ coincides with $\mathbb{E}(h(X^P_T))$. This PDE has similarities with that of Hagan (1.22) (see Chapter 1 Subsection 1.3.4.4) but it differs here, because the space variable has not been rescaled around the strike. In addition the solution of the principal PDE in the Hagan approach is a Call price in a Bachelier model, whereas $u^0(0,x_0)$ is a Call price in a Black-Scholes model.
To derive the correction terms, we shall apply a regular perturbation analysis by writing $u^\eta = u_0 + \eta u_1 + \dots$, with $u_0 = u^0$, and

$$L^\eta = \partial_t + \frac{1}{2}a^2(t, x_0 + \eta(x-x_0))(\partial^2_{x^2} - \partial_x) = L_0 + \eta\frac{1}{2}\partial_x[a^2](t,x_0)(x-x_0)(\partial^2_{x^2} - \partial_x) + \dots.$$

A formal identification of the system to PDEs to solve gives

$$\Big[\partial_t + \frac{1}{2}a^2(t,x_0)(\partial^2_{x^2} - \partial_x)\Big]u_0(t,x) = 0, \quad L_0 u_1 = -\frac{1}{2}\partial_x[a^2](t,x_0)(x-x_0)(\partial^2_{x^2} - \partial_x)u_0$$

with $u_0(T,x) = h(x)$ and $u_1(T,x) = 0$. As mentioned before, in our opinion, an explicit resolution of u_1 is difficult to exhibit without knowing the solution. However, after tedious calculus involving Gaussian kernels and convolutions, we can retrieve the corrective terms of Theorem 2.1.2.1.
Also, a PDE error analysis (which we have not been able to find in the literature in the case of irregular h) may presumably give error estimates only in powers of η (which equals 1 here!) and not as $O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}})$. Additionally, due to the form of the proxy, our intuition is that the error at $(0,x_0)$ (where we aim at computing the solution) is smaller that the error at arbitrary (t,x). All these reasons indicate that a PDE approach to derive correction terms and error analysis in our proxy setting is probably irrelevant.

2.3 Higher-order proxy approximation

In this section, we give several expansions formulas with a third order accuracy. First, we recall without proof results obtained in [Benhamou 2010a] and [Gobet 2012b] for expansions based on local volatility at spot and at strike. Second we introduce a new expansion with local volatility frozen at mid-point. Finally new expansions of implied volatility are provided.

2.3.1 Third order approximation with the local volatility at spot and at strike.

We define some integral operators useful to state the next theorems.

Definition 2.3.1.1. *If the derivatives and the integrals have a meaning, we define for a two variables function l the above operators:*

$$C_1(l;z)_0^T = \omega(l^2(z), l(z)l^{(1)}(z))_0^T, \qquad C_2(l;z)_0^T = \omega(l^2(z), (l^{(1)}(z))^2 + l(z)l^{(2)}(z))_0^T,$$
$$C_3(l;z)_0^T = \omega(l^2(z), l^2(z), (l^{(1)}(z))^2 + l(z)l^{(2)}(z))_0^T, \qquad C_4(l;z)_0^T = \omega(l^2(z), l(z)l^{(1)}(z), l(z)l^{(1)}(z))_0^T.$$

We frequently use some linear combinations of these operators:

$$\eta_1(l;z)_0^T = \frac{C_1(l;z)_0^T}{2} - \frac{C_2(l;z)_0^T}{2} - \frac{C_3(l;z)_0^T}{4} - \frac{C_4(l;z)_0^T}{2},$$
$$\eta_2(l;z)_0^T = -\frac{3C_1(l;z)_0^T}{2} + \frac{C_2(l;z)_0^T}{2} + \frac{5C_3(l;z)_0^T}{4} + \frac{7C_4(l;z)_0^T}{2} + \frac{[C_1(l;z)_0^T]^2}{8},$$
$$\eta_3(l;z)_0^T = C_1(l;z)_0^T - 2C_3(l;z)_0^T - 6C_4(l;z)_0^T - \frac{3[C_1(l;z)_0^T]^2}{4},$$
$$\eta_4(l;z)_0^T = C_3(l;z)_0^T + 3C_4(l;z)_0^T + \frac{13[C_1(l;z)_0^T]^2}{8},$$
$$\eta_5(l;z)_0^T = -\frac{3[C_1(l;z)_0^T]^2}{2}, \qquad \eta_6(l;z)_0^T = \frac{[C_1(l;z)_0^T]^2}{2},$$
$$\zeta_2(l;z)_0^T = \frac{C_2(l;z)_0^T}{2}, \qquad \zeta_3(l;z)_0^T = C_1(l;z)_0^T,$$
$$\zeta_4(l;z)_0^T = C_3(l;z)_0^T + 3C_4(l;z)_0^T, \qquad \zeta_6(l;z)_0^T = \frac{[C_1(l;z)_0^T]^2}{2}.$$

Theorem 2.3.1.1. ***(3rd order approximations for Call options with the local volatility at spot)****. Assuming ($\mathcal{H}^a$) and using the log-normal proxy, one has*

$$\mathrm{Call}(e^{x_0}, T, e^k) = \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_0}^2 T, k) + \sum_{i=1}^{6} \eta_i(a;x_0)_0^T \partial_{x^i}^i \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_0}^2 T, k) + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3 T^2).$$

Assuming ($\mathcal{H}^\Sigma$) and using the normal proxy, one has

$$\mathrm{Call}(S_0, T, K) = \mathrm{Call}^{\mathrm{BA}}(S_0, \bar{\Sigma}_{S_0}^2 T, K) + \sum_{i\in\{2,3,4,6\}} \zeta_i(\Sigma; S_0)_0^T \partial_{S^i}^i \mathrm{Call}^{\mathrm{BA}}(S_0, \bar{\Sigma}_{S_0}^2 T, K) + O(\mathcal{M}_1(\Sigma)[\mathcal{M}_0(\Sigma)]^3 T^2).$$

The operators ζ_i and η_i in the above expansions are defined in Definition 2.3.1.1.

The magnitude of the residual terms in the previous formulas justifies the label of third order approximations.

The above theorem is a straightforward application of [Benhamou 2010a, Theorems 2.2, 2.3 and 4.2], taking into account that we slightly modify the notations of the Greek coefficients. Namely, for convenience we merge certain ω operators: for instance the reader can easily check that:

$$[C_1(l;z)_0^T]^2 = [\omega(l(z)^2, l(z)l^{(1)}(z))_0^T]^2 = 4\omega(l(z)^2, l(z)^2, l(z)l^{(1)}(z), l(z)l^{(1)}(z))_0^T \\ + 2\omega(l(z)^2, l(z)l^{(1)}(z), l(z)^2, l(z)l^{(1)}(z))_0^T.$$

We should mention that it seems possible to relax the strong hypothesis $(\mathcal{H}^a)$ which appears in [Benhamou 2010a, Theorems 2.2 and 4.2]. As for the second order approximations, $(\mathcal{H}^a_{x_0})$ may be sufficient.

Using the duality argument introduced in Subsection 2.1.3 and [Benhamou 2010a, Theorems 2.2, 2.3 and 4.2], approximations using the volatility at strike are available too.

Theorem 2.3.1.2. ***(3rd order approximations for Call options with the local volatility at strike).*** *Assuming $(\mathcal{H}^a)$ and using the log-normal proxy, one has*

$$\mathrm{Call}(e^{x_0}, T, e^k) = \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_k^2 T, k) + \sum_{i=1}^{6} \eta_i(\widetilde{a}; k)_0^T \partial^i_{z^i} \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_k^2 T, k) + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3 T^2),$$

Assuming $(\mathcal{H}^\Sigma)$and using the normal proxy, one has

$$\mathrm{Call}(S_0, T, K) = \mathrm{Call}^{\mathrm{BA}}(S_0, \bar{\Sigma}_K^2 T, K) + \sum_{i \in \{2,3,4,6\}} \zeta_i(\widetilde{\Sigma}; K)_0^T \partial^i_{Z^i} \mathrm{Call}^{\mathrm{BA}}(S_0, \bar{\Sigma}_K^2 T, K) + O(\mathcal{M}_1(\Sigma)[\mathcal{M}_0(\Sigma)]^3 T^2).$$

The operators ζ_i and η_i in the above expansions are defined in Definition 2.3.1.1.

2.3.2 Third order approximation with the local volatility at mid-point.

We now state a new result related to third order expansions based on the local volatility at mid-point x_{avg} or S_{avg}. For a clearer proof, we change the presentation of the corrective terms in comparison with Theorems 2.3.1.1 and 2.3.1.2: instead of gathering them according the order of the Greeks, we put them together according to the operators C_i introduced in Definition 2.3.1.1.

Theorem 2.3.2.1. ***(3rd order approximations for Call options with the local volatility at mid-point).*** *Assuming $(\mathcal{H}^a)$ and using the log-normal proxy, one has*

$$\begin{aligned}
\mathrm{Call}(e^{x_0}, T, e^k) =& \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) + \frac{C_1(a; x_{avg})_0^T - C_1(\widetilde{a}; x_{avg})_0^T}{2}(\partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) \\
&+ \frac{C_2(a; x_{avg})_0^T + C_2(\widetilde{a}; x_{avg})_0^T}{2}(\frac{1}{2}\partial^2_{x^2} - \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) \\
&+ \frac{C_3(a; x_{avg})_0^T + C_3(\widetilde{a}; x_{avg})_0^T}{2}(\partial^4_{x^4} - 2\partial^3_{x^3} + \frac{5}{4}\partial^2_{x^2} - \frac{1}{4}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) \\
&+ \frac{C_4(a; x_{avg})_0^T + C_4(\widetilde{a}; x_{avg})_0^T}{2}(3\partial^4_{x^4} - 6\partial^3_{x^3} + \frac{7}{2}\partial^2_{x^2} - \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) \\
&+ \frac{[C_1(a; x_{avg})_0^T]^2 + [C_1(\widetilde{a}; x_{avg})_0^T]^2}{2}(\frac{1}{2}\partial^6_{x^6} - \frac{3}{2}\partial^5_{x^5} + \frac{13}{8}\partial^4_{x^4} - \frac{3}{4}\partial^3_{x^3} + \frac{1}{8}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) \\
&- (x_0 - k)^2 C_5(a; x_{avg})_0^T (\frac{1}{8}\partial^2_{x^2} - \frac{1}{8}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) \\
&- (x_0 - k)^2 C_6(a; x_{avg})_0^T (\frac{1}{4}\partial^4_{x^4} - \frac{1}{2}\partial^3_{x^3} + \frac{1}{4}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}^2 T, k) \\
&+ O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3 T^2),
\end{aligned} \tag{2.24}$$

where the operators C_i for $i = 1\ldots 4$ are defined in Definition 2.3.1.1 and where the time reversal invariant[2] *operators C_5 and C_6 are defined by:*

$$C_5(l;z)_0^T = \omega((l^{(1)}(z))^2 + l(z)l^{(2)}(z))_0^T, \qquad C_6(l;z)_0^T = \omega(l(z)l^{(1)}(z), l(z)l^{(1)}(z))_0^T.$$

Assuming $(\mathcal{H}^\Sigma)$and using the normal proxy, one has

$$\begin{aligned}
\mathrm{Call}(S_0,T,K) =& \mathrm{Call}^{\mathrm{BA}}(S_0,\bar{\Sigma}^2_{S_{avg}}T,K) + \frac{C_1(\Sigma;S_{avg})_0^T - C_1(\widetilde{\Sigma};S_{avg})_0^T}{2}\partial^3_{S^3}\mathrm{Call}^{\mathrm{BA}}(S_0,\bar{\Sigma}^2_{S_{avg}}T,K)\\
&+\frac{C_2(\Sigma;S_{avg})_0^T + C_2(\widetilde{\Sigma};S_{avg})_0^T}{4}\partial^2_{S^2}\mathrm{Call}^{\mathrm{BA}}(S_0,\bar{\Sigma}^2_{S_{avg}}T,K)\\
&+\frac{C_3(\Sigma;S_{avg})_0^T + C_3(\widetilde{\Sigma};S_{avg})_0^T}{2}\partial^4_{S^4}(S_0,\bar{\Sigma}^2_{S_{avg}}T,K)\\
&+3\frac{C_4(\Sigma;S_{avg})_0^T + C_4(\widetilde{\Sigma};S_{avg})_0^T}{2}\partial^4_{S^4}(S_0,\bar{\Sigma}^2_{S_{avg}}T,K)\\
&+\frac{[C_1(\Sigma;S_{avg})_0^T]^2 + [C_1(\widetilde{\Sigma};S_{avg})_0^T]^2}{4}\partial^6_{S^6}(S_0,\bar{\Sigma}^2_{S_{avg}}T,K)\\
&-(x_0-k)^2\frac{C_5(\Sigma;S_{avg})_0^T}{8}\partial^2_{S^2}\mathrm{Call}^{\mathrm{BA}}(S_0,\bar{\Sigma}^2_{S_{avg}}T,K)\\
&-(x_0-k)^2\frac{C_6(\Sigma;S_{avg})_0^T}{4}\partial^4_{S^4}(S_0,\bar{\Sigma}^2_{S_{avg}}T,K)\\
&+O(\mathcal{M}_1(\Sigma)[\mathcal{M}_0(\Sigma)]^3T^2).
\end{aligned}$$

Proof. We only prove the result for the log-normal proxy. The case of normal proxy is similar, and it is left to the reader as an exercise. The idea is again to consider the average of the third order formulas in spot and strike provided in Theorems 2.3.1.1 and 2.3.1.2 and to perform an expansion around the mid-point.

▷ **Step 1: expansion of the leading term.** Firstly we aim at showing that:

$$\begin{aligned}
&(\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k) + \mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_kT,k))/2\\
=&\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) + \frac{(x_0-k)^2}{4}C_5(a;x_{avg})_0^T\partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\\
&+(x_0-k)^2C_6(a;x_{avg})_0^T\partial^2_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2),
\end{aligned}\tag{2.25}$$

where the operators C_5 and C_6 are defined in Theorem 2.3.2.1. Perform Taylor expansions to obtain:

$$\begin{aligned}
\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k) =& \mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) + \partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)T(\bar{a}^2_{x_0} - \bar{a}^2_{x_{avg}})\\
&+\frac{1}{2}\partial_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)T^2(\bar{a}^2_{x_0} - \bar{a}^2_{x_{avg}})^2 + R_1\\
=&\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) + \partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\omega(a(x_{avg})a^{(1)}(x_{avg}))_0^T(x_0-k)\\
&+\frac{1}{4}\partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)C_5(a;x_{avg})_0^T(x_0-k)^2\\
&+\partial^2_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)C_6(a;x_{avg})_0^T(x_0-k)^2 + R_1 + R_2 + R_3,
\end{aligned}\tag{2.26}$$

where we have used the relations $\partial_z\omega(l^2(z))_0^T = 2\omega(l(z)l^{(1)}(z))_0^T$, $\partial^2_{z^2}\omega(l^2(z))_0^T = 2C_5(l;z)_0^T$, $[\omega(l(z)l^{(1)}(z))_0^T]^2 = 2C_6(l;z)_0^T$ and where R_1, R_2 and R_3 are defined by:

$$R_1 = T^3(\bar{a}^2_{x_0} - \bar{a}^2_{x_{avg}})^3\int_0^1(\partial^3_{y^3}\mathrm{Call}^{\mathrm{BS}}(x_0,y,k))|_{y=T(\lambda\bar{a}^2_{x_0}+(1-\lambda)\bar{a}^2_{x_{avg}})}\frac{(1-\lambda)^2}{2}\mathrm{d}\lambda,$$

[2] that is $C_5(\tilde{l};z)_0^T = C_5(l;z)_0^T$ and $C_6(\tilde{l};z)_0^T = C_6(l;z)_0^T$ using (2.1).

$$
\begin{aligned}
R_2 =&\partial_y \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_{avg}}T, k)\frac{(x_0-k)^3}{8}\int_0^1 (\partial^3_x \omega(a^2(x))^T_0)|_{x=\lambda x_0+(1-\lambda)x_{avg}}\frac{(1-\lambda)^2}{2}\mathrm{d}\lambda,\\
R_3 =&\frac{1}{2}\partial^2_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_{avg}}T, k)\frac{(x_0-k)^2}{4}\int_0^1 (\partial^2_x \omega(a^2(x))^T_0)_{x=\lambda x_0+(1-\lambda)x_{avg}}(1-\lambda)\mathrm{d}\lambda\\
&\times\Big[\frac{(x_0-k)^2}{4}\int_0^1 (\partial^2_x \omega(a^2(x))^T_0)_{x=\lambda x_0+(1-\lambda)x_{avg}}(1-\lambda)\mathrm{d}\lambda + 2\omega(a(x_{avg})a^{(1)}(x_{avg}))^T_0(x_0-k)\Big].
\end{aligned}
$$

Using (1.8) in Chapter 1 Section 1.2, Corollary 2.6.1.1 and $(\mathcal{H}^a)$ we obtain

$$
|R_1+R_2+R_3| \le_c \mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2.
$$

Similarly, we show that:

$$
\begin{aligned}
\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_k T, k) =&\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}_{x_{avg}}, k) - \partial_y \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_{avg}}T, k)\omega(a(x_{avg})a^{(1)}(x_{avg}))^T_0(x_0-k)\\
&+\frac{1}{4}\partial_y \mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_{avg}}T, k)C_5(a; x_{avg})^T_0(x_0-k)^2\\
&+\partial^2_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_{avg}}T, k)C_6(a; x_{avg})^T_0(x_0-k)^2 + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2).
\end{aligned}
$$

Combine this with (2.26) to obtain (2.25).

▷ **Step 2: expansion of the corrective terms.** Firstly we treat the corrective terms with the operators C_2, C_3, C_4 and $[C_1]^2$ in Theorems 2.3.1.1 and 2.3.1.2. We let the reader verify that in the formula with volatility at spot (respectively in strike), we can replace the point x_0 (respectively k) by the point x_{avg} in all the corrective terms involving these operators: indeed it induces an extra error of order $\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2$. This is very similar to the proof of Lemma 2.1.3.1 so we skip it. Then we can replace derivatives w.r.t. z with derivatives w.r.t. x in $\mathrm{Call}^{\mathrm{BS}}$ thanks to Proposition 2.6.1.3, equations (2.49)-(2.51)-(2.52)-(2.53). That leads to:

$$
\begin{aligned}
&\frac{1}{2}C_2(a; x_0)^T_0(\frac{1}{2}\partial^2_{x^2}-\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_0}T, k) + \frac{1}{2}C_2(\overline{a}; k)^T_0(\frac{1}{2}\partial^2_{z^2}-\frac{1}{2}\partial_z)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_k T, k) \qquad (2.27)\\
&+\frac{1}{2}C_3(a; x_0)^T_0(\partial^4_{x^4}-2\partial^3_{x^3}+\frac{5}{4}\partial^2_{x^2}-\frac{1}{4}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_0}T, k)\\
&+\frac{1}{2}C_3(\overline{a}; k)^T_0(\partial^4_{z^4}-2\partial^3_{z^3}+\frac{5}{4}\partial^2_{z^2}-\frac{1}{4}\partial_z)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_k T, k)\\
&+\frac{1}{2}C_4(a; x_0)^T_0(3\partial^4_{x^4}-6\partial^3_{x^3}+\frac{7}{2}\partial^2_{x^2}-\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_0}T, k)\\
&+\frac{1}{2}C_4(\overline{a}; k)^T_0(3\partial^4_{z^4}-6\partial^3_{z^3}+\frac{7}{2}\partial^2_{z^2}-\frac{1}{2}\partial_z)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_k T, k)\\
&+\frac{1}{2}[C_1(a; x_0)^T_0]^2(\frac{1}{2}\partial^6_{x^6}-\frac{3}{2}\partial^5_{x^5}+\frac{13}{8}\partial^4_{x^4}-\frac{3}{4}\partial^3_{x^3}+\frac{1}{8}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_0}T, k)\\
&+\frac{1}{2}[C_1(\overline{a}; k)^T_0]^2(\frac{1}{2}\partial^6_{z^6}-\frac{3}{2}\partial^5_{z^5}+\frac{13}{8}\partial^4_{z^4}-\frac{3}{4}\partial^3_{z^3}+\frac{1}{8}\partial^2_{z^2})\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_k T, k)\\
=&\frac{C_2(a; x_{avg})^T_0 + C_2(\overline{a}; x_{avg})^T_0}{2}(\frac{1}{2}\partial^2_{x^2}-\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_{avg}}T, k)\\
&+\frac{C_3(a; x_{avg})^T_0 + C_3(\overline{a}; x_{avg})^T_0}{2}(\partial^4_{x^4}-2\partial^3_{x^3}+\frac{5}{4}\partial^2_{x^2}-\frac{1}{4}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_{avg}}T, k)\\
&+\frac{C_4(a; x_{avg})^T_0 + C_4(\overline{a}; x_{avg})^T_0}{2}(3\partial^4_{x^4}-6\partial^3_{x^3}+\frac{7}{2}\partial^2_{x^2}-\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_{avg}}T, k)\\
&+\frac{[C_1(a; x_{avg})^T_0]^2 + [C_1(\overline{a}; x_{avg})^T_0]^2}{2}(\frac{1}{2}\partial^6_{x^6}-\frac{3}{2}\partial^5_{x^5}+\frac{13}{8}\partial^4_{x^4}-\frac{3}{4}\partial^3_{x^3}+\frac{1}{8}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{a}^2_{x_{avg}}T, k)\\
&+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2).
\end{aligned}
$$

Secondly, we pass to the corrective terms in which appears the operator C_1. For the sake of clarity, we introduce the following notation $\mathcal{A}_x = \partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x$ and $\mathcal{A}_z = \partial^3_{z^3} - \frac{3}{2}\partial^2_{z^2} + \frac{1}{2}\partial_z$. For example $\partial_y\mathcal{A}_x$ stands for the differential operator $\partial^4_{yx^3} - \frac{3}{2}\partial^3_{yx^2} + \frac{1}{2}\partial^2_{yx}$ and similarly for $\partial^2_{y^2}\mathcal{A}_x$. We recall the following relation $\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}} = -\mathcal{A}_z\mathrm{Call}^{\mathrm{BS}}$ (see (2.49) in Proposition 2.6.1.3). Our purpose is to prove that:

$$\begin{aligned}
&\frac{1}{2}C_1(a;x_0)_0^T\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k) + \frac{1}{2}C_1(\widetilde{a};k)_0^T\mathcal{A}_z\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{k}T,k) \\
=&\frac{1}{2}(C_1(a;x_{avg})_0^T - C_1(\widetilde{a};x_{avg})_0^T)\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) \\
&+\frac{(x_0-k)}{4}[4C_6(a;x_{avg})_0^T + C_2(a;x_{avg})_0^T + C_2(\widetilde{a};x_{avg})_0^T]\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) \\
&+[C_4(a;x_{avg})_0^T + C_4(\widetilde{a};x_{avg})_0^T + \omega(a(x_{avg})a^{(1)}(x_{avg}),a^2(x_{avg}),a(x_{avg})a^{(1)}(x_{avg}))_0^T] \\
&\quad\times(x_0-k)\partial_y\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2).
\end{aligned}\tag{2.28}$$

Perform a second order Taylor expansion for the function $y \mapsto \mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,y,k)$ at $y = \bar{a}^2_{x_0}T = \omega(a^2(x_0))_0^T$ around $y = \bar{a}^2_{x_{avg}}T = \omega(a^2(x_{avg}))_0^T$ and for the function $x \mapsto C_1(a;x)_0^T$ at $x = x_0$ around $x = x_{avg}$:

$$\begin{aligned}
&C_1(a;x_0)_0^T\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k) \\
=&\{C_1(a;x_{avg})_0^T + \partial_x(C_1(a;x)_0^T)|_{x=x_{avg}}\frac{(x_0-k)}{2} + R_1\} \\
&\times\{\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) + \partial_y\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)T(\bar{a}^2_{x_0} - \bar{a}^2_{x_{avg}}) + R_2\} \\
=&\{C_1(a;x_{avg})_0^T + \partial_x(C_1(a;x)_0^T)|_{x=x_{avg}}\frac{(x_0-k)}{2} + R_1\} \\
&\times\{\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) + R_3 + R_2 + \partial_y\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\omega(a(x_{avg})a^{(1)}(x_{avg}))_0^T(x_0-k)\} \\
=&C_1(a;x_{avg})_0^T\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) \\
&+\partial_x(C_1(a;x)_0^T)|_{x=x_{avg}}\frac{(x_0-k)}{2}\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) \\
&+C_1(a;x_{avg})_0^T\partial_y\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\omega(a(x_{avg})a^{(1)}(x_{avg}))_0^T(x_0-k) \\
&+R,
\end{aligned}\tag{2.29}$$

where:

$$\begin{aligned}
R =&C_1(a;x_0)_0^T[R_3+R_2] + R_1\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) \\
&+(x_0-k)\partial_y\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\omega(a(x_{avg})a^{(1)}(x_{avg}))_0^T[R_1 + \partial_x(C_1(a;x)_0^T)|_{x=x_{avg}}\frac{(x_0-k)}{2}], \\
R_1 =&\frac{(x_0-k)^2}{4}\int_0^1(\partial^2_{x^2}(C_1(a;x)_0^T))|_{x=\lambda x_0+(1-\lambda)x_{avg}}(1-\lambda)\mathrm{d}\lambda, \\
R_2 =&T^2(\bar{a}^2_{x_0} - \bar{a}_{x_{avg}})^2\int_0^1(\partial^2_{y^2}\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,y,k))|_{y=T(\lambda\bar{a}^2_{x_0}+(1-\lambda)\bar{a}_{x_{avg}})}(1-\lambda)\mathrm{d}\lambda, \\
R_3 =&\partial_y\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\frac{(x_0-k)^2}{4}\int_0^1(\partial^2_{x^2}(\omega(a^2(x))_0^T))|_{x=\lambda x_0+(1-\lambda)x_{avg}}(1-\lambda)\mathrm{d}\lambda.
\end{aligned}$$

On the one hand, we have:

$$\begin{aligned}
\partial_z(C_1(l;z)_0^T) &= 2C_6(l;z)_0^T + C_2(l;z)_0^T, \\
C_1(l;z)_0^T\omega(l(z)l^{(1)}(z))_0^T &= 2C_4(l;z)_0^T + \omega(l(z)l^{(1)}(z),l^2(z),l(z)l^{(1)}(z))_0^T,
\end{aligned}$$

and on the other hand, with (1.8) in Chapter 1 Section 1.2, Corollary 2.6.1.1 and $(\mathcal{H}^a)$, it comes:

$$|R| \leq_c \mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2.$$

We skip further details. Consequently we can write (2.29) as follows:

$$\begin{aligned}&C_1(a;x_0)_0^T\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k) \qquad (2.30)\\
=&C_1(a;x_{avg})_0^T\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)+\frac{(x_0-k)}{2}[2C_6(a;x_{avg})_0^T+C_2(a;x_{avg})_0^T]\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\\
&+[2C_4(a;x_{avg})_0^T+\omega(a(x_{avg})a^{(1)}(x_{avg}),a^2(x_{avg}),a(x_{avg})a^{(1)}(x_{avg}))_0^T](x_0-k)\partial_y\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\\
&+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2).\end{aligned}$$

Then using the relation $\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}=-\mathcal{A}_z\mathrm{Call}^{\mathrm{BS}}$, the time reversal invariance of $l\mapsto C_6(l,z)_0^T$ and $l\mapsto\omega(l(z)l^{(1)}(z),l^2(z),l(z)l^{(1)}(z))_0^T$ (for any z), one obtains similarly:

$$\begin{aligned}&C_1(\overline{a};k)_0^T\mathcal{A}_z\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k) \qquad (2.31)\\
=&-C_1(\overline{a};x_{avg})_0^T\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)+\frac{(x_0-k)}{2}[2C_6(a;x_{avg})_0^T+C_2(\overline{a};x_{avg})_0^T]\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\\
&+[2C_4(\overline{a};x_{avg})_0^T+\omega(a(x_{avg})a^{(1)}(x_{avg}),a^2(x_{avg}),a(x_{avg})a^{(1)}(x_{avg}))_0^T](x_0-k)\partial_y\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\\
&+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2).\end{aligned}$$

Compute the average of (2.30) and (2.31) to complete the proof of (2.28).

▹ **Step 3: mathematical reductions.** We gather terms coming from (2.25) and (2.28). In view of (1.8) in Chapter 1 Section 1.2 and equations (2.48) and (2.49) in Proposition 2.6.1.3, we have:

$$\begin{aligned}&\frac{(x_0-k)^2}{4}C_5(a;x_{avg})_0^T\partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)+\frac{(x_0-k)}{4}[C_2(a;x_{avg})_0^T+C_2(\overline{a};x_{avg})_0^T]\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\\
=&\frac{(x_0-k)^2}{4}\partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\Big(C_5(a;x_{avg})_0^T-2\frac{[C_2(a;x_{avg})_0^T+C_2(\overline{a};x_{avg})_0^T]}{\omega(a^2(x_{avg}))_0^T}\Big)\\
=&\frac{(x_0-k)^2}{4}\partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)(C_5(a;x_{avg})_0^T-2C_5(a;x_{avg})_0^T)\\
=&-\frac{(x_0-k)^2}{8}C_5(a;x_{avg})_0^T(\partial^2_{x^2}-\partial_x)\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k), \qquad (2.32)\end{aligned}$$

where we have used at the second equality the relation $C_5(l;z)_0^T\omega(l^2(z))_0^T=C_2(l;z)_0^T+C_2(\overline{l};z)_0^T$ obtained easily with (2.1). Then (1.8) in Chapter 1 Section 1.2, (2.48) and (2.49) yield

$$\begin{aligned}\partial_y\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)=&\partial_y((-\frac{2(x_0-k)}{y}\partial_y)\mathrm{Call}^{\mathrm{BS}}(x_0,y,k))|_{y=\bar{a}^2_{x_{avg}}T}\\
=&\frac{2(x_0-k)}{\omega(a^2(x_{avg}))_0^T}\Big[\frac{\partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)}{\omega(a^2(x_{avg}))_0^T}-\partial^2_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\Big],\end{aligned}$$

and straightforward calculus allows to obtain with (2.1):

$$C_6(l;z)_0^T\omega(l^2(z))_0^T=C_4(l;z)_0^T+C_4(\overline{l};z)_0^T+\omega(l(z)l^{(1)}(z),l^2(z),l(z)l^{(1)}(z))_0^T.$$

These two intermediate results give:

$$\begin{aligned}&(x_0-k)C_6(a;x_{avg})_0^T[\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)+(x_0-k)\partial^2_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)]\\
&+[C_4(a;x_{avg})_0^T+C_4(\overline{a};x_{avg})_0^T+\omega(a(x_{avg})a^{(1)}(x_{avg}),a^2(x_{avg}),a(x_{avg})a^{(1)}(x_{avg}))_0^T]\\
&\times(x_0-k)\partial_y\mathcal{A}_x\mathrm{Call}^{\mathrm{BS}}(x_0,\omega(a^2(x_{avg}))_0^T,k)\\
=&-2\frac{(x_0-k)^2}{\omega(a^2(x_{avg}))_0^T}C_6(a;x_{avg})_0^T\partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\end{aligned}$$

$$
\begin{aligned}
&+(x_0-k)^2 C_6(a;x_{avg})_0^T \partial^2_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\\
&+(x_0-k)[C_6(a;x_{avg})_0^T\omega(a^2(x_{avg}))_0^T]\frac{2(x_0-k)}{\omega(a^2(x_{avg}))_0^T}\\
&\times[\frac{\partial_y\mathrm{Call}^{\mathrm{BS}}(x_0,\omega(a^2(x_{avg}))_0^T,k)}{\omega(a^2(x_{avg}))_0^T}-\partial^2_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0,\omega(a^2(x_{avg}))_0^T,k)]\\
=&-(x_0-k)^2 C_6(a;x_{avg})_0^T \partial^2_{y^2}\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\\
=&-\frac{(x_0-k)^2}{4}C_6(a;x_{avg})_0^T(\partial^4_{x^4}-2\partial^3_{x^3}+\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k).
\end{aligned}
\tag{2.33}
$$

Finally, sum the relations (2.25-2.28-2.27) taking into account the simplifications (2.32-2.33) and apply Theorems 2.3.1.1 and 2.3.1.2 to obtain the announced result (2.24). □

2.3.3 Third order expansion of the implied volatility

We define extra integral operators in order to state a new result about third order expansions of the implied volatility.

Definition 2.3.3.1. *Provided that the derivatives and the integrals below have a meaning, we define the following operators for a two variables non-negative function l such that $\bar{l}_z > 0$:*

$$
\begin{aligned}
\gamma_0(l;z)_0^T =&\bar{l}_z+\frac{C_2(l;z)_0^T}{2\bar{l}_zT}-\frac{C_4(l;z)_0^T}{4\bar{l}_zT}-\frac{C_3(l;z)_0^T}{\bar{l}_z^3T^2}-\frac{3C_4(l;z)_0^T}{\bar{l}_z^3T^2}+\frac{[C_1(l;z)_0^T]^2}{8\bar{l}_z^3T^2}\\
&+\frac{3[C_1(l;z)_0^T]^2}{2\bar{l}_z^5T^3},\\
\gamma_1(l;z)_0^T =&\frac{C_1(l;z)_0^T}{\bar{l}_z^3T^2},\qquad \gamma_2(l;z)_0^T=\frac{C_3(l;z)_0^T}{\bar{l}_z^5T^3}+3\frac{C_4(l;z)_0^T}{\bar{l}_z^5T^3}-\frac{3[C_1(l;z)_0^T]^2}{\bar{l}_z^7T^4};\\
\pi_0(l;z)_0^T =&\frac{\gamma_0(l;z)_0^T+\gamma_0(\widetilde{l};z)_0^T}{2},\qquad \pi_1(l;z)_0^T=\frac{\gamma_1(\widetilde{l};z)_0^T-\gamma_1(l;z)_0^T}{2},\\
\pi_2(l;z)_0^T =&\frac{\gamma_2(l;z)_0^T+\gamma_2(\widetilde{l};z)_0^T}{2}-\frac{C_5(l;z)_0^T}{8\bar{l}_zT}+\frac{C_6(l;z)_0^T}{4\bar{l}_z^3T^2};\\
\chi_0(l;z)_0^T =&\bar{l}_z+\frac{C_2(l;z)_0^T}{2\bar{l}_zT}-\frac{C_3(l;z)_0^T}{\bar{l}_z^3T^2}-\frac{3C_4(l;z)_0^T}{\bar{l}_z^3T^2}+\frac{3[C_1(l;z)_0^T]^2}{2\bar{l}_z^5T^3},\\
\chi_1(l;z)_0^T =&\gamma_1(l;z)_0^T,\qquad \chi_2(l;z)_0^T=\gamma_2(l;z)_0^T;\\
\Xi_0(l;z)_0^T =&\frac{\chi_0(l;z)_0^T+\chi_0(\widetilde{l};z)_0^T}{2},\qquad \Xi_1(l;z)_0^T=\pi_1(l;z)_0^T,\qquad \Xi_2(l;z)_0^T=\pi_2(l;z)_0^T.
\end{aligned}
$$

Theorem 2.3.3.1. *(**3rd order expansions of the implied volatility**). Assume ($\mathcal{H}^a$). We have:*

$$\sigma_{\mathrm{I}}(x_0,T,k)=\gamma_0(a;x_0)_0^T-\gamma_1(a;x_0)_0^Tm+\gamma_2(a;x_0)_0^Tm^2+\mathrm{Error}^{\mathrm{I}}_{3,\mathrm{x}_0},\tag{2.34}$$

$$\sigma_{\mathrm{I}}(x_0,T,k)=\gamma_0(\widetilde{a};k)_0^T+\gamma_1(\widetilde{a};k)_0^Tm+\gamma_2(\widetilde{a};k)_0^Tm^2+\mathrm{Error}^{\mathrm{I}}_{3,\mathrm{k}},\tag{2.35}$$

$$\sigma_{\mathrm{I}}(x_0,T,k)=\pi_0(a;x_{avg})_0^T+\pi_1(a;x_{avg})_0^Tm+\pi_2(a;x_{avg})_0^Tm^2+\mathrm{Error}^{\mathrm{I}}_{3,\mathrm{x_{avg}}}.\tag{2.36}$$

Under ($\mathcal{H}^\Sigma$) we have

$$\Sigma_{\mathrm{I}}(S_0,T,K)=\chi_0(\Sigma;S_0)_0^T-\chi_1(\Sigma;S_0)_0^TM+\chi_2(\Sigma;S_0)_0^TM^2+\mathrm{Error}^{\mathrm{I}}_{3,\mathrm{S}_0},\tag{2.37}$$

$$\Sigma_{\mathrm{I}}(S_0,T,K) = \chi_0(\widetilde{\Sigma};K)_0^T + \chi_1(\widetilde{\Sigma};K)_0^T M + \chi_2(\widetilde{\Sigma};K)_0^T M^2 + \mathrm{Error}^{\mathrm{I}}_{3,\mathrm{K}}, \tag{2.38}$$

$$\Sigma_{\mathrm{I}}(S_0,T,K) = \Xi_0(\Sigma;S_{avg})_0^T + \Xi_1(\Sigma;S_{avg})_0^T M + \Xi_2(\Sigma;S_{avg})_0^T M^2 + \mathrm{Error}^{\mathrm{I}}_{3,\mathrm{S_{avg}}}. \tag{2.39}$$

The operators γ_i, π_i, χ_i and Ξ_i used in the above expansions are defined in Definition 2.3.3.1.

Remark 2.3.3.1. *We have obtained Black-Scholes (respectively Bachelier) implied volatility approximations which are written as a quadratic function w.r.t. the Black-Scholes log-moneyness (respectively w.r.t. the Bachelier moneyness). At the money, observe that the corresponding approximations are not equal to the local volatility function computed at spot. However, in view of the definition of the operators C_1, C_2, C_3 and C_4 (see Definition 2.3.1.1) and the operators γ_0 and χ_0 (see Definition 2.3.3.1), we easily obtain the estimate:*

$$|\gamma_0(a;x_0)_0^T - \bar{a}_{x_0}| + |\gamma_0(\widetilde{a};x_0)_0^T - \bar{a}_{x_0}| + |\chi_0(\Sigma;S_0) - \bar{\Sigma}_{S_0}| + |\chi_0(\widetilde{\Sigma};S_0) - \bar{\Sigma}_{S_0}| \leq_c T.$$

It shows that when the maturity tends to zero, our implied volatility approximations at the money become equal to the local volatility function frozen at spot. We therefore interpret the difference between our implied volatility approximations ATM and the local volatility function frozen at spot as a maturity bias.

Proof. We first focus on the formula (2.34), the treatment of (2.35-2.37-2.38) being similar. Start from Theorem 2.3.1.1 and apply Proposition 2.6.1.3 in order to write the Greeks w.r.t. x (for each operator C_i) in terms of the $\mathrm{Vega}^{\mathrm{BS}}$ and the $\mathrm{Vomma}^{\mathrm{BS}}$. Thus the third order expansion formula based on the ATM local volatility with log-normal proxy can be transformed into:

$$\begin{aligned}
&\mathrm{Call}(e^{x_0},T,e^k)\\
=&\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k) + \mathrm{Vega}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k)\Big[-\frac{C_1(a;x_0)_0^T m}{\bar{a}^3_{x_0}T^2} + \frac{C_2(a;x_0)_0^T}{2\bar{a}_{x_0}T} + \frac{C_3(a;x_0)_0^T m^2}{\bar{a}^5_{x_0}T^3} - \frac{C_3(a;x_0)_0^T}{\bar{a}^3_{x_0}T^2}\\
&+\frac{3C_4(a;x_0)_0^T m^2}{\bar{a}^5_{x_0}T^3} - \frac{3C_4(a;x_0)_0^T}{\bar{a}^3_{x_0}T^2} - \frac{C_4(a;x_0)_0^T}{4\bar{a}_{x_0}T} + \frac{[C_1(a;x_0)_0^T]^2}{8\bar{a}^3_{x_0}T^2} + \frac{3[C_1(a;x_0)_0^T]^2}{2\bar{a}^5_{x_0}T^3} - \frac{3[C_1(a;x_0)_0^T]^2 m^2}{\bar{a}^7_{x_0}T^4}\Big]\\
&+\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k)\Big(\frac{C_1(a;x_0)_0^T m}{\bar{a}^3_{x_0}T^2}\Big)^2 + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2)\\
=&\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k) + \mathrm{Vega}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k)[\gamma_0(a;x_0)_0^T - \bar{a}_{x_0} - \gamma_1(a;x_0)_0^T m + \gamma_2(a;x_0)_0^T m^2]\\
&+\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_0}T,k)[\gamma_1(a;x_0)_0^T m]^2 + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2)\\
\approx&\mathrm{Call}^{\mathrm{BS}}(x_0,[\gamma_0(a;x_0)_0^T - \gamma_1(a;x_0)_0^T m + \gamma_2(a;x_0)_0^T m^2]^2T,k).
\end{aligned}$$

This reads as an expansion of the implied volatility and achieves the proof of (2.34).

Now we give the main lines of the derivation of the error estimate in (2.36), while (2.39) is left to the reader. Again, we apply Theorem 2.3.2.1 and Proposition 2.6.1.3 in order to replace the x_0-Greeks with the $\mathrm{Vega}^{\mathrm{BS}}$ and the $\mathrm{Vomma}^{\mathrm{BS}}$. One obtains similarly:

$$\begin{aligned}
&\mathrm{Call}(e^{x_0},T,e^k) \qquad (2.40)\\
=&\mathrm{Call}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k) + \mathrm{Vega}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)\Big[\frac{\gamma_0(a;x_{avg})_0^T + \gamma_0(\widetilde{a};x_{avg})_0^T}{2} - \bar{a}_{x_{avg}} + \frac{\gamma_1(\widetilde{a};x_{avg})_0^T - \gamma_1(a;x_{avg})_0^T}{2}m\\
&+\frac{\gamma_2(a;x_{avg})_0^T + \gamma_2(\widetilde{a};x_{avg})_0^T}{2}m^2 - \frac{C_5(a;x_{avg})_0^T}{8\bar{a}_{x_{avg}}T}m^2 + \frac{C_6(a;x_{avg})_0^T}{4\bar{a}^3_{x_{avg}}T^2}m^2 + \frac{C_6(a;x_{avg})_0^T}{16\bar{a}_{x_{avg}}T}m^2 - \frac{C_6(a;x_{avg})_0^T}{4\bar{a}^5_{x_{avg}}T^3}m^4\Big]\\
&+\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)m^2\Big(\frac{[\gamma_1(a;x_{avg})_0^T]^2 + [\gamma_1(\widetilde{a};x_{avg})_0^T]^2}{2}\Big)
\end{aligned}$$

$+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2)$.

Then write $\left(\frac{[\gamma_1(a;x_{avg})_0^T]^2+[\gamma_1(\widetilde{a};x_{avg})_0^T]^2}{2}\right) = \left(\frac{\gamma_1(\widetilde{a};x_{avg})_0^T-\gamma_1(a;x_{avg})_0^T}{2}\right)^2 + \left(\frac{\gamma_1(\widetilde{a};x_{avg})_0^T+\gamma_1(a;x_{avg})_0^T}{2}\right)^2$, use the fact that (see the definition (2.43) of $\mathrm{Vomma}^{\mathrm{BS}}$ and the definition of γ_1 in Definition 2.3.3.1)

$$\begin{aligned}&\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)m^2(\frac{\gamma_1(\widetilde{a};x_{avg})_0^T+\gamma_1(a;x_{avg})_0^T}{2})^2\\ =&\mathrm{Vega}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)(C_1(\widetilde{a};x_{avg})_0^T+C_1(a;x_{avg})_0^T)^2[-\frac{m^2}{32\bar{a}^5_{x_{avg}}T^3}+\frac{m^4}{8\bar{a}^9_{x_{avg}}T^5}],\end{aligned}$$

and finally, use the above identity (obtained with the definitions of C_1, C_6 and with the relation (2.1)):

$$\begin{aligned}(C_1(\widetilde{l};x)_0^T+C_1(l;x)_0^T)^2 =&2[\omega(l^2(z))_0^T]^2C_6(l;z)_0^T = 4\omega(l^2(z),l^2(z))_0^T\omega(l(z)l^{(1)}(z),l(z)l^{(1)}(z))_0^T\\ =&4[\omega(l(z)l^{(1)}(z),l(z)l^{(1)}(z),l^2(z),l^2(z))_0^T+\omega(l^2(z),l^2(z),l(z)l^{(1)}(z),l(z)l^{(1)}(z))_0^T\\ &+\omega(l^2(z),l(z)l^{(1)}(z),l(z)l^{(1)}(z),l^2(z))_0^T+\omega(l(z)l^{(1)}(z),l^2(z),l^2(z),l(z)l^{(1)}(z))_0^T\\ &+\omega(l^2(z),l(z)l^{(1)}(z),l^2(z),l(z)l^{(1)}(z))_0^T+\omega(l(z)l^{(1)}(z),l^2(z),l(z)l^{(1)}(z),l^2(z))_0^T],\end{aligned}$$

to cancel the terms $\frac{C_6(a;x_{avg})_0^T}{16\bar{a}_{x_{avg}}T}m^2$, $-\frac{C_6(a;x_{avg})_0^T}{4\bar{a}^5_{x_{avg}}T^3}m^4$ and $\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x_0,\bar{a}^2_{x_{avg}}T,k)m^2(\frac{\gamma_1(\widetilde{a};x_{avg})_0^T+\gamma_1(a;x_{avg})_0^T}{2})^2$ in (2.40). That achieves the proof of (2.36). □

In addition to these implied volatility expansions, one can under additional technical assumptions upper bound the residuals terms. For instance, let us consider (2.34), for which we can prove

$$\mathrm{Error}^{\mathrm{I}}_{3,x_0} = O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^{\frac{3}{2}}), \tag{2.41}$$

which justifies the label of third order expansion. This is available under the assumptions that $|m| \le \xi\mathcal{M}_0(a)\sqrt{T}$ (for a given $\xi \ge 0$) and that $\mathcal{M}_0(a)$, $\mathcal{M}_1(a)$ and T are globally small enough to ensure that the implied volatility approximation $\gamma_0(a;x_0)_0^T - \gamma_1(a;x_0)_0^Tm + \gamma_2(a;x_0)_0^Tm^2$ is bounded away from 0. The method of proof is analogous to that in Subsection 2.1.4, by performing a third order expansion of the Black-Scholes price w.r.t. the volatility, using the estimate (2.47) on $\mathrm{Ultima}^{\mathrm{BS}}$ (see Corollary 2.6.1.2), and carefully gathering terms and evaluating their magnitudes.

2.4 Approximation of the Delta

In this section, we investigate the approximation of the delta of the Call price, i.e. the derivative w.r.t. the spot, by deriving similar expansion formulas. For the sake of brevity we present only results using a log-normal proxy. The results are new.
To achieve this goal, we follow again the Dupire approach taking advantage of the symmetry between spot and strike. We start from the Feynman-Kac representation (2.11) which leads to a nice expression for the delta:

$$\delta(S_0,T,K) = \partial_{S_0}\mathbb{E}[(S_0-e^{k_T})_+] = \mathbb{P}(e^{k_T} < S_0) = \mathbb{P}(k_T < x_0).$$

Thus we are reduced to compute the price of a binary option on the fictitious asset $(k_t)_t$. This binary payoff is not anymore differentiable, but we can however apply directly [Benhamou 2010a, Theorems 2.1, 2.2 and 4.3] to obtain

Theorem 2.4.0.2. ***(1st and 2nd order approximations for delta using local volatility at strike)****. Assume* $(\mathcal{H}^a)$*. Then we have:*

$$\delta(e^{x_0},T,e^k) = \delta^{\mathrm{BS}}(x_0,\bar{a}_k^2T,k) + C_1(\widetilde{a};k)_0^T(\partial^3_{z^3} - \frac{3}{2}\partial^2_{z^2} + \frac{1}{2}\partial_z)\delta^{\mathrm{BS}}(x_0,\bar{a}_k^2T,k) + O(\mathcal{M}_1(a)\mathcal{M}_0(a)T),$$

$$\delta(e^{x_0},T,e^k) = \delta^{\mathrm{BS}}(x_0,\bar{a}_k^2T,k) + \sum_{i=1}^{6}\eta_i(\widetilde{a};k)_0^T\partial^i_{k^i}\delta^{\mathrm{BS}}(x_0,\bar{a}_k^2T,k) + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}),$$

where δ^{BS} *is Black-Scholes delta function defined by* $\delta^{\mathrm{BS}}(x,y,z) = \mathcal{N}(d_1(x,y,z))$*, with* x *the log-spot,* y *the total variance and* z *the log-strike.*

Remark 2.4.0.2. *In view of the error estimate, observe that the corresponding second and third order formulas for vanilla payoffs are respectively first and second order approximations for binary payoffs. This is due to the lack of regularity of the payoff (see our discussion in Chapter 1 Subsection 1.3.5).*

Like in the previous price approximation formulas, it is possible to perform additional Taylor expansions in order to obtain similar formulas using local volatility function frozen at spot or at mid-point. We announce two Lemmas which proof is very similar to those of Lemmas 2.1.3.1 and 2.1.3.2 and Theorem 2.3.2.1 and consequently is left to the reader. Extra technical results are postponed in Appendix, Subsection 2.6.3.

Lemma 2.4.0.1. *Let* $x \in \{x_0, x_{avg}\}$*. Assume* $(\mathcal{H}^a)$*, then we have*

$$\begin{aligned}\delta^{\mathrm{BS}}(x_0,\bar{a}_k^2T,k) &= \delta^{\mathrm{BS}}(x_0,\overline{a}_x^2T,k) + 2\partial_y\delta^{\mathrm{BS}}(x_0,\overline{a}_x^2T,k)(k-x)C_7(a;x)_0^T + O(\mathcal{M}_1(a)\mathcal{M}_0(a)T),\\ &= \delta^{\mathrm{BS}}(x_0,\overline{a}_x^2T,k) + [2(k-x)C_7(a;x)_0^T + (k-x)^2C_5(a;x)_0^T]\partial_y\delta^{\mathrm{BS}}(x_0,\overline{a}_x^2T,k)\\ &\quad + 4\partial^2_{y^2}\delta^{\mathrm{BS}}(x_0,\overline{a}_x^2T,k)(k-x)^2C_6(a;x)_0^T + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}),\end{aligned}$$

where $C_7(l;z)_0^T = \omega(l(z)l^{(1)}(z))_0^T$.

Lemma 2.4.0.2. *Let* $x \in \{x_0, x_{avg}\}$*. Assume* $(\mathcal{H}^a)$*, then we have*

$$C_1(\widetilde{a};k)_0^T(\partial^3_{z^3} - \frac{3}{2}\partial^2_{z^2} + \frac{1}{2}\partial_z)\delta^{\mathrm{BS}}(x_0,\bar{a}_k^2T,k) = C_1(\widetilde{a};x)_0^T(\partial^3_{z^3} - \frac{3}{2}\partial^2_{z^2} + \frac{1}{2}\partial_z)\delta^{\mathrm{BS}}(x_0,\overline{a}_x^2T,k) + O(\mathcal{M}_1(a)\mathcal{M}_0(a)T),$$

$$\begin{aligned}\sum_{i=1}^{6}\eta_i(\widetilde{a};k)_0^T\partial^i_{z^i}\delta^{\mathrm{BS}}(x_0,\bar{a}_k^2T,k) &= \sum_{i=1}^{6}\eta_i(\widetilde{a};x)_0^T\partial^i_{z^i}\delta^{\mathrm{BS}}(x_0,\overline{a}_x^2T,k)\\ &\quad + [2C_6(\widetilde{a};x)_0^T + C_2(\widetilde{a};x)_0^T](k-x)(\partial^3_{z^3} - \frac{3}{2}\partial^2_{z^2} + \frac{1}{2}\partial_z)\delta^{\mathrm{BS}}(x_0,\overline{a}_x^2T,k)\\ &\quad + 2(k-x)C_1(\widetilde{a};x)_0^TC_7(\widetilde{a};x)_0^T(\partial^4_{yz^3} - \frac{3}{2}\partial^3_{yz^2} + \frac{1}{2}\partial^2_{yz})\delta^{\mathrm{BS}}(x_0,\overline{a}_x^2T,k)\\ &\quad + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}).\end{aligned}$$

Then remark that:

$$C_1(\widetilde{a};x)_0^TC_7(\widetilde{a};x)_0^T = 2C_4(\widetilde{a};x)_0^T + C_8(\widetilde{a};x)_0^T,$$

where the operator C_8 is defined as follows:

$$C_8(l;z)_0^T = C_8(\widetilde{l};z)_0^T = \omega(l(z)l^{(1)}(z),l^2(z),l(z)l^{(1)}(z))_0^T.$$

An application of Proposition 2.6.3.2 finally yields the theorem below.

Theorem 2.4.0.3. ***(1st and 2nd order approximations for delta using local volatility at spot and mid-point)***. *Assume* $(\mathcal{H}^a)$ *and let* $x \in \{x_0, x_{avg}\}$. *We have:*

$$\begin{aligned}\delta(e^{x_0}, T, e^k) =& \delta^{\mathrm{BS}}(x_0, \overline{a}_x^2 T, k) + C_1(\overline{a}; x)_0^T (\partial^3_{z^3} - \frac{3}{2}\partial^2_{z^2} + \frac{1}{2}\partial_z)\delta^{\mathrm{BS}}(x_0, \overline{a}_x^2 T, k)\\ &+ (k-x)C_7(a;x)_0^T(\partial^2_{z^2} - \partial_z)\delta^{\mathrm{BS}}(x_0, \overline{a}_x^2 T, k) + O(\mathcal{M}_1(a)\mathcal{M}_0(a)T),\end{aligned}$$

$$\begin{aligned}\delta(e^{x_0}, T, e^k) =& \delta^{\mathrm{BS}}(x_0, \overline{a}_x^2 T, k) + \sum_{i=1}^{6} \eta_i(\overline{a}; k)_0^T \partial^i_{z^i} \delta^{\mathrm{BS}}(x_0, \overline{a}_x^2 T, k)\\ &+ (k-x)[C_7(a;x)_0^T + \frac{(k-x)}{2} C_5(a;x)_0^T](\partial^2_{z^2} - \partial_z)\delta^{\mathrm{BS}}(x_0, \overline{a}_x^2 T, k)\\ &+ (k-x)^2 C_6(a;x)_0^T(\partial^4_{z^4} - 2\partial^3_{z^3} + \partial^2_{z^2})\delta^{\mathrm{BS}}(x_0, \overline{a}_x^2 T, k)\\ &+ (k-x)[2C_6(\overline{a};x)_0^T + C_2(\overline{a};x)_0^T](\partial^3_{z^3} - \frac{3}{2}\partial^2_{z^2} + \frac{1}{2}\partial_z)\delta^{\mathrm{BS}}(x_0, \overline{a}_x^2 T, k)\\ &+ (k-x)[2C_4(\overline{a};x)_0^T + C_8(\overline{a};x)_0^T](\partial^5_{z^5} - \frac{5}{2}\partial^4_{z^4} + 2\partial^3_{z^3} - \frac{1}{2}\partial^2_{z^2})\delta^{\mathrm{BS}}(x_0, \overline{a}_x^2 T, k)\\ &+ O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}}).\end{aligned}$$

2.5 Numerical experiments

2.5.1 The set of tests

For the numerical experiments, we consider a CEV model with constant parameters: $\sigma(t, S) = \nu S^{\beta-1}$. We choose a spot value S_0 equal to 1 and we test two values of ν (a parameter interpreted as a level of volatility): firstly we set $\nu = 0.25$ and we consider either $\beta = 0.8$ (a priori close to the log-normal case) or $\beta = 0.2$ (a priori close to the normal case). Then we investigate the case of a larger volatility with $\nu = 0.4$ and $\beta = 0.5$. For the sake of completeness, we give in Appendix 2.6.5 the expressions of corrective coefficients allowing the computation of our various approximation formulas proposed throughout the Chapter.

We compare the accuracy of different approximations, for various maturities and various strikes gathered in 5 categories. The strikes evolve approximately as $S_0 \exp(c\nu\sqrt{T})$ where c takes the value

Table 2.1: Set of maturities and strikes for the numerical experiments

T/K	far ITM		ITM			ATM			OTM			far OTM	
3M	0.70	0.75	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.25	1.30	1.35
6M	0.65	0.75	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.20	1.25	1.35	1.50
1Y	0.55	0.65	0.75	0.80	0.90	0.95	1.00	1.05	1.15	1.25	1.40	1.50	1.80
1.5Y	0.50	0.60	0.70	0.75	0.85	0.95	1.00	1.10	1.15	1.30	1.50	1.65	2.00
2Y	0.45	0.55	0.65	0.75	0.85	0.90	1.00	1.10	1.20	1.35	1.55	1.80	2.30
3Y	0.35	0.50	0.55	0.70	0.80	0.90	1.00	1.10	1.25	1.45	1.75	2.05	2.70
5Y	0.25	0.40	0.50	0.60	0.75	0.85	1.00	1.15	1.35	1.60	2.05	2.50	3.60
10Y	0.15	0.25	0.35	0.50	0.65	0.80	1.00	1.20	1.50	1.95	2.75	3.65	6.30

of various quantiles of the standard Gaussian law (1%-5%-10%-20%-30%-40%-50%-60%-70%-80%-90%-95%-99%) which allows to cover far ITM and far OTM options. We report in Tables 2.2, 2.3 and 2.4 the Black-Scholes implied volatilities corresponding to the exact Call prices with constant parameters [Schroder 1989].

Table 2.2: CEV model ($\beta = 0.8, \nu = 0.25$): BS implied volatilities in %.

3M	25.90	25.73	25.56	25.41	25.26	25.13	25.00	24.88	24.76	24.65	24.45	24.35	24.26
6M	26.09	25.73	25.56	25.41	25.27	25.13	25.00	24.88	24.76	24.55	24.45	24.26	24.00
1Y	26.53	26.10	25.73	25.56	25.27	25.13	25.00	24.88	24.65	24.45	24.17	24.00	23.56
1.5Y	26.78	26.30	25.91	25.73	25.41	25.13	25.00	24.77	24.66	24.35	24.00	23.77	23.31
2Y	27.06	26.53	26.10	25.73	25.41	25.27	25.01	24.77	24.55	24.26	23.92	23.56	22.98
3Y	27.73	26.78	26.53	25.91	25.57	25.27	25.01	24.77	24.45	24.09	23.63	23.25	22.60
5Y	28.64	27.38	26.79	26.31	25.74	25.42	25.01	24.66	24.27	23.85	23.26	22.79	21.94
10Y	30.08	28.66	27.75	26.80	26.12	25.59	25.02	24.57	24.02	23.39	22.57	21.92	20.69

Table 2.3: CEV model ($\beta = 0.2, \nu = 0.25$): BS implied volatilities in %.

3M	28.75	28.00	27.31	26.67	26.08	25.53	25.01	24.53	24.07	23.64	22.84	22.48	22.13
6M	29.59	28.02	27.32	26.69	26.09	25.54	25.02	24.53	24.08	23.24	22.85	22.13	21.18
1Y	31.54	29.62	28.05	27.35	26.12	25.56	25.04	24.55	23.66	22.87	21.81	21.19	19.60
1.5Y	32.71	30.57	28.83	28.07	26.74	25.58	25.06	24.11	23.68	22.51	21.20	20.36	18.73
2Y	34.03	31.62	29.69	28.10	26.76	26.16	25.08	24.13	23.29	22.18	20.92	19.62	17.62
3Y	37.34	32.84	31.70	28.92	27.46	26.21	25.12	24.17	22.93	21.55	19.88	18.56	16.40
5Y	42.07	35.80	33.00	30.82	28.27	26.91	25.20	23.80	22.26	20.71	18.59	17.01	14.38
10Y	47.85	41.60	37.46	33.14	30.08	27.76	25.38	23.53	21.41	19.09	16.35	14.32	10.99

Table 2.4: CEV model ($\beta = 0.5, \nu = 0.4$): BS implied volatilities in %.

3M	43.69	42.97	42.29	41.67	41.08	40.53	40.02	39.53	39.07	38.63	37.82	37.45	37.09
6M	44.51	42.99	42.31	41.68	41.10	40.55	40.03	39.55	39.09	38.23	37.84	37.10	36.11
1Y	46.38	44.55	43.03	42.35	41.13	40.58	40.06	39.58	38.68	37.86	36.78	36.13	34.45
1.5Y	47.49	45.46	43.80	43.06	41.75	40.61	40.10	39.14	38.71	37.51	36.15	35.27	33.52
2Y	48.73	46.47	44.63	43.10	41.79	41.20	40.13	39.17	38.31	37.17	35.87	34.49	32.31
3Y	51.76	47.62	46.55	43.90	42.48	41.26	40.18	39.22	37.97	36.54	34.79	33.36	30.97
5Y	55.94	50.30	47.73	45.69	43.27	41.95	40.28	38.87	37.30	35.69	33.42	31.68	28.66
10Y	60.86	55.20	51.48	47.60	44.80	42.63	40.36	38.55	36.41	33.98	30.97	28.64	24.51

The purpose of the numerical tests is to compare the following approximations:

1. `ImpVol(AppPriceLN(2,z))` and `ImpVol(AppPriceN(2,z))`: the BS implied volatility of the second order expansions based respectively on the log-normal and normal proxy with local volatility frozen at point z, z being respectively equal to x_0, k or x_{avg} and to S_0, K or S_{avg}. See Theorems 2.1.2.3-2.1.3.1-2.1.3.2.

2. `AppImpVolLN(2,z)` and `AppImpVolN(2,z)`: the second order implied volatility expansions (Theorem 2.1.4.1). All the results are converted into Black-Scholes implied volatility. Namely, for the normal proxy, once we have computed Bachelier implied volatility expansions, we first evaluate the price with the Bachelier formula and then compute the related implied Black-Scholes volatility.

3. `ImpVol(AppPriceLN(3,z))` and `ImpVol(AppPriceN(3,z))`: the implied volatility of the third order expansions (Theorems 2.3.1.1 and 2.3.1.2). In addition for the log-normal proxy, we test the average of approximations based on strike and on spot and we denote it by `Av.ImpVol(AppPriceLN(3,.))`.

4. `AppImpVolLN(3,z)` and `AppImpVolN(3,z)`: the third order implied volatility expansions (The-

orem 2.3.3.1). We use the notation `Av.AppImpVolLN(3,.)` for the average of the expansions in strike and in spot.

5. Hagan and Henry-Labordère formulas denoted by (HF) and (HLF) in the following: benchmark implied volatility approximations of Hagan et al. [Hagan 1999, formula (7) p.149] and Henry-Labordère [Henry-Labordère 2008, formula (5.41) p.141]. For the sake of completeness, we recall these well-know implied volatility approximations in the CEV model:

$$\sigma_{\mathrm{I}}(x_0,T,k) \approx \nu\Big(\frac{S_0+K}{2}\Big)^{\beta-1}\Big(1+\frac{(1-\beta)(2+\beta)}{6}\Big(\frac{S_0-K}{S_0+K}\Big)^2+\frac{(\beta-1)^2\nu^2T}{24}\Big(\frac{S_0+K}{2}\Big)^{2\beta-2}\Big), \tag{HF}$$

$$\sigma_{\mathrm{I}}(x_0,T,k) \approx \frac{\nu(1-\beta)\log(\frac{S_0}{K})}{S_0^{1-\beta}-K^{1-\beta}}\Big(1+\frac{(\beta-1)^2\nu^2T}{24}\Big(\frac{S_0+K}{2}\Big)^{2\beta-2}\Big). \tag{HLF}$$

We recall that these formulas are essentially available for time-independent volatility, while our formulas allow time dependency.

Our goal is to demonstrate the interest of our approximation formulas in comparison to those of Hagan and Henry-Labordere. We are rather exhaustive with our numerical experiments in order to, on the one hand, select the best approximation formulas among ours, and on the other hand to show that our methods with log-normal proxy involving the mid-point generally outperform Hagan and Henry-Labordère formulas. Full details allow the reader to easily reproduce the results.

In Tables 2.7 and 2.9, we report the errors expressed in bps (basis points) on implied volatility for $(\beta,\nu) = (0.8,0.25)$ using the second and the third order price expansions. Tables 2.8 and 2.10 give results for the second and the third order implied volatility expansions. Next in Table 2.11, we report the errors in bps obtained with the averaged expansions `Av.ImpVol(AppPriceLN(3,.))` and `Av.AppImpVolLN(3,.)` and the benchmarks (HF) and (HLF). Then in Table 2.12, we compare `Av.ImpVol(AppPriceLN(3,.))`, `ImpVol(AppPriceLN(3,`x_{avg}`))`, `Av.AppImpVolLN(3,.)` and `AppImpVolLN(3,`x_{avg}`)` with the benchmarks (HF) and (HLF).
After we analyse the case $(\beta,\nu) = (0.2,0.25)$ and we report in Tables 2.13 and 2.14 the errors using `ImpVol(AppPriceLN(3,`x_{avg}`))`, `ImpVol(AppPriceN(3,`S_{avg}`))`, `AppImpVolLN(3,`x_{avg}`)`, `AppImpVolN(3,`S_{avg}`)` and the benchmarks (HF) and (HLF). Because the other methods in general give globally less accurate results, we just report and compare the best approximations.
Finally in Tables 2.15 and 2.16 we establish a comparison between `ImpVol(AppPriceLN(3,`x_{avg}`))`, `AppImpVolLN(3,`x_{avg}`)` and the benchmarks Hagan and (HLF) for $(\beta,\nu) = (0.5,0.4)$.

For example, on the first row of Table 2.7, the value -12 corresponds to the approximation error of `ImpVol(AppPriceLN(2,`x_0`))` for the first strike of the maturity $T = 3M$ (i.e. $K = 0.7$), whereas on the second row, the value -3 corresponds to the approximation error of `ImpVol(AppPriceLN(2,k))` for the third strike of the maturity $T = 6M$ (i.e. $K = 0.8$). If the price approximation does not belong to the non-arbitrage interval for Call options (it may happen for extreme strikes) we just report ND in the tabular.

2.5.2 Analysis of results

▷ **Influence of T and K.** We notice in Tables 2.7, 2.8, 2.9, 2.10 that errors are increasing w.r.t. T for all the different approximations: this is coherent with the $T^{3/2}$ or T^2-factor of our theoretical error estimates. For ATM options, all the approximations are excellent and errors remain small for a large range of strikes and maturities: with the log-normal proxy, usually smaller than 10 bps up to $10Y$ for strikes corresponding to the Gaussian quantiles range [10%, 90%].

▷ **Influence of the proxy.** As expected, approximations based on log-normal proxy perform better than approximations based on normal proxy. On the one hand, we obtain simpler approximation formulas with the normal proxy: on the other hand, the errors become significant when considering slightly OTM or ITM options, even for short maturities and for advanced methods (order 3, local volatility frozen at the mid-point…).

▷ **Influence of the order.** Regarding firstly Tables 2.7-2.9 and then Tables 2.8-2.10, we notice that as expected, third order approximations are more accurate than second order ones. In addition, for the log-normal proxy case, second order approximations in spot or strike often underestimate the true implied volatility values whereas third order approximations in spot overestimate the true values for OTM options and yield underestimation for ITM options; the converse occurs for the third order approximations in strike. Because the errors have approximately the same magnitude but with opposite signs, approximations are improved by considering the average between the approximations. It is discussed below.

▷ **Influence of the point.** Unquestionably, methods using the local volatility at mid-point systematically give the best results. With `ImpVol(AppPriceLN(2,`x_{avg}`))` (Table 2.7), errors do not exceed 15 bps for the whole set of strikes and maturities, which is already really good, whereas `ImpVol(AppPriceLN(3,`x_{avg}`))` and `AppImpVolLN(3,`x_{avg}`)` provide errors close to 0 proving an extreme accuracy.

▷ **Price expansions vs implied volatility expansions.** Generally speaking, the implied volatility expansions are more precise and stable. This can be easily observed by comparing on the one hand Tables 2.7 and 2.8 and on the other hand Tables 2.9 and 2.10. Sometimes, especially for extreme strikes, a simple direct second order second implied volatility expansion is more accurate than the corresponding third order price expansion. Since in addition the formulas are easier to compute, we recommend the use of implied volatility expansions. Moreover, the difference between `ImpVol(AppPriceLN(3,`x_{avg}`))` and `AppImpVolLN(3,`x_{avg}`)` is not clear, both methods giving similar and excellent results (see Tables 2.12 or 2.13) although the direct implied volatility expansion remains more stable especially for $\beta = 0.2$ and/or for large maturities. Last, when the local volatility function is frozen at spot or at strike, there is really an improvement in using implied volatility expansions instead of the corresponding price expansions.

▷ **Comparison with the benchmarks.** In Table 2.11, we report the performance of the methods `Av.ImpVol(AppPriceLN(3,.))`, `Av.AppImpVolLN(3,.)` and the benchmarks (`HF`) and (`HLF`). Errors on the implied volatility are equal to zero bp for the whole range of maturities and strikes for `Av.AppImpVolLN(3,.)` and the (`HLF`) approximation, whereas `Av.ImpVol(AppPriceLN(3,.))` and (`HF`) provide errors smaller than 45 and 70 bps in absolute value respectively. In Table 2.12, we compare `Av.ImpVol(AppPriceLN(3,.))`, `ImpVol(AppPriceLN(3,`x_{avg}`))`, `Av.AppImpVolLN(3,.)` and `AppImpVolLN(3,`x_{avg}`)` with the benchmarks (`HF`) and (`HLF`). In order to observe more clearly the accuracy of the different methods, we partially gather the results and we report the average of errors for different categories of strike (far ITM, ITM, ATM, OTM and far OTM, see Table 2.1), using a scientific notation for the errors. Computing the average per categories of strikes gives an advantage to methods which errors have non constant sign. These methods may be more reliable than those giving a systematic over/under-estimation.
The best method is clearly `AppImpVolLN(3,`x_{avg}`)` which yields errors of 10^{-5} bps for short maturities and 10^{-2} bps for long maturities. The method gives better results than the excellent approximation proposed by Henry-Labordère (errors of 10^{-4} bps for short maturities and 10^{-1} bps for long maturities). `ImpVol(AppPriceLN(3,`x_{avg}`))` seems to be slightly better than (`HLF`) but is less robust for extreme

strikes than AppImpVolLN(3,x_{avg}). Significantly better results are obtained by averaging the expansions in spot and strike, thanks to the symmetrical roles played by these two variables. The results are close to those of the corresponding expansions with the mid-point, but they remain less accurate and less robust for extreme strikes. The problem of this averaging method is the risk of huge inaccuracy if one of two approximations in spot and strike fails. (HF) is clearly less accurate than all the other approximations.

▷ **Influence of β.** In the Table 2.13, as expected the log-normal proxy provides larger errors than for $\beta = 0.8$. Although the results of the normal proxy are better in comparison with the case $\beta = 0.8$, they remain less accurate and less robust than those obtained with the log-normal proxy. Up to the maturity $5Y$, AppImpVolLN(3,x_{avg}) yields errors in bps smaller than 7 bps which is truly excellent. (HLF) gives comparable results. (HF) seems less accurate and cruder for extreme strikes. For the maturity $10Y$, we observe that AppImpVolLN(3,x_{avg}) (maximal error close to 159 bps) behaves better than (HLF) (maximal error close to 271 bps) for very small strikes, whereas for very large strikes (HLF) is slightly better (-5 bps for AppImpVolLN(3,x_{avg}) versus -1 bp for (HLF)). Surprisingly (HF) yields the smallest maximal error (close to 112 bps) but is more inaccurate for OTM. (HLF) and AppImpVolLN(3,x_{avg}) give excellent results with errors of the order of 10^{-3} bps for short maturity ($3M$) and 10^{-1} bps for the maturity $3Y$. We nevertheless notice that ATM, (HLF) is better.

▷ **Impact of ν.** The level of volatility ν plays a similar role to $\sqrt{T}$, and in Tables 2.15 and 2.16, we analyse the impact of a larger volatility on our approximations. We take $\nu = 40\%$ and $\beta = 0.5$. We notice that up to the maturity $5Y$, the errors in bps do not exceed 6 bps for the methods ImpVol(AppPriceLN(3,x_{avg})) or AppImpVolLN(3,x_{avg}) with a maximal error of 92 bps for the maturity $10Y$. Their accuracy is better than those of (HF) or (HLF) for short and long maturities. (HLF) is much more inaccurate ITM for the maturity $10Y$ (maximal error of 286 bps). In Table 2.16, we aggregate the results per categories of strike up to the maturity $3Y$ and we observe a good accuracy of ImpVol(AppPriceLN(3,x_{avg})) and AppImpVolLN(3,x_{avg}): 10^{-3} bps for the maturity $3M$ and 10^{-1} for the maturity $3Y$. In particular we notice that ATM, (HF) and (HLF) are less accurate.

In view of all these tests, we may conclude that ImpVol(AppPriceLN(3,x_{avg})) and particulary AppImpVolLN(3,x_{avg}) give very satisfying results, being at least as good as the Henry-Labordère formula in the worst situations ($\beta = 0.2$ or $\nu = 0.4$) and being often better in the case $\beta = 0.8$. The different current tests prove that our direct implied volatility approximations outperform the corresponding price approximations. In addition, a normal proxy seems not to be the most appropriate for the approximation of a CEV model, in view of the large errors obtained especially for very small strikes. This presumably explains why the Hagan formula is much less accurate than our approximations with log-normal proxy and than that of Henry-Labordère. The Hagan formula is namely close in the spirit to our approximation formulas with normal proxy.
To conclude, our approximations maintain very tight error estimates and allow to deal naturally with general time-dependent local volatility (or with stochastic interest rates, see [Benhamou 2012]) which is a significant advantage compared to other approaches.

2.5.3 CEV Delta approximations

Now we test our approximation formulas for the deltas, by choosing again a CEV model with spot value $S_0 = 1$ and constant parameters. We test the values $(\beta, \nu) = (0.8, 0.25)$ and $(\beta, \nu) = (0.2, 0.25)$. We report in Tables 2.5 and 2.6 the exact delta values for the set of maturities and strikes defined in Table 2.1.

Table 2.5: CEV model ($\beta = 0.8, \nu = 0.25$): deltas in %.

3M	99.75	98.89	96.38	90.83	81.18	67.67	51.99	36.60	23.57	13.91	3.78	1.76	0.76
6M	99.20	95.10	90.44	83.56	74.58	64.05	52.82	41.80	31.77	16.33	11.09	4.64	1.01
1Y	99.09	96.04	88.90	83.53	69.78	61.97	53.98	46.15	31.98	20.75	9.76	5.55	0.82
1.5Y	98.76	95.68	89.46	85.06	74.11	61.40	54.88	42.35	36.62	22.35	10.35	5.44	1.03
2Y	98.75	95.94	90.58	82.55	72.45	66.93	55.63	44.75	34.97	23.04	12.30	5.15	0.75
3Y	99.13	95.36	93.13	83.51	75.22	66.13	56.88	47.99	36.04	23.46	11.44	5.25	0.86
5Y	99.23	95.93	91.89	86.53	76.75	69.63	58.86	48.70	36.85	25.21	12.05	5.54	0.79
10Y	99.13	97.10	93.95	87.68	80.34	72.60	62.46	53.06	40.92	27.19	12.88	5.57	0.54

Table 2.6: CEV model ($\beta = 0.2, \nu = 0.25$): deltas in %.

3M	99.37	98.08	95.04	89.11	79.43	66.09	50.50	34.98	21.76	12.08	2.60	1.01	0.34
6M	98.13	92.97	88.01	81.08	72.24	61.88	50.71	39.59	29.35	13.70	8.60	2.85	0.35
1Y	97.34	93.15	85.45	80.09	66.66	58.97	51.00	43.08	28.51	17.00	6.39	2.91	0.14
1.5Y	96.23	92.03	85.29	80.89	70.22	57.76	51.23	38.45	32.50	17.73	6.20	2.35	0.13
2Y	95.71	91.77	85.84	77.85	68.06	62.67	51.43	40.30	30.09	17.68	7.29	1.81	0.04
3Y	95.62	90.10	87.57	77.87	69.86	61.01	51.76	42.61	30.02	16.82	5.53	1.36	0.03
5Y	94.70	89.45	84.88	79.47	70.04	63.12	52.32	41.69	28.91	16.52	4.53	0.87	0.00
10Y	94.37	90.36	86.11	79.22	71.81	64.05	53.52	43.32	29.65	14.59	2.76	0.23	0.00

We test the 6 following approximations:

1. `AppDeltaLN(1,`x_0`)`, `AppDeltaLN(1,k)` and `AppDeltaLN(1,`x_{avg}`)`: first order delta expansions based on the log-normal proxy with local volatility frozen at point x_0, k and x_{avg}.

2. `AppDeltaLN(2,`x_0`)`, `AppDeltaLN(2,k)` and `AppDeltaLN(2,`x_{avg}`)`: second order delta expansions based on the log-normal proxy with local volatility frozen at point x_0, k and x_{avg}.

Tables 2.17-2.18 (respectively 2.19) give errors on deltas (expressed in bps) using all the approximations with $\beta = 0.8$ (respectively $\beta = 0.2$).
Regarding the results, the accuracy for $\beta = 0.8$ is excellent because, except for `AppDeltaLN(1,`x_0`)`, we obtain a maximal error (in absolute value) equal to 36 bps. Generally speaking, approximations with local volatility at spot are not as good as related approximations at strike. In addition, for second order formulas, we do not observe any symmetry between the spot and strike approximations (which often overestimate the exact delta), whereas the symmetry slightly appears for the first order expansions (not exactly with the same magnitude but opposite signs). Maybe in this situation, the optimal expansion point is not exactly the convex combination $x_{avg} = (x_0 + k)/2$. However the methods with the mid-point are truly excellent, in particular `AppDeltaLN(2,`x_{avg}`)` which yields a maximal error (in absolute value) close to 1 bps. From Table 2.18, we observe that in average, the errors for `AppDeltaLN(2,`x_{avg}`)` range from 10^{-3} for short maturities to 10^{-1} for long maturities.

In Table 2.19 ($\beta = 0.2$), without surprise the errors are larger compared to $\beta = 0.8$. The best approximation is still `AppDeltaLN(2,`x_{avg}`)` which provides errors smaller than 27 bps up to $5Y$ with a global maximal error of 157 bps, which remains quite good. Curiously, for ATM options, the first order approximation may give better estimates even if the related errors quickly increase for large or small strikes in comparison with the second order approximations.

2.6 Appendix

2.6.1 Computations of derivatives of $\mathrm{Call}^{\mathrm{BS}}$ w.r.t the log spot, the log strike and the total variance

In the following Proposition, we make explicit the formula for the derivatives at any order of $\mathrm{Call}^{\mathrm{BS}}$ w.r.t. x and z:

Proposition 2.6.1.1. *Let $x, z \in \mathbb{R}$ and $y > 0$. For any integer $n \geq 1$, we have:*

$$\partial^n_{x^n}\mathrm{Call}^{\mathrm{BS}}(x,y,z) = e^x\mathcal{N}(d_1(x,y,z)) + \mathbb{1}_{n\geq 2}\, e^x\mathcal{N}'(d_1(x,y,z))\sum_{k=1}^{n-1}\binom{n-1}{k}(-1)^{k-1}\frac{H_{k-1}(d_1(x,y,z))}{y^{\frac{k}{2}}},$$

$$\partial^z_{z^n}\mathrm{Call}^{\mathrm{BS}}(x,y,z) = -\,e^z\mathcal{N}(d_2(x,y,z)) + \mathbb{1}_{n\geq 2}\, e^z\mathcal{N}'(d_2(x,y,z))\sum_{k=1}^{n-1}\binom{n-1}{k}\frac{H_{k-1}(d_2(x,y,z))}{y^{\frac{k}{2}}},$$

where $(H_k)_{k\in\mathbb{N}}$ are the Hermite polynomials defined for any $n \in \mathbb{N}$ and for any $x \in \mathbb{R}$ by:

$$H_n(x) = (-1)^n e^{x^2/2}\partial^n_{x^n}(e^{-x^2/2})$$

Proof. For $n = 1$ the formulas are easy to obtain. For $n \geq 2$, apply the Leibniz formula to the products $e^x\mathcal{N}(d_1(x,y,z))$ and $e^z\mathcal{N}(d_2(x,y,z))$. □

We deduce a very useful Corollary:

Corollary 2.6.1.1. *Let $x, z \in \mathbb{R}$ and $y > 0$. For any integers $n \geq 1$ and $m \geq 1$, we have:*

$$|\partial^n_{x^n}\mathrm{Call}^{\mathrm{BS}}(x,y,z)| + |\partial^n_{z^n}\mathrm{Call}^{\mathrm{BS}}(x,y,z)| \leq_c y^{\frac{1-n}{2}},$$

$$|x-z|^m|\partial^n_{x^n}\mathrm{Call}^{\mathrm{BS}}(x,y,z) - e^x\mathcal{N}(d_1(x,y,z))| \leq_c y^{\frac{1-n+m}{2}},$$

$$|x-z|^m|\partial^n_{z^n}\mathrm{Call}^{\mathrm{BS}}(x,y,z) + e^z\mathcal{N}(d_2(x,y,z))| \leq_c y^{\frac{1-n+m}{2}},$$

where the generic constants depend polynomially on y.

Remark 2.6.1.1. *In practice the two last estimates are used when we want to bound $(x-z)^m\sum_{i=1}^{n}\alpha_i\partial^i_{x^i}\mathrm{Call}^{\mathrm{BS}}(x,y,z)$ or $(x-z)^m\sum_{i=1}^{n}\alpha_i\partial^i_{z^i}\mathrm{Call}^{\mathrm{BS}}(x,y,z)$ (with $\sum_{i=1}^{n}\alpha_i = 0$) by a power of y with the highest possible degree.*

Proof. We recall that for any polynomial function $\mathcal{P}$, $x \mapsto \mathcal{P}(x)\mathcal{N}'(x)$ is a bounded function. Then the first inequality follows directly from Proposition 2.6.1.1. For the second and the third, write $(x-z) = d_1(x,y,z)\sqrt{y} - \frac{1}{2}y = d_2(x,y,z)\sqrt{y} + \frac{1}{2}y$ and conclude similarly. □

In the next Proposition, we provide the formulas of the first, the second and the third derivatives of $\mathrm{Call}^{\mathrm{BS}}$ w.r.t. a positive volatility:

Proposition 2.6.1.2. *Let $x, z \in \mathbb{R}$, $\nu > 0$ and $T > 0$. We have:*

$$\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z) = \partial_\nu\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z) = e^x\sqrt{T}\mathcal{N}'(d_1(x,\nu^2T,z)) = e^z\sqrt{T}\mathcal{N}'(d_2(x,\nu^2T,z)), \tag{2.42}$$

$$\begin{aligned}\mathrm{Vomma}^{\mathrm{BS}}(x,\nu^2T,z) &= \partial_\nu\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z) = \frac{\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)}{\nu}d_1(x,\nu^2T,z)d_2(x,\nu^2T,z)\\ &= \frac{\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)}{\nu}\left[\frac{(x-z)^2}{\nu^2T} - \frac{\nu^2T}{4}\right],\end{aligned} \tag{2.43}$$

$$\begin{aligned}\mathrm{Ultima}^{\mathrm{BS}}(x,\nu^2T,z) =&\partial_\nu \mathrm{Vomma}^{\mathrm{BS}}(x,\nu^2T,z) = -\frac{\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)}{\nu^2}[d_1d_2(1-d_1d_2)+d_1^2+d_2^2](x,\nu^2T,z)\\ =&-\frac{\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)}{\nu^2}[\frac{(x-z)^2}{2}+\frac{3(x-z)^2}{\nu^2T}+\frac{\nu^2T}{4}-\frac{(x-z)^4}{\nu^4T^2}-\frac{\nu^4T^2}{16}].\end{aligned} \tag{2.44}$$

The above Proposition directly implies the following result:

Corollary 2.6.1.2. *Let $x,z\in\mathbb{R}$, $\nu>0$ and $T>0$. We have the following estimates:*

$$0<\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)\le_c\sqrt{T}, \tag{2.45}$$

$$|\mathrm{Vomma}^{\mathrm{BS}}(x,\nu^2T,z)|\le_c\frac{\sqrt{T}}{\nu}, \tag{2.46}$$

$$|\mathrm{Ultima}^{\mathrm{BS}}(x,\nu^2T,z)|\le_c\frac{\sqrt{T}}{\nu^2}, \tag{2.47}$$

where the generic constants depend polynomially of ν.

We finally state relations between the derivatives w.r.t. x or z, the $\mathrm{Vega}^{\mathrm{BS}}$ and the $\mathrm{Vomma}^{\mathrm{BS}}$. These relations allow on the one hand to replace derivatives w.r.t. z with derivatives w.r.t. x and on the other hand to write the differential operators w.r.t. x or z in terms of the $\mathrm{Vega}^{\mathrm{BS}}$ and the $\mathrm{Vomma}^{\mathrm{BS}}$. The verification of these identities is tedious but without mathematical difficulties. For instance, we have used Mathematica to check these relations.

Proposition 2.6.1.3. *Let $x,z\in\mathbb{R}$, $\nu>0$ and $T>0$. We have:*

$$\begin{aligned}(\partial^2_{x^2}-\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z) =&(\partial^2_{z^2}-\partial_z)\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z)\\ =&\frac{e^x}{\nu\sqrt{T}}\mathcal{N}'(d_1(x,\nu^2T,z))=\frac{\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)}{\nu T},\end{aligned} \tag{2.48}$$

$$\begin{aligned}(\partial^3_{x^3}-\frac{3}{2}\partial^2_{x^2}+\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z) =&-(\partial^3_{z^3}-\frac{3}{2}\partial^2_{z^2}+\frac{1}{2}\partial_z)\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z)\\ =&-\frac{e^x(x-z)}{\nu^3T^{\frac{3}{2}}}\mathcal{N}'(d_1(x,\nu^2T,z))=-\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)\frac{(x-z)}{\nu^3T^2},\end{aligned} \tag{2.49}$$

$$\begin{aligned}(\frac{1}{4}\partial^4_{x^4}-\frac{1}{2}\partial^3_{x^3}+\frac{1}{4}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z) =&(\frac{1}{4}\partial^4_{z^4}-\frac{1}{2}\partial^3_{z^3}+\frac{1}{4}\partial^2_{z^2})\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z)\\ =&e^x\mathcal{N}'(d_1(x,\nu^2T,z))[\frac{(x-z)^2}{4\nu^5T^{\frac{5}{2}}}-\frac{1}{16\nu\sqrt{T}}-\frac{1}{4\nu^3T^{\frac{3}{2}}}]\\ =&\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)[\frac{(x-z)^2}{4\nu^5T^3}-\frac{1}{16\nu T}-\frac{1}{4\nu^3T^2}],\end{aligned} \tag{2.50}$$

$$\begin{aligned}(\partial^4_{x^4}-2\partial^3_{x^3}+\frac{5}{4}\partial^2_{x^2}-\frac{1}{4}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z) =&(\partial^4_{z^4}-2\partial^3_{z^3}+\frac{5}{4}\partial^2_{z^2}-\frac{1}{4}\partial_z)\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z)\\ =&e^x\mathcal{N}'(d_1(x,\nu^2T,z))[\frac{(x-z)^2}{\nu^5T^{\frac{5}{2}}}-\frac{1}{\nu^3T^{\frac{3}{2}}}]\\ =&\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)[\frac{(x-z)^2}{\nu^5T^3}-\frac{1}{\nu^3T^2}],\end{aligned} \tag{2.51}$$

$$\begin{aligned}(3\partial^4_{x^4}-6\partial^3_{x^3}+\frac{7}{2}\partial^2_{x^2}-\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z) =&(3\partial^4_{z^4}-6\partial^3_{z^3}+\frac{7}{2}\partial^2_{z^2}-\frac{1}{2}\partial_z)\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z)\\ =&e^x\mathcal{N}'(d_1(x,\nu^2T,z))[3\frac{(x-z)^2}{\nu^5T^{\frac{5}{2}}}-\frac{3}{\nu^3T^{\frac{3}{2}}}-\frac{1}{4\nu\sqrt{T}}]\end{aligned}$$

$$= \mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)[3\frac{(x-z)^2}{\nu^5T^3}-\frac{3}{\nu^3T^2}-\frac{1}{4\nu T}], \tag{2.52}$$

$$\begin{aligned}
&(\frac{1}{2}\partial^6_{x^6}-\frac{3}{2}\partial^5_{x^5}+\frac{13}{8}\partial^4_{x^4}-\frac{3}{4}\partial^3_{x^3}+\frac{1}{8}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z)\\
=&(\frac{1}{2}\partial^6_{z^6}-\frac{3}{2}\partial^5_{z^5}+\frac{13}{8}\partial^4_{z^4}-\frac{3}{4}\partial^3_{z^3}+\frac{1}{8}\partial^2_{z^2})\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,z)\\
=&e^x\mathcal{N}'(d_1(x,\nu^2T,z))[\frac{(x-z)^4}{2\nu^9T^{\frac{9}{2}}}-\frac{(x-z)^2}{8\nu^5T^{\frac{5}{2}}}-3\frac{(x-z)^2}{\nu^7T^{\frac{7}{2}}}+\frac{1}{8\nu^3T^{\frac{3}{2}}}+\frac{3}{2\nu^5T^{\frac{5}{2}}}]\\
=&\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,z)[-3\frac{(x-z)^2}{\nu^7T^4}+\frac{1}{8\nu^3T^2}+\frac{3}{2\nu^5T^3}]+\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x,\nu^2T,z)\frac{(x-z)^2}{\nu^6T^4}.
\end{aligned} \tag{2.53}$$

2.6.2 Derivatives of $\mathrm{Call}^{\mathrm{BA}}$ w.r.t the spot, the strike and the total variance

Proposition 2.6.2.1. *Let $S,Z\in\mathbb{R}$ and $Y>0$. For any integer $n\geq 1$, we have:*

$$\begin{aligned}
\partial^n_{S^n}\mathrm{Call}^{\mathrm{BA}}(S,Y,Z) =&\mathbb{1}_{n=1}\mathcal{N}(\frac{S-Z}{\sqrt{Y}})+\mathbb{1}_{n\geq 2}\mathcal{N}'(\frac{S-Z}{\sqrt{Y}})(-1)^{n-2}\frac{H_{n-2}(\frac{S-Z}{\sqrt{Y}})}{Y^{\frac{n-1}{2}}},\\
\partial^n_{Z^n}\mathrm{Call}^{\mathrm{BA}}(S,Y,Z) =&-\mathbb{1}_{n=1}\mathcal{N}(\frac{S-Z}{\sqrt{Y}})+\mathbb{1}_{n\geq 2}\mathcal{N}'(\frac{S-Z}{\sqrt{Y}})\frac{H_{n-2}(\frac{S-Z}{\sqrt{Y}})}{Y^{\frac{n-1}{2}}}.
\end{aligned}$$

Corollary 2.6.2.1. *Let $S,Z\in\mathbb{R}$ and $Y>0$. For any integers $n\geq 2$ and $m\geq 1$, we have:*

$$|S-Z|^m(|\partial^n_{S^n}\mathrm{Call}^{\mathrm{BA}}(S,Y,Z)|+|\partial^n_{Z^n}\mathrm{Call}^{\mathrm{BA}}(S,Y,Z)|)\leq_c Y^{\frac{1-n+m}{2}},$$

where the generic constants depend polynomially on Y.

Proposition 2.6.2.2. *Let $S,Z\in\mathbb{R}$, $V>0$ and $T>0$. We have:*

$$\begin{aligned}
\mathrm{Vega}^{\mathrm{BA}}(S,V^2T,Z) =&\partial_V\mathrm{Call}^{\mathrm{BA}}(S,V^2T,Z) = \sqrt{T}\mathcal{N}'(\frac{S-Z}{V\sqrt{T}}),\\
\mathrm{Vomma}^{\mathrm{BA}}(S,V^2T,Z) =&\partial_V\mathrm{Vega}^{\mathrm{BA}}(S,V^2T,Z) = \frac{\mathrm{Vega}^{\mathrm{BA}}(S,V^2T,Z)}{V}\frac{(S-Z)^2}{V^2T},\\
\mathrm{Ultima}^{\mathrm{BA}}(S,V^2T,Z) =&\partial_V\mathrm{Vomma}^{\mathrm{BA}}(S,V^2T,Z) = -\frac{\mathrm{Vega}^{\mathrm{BA}}(S,V^2T,Z)}{V^2}[\frac{3(S-Z)^2}{V^2T}-\frac{(S-Z)^4}{V^4T^2}].
\end{aligned}$$

Corollary 2.6.2.2. *Let $S,Z\in\mathbb{R}$ $V>0$ and $T>0$. We have the following estimates:*

$$\begin{aligned}
0<\mathrm{Vega}^{\mathrm{BA}}(S,V^2T,Z)\leq_c& \sqrt{T},\\
|\mathrm{Vomma}^{\mathrm{BA}}(S,V^2T,Z)|\leq_c&\frac{\sqrt{T}}{V},\\
|\mathrm{Ultima}^{\mathrm{BA}}(S,V^2T,Z)|\leq_c&\frac{\sqrt{T}}{V^2},
\end{aligned}$$

where the generic constants depend polynomially on V.

Proposition 2.6.2.3. *Let $S,Z\in\mathbb{R}$ $V>0$ and $T>0$. We have:*

$$\begin{aligned}
\partial^2_{S^2}\mathrm{Call}^{\mathrm{BA}}(S,V^2T,Z) =& \partial^2_{Z^2}\mathrm{Call}^{\mathrm{BA}}(x,V^2T,Z) = \frac{\mathrm{Vega}^{\mathrm{BA}}(S,V^2T,Z)}{VT},\\
\partial^3_{S^3}\mathrm{Call}^{\mathrm{BA}}(S,V^2T,Z) =& -\partial^3_{Z^3}\mathrm{Call}^{\mathrm{BA}}(S,V^2T,Z) = -\mathrm{Vega}^{\mathrm{BA}}(S,V^2T,Z)\frac{(S-Z)}{V^3T^2},
\end{aligned}$$

$$\partial^3_{S^4}\mathrm{Call}^{\mathrm{BA}}(S,V^2T,Z) = \partial^4_{Z^4}\mathrm{Call}^{\mathrm{BA}}(S,V^2T,Z) = \mathrm{Vega}^{\mathrm{BA}}(S,V^2T,Z)[\frac{(S-Z)^2}{V^5T^3} - \frac{1}{V^3T^2}],$$
$$\partial^6_{S^6}\mathrm{Call}^{\mathrm{BA}}(S,V^2T,Z) = \partial^6_{Z^6}\mathrm{Call}^{\mathrm{BA}}(S,V^2T,Z) = \mathrm{Vega}^{\mathrm{BA}}(S,V,T,Z)[-6\frac{(S-Z)^2}{V^7T^4} + \frac{3}{V^5T^3}]$$
$$+\mathrm{Vomma}^{\mathrm{BA}}(S,V^2T,Z)\frac{(S-Z)^2}{V^6T^4}.$$

2.6.3 Derivatives of δ^{BS} w.r.t the log spot, the log strike and the total variance

Proposition 2.6.3.1. *Let $x,z \in \mathbb{R}$ and $y > 0$. For any integer $n \geq 1$, we have:*

$$\partial^n_{x^n}\delta^{\mathrm{BS}}(x,y,z) = (-1)^{n-1}\mathcal{N}'(d_1(x,y,z))\frac{H_{n-1}(d_1(x,y,z))}{y^{\frac{n}{2}}},$$
$$\partial^n_{z^n}\delta^{\mathrm{BS}}(x,y,z) = -\mathcal{N}'(d_1(x,y,z))\frac{H_{n-1}(d_1(x,y,z))}{y^{\frac{n}{2}}}.$$

Corollary 2.6.3.1. *Let $x,z \in \mathbb{R}$ and $y > 0$. For any integers $n \geq 1$ and $m \geq 1$, we have:*

$$|x-z|^m(|\partial^n_{x^n}\delta^{\mathrm{BS}}(x,y,z)| + |\partial^n_{z^n}\delta^{\mathrm{BS}}(x,y,z)|) \leq_c y^{\frac{m-n}{2}},$$

where the generic constants depend polynomially on y.

Proposition 2.6.3.2. *Let $x,z \in \mathbb{R}$ and $y > 0$. We have:*

$$\partial_y\delta^{\mathrm{BS}}(x,y,z) = \frac{1}{2}(\partial^2_{z^2} - \partial_z)\delta^{\mathrm{BS}}(x,y,z) = -\frac{\mathcal{N}'(d_1(x,y,z))}{2y}d_2(x,y,z).$$

2.6.4 Proof of Lemma 2.1.2.1

We proceed by induction. The key is to prove the above technical result:

Lemma 2.6.4.1. *Let $(m_t)_{t\in[0,T]}$ be a square integrable and predictable process, $(\lambda_t)_{t\in[0,T]}$ be a measurable and bounded deterministic function and φ be a C^∞_b function. Then, we have:*

$$\mathbb{E}\Big(\varphi(\int_0^T \lambda_t \mathrm{d}W_t)\int_0^T m_t \mathrm{d}W_t\Big) = \mathbb{E}\Big(\varphi^{(1)}(\int_0^T \lambda_t \mathrm{d}W_t)\int_0^T \lambda_t m_t \mathrm{d}t\Big).$$

Proof. We propose two proofs: firstly we employ a PDE argument and secondly we show that this is a straightforward application of the Malliavin calculus theory. In the two points of view, we use the common notation for the diffusion process $(Z_t)_{t\in[0,T]} = (\int_0^t \lambda_s \mathrm{d}W_s)_{t\in[0,T]}$ and we recall that $(\mathcal{F}_t)_{t\in[0,T]}$ denotes the augmented filtration of the Brownian motion W.

▷ **PDE argument.** We introduce $u(t,x) = \mathbb{E}[\varphi(Z_T)|Z_t = x]$ which solves the following PDE with terminal condition:

$$\begin{cases} \partial_t u(t,x) + \frac{1}{2}\lambda_t^2\partial^2_{xx}u(t,x) = 0, \quad (t,x) \in]0,T[\times\mathbb{R}, \\ u(T,x) = \varphi(x), \quad x \in \mathbb{R}. \end{cases}$$

Thanks to the above PDE and the assumption on φ, $\forall i \in \mathbb{N}$, $\partial^i_{x^i}(u(t,Z_t))_{t\in[0,T]}$ is a martingale and $\forall t \in [0,T]$, we have:

$$\partial^i_{x^i}u(t,Z_t) = \mathbb{E}[\varphi^{(i)}(Z_T)|\mathcal{F}_t] = \mathbb{E}[\varphi^{(i)}(Z_T)] + \int_0^t \partial^{i+1}_{x^{i+1}}u(s,Z_s)\lambda_s \mathrm{d}W_s.$$

Then applying the L^2-isometry for the product $u(T,Z_T)\int_0^T m_t \mathrm{d}W_t = \varphi(Z_T)\int_0^T m_t \mathrm{d}W_t$, it readily comes:

$$\mathbb{E}\Big(\varphi(Z_T)\int_0^T m_t \mathrm{d}W_t\Big) = \int_0^T \mathbb{E}[\partial_x u(t,Z_t)\lambda_t m_t]\mathrm{d}t = \mathbb{E}\Big(\varphi^{(1)}(Z_T)\int_0^T \lambda_t m_t \mathrm{d}t\Big),$$

where at the last equality we have used the martingale property of $\partial_x(u(t,Z_t))_{t\in[0,T]}$.

$\triangleright$ **Malliavin calculus approach.** The result directly comes from the duality relationship of Malliavin calculus (see [Nualart 2006, Lemma 1.2.1]) identifying the Itô integral $\int_0^T m_t \mathrm{d}W_t$ with the Skorohod operator and observing that $(\varphi^{(1)}(Z_T)\lambda_t)_{t\in[0,T]}$ is the first Malliavin derivative of $\varphi(Z_T)$. $\square$

Lemma 2.6.4.1 is a particular case of Lemma 2.1.2.1 for $N = 1$ and $I_N = 1$ noting that $\forall i \in \mathbb{N}$, $\mathbb{E}\Big(\varphi^{(i)}(\int_0^T \lambda_t \mathrm{d}W_t)\Big) = \partial^i_{\varepsilon^i}\mathbb{E}\Big(\varphi(\int_0^T \lambda_t \mathrm{d}W_t + \varepsilon)\Big)|_{\varepsilon=0}$, thanks to the regularity of φ. For $N = 1$ and $I_N = 0$, there is nothing to prove. Suppose that the formula (2.9) is true for $N \geq 2$. Then apply Lemma 2.6.4.1 if $I_{N+1} = 1$ to obtain:

$$\begin{aligned}
&\mathbb{E}\Big(\varphi(\int_0^T \lambda_t \mathrm{d}W_t)\int_0^T l_{N+1,t_{N+1}}\int_0^{t_{N+1}} l_{N,t_N}\cdots\int_0^{t_2} l_{1,t_1}\mathrm{d}W^{I_1}_{t_1}\ldots\mathrm{d}W^{I_N}_{t_N}\mathrm{d}W^{I_{N+1}}_{t_{N+1}}\Big)\\
=&\mathbb{E}\Big(\varphi^{(I_{N+1})}(\int_0^T \lambda_t \mathrm{d}W_t)\int_0^T \widehat{l}_{N+1,t_{N+1}}\int_0^{t_{N+1}} l_{N,t_N}\int_0^{t_N}\cdots\int_0^{t_2} l_{1,t_1}\mathrm{d}W^{I_1}_{t_1}\ldots\mathrm{d}W^{I_N}_{t_N}\mathrm{d}t_{N+1}\Big)\\
=&\mathbb{E}\Big(\varphi^{(I_{N+1})}(\int_0^T \lambda_t \mathrm{d}W_t)\int_0^T (\int_{t_N}^T \widehat{l}_{N+1,s}\mathrm{d}s) l_{N,t_N}\int_0^{t_N}\cdots\int_0^{t_2} l_{1,t_1}\mathrm{d}W^{I_1}_{t_1}\ldots\mathrm{d}W^{I_N}_{t_N}\Big),
\end{aligned}$$

where at the last equality we have used the fact that $\int_0^T f_t Z_t \mathrm{d}t = \int_0^T (\int_t^T f_s \mathrm{d}s)\mathrm{d}Z_t$ for any continuous semimartingale Z starting from 0 and any measurable and bounded deterministic function f (apply the Itô's formula to the product $(\int_t^T f_s \mathrm{d}s)Z_t$). We easily conclude with the induction hypothesis and leave the details to the reader.

2.6.5 Applications of the expansions for time-independent CEV model

We specify in this section the results and the practical calculus of the various expansion coefficients when the volatility has the form:

$$\sigma(S) = \nu S^{\beta-1},$$

i.e. a CEV-type time-independent volatility with a level ν and a skew $\beta \leq 1$. Although the volatility and its derivatives are not bounded, we expect that our expansions can be generalized to that model. Alternatively, to fit our assumptions, we would need to modify the CEV volatility function σ near 0 and $+\infty$, so that the ellipticity and regularity conditions are met. The impact of such a modification has been studied in the case of Limited CEV model in [Andersen 2000] where the authors show a very small impact on prices. Observe in addition that the correction terms in our expansions do no depend on the modification of σ at 0 and $+\infty$.

To apply our different expansion theorems, we need to give the expressions of the coefficients $(C_i)_{1\leq i\leq 8}$ defined in Definition 2.3.1.1, in Theorem 2.3.2.1 and in Lemmas 2.4.0.1-2.4.0.2. A straightforward calculus leads to:

$$\begin{aligned}
a(x) &= \nu e^{x(\beta-1)}, & a^{(1)}(x) &= (\beta-1)a(x), & a^{(2)}(x) &= (\beta-1)^2 a(x),\\
a(x_0) &= \nu S_0^{\beta-1}, & a(k) &= \nu K^{\beta-1}, & a(x_{avg}) &= \nu(S_0 K)^{\frac{\beta-1}{2}},\\
\Sigma(S) &= \nu S^{\beta}, & \Sigma^{(1)}(S) &= \beta\frac{\Sigma(S)}{S}, & \Sigma^{(2)}(S) &= \beta(\beta-1)\frac{\Sigma(S)}{S^2},
\end{aligned}$$

$$\Sigma(S_0) = \nu S_0^{\beta}, \qquad \Sigma(K) = \nu K^{\beta}, \qquad \Sigma(S_{avg}) = \nu(\frac{S_0+K}{2})^{\beta}.$$

Thus for $\beta \in [0,1]$, the magnitudes of $\mathcal{M}_0(a)$ and $\mathcal{M}_1(a)$ are mainly linked to those of ν and $\nu(\beta-1)$. At the limit case $\beta = 1$, the model coincides with the log-normal proxy and $\mathcal{M}_1(a) = 0$. In the same spirit, ν and $\nu\beta$ are respectively linked to $\mathcal{M}_0(\Sigma)$ and $\mathcal{M}_1(\Sigma)$. At the limit case $\beta = 0$, the model coincides with the normal proxy and $\mathcal{M}_1(\Sigma) = 0$.

Finally, the expression of the coefficients $(C_i)_{1\le i\le 8}$ are:

$$\begin{aligned}
C_1(a;x)_0^T &= (\beta-1)a^4(x)\frac{T^2}{2}, & C_2(a;x)_0^T &= (\beta-1)^2a^4(x)T^2,\\
C_3(a;x)_0^T &= (\beta-1)^2a^6(x)\frac{T^3}{3}, & C_4(a;x)_0^T &= C_8(a;x)_0^T = (\beta-1)^2a^6(x)\frac{T^3}{6},\\
C_5(a;x)_0^T &= 2(\beta-1)^2a^2(x)T, & C_6(a;x)_0^T &= (\beta-1)^2a^4(x)\frac{T^2}{2},\\
C_7(a;x)_0^T &= (\beta-1)a^2(x)T, & &\\
C_1(\Sigma;S)_0^T &= \beta\frac{\Sigma^4(S)}{S}\frac{T^2}{2}, & C_2(\Sigma;S)_0^T &= \beta(2\beta-1)\frac{\Sigma^4(S)}{S^2}\frac{T^2}{2},\\
C_3(\Sigma;S)_0^T &= \beta(2\beta-1)\frac{\Sigma^6(S)}{S^2}\frac{T^3}{6}, & C_4(\Sigma;S)_0^T &= C_8(\Sigma;S)_0^T = \beta^2\frac{\Sigma^6(S)}{S^2}\frac{T^3}{6},\\
C_5(\Sigma;S)_0^T &= \beta(2\beta-1)\frac{\Sigma^2(S)}{S^2}T, & C_6(\Sigma;S)_0^T &= \beta^2\frac{\Sigma^4(S)}{S^2}\frac{T^2}{2},\\
C_7(\Sigma;S)_0^T &= \beta\frac{\Sigma^2(S)}{S}T, & &
\end{aligned}$$

where $x = x_0, k, x_{avg}$ and $S = S_0, K, S_{avg}$.

We now give the expressions of the coefficients γ_i, π_i, χ_i and Ξ_i defined in Definition 2.3.3.1 useful to compute the implied volatility expansions:

$$\begin{aligned}
\gamma_0(a;x)_0^T &= \frac{(\beta-1)^2}{24}a^3(x)T[1-\frac{a^2(x)T}{4}], & \gamma_1(a;x)_0^T &= \frac{(\beta-1)}{2}a(x),\\
\gamma_2(a;x)_0^T &= \frac{(\beta-1)^2}{12}a(x), & \pi_0(a;x)_0^T &= \gamma_0(a;x)_0^T,\\
\pi_1(a;x)_0^T &= 0, & \pi_2(a;x)_0^T &= -\frac{(\beta-1)^2}{24}a(x),\\
\chi_1(\Sigma;S)_0^T &= \frac{\beta\Sigma(S)}{2S}, & \chi_0(\Sigma;S)_0^T &= \frac{\beta(\beta-2)}{24S^2}\Sigma^3(S)T,\\
\chi_2(\Sigma;S)_0^T &= \frac{\beta(\beta-2)}{12S^2}\Sigma(S), & \Xi_0(\Sigma;S)_0^T &= \chi_0(\Sigma;S)_0^T,\\
\Xi_2(\Sigma;S)_0^T &= -\frac{\beta(\beta+1)}{24S^2}\Sigma(S), & \Xi_1(\Sigma;S)_0^T &= 0.
\end{aligned}$$

For example, the second and third order Black-Scholes and Bachelier implied volatility expansions based on the mid-points are explicitly given by:

$$\begin{aligned}
&\sigma_{\mathrm{I}}(x_0,T,k) \approx \nu(S_0K)^{\frac{\beta-1}{2}},\\
&\sigma_{\mathrm{I}}(x_0,T,k) \approx \nu(S_0K)^{\frac{\beta-1}{2}}[1+\frac{(\beta-1)^2\nu^2T}{24}(S_0K)^{\beta-1}(1-\frac{\nu^2T(S_0K)^{\beta-1}}{4})-\frac{(\beta-1)^2}{24}\log^2(\frac{S_0}{K})],\\
&\Sigma_{\mathrm{I}}(S_0,T,K) \approx \nu(\frac{S_0+K}{2})^{\beta},\\
&\Sigma_{\mathrm{I}}(S_0,T,K) \approx \nu(\frac{S_0+K}{2})^{\beta}[1+\frac{\beta(\beta-2)\nu^2T}{24}(\frac{S_0+K}{2})^{2\beta-2}-\frac{\beta(\beta+1)}{6}(\frac{S_0-K}{S_0+K})^2],
\end{aligned}$$

which are very simple formulas. The last formula coincides with the intermediate equation (A.28b) in [Hagan 1999].

Table 2.7: CEV model ($\beta = 0.8$, $\nu = 0.25$): errors in bps on the BS implied volatility using the 6 second order price approximations ImpVol(AppPriceLN(2,x_0)), ImpVol(AppPriceLN(2,k)), ImpVol(AppPriceLN(2,x_{avg})), ImpVol(AppPriceN(2,S_0)), ImpVol(AppPriceN(2,K)) and ImpVol(AppPriceN(2,S_{avg})).

3M	−12	−6	−2	−1	0	0	0	0	0	−1	−3	−5	−8
	−17	−7	−3	−1	0	0	0	0	0	−1	−2	−4	−7
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−577	−79	−18	−1	2	2	2	2	2	0	−22	−44	−73
	−124	−59	−20	−2	2	2	2	2	2	1	−21	−53	−125
	21	14	9	6	3	2	2	2	3	4	9	11	15
6M	−13	−3	−1	−1	0	0	0	0	0	−1	−2	−4	−15
	−17	−4	−2	−1	0	0	0	0	0	−1	−2	−4	−11
	1	0	0	0	0	0	0	0	0	0	0	0	1
	−269	−18	0	5	5	4	3	4	4	1	−5	−35	−117
	−138	−26	−4	4	5	4	3	4	4	3	−2	−30	−238
	33	16	11	7	5	4	3	3	4	8	10	16	26
1Y	−23	−8	−2	−1	0	0	0	0	0	−1	−4	−8	−37
	−34	−9	−2	−1	0	0	0	0	0	−1	−4	−7	−23
	1	1	0	0	0	0	0	0	0	0	0	0	1
	−848	−48	6	10	8	7	6	7	9	6	−21	−59	−235
	−240	−69	−1	8	8	7	6	7	9	9	−9	−45	ND
	64	36	20	14	8	7	6	7	9	13	22	29	54
1.5Y	−28	−11	−3	−2	−1	0	0	0	−1	−2	−6	−12	−50
	−41	−12	−4	−2	−1	0	0	0	−1	−2	−5	−10	−30
	2	1	0	0	0	0	0	0	0	0	0	1	2
	−644	−50	10	16	14	10	9	11	12	10	−31	−95	−299
	−291	−90	−4	10	14	10	9	11	12	13	−9	−71	ND
	89	52	30	23	14	10	9	11	12	19	32	44	74
2Y	−36	−14	−5	−2	−1	−1	−1	−1	−1	−2	−6	−18	−91
	−56	−17	−6	−2	−1	−1	−1	−1	−1	−2	−6	−14	−44
	2	1	0	0	0	0	−1	0	0	0	0	1	2
	ND	−65	12	22	18	15	13	14	17	13	−27	−138	−418
	−373	−129	−13	18	18	15	13	14	17	18	−1	−107	ND
	119	72	44	27	17	15	13	14	17	25	39	60	105
3Y	−64	−18	−11	−3	−1	−1	−1	−1	−1	−3	−10	−27	−141
	−122	−21	−13	−3	−1	−1	−1	−1	−1	−3	−9	−21	−57
	4	1	1	0	−1	−1	−1	−1	0	0	1	1	3
	ND	−43	6	35	28	22	19	20	25	17	−62	−208	−534
	−638	−154	−68	27	28	22	19	20	25	27	−8	−159	−2260
	208	102	81	41	28	21	19	20	25	37	61	87	146
5Y	−106	−31	−13	−6	−2	−1	−1	−1	−2	−5	−17	−45	−472
	−256	−38	−14	−6	−2	−1	−1	−1	−2	−5	−15	−32	−88
	7	2	1	0	−1	−1	−1	−1	−1	0	1	2	5
	ND	−41	53	64	49	39	32	35	41	25	−116	−334	−753
	−1000	−295	−55	34	48	39	32	35	43	47	−4	−249	ND
	377	183	119	80	49	38	32	34	43	60	99	140	233
10Y	−172	−69	−30	−10	−4	−3	−2	−2	−4	−9	−35	−103	ND
	ND	−95	−34	−10	−4	−3	−2	−2	−4	−9	−28	−61	−159
	15	6	2	0	−2	−2	−2	−2	−1	0	2	5	10
	ND	40	158	146	109	82	67	70	80	32	−271	−625	−1100
	−1531	−762	−232	75	103	82	67	70	85	95	16	−781	ND
	786	451	289	166	108	80	67	68	84	120	192	267	439

Table 2.8: CEV model ($\beta = 0.8$, $\nu = 0.25$): errors in bps on the BS implied volatility using the 6 second order implied volatility approximations AppImpVolLN(2,x_0), AppImpVolLN(2,k), AppImpVolLN(2,x_{avg}), AppImpVolN(2,S_0), AppImpVolN(2,K) and AppImpVolN(2,S_{avg}).

3M	−1	−1	0	0	0	0	0	0	0	0	0	−1	−1
	−1	−1	−1	0	0	0	0	0	0	0	0	−1	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
	27	18	12	7	4	2	2	2	3	5	11	15	19
	29	19	12	7	4	2	2	2	3	5	11	15	18
	21	14	9	6	3	2	2	2	3	4	9	11	15
6M	−2	−1	−1	0	0	0	0	0	0	0	−1	−1	−1
	−2	−1	−1	0	0	0	0	0	0	0	−1	−1	−1
	1	0	0	0	0	0	0	0	0	0	0	0	1
	41	20	13	9	5	4	3	4	5	10	13	21	36
	44	21	14	9	6	4	3	4	5	9	12	20	33
	33	16	11	7	5	4	3	3	4	8	10	16	26
1Y	−3	−2	−1	−1	0	0	0	0	0	−1	−1	−2	−3
	−4	−2	−1	−1	0	0	0	0	0	−1	−1	−1	−3
	1	1	0	0	0	0	0	0	0	0	0	0	1
	79	44	24	17	9	7	6	7	10	16	28	39	74
	88	48	25	17	9	7	6	7	10	15	27	36	67
	64	36	20	14	8	7	6	7	9	13	22	29	54
1.5Y	−4	−3	−1	−1	−1	0	0	0	−1	−1	−2	−2	−4
	−5	−3	−2	−1	−1	0	0	0	−1	−1	−2	−2	−4
	2	1	0	0	0	0	0	0	0	0	0	1	2
	108	64	36	27	15	10	9	11	13	23	42	59	104
	123	69	38	28	15	10	9	11	13	22	39	54	91
	89	52	30	23	14	10	9	11	12	19	32	44	74
2Y	−6	−4	−2	−1	−1	−1	−1	−1	−1	−1	−2	−3	−6
	−7	−4	−2	−1	−1	−1	−1	−1	−1	−1	−2	−3	−5
	2	1	0	0	0	0	−1	0	0	0	0	1	2
	145	87	52	31	19	15	13	14	19	30	50	80	149
	167	97	55	32	19	15	13	14	18	29	47	72	127
	119	72	44	27	17	15	13	14	17	25	39	60	105
3Y	−10	−5	−4	−2	−1	−1	−1	−1	−1	−2	−3	−5	−9
	−12	−6	−4	−2	−1	−1	−1	−1	−1	−2	−3	−4	−7
	4	1	1	0	−1	−1	−1	−1	0	0	1	1	3
	249	121	96	47	31	22	19	20	28	45	80	120	212
	301	136	105	49	31	22	19	20	28	43	73	106	176
	208	102	81	41	28	21	19	20	25	37	61	87	146
5Y	−18	−9	−5	−4	−2	−1	−1	−1	−2	−3	−5	−8	−14
	−23	−10	−6	−4	−2	−1	−1	−1	−2	−3	−5	−7	−11
	7	2	1	0	−1	−1	−1	−1	−1	0	1	2	5
	443	216	140	92	52	39	32	35	48	74	132	196	352
	571	252	155	98	53	40	32	35	47	69	117	167	277
	377	183	119	80	49	38	32	34	43	60	99	140	233
10Y	−33	−19	−12	−7	−4	−3	−2	−2	−3	−6	−10	−16	−29
	−47	−24	−14	−7	−4	−3	−2	−2	−3	−5	−9	−12	−21
	15	6	2	0	−2	−2	−2	−2	−1	0	2	5	10
	904	522	333	188	117	82	67	70	94	149	264	394	725
	1289	660	390	203	120	83	67	70	91	137	224	313	510
	786	451	289	166	108	80	67	68	84	120	192	267	439

Table 2.9: CEV model ($\beta = 0.8$, $\nu = 0.25$): errors in bps on the BS implied volatility using the 6 third order price approximations `ImpVol(AppPriceLN(3,`x_0`))`, `ImpVol(AppPriceLN(3,k))`, `ImpVol(AppPriceLN(3,`x_{avg}`))`, `ImpVol(AppPriceN(3,`S_0`))`, `ImpVol(AppPriceN(3,K))` and `ImpVol(AppPriceN(3,`S_{avg}`))`.

3M	−1	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	6	−4	−2	−1	0	0	0	0	0	0	2	0	−6
	−22	−1	2	1	0	0	0	0	0	0	−2	−3	−2
	−1	0	0	0	0	0	0	0	0	0	0	0	0
6M	−1	0	0	0	0	0	0	0	0	0	0	0	1
	1	0	0	0	0	0	0	0	0	0	0	0	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−17	−4	−1	0	0	0	0	0	0	1	2	4	−12
	−13	4	2	0	0	0	0	0	0	−1	−1	−5	−12
	−1	0	0	0	0	0	0	0	0	0	0	0	−1
1Y	1	1	0	0	0	0	0	0	0	0	0	0	1
	−2	0	0	0	0	0	0	0	0	0	0	0	3
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−62	−15	−2	−1	0	0	0	0	0	1	6	9	−56
	−34	13	3	1	0	0	0	0	0	−1	−5	−12	−45
	−3	0	0	0	0	0	0	0	0	0	0	0	−3
1.5Y	−2	0	0	0	0	0	0	0	0	0	0	0	4
	3	0	0	0	0	0	0	0	0	0	0	0	−3
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−100	−23	−4	−2	0	0	0	0	0	2	11	14	−83
	−36	21	7	3	0	0	0	0	0	−1	−8	−23	−98
	−4	−1	0	0	0	0	0	0	0	0	0	−1	−5
2Y	−3	−1	0	0	0	0	0	0	0	0	0	1	10
	4	1	0	0	0	0	0	0	0	0	0	−1	−7
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−163	−36	−8	−1	0	0	0	0	0	3	14	17	−177
	−63	31	14	2	0	0	0	0	0	−2	−9	−38	−243
	−7	−1	0	0	0	0	0	0	0	0	0	−1	−12
3Y	−9	−1	0	0	0	0	0	0	0	0	0	1	16
	11	1	0	0	0	0	0	0	0	0	0	−1	−10
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−818	−48	−23	−2	−1	0	0	0	1	5	29	16	−275
	−250	52	41	4	1	0	0	0	−1	−3	−19	−69	ND
	−24	−2	−1	0	0	0	0	0	0	0	−1	−3	−23
5Y	−18	−1	0	0	0	0	0	0	0	0	1	2	39
	23	1	0	0	0	0	0	0	0	0	−1	−2	−21
	0	0	0	0	0	0	0	0	0	0	0	0	0
	ND	−103	−26	−7	−2	−1	0	0	1	12	57	−10	−522
	−597	90	65	17	1	1	0	−1	−2	−5	−38	−143	ND
	−73	−5	−1	−1	0	0	0	0	0	0	−2	−7	−76
10Y	−34	−5	−1	0	0	0	0	0	0	0	1	7	147
	28	3	1	0	0	0	0	0	0	0	−1	−7	−59
	0	0	0	0	0	0	0	0	0	0	0	0	0
	ND	−340	−83	−16	−6	−4	−1	1	4	42	109	−222	−987
	−1200	−42	230	49	5	1	−1	−3	−4	−13	−94	−588	ND
	−214	−29	−7	−3	−2	−1	−1	−1	−1	−1	−6	−28	ND

Table 2.10: CEV model ($\beta = 0.8$, $\nu = 0.25$): errors in bps on the BS implied volatility using the 6 third order implied volatility approximations `AppImpVolLN(3,`x_0`)`, `AppImpVolLN(3,k)`, `AppImpVolLN(3,`x_{avg}`)`, `AppImpVolN(3,S0)`, `AppImpVolN(3,K)` and `AppImpVolN(3,`S_{avg}`)`.

3M	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	4	2	1	0	0	0	0	0	0	0	−1	−2	−3
	−5	−3	−1	0	0	0	0	0	0	0	1	1	2
	0	0	0	0	0	0	0	0	0	0	0	0	0
6M	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	7	2	1	0	0	0	0	0	0	−1	−1	−3	−8
	−10	−3	−1	0	0	0	0	0	0	0	1	2	5
	0	0	0	0	0	0	0	0	0	0	0	0	0
1Y	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	17	6	2	1	0	0	0	0	0	−1	−4	−7	−25
	−30	−10	−2	−1	0	0	0	0	0	1	3	5	14
	1	0	0	0	0	0	0	0	0	0	0	0	1
1.5Y	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	25	10	3	2	0	0	0	0	0	−2	−7	−15	−43
	−49	−17	−5	−2	0	0	0	0	0	1	5	9	22
	2	1	0	0	0	0	0	0	0	0	0	1	2
2Y	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	37	16	6	1	0	0	0	0	0	−2	−9	−24	−79
	−80	−29	−9	−2	0	0	0	0	0	2	6	14	35
	4	1	0	0	0	0	0	0	0	0	0	1	4
3Y	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	77	23	15	2	0	0	0	0	−1	−5	−20	−47	−143
	−214	−48	−28	−4	0	0	0	0	0	3	11	23	55
	12	2	1	0	0	0	0	0	0	0	1	2	8
5Y	−1	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	160	50	22	7	0	−1	0	0	−2	−10	−47	−108	−350
	−618	−130	−47	−15	−1	0	0	−1	0	6	22	43	102
	36	7	2	0	0	0	0	0	0	0	2	6	21
10Y	−1	−1	−1	0	0	0	0	0	0	0	1	1	1
	2	1	1	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	362	155	68	16	0	−3	−1	1	−4	−34	−153	−368	−1307
	−2195	−631	−216	−44	−6	1	−1	−3	0	14	53	102	231
	113	34	10	0	−1	−1	−1	−1	−1	1	8	21	70

Table 2.11: CEV model ($\beta = 0.8$, $\nu = 0.25$): errors in bps on the BS implied volatility using the 4 approximations `Av.ImpVol(AppPriceLN(3,.))`, `Av.AppImpVolLN(3,.)`, (HF), and (HLF).

3M	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
6M	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
1Y	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−1	0	0	0	0	0	0	0	0	0	0	0	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
1.5Y	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−2	−1	0	0	0	0	0	0	0	0	0	0	−2
	0	0	0	0	0	0	0	0	0	0	0	0	0
2Y	0	0	0	0	0	0	0	0	0	0	0	0	2
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−3	−1	0	0	0	0	0	0	0	0	0	−1	−3
	0	0	0	0	0	0	0	0	0	0	0	0	0
3Y	1	0	0	0	0	0	0	0	0	0	0	0	3
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−9	−2	−1	0	0	0	0	0	0	0	−1	−2	−6
	0	0	0	0	0	0	0	0	0	0	0	0	0
5Y	2	0	0	0	0	0	0	0	0	0	0	0	9
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−24	−5	−2	0	0	0	0	0	0	0	−2	−4	−14
	0	0	0	0	0	0	0	0	0	0	0	0	0
10Y	−3	−1	0	0	0	0	0	0	0	0	0	0	44
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−69	−24	−8	−1	0	0	0	0	0	−1	−6	−14	−44
	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 2.12: CEV model ($\beta = 0.8$, $\nu = 0.25$): average per categories of strikes of errors in bps on the BS implied volatility using the 6 approximations `Av.ImpVol(AppPriceLN(3,.))`, `ImpVol(AppPriceLN(3,`x_{avg}`))`, `Av.AppImpVolLN(3,.)`, `AppImpVolLN(3,`x_{avg}`)`, (`HF`), and (`HLF`).

	far ITM	ITM	ATM	OTM	far OTM
3M	6.7E−2	1.4E−4	−7.9E−6	1.8E−4	1.5E−2
	−2.4E−4	−1.4E−5	−1.3E−5	−1.3E−5	−9.9E−5
	4.3E−4	9.1E−5	−8.2E−6	7.3E−5	2.6E−4
	−3.4E−5	−9.6E−6	−1.3E−5	−8.7E−6	−1.7E−5
	−9.1E−2	−8.6E−3	2.0E−4	−7.3E−3	−4.8E−2
	−6.4E−5	1.7E−4	2.4E−4	1.5E−4	−9.8E−6
6M	3.6E−2	1.6E−4	−4.3E−5	1.5E−4	3.2E−2
	−3.0E−4	−4.3E−5	−5.3E−5	−3.6E−5	−2.4E−4
	1.0E−3	1.5E−4	−4.3E−5	1.4E−4	7.8E−4
	−6.5E−5	−4.3E−5	−5.3E−5	−3.5E−5	−4.9E−5
	−1.6E−1	−7.9E−3	9.2E−4	−8.6E−3	−1.3E−1
	2.5E−4	8.6E−4	9.5E−4	7.1E−4	1.3E−4
1Y	9.6E−2	5.1E−4	−1.9E−4	4.4E−4	1.9E−1
	−1.1E−3	−1.8E−4	−2.1E−4	−1.4E−4	−1.1E−3
	4.2E−3	4.3E−4	−1.9E−4	5.5E−4	3.0E−3
	−2.5E−4	−1.8E−4	−2.1E−4	−1.4E−4	−2.1E−4
	−6.3E−1	−2.2E−2	3.8E−3	−3.6E−2	−5.1E−1
	1.2E−3	3.7E−3	3.8E−3	2.8E−3	4.4E−4
1.5Y	9.2E−2	1.2E−3	−3.9E−4	7.5E−4	2.8E−1
	−1.8E−3	−4.0E−4	−4.5E−4	−3.2E−4	−2.1E−3
	8.5E−3	1.1E−3	−3.9E−4	1.1E−3	6.1E−3
	−4.9E−4	−3.9E−4	−4.5E−4	−3.0E−4	−4.0E−4
	−1.2E+0	−5.5E−2	8.3E−3	−7.3E−2	−9.9E−1
	3.9E−3	8.4E−3	8.6E−3	6.2E−3	1.1E−3
2Y	1.4E−1	1.7E−3	−6.4E−4	1.4E−3	8.7E−1
	−3.1E−3	−7.1E−4	−7.9E−4	−5.3E−4	−4.7E−3
	1.5E−2	1.8E−3	−6.4E−4	1.7E−3	1.1E−2
	−8.9E−4	−7.1E−4	−7.9E−4	−5.1E−4	−8.1E−4
	−2.0E+0	−9.6E−2	1.5E−2	−1.0E−1	−1.9E+0
	7.5E−3	1.5E−2	1.5E−2	1.1E−2	1.2E−3
3Y	5.8E−1	2.5E−3	−1.5E−3	1.9E−3	1.5E+0
	−9.9E−3	−1.6E−3	−1.7E−3	−1.1E−3	−9.4E−3
	3.7E−2	5.3E−3	−1.5E−3	3.8E−3	2.2E−2
	−2.4E−3	−1.5E−3	−1.7E−3	−1.1E−3	−1.6E−3
	−5.2E+0	−3.5E−1	3.4E−2	−2.6E−1	−3.8E+0
	1.8E−2	3.5E−2	3.5E−2	2.4E−2	3.1E−3
5Y	1.2E+0	6.2E−3	−3.6E−3	1.2E−3	4.5E+0
	−3.0E−2	−4.0E−3	−4.2E−3	−2.7E−3	−2.7E−2
	1.0E−1	1.2E−2	−3.5E−3	9.4E−3	5.4E−2
	−7.2E−3	−3.9E−3	−4.3E−3	−2.5E−3	−4.4E−3
	−1.5E+1	−6.9E−1	9.4E−2	−6.7E−1	−9.1E+0
	6.5E−2	1.1E−1	9.7E−2	6.3E−2	7.6E−3
10Y	−2.1E+0	−4.6E−2	−1.1E−2	−3.2E−2	2.2E+1
	−9.7E−2	−1.1E−2	−1.2E−2	−7.5E−3	−1.3E−1
	2.8E−1	3.8E−2	−1.1E−2	2.9E−2	1.7E−1
	−2.3E−2	−1.1E−2	−1.3E−2	−6.9E−3	−1.8E−2
	−4.6E+1	−3.1E+0	3.8E−1	−2.4E+0	−2.9E+1
	4.4E−1	4.8E−1	4.0E−1	2.3E−1	2.2E−2

Table 2.13: CEV model ($\beta = 0.2$, $\nu = 0.25$): errors in bps on the BS implied volatility using the 6 approximations ImpVol(AppPriceLN(3,x_{avg})), AppImpVolLN(3,x_{avg}), ImpVol(AppPriceN(3,S_{avg})), AppImpVolN(3,S_{avg}), (HF) and (HLF).

3M	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−1	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
6M	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−1	0	0	0	0	0	0	0	0	0	0	0	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
1Y	0	0	0	0	0	0	0	0	0	0	0	0	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0
	−5	−1	0	0	0	0	0	0	0	0	0	−1	−2
	0	0	0	0	0	0	0	0	0	0	0	0	0
1.5Y	−1	0	0	0	0	0	0	0	0	0	0	0	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	0	0	0	0	0	0	0	0	0	0	0
	−9	−3	−1	0	0	0	0	0	0	0	−1	−1	−4
	−1	0	0	0	0	0	0	0	0	0	0	0	0
2Y	−1	0	0	0	0	0	0	0	0	0	0	0	−3
	−1	0	0	0	0	0	0	0	0	0	0	0	0
	2	1	0	0	0	0	0	0	0	0	0	0	0
	2	1	1	0	0	0	0	0	0	0	0	0	1
	−15	−5	−1	0	0	0	0	0	0	0	−1	−3	−8
	−1	−1	0	0	0	0	0	0	0	0	0	0	0
3Y	−4	−1	−1	−1	−1	−1	0	0	0	0	0	−1	−8
	−2	−1	−1	−1	−1	−1	0	0	0	0	0	0	−1
	7	2	2	1	0	0	0	0	0	0	0	0	0
	8	3	2	1	0	0	0	0	0	0	0	1	1
	−45	−9	−5	−1	0	0	0	0	0	−1	−2	−5	−15
	−4	−1	−1	0	0	0	0	0	0	0	0	0	0
5Y	2	−1	−2	−2	−2	−2	−1	−1	−1	0	0	−2	−37
	7	−1	−2	−2	−2	−2	−1	−1	−1	0	0	0	−1
	47	13	6	4	2	1	1	1	1	1	1	1	0
	50	13	7	4	2	1	1	1	1	1	1	1	3
	−117	−26	−9	−3	0	0	0	0	0	−1	−6	−12	−31
	4	0	0	0	0	0	0	0	0	0	−1	−1	−1
10Y	148	84	41	12	2	−2	−3	−2	−2	−1	−2	−13	ND
	159	85	41	12	2	−2	−3	−2	−2	−1	−1	−1	−5
	530	221	109	45	21	11	6	3	2	2	2	2	−8
	541	224	111	45	21	11	6	3	2	2	2	3	6
	−112	−6	18	18	12	8	4	2	−1	−5	−17	−33	−73
	271	123	65	29	14	8	4	2	0	−1	−1	−1	−1

Table 2.14: CEV model ($\beta = 0.2$, $\nu = 0.25$): average per categories of strikes of errors in bps on the BS implied volatility using the 6 approximations `ImpVol(AppPriceLN(3,`x_{avg}`))`, `AppImpVolLN(3,`x_{avg}`)`, `ImpVol(AppPriceN(3,`S_{avg}`))`, `AppImpVolN(3,`S_{avg}`)`, (HF) and (HLF).

	far ITM	ITM	ATM	OTM	far OTM
3M	−5.4E−2	−4.0E−3	−3.4E−3	−3.3E−3	−2.9E−2
	−9.3E−3	−3.1E−3	−3.4E−3	−1.9E−3	−4.1E−3
	2.5E−2	6.3E−3	1.8E−3	3.7E−3	8.9E−3
	3.9E−2	7.2E−3	1.8E−3	4.5E−3	1.6E−2
	−3.8E−1	−3.9E−2	7.9E−5	−2.9E−2	−1.6E−1
	−2.1E−2	−4.6E−3	3.2E−4	−2.6E−3	−7.9E−3
6M	−6.9E−2	−1.5E−2	−1.4E−2	−7.5E−3	−7.4E−2
	−2.1E−2	−1.4E−2	−1.4E−2	−7.2E−3	−1.1E−2
	6.8E−2	1.6E−2	6.9E−3	9.6E−3	2.5E−2
	9.0E−2	1.7E−2	6.9E−3	1.0E−2	4.3E−2
	−7.2E−1	−4.5E−2	1.5E−3	−3.9E−2	−4.3E−1
	−5.6E−2	−7.4E−3	1.8E−3	−5.4E−3	−2.2E−2
1Y	−2.5E−1	−6.6E−2	−5.5E−2	−2.8E−2	−3.8E−1
	−9.9E−2	−6.6E−2	−5.5E−2	−2.6E−2	−4.6E−2
	3.2E−1	6.4E−2	2.8E−2	3.5E−2	7.8E−2
	4.1E−1	6.7E−2	2.8E−2	3.8E−2	1.5E−1
	−3.0E+0	−1.4E−1	8.2E−3	−1.5E−1	−1.6E+0
	−2.6E−1	−2.3E−2	8.8E−3	−1.9E−2	−7.2E−2
1.5Y	−4.4E−1	−1.6E−1	−1.2E−1	−6.2E−2	−7.4E−1
	−2.4E−1	−1.6E−1	−1.2E−1	−5.6E−2	−8.8E−2
	7.4E−1	1.7E−1	6.2E−2	7.1E−2	1.5E−1
	8.8E−1	1.8E−1	6.2E−2	7.8E−2	2.9E−1
	−5.6E+0	−3.5E−1	1.7E−2	−3.0E−1	−3.0E+0
	−5.6E−1	−5.8E−2	1.9E−2	−3.5E−2	−1.3E−1
2Y	−8.0E−1	−3.1E−1	−2.2E−1	−9.6E−2	−1.9E+0
	−4.9E−1	−3.1E−1	−2.2E−1	−9.0E−2	−1.8E−1
	1.5E+0	3.2E−1	1.2E−1	1.1E−1	2.1E−1
	1.7E+0	3.3E−1	1.2E−1	1.2E−1	5.1E−1
	−1.0E+1	−6.1E−1	3.0E−2	−4.3E−1	−5.5E+0
	−1.1E+0	−9.2E−2	3.6E−2	−5.5E−2	−2.1E−1
3Y	−2.3E+0	−8.0E−1	−4.8E−1	−2.0E−1	−4.3E+0
	−1.4E+0	−8.0E−1	−4.8E−1	−1.7E−1	−3.5E−1
	4.8E+0	9.8E−1	2.8E−1	2.4E−1	3.4E−1
	5.4E+0	1.0E+0	2.8E−1	2.6E−1	9.3E−1
	−2.7E+1	−2.1E+0	8.9E−2	−1.0E+0	−9.9E+0
	−2.7E+0	−2.7E−1	9.7E−2	−1.1E−1	−3.7E−1

Table 2.15: CEV model ($\beta = 0.5$, $\nu = 0.4$): errors in bps on the BS implied volatility using the 4 approximations ImpVol(AppPriceLN(3,x_{avg})), AppImpVolLN(3,x_{avg}), (HF) and (HLF).

3M	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−1	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
6M	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−1	0	0	0	0	0	0	0	0	0	0	0	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
1Y	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−4	−1	0	0	0	0	0	0	0	0	0	−1	−3
	0	0	0	0	0	0	0	0	0	0	0	0	0
1.5Y	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−8	−2	0	0	0	0	0	0	0	0	−1	−2	−5
	0	0	1	1	0	0	0	0	0	0	0	0	0
2Y	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	−13	−4	0	1	1	1	1	1	1	0	−1	−3	−10
	1	1	1	1	1	1	1	1	1	0	0	0	0
3Y	−1	−1	−1	−1	−1	−1	−1	−1	0	0	0	0	0
	−1	−1	−1	−1	−1	−1	−1	−1	0	0	0	0	0
	−37	−5	−2	2	2	2	2	2	1	0	−2	−6	−19
	3	3	3	3	2	2	2	2	1	1	0	0	0
5Y	6	1	0	−1	−1	−1	−1	−1	−1	−1	0	0	−1
	6	1	0	−1	−1	−1	−1	−1	−1	−1	0	0	0
	−88	−9	4	7	8	7	6	5	3	1	−5	−14	−43
	25	15	12	10	8	7	6	5	3	2	1	0	−1
10Y	92	61	40	22	13	8	4	2	1	0	0	−1	−8
	91	61	40	22	13	8	4	2	1	0	0	0	−1
	−58	54	76	69	54	42	31	23	15	4	−17	−44	−118
	286	173	120	79	56	42	31	23	16	9	4	1	−2

Table 2.16: CEV model ($\beta = 0.5$, $\nu = 0.4$): average per categories of strikes of errors in bps on the BS implied volatility using the 4 approximations `ImpVol(AppPriceLN(3,`x_{avg}`))`, `AppImpVolLN(3,`x_{avg}`)`, (`HF`) and (`HLF`).

	far ITM	ITM	ATM	OTM	far OTM
3M	−6.9E−3	−5.3E−3	−5.4E−3	−3.6E−3	−3.2E−3
	−4.6E−3	−5.3E−3	−5.4E−3	−3.6E−3	−2.2E−3
	−3.8E−1	−2.9E−2	1.1E−2	−2.4E−2	−1.8E−1
	−5.9E−3	8.0E−3	1.1E−2	5.4E−3	−3.0E−3
6M	−2.3E−2	−2.3E−2	−2.1E−2	−1.5E−2	−1.1E−2
	−2.1E−2	−2.3E−2	−2.1E−2	−1.5E−2	−8.3E−3
	−6.8E−1	4.2E−3	4.4E−2	−1.1E−2	−4.8E−1
	1.1E−2	4.4E−2	4.5E−2	2.8E−2	−3.4E−3
1Y	−9.9E−2	−9.5E−2	−8.2E−2	−5.3E−2	−4.2E−2
	−9.5E−2	−9.4E−2	−8.2E−2	−5.3E−2	−2.8E−2
	−2.6E+0	8.3E−2	1.8E−1	−5.3E−2	−1.8E+0
	8.7E−2	2.0E−1	1.8E−1	1.0E−1	−1.5E−2
1.5Y	−2.3E−1	−2.1E−1	−1.7E−1	−1.1E−1	−7.9E−2
	−2.2E−1	−2.1E−1	−1.7E−1	−1.1E−1	−5.2E−2
	−4.7E+0	2.0E−1	4.1E−1	−7.7E−2	−3.5E+0
	3.4E−1	5.0E−1	4.1E−1	2.4E−1	−2.5E−2
2Y	−4.0E−1	−3.6E−1	−2.9E−1	−1.8E−1	−1.5E−1
	−3.9E−1	−3.6E−1	−2.9E−1	−1.8E−1	−8.1E−2
	−8.2E+0	4.3E−1	7.6E−1	−3.7E−2	−6.7E+0
	8.1E−1	9.7E−1	7.7E−1	4.2E−1	−6.8E−2
3Y	−7.0E−1	−7.2E−1	−5.7E−1	−3.3E−1	−3.0E−1
	−6.6E−1	−7.2E−1	−5.8E−1	−3.3E−1	−1.4E−1
	−2.1E+1	7.9E−1	1.8E+0	−2.2E−1	−1.3E+1
	3.0E+0	2.6E+0	1.8E+0	8.8E−1	−1.2E−1

Table 2.17: CEV model ($\beta = 0.8$, $\nu = 0.25$): errors in bps on the deltas using the 6 approximations `AppDeltaLN(1,`x_0`)`, `AppDeltaLN(1,k)`, `AppDeltaLN(1,`x_{avg}`)`, `AppDeltaLN(2,`x_0`)`, `AppDeltaLN(2,k)` and `AppDeltaLN(2,`x_{avg}`)`.

3M	1	2	2	1	1	0	0	0	−1	−1	−2	−2	−2
	1	0	0	−1	−1	0	0	0	0	1	0	0	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
6M	3	3	2	1	1	1	0	−1	−1	−2	−3	−4	−4
	1	−1	−1	−1	−1	0	0	0	1	1	1	0	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
1Y	5	5	3	3	1	1	0	−1	−2	−4	−6	−8	−9
	1	−2	−3	−2	−1	0	0	0	1	2	2	1	−3
	0	0	0	1	0	0	0	0	−1	−1	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
1.5Y	7	7	5	4	2	1	0	−2	−3	−5	−9	−13	−15
	0	−3	−4	−3	−2	0	0	1	1	3	3	1	−4
	0	0	1	1	1	0	0	−1	−1	−1	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	1
	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
2Y	9	9	7	5	3	2	0	−2	−4	−7	−12	−18	−20
	0	−4	−6	−4	−2	−1	0	1	2	4	5	0	−6
	0	0	1	1	1	1	0	−1	−1	−1	−1	0	0
	1	0	0	0	0	0	0	0	0	0	0	1	2
	1	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
3Y	12	12	10	6	4	2	0	−3	−6	−11	−19	−28	−32
	0	−8	−8	−6	−3	−1	0	1	3	6	6	−1	−9
	0	0	1	1	1	1	0	−1	−2	−2	−1	0	1
	1	0	0	0	0	0	0	0	0	0	1	1	4
	1	1	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	0	0	0
5Y	16	17	13	10	7	4	0	−5	−10	−18	−34	−51	−61
	−3	−13	−14	−10	−5	−2	0	1	5	10	9	−4	−15
	0	1	2	2	2	1	0	−2	−3	−3	−1	0	1
	3	1	1	0	0	0	0	0	0	1	1	3	10
	3	1	0	0	0	0	0	0	0	0	1	2	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
10Y	24	26	21	16	12	7	−1	−8	−19	−37	−76	−118	−145
	−16	−26	−26	−17	−9	−4	−1	2	8	20	13	−17	−23
	−1	1	3	4	4	2	−1	−3	−5	−6	−2	1	2
	4	2	1	1	1	1	1	1	1	2	3	8	36
	6	2	0	0	0	1	1	1	0	0	3	5	−6
	0	0	0	0	0	1	1	1	1	0	0	−1	−1

Table 2.18: CEV model ($\beta = 0.8$, $\nu = 0.25$): average per categories of strikes of errors in bps on the deltas using the 2 approximations `AppDeltaLN(1,`x_{avg}`)` and `AppDeltaLN(2,`x_{avg}`)`.

	far ITM	ITM	ATM	OTM	far OTM
3M	−3.1E−2	7.1E−2	−1.9E−3	−8.0E−2	4.0E−2
	−1.8E−3	2.4E−4	3.6E−3	4.7E−4	−2.4E−3
6M	−1.1E−2	2.3E−1	−6.0E−3	−2.3E−1	4.8E−2
	−4.6E−3	5.0E−3	1.1E−2	4.4E−3	−6.1E−3
1Y	−4.0E−2	4.5E−1	−1.9E−2	−4.5E−1	7.9E−2
	−1.4E−2	1.7E−2	3.0E−2	1.2E−2	−1.6E−2
1.5Y	−1.9E−2	6.8E−1	−1.2E−1	−6.6E−1	1.4E−1
	−2.4E−2	2.8E−2	5.3E−2	2.4E−2	−3.2E−2
2Y	−2.4E−2	8.7E−1	−4.8E−2	−9.6E−1	2.1E−1
	−3.6E−2	4.4E−2	7.9E−2	3.9E−2	−4.9E−2
3Y	3.9E−2	1.2E+0	−9.4E−2	−1.4E+0	3.6E−1
	−5.7E−2	6.7E−2	1.4E−1	6.2E−2	−9.5E−2
5Y	6.3E−2	2.0E+0	−1.9E−1	−2.3E+0	6.5E−1
	−1.2E−1	1.4E−1	2.8E−1	1.2E−1	−2.1E−1
10Y	8.9E−2	3.4E+0	−5.6E−1	−4.3E+0	1.5E+0
	−3.4E−1	2.5E−1	6.2E−1	2.3E−1	−6.1E−1

Table 2.19: CEV model ($\beta = 0.2$, $\nu = 0.25$): errors in bps on the deltas using the 6 approximations `AppDeltaLN(1,`x_0`)`, `AppDeltaLN(1,k)`, `AppDeltaLN(1,`x_{avg}`)`, `AppDeltaLN(2,`x_0`)`, `AppDeltaLN(2,k)` and `AppDeltaLN(2,`x_{avg}`)`.

3M	19	27	25	18	11	6	0	−6	−11	−17	−30	−31	−27
	9	−1	−10	−13	−9	−3	0	3	7	9	−3	−7	−6
	−1	−1	0	2	2	2	0	−2	−2	−1	0	0	0
	5	3	2	1	1	0	0	0	0	1	2	3	5
	3	3	2	0	0	0	0	0	0	0	1	0	−1
	0	0	0	0	0	0	0	0	0	0	0	0	0
6M	51	45	34	24	16	9	0	−8	−15	−31	−41	−59	−58
	−7	−29	−27	−19	−10	−3	0	4	9	17	13	−6	−11
	−2	2	4	5	5	3	0	−2	−4	−3	−2	0	1
	10	3	2	2	1	1	1	1	1	2	2	5	13
	10	3	1	1	1	1	1	1	0	1	2	3	−2
	−1	0	0	0	1	1	1	1	1	0	0	0	0
1Y	99	90	59	46	24	13	1	−10	−32	−53	−93	−117	−113
	−54	−69	−52	−37	−11	−3	1	5	19	32	12	−14	−9
	−3	3	9	11	8	5	1	−3	−7	−7	−2	1	1
	24	10	5	4	3	2	2	2	3	4	7	12	41
	28	9	2	2	2	2	2	2	1	1	6	6	−4
	−2	0	1	1	2	2	2	2	1	1	0	−1	0
1.5Y	146	126	86	70	43	16	2	−26	−38	−77	−140	−183	−180
	−123	−114	−81	−60	−23	−3	2	13	21	45	12	−27	−10
	−2	7	15	17	15	7	2	−7	−9	−10	−3	1	1
	35	16	10	8	5	4	4	4	4	7	13	24	72
	37	12	3	3	4	4	4	3	2	2	11	7	−6
	−3	0	2	3	4	4	4	3	3	1	−1	−1	−1
2Y	189	163	17	81	51	36	3	−29	−58	−102	−173	−254	−224
	−197	−164	−115	−63	−23	−11	3	14	33	58	23	−40	−4
	0	11	20	23	19	14	3	−7	−13	−13	−5	2	1
	50	24	15	11	7	6	5	6	7	11	18	40	125
	45	15	5	5	6	6	5		3	3	14	4	−3
	−4	0	3	5	6	6	5	5	4	2	−1	−2	−1
3Y	262	218	188	118	82	45	6	−33	−86	−157	−286	−393	−354
	−377	−245	−204	−93	−40	−10	6	18	49	82	−3	−54	−3
	0	24	30	36	31	20	6	−7	−19	−18	−4	3	1
	95	36	29	19	14	11	10	10	13	0	35	77	230
	51	13	9	9	11	11	10	8	5	6	26	−4	−2
	−8	2	5	9	10	10	10	9	6	3	−2	−4	−1
5Y	373	323	253	197	131	85	11	−61	−148	−258	−497	−684	−592
	−641	−409	−276	−169	−60	−19	11	34	84	121	−43	−59	0
	0	43	59	64	53	38	11	−13	−30	−28	−3	4	0
	156	62	47	39	28	23	20	20	26	38	73	170	533
	−156	−19	3	15	22	23	20	15	8	13	38	−22	0
	−27	2	13	19	22	21	20	17	12	5	−6	−6	0
10Y	390	381	325	262	06	134	18	−102	−269	−529	−1075	−1477	−1164
	−174	−486	−447	−269	−123	−36	18	56	49	175	−127	−23	0
	−71	−8	39	76	78	58	18	−21	−53	−44	0	4	0
	75	−5	3	25	32	34	38	43	57	88	189	536	1725
	−1252	−598	−237	−43	18	7	38	32	15	47	5	−21	0
	−157	−97	−49	−3	23	34	38	36	26	6	−16	−10	0

CHAPTER 3

Forward implied volatility expansions in local volatility models

Contents

In this Chapter, we provide analytical approximations to efficiently price forward start option on equity in the local volatility model. We use a conditional argument to represent the price as an expectation of a Black-Scholes formula computed with a stochastic implied volatility depending on the value of the equity at the forward date. Then we perform a volatility expansion to derive analytical approximations of the forward implied volatility with a precise error estimate. We also illustrate the accuracy of our formulas with some numerical experiments.

3.1 Introduction

▷ **The forward volatility risk and associated derivative products.** The volatility allows to quantify the risk associated to the return of an underlying asset. Many products are actively traded on financial markets to manage the observed volatility smile and skew whereas a lot of models (as the local volatility model [Dupire 1994], the stochastic volatility model [Heston 1993] or a mixture of both of them known as the stochastic local volatility model [Britton-Jones 2000]) have emerged in the two last decades to try to reproduce these phenomena. Thus practitioners and researchers began to have a good intuition of the implied volatility behaviour associated to the pricing of plain vanilla options with the Black-Scholes (see [Black 1973]) formula. Despite the significant research on implied volatility asymptotics, only a few studies have been carried out on the asymptotic of the forward smile. The forward volatility risk is harder to manage and forward skew and smile shapes are still open to research. Recently a large class of new exotic options have emerged in order to take a bet or to hedge its exposure on the behaviour of the forward volatility surface, as the family of cliquet options which are commonly interpreted through levels of implied volatility at some future date and at some future level of the underlying asset. The evaluation and the hedging of such products are by far not trivial and the market has not yet settled for an agreed reference model. In this article, we focus on the pricing for the forward start option, which constitutes the building block for more complex cliquet structures as the Napoleons, the Multiplicative cliquets or the reverse cliquets (see [Gatheral 2006]). Basically it can be considered as a forward on an option. More precisely this is an option which begins at some specified future date $t_i > 0$, the forward date and with an expiration further in the future $t_i + T$ with $T > 0$, the premium being paid in advance at the initial date $t_0 = 0$. Denoting by S_t the price at time t of the underlying asset, we can distinguish two types of payoffs:

- $(\frac{S_{t_i+T}}{S_{t_i}} - K)_+$ (type A) for a given strike $K > 0$. It is essentially an option on the return of the asset between the dates t_i and $t_i + T$.
- $(S_{t_i+T} - KS_{t_i})_+$ (type B) with $K > 0$, which can be view as an option with a stochastic strike which will be determined at the forward date t_i. This looks like a spread option with the same underlying but considered at different dates.

From these payoffs one can build more complex structures of derivatives products. For instance a serie of consecutive forward start options creates a cliquet option with payoffs of the forms:

$$\sum_{i=1}^{n} (\frac{S_{t_i+1}}{S_{t_i}} - K_i)_+ \quad \text{or} \quad \sum_{i=1}^{n} (S_{t_i+1} - K_i S_{t_i})_+,$$

the valuation of such products being easily obtained summing the value of every legs.

▷**Literature review on the forward start options pricing.** Regarding the pricing of forward start options, many approaches could be considered as in the plain vanilla case. Basically on has to choose the mathematical modelling employed for the underlying asset (local volatility model, stochastic volatility model etc) and the analytical approximation methodology to be performed, closed-form formulas being available only in some very particular cases like in Gaussian or log-normal models. However it seems that many authors to have been interested by the pricing of forward start options have mainly considered the case of models with stochastic volatility like the Heston model: see for instance [Lucic 2003], [Hong 2004] or [Kruse 2005]. Brigo and Mercurio consider in the context of interest rates the Hull-White model in [Brigo 2006]. In all these works, owing to the properties of the affine models, it is possible if the model parameters are time-homogeneous to compute the forward characteristic function using the tower property for conditional expectations. Thus one can derive, up to numerical integration, (semi) analytical formulas. We also cite the work of Glasserman an Wu [Glasserman 2011] where the authors investigate the notion of forward implied volatility in the framework of stochastic volatility models applied to the currency markets. Then using the analytical approximation of the implied volatility in the SABR model (see [Hagan 2002]) and the asymptotic expansion for the bivariate density of both the underlying and its stochastic volatility developed in [Wu 2010], they provide tools for fast computation of the conditional expectations arising in the estimation of the forward implied volatility.
An alternative modeling is the use of Lévy processes proposed for instance in [Beyer 2008]. If the simple exponential Lévy model induces the same forward volatility curve for all futures times, a non trivial subordinator changes its dynamic. The authors derive the forward characteristic function and employ a Fourier transform machinery to obtain analytical pricing formulas for forward start options in various models including the Variance Gamma model and the NIG model subordinated by a CIR process. We also cite the work of Keller-Ressel and Kilin [Keller-Ressel 2008] who derive a semi-analytical formula for the pricing of forward start option in the Barndoff-Nielsen-Stephard (see [Barndorff-Nielson 2001]) model using its affine property.

We finally mention the very recent work of Jacquier and Roome [Jacquier 2012] in which is provided an expansion formula of the forward implied volatility using calculations based on the forward characteristic function and large deviations techniques. Remarkably their results can be applied for both small and large maturities in a large class of models, from the Heston model passing to the time-changed exponential Lévy processes.
In the class of models mentioned above, we start with a price process or a joint process price-volatility and deduce more and less the dynamic of the future volatility. To get a better control on the dynamic of the implied volatility, the most natural modeling approach is to model directly the implied volatility surface. This idea has been explored by Schonbucher [Schonbucher 1999]. But the method has been hampered by the difficulty of ensuring no arbitrage in future smiles or skew. To overcome the difficulty, Bergomi modelises jointly the dynamic for the forward variance swap and the spot consistently and discuss about the calibration and pricing in [Bergomi 2005], [Bergomi 2008].
Such an enthusiasm for the stochastic volatility models or more generally for two or more factors models in the literature can be explained by at least two reasons:

- At first glance the use of local volatility models, in which the volatility is a deterministic function of the random asset price, could seem inadequate to price forward start options which values appear to depend specifically on the random nature of the volatility itself. In addition it is well known that Skew/Smile generated by the non-constant local volatility function flattens for long maturities. As the forward start option depends on $\sigma(t, S_t)_{t\in[t_i,t_i+T]}$, we can expect it to be almost constant in S for large forward date t_i.

- The stochastic volatility models seem to induce more forward smile on $[t_i, t_i + T]$, which depends on the time-averaged stochastic volatility on $[t_i, t_i + T]$, than the implied volatility curve on $[0, t_i + T]$. Besides the availability of closed-form formulas is very attractive for practical uses.

Such important quantity of theoretical and practical studies show that we are far from having a reference model and that the model uncertainty associated to the pricing and hedging for the forward start option is high. We find important to consider a family of models, which are intuitive to the user and are able to replicate the market observed skew-smile as much as possible, in order to quantify properly the model risk. Indeed, Glassermann and Wu in ([Glasserman 2011]) show that the observed implied volatility does not contain relevant information for the future volatility. Moreover, even perfectly calibrated models can give quite different prices for exotic products, especially for the forward start option, because they generally generate different transition probabilities.
So it seems to us that there is a theoretical and a practical interest to provide analytical formulas for the forward implied volatility generated by local volatility models. First this is a challenge because as previously mentioned there is no closed-form formula and only few authors have been focused on the question. Second there is a risk of underestimation of the forward smile which can adversely affect when pricing forward start options too cheap as mentioned in [Gatheral 2003]. The purpose of this work is to overcome these drawbacks and to propose an accurate and tractable analytical approximation of the forward implied volatility in local volatility models with a precise estimation of the error.

▷**Formulation of the problem and contribution of our study.** In this work, we consider financial products in a world with null interest rate written w.r.t. a single asset which price at time t denoted by S_t assumed to pay no dividend. We consider a linear Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ where $(\mathcal{F}_t)_{t\geq 0}$ is the completion of the natural filtration of W. We suppose that S follows the local volatility model under the measure $\mathbb{P}$, i.e. it is solution of the next SDE:

$$\mathrm{d}S_t = S_t \sigma(t, S_t)\mathrm{d}W_t, \; S_0 > 0. \tag{3.1}$$

We assume that S is positive, we define the log-asset X by posing $X = \log(S)$ which satisfies:

$$\mathrm{d}X_t = a(t,X_t)\mathrm{d}W_t - \frac{1}{2}a^2(t,X_t)\mathrm{d}t, \quad x_0 = \log(S_0), \quad a(t,x) = \sigma(t,e^x). \tag{3.2}$$

We are interested by the price at time 0 of a forward start call option of type A and B written as:

$$\begin{cases} \mathrm{Call}^{\mathrm{FS},A}(S_0,t_i,T,K) = & \mathbb{E}[(\frac{S_{t_i+T}}{S_{t_i}} - K)_+] = \mathbb{E}[(e^{X_{t_i+T}-X_{t_i}} - e^k)_+], \\ \mathrm{Call}^{\mathrm{FS},B}(S_0,t_i,T,K) = & \mathbb{E}[(S_{t_i+T} - S_{t_i}K)_+] = \mathbb{E}[(e^{X_{t_i+T}} - e^{k+X_{t_i}})_+], \end{cases} \tag{3.3}$$

where $t_i > 0$ is the forward date, $T > 0$ the forward maturity and $K = e^k > 0$ the strike and $\mathbb{E}$ stands for the expectation operator. Notice that if S follows a log-normal model with deterministic volatility $(\sigma_t)_{t\in[0,T]}$, we have an analytical formula for the price of the forward start Call option of type A. The price is given in term of the Black-Scholes formula by:

$$\mathbb{E}[(\frac{S_{t_i+T}}{S_{t_i}} - K)_+] = \mathrm{Call}^{\mathrm{BS}}(0, \int_{t_i}^{t_i+T} \sigma_t^2 \mathrm{d}t, k), \tag{3.4}$$

where $\mathrm{Call}^{\mathrm{BS}}(x,y,z)$ denotes the *Black-Scholes Call price* function depending on log-spot x, total variance y and log-strike z and defined by:

$$\mathrm{Call}^{\mathrm{BS}}(x,y,z) = e^x\mathcal{N}(d_1(x,y,z)) - e^z\mathcal{N}(d_2(x,y,z)),$$
$$\mathcal{N}(x) = \int_{-\infty}^{x} \frac{e^{-u^2/2}}{\sqrt{2\pi}}\mathrm{d}u, \quad d_1(x,y,z) = \frac{x-z}{\sqrt{y}} + \frac{1}{2}\sqrt{y}, \quad d_2(x,y,z) = d_1(x,y,z) - \sqrt{y}.$$

For forward start options of type B in the Black-Scholes framework, we readily have using the tower property for the expectations, the Markov property of S and the independence of $\frac{S_{t_i+T}}{S_{t_i}}$ and S_{t_i}:

$$\begin{aligned} \mathbb{E}[(S_{t_i+T} - KS_{t_i})_+] =& \mathbb{E}[S_{t_i}(\frac{S_{t_i+T}}{S_{t_i}} - K)_+] = \mathbb{E}[S_{t_i}\mathbb{E}[(\frac{S_{t_i+T}}{S_{t_i}} - K)_+|S_{t_i}]] = \mathbb{E}[S_{t_i}]\mathrm{Call}^{\mathrm{BS}}(0, \int_{t_i}^{t_i+T} \sigma_t^2\mathrm{d}t, k) \\ =& S_0\mathrm{Call}^{\mathrm{BS}}(0, \int_{t_i}^{t_i+T} \sigma_t^2\mathrm{d}t, k) = \mathrm{Call}^{\mathrm{BS}}(x_0, \int_{t_i}^{t_i+T} \sigma_t^2\mathrm{d}t, x_0 + k). \end{aligned} \tag{3.5}$$

with $x_0 = \log(S_0)$. As a conclusion, the log-normal assumption on S leads to analytical formulas using the Black-Scholes pricer with the quadratic mean of the deterministic volatility on $[t_i, t_i + T]$. If the choice of a local volatility model allows to take into account the implied volatility skew usually observed on the equity market, no closed-form formulas are available in general even for the plain vanilla case (i.e. $t_i = 0$). Instead of resorting to time-costing (especially for large forward maturity T) numerical method like PDE techniques or Monte Carlo simulations, we aim at providing an accurate analytical approximation involving the same computational time than the application of the Black-Scholes formula. For this, we use the proxy principle introduced in [Benhamou 2009] and [Benhamou 2010a]. A broad overview of this non-asymptotic methodology as well comparisons with existing approximation methods to derive analytical formulas for quick an accurate evaluation of option prices is given in Chapters 1 and 2 of the thesis. We notably provide in Chapter 2 new approximations of the implied volatility very accurate and this is the starting point of our work. To derive an approximation of the forward implied volatility for the pricing of forward start option of type A, we use a conditioning argument and the results of Chapter 2 to express the price (3.3) of the forward start option as an expectation of the Black-Scholes price function with a stochastic volatility argument involving the local volatility function frozen at X_{t_i}, plus an error. Then we perform a volatility expansion to consider the local volatility function frozen at some deterministic point. A change of probability measure argument allows to adapt the results and to deduce

forward implied volatility of type B approximations.
The Chapter is organised as follows. First we give in Section 3.2 some usefull notations. Then we expose in Section 3.3 the main results of the Chapter with a second and third order forward implied volatility expansions of type A provided in Theorems 3.3.1.1-3.3.2.1 and forward implied volatility approximations of type B in Theorem 3.3.3.1. Section 3.4 is devoted to numerical experiments illustrating the behaviour of the forward implied volatilities and the accuracy of our approximations. We give in Appendix 3.5 some technical results and proofs as well as the explicit form of the forward implied volatility approximations in the context of time-independent local volatility function.

3.2 Notations

▷ **Assumptions on a. ($\mathcal{H}_a$)**: a is a bounded measurable function of $(t,x)\in[0,T+t_i]\times\mathbb{R}$, and five times continuously differentiable in x with bounded derivatives. Set

$$\mathcal{M}_1(a)=\max_{1\le i\le 5}\sup_{(t,x)\in[0,T+t_i]\times\mathbb{R}}|\partial^i_{x^i}a(t,x)| \text{ and } \mathcal{M}_0(a)=\max_{0\le i\le 5}\sup_{(t,x)\in[0,T+t_i]\times\mathbb{R}}|\partial^i_{x^i}a(t,x)|.$$

In addition, there exists $C_a\in]0,1]$ such that $|a(t,x)|\ge C_a\mathcal{M}_0(a)>0$ for any $(t,x)\in[0,T+t_i]\times\mathbb{R}$.

▷ **Temporal shift of the local volatility function a.** We introduce the time-shifted local volatility function α defined by $\alpha(t,x)=a(t+t_i,x)$ for any $(t,x)\in[0,T]\times\mathbb{R}$.

▷ **Time-space shifted local volatility process.** We introduce the time-space shifted local volatility process starting from 0 defined for any $x\in\mathbb{R}$ by the following SDE:

$$\mathrm{d}Z^x_t=\alpha(t,Z^x_t+x)\mathrm{d}W_t-\frac{1}{2}\alpha^2(t,Z^x_t+x)\mathrm{d}t,\quad z_0=0. \tag{3.6}$$

▷ **Integral operators.** For any $n\ge 1$, any $l_1,..,l_n$ measurable and bounded functions of $t\in[0,T+t_i]$, any $0\le s<t\le T+t_i$, we set:

$$\omega(l_1,..,l_n)^t_s=\int_s^t l_1(r_1)\int_{r_1}^t l_2(r_2)...\int_{r_{n-1}}^t l_n(r_n)\mathrm{d}r_n\mathrm{d}r_{n-1}..\mathrm{d}r_2\mathrm{d}r_1.$$

▷ **Quadratic mean and total variance.** We define the quadratic mean for any bounded measurable function g of $(t,x)\in[0,T+t_i]\times\mathbb{R}$ and for any non empty $[s,t]\subseteq[0,T+t_i]$ at the spatial point x on $[s,t]$ by setting:

$$\overline{g}^{s,t}_x=\sqrt{\frac{1}{t-s}\int_s^t g^2(r,x)\mathrm{d}r}.$$

For any bounded measurable function g of $(t,x)\in[0,T+t_i]\times\mathbb{R}$ and for any non empty $[s,t]\subseteq[0,T+t_i]$, we define the total variance of g at x on $[s,t]$ as:

$$\mathcal{V}(g;x)^t_s=\int_s^t g^2(r,x)\mathrm{d}r=(\overline{g}^{s,t}_x)^2(t-s).$$

We finally introduce some integral operators C, γ and π already used in Chapter 2:

Definition 3.2.0.1. *If the derivatives and the integrals have a meaning, we define for any non empty $[s,t]\subseteq[0,T+t_i]$ and for any bounded measurable function $(l(t,z))_{(t,z)\in[0,T+t_i]\times\mathbb{R}}$ the next operators:*

$$C_1(l;z)^t_s=\omega(l^2(z),l(z)l^{(1)}(z))^t_s,\qquad C_2(l;z)^t_s=\omega(l^2(z),(l^{(1)}(z))^2+l(z)l^{(2)}(z))^t_s,$$

$$C_3(l;z)_s^t = \omega(l^2(z), l^2(z), (l^{(1)}(z))^2 + l(z)l^{(2)}(z))_s^t, \qquad C_4(l;z)_s^t = \omega(l^2(z), l(z)l^{(1)}(z), l(z)l^{(1)}(z))_s^t,$$
$$C_5(l;z)_s^t = \omega((l^{(1)}(z))^2 + l(z)l^{(2)}(z))_t^s, \qquad C_6(l;z)_s^t = \omega(l(z)l^{(1)}(z), l(z)l^{(1)}(z))_t^s,$$
$$C_7(l;z)_s^t = \omega(l(z)l^{(1)}(z))_t^s.$$

We can define similarly the reverse operators $\widetilde{C}$ obtained by changing the order of integration. For example $\widetilde{C}_1(l;z)_s^t = \omega(l(z)l^{(1)}(z), l^2(z))_s^t$.
Supposing in addition that l is non-negative such that $\bar{l}_z^{s,t} > 0$, we define the following operators:

$$\gamma_0(l;z)_s^t = \bar{l}_z^{s,t} + \frac{C_2(l;z)_s^t}{2\bar{l}_z^{s,t}(t-s)} - \frac{C_4(l;z)_s^t}{4\bar{l}_z^{s,t}(t-s)} - \frac{C_3(l;z)_s^t}{(\bar{l}_z^{s,t})^3(t-s)^2} - \frac{3C_4(l;z)_s^t}{(\bar{l}_z^{s,t})^3(t-s)^2} + \frac{[C_1(l;z)_s^t]^2}{8(\bar{l}_z^{s,t})^3(t-s)^2} + \frac{3[C_1(l;z)_s^t]^2}{2(\bar{l}_z^{s,t})^5(t-s)^3},$$
$$\gamma_1(l;z)_s^t = \frac{C_1(l;z)_s^t}{(\bar{l}_z^{s,t})^3(t-s)^2}, \qquad \gamma_2(l;z)_s^t = \frac{C_3(l;z)_s^t}{(\bar{l}_z^{s,t})^5(t-s)^3} + 3\frac{C_4(l;z)_s^t}{(\bar{l}_z^{s,t})^5(t-s)^3} - \frac{3[C_1(l;z)_s^t]^2}{(\bar{l}_z^{s,t})^7(t-s)^4},$$
$$\pi_0(l;z)_s^t = \frac{\gamma_0(l;z)_s^t + \widetilde{\gamma}_0(l;z)_s^t}{2}, \qquad \pi_1(l;z)_s^t = \frac{\widetilde{\gamma}_1(l;z)_s^t - \gamma_1(l;z)_s^t}{2},$$
$$\pi_2(l;z)_s^t = \frac{\gamma_2(l;z)_s^t + \widetilde{\gamma}_2(l;z)_s^t}{2} - \frac{C_5(l;z)_s^t}{8\bar{l}_z^{s,t}(t-s)} + \frac{C_6(l;z)_s^t}{4(\bar{l}_z^{s,t})^3(t-s)^2},$$

where the reverse operators $\widetilde{\gamma}$ are obtained using the reverse operators $\widetilde{C}$.

Remark 3.2.0.1. *Any of the previously defined operators applied with the function α and the spatial point $x \in \mathbb{R}$ between the dates 0 and T gives the same result as that obtained with a and x between the dates t_i and $T + t_i$. For example:*

$$\begin{aligned}
C_1(\alpha;x)_0^T &= \omega(\alpha^2(x), \alpha(x)\alpha^{(1)}(x))_0^T = \int_0^T \alpha_t^2(x) \int_t^T \alpha_s(x)\alpha_s^{(1)}(x)\mathrm{d}s\mathrm{d}t \\
&= \int_0^T a_{t+t_i}^2(x) \int_t^T a_{s+t_i}(x)a_{s+t_i}^{(1)}(x)\mathrm{d}s\mathrm{d}t = \int_{t_i}^{T+t_i} a_t^2(x) \int_t^{T+t_i} a_s(x)a_s^{(1)}(x)\mathrm{d}s\mathrm{d}t \\
&= \omega(a(x), a(x)a^{(1)}(x))_{t_i}^{T+t_i} = C_1(a;x)_{t_i}^{T+t_i}.
\end{aligned}$$

▷ **Forward implied Black-Scholes volatility of type A and B.** For (x_0, t_i, T, k) given, the forward implied Black Scholes volatilities of type A and B are the unique non-negative parameters $\sigma_{\mathrm{I,F,A}}(x_0, t_i, T, k)$ and $\sigma_{\mathrm{I,F,B}}(x_0, t_i, T, k)$ such that:

$$\begin{cases} \mathrm{Call}^{\mathrm{FS},A}(e^{x_0}, t_i, T, e^k) = \mathrm{Call}^{\mathrm{BS}}(0, \sigma_{\mathrm{I,F,A}}^2(x_0, t_i, T, k)T, k), \\ \mathrm{Call}^{\mathrm{FS},B}(e^{x_0}, t_i, T, e^k) = \mathrm{Call}^{\mathrm{BS}}(x_0, \sigma_{\mathrm{I,F,B}}^2(x_0, t_i, T, k)T, x_0 + k). \end{cases} \tag{3.7}$$

If a does not depend on the spatial component (Black-Scholes framework), we have

$$\sigma_{\mathrm{I,F,A}}(x_0, t_i, T, k) = \sigma_{\mathrm{I,F,B}}(x_0, t_i, T, k) = \sqrt{\frac{1}{T}\int_{t_i}^{t_i+T} a_t^2 \mathrm{d}t} = \bar{a}^{t_i, t_i+T} = \bar{\alpha}^{0,T}.$$

▷ **New log-strike and new mid-point.** We use the notation $k' = \frac{k}{2}$ and $x'_{avg} = x_0 + k' = x_0 + \frac{k}{2}$.
▷ **About the constants.** All our error estimates are stated throughout the paper using the notations:

- "$A = O(B)$" means that $|A| \leq CB$ where C stands for a generic constant that is a non-negative increasing function of T, t_i, $\mathcal{M}_0(a)$, $\mathcal{M}_1(a)$ and the oscillation ratio $\frac{1}{C_a}$.
- Similarly, if A is non-negative, $A \leq_c B$ means that $A \leq CB$ for a generic constant C.

3.3 Second and third order forward implied volatility expansions

3.3.1 Second order forward implied volatility expansion of type A

We provide in the next Theorem a second order forward volatility expansion of type A:

Theorem 3.3.1.1. *(**2nd order expansion of the forward implied volatility of type A**). Assume $(\mathcal{H}_a)$ and suppose that $\mathcal{M}_1(a)$, $\mathcal{M}_0(a)$, T, t_i and $|k|$ are globally small enough to ensure the existence of $C_1(k) > 0$ and $C_2(k) > 0$ such that:*

$$\bar{\alpha}_x^{0,T} - k\pi_1(\alpha; x)_0^T = \bar{a}_x^{t_i,T+t_i} - k\pi_1(a; x)_{t_i}^{T+t_i} > C_1(k)\mathcal{M}_0(a) > 0, \quad \forall x \in \mathbb{R}, \tag{3.8}$$

$$\tilde{\sigma}_{\mathrm{I,F,A}}^{2,x'_{avg}}(x_0, t_i, T, k) = \bar{a}_{x'_{avg}}^{t_i,T+t_i} - \frac{1}{2}\frac{C_7(a; x'_{avg})_{t_i}^{T+t_i}}{\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}\mathcal{V}(a; x_0)_0^{t_i} - k\pi_1(a; x'_{avg})_{t_i}^{T+t_i} > C_2(k)\mathcal{M}_0(a) > 0, \tag{3.9}$$

where the operator π_1 is defined in Definition 3.2.0.1. Then $\tilde{\sigma}_{\mathrm{I,F,A}}^{2,x'_{avg}}(x_0, t_i, T, k)$ is a second order approximation of $\sigma_{\mathrm{I,F,A}}(x_0, t_i, T, k)$ in the following sense:

$$\mathrm{Call}^{\mathrm{FS},A}(e^{x_0}, t_i, T, e^k) = \mathrm{Call}^{\mathrm{BS}}\Big(0, (\tilde{\sigma}_{\mathrm{I,F,A}}^{2,x'_{avg}}(x_0, t_i, T, k))^2 T, k\Big) + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2\sqrt{T}(\sqrt{t_i} + \sqrt{T})^2), \tag{3.10}$$

where the constant in the above estimate notably depends on $C_1(k)$ and $C_2(k)$.

Remark 3.3.1.1. *If T tends to 0, the error in (3.10) becomes null. This is coherent with the fact that $\mathrm{Call}^{\mathrm{FS},A}(e^{x_0}, t_i, 0, e^k) = (1 - e^k)_+$. If t_i tends to 0 (or $\mathcal{M}_1(a)$ tends to 0) the term $-\frac{1}{2}\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}\mathcal{V}(a; x_0)_0^{t_i}$ vanishes. We namely have:*

$$\left|\frac{1}{2}\frac{C_7(a; x'_{avg})_{t_i}^{T+t_i}}{\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}\mathcal{V}(a; x_0)_0^{t_i}\right| \leq_c \mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 t_i. \tag{3.11}$$

Hence for $x_0 = 0$, we retrieve the expansion of the vanilla case (see Theorem 2.1.4.1 in Chapter 2). This additional term due to the forward start is therefore interpreted as a forward bias.

Remark 3.3.1.2. *In view of the magnitude of the term $-\frac{1}{2}\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}\mathcal{V}(a; x_0)_0^{t_i}$ (see (3.11)) and owing to the Lemma 3.5.1.1, one can replace $\mathcal{V}(a; x_0)_0^{t_i}$ by $\mathcal{V}(a; x'_{avg})_0^{t_i}$ without changing the magnitude of the error. Thus the approximation of the forward implied volatility with the local volatility frozen at x'_{avg} reads:*

$$\sigma_{\mathrm{I,F,A}}(x_0, t_i, T, k) \approx \bar{a}_{x'_{avg}}^{t_i,T+t_i} - \frac{1}{2}\frac{C_7(a; x'_{avg})_{t_i}^{T+t_i}}{\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}\mathcal{V}(a; x'_{avg})_0^{t_i} - k\pi_1(a; x'_{avg})_{t_i}^{T+t_i}. \tag{3.12}$$

Proof. We use the Markov property of the process X to get:

$$\mathrm{Call}^{\mathrm{FS},A}(e^{x_0}, t_i, T, e^k) = \mathbb{E}[\mathbb{E}[(e^{X_{t_i+T} - X_{t_i}} - e^k)_+ | X_{t_i}]]. \tag{3.13}$$

Then using the deterministic time change $t \mapsto t + t_i$ for any $t \in [0, T]$ and [Revuz 1999, Propositions 5.1.4 and 5.1.5], we easily see that under the conditional knowledge of X_{t_i}, $X_{t_i+T} - X_{t_i}$ has the same law that $Z_T^{X_{t_i}}$ where $(Z_t^x)_{t\in[0,T]}$ is the solution of the SDE (3.6). Hence we can write:

$$\mathbb{E}[(e^{X_{t_i+T} - X_{t_i}} - e^k)_+ | X_{t_i}] = \mathbb{E}[(e^{Z_T^{X_{t_i}}} - e^k)_+ | X_{t_i}]. \tag{3.14}$$

Next remark that $\mathbb{E}[(e^{Z_T^{X_{t_i}}}-e^k)_+|X_{t_i}]$ is nothing else but the price of a Call option at time 0 with maturity T, strike e^k, spot 1 written on the log-asset $Z_T^{X_{t_i}}$ with local volatility function $(t,x)\mapsto \alpha(t,x+X_{t_i})$.
Then owing to the assumed positivity of $\bar{\alpha}_x^{0,T}-k\pi_1(\alpha;x)_0^T$, $\forall x\in\mathbb{R}$, we can follow the proof of Theorem 2.1.4.1 in Chapter 2 and one easily obtains with Lemma 3.5.1.1:

$$\mathbb{E}[(e^{Z_T^{X_{t_i}}}-e^k)_+|X_{t_i}]=\mathrm{Call}^{\mathrm{BS}}(0,(\bar{\alpha}_{k'+X_{t_i}}^{0,T}-k\pi_1(\alpha;k'+X_{t_i})_0^T)^2T,k)+\mathrm{Error}_{2,k'}(X_{t_i}), \tag{3.15}$$

where $\bar{\alpha}_{k'+X_{t_i}}^{0,T}-k\pi_1(\alpha;k'+X_{t_i})_0^T$ is the order 2 implied volatility expansion and where a.s. $\left|\mathrm{Error}_{2,k'}(X_{t_i})\right|\leq_c \mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}$ (see remark 3.3.1.3 below). Then the price (3.13) can be written as:

$$\begin{aligned}\mathrm{Call}^{\mathrm{FS},A}(e^{x_0},t_i,T,e^k)=&\mathbb{E}[\mathrm{Call}^{\mathrm{BS}}(0,(\bar{\alpha}_{k'+X_{t_i}}^{0,T}-k\pi_1(\alpha;k'+X_{t_i})_0^T)^2T,k)]+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}) \quad (3.16)\\ =&\mathbb{E}[\mathrm{Call}^{\mathrm{BS}}(0,(\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-k\pi_1(a;k'+X_{t_i})_{t_i}^{T+t_i})^2T,k)]+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}}).\end{aligned}$$

Then we apply Taylor expansions twice: firstly for the smooth function $\nu\mapsto \mathrm{Call}^{\mathrm{BS}}(0,\nu^2T,k)$ at $\nu=\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-k\pi_1(a;k'+X_{t_i})_{t_i}^{T+t_i}$ around $\nu=\bar{a}_{x'_{avg}}^{t_i,T+t_i}-k\pi_1(a;x'_{avg})_{t_i}^{T+t_i}$ and secondly for the functions $x\mapsto\bar{a}_x^{t_i,T+t_i}$ and $x\mapsto \pi_1(a;x)_{t_i}^{T+t_i}$ at $x=k'+X_{t_i}$ around $x=k'+x_0=x'_{avg}$:

$$\begin{aligned}&\mathbb{E}[\mathrm{Call}^{\mathrm{BS}}(0,(\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-k\pi_1(a;k'+X_{t_i})_{t_i}^{T+t_i})^2T,k)] \quad (3.17)\\ =&\mathrm{Call}^{\mathrm{BS}}(0,(\bar{a}_{x'_{avg}}^{t_i,T+t_i}-k\pi_1(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\\ &+\mathrm{Vega}^{\mathrm{BS}}(0,(\bar{a}_{x'_{avg}}^{t_i,T+t_i}-k\pi_1(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\\ &\times\mathbb{E}[(\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-\bar{a}_{x'_{avg}}^{t_i,T+t_i}+k(\pi_1(a;x'_{avg})_{t_i}^{T+t_i}-\pi_1(a;k'+X_{t_i})_{t_i}^{T+t_i}))]+R_1,\\ =&\mathrm{Call}^{\mathrm{BS}}(0,(\bar{a}_{x'_{avg}}^{t_i,T+t_i}-k\pi_1(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\\ &+\mathrm{Vega}^{\mathrm{BS}}(0,(\bar{a}_{x'_{avg}}^{t_i,T+t_i}-k\pi_1(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}\mathbb{E}[(X_{t_i}-x_0)]\\ &+R_1+R_2+R_3,\end{aligned}$$

where the operator C_7 is defined in Definition 3.2.0.1, $\mathrm{Vega}^{\mathrm{BS}}$,$\mathrm{Vomma}^{\mathrm{BS}}$ in Lemma 3.5.1.1 and where:

$$\begin{aligned}R_1=&\mathbb{E}\Big[(\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-\bar{a}_{x'_{avg}}^{t_i,T+t_i}+k(\pi_1(a;x'_{avg})_{t_i}^{T+t_i}-\pi_1(a;k'+X_{t_i})_{t_i}^{T+t_i}))^2\\ &\times\int_0^1\mathrm{Vomma}^{\mathrm{BS}}(0,\nu^2T,k)|_{\nu=(1-\lambda)(\bar{a}_{x'_{avg}}^{t_i,T+t_i}-k\pi_1(a;x'_{avg})_{t_i}^{T+t_i})+\lambda(\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-k\pi_1(a;k'+X_{t_i})_{t_i}^{T+t_i})}(1-\lambda)\mathrm{d}\lambda\Big],\\ R_2=&\mathbb{E}\Big[\mathrm{Vega}^{\mathrm{BS}}(0,(\bar{a}_{x'_{avg}}^{t_i,T+t_i}-k\pi_1(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)(X_{t_i}-x_0)^2\int_0^1\partial_{x^2}^2\bar{a}_x^{t_i,T+t_i}|_{x=k'+\lambda X_{t_i}+(1-\lambda)x_0}(1-\lambda)\mathrm{d}\lambda\Big],\\ R_3=&\mathbb{E}\Big[\mathrm{Vega}^{\mathrm{BS}}(0,(\bar{a}_{x'_{avg}}^{t_i,T+t_i}-k\pi_1(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)k(x_0-X_{t_i})\int_0^1\partial_x\pi_1(a;x)_{t_i}^{T+t_i}|_{x=k'+\lambda X_{t_i}+(1-\lambda)x_0}\mathrm{d}\lambda\Big].\end{aligned}$$

Then using Equations (2.18)-(2.6) of Chapter 2, one has the following weak approximation:

$$\mathbb{E}[X_{t_i}-x_0]=\mathbb{E}[X_{1,t_i}]+O(\mathcal{M}_1(a)\mathcal{M}_0(a)t_i)=-\frac{1}{2}\mathcal{V}(a;x_0)_0^{t_i}+O(\mathcal{M}_1(a)\mathcal{M}_0(a)t_i), \tag{3.18}$$

where $(X_{1,t})_{t\in[0,T+t_i]}$ is the corrective (Gaussian) process defined by:

$$\mathrm{d}X_{1,t}=a(t,x_0)\mathrm{d}W_t-\frac{1}{2}a^2(t,x_0)\mathrm{d}t,\ X_{1,0}=0. \tag{3.19}$$

Then using Lemma 3.5.1.1 we easily get with (3.18):

$$\mathrm{Vega}^{\mathrm{BS}}(0,(\bar{a}^{t_i,T+t_i}_{x'_{avg}}-k\pi_1(a;x'_{avg})^{T+t_i}_{t_i})^2T,k)\frac{C_7(a;x'_{avg})^{T+t_i}_{t_i}}{\bar{a}^{t_i,T+t_i}_{x'_{avg}}T}\mathbb{E}[(X_{t_i}-x_0)] \tag{3.20}$$
$$=-\frac{1}{2}\mathrm{Vega}^{\mathrm{BS}}(0,(\bar{a}^{t_i,T+t_i}_{x'_{avg}}-k\pi_1(a;x'_{avg})^{T+t_i}_{t_i})^2T,k)\frac{C_7(a;x'_{avg})^{T+t_i}_{t_i}}{\bar{a}^{t_i,T+t_i}_{x'_{avg}}T}\mathcal{V}(a;x_0)^{t_i}_0+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i\sqrt{T}).$$

Next we bound the remainders. We readily have using Lemma 3.5.1.1:

$$R_1=O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i\sqrt{T}),\quad R_2=O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i\sqrt{T}),\quad R_3=O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2\sqrt{t_i}T).$$

Combining these estimates with (3.16)-(3.17)-(3.20) and using (3.11) and Lemma 3.5.1.1 finally yields to:

$$\begin{aligned}\mathrm{Call}^{\mathrm{FS},A}(e^{x_0},t_i,T,e^k)=&\mathrm{Call}^{\mathrm{BS}}(0,(\bar{a}^{t_i,T+t_i}_{x'_{avg}}-k\pi_1(a;x'_{avg})^{T+t_i}_{t_i})^2T,k)\\&-\frac{1}{2}\mathrm{Vega}^{\mathrm{BS}}(0,(\bar{a}^{t_i,T+t_i}_{x'_{avg}}-k\pi_1(a;x'_{avg})^{T+t_i}_{t_i})^2T,k)\frac{C_7(a;x'_{avg})^{T+t_i}_{t_i}}{\bar{a}^{t_i,T+t_i}_{x'_{avg}}T}\mathcal{V}(a;x_0)^{t_i}_0\\&+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2\sqrt{T}(\sqrt{t_i}+\sqrt{T})^2)\\=&\mathrm{Call}^{\mathrm{BS}}\Big(0,(\bar{a}^{t_i,T+t_i}_{x'_{avg}}-\frac{1}{2}\frac{C_7(a;x'_{avg})^{T+t_i}_{t_i}}{\bar{a}^{t_i,T+t_i}_{x'_{avg}}T}\mathcal{V}(a;x_0)^{t_i}_0-k\pi_1(a;x'_{avg})^{T+t_i}_{t_i})^2T,k\Big)\\&+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2\sqrt{T}(\sqrt{t_i}+\sqrt{T})^2).\end{aligned}$$

Remark 3.3.1.3. *In can be easily proven that the error* $\mathrm{Error}_{2,k'}(X_{t_i})$ *defined in* (3.15) *involves a dependence w.r.t.* X_{t_i} *only throughout the local volatility function a (as a shift parameter) and its derivatives which are bounded functions. One can notice that the employed conditional argument seems a priori inadequate to handle stochastic volatility models like the Heston model. A carefully introspection of the work of Gobet et al. [Benhamou 2010b] namely reveals that the error in the Heston Call price approximation depends on* v_0^{-p} *with* v_0 *the initial variance process value and* $p>1$. *In the context of forward start options, we may bound negative moments of* v_{t_i} *which can explode, see [Bossy 2007, Lemma A.1].*

□

3.3.2 Third order forward implied volatility expansion of type A

We announce the main result of the Chapter:

Theorem 3.3.2.1. *(**3rd order expansion of the forward implied volatility of type A**). Assume* $(\mathcal{H}_a)$ *and suppose that* $\mathcal{M}_1(a)$, $\mathcal{M}_0(a)$, T, t_i *and* $|k|$ *are globally small enough to ensure the existence of* $C_3(k)>0$ *and* $C_4(k)>0$ *such that:*

$$\pi^k(\alpha;x)^T_0=\pi^k(a;x)^{t_i+T}_{t_i}=\pi_0(\alpha;x)^T_0-k\pi_1(\alpha;x)^T_0+k^2\pi_2(\alpha;x)^T_0>C_3(k)\mathcal{M}_0(a)>0,\quad\forall x\in\mathbb{R}, \tag{3.21}$$
$$\tilde{\sigma}^{3,x'_{avg}}_{\mathrm{I,F,A}}(x_0,t_i,T,k) \tag{3.22}$$
$$=\pi^k(a;x'_{avg})^{T+t_i}_{t_i}+\frac{C_7(a;x'_{avg})^{T+t_i}_{t_i}}{2\bar{a}^{t_i,T+t_i}_{x'_{avg}}T}[-\mathcal{V}(a;x_0)^{t_i}_0+C_1(a;x_0)^{t_i}_0] \tag{3.23}$$
$$+[\frac{C_5(a;x'_{avg})^{T+t_i}_{t_i}}{2\bar{a}^{t_i,T+t_i}_{x'_{avg}}T}-\frac{C_6(a;x'_{avg})^{T+t_i}_{t_i}}{(\bar{a}^{t_i,T+t_i}_{x'_{avg}})^3T^2}][\mathcal{V}(a;x_0)^{t_i}_0+\frac{1}{4}\mathcal{V}^2(a;x_0)^{t_i}_0]$$

$$
\begin{aligned}
&+\Big[\frac{k^2}{(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T}-\frac{(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T}{4}\Big]\frac{C_6(a;x'_{avg})_{t_i}^{T+t_i}}{(\bar{a}_{x'_{avg}}^{t_i,T+t_i})^2\pi^k(a;x'_{avg})_{t_i}^{T+t_i}T^2}\mathcal{V}(a;x_0)_0^{t_i}\\
&+\frac{k}{2}\Big[\frac{C_2(\tilde{a};x'_{avg})_{t_i}^{T+t_i}-C_2(a;x'_{avg})_{t_i}^{T+t_i}}{2(\bar{a}_x^{t_i,T+t_i})^3T^2}+3\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}(C_1(a;x'_{avg})_{t_i}^{T+t_i}-C_1(\tilde{a};x'_{avg})_{t_i}^{T+t_i})}{2(\bar{a}_x^{t_i,T+t_i})^5T^3}\Big]\mathcal{V}(a;x_0)_0^{t_i}\\
&>C_4(k)\mathcal{M}_0(a)>0,
\end{aligned}
$$

where the operators π_0, π_1, π_2, C_5, C_6 and C_7 are defined in Definition 3.2.0.1. Then $\tilde{\sigma}_{\mathrm{I,F,A}}^{3,x'_{avg}}(x_0,t_i,T,k)$ is a third order approximation of $\sigma_{\mathrm{I,F,A}}(x_0,t_i,T,k)$ in the following sense:

$$
\mathrm{Call}^{\mathrm{FS},A}(e^{x_0},t_i,T,e^k)=\mathrm{Call}^{\mathrm{BS}}\Big(0,(\tilde{\sigma}_{\mathrm{I,F,A}}^{3,x'_{avg}}(x_0,t_i,T,k))^2T,k\Big)+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3\sqrt{T}(\sqrt{t_i}+\sqrt{T})^3), \quad (3.24)
$$

where the constant in the above estimate notably depends on $C_3(k)$ and $C_4(k)$.

Remark 3.3.2.1. *If $t_i=0$, the above forward implied volatility approximation reduces to $\pi^k(a;x'_{avg})_{t_i}^{T+t_i}$ what is the approximation of the implied volatility expansion given in Theorem 2.3.3.1 of Chapter 2. Owing to their magnitude, all the new terms due to the forward start can be considered with the local volatility frozen at x'_{avg} instead of x_0 without changing the magnitude of the final error, thanks to Lemma 3.5.1.1.*

Proof. Using the assumption on $\pi^k(a;x)_{t_i}^{T+t_i}$, we can apply Theorem 2.3.3.1 of Chapter 2 and using the same methodology previously employed for the second order, we get

$$
\mathrm{Call}^{\mathrm{FS},A}(e^{x_0},t_i,T,e^k)=\mathbb{E}[\mathrm{Call}^{\mathrm{BS}}(0,(\pi^k(a;k'+X_{t_i})_{t_i}^{T+t_i})^2T,k)]+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2). \quad (3.25)
$$

Then apply a Taylor expansion for the smooth function $\nu\mapsto\mathrm{Call}^{\mathrm{BS}}(0,\nu^2T,k)$ at $\nu=\pi^k(a;k'+X_{t_i})_{t_i}^{T+t_i}$ around $\nu=\pi^k(a;x'_{avg})_{t_i}^{T+t_i}$:

$$
\begin{aligned}
&\mathbb{E}[\mathrm{Call}^{\mathrm{BS}}(0,(\pi^k(a;k'+X_{t_i})_{t_i}^{T+t_i})^2T,k)] \quad (3.26)\\
=&\mathrm{Call}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\\
&+\mathrm{Vega}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\mathbb{E}[\pi^k(a;k'+X_{t_i})_{t_i}^{T+t_i}-\pi^k(a;x'_{avg})_{t_i}^{T+t_i}]\\
&+\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\mathbb{E}[(\pi^k(a;k'+X_{t_i})_{t_i}^{T+t_i}-\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2]\\
&+R.
\end{aligned}
$$

where:

$$
\begin{aligned}
R=&\mathbb{E}\Big[(\pi^k(a;k'+X_{t_i})_{t_i}^{T+t_i}-\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^3\\
&\times\int_0^1\mathrm{Ultima}^{\mathrm{BS}}(0,\nu^2T,k)|_{\nu=(1-\lambda)\pi^k(a;x'_{avg})_{t_i}^{T+t_i}+\lambda\pi^k(a;k'+X_{t_i})_{t_i}^{T+t_i}}\frac{(1-\lambda)^2}{2}\mathrm{d}\lambda\Big],
\end{aligned}
$$

and where $\mathrm{Ultima}^{\mathrm{BS}}$ is defined in Lemma 3.5.1.1. First notice that we readily have using Lemma 3.5.1.1:

$$
R=O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3t_i^{\frac{3}{2}}\sqrt{T}). \quad (3.27)
$$

Then we expand the functions $x\mapsto\bar{a}_x^{t_i,T+t_i}$, $x\mapsto\pi_0(a;x)_{t_i}^{T+t_i}-\bar{a}_x^{t_i,T+t_i}$ and $x\mapsto\pi_i(a;x)_{t_i}^{T+t_i}$, $i\in\{1,2\}$. We announce three technical Lemmas, the proofs being postponed to Appendix 3.5.2. First we give in the next Lemma expansion results for $\bar{a}_x^{t_i,T+t_i}$:

Lemma 3.3.2.1. *(**Expansion of** $\bar{a}_x^{t_i,T+t_i}$). Under the hypotheses of Theorem 3.3.2.1, we have:*

$$
\begin{aligned}
&\mathrm{Vega}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\mathbb{E}[\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-\bar{a}_{x'_{avg}}^{t_i,T+t_i}]\\
=&\mathrm{Vega}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\Big\{\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{2\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}[-\mathcal{V}(a;x_0)_0^{t_i}+C_1(a;x_0)_0^{t_i}]\\
&+[\frac{C_5(a;x'_{avg})_{t_i}^{T+t_i}}{2\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}-\frac{C_6(a;x'_{avg})_{t_i}^{T+t_i}}{(\bar{a}_{x'_{avg}}^{t_i,T+t_i})^3T^2}][\mathcal{V}(a;x_0)_0^{t_i}+\frac{1}{4}\mathcal{V}^2(a;x_0)_0^{t_i}]\Big\}+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3t_i^{\frac{3}{2}}\sqrt{T}),\\
&\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\mathbb{E}[(\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-\bar{a}_{x'_{avg}}^{t_i,T+t_i})^2]\\
=&\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\Big\{(\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{2\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}\mathcal{V}(a;x_0)_0^{t_i})^2+2\frac{C_6(a;x'_{avg})_{t_i}^{T+t_i}}{(\bar{a}_{x'_{avg}}^{t_i,T+t_i})^2T^2}\mathcal{V}(a;x_0)_0^{t_i}\Big\}\\
&+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3t_i^{\frac{3}{2}}\sqrt{T}).
\end{aligned}
$$

We now announce an expansion result for $\pi_1(a;x)_{t_i}^{T+t_i}$:

Lemma 3.3.2.2. *(**Expansion of** $\pi_1(a;x)_{t_i}^{T+t_i}$). Under the hypotheses of Theorem 3.3.2.1, we have:*

$$
\begin{aligned}
&\mathrm{Vega}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)k\mathbb{E}[\pi_1(a;x'_{avg})_{t_i}^{T+t_i}-\pi_1(a;k'+X_{t_i})_{t_i}^{T+t_i}]\\
=&\mathrm{Vega}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\frac{k}{2}\mathcal{V}(a;x_0)_0^{t_i}[\frac{C_2(\tilde{a};x'_{avg})_{t_i}^{T+t_i}-C_2(a;x'_{avg})_{t_i}^{T+t_i}}{2(\bar{a}_x^{t_i,T+t_i})^3T^2}\\
&+3\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}(C_1(a;x'_{avg})_{t_i}^{T+t_i}-C_1(\tilde{a};x'_{avg})_{t_i}^{T+t_i})}{2(\bar{a}_x^{t_i,T+t_i})^5T^3}]+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3Tt_i).
\end{aligned}
$$

Last, we show that all the remaining terms are negligible. The following Lemma summarises the results:

Lemma 3.3.2.3. *Under the hypotheses of Theorem 3.3.2.1, we have:*

$$
\begin{aligned}
&\mathrm{Vega}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\\
&\times\mathbb{E}[\pi_0(a;k'+X_{t_i})_{t_i}^{T+t_i}-\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-\pi_0(a;x'_{avg})_{t_i}^{T+t_i}+\bar{a}_{x'_{avg}}^{t_i,T+t_i}+k^2(\pi_2(a;k'+X_{t_i})_{t_i}^{T+t_i}-\pi_2(a;x'_{avg})_{t_i}^{T+t_i})]\\
\le_c&\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3\sqrt{T}(\sqrt{t_i}+\sqrt{T})^3,\\
&\mathrm{Vomma}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\mathbb{E}[(\pi^k(a;k'+X_{t_i})_{t_i}^{T+t_i}-\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2-(\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-\bar{a}_{x'_{avg}}^{t_i,T+t_i})^2]\\
\le_c&\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3\sqrt{T}(\sqrt{t_i}+\sqrt{T})^3.
\end{aligned}
$$

Combining (3.25)-(3.26)-(3.27), Lemmas 3.3.2.1-3.3.2.2-3.3.2.3 and identity (3.38) yields that:

$$
\begin{aligned}
&\mathrm{Call}^{\mathrm{FS},A}(e^{x_0},t_i,T,e^k) \qquad (3.28)\\
=&\mathrm{Call}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\\
&+\mathrm{Vega}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\Big\{\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{2\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}[-\mathcal{V}(a;x_0)_0^{t_i}+C_1(a;x_0)_0^{t_i}]\\
&+[\frac{C_5(a;x'_{avg})_{t_i}^{T+t_i}}{2\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}-\frac{C_6(a;x'_{avg})_{t_i}^{T+t_i}}{(\bar{a}_{x'_{avg}}^{t_i,T+t_i})^3T^2}][\mathcal{V}(a;x_0)_0^{t_i}+\frac{1}{4}\mathcal{V}^2(a;x_0)_0^{t_i}]
\end{aligned}
$$

$$+\Big[\frac{k^2}{(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T}-\frac{(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T}{4}\Big]\frac{C_6(a;x'_{avg})_{t_i}^{T+t_i}}{(\bar{a}_{x'_{avg}}^{t_i,T+t_i})^2\pi^k(a;x'_{avg})_{t_i}^{T+t_i}T^2}\mathcal{V}(a;x_0)_0^{t_i}$$
$$+\frac{k}{2}\Big[\frac{C_2(\tilde{a};x'_{avg})_{t_i}^{T+t_i}-C_2(a;x'_{avg})_{t_i}^{T+t_i}}{2(\bar{a}_x^{t_i,T+t_i})^3T^2}+3\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}(C_1(a;x'_{avg})_{t_i}^{T+t_i}-C_1(\tilde{a};x'_{avg})_{t_i}^{T+t_i})}{2(\bar{a}_x^{t_i,T+t_i})^5T^3}\Big]\mathcal{V}(a;x_0)_0^{t_i}\Big\}$$
$$+\frac{1}{2}\text{Vomma}^{\text{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)\Big(\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{2\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}\mathcal{V}(a;x_0)_0^{t_i}\Big)^2$$
$$+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3\sqrt{T}(\sqrt{t_i}+\sqrt{T})^3),$$

what allows to easily achieve the proof using Lemma 3.5.1.1 and the magnitude of the terms. □

3.3.3 Forward implied volatility expansions of type B

In this Subsection we provide approximations of the forward implied volatilities related to the pricing of forward start options of type B. We introduce the stopped process $(\widetilde{X}_t^{t_i})_{t\in[0,t_i+T]}=(X_{t\wedge t_i})_{t\in[0,t_i+T]}$ and we interpret $\frac{e^{\widetilde{X}_{t_i+T}^{t_i}}}{e^{x_0}}=\frac{e^{X_{t_i}}}{e^{x_0}}$ as a Radon-Nikodym derivative of a new measure $\widetilde{\mathbb{P}}$ w.r.t. $\mathbb{P}$ on $\mathcal{F}_T$ under which

$$\widetilde{W}_t=W_t-\int_0^{t\wedge t_i}\sigma(s,S_s)\mathrm{d}s, \tag{3.29}$$

is a standard Brownian motion. Thus we can write the value of a forward start call option of type B as:

$$\text{Call}^{\text{FS},B}(S_0,t_i,T,K)=\mathbb{E}[(e^{X_{t_i+T}}-e^{k+X_{t_i}})_+]=\mathbb{E}[e^{\widetilde{X}_{t_i+T}^{t_i}}(e^{X_{t_i+T}-X_{t_i}}-e^k)_+]=e^{x_0}\widetilde{\mathbb{E}}[(e^{X_{t_i+T}-X_{t_i}}-e^k)_+]. \tag{3.30}$$

Thus we are reduced to compute similar expectations that those involved by the forward start options of type A but under the new measure $\widetilde{\mathbb{P}}$. This leads to the next Theorem, the proof being given in Appendix 3.5.3:

Theorem 3.3.3.1. ***(2nd and 3rd order expansions of the forward implied volatility of type B)****. Assume $(\mathcal{H}_a)$ and suppose that $\mathcal{M}_1(a)$, $\mathcal{M}_0(a)$, T, t_i and $|k|$ are globally small enough to ensure the existence of $C_1(k)>0$, $C_5(k)>0$, $C_3(k)>0$ and $C_6(k)>0$ such that:*

$$\bar{\alpha}_x^{0,T}-k\pi_1(\alpha;x)_0^T=\bar{a}_x^{t_i,T+t_i}-k\pi_1(a;x)_{t_i}^{T+t_i}>C_1(k)\mathcal{M}_0(a)>0,\quad\forall x\in\mathbb{R},$$
$$\tilde{\sigma}_{\text{I,F,B}}^{2,x'_{avg}}(x_0,t_i,T,k)=\bar{a}_{x'_{avg}}^{t_i,T+t_i}+\frac{1}{2}\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}\mathcal{V}(a;x_0)_0^{t_i}-k\pi_1(a;x'_{avg})_{t_i}^{T+t_i}>C_5(k)\mathcal{M}_0(a)>0, \tag{3.31}$$
$$\pi^k(\alpha;x)_0^T=\pi^k(a;x)_{t_i}^{t_i+T}=\pi_0(\alpha;x)_0^T-k\pi_1(\alpha;x)_0^T+k^2\pi_2(\alpha;x)_0^T>C_3(k)\mathcal{M}_0(a)>0,\quad\forall x\in\mathbb{R},$$
$$\tilde{\sigma}_{\text{I,F,B}}^{3,x'_{avg}}(x_0,t_i,T,k) \tag{3.32}$$
$$=\pi^k(a;x'_{avg})_{t_i}^{T+t_i}+\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{2\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}[+\mathcal{V}(a;x_0)_0^{t_i}+C_1(a;x_0)_0^{t_i}] \tag{3.33}$$
$$+\Big[\frac{C_5(a;x'_{avg})_{t_i}^{T+t_i}}{2\bar{a}_{x'_{avg}}^{t_i,T+t_i}T}-\frac{C_6(a;x'_{avg})_{t_i}^{T+t_i}}{(\bar{a}_{x'_{avg}}^{t_i,T+t_i})^3T^2}\Big][\mathcal{V}(a;x_0)_0^{t_i}+\frac{1}{4}\mathcal{V}^2(a;x_0)_0^{t_i}]$$
$$+\Big[\frac{k^2}{(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T}-\frac{(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T}{4}\Big]\frac{C_6(a;x'_{avg})_{t_i}^{T+t_i}}{(\bar{a}_{x'_{avg}}^{t_i,T+t_i})^2\pi^k(a;x'_{avg})_{t_i}^{T+t_i}T^2}\mathcal{V}(a;x_0)_0^{t_i}$$

$$+\frac{k}{2}\Big[\frac{C_2(\tilde{a};x'_{avg})_{t_i}^{T+t_i}-C_2(a;x'_{avg})_{t_i}^{T+t_i}}{2(\bar{a}_x^{t_i,T+t_i})^3T^2}+3\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}(C_1(a;x'_{avg})_{t_i}^{T+t_i}-C_1(\tilde{a};x'_{avg})_{t_i}^{T+t_i})}{2(\bar{a}_x^{t_i,T+t_i})^5T^3}\Big]\mathcal{V}(a;x_0)_0^{t_i}$$
$$>C_6(k)\mathcal{M}_0(a)>0.$$

Then $\tilde{\sigma}_{\mathrm{I,F,B}}^{2,x'_{avg}}(x_0,t_i,T,k)$ *and* $\tilde{\sigma}_{\mathrm{I,F,B}}^{3,x'_{avg}}(x_0,t_i,T,k)$ *are respectively a second and a third order approximation of* $\sigma_{\mathrm{I,F,B}}(x_0,t_i,T,k)$ *in the following sense:*

$$\mathrm{Call}^{\mathrm{FS},B}(e^{x_0},t_i,T,e^k)=\mathrm{Call}^{\mathrm{BS}}\Big(x_0,(\tilde{\sigma}_{\mathrm{I,F,B}}^{2,x'_{avg}}(x_0,t_i,T,k))^2T,x_0+k\Big)+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2\sqrt{T}(\sqrt{t_i}+\sqrt{T})^2),\tag{3.34}$$

$$\mathrm{Call}^{\mathrm{FS},B}(e^{x_0},t_i,T,e^k)=\mathrm{Call}^{\mathrm{BS}}\Big(x_0,(\tilde{\sigma}_{\mathrm{I,F,B}}^{3,x'_{avg}}(x_0,t_i,T,k))^2T,x_0+k\Big)+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3\sqrt{T}(\sqrt{t_i}+\sqrt{T})^3),\tag{3.35}$$

where the constants in the above estimates notably depend on $C_1(k)$*,* $C_3(k)$*,* $C_5(k)$*, and* $C_6(k)$*.*

3.4 Numerical Experiments

To illustrate the accuracy of our formulas, we choose the time-homogeneous CEV model for the stock value which the dynamic is given by:

$$\mathrm{d}S_t=\nu S_t^{\beta}\mathrm{d}W_t,\ S_0>0,$$

with $\beta\in[0,1]$. Although the asset price S_t can reach 0 with a probability strictly positive and although the hypotheses of boundedness and ellipticity are not fulfilled, we reasonably expect that our approximation formulas remain valid for this model and we consider a fictitious log-asset with local volatility given by

$$a(x)=\nu e^{(\beta-1)x}.$$

Thus our forward implied volatility approximation formulas read applying Proposition 3.5.4.1:

$$\tilde{\sigma}_{\mathrm{I,F,C}}^{2,x'_{avg}}(x_0,t_i,T,k)=\nu(S_0^2K)^{\frac{\beta-1}{2}}\{1+(\mathbb{1}_{C=B}-\mathbb{1}_{C=A})\frac{(\beta-1)\nu^2t_i}{2}S_0^{2(\beta-1)}\},\tag{3.36}$$

$$\begin{aligned}\pi^k(a;x'_{avg})_{t_i}^{t_i+T}=&\nu(S_0^2K)^{\frac{\beta-1}{2}}\Big\{1+\frac{(\beta-1)^2\nu^2T}{24}(S_0^2K)^{\beta-1}(1-\frac{\nu^2T(S_0^2K)^{\beta-1}}{4})-\frac{(\beta-1)^2}{24}\log^2(K)\Big\},\\ \tilde{\sigma}_{\mathrm{I,F,C}}^{3,x'_{avg}}(x_0,t_i,T,k)=&\pi^k(a;x'_{avg})_{t_i}^{t_i+T}+\frac{(\beta-1)}{2}\nu^3t_i(S_0^6K)^{\frac{\beta-1}{2}}\Big\{\beta-2\mathbb{1}_{C=A}+\frac{3(\beta-1)\nu^2t_i}{4}S_0^{2(\beta-1)}\\ &+\frac{(\beta-1)\nu(S_0^2K)^{\frac{\beta-1}{2}}}{\pi^k(a;x'_{avg})_{t_i}^{t_i+T}}\Big[\frac{\log^2(K)}{(\pi^k(a;x'_{avg})_{t_i}^{t_i+T})^2T}-\frac{(\pi^k(a;x'_{avg})_{t_i}^{t_i+T})^2T}{4}\Big]\Big\},\end{aligned}\tag{3.37}$$

$\forall C\in\{A,B\}$. From the formula (3.36), it seems that at least around the money (i.e. $K=1$) and for short maturity (i.e. $T<<1$), the forward implied volatility of type A is higher than the corresponding vanilla implied volatility (i.e. $t_i=0$) which is itself greater than the forward implied volatility of type B. From the third order approximation formula (3.37), we can expect that for options reasonably far from the money, the forward implied volatility of type A still remains higher than the forward implied volatility of type B. The term $[\frac{\log^2(K)}{(\pi^k(a;x'_{avg})_{t_i}^{t_i+T})^2T}-\frac{(\pi^k(a;x'_{avg})_{t_i}^{t_i+T})^2T}{4}]$ is negative around the money but can become positive far from the money. This may suggest that it is possible for the forward implied volatility of type B to be higher than the vanilla implied volatility far from the money. All these features are confirmed by the

following numerical tests.
For numerical experiments, we choose the values $S_0 = 1$, $\nu = 20\%$, $\beta = 0.5$ and we allow the maturities, the forward dates and the strikes to vary. We test the forward dates 0, 1M, 3M, 6M and 1Y and the maturities vary from 3M to 10Y. Then the strikes approximately behave as $e^{q\nu\sqrt{T}}$ where q denotes the value of various quantiles of the standard Normal law (from 1% to 99%) to cover around the money as well far from the money options. The set of strikes are given w.r.t. the set maturities in Table 3.1.
As a benchmark, we use Monte Carlo methods using the conditional argument which is the central idea of our proofs. More precisely for each sample path, we simulate with an Euler scheme the process:

$$dX_t = \nu e^{(\beta-1)X_t}(dW_t - \frac{1}{2}\nu e^{(\beta-1)X_t}dt),$$

under both the probabilities $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ up to the forward date t_i. Then we evaluate for the whole set of strikes and maturities $\mathbb{E}[(e^{X_{t_i+T}-X_{t_i}} - K)_+|X_{t_i}]$ and $\widetilde{\mathbb{E}}[(e^{X_{t_i+T}-X_{t_i}} - K)_+|X_{t_i}]$ using the closed-form formula of a Call in the CEV model provided by [Schroder 1989], using as level of volatility $\nu e^{(\beta-1)X_{t_i}}$. This method allows a quick and accurate estimation of the whole set of forward implied volatilities. For instance, using 10^6 sample paths and a time discretization of 300 steps by year, we obtain, using C++ on a Intel(R) Core(TM) i5 CPU@2.40GHz with 4 GB of ram, all the estimations of the forward implied volatilities for a given forward date with a computational time smaller than 2h and with confidence interval widths reduced to less than 1 bp[1] for all the strikes and maturities. We report in Tables 3.2-3.3 the estimations of the forward implied volatilities without indicating the confidence intervals for the sake of brevity and taking into account that errors on these estimates are lower than ±1 bp. On Tables 3.4, 3.5, we report the errors using the third order formulas. These errors as estimated by subtracting to the value obtained by the approximation formula the bound of the confidence interval of the Monte Carlo estimator corresponding to the worst case. That is if the value estimated by the approximation formula is greater (respectively lower) than the Monte Carlo estimator we subtract to this value the lower (respectively higher) bound of the confidence interval.

▷**Behaviours of the forward implied volatilities.** From Table 3.2, we notice that the forward implied volatility of type A is always higher than the vanilla implied volatility and is growing with t_i. Far from the money, the forward volatility is even greater: more than 1% of difference between $t_i = 0$ and $t_i = 1$Y. Regarding Table 3.3, we see that the forward implied volatility of type B decreases with t_i at the money and is increasing with t_i far from the money and is always smaller than the forward implied volatility of type A.

▷**Accuracy of the approximation formulas.** We notice in Table 3.4 that for the forward implied volatility of type A, the results are very good with a maximum error of 11.39 bps for $T = 10$Y and $t_i = 1$Y. The errors become higher when t_i, T or $|k|$ increase. Up to maturity 5Y, we observe that the approximation formula generally overestimates the true forward implied volatility for small values of t_i and then yields to an underestimation for large t_i. For the maturity 10Y, there is almost always an overestimation. The results are also very satisfying for the forward implied volatility of type B with a maximum error of 17.21 bps for $T = 10$Y and $t_i = 1$Y. The magnitudes of the errors are very close but we nevertheless observe a fairly frequent overestimation with slightly underestimations around the money for small value of T and medium values of t_i.

[1] 1 bp (basis point) is equal to 0.01%.

Table 3.1: Set of maturities and strikes for the numerical experiments

$T\backslash K$	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
3M	0.70	0.75	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.25	1.30	1.35
6M	0.65	0.75	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.20	1.25	1.35	1.50
1Y	0.55	0.65	0.75	0.80	0.90	0.95	1.00	1.05	1.15	1.25	1.40	1.50	1.80
1.5Y	0.50	0.60	0.70	0.75	0.85	0.95	1.00	1.10	1.15	1.30	1.50	1.65	2.00
2Y	0.45	0.55	0.65	0.75	0.85	0.90	1.00	1.10	1.20	1.35	1.55	1.80	2.30
3Y	0.35	0.50	0.55	0.70	0.80	0.90	1.00	1.10	1.25	1.45	1.75	2.05	2.70
5Y	0.25	0.40	0.50	0.60	0.75	0.85	1.00	1.15	1.35	1.60	2.05	2.50	3.60
10Y	0.15	0.25	0.35	0.50	0.65	0.80	1.00	1.20	1.50	1.95	2.75	3.65	6.30

3.5 Appendix

3.5.1 Results on Vega, the Vomma and the Ultima

We announce a technical result related to the Vega, the Vomma and the Ultima very useful for the proofs of the expansions .

Lemma 3.5.1.1. *Let $x,k\in\mathbb{R}$, $\nu>0$ and $T>0$. For any integer $m\geq 0$, we have:*

$$\begin{aligned}
|x-k|^m|\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,k)| &\leq_c \sqrt{T}(\nu\sqrt{T})^m,\\
|x-k|^m|\mathrm{Vomma}^{\mathrm{BS}}(x,\nu^2T,k)| &\leq_c T(\nu\sqrt{T})^{m-1},\\
|x-k|^m|\mathrm{Ultima}^{\mathrm{BS}}(x,\nu^2T,k)| &\leq_c T^{\frac{3}{2}}(\nu\sqrt{T})^{m-2},
\end{aligned}$$

where the generic constants depend polynomially on $\nu\sqrt{T}$.

Proof. For the first inequality, apply Proposition 2.6.1.3 of Chapter 2 to write that $\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,k) = \nu T(\partial^2_{x^2}-\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\nu^2T,k)$ and conclude with 2.6.1.1. For the second use Proposition 2.6.1.2 to write that:

$$\mathrm{Vomma}^{\mathrm{BS}}(x,\nu^2T,k) = \frac{\mathrm{Vega}^{\mathrm{BS}}(x,\nu^2T,k)}{\nu}\Big[\frac{(x-k)^2}{\nu^2T}-\frac{\nu^2T}{4}\Big] \tag{3.38}$$

and conclude again with Proposition 2.6.1.3 and Corollary 2.6.1.1.
The last inequality is handled similarly using Propositions 2.6.1.2, 2.6.1.3 and Corollary 2.6.1.1. □

3.5.2 Proofs of Lemmas 3.3.2.1-3.3.2.2-3.3.2.3

▷ **Proof of Lemma 3.3.2.1** We begin with the first assertion. Expand $\bar{a}_x^{t_i,T+t_i}$ to write:

$$\mathbb{E}[\bar{a}^{t_i,T+t_i}_{k'+X_{t_i}}-\bar{a}^{t_i,T+t_i}_{x'_{avg}}] = \frac{C_7(a;x'_{avg})^{T+t_i}_{t_i}}{\bar{a}^{t_i,T+t_i}_{x'_{avg}}T}\mathbb{E}[(X_{t_i}-x_0)]+\frac{1}{2}\partial^2_{x^2}\bar{a}^{t_i,T+t_i}_x|_{x=x'_{avg}}\mathbb{E}[(X_{t_i}-x_0)^2]+R_1, \tag{3.39}$$

$$\partial^2_{x^2}\bar{a}^{t_i,T+t_i}_x|_{x=x'_{avg}} = \frac{C_5(a;x'_{avg})^{T+t_i}_{t_i}}{\bar{a}^{t_i,T+t_i}_{x'_{avg}}T}-\frac{[C_7(a;x'_{avg})^{T+t_i}_{t_i}]^2}{(\bar{a}^{t_i,T+t_i}_{x'_{avg}})^3T^2} = \frac{C_5(a;x'_{avg})^{T+t_i}_{t_i}}{\bar{a}^{t_i,T+t_i}_{x'_{avg}}T}-2\frac{C_6(a;x'_{avg})^{T+t_i}_{t_i}}{(\bar{a}^{t_i,T+t_i}_{x'_{avg}})^3T^2}, \tag{3.40}$$

$$R_1 = \mathbb{E}\Big[(X_{t_i}-x_0)^3\int_0^1\partial^3_{x^3}\bar{a}^{t_i,T+t_i}_x|_{x=k'+\lambda X_{t_i}+(1-\lambda)x_0}\frac{(1-\lambda)^2}{2}\mathrm{d}\lambda\Big],$$

where the operators C_5, C_6 and C_7 are defined in Definition 3.2.0.1. We readily have using Lemma 3.5.1.1:

$$R_1 \mathrm{Vega}^{\mathrm{BS}}(0,(\pi^k(a;x'_{avg})_{t_i}^{T+t_i})^2T,k)=O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3t_i^{\frac{3}{2}}\sqrt{T}).$$

Then introducing the corrective process $(X_{2,t})_{t\in[0,T+t_i]}$ depending on the Gaussian process X_1 (see (3.19)):

$$\mathrm{d}X_{2,t}=2a^{(1)}(t,x_0)X_{1,t}(\mathrm{d}W_t-a(t,x_0)\mathrm{d}t),\ X_{2,0}=0,$$

we obtain using Equations (2.3)-(2.6)-(2.7)-(2.18) of Chapter 2 the following weak approximations:

$$\begin{aligned}\mathbb{E}[X_{t_i}-x_0]=&\mathbb{E}[X_{1,t_i}]+\frac{1}{2}\mathbb{E}[X_{2,t_i}]+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i^{\frac{3}{2}})\\=&-\frac{1}{2}\mathcal{V}(a;x_0)_0^{t_i}-\int_0^{t_i}a_t(x_0)a_t^{(1)}(x_0)\mathbb{E}[X_{1,t}]\mathrm{d}t+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i^{\frac{3}{2}})\\=&-\frac{1}{2}\mathcal{V}(a;x_0)_0^{t_i}+\frac{1}{2}C_1(a;x_0)_0^{t_i}+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i^{\frac{3}{2}}),\end{aligned}\tag{3.41}$$

$$\begin{aligned}\mathbb{E}[(X_{t_i}-x_0)^2]=&\mathbb{E}[X_{1,t_i}^2]+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i^{\frac{3}{2}})\\=&-\int_0^{t_i}a_t^2(x_0)\mathbb{E}[X_{1,t}]\mathrm{d}t+\mathcal{V}(a;x_0)_0^{t_i}+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i^{\frac{3}{2}})\\=&\frac{1}{4}\mathcal{V}^2(a;x_0)_0^{t_i}+\mathcal{V}(a;x_0)_0^{t_i}+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i^{\frac{3}{2}}).\end{aligned}\tag{3.42}$$

Combining (3.39)-(3.40)-(3.41)-(3.42) and Lemma 3.5.1.1 leads to the announced result.
We now pass to the second equality. Similarly we write using (3.42):

$$\begin{aligned}\mathbb{E}[(\bar a_{k'+X_{t_i}}^{t_i,T+t_i}-\bar a_{x'_{avg}}^{t_i,T+t_i})^2]=&(\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{\bar a_{x'_{avg}}^{t_i,T+t_i}T})^2\mathbb{E}[(X_{t_i}-x_0)^2]+R_2,\\=&(\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{\bar a_{x'_{avg}}^{t_i,T+t_i}T})^2(\frac{1}{4}\mathcal{V}^2(a;x_0)_0^{t_i}+\mathcal{V}(a;x_0)_0^{t_i})+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^4t_i^{\frac{3}{2}})+R_2\\=&(\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{2\bar a_{x'_{avg}}^{t_i,T+t_i}T}\mathcal{V}(a;x_0)_0^{t_i})^2+2\frac{C_6(a;x'_{avg})_{t_i}^{T+t_i}}{(\bar a_{x'_{avg}}^{t_i,T+t_i})^2T^2}\mathcal{V}(a;x_0)_0^{t_i}+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^4t_i^{\frac{3}{2}}),\end{aligned}\tag{3.43}$$

$$\begin{aligned}R_2=&\mathbb{E}[R_3^2+2R_3\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}}{\bar a_{x'_{avg}}^{t_i,T+t_i}T}(X_{t_i}-x_0)]=O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^4t_i^{\frac{3}{2}}),\\R_3=&(X_{t_i}-x_0)^2\int_0^1\partial_{x^2}^2\bar a_x^{t_i,T+t_i}|_{x=k'+\lambda X_{t_i}+(1-\lambda)x_0}(1-\lambda)\mathrm{d}\lambda.\end{aligned}$$

We achieve the proof with (3.43) and Lemma 3.5.1.1.
▷ **Step 2: Expansion of** $\pi_1(a;x)_{t_i}^{T+t_i}$**.** We have using the definition of π_1:

$$\mathbb{E}[\pi_1(a;k'+X_{t_i})_{t_i}^{T+t_i}-\pi_1(a;x'_{avg})_{t_i}^{T+t_i}]=\partial_x\pi_1(a;x)_{t_i}^{T+t_i}|_{x=x'_{avg}}\mathbb{E}[X_{t_i}-x_0]+R_4\tag{3.44}$$

$$\partial_x\pi_1(a;x)_{t_i}^{T+t_i}=\frac{1}{2}(\partial_x\frac{C_1(\tilde a;x)_{t_i}^{T+t_i}}{(\bar a_x^{t_i,T+t_i})^3T^2}-\partial_x\frac{C_1(a;x)_{t_i}^{T+t_i}}{(\bar a_x^{t_i,T+t_i})^3T^2}),\tag{3.45}$$

$$\partial_x\frac{C_1(a;x)_{t_i}^{T+t_i}}{(\bar a_x^{t_i,T+t_i})^3T^2}=\frac{C_2(a;x)_{t_i}^{T+t_i}+2C_6(a;x)_{t_i}^{T+t_i}}{(\bar a_x^{t_i,T+t_i})^3T^2}-3\frac{C_7(a;x)_{t_i}^{T+t_i}C_1(a;x)_{t_i}^{T+t_i}}{(\bar a_x^{t_i,T+t_i})^5T^3},\tag{3.46}$$

$$R_4=\mathbb{E}[(X_{t_i}-x_0)^2$$

$$\times \int_0^1 \partial^2_{x^2}\pi_1(a;x)_{t_i}^{T+t_i}|_{x=k'+\lambda X_{t_i}+(1-\lambda)x_0}(1-\lambda)\mathrm{d}\lambda].$$
$$=O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 t_i). \tag{3.47}$$

Hence using (3.18)-(3.45)-(3.46)-(3.47) and the symmetry of the operator C_6, we get for (3.44):

$$\begin{aligned}&\mathbb{E}[\pi_1(a;k'+X_{t_i})_{t_i}^{T+t_i}-\pi_1(a;x'_{avg})_{t_i}^{T+t_i}]\\ =&-\frac{1}{2}\mathcal{V}(a;x_0)_0^{t_i}[\frac{C_2(\tilde{a};x'_{avg})_{t_i}^{T+t_i}-C_2(a;x'_{avg})_{t_i}^{T+t_i}}{2(\bar{a}_x^{t_i,T+t_i})^3T^2}+3\frac{C_7(a;x'_{avg})_{t_i}^{T+t_i}(C_1(a;x'_{avg})_{t_i}^{T+t_i}-C_1(\tilde{a};x'_{avg})_{t_i}^{T+t_i})}{2(\bar{a}_x^{t_i,T+t_i})^5T^3}]\\ &+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i).\end{aligned}$$

We conclude with an application of Lemma 3.5.1.1.

▷ **Proof of Lemma 3.3.2.3** Although the explicit form of the residuals is tedious to write, there is no extra difficulty so we let it as an exercise to the reader.

3.5.3 Proof of Theorem 3.3.3.1

We only detail the main differences with the proof of Theorems 3.3.1.1-3.3.2.1. First notice that:

$$X_{t_i+T}-X_{t_i}=\int_{t_i}^{t_i+T}a(t,X_t)(\mathrm{d}W_t-\frac{1}{2}a(t,X_t)\mathrm{d}t)=X_{t_i+T}-X_{t_i}=\int_{t_i}^{t_i+T}a(t,X_t)(\mathrm{d}\widetilde{W}_t-\frac{1}{2}a(t,X_t)\mathrm{d}t),$$

in view of the definition of $\widetilde{W}$ (see (3.29)). Thus combining this observation with (3.30), the same conditional expectation argument previously employed yields:

$$\begin{aligned}\mathrm{Call}^{\mathrm{FS},A}(e^{x_0},t_i,T,e^k)=&\widetilde{\mathbb{E}}[\mathrm{Call}^{\mathrm{BS}}(x_0,(\bar{a}_{k'+X_{t_i}}^{t_i,T+t_i}-k\pi_1(a;k'+X_{t_i})_{t_i}^{T+t_i})^2T,x_0+k)]+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2T^{\frac{3}{2}},\\ =&\widetilde{\mathbb{E}}[\mathrm{Call}^{\mathrm{BS}}(x_0,(\pi^k(a;k'+X_{t_i})_{t_i}^{T+t_i})^2T,x_0+k)]+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3T^2).\end{aligned}$$

Next few changes occur after the volatility expansion when computing the expectations $\widetilde{\mathbb{E}}[(X_{t_i}-x_0)],\widetilde{\mathbb{E}}[(X_{t_i}-x_0)^2]$. The dynamic of $(X_t)_{t\in[0,t_i]}$ under $\widetilde{\mathbb{P}}$ is given by :

$$\mathrm{d}X_t=a(t,X_t)(\mathrm{d}\widetilde{W}_t+\frac{1}{2}a(t,X_t)\mathrm{d}t),\ X_0=x_0.$$

The following weak approximations come with $\begin{cases}\mathrm{d}X_{1,t}=a(t,x_0)\mathrm{d}\widetilde{W}_t+\frac{1}{2}a^2(t,x_0)\mathrm{d}t,\ X_{1,0}=0,\\ \mathrm{d}X_{2,t}=2a^{(1)}(t,x_0)X_{1,t}(\mathrm{d}\widetilde{W}_t+a(t,x_0)\mathrm{d}t),\ X_{2,0}=0,\end{cases}$

$$\begin{aligned}\widetilde{\mathbb{E}}[X_{t_i}-x_0]=&\widetilde{\mathbb{E}}[X_{1,t_i}]+O(\mathcal{M}_1(a)\mathcal{M}_0(a)t_i)=+\frac{1}{2}\mathcal{V}(a;x_0)_0^{t_i}+O(\mathcal{M}_1(a)\mathcal{M}_0(a)t_i)\\ =&\widetilde{\mathbb{E}}[X_{1,t_i}]+\frac{1}{2}\widetilde{\mathbb{E}}[X_{2,t_i}]+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i^{\frac{3}{2}})=+\frac{1}{2}\mathcal{V}(a;x_0)_0^{t_i}+\frac{1}{2}C_1(a;x_0)_0^{t_i}+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i^{\frac{3}{2}}),\\ \widetilde{\mathbb{E}}[(X_{t_i}-x_0)^2]=&\widetilde{\mathbb{E}}[X_{1,t_i}^2]+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i^{\frac{3}{2}})=\frac{1}{4}\mathcal{V}^2(a;x_0)_0^{t_i}+\mathcal{V}(a;x_0)_0^{t_i}+O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2t_i^{\frac{3}{2}}).\end{aligned}$$

Taking into account these modifications allows to easily complete the proof.

3.5.4 Forward implied volatilities in time-independent local volatility models

We give in this Subsection the explicit form of our forward implied volatility expansions given in Theorems 3.3.1.1-3.3.2.1-3.3.3.1 when a is time-independent.

Proposition 3.5.4.1. *Under the hypotheses of Theorems 3.3.1.1-3.3.2.1-3.3.3.1, one has if a is time-independent for any $C \in \{A, B\}$:*

$$\tilde{\sigma}^{2,x'_{avg}}_{\mathrm{I,F,C}}(x_0,t_i,T,k) = a(x'_{avg}) + (\mathbb{1}_{C=B} - \mathbb{1}_{C=A})\frac{1}{2}a^{(1)}(x'_{avg})a^2(x_0)t_i, \tag{3.48}$$

$$\pi^k(a;x)^{t_i+T}_{t_i} = a(x)\Big\{1 + T\Big[\frac{a(x)a^{(2)}(x)}{12} - (a^{(1)})^2(x)\Big(\frac{1}{24} + \frac{a^2(x)T}{96}\Big)\Big] + k^2\Big[\frac{a^{(2)}(x)}{24a(x)} - \frac{(a^{(1)})^2(x)}{12a^2(x)}\Big]\Big\}, \tag{3.49}$$

$$\begin{aligned}\tilde{\sigma}^{3,x'_{avg}}_{\mathrm{I,F,C}}(x_0,t_i,T,k) =& \pi^k(a;x'_{avg})^{t_i+T}_{t_i} + a^2(x_0)t_i\Big\{\frac{a^{(1)}(x'_{avg})}{2}\Big[(\mathbb{1}_{C=B} - \mathbb{1}_{C=A}) + \frac{a(x_0)a^{(1)}(x_0)t_i}{2}\Big]\\ &+ \frac{a^{(2)}(x'_{avg})}{2}\Big[1 + \frac{a^2(x_0)t_i}{4}\Big] + \frac{(a^{(1)}(x'_{avg}))^2}{2\pi^k(a;x'_{avg})^{T+t_i}_{t_i}}\Big[\frac{k^2}{(\pi^k(a;x'_{avg})^{T+t_i}_{t_i})^2 T} - \frac{(\pi^k(a;x'_{avg})^{T+t_i}_{t_i})^2 T}{4}\Big]\Big\}.\end{aligned} \tag{3.50}$$

Table 3.2: CEV model ($\beta = 0.5, \nu = 0.2$): BS forward implied volatilities of type A in % estimated by Monte Carlo.

T	$t_i \backslash K$	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
3M	0	21.84	21.48	21.14	20.83	20.53	20.26	20.00	19.76	19.53	19.31	18.91	18.72	18.54
	1M	21.96	21.56	21.20	20.87	20.57	20.29	20.03	19.78	19.56	19.35	18.97	18.80	18.64
	3M	22.21	21.75	21.34	20.97	20.64	20.34	20.08	19.84	19.63	19.44	19.11	18.97	18.85
	6M	22.59	22.03	21.54	21.12	20.75	20.43	20.16	19.92	19.73	19.56	19.32	19.23	19.17
	1Y	23.35	22.59	21.95	21.41	20.96	20.60	20.31	20.09	19.93	19.82	19.74	19.75	19.80
6M	0	22.24	21.48	21.14	20.83	20.54	20.26	20.00	19.76	19.53	19.11	18.91	18.54	18.04
	1M	22.33	21.53	21.19	20.86	20.56	20.29	20.03	19.79	19.56	19.14	18.95	18.60	18.14
	3M	22.53	21.66	21.28	20.94	20.63	20.34	20.08	19.84	19.62	19.22	19.05	18.73	18.34
	6M	22.83	21.84	21.42	21.05	20.72	20.42	20.16	19.92	19.71	19.34	19.19	18.93	18.64
	1Y	23.43	22.20	21.71	21.28	20.91	20.59	20.31	20.08	19.88	19.58	19.47	19.32	19.24
1Y	0	23.15	22.24	21.48	21.15	20.54	20.27	20.01	19.77	19.32	18.91	18.37	18.05	17.21
	1M	23.24	22.30	21.52	21.18	20.57	20.29	20.03	19.79	19.34	18.95	18.42	18.11	17.31
	3M	23.43	22.43	21.61	21.26	20.63	20.34	20.08	19.84	19.40	19.02	18.52	18.23	17.53
	6M	23.72	22.63	21.75	21.37	20.71	20.43	20.16	19.92	19.49	19.12	18.67	18.42	17.85
	1Y	24.29	23.01	22.01	21.59	20.88	20.58	20.32	20.08	19.67	19.34	18.97	18.79	18.49
1.5Y	0	23.69	22.68	21.85	21.49	20.84	20.27	20.01	19.54	19.32	18.73	18.05	17.61	16.74
	1M	23.77	22.74	21.89	21.52	20.87	20.30	20.04	19.56	19.35	18.76	18.10	17.67	16.84
	3M	23.95	22.86	21.98	21.60	20.93	20.35	20.09	19.62	19.40	18.83	18.20	17.80	17.04
	6M	24.21	23.05	22.12	21.72	21.02	20.43	20.17	19.70	19.49	18.93	18.34	17.98	17.35
	1Y	24.74	23.42	22.39	21.95	21.20	20.59	20.32	19.86	19.65	19.14	18.64	18.36	17.96
2Y	0	24.29	23.16	22.25	21.49	20.84	20.55	20.02	19.54	19.12	18.55	17.90	17.22	16.13
	1M	24.37	23.22	22.30	21.53	20.87	20.57	20.04	19.57	19.14	18.58	17.94	17.28	16.24
	3M	24.54	23.35	22.39	21.60	20.93	20.63	20.09	19.62	19.20	18.65	18.04	17.41	16.46
	6M	24.81	23.54	22.53	21.71	21.02	20.71	20.17	19.70	19.29	18.75	18.17	17.60	16.80
	1Y	25.33	23.91	22.81	21.92	21.19	20.88	20.32	19.85	19.46	18.96	18.45	17.99	17.46
3Y	0	25.76	23.71	23.18	21.87	21.17	20.56	20.02	19.55	18.93	18.22	17.35	16.64	15.45
	1M	25.85	23.76	23.22	21.90	21.19	20.58	20.05	19.57	18.95	18.25	17.39	16.70	15.56
	3M	26.03	23.88	23.33	21.98	21.26	20.64	20.10	19.63	19.01	18.32	17.49	16.83	15.78
	6M	26.32	24.06	23.48	22.09	21.35	20.72	20.18	19.70	19.09	18.42	17.63	17.03	16.11
	1Y	26.88	24.42	23.80	22.30	21.53	20.88	20.33	19.86	19.26	18.63	17.92	17.42	16.76
5Y	0	27.83	25.01	23.73	22.72	21.52	20.87	20.04	19.35	18.57	17.77	16.65	15.79	14.29
	1M	27.92	25.06	23.78	22.76	21.55	20.89	20.06	19.37	18.60	17.80	16.69	15.85	14.40
	3M	28.11	25.18	23.87	22.84	21.61	20.95	20.12	19.42	18.65	17.87	16.79	15.98	14.63
	6M	28.39	25.37	24.02	22.95	21.71	21.03	20.19	19.50	18.74	17.97	16.93	16.18	14.97
	1Y	28.96	25.73	24.30	23.19	21.89	21.20	20.35	19.65	18.90	18.17	17.22	16.57	15.66
10Y	0	31.20	27.91	25.86	23.79	22.33	21.23	20.08	19.17	18.10	16.89	15.40	14.25	12.21
	1M	31.28	27.97	25.91	23.82	22.36	21.25	20.10	19.19	18.12	16.92	15.45	14.31	12.33
	3M	31.46	28.10	26.01	23.90	22.43	21.31	20.15	19.24	18.18	16.99	15.54	14.45	12.60
	6M	31.72	28.29	26.16	24.02	22.52	21.39	20.23	19.32	18.26	17.09	15.69	14.66	12.99
	1Y	32.23	28.67	26.46	24.25	22.71	21.56	20.38	19.47	18.42	17.29	15.98	15.07	13.77

Table 3.3: CEV model ($\beta = 0.5, \nu = 0.2$): BS forward implied volatilities of type B in % estimated by Monte Carlo.

T	$t_i \backslash K$	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
3M	0	21.84	21.48	21.14	20.83	20.53	20.26	20.00	19.76	19.53	19.31	18.91	18.72	18.54
	1M	21.92	21.53	21.17	20.84	20.53	20.25	19.99	19.75	19.53	19.32	18.94	18.77	18.61
	3M	22.10	21.64	21.23	20.86	20.54	20.24	19.98	19.74	19.53	19.34	19.01	18.88	18.76
	6M	22.36	21.80	21.32	20.90	20.54	20.22	19.95	19.72	19.53	19.36	19.12	19.04	18.97
	1Y	22.85	22.11	21.49	20.97	20.53	20.18	19.90	19.68	19.52	19.41	19.32	19.34	19.38
6M	0	22.24	21.48	21.14	20.83	20.54	20.26	20.00	19.76	19.53	19.11	18.91	18.54	18.04
	1M	22.29	21.50	21.15	20.83	20.53	20.25	20.00	19.75	19.53	19.11	18.92	18.57	18.11
	3M	22.42	21.55	21.17	20.83	20.52	20.24	19.98	19.74	19.52	19.13	18.95	18.64	18.25
	6M	22.59	21.62	21.21	20.84	20.51	20.22	19.96	19.72	19.51	19.15	18.99	18.74	18.45
	1Y	22.93	21.74	21.26	20.84	20.48	20.17	19.90	19.67	19.48	19.18	19.07	18.92	18.83
1Y	0	23.15	22.24	21.48	21.15	20.54	20.27	20.01	19.77	19.32	18.91	18.37	18.05	17.21
	1M	23.20	22.27	21.49	21.15	20.53	20.26	20.00	19.76	19.31	18.91	18.39	18.08	17.28
	3M	23.31	22.32	21.50	21.15	20.52	20.24	19.98	19.74	19.31	18.92	18.43	18.14	17.44
	6M	23.48	22.39	21.53	21.15	20.50	20.22	19.96	19.72	19.29	18.93	18.48	18.23	17.66
	1Y	23.78	22.54	21.56	21.15	20.46	20.17	19.90	19.67	19.27	18.94	18.58	18.40	18.09
1.5Y	0	23.69	22.68	21.85	21.49	20.84	20.27	20.01	19.54	19.32	18.73	18.05	17.61	16.74
	1M	23.73	22.70	21.86	21.49	20.83	20.26	20.00	19.53	19.32	18.73	18.07	17.64	16.81
	3M	23.82	22.75	21.87	21.49	20.82	20.25	19.99	19.52	19.30	18.73	18.10	17.71	16.96
	6M	23.96	22.82	21.90	21.50	20.81	20.22	19.96	19.50	19.29	18.74	18.16	17.80	17.17
	1Y	24.22	22.94	21.93	21.50	20.77	20.17	19.91	19.45	19.25	18.75	18.25	17.98	17.58
2Y	0	24.29	23.16	22.25	21.49	20.84	20.55	20.02	19.54	19.12	18.55	17.90	17.22	16.13
	1M	24.33	23.18	22.26	21.49	20.84	20.54	20.01	19.54	19.11	18.55	17.91	17.25	16.21
	3M	24.42	23.23	22.28	21.49	20.82	20.53	19.99	19.52	19.10	18.56	17.94	17.32	16.38
	6M	24.55	23.30	22.30	21.49	20.81	20.50	19.97	19.50	19.09	18.56	17.99	17.42	16.62
	1Y	24.80	23.42	22.34	21.47	20.76	20.45	19.91	19.45	19.06	18.57	18.07	17.62	17.08
3Y	0	25.76	23.71	23.18	21.87	21.17	20.56	20.02	19.55	18.93	18.22	17.35	16.64	15.45
	1M	25.80	23.72	23.19	21.87	21.16	20.55	20.02	19.54	18.92	18.22	17.37	16.67	15.53
	3M	25.90	23.76	23.21	21.86	21.15	20.53	20.00	19.53	18.91	18.23	17.40	16.75	15.70
	6M	26.05	23.81	23.25	21.86	21.13	20.51	19.97	19.51	18.90	18.24	17.45	16.85	15.94
	1Y	26.32	23.91	23.31	21.85	21.09	20.46	19.92	19.45	18.87	18.25	17.55	17.06	16.40
5Y	0	27.83	25.01	23.73	22.72	21.52	20.87	20.04	19.35	18.57	17.77	16.65	15.79	14.29
	1M	27.87	25.02	23.74	22.72	21.52	20.86	20.03	19.34	18.57	17.77	16.67	15.82	14.38
	3M	27.96	25.06	23.75	22.72	21.50	20.85	20.01	19.32	18.56	17.78	16.70	15.90	14.56
	6M	28.10	25.11	23.77	22.72	21.49	20.82	19.99	19.30	18.55	17.79	16.76	16.01	14.82
	1Y	28.35	25.20	23.80	22.71	21.44	20.77	19.93	19.25	18.52	17.80	16.86	16.22	15.32
10Y	0	31.20	27.91	25.86	23.79	22.33	21.23	20.08	19.17	18.10	16.89	15.40	14.25	12.21
	1M	31.23	27.93	25.86	23.79	22.33	21.22	20.07	19.16	18.09	16.90	15.42	14.29	12.31
	3M	31.30	27.96	25.88	23.78	22.31	21.20	20.05	19.15	18.09	16.90	15.47	14.38	12.53
	6M	31.39	28.00	25.89	23.77	22.29	21.18	20.02	19.12	18.07	16.92	15.53	14.50	12.85
	1Y	31.56	28.08	25.92	23.75	22.25	21.12	19.96	19.07	18.04	16.93	15.65	14.75	13.46

Table 3.4: CEV model ($\beta = 0.5, \nu = 0.2$): error in bps on the BS forward implied volatilities of type A using the third order formula.

T	$t_i \backslash K$	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
3M	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	1M	0.65	0.49	0.38	0.30	0.24	0.21	0.20	0.21	0.23	0.26	0.37	0.45	0.54
	3M	−0.26	−0.33	−0.35	−0.34	−0.32	−0.30	−0.29	−0.29	−0.30	−0.31	−0.31	−0.29	−0.24
	6M	−0.87	−0.99	−0.96	−0.86	−0.75	−0.68	−0.65	−0.66	−0.71	−0.78	−0.88	−0.87	−0.78
	1Y	−2.00	−2.24	−2.05	−1.69	−1.33	−1.07	−0.98	−1.04	−1.23	−1.49	−1.94	−1.96	−1.78
6M	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	1M	0.53	0.35	0.29	0.25	0.23	0.21	0.20	0.20	0.21	0.25	0.28	0.36	0.52
	3M	−0.31	−0.35	−0.34	−0.33	−0.31	−0.30	−0.29	−0.29	−0.29	−0.30	−0.31	−0.30	−0.21
	6M	−0.96	−0.94	−0.86	−0.78	−0.72	−0.67	−0.65	−0.66	−0.68	−0.76	−0.80	−0.85	−0.71
	1Y	−2.19	−1.95	−1.68	−1.40	−1.18	−1.04	−0.99	−1.01	−1.11	−1.43	−1.60	−1.85	−1.65
1Y	0	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	−0.01
	1M	0.53	0.38	0.29	0.26	0.22	0.21	0.21	0.20	0.21	0.24	0.29	0.35	0.57
	3M	−0.29	−0.35	−0.34	−0.33	−0.31	−0.30	−0.30	−0.29	−0.29	−0.30	−0.30	−0.28	0.29
	6M	−0.93	−0.96	−0.86	−0.80	−0.71	−0.68	−0.66	−0.66	−0.68	−0.73	−0.80	−0.80	−0.45
	1Y	−2.10	−2.02	−1.62	−1.42	−1.12	−1.04	−1.01	−1.02	−1.13	−1.33	−1.64	−1.75	−1.09
1.5Y	0	0.02	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	−0.01
	1M	0.52	0.39	0.30	0.27	0.23	0.21	0.21	0.21	0.21	0.24	0.29	0.36	0.56
	3M	−0.28	−0.34	−0.34	−0.34	−0.32	−0.30	−0.30	−0.29	−0.29	−0.30	−0.29	−0.26	0.33
	6M	−0.91	−0.96	−0.88	−0.82	−0.74	−0.68	−0.67	−0.67	−0.68	−0.72	−0.78	−0.76	−0.35
	1Y	−2.05	−1.99	−1.66	−1.49	−1.21	−1.05	−1.03	−1.05	−1.10	−1.31	−1.60	−1.67	−0.87
2Y	0	0.05	0.06	0.05	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.01	−0.03
	1M	0.56	0.42	0.33	0.28	0.24	0.23	0.22	0.21	0.22	0.24	0.29	0.37	0.63
	3M	0.26	−0.33	−0.34	−0.33	−0.31	−0.30	−0.29	−0.29	−0.29	−0.29	−0.28	−0.23	0.49
	6M	−0.86	−0.95	−0.89	−0.80	−0.73	−0.71	−0.67	−0.67	−0.68	−0.72	−0.76	−0.71	0.64
	1Y	−1.95	−1.98	−1.71	−1.42	−1.19	−1.11	−1.04	−1.05	−1.13	−1.31	−1.54	−1.58	1.16
3Y	0	0.15	0.15	0.13	0.09	0.07	0.06	0.05	0.05	0.04	0.04	0.03	0.01	−0.06
	1M	0.73	0.49	0.44	0.33	0.28	0.26	0.24	0.23	0.24	0.25	0.31	0.38	0.63
	3M	0.48	−0.26	−0.28	−0.30	−0.29	−0.28	−0.28	−0.28	−0.28	−0.28	−0.25	−0.18	0.60
	6M	−0.60	−0.89	−0.89	−0.80	−0.74	−0.70	−0.67	−0.66	−0.68	−0.71	−0.72	−0.62	0.93
	1Y	−1.50	−1.90	−1.81	−1.45	−1.25	−1.12	−1.06	−1.06	−1.14	−1.31	−1.51	−1.40	1.82
5Y	0	0.75	0.55	0.42	0.32	0.22	0.18	0.15	0.12	0.11	0.09	0.06	0.02	−0.18
	1M	1.39	0.90	0.69	0.55	0.43	0.37	0.33	0.30	0.29	0.30	0.34	0.41	0.68
	3M	1.23	0.61	0.43	0.33	0.25	0.22	−0.21	−0.22	−0.23	−0.23	−0.20	0.24	0.95
	6M	1.09	−0.56	−0.67	−0.69	−0.66	−0.64	−0.62	−0.62	−0.64	−0.66	−0.64	−0.43	1.84
	1Y	1.05	−1.56	−1.57	−1.43	−1.21	−1.11	−1.04	−1.05	−1.13	−1.28	−1.39	−1.03	3.93
10Y	0	7.81	4.33	2.81	1.70	1.15	0.84	0.62	0.49	0.39	0.30	0.17	−0.03	−0.80
	1M	8.57	4.74	3.09	1.91	1.33	1.02	0.78	0.66	0.56	0.49	0.46	0.45	0.65
	3M	8.75	4.50	2.80	1.65	1.11	0.83	0.62	0.50	0.39	0.31	0.29	0.44	2.03
	6M	9.19	4.20	2.37	1.24	0.76	0.52	0.34	−0.32	−0.41	−0.47	−0.40	0.55	4.95
	1Y	10.86	4.10	1.94	0.82	−0.49	−0.61	−0.70	−0.78	−0.92	−1.10	−0.99	0.86	11.39

Table 3.5: CEV model ($\beta = 0.5, \nu = 0.2$): error in bps on the BS forward implied volatilities of type B using the third order formula.

T	$t_i \backslash K$	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
3M	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	1M	0.54	0.37	0.25	0.17	0.12	−0.14	−0.14	−0.13	0.11	0.15	0.27	0.35	0.44
	3M	0.58	0.38	0.28	0.22	−0.21	−0.21	−0.21	−0.21	−0.20	0.20	0.28	0.36	0.47
	6M	1.39	0.79	0.48	0.33	−0.31	−0.32	−0.32	−0.32	−0.30	0.29	0.52	0.76	1.13
	1Y	5.57	3.41	2.20	1.57	1.27	1.14	1.10	1.11	1.19	1.38	2.34	3.26	4.57
6M	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	1M	0.41	0.22	0.17	0.13	−0.13	−0.14	−0.14	−0.14	−0.12	0.14	0.17	0.25	0.43
	3M	0.44	0.26	0.22	−0.21	−0.21	−0.21	−0.21	−0.21	−0.20	0.19	0.21	0.27	0.47
	6M	0.91	0.41	0.32	−0.32	−0.33	−0.33	−0.32	−0.32	−0.31	−0.27	0.31	0.49	1.09
	1Y	3.80	1.94	1.54	1.31	1.18	1.12	1.09	1.09	1.12	1.31	1.50	2.20	4.34
1Y	0	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	−0.01
	1M	0.40	0.25	0.16	0.13	−0.13	−0.14	−0.14	−0.14	−0.12	0.12	0.19	0.24	0.48
	3M	0.45	0.28	0.22	−0.21	−0.21	−0.21	−0.21	−0.20	−0.20	−0.19	0.22	0.26	0.58
	6M	0.89	0.45	−0.31	−0.33	−0.33	−0.33	−0.32	−0.32	−0.31	−0.29	0.34	0.47	1.41
	1Y	3.64	2.08	1.45	1.30	1.15	1.11	1.09	1.09	1.11	1.21	1.60	2.10	5.25
1.5Y	0	0.02	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	−0.01
	1M	0.38	0.25	0.17	0.15	−0.12	−0.13	−0.13	−0.13	−0.12	0.12	0.19	0.25	0.47
	3M	0.43	0.29	0.23	0.22	−0.21	−0.20	−0.20	−0.20	−0.19	0.19	0.22	0.28	0.59
	6M	0.81	0.44	−0.31	−0.33	−0.33	−0.32	−0.32	−0.31	−0.31	−0.29	0.34	0.51	1.42
	1Y	3.28	2.01	1.46	1.33	1.18	1.11	1.10	1.08	1.09	1.18	1.58	2.20	5.16
2Y	0	0.05	0.06	0.05	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.01	−0.03
	1M	0.41	0.28	0.20	0.15	0.12	−0.12	−0.12	−0.12	−0.11	0.13	0.18	0.27	0.54
	3M	0.47	0.32	0.25	0.23	0.21	0.21	0.20	0.19	0.19	0.19	0.21	0.30	0.77
	6M	0.84	0.47	0.33	−0.33	−0.32	−0.32	−0.31	−0.31	−0.30	−0.28	0.31	0.56	1.89
	1Y	3.30	2.06	1.52	1.28	1.17	1.14	1.10	1.08	1.09	1.16	1.47	2.36	6.48
3Y	0	0.15	0.15	0.13	0.09	0.07	0.06	0.05	0.05	0.04	0.04	0.03	0.01	−0.06
	1M	0.55	0.35	0.30	0.20	0.16	0.14	0.13	0.12	0.13	0.15	0.20	0.28	0.54
	3M	0.67	0.39	0.34	0.28	0.25	0.24	0.23	0.22	0.21	0.20	0.23	0.33	0.86
	6M	1.16	0.50	0.42	0.32	−0.30	−0.29	−0.29	−0.28	−0.28	−0.26	0.35	0.63	2.12
	1Y	3.99	1.99	1.72	1.32	1.22	1.16	1.13	1.10	1.09	1.15	1.57	2.52	6.97
5Y	0	0.75	0.55	0.42	0.32	0.22	0.18	0.15	0.12	0.11	0.09	0.06	0.02	−0.18
	1M	1.18	0.75	0.56	0.43	0.31	0.26	0.22	0.20	0.19	0.20	0.24	0.31	0.58
	3M	1.37	0.78	0.60	0.49	0.40	0.36	0.32	0.29	0.26	0.24	0.27	0.39	1.18
	6M	1.93	0.88	0.64	0.52	0.43	0.40	0.36	0.33	0.29	0.28	0.40	0.76	3.02
	1Y	4.99	2.40	1.83	1.55	1.36	1.29	1.21	1.15	1.11	1.15	1.63	2.81	9.22
10Y	0	7.81	4.33	2.81	1.70	1.15	0.84	0.62	0.49	0.39	0.30	0.17	−0.03	−0.80
	1M	8.27	4.55	2.95	1.80	1.23	0.92	0.69	0.56	0.47	0.40	0.36	0.35	0.54
	3M	8.64	4.62	2.98	1.86	1.32	1.02	0.79	0.66	0.53	0.43	0.40	0.56	2.23
	6M	9.43	4.78	3.02	1.87	1.35	1.06	0.83	0.69	0.55	0.45	0.58	1.20	6.17
	1Y	13.11	6.65	4.34	2.93	2.31	1.97	1.70	1.51	1.33	1.28	1.91	3.85	17.21

CHAPTER 4

Discussion on the parameterization and on the proxy model

Contents

In this Chapter, we discuss the choice of the parameterization and of the proxy model. First we show that different parameterizations can lead to the same expansions and we compare their advantages and drawbacks. Then we see how to directly obtain a log-normal proxy model without performing the classical exponential change of function for the payoff function and for the diffusion coefficients. This allows a comparison between the use of a normal proxy on the log-asset and the use of a log-normal proxy on the asset. Next we investigate the use of a displaced log-normal proxy in the time homogeneous framework and we finally gather some numerical results illustrating our discussions. We keep the notations of the Chapters 1 (see Section 1.2) and 2 (see Subsection 2.1.1 and Definition 2.3.1.1).

4.1 Revisiting the parameterization

In this section we assume that the local volatility function a is bounded, infinitely differentiable with bounded derivatives. We recall the interpolated process used to link the initial model and the proxy:

$$\mathrm{d}X_t^\eta = \eta(-\frac{1}{2}a^2(t,X_t^\eta)\mathrm{d}t + a(t,X_t^\eta)\mathrm{d}W_t), \quad X_t^\eta = x_0, \tag{4.1}$$

for an interpolation parameter η lying in $[0,1]$. We have seen that the Gaussian proxy process (obtained by freezing at x_0 the local volatility function a) is equal to $X_t^P = X_t^0 + \partial_\eta X_t^\eta|_{\eta=0}$ and thus we approximate at the order n the difference $X_t - X_t^P$ with the partial sum $\sum_{j=2}^{n}\frac{1}{j!}(\partial^j_{\eta^j}X_t^\eta)|_{\eta=0}$.

We could consider instead of X_t^η, the following interpolated process:

$$\mathrm{d}\hat{X}_t^\eta = -\frac{1}{2}a^2(t,x_0+\eta(\hat{X}_t^\eta - x_0))\mathrm{d}t + a(t,x_0+\eta(\hat{X}_t^\eta - x_0))\mathrm{d}W_t, \quad \hat{X}_0^\eta = x_0. \tag{4.2}$$

Observe that this parametization is more user friendly because there is no need to derive w.r.t. η to see the link between X_t and X_t^P. Actually X_t is nothing else but $\hat{X}_t^1$ whereas $X_t^P = \hat{X}_t^0$. This parameterization

seems more natural than (4.1) and simply consists in doing an interpolation of the spatial argument of the local volatility function between X_t and x_0.
The question is: do the two interpolated processes lead to the same approximations? The answer is yes because the two processes are related by the relation $X_t^\eta = x_0 + \eta(\hat{X}_t^\eta - x_0)$. Posing $Y_t^\eta = x_0 + \eta(\hat{X}_t^\eta - x_0)$, an application of the Itô's formula namely yields to:

$$\begin{aligned}\mathrm{d}Y_t^\eta =&\eta \mathrm{d}\hat{X}^\eta = \eta(-\frac{1}{2}a^2(t, x_0 + \eta(\hat{X}_t^\eta - x_0))\mathrm{d}t + a(t, x_0 + \eta(\hat{X}_t^\eta - x_0))\mathrm{d}W_t)\\ =&\eta(-\frac{1}{2}a^2(t, Y_t^\eta)\mathrm{d}t + a(t, Y_t^\eta)\mathrm{d}W_t), \quad Y_t^\eta = 0.\end{aligned}$$

We retrieve exactly the dynamic (4.1) with the same initial condition and we conclude with an uniqueness argument.
Thus we have: $X_t^0 = 0$, $(\partial_\eta X_t^\eta)|_{\eta=0} = \hat{X}_t^0 - x_0 = X_t^P - x_0$ and for any $i \geq 2$:

$$(\partial^i_{\eta^i} X_t^\eta)|_{\eta=0} = [\partial^i_{\eta^i}\eta(\hat{X}_t^\eta - x_0)]|_{\eta=0} = i(\partial^{i-1}_{\eta^{i-1}}\hat{X}^\eta)|_{\eta=0}.$$

Consequently, if one considers a n-th order expansion, one has:

$$X_t - X_t^P \approx \sum_{j=2}^{n} \frac{1}{j!}(\partial^j_{\eta^j} X_t^\eta)|_{\eta=0} = \sum_{j=1}^{n-1} \frac{1}{j!}(\partial^j_{\eta^j}\hat{X}_t^\eta)|_{\eta=0}.$$

An other advantage of this equivalent parameterization (4.2), is that to perform a n-th order expansion, one only needs to compute the n first derivatives of $\hat{X}_t^\eta$ w.r.t. η (the n-th is for the error estimate) instead of the $n+1$ first derivatives of X_t^η w.r.t. η. As a consequence, it is sufficient to assume that the local volatility function is $C^{n+1}(\mathbb{R},\mathbb{R})$ instead of being $C^{n+2}(\mathbb{R},\mathbb{R})$ (see [Kunita 1984, Theorem 2.3]).
For all these advantages, it seems better to use the interpolated process (4.2) and we use a similar parameterization in the next Parts of the thesis.
There are doubtless many other suitable parameterizations allowing to perform similar expansions. For example we mention:

$$\mathrm{d}\tilde{X}_t^\eta = -\frac{1}{2}(\eta a^2(t, \tilde{X}^\eta) + (1-\eta)a^2(t, x_0))\mathrm{d}t + (\eta a(t, \tilde{X}_t^\eta) + (1-\eta)a(t, x_0))\mathrm{d}W_t, \quad \tilde{X}_0^\eta = x_0. \tag{4.3}$$

We have obviously $\tilde{X}_1 = X_t$ and $\tilde{X}_t^0 = X_t^P$. But this interpolation is less convenient than (4.2) because the first corrective process $(\partial_\eta \tilde{X}_t^\eta)|_{\eta=0}$ is not explicit. We have indeed with the notation $(\partial_\eta \tilde{X}_t^\eta)|_{\eta=0} = \tilde{X}_{1,t}$:

$$\mathrm{d}\tilde{X}_{1,t} = -\frac{1}{2}(a^2(t, X_t^P) - a^2(t, x_0))\mathrm{d}t + (a(t, X_t^P) - a(t, x_0))\mathrm{d}W_t, \quad \tilde{X}_{1,0}^\eta = 0.$$

One could approximate $\tilde{X}_{1,t}$ by $\hat{X}_{1,t}$ solution of:

$$\mathrm{d}\hat{X}_{1,t} = -a(t, x_0)a^{(1)}(t, x_0)(X_t^P - x_0)\mathrm{d}t + a^{(1)}(t, x_0)(X_t^P - x_0)\mathrm{d}W_t, \quad \hat{X}_{1,0}^\eta = 0.$$

A careful analysis of the SDE satisfied by the difference process $(\tilde{X}_{1,t} - \hat{X}_{1,t})_{t\in[0,T]}$ leads without difficulty to the estimate $\sup_{t\in[0,T]} \|\tilde{X}_{1,t} - \hat{X}_{1,t}\|_p = O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}})$ for any $p \geq 1$, what is the magnitude of the error when performing a second order expansion.
It finally seems more straightforward to directly consider the interpolated process (4.2).

4.2 Normal proxy on $\log(S)$ or Log-normal proxy on S

To consider a log-normal proxy, we worked in Chapter 2 with the log-asset $X = \log(S)$ where S is assumed to be a positive process solution of the next SDE:

$$\mathrm{d}S_t = S_t\sigma(t,S_t)\mathrm{d}W_t, \quad S_0 = e^{x_0} > 0, \tag{4.4}$$

and we envisaged a Gaussan proxy process for X. One had to carefully change the volatility function (a in (4.1) is defined by $a(t,x) = \sigma(t,e^x)$) and the payoff function: $\mathbb{E}[g(S_T)] = \mathbb{E}[h(X_T)]$ with $h(x) = g(e^x)$. As a result, we have obtained an expansion formula with corrective terms which are sensitivities w.r.t. the X-variable. We recall the second order formula with the volatility frozen at x_0:

Theorem 4.2.0.1. *(**2nd order log-normal approximation with Greeks w.r.t. to the X-variable**). Assume $(\mathcal{H}^a_{x_0})$ and suppose that h is locally Lipschitz with exponential growth in the following sense: for some constant $C_h \geq 0$,*

$$|h(x)| \leq C_h e^{C_h|x|}, \qquad |\frac{h(y)-h(x)}{y-x}| \leq C_h e^{C_h(|x|+|y|)} \quad (\forall y \neq x).$$

Then

$$\begin{aligned}\mathbb{E}[h(X_T)] =&\mathbb{E}[h(X^P_T)] + [\int_0^T a^2(t,x_0)(\int_t^T a(s,x_0)a^{(1)}(s,x_0)\mathrm{d}s)\mathrm{d}t](\partial^3_{x_0^3} - \frac{3}{2}\partial^2_{x_0^2} + \frac{1}{2}\partial_{x_0})\mathbb{E}[h(X^P_T)]\\ &+ O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}}).\end{aligned}$$

One could prefer an expression in terms of (usual) Greeks w.r.t. the S-variable (this is done for instance in [Gobet 2012b] up to the third order formula). We define the log-normal proxy S^P, proxy model of the process S by:

$$\mathrm{d}S^P_t = S^P_t\sigma(t,S_0)\mathrm{d}W_t. \tag{4.5}$$

One has:

$$\partial_{x_0}\mathbb{E}[h(X^P_T)] = \partial_{x_0}\mathbb{E}[g(e^{X^P_T})] = \partial_{x_0}\mathbb{E}[g(\frac{S^P_T}{S_0}e^{x_0})] = e^{x_0}\partial_{S_0}\mathbb{E}[g(\frac{S^P_T}{S_0}S_0)] = S_0\partial_{S_0}\mathbb{E}[g(S^P_T)].$$

Using the Faà di Bruno's formula (see [Faà di Bruno 1857]), one obtains more generally for any $n \geq 1$:

$$\partial^n_{x_0^n}\mathbb{E}[h(X^P_T)] = \sum_{\substack{k=(k_1,\cdots,k_n)\in\mathbb{N}^n\\ \sum_{j=1}^n jk_j=n}} \frac{n!}{\prod_{j=1}^n k_j!(j!)^{k_j}} S_0^{\sum_{j=1}^n k_j}\partial^{\sum_{j=1}^n k_j}_{S_0^{\sum_{j=1}^n k_j}}\mathbb{E}[g(S^P_T)]. \tag{4.6}$$

Applying this formula to the expansion proposed in Theorem 4.2.0.1, we obtain:

Theorem 4.2.0.2. *(**2nd order log-normal approximation with Greeks w.r.t. to the S-variable**). Assume $(\mathcal{H}^a_{x_0})$ and suppose that $h = g \circ \exp$ is locally Lipschitz with exponential growth (in the sense described in Theorem 4.2.0.1). Then:*

$$\begin{aligned}\mathbb{E}[g(S_T)] =&\mathbb{E}[g(S^P_T)] + [\int_0^T \sigma^2(t,S_0)(\int_t^T \sigma(s,S_0)S_0\sigma^{(1)}(s,S_0)\mathrm{d}s)\mathrm{d}t](S_0^3\partial^3_{S_0^3} + \frac{3}{2}S_0^2\partial^2_{S_0^2})\mathbb{E}[g(X^P_T)]\\ &+ O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}}).\end{aligned}$$

We would like to find a parameterization directly on the process S allowing to retrieve the results of the above Theorem. In the spirit of the interpolated process (4.2), we propose an interpolation based on the geometrical mean between S_t and S_0 in the function σ and we define for any $\eta \in [0,1]$:

$$dS^{\eta}_t = S^{\eta}_t \sigma(t, (S^{\eta}_t)^{\eta} S_0^{1-\eta}) dW_t, \quad S^{\eta}_0 = e^{x_0} > 0, \tag{4.7}$$

so that $S^1_t = S_t$ and $S^0_t = S^P_t$. It's not hard to verify that $(e^{\hat{X}^{\eta}_t})_{t\in[0,T]}$ (see (4.2)) follows the dynamic (4.7) and hence that $S^{\eta}_t = e^{\hat{X}^{\eta}_t}$ for any $t \in [0,T]$ and any $\eta \in [0,1]$. If we work with $h(\hat{X}^{\eta}_T)$, we retrieve the results of the Theorem 4.2.0.1 (see Section 4.1). To directly obtain the corrective terms of the expansion proposed in Theorem 4.2.0.2, we focus on S^{η}. Assume:

- $(\mathcal{H}^{\sigma}_{S_0})$: σ is a bounded measurable function of $(t,S) \in [0,T] \times \mathbb{R}^+$, and five times continuously differentiable in S with bounded derivatives. Set

$$\mathcal{M}_1(\sigma) = \max_{1\le i\le 5} \sup_{(t,S)\in[0,T]\times\mathbb{R}^+} |\partial^i_{S^i}\sigma(t,S)| \text{ and } \mathcal{M}_0(\sigma) = \max_{0\le i\le 5} \sup_{(t,S)\in[0,T]\times\mathbb{R}^+} |\partial^i_{S^i}\sigma(t,S)|.$$

  In addition we assume that $\int_0^T \sigma^2(t,S_0)dt > 0$.

A straightforward differentiation of (4.7) w.r.t. η gives using the notation $S^{\eta}_{1,t} = \partial_\eta S^{\eta}_t$:

$$dS^{\eta}_{1,t} = \{S^{\eta}_{1,t}\sigma(t,(S^{\eta}_t)^{\eta}S_0^{1-\eta}) + S^{\eta}_t\sigma^{(1)}(t,(S^{\eta}_t)^{\eta}S_0^{1-\eta})(S^{\eta}_t)^{\eta}S_0^{1-\eta}[(\hat{X}^{\eta}_t - x_0) + \eta\hat{X}^{\eta}_{1,t}]\}dW_t, \quad S^{\eta}_{1,0} = 0,$$

and thus we get for $\eta = 0$, noting $S_{1,t} = (\partial_\eta S^{\eta}_t)|_{\eta=0}$:

$$dS_{1,t} = \{S_{1,t}\sigma(t,S_0) + S^P_t S_0 \sigma^{(1)}(t,S_0)(X^P_T - x_0)\}dW_t, \quad S_{1,0} = 0.$$

This linear equation can be solved in closed-form (see [Protter 2004, Theorem 52]):

$$S_{1,t} = S^P_t \int_0^t S_0\sigma^{(1)}(s,S_0)(X^P_s - x_0)(dW_s - \sigma(s,S_0)ds).$$

For the formal following calculus, we assume that g is a smooth function with compact support (on can relax this assumption and apply a regularization argument, we skip details). One has to compute the corrective term $\mathbb{E}[g^{(1)}(S^P_T)S_{1,T}]$. We have:

- using the Malliavin duality relationship,
- identifying the Itô integrals with the Skorokhod integrals,
- using that for any $t \le T$, and any $n,m \ge 1$,

$$\begin{aligned} D_t(g^{(n)}(S^P_T)(S^P_T)^m) =& g^{(n+1)}(S^P_T)(S^P_T)^m D_t(S^P_T) + m g^{(n)}(S^P_T)(S^P_T)^{m-1}D_t(S^P_T) \\ =& \{g^{(n+1)}(S^P_T)(S^P_T)^{m+1} + m g^{(n)}(S^P_T)(S^P_T)^m\}\sigma(t,S_0), \end{aligned}$$

that:

$$\begin{aligned} \mathbb{E}[g^{(1)}(S^P_T)S_{1,T}] =& \mathbb{E}[g^{(1)}(S^P_T)S^P_T \int_0^T S_0\sigma^{(1)}(t,S_0)(X^P_t - x_0)dW_t] \\ &- \mathbb{E}[g^{(1)}(S^P_T)S^P_T \int_0^T \sigma(s,S_0)S_0\sigma^{(1)}(t,S_0)(X^P_t - x_0)dt] \\ =& \mathbb{E}[(g^{(2)}(S^P_T)(S^P_T)^2 + g^{(1)}(S^P_T)S^P_T) \int_0^T \sigma(t,S_0)S_0\sigma^{(1)}(t,S_0)(X^P_t - x_0)dt] \end{aligned}$$

$$
\begin{aligned}
&-\mathbb{E}[g^{(1)}(S_T^P)S_T^P\int_0^T\sigma(s,S_0)S_0\sigma^{(1)}(t,S_0)(X_t^P-x_0)\mathrm{d}t]\\
=&\mathbb{E}[g^{(2)}(S_T^P)(S_T^P)^2\int_0^T\sigma(t,S_0)S_0\sigma^{(1)}(t,S_0)(X_t^P-x_0)\mathrm{d}t]
\end{aligned}
$$

Then we write using the Itô's formula for the product $(X_t^P-x_0)\int_t^T\sigma(s,S_0)S_0\sigma^{(1)}(s,S_0)\mathrm{d}s$:

$$
\begin{aligned}
\int_0^T\sigma(t,S_0)S_0\sigma^{(1)}(t,S_0)(X_t^P-x_0)\mathrm{d}t &= \int_0^T(\int_t^T\sigma(s,S_0)S_0\sigma^{(1)}(s,S_0)\mathrm{d}s)\mathrm{d}(X_t^P-x_0)\\
&=\int_0^T(\int_t^T\sigma(s,S_0)S_0\sigma^{(1)}(s,S_0)\mathrm{d}s)(-\frac{1}{2}\sigma^2(t,S_0)\mathrm{d}t+\sigma(t,S_0)\mathrm{d}W_t).
\end{aligned}
$$

Thus we obtain applying again the same tools:

$$
\begin{aligned}
&\mathbb{E}[g^{(2)}(S_T^P)(S_T^P)^2\int_0^T\sigma(t,S_0)S_0\sigma^{(1)}(t,S_0)(X_t^P-x_0)\mathrm{d}t]\\
=&-\frac{1}{2}\int_0^T\sigma^2(t,S_0)(\int_t^T\sigma(s,S_0)S_0\sigma^{(1)}(s,S_0)\mathrm{d}s)\mathrm{d}t\mathbb{E}[g^{(2)}(S_T^P)(S_T^P)^2]\\
&+\mathbb{E}[g^{(2)}(S_T^P)(S_T^P)^2\int_0^T(\int_t^T\sigma(s,S_0)S_0\sigma^{(1)}(s,S_0)\mathrm{d}s)\sigma(t,S_0)\mathrm{d}W_t]\\
=&-\frac{1}{2}\int_0^T\sigma^2(t,S_0)(\int_t^T\sigma(s,S_0)S_0\sigma^{(1)}(s,S_0)\mathrm{d}s)\mathrm{d}t\mathbb{E}[g^{(2)}(S_T^P)(S_T^P)^2]\\
&+\mathbb{E}[\{g^{(3)}(S_T^P)(S_T^P)^3+2g^{(2)}(S_T^P)(S_T^P)^2\}\int_0^T(\int_t^T\sigma(s,S_0)S_0\sigma^{(1)}(s,S_0)\mathrm{d}s)\sigma^2(t,S_0)\mathrm{d}t]\\
=&\int_0^T\sigma^2(t,S_0)(\int_t^T\sigma(s,S_0)S_0\sigma^{(1)}(s,S_0)\mathrm{d}s)\mathrm{d}t(\mathbb{E}[g^{(3)}(S_T^P)(S_T^P)^3]+\frac{3}{2}\mathbb{E}[g^{(2)}(S_T^P)(S_T^P)^2]).
\end{aligned}
$$

The final trick lies in the observation that

$$
\partial^n_{S_0^n}\mathbb{E}[g(S_T^P)]=\partial^{n-1}_{S_0^{n-1}}\mathbb{E}[g^{(1)}(S_T^P)\frac{S_T^P}{S_0}]=\cdots=\mathbb{E}[g^{(n)}(S_T^P)(\frac{S_T^P}{S_0})^n],
$$

for any $n\geq 1$. We exactly retrieve the corrective terms of Theorem 4.2.0.2.
We nevertheless mention that the explicit derivation of the expansion coefficients is a little bit more tricky. First, the corrective processes $(\partial^n_{\eta^n}S^\eta)|_{\eta=0}$ are not anymore a sequence of iterated Wiener integrals but a sequence of nested processes solutions of linear SDEs. If an explicit form remains available, the manipulation of such processes seems a little less tractable. Second one has to carefully apply the Malliavin calculus on log-normal processes. Finally the identification of the Greeks is slightly less straightforward and relies on knowledge of martingales properties (for any $n\in\mathbb{N}$, $((S_t)^n\partial^n_{S^n}v(t,S_t^P))_{t\in[0,T[}$ is a martingale where $v(t,S)=\mathbb{E}[g(S_T^P)|S_t^P=S]$) or flow properties.
If we wanted to perform a rigorous error analysis in terms of $\mathcal{M}_0(\sigma),\mathcal{M}_1(\sigma)$, T and the growth assumptions on the function g (a good hypothesis could be to assume that g has polynomial growth), it would be necessary to estimate the L^p norms of $S_t-S_t^P-S_{1,t}^P=\int_0^1(1-\eta)\partial^2_{\eta^2}S_t^\eta\mathrm{d}\eta$ and the Malliavin derivatives of $S_t-S_t^P=\int_0^1S_{1,t}^\eta\mathrm{d}\eta$. Although the methodology is the same as that previously employed with the Gaussian proxys, the calculus seem a little bit more tedious similarly to the corrective terms derivation step.
Finally we should mention that the parameterization (4.7) is not very natural without the knowledge of the interpolated process (4.2) for the log-asset. This parameterization (4.7) fortunately allows an explicit calculus of the expansion terms (this is its *raison d'être*) but this is not obvious to guess it directly.

As a conclusion, when considering a proxy process which is a function of a Gaussian process (as the log-normal process), it seems more convenient to apply a transformation on the initial process and on the payoff function to go back to Gaussian proxys.

4.3 Towards a displaced log-normal proxy

So far we only have considered Gaussian and log-normal proxys. Such proxys are obtained with a zero order approximation of the diffusion coefficients. In order to reduce the number of corrective terms in the expansions, we could envisage higher order proxys. In the context of time-homogeneous[1] local volatility, a natural surrogate is the displaced log-normal proxy obtained with a first order expansion of the diffusion coefficients. We consider the solution of the following SDE:

$$\mathrm{d}S_t = \Sigma(S_t)\mathrm{d}W_t, \quad S_0 > 0, \tag{4.8}$$

where we assume that Σ satisfies:

- $(\tilde{\mathcal{H}}^{\Sigma})$: Σ is a bounded measurable function of $S \in \mathbb{R}^+$, and five times continuously differentiable with bounded derivatives. Set

  $$\mathcal{M}_1(\Sigma) = \max_{1\leq i\leq 5} \sup_{S\in\mathbb{R}^+} |\Sigma^{(i)}(S)| \text{ and } \mathcal{M}_0(\Sigma) = \max_{0\leq i\leq 5} \sup_{S\in\mathbb{R}^+} |\Sigma^{(i)}(S)|.$$

  In addition, there exists a constant $c_\Sigma > 0$ such that $\min(\Sigma^{(1)}(S_0), \inf_{S\in\mathbb{R}^+} \Sigma(S)) > c_\Sigma$.

Then the corresponding displaced log-normal proxy is defined by:

$$\mathrm{d}S^P_t = [\Sigma(S_0) + \Sigma^{(1)}(S_0)(S^P_t - S_0)]\mathrm{d}W_t, \quad S^P_0 = S_0 \quad (\text{Displaced Log-Normal Proxy}). \tag{4.9}$$

Remark 4.3.0.1. *We assume that $\Sigma^{(1)}(S_0) \neq 0$, otherwise the proxy (4.9) is equivalent to the Gaussian proxy. We suppose w.l.g. that $\Sigma^{(1)}(S_0) > 0$, the case $\Sigma^{(1)}(S_0) < 0$ being handled similarly.*

Using standard inequalities and Gronwall's Lemma, one easily obtains the estimate $\sup_{t\in[0,T]} \|S_t - S^P_t\|_p = O(\mathcal{M}_1(\sigma)[\mathcal{M}_0(\sigma)]^2 T^{\frac{3}{2}})$, for any $p \geq 1$. This error magnitude means that when considering a payoff function g locally Lipschitz with polynomial growth, the approximation $\mathbb{E}[g(S_T)] \approx \mathbb{E}[g(S^P_t)]$ is already of order 2 w.r.t. the model parameters. A closed-form formula is available for $\mathbb{E}[g(S^P_t)]$ because $(S^P_t)_{t\in[0,T]}$ is connected to a log-normal process $(Y^P_t)_{t\in[0,T]}$ through the relations:

$$Y^P_t = \Sigma(S_0) + \Sigma^{(1)}(S_0)(S^P_t - S_0), \tag{4.10}$$

$$\mathrm{d}Y^P_t = \Sigma^{(1)}(S_0)\mathrm{d}S^P_t = \Sigma^{(1)}(S_0)Y^P_t\mathrm{d}W_t, \quad Y^P_0 = \Sigma(S_0). \tag{4.11}$$

As discussed in the Section 4.2, to find a interpolated process allowing to connect S and S^P and to compute corrective terms in order to improve the approximation $\mathbb{E}[g(S_T)] \approx \mathbb{E}[g(S^P_t)]$ seems not to be an easy task at first glance.

▷**A direct approach.** A first attempt could be to consider the next parameterization:

$$\mathrm{d}S^\eta_t = \{\eta\Sigma(S^\eta_t) + (1-\eta)[\Sigma(S_0) + \Sigma^{(1)}(S_0)(S^\eta_t - S_0)]\}\mathrm{d}W_t, \quad S^\eta_0 = S_0.$$

With the notation $(\partial_\eta S^\eta_t)|_{\eta=0} = S_{1,t}$, it comes:

$$\mathrm{d}S_{1,t} = \{\Sigma^{(1)}(S_0)S_{1,t} + \Sigma(S^P_t) - [\Sigma(S_0) + \Sigma^{(1)}(S_0)(S^P_t - S_0)]\}\mathrm{d}W_t, \quad S_{1,0} = 0.$$

[1] We consider time-independent parameters to make explicit the calculus using the displaced log-normal proxy.

The solution of this linear SDE does not lead to illuminating computations. In view of the discussion at the end of the Section 4.1, we consider:

$$d\tilde{S}_{1,t} = \{\Sigma^{(1)}(S_0)\tilde{S}_{1,t} + \frac{\Sigma^{(2)}(S_0)}{2}(S^P_t - S_0)^2\}dW_t, \quad \tilde{S}_{1,0} = 0,$$

and once again classical inequalities combined with the Gronwall's Lemma yield to the estimate $\sup_{t\in[0,T]} \|S_{1,t} - \tilde{S}_{1,t}\|_p = O(\mathcal{M}_1(\sigma)[\mathcal{M}_0(\sigma)]^3 T^2)$. Then we write in closed-form $\tilde{S}_{1,t}$ using (4.11) and (4.10):

$$\begin{aligned}\tilde{S}_{1,t} =& Y^P_t \int_0^t (Y^P_s)^{-1} \frac{\Sigma^{(2)}(S_0)}{2}(S^P_s - S_0)^2(dW_s - \Sigma^{(1)}(S_0)ds)\\ =& Y^P_t \int_0^t (Y^P_s)^{-1} \frac{\Sigma^{(2)}(S_0)}{2(\Sigma^{(1)}(S_0))^2}(Y^P_s - \Sigma(S_0))^2(dW_s - \Sigma^{(1)}(S_0)ds).\end{aligned}$$

For the formal following calculus, we assume that the function g is smooth with compact support. Using the fact that and $D_t Y^P_T = \Sigma^{(1)}(S_0)Y^P_T$ and $D_t S^P_T = \frac{D_t Y^P_T}{\Sigma^{(1)}(S_0)} = Y^P_T$ for any $t \le T$ (see (4.10)) and the Malliavin tools developed in the Section 4.2, we obtain without difficulty:

$$\mathbb{E}[g^{(1)}(S^P_T)\tilde{S}_{1,T}] = \frac{\Sigma^{(2)}(S_0)}{2(\Sigma^{(1)}(S_0))^2}\mathbb{E}[g^{(2)}(S^P_T)(Y^P_T)^2 \int_0^T (Y^P_s)^{-1}(Y^P_s - \Sigma(S_0))^2 ds].$$

But how to continue the computation is not clear because Y satisfies a linear equation as a log-normal process and consequently the application of the same arguments would be endless. An idea could be to consider a Gaussian approximation of $(Y^P_s)^{-1}(Y^P_s - \Sigma(S_0))^2$.

▷**Using a transformation on the model.** As an alternative, as explained in Section 4.2, we could apply a transformation to the payoff function and to the initial process.
The heuristic is the following: if S is closed to the displaced log-normal process S^P, the process $Y = \Sigma(S_0) + \Sigma^{(1)}(S_0)(S_t - S_0)$ behaves closely to the log-normal process Y^P. Y is solution of the next SDE:

$$dY_t = \Sigma^{(1)}(S_0)dS_t = \Sigma^{(1)}(S_0)\Sigma(S_t)dW_t = \Sigma^{(1)}(S_0)\Sigma(S_0 + \frac{Y_t - \Sigma(S_0)}{\Sigma^{(1)}(S_0)})dW_t, \quad Y_0 = \Sigma(S_0). \tag{4.12}$$

Then assuming that Y is strictly positive, we define the process $(X_t = \log(Y_t))_{t\in[0,T]}$ solution of:

$$dX_t = a(X_t)(dW_t - \frac{1}{2}a(X_t)dt), \quad x_0 = \log(\Sigma(S_0)). \tag{4.13}$$

where $a(x) = e^{-x}\Sigma^{(1)}(S_0)\Sigma(S_0 + \frac{e^x - \Sigma(S_0)}{\Sigma^{(1)}(S_0)})$. Now assume $(\mathcal{H}^a)$. Straightforward calculus show that:

$$a(x_0) = \Sigma^{(1)}(S_0), \qquad a^{(1)}(x_0) = 0, \qquad a^{(2)}(x_0) = \frac{\Sigma(S_0)\Sigma^{(2)}(S_0)}{\Sigma^{(1)}(S_0)}.$$

Thus if we want to price a Call option written on the asset S with strike K, we write using the strict positivity of $\Sigma^{(1)}(S_0)$:

$$\mathbb{E}[(S_T - K)_+] = \frac{1}{\Sigma^{(1)}(S_0)}\mathbb{E}[(e^{X_T} - [\Sigma(S_0) + \Sigma^{(1)}(S_0)(K - S_0)])_+]$$

Next we suppose that the new strike $K_d = [\Sigma^{(1)}(S_0)(K - S_0) + \Sigma(S_0)] > 0$ (this is the case for OTM and ATM options) otherwise the price is equal to $S_0 - K$ and we apply the Theorem 2.3.1.1 of Chapter 2 Section 2.3 to directly obtain the following result:

Theorem 4.3.0.3. *(2nd and 3rd order displaced log-normal approximations for Call options). Assume* $(\tilde{\mathcal{H}}^{\Sigma})$ *and* $(\mathcal{H}^{a})$ *where* $a(x) = e^{-x}\Sigma^{(1)}(S_0)\Sigma(S_0 + \frac{e^x - \Sigma(S_0)}{\Sigma^{(1)}(S_0)})$. *Suppose that* $K_d = [\Sigma^{(1)}(S_0)(K - S_0) + \Sigma(S_0)] > 0$ *and let* $x_d = x_0 = \log(\Sigma(S_0))$, $a_d = a(x_d) = \Sigma^{(1)}(S_0)$, *and* $k_d = \log(K_d)$. *Then we have using the displaced log-normal proxy:*

$$
\begin{aligned}
\mathrm{Call}(S_0, T, K) =& \frac{1}{\Sigma^{(1)}(S_0)}\mathrm{Call}^{\mathrm{BS}}(x_d, a_d^2 T, k_d) + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^2 T^{\frac{3}{2}}), && (4.14)\\
=& \frac{1}{\Sigma^{(1)}(S_0)}\mathrm{Call}^{\mathrm{BS}}(x_d, a_d^2 T, k_d) + \frac{C_2(a; x_d)_0^T}{\Sigma^{(1)}(S_0)}(-\frac{1}{2}\partial_x + \frac{1}{2}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x_d, a_d^2 T, k_d) && (4.15)\\
&+ \frac{C_3(a; x_d)_0^T}{\Sigma^{(1)}(S_0)}(-\frac{1}{4}\partial_x + \frac{5}{4}\partial^2_{x^2} - 2\partial^3_{x^3} + \partial^4_{x^4})\mathrm{Call}^{\mathrm{BS}}(x_d, a_d^2 T, k_d) + O(\mathcal{M}_1(a)[\mathcal{M}_0(a)]^3 T^2),
\end{aligned}
$$

where:

$$
\begin{aligned}
C_2(a; x_d)_0^T =& \frac{T^2}{2}a^3(x_d)a^{(2)}(x_d) = \frac{T^2}{2}\Sigma(S_0)(\Sigma^{(1)}(S_0))^2\Sigma^{(2)}(S_0),\\
C_3(a; x_d)_0^T =& \frac{T^3}{6}a^5(x_d)a^{(2)}(x_d) = \frac{T^2}{2}\Sigma(S_0)(\Sigma^{(1)}(S_0))^4\Sigma^{(2)}(S_0).
\end{aligned}
$$

We notice that the above expansions contain less terms (owing to the fact that $a^{(1)}(x_0) = a^{(1)}(x_d) = 0$) than the corresponding order 2 and order 3 approximations using the log-normal proxy. The second order formula (4.14) reduces to only a suitably rescaled Black-Scholes price $\frac{1}{\Sigma^{(1)}(S_0)}\mathrm{Call}^{\mathrm{BS}}(x_d, a_d^2 T, k_d)$. The simplicity of the formulas is a significant advantage.
In addition, if $\Sigma^{(2)}$ is null (i.e. if Σ is affine and thus $\Sigma^{(1)}$ is constant), the corrective terms are equal to zero what is coherent because the initial model coincides with the displaced log-normal proxy model. Moreover the error term is obviously equal to zero because in this case, for any $x \in \mathbb{R}$, we have:

$$
a'(x) = -\Sigma^{(1)}(S_0)e^{-x}\Sigma(S_0 + \frac{e^x - \Sigma(S_0)}{\Sigma^{(1)}(S_0)}) + \Sigma^{(1)}(S_0 + \frac{e^x - \Sigma(S_0)}{\Sigma^{(1)}(S_0)}) = -\Sigma^{(1)}(S_0) + \Sigma^{(1)}(S_0) = 0,
$$

and hence $\mathcal{M}_1(a) = 0$. Numerical tests of these approximations and comparisons with the normal and log-normal proxys are postponed to Subsection 4.4.2.

4.4 Numerical experiments

4.4.1 Comparison of the implied volatility behaviors in the displaced log-normal and CEV models

In this Subsection, we compare the behaviors of the implied volatility in the displaced log-normal and the CEV models and the accuracy of the approximation formulas developed in Chapter 2. We consider the processes S^d (displaced log-normal) and S^c (CEV) solution of:

$$
\begin{aligned}
\mathrm{d}S^d_t =& \nu(1 - \beta + \beta S^d_t)\mathrm{d}W_t, \quad S^d_0 = 1,\\
\mathrm{d}S^c_t =& \nu(S^c_t)^{\beta}, \quad S^d_0 = 1,
\end{aligned}
$$

for $\beta \in [0, 1]$ and $\nu > 0$. With the notation $\Sigma_d(S) = \nu(1 - \beta + \beta S)$, $a_d(x) = \nu(1 - \beta + \beta e^x)e^{-x}$, $\Sigma_c(S) = \nu S^{\beta}$ and $a_c(x) = \nu e^{(\beta-1)x}$, we readily have:

$$
\begin{aligned}
&\Sigma_d(1) = \nu, \quad \Sigma_d^{(1)}(1) = \nu\beta, \quad \Sigma_d^{(2)}(1) = 0, \quad && a_d(0) = \nu, \quad a_d^{(1)}(0) = \nu(\beta - 1), \quad a_d^{(2)}(0) = \nu(1 - \beta)\\
&\Sigma_c(1) = \nu, \quad \Sigma_c^{(1)}(1) = \nu\beta, \quad \Sigma_c^{(2)}(1) = \nu\beta(\beta - 1), \quad && a_c(0) = \nu, \quad a_c^{(1)}(0) = \nu(\beta - 1), \quad a_c^{(2)}(0) = \nu(\beta - 1)^2.
\end{aligned}
$$

We deduce from these simple calculus that the two models induce a similar short term implied volatility ATM because the corresponding local volatility functions are equal at spot (directly connected to the level of the implied volatility ATM for short maturity) and have the same slope at spot (approximately equal to the double of the slope of the implied volatility ATM for short maturity). Besides the order 2 approximations (see Theorem 2.1.2.3 of Chapter 2 Section 2.1) give the same price approximations for the two models. However the curvature of the local volatility functions at spot are different and we can consequently expect a different behaviour far ITM and far OTM and/or for long maturities.
For the numerical experiments, we set $\nu = 0.25$ and we consider either $\beta = 0.8$ (a priori close to the log-normal case) or $\beta = 0.2$ (a priori close to the normal case) and compare the true values of the implied volatility in the two models as well their approximations with the order 2 and 3 price approximations at spot (and then evaluating the corresponding implied volatilities) using the normal and log-normal proxys (see Theorems 2.1.2.3 and 2.3.1.1 of Chapter 2 Section 2.1) for various maturities and strikes.
We use the following maturities T: 6 months, 1 year, 2 years and 5 years and the strikes evolve approximately as $S_0 \exp(c\nu\sqrt{T})$ where c takes the value of various quantiles of the standard Gaussian law from 1% to 99% in order to cover far ITM and far OTM options.
We report in Tables 4.1-4.2-4.3-4.4 for the value $\beta = 0.8$ (one table for each maturity):

- The true values of the implied volatility in the displaced log-normal model denoted by `True LND`, the implied volatility approximations obtained with the second order price approximation at spot with the normal and log-normal proxys using the notations $\mathtt{AppN}_2$ and $\mathtt{AppLN}_2$ and with the third order price approximation at spot with the normal and log-normal proxys denoted by $\mathtt{AppN}_3^{\mathtt{DLN}}$ and $\mathtt{AppLN}_3^{\mathtt{DLN}}$,

- The true values of the implied volatility in the CEV model denoted by `True CEV`, the implied volatility approximations obtained with the second order price approximation at spot with the normal and log-normal proxys using the notations $\mathtt{AppN}_2$ and $\mathtt{AppLN}_2$ (similar to the displaced log-normal model) and with the third order price approximation at spot with the normal and log-normal proxys denoted by $\mathtt{AppN}_3^{\mathtt{CEV}}$ and $\mathtt{AppLN}_3^{\mathtt{CEV}}$,

We do the same work in Tables 4.5-4.6-4.7-4.8 for the value $\beta = 0.2$. When the price approximation does not belong to the non-arbitrage interval $](S_0 - K)_+, S_0[$, we just report ND in the tabular.

Table 4.1: True Implied Black-Scholes volatilities (%) for the displaced log-normal and CEV models and order 2 and 3 approximations using normal and log-normal proxys for $\nu = 0.25, \beta = 0.8$ and $T = 6M$.

Strikes	0.65	0.75	0.80	0.85	0.90	0.95	1	1.05	1.10	1.20	1.25	1.35	1.50
`True DLN`	26.24	25.80	25.61	25.44	25.28	25.14	25.01	24.89	24.78	24.58	24.49	24.32	24.11
$\mathtt{AppN}_2$	23.41	25.54	25.56	25.46	25.31	25.17	25.03	24.91	24.81	24.56	24.39	23.91	22.83
$\mathtt{AppLN}_2$	25.96	25.69	25.55	25.40	25.26	25.13	25.00	24.88	24.76	24.54	24.43	24.21	23.85
$\mathtt{AppN}_3^{\mathtt{DLN}}$	26.17	25.76	25.60	25.44	25.28	25.14	25.01	24.89	24.78	24.59	24.50	24.35	23.93
$\mathtt{AppLN}_3^{\mathtt{DLN}}$	26.19	25.79	25.61	25.44	25.28	25.14	25.01	24.89	24.78	24.58	24.49	24.34	24.16
`True CEV`	26.09	25.73	25.56	25.41	25.27	25.13	25.00	24.88	24.76	24.55	24.45	24.26	24.00
$\mathtt{AppN}_2$	23.41	25.54	25.56	25.46	25.31	25.17	25.03	24.91	24.81	24.56	24.39	23.91	22.83
$\mathtt{AppLN}_2$	25.96	25.69	25.55	25.40	25.26	25.13	25.00	24.88	24.76	24.54	24.43	24.21	23.85
$\mathtt{AppN}_3^{\mathtt{CEV}}$	25.92	25.69	25.55	25.41	25.27	25.13	25.00	24.88	24.76	24.56	24.47	24.30	23.88
$\mathtt{AppLN}_3^{\mathtt{CEV}}$	26.09	25.73	25.56	25.41	25.27	25.13	25.00	24.88	24.76	24.55	24.45	24.26	24.01

Table 4.2: True Implied Black-Scholes volatilities (%) for the displaced log-normal and CEV models and order 2 and 3 approximations using normal and log-normal proxys for $\nu = 0.25$, $\beta = 0.8$ and $T = 1Y$.

Strikes	0.55	0.65	0.75	0.80	0.90	0.95	1	1.05	1.15	1.25	1.40	1.50	1.80
`True DLN`	26.83	26.26	25.81	25.62	25.30	25.15	25.02	24.90	24.69	24.50	24.26	24.12	23.78
$\texttt{AppN}_2$	18.04	25.62	25.79	25.67	25.35	25.20	25.07	24.95	24.74	24.51	23.96	23.41	21.21
$\texttt{AppLN}_2$	26.29	26.02	25.71	25.55	25.26	25.13	25.00	24.88	24.65	24.44	24.13	23.93	23.19
$\texttt{AppN}_3^{\texttt{DLN}}$	26.47	26.14	25.79	25.61	25.30	25.15	25.02	24.90	24.69	24.51	24.31	24.18	23.05
$\texttt{AppLN}_3^{\texttt{DLN}}$	26.69	26.22	25.80	25.62	25.30	25.15	25.02	24.90	24.69	24.50	24.27	24.15	23.93
`True CEV`	26.53	26.10	25.73	25.56	25.27	25.13	25.00	24.88	24.65	24.45	24.17	24.00	23.56
$\texttt{AppN}_2$	18.04	25.62	25.79	25.67	25.35	25.20	25.07	24.95	24.74	24.51	23.96	23.41	21.21
$\texttt{AppLN}_2$	26.29	26.02	25.71	25.55	25.26	25.13	25.00	24.88	24.65	24.44	24.13	23.93	23.19
$\texttt{AppN}_3^{\texttt{CEV}}$	25.91	25.95	25.71	25.56	25.27	25.13	25.00	24.88	24.66	24.46	24.24	24.10	23.00
$\texttt{AppLN}_3^{\texttt{CEV}}$	26.51	26.09	25.73	25.56	25.27	25.13	25.00	24.88	24.65	24.45	24.17	24.00	23.59

Table 4.3: True Implied Black-Scholes volatilities (%) for the displaced log-normal and CEV models and order 2 and 3 approximations using normal and log-normal proxys for $\nu = 0.25$, $\beta = 0.8$ and $T = 2Y$.

Strikes	0.45	0.55	0.65	0.75	0.85	0.90	1	1.10	1.20	1.35	1.55	1.80	2.30
`True DLN`	27.63	26.86	26.29	25.84	25.48	25.32	25.05	24.81	24.61	24.35	24.07	23.80	23.39
$\texttt{AppN}_2$	ND	25.88	26.22	25.95	25.59	25.42	25.13	24.91	24.72	24.39	23.65	22.18	18.80
$\texttt{AppLN}_2$	26.69	26.39	26.05	25.71	25.41	25.26	25.00	24.76	24.54	24.24	23.86	23.39	22.07
$\texttt{AppN}_3^{\texttt{DLN}}$	26.59	26.56	26.21	25.82	25.47	25.32	25.05	24.82	24.62	24.38	24.20	23.86	21.25
$\texttt{AppLN}_3^{\texttt{DLN}}$	27.34	26.76	26.25	25.83	25.47	25.32	25.05	24.82	24.62	24.37	24.11	23.89	23.81
`True CEV`	27.06	26.53	26.10	25.73	25.41	25.27	25.01	24.77	24.55	24.26	23.92	23.56	22.98
$\texttt{AppN}_2$	ND	25.88	26.22	25.95	25.59	25.42	25.13	24.91	24.72	24.39	23.65	22.18	18.80
$\texttt{AppLN}_2$	26.69	26.39	26.05	25.71	25.41	25.26	25.00	24.76	24.54	24.24	23.86	23.39	22.07
$\texttt{AppN}_3^{\texttt{CEV}}$	25.43	26.17	26.02	25.72	25.41	25.27	25.00	24.77	24.55	24.29	24.07	23.73	21.21
$\texttt{AppLN}_3^{\texttt{CEV}}$	27.02	26.53	26.10	25.73	25.41	25.27	25.01	24.77	24.55	24.26	23.93	23.57	23.08

Table 4.4: True Implied Black-Scholes volatilities (%) for the displaced log-normal and CEV models and order 2 and 3 approximations using normal and log-normal proxys for $\nu = 0.25, \beta = 0.8$ and $T = 5Y$.

Strikes	0.25	0.40	0.50	0.60	0.75	0.85	1	1.15	1.35	1.60	2.05	2.50	3.60
True DLN	30.59	28.25	27.33	26.66	25.93	25.56	25.12	24.77	24.41	24.07	23.62	23.31	22.83
$\texttt{AppN}_2$	ND	26.97	27.32	26.95	26.23	25.81	25.33	25.01	24.68	24.11	22.10	19.45	14.41
$\texttt{AppLN}_2$	27.58	27.07	26.66	26.25	25.72	25.41	25.00	24.65	24.25	23.81	23.09	22.33	17.23
$\texttt{AppN}_3^{\texttt{DLN}}$	ND	27.38	27.04	26.56	25.90	25.54	25.12	24.78	24.44	24.20	24.09	22.83	16.74
$\texttt{AppLN}_3^{\texttt{DLN}}$	29.12	27.89	27.17	26.58	25.90	25.54	25.12	24.78	24.43	24.11	23.76	23.63	24.13
True CEV	28.64	27.38	26.79	26.31	25.74	25.42	25.01	24.66	24.27	23.85	23.26	22.79	21.94
$\texttt{AppN}_2$	ND	26.97	27.32	26.95	26.23	25.81	25.33	25.01	24.68	24.11	22.10	19.45	14.41
$\texttt{AppLN}_2$	27.58	27.07	26.66	26.25	25.72	25.41	25.00	24.65	24.25	23.81	23.09	22.33	17.23
$\texttt{AppN}_3^{\texttt{CEV}}$	ND	26.35	26.52	26.24	25.72	25.41	25.01	24.67	24.28	23.97	23.83	22.69	16.72
$\texttt{AppLN}_3^{\texttt{CEV}}$	28.46	27.36	26.78	26.31	25.74	25.42	25.01	24.66	24.27	23.86	23.26	22.81	22.33

Table 4.5: True Implied Black-Scholes volatilities (%) for the displaced log-normal and CEV models and order 2 and 3 approximations using normal and log-normal proxys for $\nu = 0.25, \beta = 0.2$ and $T = 6M$.

Strikes	0.65	0.75	0.80	0.85	0.90	0.95	1	1.05	1.10	1.20	1.25	1.35	1.50
True DLN	29.73	28.09	27.37	26.72	26.11	25.55	25.03	24.55	24.09	23.27	22.89	22.20	21.29
$\texttt{AppN}_2$	29.66	28.08	27.37	26.72	26.11	25.55	25.03	24.55	24.09	23.27	22.88	22.16	21.15
$\texttt{AppLN}_2$	28.07	27.54	27.10	26.59	26.05	25.51	25.00	24.51	24.04	23.10	22.57	21.15	ND
$\texttt{AppN}_3^{\texttt{DLN}}$	29.73	28.09	27.37	26.71	26.11	25.55	25.03	24.55	24.09	23.27	22.89	22.20	21.28
$\texttt{AppLN}_3^{\texttt{DLN}}$	29.34	28.03	27.35	26.71	26.11	25.55	25.03	24.55	24.10	23.28	22.91	22.31	22.37
True CEV	29.59	28.02	27.32	26.69	26.09	25.54	25.02	24.53	24.08	23.24	22.85	22.13	21.18
$\texttt{AppN}_2$	29.66	28.08	27.37	26.72	26.11	25.55	25.03	24.55	24.09	23.27	22.88	22.16	21.15
$\texttt{AppLN}_2$	28.07	27.54	27.10	26.59	26.05	25.51	25.00	24.51	24.04	23.10	22.57	21.15	ND
$\texttt{AppN}_3^{\texttt{CEV}}$	29.60	28.02	27.33	26.69	26.09	25.54	25.02	24.53	24.08	23.24	22.85	22.13	21.18
$\texttt{AppLN}_3^{\texttt{CEV}}$	29.28	27.98	27.31	26.68	26.09	25.54	25.02	24.54	24.08	23.24	22.86	22.21	22.13

Table 4.6: True Implied Black-Scholes volatilities (%) for the displaced log-normal and CEV models and order 2 and 3 approximations using normal and log-normal proxys for $\nu = 0.25, \beta = 0.2$ and $T = 1Y$.

Strikes	0.55	0.65	0.75	0.80	0.90	0.95	1	1.05	1.15	1.25	1.40	1.50	1.80
True DLN	31.83	29.79	28.13	27.41	26.15	25.59	25.06	24.58	23.69	22.92	21.90	21.31	19.83
$\texttt{AppN}_2$	31.72	29.76	28.13	27.42	26.15	25.59	25.07	24.58	23.70	22.92	21.88	21.24	19.49
$\texttt{AppLN}_2$	28.96	28.57	27.70	27.16	26.05	25.51	25.00	24.51	23.59	22.68	21.04	19.15	ND
$\texttt{AppN}_3^{\texttt{DLN}}$	31.82	29.79	28.13	27.41	26.15	25.59	25.06	24.58	23.69	22.92	21.90	21.31	19.70
$\texttt{AppLN}_3^{\texttt{DLN}}$	31.00	29.61	28.09	27.39	26.14	25.58	25.06	24.58	23.70	22.94	21.99	21.58	23.43
True CEV	31.54	29.62	28.05	27.35	26.12	25.56	25.04	24.55	23.66	22.87	21.81	21.19	19.60
$\texttt{AppN}_2$	31.72	29.76	28.13	27.42	26.15	25.59	25.07	24.58	23.70	22.92	21.88	21.24	19.49
$\texttt{AppLN}_2$	28.96	28.57	27.70	27.16	26.05	25.51	25.00	24.51	23.59	22.68	21.04	19.15	ND
$\texttt{AppN}_3^{\texttt{CEV}}$	31.57	29.65	28.06	27.36	26.12	25.56	25.04	24.55	23.66	22.86	21.80	21.18	19.62
$\texttt{AppLN}_3^{\texttt{CEV}}$	30.90	29.51	28.02	27.34	26.11	25.56	25.04	24.56	23.67	22.88	21.86	21.37	23.04

Table 4.7: True Implied Black-Scholes volatilities (%) for the displaced log-normal and CEV models and order 2 and 3 approximations using normal and log-normal proxys for $\nu = 0.25, \beta = 0.2$ and $T = 2Y$.

Strikes	0.45	0.55	0.65	0.75	0.85	0.90	1	1.10	1.20	1.35	1.55	1.80	2.30
True DLN	34.59	31.96	29.89	28.22	26.83	26.22	25.13	24.18	23.34	22.26	21.07	19.86	18.06
AppN_2	34.48	31.93	29.89	28.23	26.84	26.23	25.13	24.18	23.35	22.27	21.03	19.68	17.30
AppLN_2	30.08	29.74	28.90	27.79	26.62	26.05	25.00	24.05	23.16	21.85	19.67	ND	ND
$\mathrm{AppN}_3^{\mathrm{DLN}}$	34.59	31.95	29.89	28.22	26.83	26.22	25.13	24.18	23.34	22.27	21.07	19.86	17.87
$\mathrm{AppLN}_3^{\mathrm{DLN}}$	33.13	31.53	29.75	28.16	26.80	26.20	25.12	24.19	23.37	22.32	21.24	20.90	25.77
True CEV	34.03	31.62	29.69	28.10	26.76	26.16	25.08	24.13	23.29	22.18	20.92	19.62	17.62
AppN_2	34.48	31.93	29.89	28.23	26.84	26.23	25.13	24.18	23.35	22.27	21.03	19.68	17.30
AppLN_2	30.08	29.74	28.90	27.79	26.62	26.05	25.00	24.05	23.16	21.85	19.67	ND	ND
$\mathrm{AppN}_3^{\mathrm{CEV}}$	34.16	31.69	29.73	28.12	26.77	26.17	25.08	24.13	23.28	22.16	20.90	19.61	17.65
$\mathrm{AppLN}_3^{\mathrm{CEV}}$	32.98	31.36	29.62	28.07	26.74	26.15	25.08	24.14	23.30	22.21	21.01	20.39	25.32

Table 4.8: True Implied Black-Scholes volatilities (%) for the displaced log-normal and CEV models and order 2 and 3 approximations using normal and log-normal proxys for $\nu = 0.25, \beta = 0.2$ and $T = 5Y$.

Strikes	0.25	0.40	0.50	0.60	0.75	0.85	1	1.15	1.35	1.60	2.05	2.50	3.60
True DLN	44.35	36.83	33.65	31.24	28.50	27.07	25.32	23.91	22.40	20.91	18.96	17.56	15.36
AppN_2	44.17	36.81	33.66	31.26	28.52	27.09	25.33	23.93	22.41	20.91	18.77	16.97	13.50
AppLN_2	31.86	31.87	31.02	29.81	27.85	26.63	25.00	23.60	21.95	19.90	ND	ND	ND
$\mathrm{AppN}_3^{\mathrm{DLN}}$	44.33	36.82	33.65	31.24	28.50	27.07	25.32	23.91	22.40	20.92	18.97	17.51	14.36
$\mathrm{AppLN}_3^{\mathrm{DLN}}$	38.18	35.47	33.04	30.89	28.34	26.97	25.30	23.94	22.48	21.09	19.84	21.57	29.94
True CEV	42.07	35.80	33.00	30.82	28.27	26.91	25.20	23.80	22.26	20.71	18.59	17.01	14.38
AppN_2	44.17	36.81	33.66	31.26	28.52	27.09	25.33	23.93	22.41	20.91	18.77	16.97	13.50
AppLN_2	31.86	31.87	31.02	29.81	27.85	26.63	25.00	23.60	21.95	19.90	ND	ND	ND
$\mathrm{AppN}_3^{\mathrm{CEV}}$	43.13	36.20	33.23	30.94	28.33	26.94	25.21	23.80	22.24	20.65	18.53	17.04	14.21
$\mathrm{AppLN}_3^{\mathrm{CEV}}$	37.93	35.11	32.72	30.65	28.18	26.85	25.19	23.82	22.31	20.80	19.09	20.42	29.37

We observe as expected that whatever is the value of β, the behaviors of the true implied volatilities of the displaced log-normal and the CEV models are very similar for short maturity and around the money. We notice a negative skew (the implied volatilities are larger ITM than ATM and smaller OTM than ATM) and the relative skew (w.r.t. the relative strikes) becomes more important when the maturity increases or when the slope of the local volatility function at spot (connected to $\beta - 1$) increases (in absolute value) corresponding to values of β close to 0 as $\beta = 0.2$.
However we can remark that the skew and the general level of the implied volatility seem more important for the displaced log-normal model. The difference is notably significant for long maturity and/or for far ITM or far OTM options. For the two values of β and for the maturity $5Y$, the implied volatility corresponding to the displaced log-normal model far ITM (respectively far OTM) is approximately 200 bps[2] higher (respectively 100 bps higher) than the implied volatility corresponding to the CEV model. This divergence for high maturities and very large moneyness in absolute value can be deduced from the difference in the curvatures of the local volatilities function. More precisely, if we refer to Theorem

[2] 1 bp (basis point) is equal to 0.01%.

2.3.3.1 of Chapter 2 Section 2.3, the third order Black-Scholes implied volatility at spot reads:

$$\sigma_I(x_0, T, k) = \gamma_0(a; x_0)_0^T - \gamma_1(a; x_0)_0^T m + \gamma_2(a; x_0)_0^T m^2$$

where we recall that x_0 is the log spot, k the log-strike and $m = x_0 - k$ the log moneyness. Using the definition of the coefficients γ (see Definitions 2.3.1.1 and 2.3.3.1 of Chapter 2 Section 2.3), one obtains that:

1. the contribution of $a^{(2)}(x_0)$ in γ_0 is exactly $\frac{a^2(x_0)a^{(2)}(x_0)T}{12}$,
2. the contribution of $a^{(2)}(x_0)$ in γ_1 is null,
3. the contribution of $a^{(2)}(x_0)$ in γ_2 is exactly $\frac{a^{(2)}(x_0)}{6}$.

Now using $a_d^{(2)}(0) = \nu(1-\beta) \geq a_c^{(2)}(0) = \nu(\beta-1)^2 \geq 0$ for any $\beta \in [0,1]$, we deduce that the maturity bias is higher in the displaced log-normal model (i.e. the smile is shifted from maturity to maturity with a higher value) as well the moneyness bias (i.e. the implied volatility in the displaced log-normal model is higher ITM and OTM than the corresponding CEV implied volatility).
The fact that the displaced log-normal model induces more skew yields to less accurate approximation formulas what can be observed in the tests especially for $\beta = 0.8$. As the order 2, the approximations (for both normal and log-normal proxys) generally underestimate the true CEV implied volatility, the results are even less accurate for the displaced log-normal model which induces higher implied volatility than the CEV model. For the order 3 approximation with log-normal proxy, we observe for the maturity $5Y$ errors greater than 100 bps when considering the displaced log-normal proxy for a maximal error of 39 bps for the CEV model.
For $\beta = 0.2$, curiously the order 2 approximation with normal proxy is closer to the true displaced implied volatility than the true CEV volatility far ITM but closer to the true CEV volatility far OTM. We similarly observe that the corresponding order 3 approximations yield to small errors far ITM (2 or less bps) for the displaced log-normal model and to large errors far OTM (up to 100 bps for $T = 5Y$), the converse being realised for the CEV model.

4.4.2 Comparison of the Gaussian, log-normal and displaced log-normal proxys for the CEV model

In this Subsection we compare the performance of the approximations at spot using a normal, a log-normal and a displaced log-normal proxy for the pricing in the CEV model. Note that for any $\beta \in [0,1]$, any $S_0 \in]0,1]$ and any $K > 0$ we have $\Sigma(S_0) = S_0^\beta$, $\Sigma^{(1)}(S_0) = \beta S_0^{\beta-1} > 0$ and
$K_d = \Sigma(S_0) + \Sigma^{(1)}(S_0)(K - S_0) = \beta K + S_0^\beta - \beta S_0 > \beta K + (1-\beta)S_0 > 0$ as required in the hypotheses of Theorem 4.3.0.3.
For the numerical experiments, we set again $\nu = 0.25$, $S_0 = 1$ and we allow β to vary by choosing the two values $\beta = 0.8$ and $\beta = 0.2$. We keep the same sets of maturities and strikes and report in Tables 4.9-4.10-4.11-4.12-4.13-4.14-4.15-4.16:

- The true value of the implied volatility in the CEV model denoted by `True CEV`.
- The implied volatility approximations obtained with the second order price approximations using the normal, the log-normal and the displaced log-normal proxys (see Theorem 4.3.0.3 equation (4.14)) denoted respectively by $\mathtt{AppN}_2^{\mathtt{CEV}}$, $\mathtt{AppLN}_2^{\mathtt{CEV}}$ and $\mathtt{AppDLN}_2^{\mathtt{CEV}}$.
- The implied volatility approximations obtained with the third order price approximations using the normal, the log-normal and the displaced log-normal proxys (see Theorem 4.3.0.3 equation (4.15)) denoted by $\mathtt{AppN}_3^{\mathtt{CEV}}$, $\mathtt{AppLN}_3^{\mathtt{CEV}}$ and $\mathtt{AppDLN}_3^{\mathtt{CEV}}$.

Table 4.9: True Implied Black-Scholes volatilities (%) for the CEV model and order 2 and 3 approximations using normal, log-normal and displaced lognormal proxys for $\nu = 0.25, \beta = 0.8$ and $T = 6M$.

Strikes	0.65	0.75	0.80	0.85	0.90	0.95	1	1.05	1.10	1.20	1.25	1.35	1.50
True CEV	26.09	25.73	25.56	25.41	25.27	25.13	25.00	24.88	24.76	24.55	24.45	24.26	24.00
$\mathtt{AppN}_2^{\mathtt{CEV}}$	23.41	25.54	25.56	25.46	25.31	25.17	25.03	24.91	24.81	24.56	24.39	23.91	22.83
$\mathtt{AppLN}_2^{\mathtt{CEV}}$	25.96	25.69	25.55	25.40	25.26	25.13	25.00	24.88	24.76	24.54	24.43	24.21	23.85
$\mathtt{AppDLN}_2^{\mathtt{CEV}}$	26.24	25.80	25.61	25.44	25.28	25.14	25.01	24.89	24.78	24.58	24.49	24.32	24.11
$\mathtt{AppN}_3^{\mathtt{CEV}}$	25.92	25.69	25.55	25.41	25.27	25.13	25.00	24.88	24.76	24.56	24.47	24.30	23.88
$\mathtt{AppLN}_3^{\mathtt{CEV}}$	26.09	25.73	25.56	25.41	25.27	25.13	25.00	24.88	24.76	24.55	24.45	24.26	24.01
$\mathtt{AppDLN}_3^{\mathtt{CEV}}$	26.11	25.73	25.57	25.41	25.27	25.13	25.00	24.88	24.76	24.55	24.44	24.25	23.98

Table 4.10: True Implied Black-Scholes volatilities (%) for the CEV model and order 2 and 3 approximations using normal, log-normal and displaced lognormal proxys for $\nu = 0.25, \beta = 0.8$ and $T = 1Y$.

Strikes	0.55	0.65	0.75	0.80	0.90	0.95	1	1.05	1.15	1.25	1.40	1.50	1.80
True CEV	26.53	26.10	25.73	25.56	25.27	25.13	25.00	24.88	24.65	24.45	24.17	24.00	23.56
$\mathtt{AppN}_2^{\mathtt{CEV}}$	18.04	25.62	25.79	25.67	25.35	25.20	25.07	24.95	24.74	24.51	23.96	23.41	21.21
$\mathtt{AppLN}_2^{\mathtt{CEV}}$	26.29	26.02	25.71	25.55	25.26	25.13	25.00	24.88	24.65	24.44	24.13	23.93	23.19
$\mathtt{AppDLN}_2^{\mathtt{CEV}}$	26.83	26.26	25.81	25.62	25.30	25.15	25.02	24.90	24.69	24.50	24.26	24.12	23.78
$\mathtt{AppN}_3^{\mathtt{CEV}}$	25.91	25.95	25.71	25.56	25.27	25.13	25.00	24.88	24.66	24.46	24.24	24.10	23.00
$\mathtt{AppLN}_3^{\mathtt{CEV}}$	26.51	26.09	25.73	25.56	25.27	25.13	25.00	24.88	24.65	24.45	24.17	24.00	23.59
$\mathtt{AppDLN}_3^{\mathtt{CEV}}$	26.58	26.12	25.74	25.57	25.27	25.13	25.00	24.88	24.65	24.44	24.16	23.98	23.51

Table 4.11: True Implied Black-Scholes volatilities (%) for the CEV model and order 2 and 3 approximations using normal, log-normal and displaced lognormal proxys for $\nu = 0.25, \beta = 0.8$ and $T = 2Y$.

Strikes	0.45	0.55	0.65	0.75	0.85	0.90	1	1.10	1.20	1.35	1.55	1.80	2.30
True CEV	27.06	26.53	26.10	25.73	25.41	25.27	25.01	24.77	24.55	24.26	23.92	23.56	22.98
$\mathtt{AppN}_2^{\mathtt{CEV}}$	ND	25.88	26.22	25.95	25.59	25.42	25.13	24.91	24.72	24.39	23.65	22.18	18.80
$\mathtt{AppLN}_2^{\mathtt{CEV}}$	26.69	26.39	26.05	25.71	25.41	25.26	25.00	24.76	24.54	24.24	23.86	23.39	22.07
$\mathtt{AppDLN}_2^{\mathtt{CEV}}$	27.63	26.86	26.29	25.84	25.48	25.32	25.05	24.81	24.61	24.35	24.07	23.79	23.39
$\mathtt{AppN}_3^{\mathtt{CEV}}$	25.43	26.17	26.02	25.72	25.41	25.27	25.00	24.77	24.55	24.29	24.07	23.73	21.21
$\mathtt{AppLN}_3^{\mathtt{CEV}}$	27.02	26.53	26.10	25.73	25.41	25.27	25.01	24.77	24.55	24.26	23.93	23.57	23.08
$\mathtt{AppDLN}_3^{\mathtt{CEV}}$	27.18	26.59	26.13	25.74	25.42	25.27	25.01	24.77	24.55	24.25	23.90	23.51	22.81

Table 4.12: True Implied Black-Scholes volatilities (%) for the CEV model and order 2 and 3 approximations using normal, log-normal and displaced lognormal proxys for $\nu = 0.25, \beta = 0.8$ and $T = 5Y$.

Strikes	0.25	0.40	0.50	0.60	0.75	0.85	1	1.15	1.35	1.60	2.05	2.50	3.60
True CEV	28.64	27.38	26.79	26.31	25.74	25.42	25.01	24.66	24.27	23.85	23.26	22.79	21.94
$\mathtt{AppN}_2^{\mathtt{CEV}}$	ND	26.97	27.32	26.95	26.23	25.81	25.33	25.01	24.68	24.11	22.10	19.45	14.41
$\mathtt{AppLN}_2^{\mathtt{CEV}}$	27.58	27.07	26.66	26.25	25.72	25.41	25.00	24.65	24.25	23.81	23.09	22.33	17.23
$\mathtt{AppDLN}_2^{\mathtt{CEV}}$	30.59	28.25	27.33	26.66	25.93	25.56	25.12	24.77	24.41	24.07	23.62	23.31	22.83
$\mathtt{AppN}_3^{\mathtt{CEV}}$	ND	26.35	26.52	26.24	25.72	25.41	25.01	24.67	24.28	23.97	23.83	22.69	16.72
$\mathtt{AppLN}_3^{\mathtt{CEV}}$	28.46	27.36	26.78	26.31	25.74	25.42	25.01	24.66	24.27	23.86	23.26	22.81	22.33
$\mathtt{AppDLN}_3^{\mathtt{CEV}}$	29.25	27.61	26.91	26.38	25.77	25.43	25.01	24.66	24.25	23.81	23.15	22.57	21.18

Table 4.13: True Implied Black-Scholes volatilities (%) for the CEV model and order 2 and 3 approximations using normal, log-normal and displaced lognormal proxys for $\nu = 0.25, \beta = 0.2$ and $T = 6M$.

Strikes	0.65	0.75	0.80	0.85	0.90	0.95	1	1.05	1.10	1.20	1.25	1.35	1.50
True CEV	29.59	28.02	27.32	26.69	26.09	25.54	25.02	24.53	24.08	23.24	22.85	22.13	21.18
$\mathtt{AppN}_2^{\mathtt{CEV}}$	29.66	28.08	27.37	26.72	26.11	25.55	25.03	24.55	24.09	23.27	22.88	22.16	21.15
$\mathtt{AppLN}_2^{\mathtt{CEV}}$	28.07	27.54	27.10	26.59	26.05	25.51	25.00	24.51	24.04	23.10	22.57	21.15	ND
$\mathtt{AppDLN}_2^{\mathtt{CEV}}$	29.73	28.09	27.37	26.72	26.11	25.55	25.03	24.55	24.09	23.27	22.89	22.20	21.29
$\mathtt{AppN}_3^{\mathtt{CEV}}$	29.60	28.02	27.33	26.69	26.09	25.54	25.02	24.53	24.08	23.24	22.85	22.13	21.18
$\mathtt{AppLN}_3^{\mathtt{CEV}}$	29.28	27.98	27.31	26.68	26.09	25.54	25.02	24.54	24.08	23.24	22.86	22.21	22.13
$\mathtt{AppDLN}_3^{\mathtt{CEV}}$	29.62	28.03	27.33	26.69	26.09	25.54	25.02	24.53	24.08	23.23	22.85	22.12	21.15

Table 4.14: True Implied Black-Scholes volatilities (%) for the CEV model and order 2 and 3 approximations using normal, log-normal and displaced lognormal proxys for $\nu = 0.25, \beta = 0.2$ and $T = 1Y$.

Strikes	0.55	0.65	0.75	0.80	0.90	0.95	1	1.05	1.15	1.25	1.40	1.50	1.80
True CEV	31.54	29.62	28.05	27.35	26.12	25.56	25.04	24.55	23.66	22.87	21.81	21.19	19.60
$\mathtt{AppN}_2^{\mathtt{CEV}}$	31.72	29.76	28.13	27.42	26.15	25.59	25.07	24.58	23.70	22.92	21.88	21.24	19.49
$\mathtt{AppLN}_2^{\mathtt{CEV}}$	28.96	28.57	27.70	27.16	26.05	25.51	25.00	24.51	23.59	22.68	21.04	19.15	ND
$\mathtt{AppDLN}_2^{\mathtt{CEV}}$	31.83	29.79	28.13	27.41	26.15	25.59	25.06	24.58	23.69	22.92	21.90	21.31	19.83
$\mathtt{AppN}_3^{\mathtt{CEV}}$	31.57	29.65	28.06	27.36	26.12	25.56	25.04	24.55	23.66	22.86	21.80	21.18	19.62
$\mathtt{AppLN}_3^{\mathtt{CEV}}$	30.90	29.51	28.02	27.34	26.11	25.56	25.04	24.56	23.67	22.88	21.86	21.37	23.04
$\mathtt{AppDLN}_3^{\mathtt{CEV}}$	31.61	29.66	28.06	27.36	26.12	25.56	25.04	24.55	23.66	22.86	21.80	21.16	19.52

$\triangleright$**Second order approximations.** For $\beta = 0.8$, the results of the second order approximation using the displaced log-normal proxy are very close to the results of the second order approximation using the log-normal proxy. Generally the use of the log-normal proxy underestimates the true implied volatility whereas the use of the displaced log-normal proxy overestimates the true implied volatility. We remark that up to the maturity $5Y$, the log-normal proxy yields to smaller errors far ITM and to larger errors OTM than those obtained with the displaced log-normal proxy.
For $\beta = 0.2$, the behaviour of the displaced log-normal proxy is closed to the behaviour of the normal proxy. Although the use of the Gaussian proxy yields to slightly better results, the use of the displaced log-normal proxy remains a good alternative.

Table 4.15: True Implied Black-Scholes volatilities (%) for the CEV model and order 2 and 3 approximations using normal, log-normal and displaced lognormal proxys for $\nu = 0.25, \beta = 0.2$ and $T = 2Y$.

Strikes	0.45	0.55	0.65	0.75	0.85	0.90	1	1.10	1.20	1.35	1.55	1.80	2.30
True CEV	34.03	31.62	29.69	28.10	26.76	26.16	25.08	24.13	23.29	22.18	20.92	19.62	17.62
$\mathrm{AppN}_2^{\mathrm{CEV}}$	34.48	31.93	29.89	28.23	26.84	26.23	25.13	24.18	23.35	22.27	21.03	19.68	17.30
$\mathrm{AppLN}_2^{\mathrm{CEV}}$	30.08	29.74	28.90	27.79	26.62	26.05	25.00	24.05	23.16	21.85	19.67	ND	ND
$\mathrm{AppDLN}_2^{\mathrm{CEV}}$	34.59	31.96	29.89	28.22	26.83	26.22	25.13	24.18	23.34	22.26	21.07	19.86	18.06
$\mathrm{AppN}_3^{\mathrm{CEV}}$	34.16	31.69	29.73	28.12	26.77	26.17	25.08	24.13	23.28	22.16	20.90	19.61	17.65
$\mathrm{AppLN}_3^{\mathrm{CEV}}$	32.98	31.36	29.62	28.07	26.74	26.15	25.08	24.14	23.30	22.21	21.01	20.39	25.32
$\mathrm{AppDLN}_3^{\mathrm{CEV}}$	34.21	31.71	29.74	28.13	26.77	26.17	25.08	24.13	23.28	22.16	20.88	19.54	17.39

Table 4.16: True Implied Black-Scholes volatilities (%) for the CEV model and order 2 and 3 approximations using normal, log-normal and displaced lognormal proxys for $\nu = 0.25, \beta = 0.2$ and $T = 5Y$.

Strikes	0.25	0.40	0.50	0.60	0.75	0.85	1	1.15	1.35	1.60	2.05	2.50	3.60
True CEV	42.07	35.80	33.00	30.82	28.27	26.91	25.20	23.80	22.26	20.71	18.59	17.01	14.38
$\mathrm{AppN}_2^{\mathrm{CEV}}$	44.17	36.81	33.66	31.26	28.52	27.09	25.33	23.93	22.41	20.91	18.77	16.97	13.50
$\mathrm{AppLN}_2^{\mathrm{CEV}}$	31.86	31.87	31.02	29.81	27.85	26.63	25.00	23.60	21.95	19.90	ND	ND	ND
$\mathrm{AppDLN}_2^{\mathrm{CEV}}$	44.35	36.83	33.65	31.24	28.50	27.07	25.32	23.91	22.40	20.91	18.96	17.56	15.36
$\mathrm{AppN}_3^{\mathrm{CEV}}$	43.13	36.20	33.23	30.94	28.33	26.94	25.21	23.80	22.24	20.65	18.53	17.04	14.21
$\mathrm{AppLN}_3^{\mathrm{CEV}}$	37.93	35.11	32.72	30.65	28.18	26.85	25.19	23.82	22.31	20.80	19.09	20.42	29.37
$\mathrm{AppDLN}_3^{\mathrm{CEV}}$	43.29	36.24	33.24	30.95	28.33	26.94	25.21	23.79	22.23	20.64	18.41	16.68	13.49

▷**Third order approximations.** The observations are similar to the second order case.
For $\beta = 0.8$, the results of the displaced log-normal proxy are almost so good as the results of the log-normal proxy. The use of the log-normal proxy yields to underestimation ITM and overestimation OTM, the converse being realised for the displaced log-normal proxy.
Similarly for $\beta = 0.2$, the use of the displaced log-normal proxy yields to errors almost of the same magnitude than those obtained with the normal proxy. The approximation with the displaced log-normal overestimates ITM the true implied volatility and underestimates OTM the true implied volatility.

Consequently the approximation formulas using the displaced log-normal proxy yield to very good results for large as well for small values of β and we never obtain price outside the non-arbitrage bounds. This stability is a significant advantage because it is safe to use a method which is competitive in all situations instead of employing an approximation which is excellent in some situations but leads to poor results in other situations. In addition we pay attention to the simplicity of the second order approximation using the displaced log-normal proxy which reduces to a suitable rescaled Black-Scholes price. The third order approximation using the displaced log-normal proxy contains less terms than the corresponding formulas using the normal and log-normal proxys. For all these advantages the use of the displaced log-normal proxy is very promising.
It could be interesting to adapt for the displaced log-normal proxy some heuristics developed in the previous Chapter 2 (to freeze the local volatility at some intermediate point between the strike and the spot, to translate the price approximation in volatility expansion...) in order to achieve better numerical accuracy. These extensions are left for further research.

Part II

Models combining local and stochastic volatility

CHAPTER 5

Price expansion formulas for model combining local and stochastic volatility

Contents

This Chapter consists in introducing an option price expansion for model combining local and stochastic volatility with tight error estimates. The local volatility part is considered as general but has to satisfy some growth and boundedness assumptions. For the stochastic part, we choose a square root process, which is usually used for modelling the behaviour of the variance process. In the particular case of Call options, we also provide expansions of the Black-Scholes implied volatility which allow to obtain very simple and rapid formulas in comparison to the Monte Carlo approach while maintaining a very competitive accuracy.

5.1 Introduction

Models combining local and stochastic volatility have emerged in the last decade to offer more flexibility in the skew and smile management. This includes for instance the well known SABR model introduced by Hagan et al. in [Hagan 2002] or more recently the CEV-Heston model studied notably by Forde et al. [Forde 2012b] (see [Cox 1975] for the CEV model and [Heston 1993] for the Heston model). If the interest of such models is increasing, their use is still challenging because of the lack of closed-form formulas. The price to be paid for the more realistic dynamic is the use of time costly numerical methods like PDE engine or Monte Carlo simulations. In this work we focus on models with general local volatility function and a stochastic variance modelled by a square root process and in a perspective of real time calibration procedures, we aim at providing analytical approximations.

▹**Comparison with the literature.** In the two last decades, an impressive number of papers have been devoted to the analytical approximations and their applications to finance. Although the large body of the existing literature is mainly focusing on pure local volatility or pure stochastic volatility models, we count recently some studies focusing on hybrid local and stochastic volatility models. We cite among them regrouping the similar approaches:
Geodesic approach and small maturity expansions: we refer to Hagan et al. [Hagan 2002], Berestycki et al. [Berestycki 2004], Henry-Labordère [Henry-Labordère 2005]-[Henry-Labordère 2008] and Lewis [Lewis 2007] who used an explicit computation of the geodesic distance in the SABR model to derive short maturity implied volatility expansions. More recently we cite the work of Jordan et al. [Jordan 2011] who utilise the WKB or ray method (see [Keller 1978]) and boundary layer corrections to derive the asymptotic behavior of the density function in the SABR model for small maturities. We finally cite Forde et al. [Forde 2012a]: using small noise expansions inspired by [Freidlin 1998] and large deviation arguments, the authors provide small-time implied volatility expansions in general local and stochastic volatility models but under restrictive conditions: null correlation, strong hypotheses of the stochastic volatility coefficients (excluding square root processes) and uniform ellipticity condition for the local volatility function. Drawbacks of the geodesic approach are: 1) accuracy restricted to short maturities; 2) validity only for time homogeneous parameters.
Long maturities point of view with fixed strike/large strike regime: Forde et al. study in [Forde 2012b] the large-time asymptotic of SABR and CEV-Heston models in different regime of strikes using the large deviation theory and saddlepoint methods. But limitations are: 1) only available for time-independent parameters; 2) limited to null correlation.
Ergodic approach: see Fouque et al. [Choi 2010] where an asymptotic expansion w.r.t. a fast mean reversion parameter of the volatility is performed in a particular hybrid model built on a CEV-type local volatility and a stochastic volatility driven by an Ornstein Uhlenbeck process.
Perturbation methods: we cite the very recent paper of Pascucci et al. [Pagliarani 2013a] which provides an expansion of the characteristic function in a general local and stochastic volatility model (possibly incorporating also Lévy jumps). To perform the approximation of the characteristic function, the authors used their so-called Adjoint Expansion PDE method which is inspired of the well-known singular perturbations in the work of Hagan et al. [Hagan 1999] for the CEV model. Then they obtain option price approximation formulas using Fourier methods. Drawbacks of the method are: 1) error estimates only available under condition of uniform parabolic PDEs; 2) necessary to perform finely numerical integrations in the Fourier inversion step.
Some improvements of the methodology have been proposed in [Lorig 2013a]: 1) more general expansions are considered extending the framework to the multidimensional case; 2) expansion coefficients are fully explicit without numerical integration; 3) analytical approximation of implied volatilities are provided.

As a difference with several quoted papers which doesn't satisfy all the following conditions, we aim at giving an explicit and accurate analytical formula:

1. covering both short and long maturities,

2. handling general local volatility function, non-null correlation as well time-dependent parameters,
3. with computational time close to zero,
4. with complete mathematical justification.

To achieve this, we use the so called Proxy principle introduced in [Benhamou 2009] and [Benhamou 2010a] to perform a non-asymptotic expansion with the help of a Proxy process.

▹**Comparison with previous works and contribution of the Chapter.** The approach still consists in expanding the price w.r.t. parameters of the model using a smart parameterization and in computing the correction terms using Malliavin calculus in the Gaussian framework. As a difference with the work on Heston models [Benhamou 2010b] for which one can use a conditioning argument to represent the price as a simple expectation related to the variance process, we follow a direct approach like in [Benhamou 2010a] or [Benhamou 2009], with a suitable parameterization of both the price and variance processes. We provide an explicit third order formula order w.r.t. the interest parameters, the leading term being a suitable Black-Scholes price, while the other terms are sensitivities in the Black-Scholes framework weighted with functionals of the model parameters. This allows in particular to retrieve the results of [Benhamou 2010a] for pure local volatility models and of [Benhamou 2010b] for pure Heston models. To go even further than the cited references, we also provide implied volatility expansions for the particular case of Call options.
Note also that the Malliavin differentiability of local and stochastic volatility models is not standard and may fail for high order (see [Alòs 2008] for Heston models). To overcome this difficulty, we use the Malliavin calculus on smooth processes very close in L^p to the initial one in order to prove the accuracy of our formulas.

▹**Formulation of the problem.** We are given a maturity $T > 0$ (typically the maturity of the financial product we attempt to price) and we consider the solution of the stochastic differential equation (SDE):

$$\mathrm{d}X_t = \sigma(t, X_t)\sqrt{V_t}\mathrm{d}W_t - \frac{1}{2}\sigma^2(t, X_t)V_t\mathrm{d}t,\ X_0 = x_0, \tag{5.1}$$

$$\mathrm{d}V_t = \alpha_t\mathrm{d}t + \xi_t\sqrt{V_t}\mathrm{d}B_t,\ V_0 = v_0, \tag{5.2}$$

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

where $(B_t, W_t)_{0\le t\le T}$ is a two-dimensional correlated Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P})$ with the usual assumptions on the filtration $(\mathcal{F}_t)_{0\le t\le T}$. In our setting, $(X_t)_{t\in[0,T]}$ is the log of the forward price, σ the local volatility function and $(V_t)_{t\in[0,T]}$ is a square root process with an initial value $v_0 > 0$, a measurable, positive and bounded drift function $(\alpha_t)_{t\in[0,T]}$ and a measurable, positive and bounded volatility of volatility function $(\xi_t)_{t\in[0,T]}$.
We work with the model of stochastic variance (5.2) for the sake of clarity in the calculus, but the results developed throughout the Chapter can be adapted for a time-dependent CIR process:

$$\mathrm{d}Y_t = \kappa_t(\theta_t - Y_t)\mathrm{d}t + \gamma_t\sqrt{Y_t}\mathrm{d}B_t.$$

A simple space-change $y \mapsto e^{\int_0^t \kappa_s \mathrm{d}s} y$ allows us namely to retrieve the formulation (5.2). See details in Appendix 5.6.1.
Our aim is to give an accurate analytical approximation of any European option price of the form:

$$\mathbb{E}[h(X_T)], \tag{5.3}$$

where $\mathbb{E}$ stands for the standard expectation operator (under a risk neutral probability measure) and h is a given Lipschitz bounded payoff function. To accomplish this, we choose a proxy model in which analytical calculus are possible. At first glance we approximate the process $(X_t, V_t)_{t\in[0,T]}$ defined in (5.1)-(5.2) by the following Gaussian process:

$$\mathrm{d}X_t^P = \sigma(t, x_0)\sqrt{v_t}\mathrm{d}W_t - \frac{1}{2}\sigma^2(t, x_0)v_t\mathrm{d}t, \tag{5.4}$$

$$v_t = v_0 + \int_0^t \alpha_s \mathrm{d}s. \tag{5.5}$$

Such an approximation can be justified if one of the two following situations holds: i) the volatility of volatility is quite small leading to the approximation $V_t \approx v_t$ (which is realistic in practice) **and** the local volatility function $\sigma(t,\cdot)$ has small variations, which means that $\sigma(t, X_t) \approx \sigma(t, x_0)$; ii) the local part of the diffusion component is small (i.e. $|\sigma|_\infty$ small) which implies $X_t \approx x_0$, and thus $\sigma(t, X_t) \approx \sigma(t, x_0)$. Besides we expect to have additionally better approximations for small maturities (leading to $X_t \approx x_0$ and $V_t \approx v_0$, $t \in [0, T]$).

Remark 5.1.0.1. *The proposed Proxy is one-dimensional and thus its single use without correction term does not capture the effect of the correlation. It could be interesting to consider a two-dimensional Gaussian Proxy. This is left for further research.*

To link the initial process (5.1)-(5.2) and the proxy process (5.4)-(5.5), we introduce a two-dimensional parameterized process given by:

$$\mathrm{d}X_t^\eta = \sigma(t, \eta X_t^\eta + (1-\eta)x_0)\sqrt{V_t^\eta}\mathrm{d}W_t - \frac{1}{2}\sigma^2(t, \eta X_t^\eta + (1-\eta)x_0)V_t^\eta\mathrm{d}t, \ \ X_0^\eta = x_0, \tag{5.6}$$

$$\mathrm{d}V_t^\eta = \alpha_t\mathrm{d}t + \eta\xi_t\sqrt{V_t^\eta}\mathrm{d}B_t, \ v_0, \tag{5.7}$$

where η is an interpolation parameter lying in the range $[0, 1]$, such that on the one hand for $\eta = 1$, $X_t^1 = X_t$ and $V_t^1 = V_t$, and on the other hand for $\eta = 0$, $X_t^0 = X_t^P$ and $V_t^0 = v_t$. This parameterization is only a way to connect X and the proxy model X^P and to derive successive corrective processes in order to obtain a tractable approximation formula.

▷**Outline of the Chapter.** The Chapter is organised as follows. In Section 5.2 we present a third order price approximation formula in Theorem 5.2.2.1 which is the main result of the Chapter. We also provide the magnitude of the error term. The result is followed by an outline of the proof to present in an heuristic way the methodology to perform the expansion and to draw the attention of the reader to the main difficulties. The explicit calculus of the expansion coefficients is postponed to Appendix 5.6.2. Section 5.3 is devoted to the complete proof of the error estimate. Analyse the accuracy of the formula is far from straightforward and constitutes the technical core of the Chapter. In the Section 5.4 we apply our expansion formula to the particular case of Call/Put options to derive implied volatility expansions with local volatility frozen at spot and at mid-point between the strike and the spot. Results are stated in Theorems 5.4.1.1 and 5.4.2.1. Section 5.5 is gathering numerical experiments illustrating the performance and the rapidity of our implied volatility formulas in comparison to the Monte Carlo simulations. In Appendix 5.6, we give intermediate and complementary results.

5.2 Main Result

5.2.1 Notations and definitions

The following notations and definitions are frequently used in the following.

▷**Extremes of deterministic functions.** For measurable and bounded functions $f : [0, T] \to \mathbb{R}$

and $g : [0,T]\times\mathbb{R} \to \mathbb{R}$, we denote $f_{inf} = \inf_{t\in[0,T]} f_t$, $f_{sup} = \sup_{t\in[0,T]} f_t$, $|g|_\infty = \sup_{t\in[0,T],x\in\mathbb{R}} |g(t,x)|$ and $g_{\inf} = \inf_{t\in[0,T],x\in\mathbb{R}} g(t,x)$.

▷**Differentiation and Integration.** If these derivatives have a meaning, we write: $\psi_t^{(i)}(x) = \partial^i_{x^i}\psi(t,x)$ for any measurable function ψ of $(t,x)\in[0,T]\times\mathbb{R}$. When considering the spatial point x_0, we oftenly use if unambiguous the notations $\psi_t = \psi_t(x_0)$ and $\psi_t^{(i)} = \psi_t^{(i)}(x_0)$.

Definition 5.2.1.1. ***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 $t\in[0,T]$.

Definition 5.2.1.2. ***Greeks.*** *Let Z be a random variable and h a payoff function. The i-th Greek for the variable Z is defined by the quantity (when it has a meaning):*

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

Given appropriate smoothness assumptions concerning h, one also has:

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

▷**Assumptions on σ and $(V_t)_{t\le T}$.**

- $(\mathcal{H}_{x_0}^\sigma)$: σ is a bounded measurable function of $(t,x)\in[0,T]\times\mathbb{R}$ and three times continuously differentiable w.r.t. x with bounded[1] derivatives. Set

$$\mathcal{M}_1(\sigma) = \max_{1\le i\le 3}|\partial^i_{x^i}\sigma(t,x)|_\infty \text{ and } \mathcal{M}_0(\sigma) = \max_{0\le i\le 3}|\partial^i_{x^i}\sigma(t,x)|_\infty.$$

  In addition, we assume the following ellipticity condition: $\int_0^T \sigma_t^2 v_t \mathrm{d}t > 0$ (local non-degeneracy condition).

- (P): α and ξ are measurable, bounded on $[0,T]$ and positive. In addition $\xi_{\inf} > 0$ and $2(\frac{\alpha}{\xi^2})_{\inf} \ge 1$.

Remark 5.2.1.1. *Because there exists a unique process $(V_t)_{t\le T}$ satisfying the SDE (5.2), $(\mathcal{H}_{x_0}^\sigma)$ guarantees the existence and the uniqueness of a solution for (5.1), considering generalized stochastic integration w.r.t. semi-martingales (see [Protter 2004, Theorem 6 p. 249]). In addition (P) implies that $\forall\eta\in[0,1]$, $\mathbb{P}(\forall t\in[0,T]: V_t^\eta > 0) = 1$ (See Lemma [Benhamou 2010b, Lemma 4.2], and replace in the original paper κ by 0 and $\kappa\theta_t$ by α_t).*

We define the stochastic volatility process:

Definition 5.2.1.3. $\Lambda_t^\eta = \sqrt{V_t^\eta}$, $\forall t\in[0,T]$, $\forall\eta\in[0,1]$.

[1] the boundedness assumption of σ and its derivatives could be weakened to L^p-integrability conditions, up to extra works.

In addition, we introduce $(\lambda_t)_{t\in[0,T]}$ defined for any $t \in [0,T]$ by:

$$\lambda_t = \Lambda_t^{\eta=0} = \sqrt{V_t^0} = \sqrt{v_t} = \sqrt{v_0 + \int_0^t \alpha_s \mathrm{d}s}. \tag{5.8}$$

▷**Assumptions on the payoff function h.** We denote by $C_0^\infty(\mathbb{R})$, the space of real-valued infinitely differentiable functions with compact support. For practical applications in finance, assuming that $h \in C_0^\infty(\mathbb{R})$ is too strong and we introduce $\mathrm{Lip_b}(\mathbb{R})$, the space of Lipschitz bounded real-valued functions in the following sense: for some positive constants C_h and L_h:

$$\begin{cases} |h(x)| \le C_h & \forall x \in \mathbb{R}, \\ |\frac{h(y)-h(x)}{y-x}| \le L_h & \forall (x,y) \in \mathbb{R}^2, x \neq y, \end{cases}$$

This space includes the classical Put payoff function $x \mapsto (K-e^x)_+$ with strike K. Assume that h and/or its first derivative (defined a.e.) is exponentially bounded could lead to technical difficulties in the L^p-estimates because exponential moments of integrated square root processes explode (see [Andersen 2006]).

▷**Generic constants and upper bounds.** We keep the same notation C for all non-negative constants depending on: universal constants, on a number $p \ge 1$ arising in L^p-estimates, in a non decreasing way on $\xi_{\sup}$, $\mathcal{M}_0(\sigma)$, $\mathcal{M}_1(\sigma)$, T, $\frac{|\sigma|_\infty^2 T}{\int_0^T \sigma_t^2 v_t \mathrm{d}t}$ and $\alpha_{\sup}$. Usually, a generic constant may depend on v_0 (and notably depends on negative powers of v_0). A generic constant **does not depend** on x_0 or remains uniformly bounded in this variable.
We frequently use the short notation $A \le_c B$ for positive A which means that $A \le CB$ for a generic constant C. Similarly "$A = O(B)$" means that $|A| \le CB$ for a generic constant C.

▷**Miscellaneous.** The L^p-norm of a random variable is denoted, as usual, by $\|.\|_p$

5.2.2 Third order approximation price formula

We state the main result of the Chapter:

Theorem 5.2.2.1. *(3rd order approximation price formula.) Assume $(\mathcal{H}_{x_0}^\sigma)$ and (P). Then for any $h \in \mathrm{Lip_b}(\mathbb{R})$, we have:*

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] + \sum_{i=1}^{6} \eta_{i,T} \mathcal{G}_i^h(X_T^P) + \mathrm{Error}_{3,\mathrm{h}}, \tag{5.9}$$

where:

$$\begin{aligned}
\eta_{1,T} =& \frac{C_{1,T}^l}{2} - \frac{C_{2,T}^l}{2} - \frac{C_{3,T}^l}{4} - \frac{C_{4,T}^l}{2} - C_{1,T}^{ls}, \\
\eta_{2,T} =& -\frac{3C_{1,T}^l}{2} + \frac{C_{2,T}^l}{2} + \frac{5C_{3,T}^l}{4} + \frac{7C_{4,T}^l}{2} + \frac{(C_{1,T}^l)^2}{8} - \frac{C_{1,T}^s}{2} \\
&+ \frac{C_{3,T}^s}{4} + C_{1,T}^{ls} + \frac{C_{2,T}^{ls}}{2} + \frac{C_{3,T}^{ls}}{2} + C_{4,T}^{ls} + \frac{C_{5,T}^{ls}}{2} + \frac{C_{6,T}^{ls}}{4}, \\
\eta_{3,T} =& C_{1,T}^l - 2C_{3,T}^l - 6C_{4,T}^l - \frac{3(C_{1,T}^l)^2}{4} + \frac{C_{1,T}^s}{2} - \frac{C_{2,T}^s}{2} - \frac{C_{3,T}^s}{2} \\
&- \frac{3C_{2,T}^{ls}}{2} - \frac{3C_{3,T}^{ls}}{2} - \frac{5C_{4,T}^{ls}}{2} - C_{5,T}^{ls} - \frac{3C_{6,T}^{ls}}{4} - \frac{C_{1,T}^l C_{1,T}^s}{4},
\end{aligned}$$

$$\begin{aligned}
\eta_{4,T} =& C^l_{3,T} + 3C^l_{4,T} + \frac{13(C^l_{1,T})^2}{8} + \frac{C^s_{2,T}}{2} + \frac{C^s_{3,T}}{4} + C^{ls}_{2,T} \\
&+ C^{ls}_{3,T} + \frac{3C^{ls}_{4,T}}{2} + \frac{C^{ls}_{5,T}}{2} + \frac{C^{ls}_{6,T}}{2} + \frac{(C^s_{1,T})^2}{8} + C^l_{1,T}C^s_{1,T}, \\
\eta_{5,T} =& -\frac{3(C^l_{1,T})^2}{2} - \frac{(C^s_{1,T})^2}{4} - \frac{5C^l_{1,T}C^s_{1,T}}{4}, \\
\eta_{6,T} =& \frac{(C^l_{1,T})^2}{2} + \frac{(C^s_{1,T})^2}{8} + \frac{C^l_{1,T}C^s_{1,T}}{2},
\end{aligned}$$

and:

$$\begin{aligned}
&C^l_{1,T} = \omega(\sigma^2 v, \sigma\sigma^{(1)} v)^T_0, && C^l_{2,T} = \omega(\sigma^2 v, ((\sigma^{(1)})^2 + \sigma\sigma^{(2)})v)^T_0, && \\
&C^l_{3,T} = \omega(\sigma^2 v, \sigma^2 v, ((\sigma^{(1)})^2 + \sigma\sigma^{(2)})v)^T_0, && C^l_{4,T} = \omega(\sigma^2 v, \sigma\sigma^{(1)} v, \sigma\sigma^{(1)} v)^T_0, && C^s_{1,T} = \omega(\rho\xi\sigma v, \sigma^2)^T_0, \\
&C^s_{2,T} = \omega(\rho\xi\sigma v, \rho\xi\sigma, \sigma^2)^T_0, && C^s_{3,T} = \omega(\xi^2 v, \sigma^2, \sigma^2)^T_0, && C^{ls}_{1,T} = \omega(\rho\xi\sigma v, \sigma\sigma^{(1)})^T_0, \\
&C^{ls}_{2,T} = \omega(\rho\xi\sigma v, \sigma^2 v, \sigma\sigma^{(1)})^T_0, && C^{ls}_{3,T} = \omega(\sigma^2 v, \rho\xi\sigma v, \sigma\sigma^{(1)})^T_0, && C^{ls}_{4,T} = \omega(\rho\xi\sigma v, \sigma^2, \sigma\sigma^{(1)} v)^T_0, \\
&C^{ls}_{5,T} = \omega(\rho\xi\sigma v, \sigma\sigma^{(1)} v, \sigma^2)^T_0, && C^{ls}_{6,T} = \omega(\sigma^2 v, \rho\xi\sigma^{(1)} v, \sigma^2)^T_0. &&
\end{aligned}$$

Then the approximation error is estimated as follows:

$$\text{Error}_{3,\text{h}} = O(L_h|\sigma|_\infty[\xi^3_{\sup} + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2]T^2). \tag{5.10}$$

Remark 5.2.2.1. *Under $(\mathcal{H}^\sigma_{x_0})$, X^P_T is a non-degenerate normal random variable and consequently, whatever is the regularity of h, the Greeks $\mathcal{G}^i_h(X^P_T)$ introduced in (5.2.1.2) exist and are well defined for any integer i. Note also that on the contrary to [Benhamou 2010b, Theorem 2.2], we do not assume anymore that the correlation is bounded away to -1 and 1.*

Remark 5.2.2.2. *The magnitude of* $\text{Error}_{3,\text{h}}$ *provided in (5.10) justifies the label of third order approximation formula because using the notation $M = \max(\mathcal{M}_0(\sigma), \xi_{\sup})$, we readily have* $\text{Error}_{3,\text{h}} = O((M\sqrt{T})^4)$. *Besides, making reference to the introduction, we retrieve that if $|\sigma|_\infty = 0$ or* $\max(\mathcal{M}_1(\sigma), \xi_{\sup}) = 0$ *or $T = 0$, the approximation formula (5.9) is exact (the model and the proxy coincide and the C coefficients vanish). In addition if $L_h = 0$ (i.e. h is constant), the error is equal to zero as well the sensitivities.*

Remark 5.2.2.3. *If one prefers to restrict to a second order approximation, it simply writes:*

$$\begin{aligned}
\mathbb{E}[h(X_T)] =& \mathbb{E}[h(X^P_T)] + C^l_{1,T}[\frac{1}{2}\mathcal{G}^h_1(X^P_T) - \frac{3}{2}\mathcal{G}^h_2(X^P_T) + \mathcal{G}^h_3(X^P_T)] + \frac{C^s_{1,T}}{2}[-\mathcal{G}^h_2(X^P_T) + \mathcal{G}^h_3(X^P_T)] \\
&+ O(L_h|\sigma|_\infty[\xi^2_{\sup} + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})]T^{\frac{3}{2}}).
\end{aligned}$$

We let the reader verify that the additional corrective terms of the expansion (5.9) are bounded up to generic constants by $L_h|\sigma|_\infty[\xi^2_{\sup} + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})]T^{\frac{3}{2}}$ using standard upper bounds for the derivatives of the Gausssian density and the magnitude of the additional coefficients C.

5.2.3 Corollaries and outline of the proof

▷Particular cases of pure local volatility models and pure stochastic volatility models.

a) Observe that if $\xi_{\sup}$ is equal to zero, all the coefficients C^s and C^{ls} are null and then we exactly retrieve the expansion of the pure local volatility model proposed in [Benhamou 2010a] (taking into account the contribution of λ_t). The terms C^l therefore read as purely local contributions.

b) If $\mathcal{M}_1(\sigma)=0$ (case of pure Heston model), all the coefficients C^l and C^{ls} are equal to zero: we retrieve the development that was found in [Benhamou 2010b] (with the contributions of σ_t and considering that $\kappa=0$ whereas $\kappa\theta_t=\alpha_t$). To see this, one has to transform the sensitivities w.r.t. the total variance which appear in [Benhamou 2010b, Theorem 2.2] in terms of sensitivities w.r.t. the log-spot.
c) Finally we interpret the coefficients C^{ls} as a mixture contribution of both the local and stochastic parts of the volatility. All these terms notably depend on the correlation.
In case of independence of W and B, all the coefficients are equal to 0 except the C^l terms and $C^s_{3,T}$.

▷**Applications to Call payoff function.** One can directly apply this Theorem for the Put payoff function $h(x)=(K-e^x)_+$. The reader should remark that the above expansion formula is exact for the particular payoff function $h(x)=\exp(x)$ (indeed $\mathbb{E}[h(X_T)]=\mathbb{E}[h(X_T^P)]=\mathcal{G}_i^{\exp}(X_T^P)=e^{x_0}$ and the sum of the corrective terms is equal to zero). This implies that the expansion remains valid for the Call payoff function $h(x)=(e^x-K)_+$ although $h\notin \mathrm{Lip_b}(\mathbb{R})$ (one can replace C_h and L_h by the strike K) and that the Call/Put parity relationship is preserved within these approximations.

▷**Outline of the proof.** We present here a sketch of proof in order to fix the main ideas and to point the finger at the principal difficulties.
The first step is to construct corrective processes to approximate X_T in L^p. Consider the parameterized process defined in (5.6)-(5.7). We recall that the Gaussian proxy process $(X_t^P)_{t\in[0,T]}$ defined in (5.4) is obtained by setting $\eta=0$. The next corrective processes $(X_{i,t})_{t\in[0,T]}$-$(V_{i,t})_{t\in[0,T]}$-$(\Lambda_{i,t})_{t\in[0,T]}$ for $i\in\{1,2\}$ are obtained by a formal i-times differentiation of (5.6)-(5.7) w.r.t. η and by taking $\eta=0$ thereafter. For the first corrective processes, we obtain:

$$
\begin{aligned}
\mathrm{d}X_{1,t} =& [(X_t^P-x_0)\sigma_t^{(1)}\lambda_t+\Lambda_{1,t}\sigma_t](\mathrm{d}W_t-\sigma_t\lambda_t\mathrm{d}t), \qquad X_{2,0}=0, && (5.11)\\
V_{1,t} =& \int_0^t \xi_s\lambda_s \mathrm{d}B_s, && (5.12)\\
\Lambda_{1,t} =& \frac{V_{1,t}}{2\lambda_t}. && (5.13)
\end{aligned}
$$

The second corrective processes are:

$$
\begin{aligned}
\mathrm{d}X_{2,t} =& \{\lambda_t[(X_t^P-x_0)^2\sigma_t^{(2)}+2X_{1,t}\sigma_t^{(1)}]+2(X_t^P-x_0)\Lambda_{1,t}\sigma_t^{(1)}\}(\mathrm{d}W_t-\sigma_t\lambda_t\mathrm{d}t) && (5.14)\\
&+\{\Lambda_{2,t}\sigma_t\mathrm{d}W_t-[(X_t^P-x_0)V_{1,t}\sigma_t^{(1)}\sigma_t+(X_t^P-x_0)^2(\sigma_t^{(1)})^2v_t+\frac{V_{2,t}}{2}\sigma_t^2]\mathrm{d}t\},\ X_{2,0}=0,\\
V_{2,t} =& \int_0^t \xi_s\frac{V_{1,s}}{\lambda_t}\mathrm{d}B_s, && (5.15)\\
\Lambda_{2,t} =& \frac{V_{2,t}}{2\lambda_t}-\frac{V_{1,t}^2}{4(\lambda_t)^3}. && (5.16)
\end{aligned}
$$

The reader will notice that under $(\mathcal{H}_{x_0}^{\sigma})$, these corrective processes $(X_{i,t})_{t\in[0,T]}$-$(V_{i,t})_{t\in[0,T]}$-$(\Lambda_{i,t})_{t\in[0,T]}$ for $i\in\{1,2\}$ are well defined.
The second step is to compute the corrective terms. Assuming that $h\in C_0^\infty(\mathbb{R})$, we perform a third order Taylor expansion for the function h at $x=X_T$ around $x=X_T^P$:

$$
\begin{aligned}
\mathbb{E}[h(X_T)] =& \mathbb{E}[h(X_T^P)]+\mathbb{E}[h^{(1)}(X_T^P)(X_T-X_T^P)]+\frac{1}{2}\mathbb{E}[h^{(2)}(X_T^P)(X_T-X_T^P)^2] && (5.17)\\
&+\mathbb{E}[(X_T-X_T^P)^3\int_0^1 h^{(3)}(X_T^P+\eta(X_T-X_T^P))\frac{(1-\eta)^2}{2}\mathrm{d}\eta]\\
=&\mathbb{E}[h(X_T^P)]+\mathbb{E}[h^{(1)}(X_T^P)X_{1,T}]+\mathbb{E}[h^{(1)}(X_T^P)\frac{X_{2,T}}{2}]+\frac{1}{2}\mathbb{E}[h^{(2)}(X_T^P)X_{1,T}^2]+\mathrm{Error}_{3,\mathrm{h}},
\end{aligned}
$$

$$\begin{aligned}\mathrm{Error}_{3,\mathrm{h}} =&\mathbb{E}[h^{(1)}(X_T^P)(X_T-\sum_{j=0}^{2}\frac{X_{j,T}}{j!})]+\frac{1}{2}\mathbb{E}[h^{(2)}(X_T^P)(X_T-X_T^P-X_{1,T})(X_T-X_T^P+X_{1,T})]\\&+\mathbb{E}[(X_T-X_T^P)^3\int_0^1 h^{(3)}(X_T^P+\eta(X_T-X_T^P))\frac{(1-\eta)^2}{2}\mathrm{d}\eta],\end{aligned}$$

with the convention $X_T^P=X_T^0=X_{0,T}$. Then we transform the terms $\mathbb{E}[h^{(1)}(X_T^P)X_{1,T}]$, $\mathbb{E}[h^{(1)}(X_T^P)\frac{X_{2,T}}{2}]$ and $\frac{1}{2}\mathbb{E}[h^{(2)}(X_T^P)X_{1,T}^2]$ into a weighted sum of sensitivities. To achieve this transformation, we apply a key lemma which proof is postponed to Appendix 5.6.2:

Lemma 5.2.3.1. *Let φ be a $C_0^\infty(\mathbb{R})$ function and $(f_t)_t$ be a measurable and bounded deterministic function. Let $N\geq 1$ be fixed, and consider measurable and bounded deterministic functions $t\mapsto l_{i,t}$ for $i=1,\dots,N$. Then, using the convention $\mathrm{d}W_t^0=\mathrm{d}t$, $\mathrm{d}W_t^1=\mathrm{d}W_t$ and $\mathrm{d}W_t^2=\mathrm{d}B_t$, for any $(I_1,\dots,I_N)\in\{0,1,2\}^N$ we have:*

$$\begin{aligned}&\mathbb{E}\Big(\varphi(\int_0^T f_t\mathrm{d}W_t)\int_0^T l_{N,t_N}\int_0^{t_N} l_{N-1,t_{N-1}}\dots\int_0^{t_2} l_{1,t_1}\mathrm{d}W_{t_1}^{I_1}\dots\mathrm{d}W_{t_{N-1}}^{I_{N-1}}\mathrm{d}W_{t_N}^{I_N}\Big)\\&=\omega(\widehat{l}_1,\dots,\widehat{l}_N)\mathcal{G}^{\varphi}_{\#\{k:I_k\neq 0\}}(\int_0^T f_t\mathrm{d}W_t),\end{aligned}\tag{5.18}$$

where $\widehat{l}_{k,t}:=\begin{cases} l_{k,t} & \text{if } I_k=0,\\ f_t l_{k,t} & \text{if } I_k=1,\\ f_t\rho_t l_{k,t} & \text{if } I_k=2.\end{cases}$

Details of the complete derivation of the corrective terms appearing in (5.9) are given in Appendix 5.6.2. Remind that these weighted sensitivities are well defined even if h is not smooth.
Last but not least, one has to estimate the residual term. In the smooth case, owing to (5.17), it is sufficient to estimate the L^p-norms of the residual processes $X_T-\sum_{j=0}^{i}\frac{X_{j,T}}{j!}$ for $i\in\{1,2\}$. Under the sole assumption that $h\in\mathrm{Lip_b}(\mathbb{R})$, the reflex is to regularize h and to try to employ some Malliavin integration by parts formula like in [Benhamou 2010b]. But a straightforward application of this methodology using the representation (5.17) fails because the random variable $X_T^P+\eta(X_T-X_T^P)$ does not belong to the space $\mathbb{D}^\infty$ for $\eta\neq 0$:

- The coefficient function of the square root model does not satisfy the standard assumptions. Malliavin differentiability is studied by hand in [Alòs 2008] up to the second order.
- There are moments explosion for processes having a stochastic volatility part and a local volatility function at least linear. See for instance the Heston model in [Andersen 2006].

To overcome this difficulty, the trick is to replace X_T by the smooth random variable (in Malliavin sense) $X_T^P+X_{1,T}+\frac{X_{2,T}}{2}$ close to X_T in L^p. Considering a regularization h_δ of h (which will be specified in (5.37)), we can write:

$$\begin{aligned}\mathbb{E}[h_\delta(X_T)] =&\mathbb{E}[h_\delta(X_T^P+X_{1,T}+\frac{X_{2,T}}{2})]+\mathbb{E}[(X_T-\sum_{j=0}^{2}\frac{X_{j,T}}{j!})\int_0^1 h_\delta^{(1)}((1-\eta)\sum_{j=0}^{2}\frac{X_{j,T}}{j!}+\eta X_T)\mathrm{d}\eta]\\=&\mathbb{E}[h_\delta(X_T^P)]+\mathbb{E}[h_\delta^{(1)}(X_T^P)(X_{1,T}+\frac{X_{2,T}}{2})]+\frac{1}{2}\mathbb{E}[h_\delta^{(2)}(X_T^P)X_{1,T}^2]+\mathrm{Error}_{3,\mathrm{h}_\delta},\end{aligned}\tag{5.19}$$

$$\mathrm{Error}_{3,\mathrm{h}_\delta} =\mathbb{E}[(X_T-\sum_{j=0}^{2}\frac{X_{j,T}}{j!})\int_0^1 h_\delta^{(1)}((1-\eta)\sum_{j=0}^{2}\frac{X_{j,T}}{j!}+\eta X_T)\mathrm{d}\eta]+\frac{1}{2}\mathbb{E}[h_\delta^{(2)}(X_T^P)(X_{1,T}X_{2,T}+\frac{X_{2,T}^2}{4})]$$

$$+\mathbb{E}[(X_{1,T}+\frac{X_{2,T}}{2})^3\int_0^1\frac{(1-\eta)^2}{2}h_\delta^{(3)}(X_T^P+\eta(X_{1,T}+\frac{X_{2,T}}{2}))\mathrm{d}\eta]. \tag{5.20}$$

As h is supposed Lipschitz bounded, the first term of (5.20) which involves only the first derivative of h_δ can be handled without Malliavin calculus. The two last terms of (5.20) contain higher derivatives of h_δ with the random variables X_T^P, $X_{1,T}$ and $X_{2,T}$ belonging to $\mathbb{D}^\infty$ but $X_T^P+\eta(X_{1,T}+\frac{X_{2,T}}{2})$ suffers from degeneracy (in the Malliavin sense) for $\eta\neq 0$. To fix this last problem, we use a standard Malliavin Calculus routine which consists in adding a small noise perturbation (see for instance [Gobet 2005] or [Gobet 2012a]). A nice feature of our methodology is that the regularization of h is done in such a way that there is no loss of accuracy.
The complete analyse of the error is given in the following subsection.

5.3 Error analysis

We establish the estimate (5.10) in several steps:

1. L^p-norms estimates of the residual processes,
2. small noise perturbation to smooth the function h,
3. careful use of Malliavin integration by parts formulas to achieve the proof.

5.3.1 Approximation of X, V, Λ and error estimates

Approximation of V, Λ and error estimates

Definition 5.3.1.1. *Assume* (P)*. We introduce for* $i\in\{0,1,2\}$ *the* Λ*-residual processes defined by*

$$(R^\Lambda_{i,t}=\Lambda_t-\sum_{j=0}^{i}\frac{\Lambda_{j,t}}{j!})_{t\in[0,T]}$$

where by convention $\Lambda_{0,t}=\lambda_t$ *and the corrective processes* $\big((\Lambda_{j,t})_{t\in[0,T]}\big)_{j\in\{1,2\}}$ *are defined in* (5.13)-(5.16)*. Replacing* Λ *by* V*, we define similarly the* V*-residual processes using the notation* R^V*.*

Proposition 5.3.1.1. *Assume* (P)*. Then for any* $p\geq 1$*, we have:*

$$\sqrt{v_0}\leq\lambda_{\inf}\leq\lambda_{\sup}\leq\sqrt{v_0+T\alpha_{\sup}}, \tag{5.21}$$

$$\sup_{t\in[0,T]}\|\Lambda_{i,t}\|_p\leq_c(\xi_{\sup}\sqrt{T})^i,\ \forall i\in\{1,2\}, \tag{5.22}$$

$$\sup_{t\in[0,T]}\|R^\Lambda_{i,t}\|_p\leq_c(\xi_{\sup}\sqrt{T})^{i+1},\ \forall i\in\{0,1,2\}. \tag{5.23}$$

Proof. (5.21) is obvious in view of (5.8). The proofs of (5.22) and (5.23) can be found in [Benhamou 2010b, Propositions 4.6, 4.7 and 4.8] replacing in the original paper κ by zero and $\kappa\theta_t$ by α_t. □

Corollary 5.3.1.1. *Assume* (P)*. Then one has for any* $p\geq 1$*:*

$$v_0\leq v_{\inf}\leq v_{\sup}\leq v_0+T\alpha_{\sup}, \tag{5.24}$$

$$\sup_{t\in[0,T]}\|V_t\|_p\leq_c 1+v_0, \tag{5.25}$$

$$\sup_{t\in[0,T]} \|V_{i,t}\|_p \leq_c (\xi_{\sup}\sqrt{T})^i,\ \forall i\in\{1,2\}, \tag{5.26}$$

$$\sup_{t\in[0,T]} \|R^V_{i,t}\|_p \leq_c (\xi_{\sup}\sqrt{T})^{i+1},\ \forall i\in\{0,1,2\}. \tag{5.27}$$

Proof. The proof of (5.24) and (5.25) are easy. (5.26) are obtained readily with (5.12), (5.15) and (5.22). Proofs of (5.27) are available in [Benhamou 2010b, Corollary 4.9] replacing in the original paper κ by zero and $\kappa\theta_t$ by α_t. □

Approximation of X and error estimates.

Definition 5.3.1.2. *Assume $(\mathcal{H}^{\sigma}_{x_0})$. We introduce for $i\in\{0,1,2\}$ the X-residual processes defined by*

$$(R^X_{i,t} = X_t - \sum_{j=0}^{i}\frac{X_{j,t}}{j!})_{t\in[0,T]}$$

where by convention $X_{0,t} = X^0_t = X^P_t$ and the corrective processes $\left((X_{j,t})_{t\in[0,T]}\right)_{j\in\{1,2\}}$ are defined in (5.11)-(5.14). When writing a Taylor expansion of $\sigma_t(.)$ at $x = X_t$ around $x = x_0$, we denote by $R_{n,\sigma}(X_t)$ the n^{th} Taylor residual:

$$R_{n,\sigma}(X_t) = \sigma_t(X_t) - \sum_{i=0}^{n}\frac{(X_t - x_0)^i}{i!}\sigma^{(i)}_t. \tag{5.28}$$

Replacing σ by σ^2, we use the similar notation $R_{n,\sigma^2}(X_t)$.

Standard computations involving Burkholder-Davis-Gundy and Hölder inequalities yield:

$$\begin{aligned}\|X_t - x_0\|^p_p \leq_c & t^{\frac{p}{2}-1}\int_0^t \|\sigma_s(X_s)\sqrt{V_s}\|^p_p \mathrm{d}s + t^{p-1}\int_0^t \|\sigma^2_s(X_s)V_s\|^p_p \mathrm{d}s\\ \leq_c & t^{\frac{p}{2}-1}|\sigma|^p_\infty\int_0^t \mathbb{E}[V_s^{p/2}]\mathrm{d}s + t^{p-1}|\sigma|^{2p}_\infty\int_0^t \mathbb{E}[V^p_s]\mathrm{d}s \leq_c (|\sigma|_\infty\sqrt{T})^p,\end{aligned} \tag{5.29}$$

for any $p\geq 2$, where we have applied the estimate (5.25) at the last line. We now intend to handle X-residual processes, and the next results are intermediate steps. In the next Lemma, we provide L^p-estimates of $X^P_t - x_0$, $X_{1,t}$ and $X_{2,t}$.

Lemma 5.3.1.1. *Assume $(\mathcal{H}^{\sigma}_{x_0})$ and (P). For any $p\geq 1$:*

$$\sup_{t\in[0,T]}\|X^P_t - x_0\|_p \leq_c |\sigma|_\infty\sqrt{T}, \tag{5.30}$$

$$\sup_{t\in[0,T]}\|X_{i,t}\|_p \leq_c |\sigma|_\infty[\xi^i_{\sup} + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma)+\xi_{\sup})^{i-1}]T^{\frac{i+1}{2}},\ \forall i\in\{1,2\}. \tag{5.31}$$

Proof. (5.30) is similar to (5.29). For (5.31) $i = 1$: starting from (5.11), the same computations as before give:

$$\|X_{1,t}\|_p \leq_c \mathcal{M}_1(\sigma)\sqrt{T}(1+\mathcal{M}_0(\sigma)\sqrt{T})\sup_{t\in[0,T]}\|X^P_t - x_0\|_p + |\sigma|_\infty\sqrt{T}(1+\mathcal{M}_0(\sigma)\sqrt{T})\sup_{t\in[0,T]}\|V_{1,t}\|_p.$$

We conclude using (5.30) and (5.26). For (5.31) $i = 2$, one has from (5.14):

$$\begin{aligned}\|X_{2,t}\|_p \leq_c & \mathcal{M}_1(\sigma)\sqrt{T}(1+\mathcal{M}_0(\sigma)\sqrt{T})(\sup_{t\in[0,T]}\|(X^P_t - x_0)^2\|_p + \sup_{t\in[0,T]}\|X_{1,t}\|_p)\\ & + |\sigma|_\infty\sqrt{T}(1+\mathcal{M}_0(\sigma)\sqrt{T})\sup_{t\in[0,T]}\|V_{2,t}\|_p + |\sigma|_\infty\sqrt{T}\sup_{t\in[0,T]}\|V^2_{1,t}\|_p\\ & + \mathcal{M}_1(\sigma)\sqrt{T}(1+\mathcal{M}_0(\sigma)\sqrt{T})\sup_{t\in[0,T]}\|X^P_t - x_0\|_{2p}\sup_{t\in[0,T]}\|V_{1,t}\|_{2p}.\end{aligned}$$

We conclude using (5.30), (5.31) $i = 1$ and (5.26). □

We give in the following Lemma the explicit equations solved by the X-residual processes:

Lemma 5.3.1.2. *Assume* ($\mathcal{H}^{\sigma}_{x_0}$) *and* ($P$). *One has:*

$$\mathrm{d}R^X_{0,t} =[\lambda_t R_{0,\sigma}(X_t)+\sigma_t(X_t)R^{\Lambda}_{0,t}]\mathrm{d}W_t-\frac{1}{2}[v_t R_{0,\sigma^2}(X_t)+\sigma^2_t(X_t)R^V_{0,t}]\mathrm{d}t,\ R^X_{0,0}=0, \tag{5.32}$$

$$\begin{aligned}\mathrm{d}R^X_{1,t} =&[\lambda_t R_{1,\sigma}(X_t)+\Lambda_{1,t}R_{0,\sigma}(X_t)+\sigma_t(X_t)R^{\Lambda}_{1,t}+\lambda_t\sigma^{(1)}_t R^X_{0,t}]\mathrm{d}W_t\\ &-\frac{1}{2}[v_t R_{1,\sigma^2}(X_t)+V_{1,t}R_{0,\sigma^2}(X_t)+\sigma^2_t(X_t)R^V_{1,t}+2v_t\sigma_t\sigma^{(1)}_t R^X_{0,t}]\mathrm{d}t,\ R^X_{1,0}=0,\end{aligned} \tag{5.33}$$

$$\begin{aligned}\mathrm{d}R^X_{2,t} =&[\lambda_t R_{2,\sigma}(X_t)+\Lambda_{1,t}R_{1,\sigma}(X_t)+\frac{\Lambda_{2,t}}{2}R_{0,\sigma}(X_t)+\sigma_t(X_t)R^{\Lambda}_{2,t}+\lambda_t\sigma^{(1)}_t R^X_{1,t}+\lambda_t\frac{\sigma^{(2)}_t}{2}R^X_{0,t}(X_t+X^P_t)\\ &+\Lambda_{1,t}\sigma^{(1)}_t R^X_{0,t}]\mathrm{d}W_t-\frac{1}{2}[v_t R_{2,\sigma^2}(X_t)+V_{1,t}R_{1,\sigma^2}(X_t)+\frac{V_{2,t}}{2}R_{0,\sigma^2}(X_t)+\sigma^2_t(X_t)R^V_{2,t}+2v_t\sigma_t\sigma^{(1)}_t R^X_{1,t}\\ &+v_t((\sigma^{(1)}_t)^2+\sigma^{(2)}_t\sigma_t)R^X_{0,t}(X_t+X^P_t)+2V_{1,t}\sigma_t\sigma^{(1)}_t R^X_{0,t}]\mathrm{d}t,\ R^X_{2,0}=0.\end{aligned} \tag{5.34}$$

Proof. The verification of these identities is tedious but without mathematical difficulties. For convenience, we detail some computations. To obtain (5.32), start from (5.1) and (5.4) and write:

$$\begin{aligned}\mathrm{d}R^X_{0,t} =&[\sigma_t(X_t)\Lambda_t-\sigma_t\lambda_t]\mathrm{d}W_t-\frac{1}{2}[\sigma^2_t(X_t)V_t-\sigma^2_t v_t]\mathrm{d}t\\ =&[\lambda_t(\sigma_t(X_t)-\sigma_t)+\sigma_t(X_t)(\Lambda_t-\lambda_t)]\mathrm{d}W_t-\frac{1}{2}[v_t(\sigma^2_t(X_t)-\sigma^2_t)+\sigma^2_t(X_t)(V_t-v_t)]\mathrm{d}t\\ =&[\lambda_t R_{0,\sigma}(X_t)+\sigma_t(X_t)R^{\Lambda}_{0,t}]\mathrm{d}W_t-\frac{1}{2}[v_t R_{0,\sigma^2}(X_t)+\sigma^2_t(X_t)R^V_{0,t}]\mathrm{d}t.\end{aligned}$$

Similarly for (5.33), using (5.32) and (5.11), we get:

$$\begin{aligned}\mathrm{d}R^X_{1,t} =&\mathrm{d}R^X_{0,t}-\mathrm{d}X_{1,t}\\ =&[\lambda_t R_{0,\sigma}(X_t)+\sigma_t(X_t)R^{\Lambda}_{0,t}]\mathrm{d}W_t-\frac{1}{2}[v_t R_{0,\sigma^2}(X_t)+\sigma^2_t(X_t)R^V_{0,t}]\mathrm{d}t\\ &-[(X^P_t-x_0)\sigma^{(1)}_t\lambda_t+\Lambda_{1,t}\sigma_t](\mathrm{d}W_t-\sigma_t\lambda_t\mathrm{d}t)\\ =&[\lambda_t R_{0,\sigma}(X_t)+\Lambda_{1,t}R_{0,\sigma}(X_t)+\sigma_t(X_t)R^{\Lambda}_{1,t}]\mathrm{d}W_t-\frac{1}{2}[v_t R_{0,\sigma^2}(X_t)+V_{1,t}R_{0,\sigma^2}(X_t)+\sigma^2_t(X_t)R^V_{1,t}]\mathrm{d}t\\ &-(X^P_t-x_0)\sigma^{(1)}_t\lambda_t(\mathrm{d}W_t-\sigma_t\lambda_t\mathrm{d}t)\\ =&[\lambda_t R_{1,\sigma}(X_t)+\Lambda_{1,t}R_{0,\sigma}(X_t)+\sigma_t(X_t)R^{\Lambda}_{1,t}+\lambda_t\sigma^{(1)}_t R^X_{0,t}]\mathrm{d}W_t\\ &-\frac{1}{2}[v_t R_{1,\sigma^2}(X_t)+V_{1,t}R_{0,\sigma^2}(X_t)+\sigma^2_t(X_t)R^V_{1,t}+2v_t\sigma_t\sigma^{(1)}_t R^X_{0,t}]\mathrm{d}t.\end{aligned}$$

Now consider (5.34). Start from (5.33)-(5.14) and write:

$$\begin{aligned}\mathrm{d}R^X_{2,t} =&\mathrm{d}R^X_{1,t}-\frac{1}{2}\mathrm{d}X_{2,t}\\ =&[\lambda_t R_{1,\sigma}(X_t)+\Lambda_{1,t}R_{0,\sigma}(X_t)+\sigma_t(X_t)R^{\Lambda}_{1,t}+\lambda_t\sigma^{(1)}_t R^X_{0,t}]\mathrm{d}W_t\\ &-\frac{1}{2}[v_t R_{1,\sigma^2}(X_t)+V_{1,t}R_{0,\sigma^2}(X_t)+\sigma^2_t(X_t)R^V_{1,t}+2v_t\sigma_t\sigma^{(1)}_t R^X_{0,t}]\mathrm{d}t\\ &-\frac{1}{2}\{\lambda_t[(X^P_t-x_0)^2\sigma^{(2)}_t+2\sigma^{(1)}_t X_{1,t}]+2(X^P_t-x_0)\Lambda_{1,t}\sigma^{(1)}_t\}(\mathrm{d}W_t-\sigma_t\lambda_t\mathrm{d}t)\\ &-\frac{1}{2}\{\Lambda_{2,t}\sigma_t\mathrm{d}W_t-[(X^P_t-x_0)V_{1,t}\sigma^{(1)}_t\sigma_t+(X^P_t-x_0)^2(\sigma^{(1)}_t)^2v_t+\frac{V_{2,t}}{2}\sigma^2_t]\mathrm{d}t\}\\ =&[\lambda_t R_{1,\sigma}(X_t)+\Lambda_{1,t}R_{0,\sigma}(X_t)+\frac{\Lambda_{2,t}}{2}R_{0,\sigma}(X_t)+\sigma_t(X_t)R^{\Lambda}_{2,t}+\lambda_t\sigma^{(1)}_t R^X_{1,t}]\mathrm{d}W_t\end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{2}[v_t R_{1,\sigma^2}(X_t) + V_{1,t}R_{0,\sigma^2}(X_t) + \frac{V_{2,t}}{2}R_{0,\sigma^2}(X_t) + \sigma_t^2(X_t)R^V_{2,t} + 2v_t\sigma_t\sigma_t^{(1)}R^X_{1,t}]\mathrm{d}t \\
&-\frac{1}{2}\{\lambda_t(X_t^P - x_0)^2\sigma_t^{(2)} + 2(X_t^P - x_0)\Lambda_{1,t}\sigma_t^{(1)}\}(\mathrm{d}W_t - \sigma_t\lambda_t\mathrm{d}t) \\
&+\frac{1}{2}[(X_t^P - x_0)V_{1,t}\sigma_t^{(1)}\sigma_t + (X_t^P - x_0)^2(\sigma_t^{(1)})^2 v_t]\mathrm{d}t \\
=&[\lambda_t R_{2,\sigma}(X_t) + \Lambda_{1,t}R_{0,\sigma}(X_t) + \frac{\Lambda_{2,t}}{2}R_{0,\sigma}(X_t) + \sigma_t(X_t)R^\Lambda_{2,t} + \lambda_t\sigma_t^{(1)}R^X_{1,t} + \lambda_t\frac{\sigma_t^{(2)}}{2}R^X_{0,t}(X_t + X_t^P)]\mathrm{d}W_t \\
&-\frac{1}{2}[v_t R_{2,\sigma^2}(X_t) + V_{1,t}R_{0,\sigma^2}(X_t) + \frac{V_{2,t}}{2}R_{0,\sigma^2}(X_t) + \sigma_t^2(X_t)R^V_{2,t} + v_t((\sigma_t^{(1)})^2 + \sigma_t^{(2)}\sigma_t)R^X_{0,t}(X_t + X_t^P) \\
&+ 2v_t\sigma_t\sigma_t^{(1)}R^X_{1,t}]\mathrm{d}t - (X_t^P - x_0)\Lambda_{1,t}\sigma_t^{(1)}(\mathrm{d}W_t - \sigma_t\lambda_t\mathrm{d}t) + \frac{1}{2}(X_t^P - x_0)V_{1,t}\sigma_t^{(1)}\sigma_t\mathrm{d}t \\
=&[\lambda_t R_{2,\sigma}(X_t) + \Lambda_{1,t}R_{1,\sigma}(X_t) + \frac{\Lambda_{2,t}}{2}R_{0,\sigma}(X_t) + \sigma_t(X_t)R^\Lambda_{2,t} + \lambda_t\sigma_t^{(1)}R^X_{1,t} + \lambda_t\frac{\sigma_t^{(2)}}{2}R^X_{0,t}(X_t + X_t^P) \\
&+ \Lambda_{1,t}\sigma_t^{(1)}R^X_{0,t}]\mathrm{d}W_t - \frac{1}{2}[v_t R_{2,\sigma^2}(X_t) + V_{1,t}R_{1,\sigma^2}(X_t) + \frac{V_{2,t}}{2}R_{0,\sigma^2}(X_t) + \sigma_t^2(X_t)R^V_{2,t} + 2v_t\sigma_t\sigma_t^{(1)}R^X_{1,t} \\
&+ v_t((\sigma_t^{(1)})^2 + \sigma_t^{(2)}\sigma_t)R^X_{0,t}(X_t + X_t^P) + 2V_{1,t}\sigma_t\sigma_t^{(1)}R^X_{0,t}]\mathrm{d}t.
\end{aligned}
$$

□

An intermediate result is the estimates of $R_{n,\sigma}(X_t)$ and $R_{n,\sigma^2}(X_t)$. Assuming $(\mathcal{H}^\sigma_{x_0})$, from the Taylor-Lagrange inequality, we have $|R_{n,\sigma}(X_t)| \le_c |X_t - x_0|^{n+1}\mathcal{M}_1(\sigma)$ and $|R_{n,\sigma^2}(X_t)| \le_c |X_t - x_0|^{n+1}\mathcal{M}_0(\sigma)\mathcal{M}_1(\sigma)$. Combined with (5.29), this readily gives $\forall p \ge 2$ and $\forall j \in \{0,..,2\}$:

$$
\sup_{t\in[0,T]} \|R_{j,\sigma}(X_t)\|_p \le_c (|\sigma|_\infty \sqrt{T})^{j+1}\mathcal{M}_1(\sigma), \quad \sup_{t\in[0,T]} \|R_{j,\sigma^2}(X_t)\|_p \le_c (|\sigma|_\infty \sqrt{T})^{j+1}\mathcal{M}_0(\sigma)\mathcal{M}_1(\sigma). \quad (5.35)
$$

We now state the result related to the estimates of the residuals processes:

Proposition 5.3.1.2. *Assume that* $(\mathcal{H}^\sigma_{x_0})$ *and* (P) *hold. Then for any* $p \ge 1$*, we have:*

$$
\sup_{t\in[0,T]} \|R^X_{j,t}\|_p \le_c |\sigma|_\infty\{\xi_{\sup}^{j+1} + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^j\}T^{\frac{j}{2}+1}, \ \forall j \in \{0,1,2\}. \quad (5.36)
$$

Proof. We leverage the explicit equations solved by the residuals $(R^X_{j,t})_{t\in[0,T]}$ (see Lemma 5.3.1.2). We begin with $R^X_{0,t}$. Starting from (5.32) and using standard inequalities, it readily follows:

$$
\|R^X_{0,t}\|_p \le_c \sqrt{T}[\sup_{t\le T}\|R_{0,\sigma}(X_t)\|_p + |\sigma|_\infty \sup_{t\le T}\|R^\Lambda_{0,t}\|_p] + T[\sup_{t\le T}\|R_{0,\sigma^2}(X_t)\|_p + |\sigma|^2_\infty \sup_{t\le T}\|R^V_{0,t}\|_p].
$$

We conclude using (5.35)-(5.23)-(5.27). Similarly for $R^X_{1,t}$ given in (5.33), we obtain:

$$
\begin{aligned}
\|R^X_{1,t}\|_p \le_c \sqrt{T}\{&\sup_{t\le T}\|R_{1,\sigma}(X_t)\|_p + \sup_{t\le T}\|\Lambda_{1,t}R_{0,\sigma}(X_t)\|_p + |\sigma|_\infty \sup_{t\le T}\|R^\Lambda_{1,t}\|_p + \mathcal{M}_1(\sigma)\sup_{t\le T}\|R^X_{0,t}\|_p\} \\
&+ T\{\sup_{t\le T}\|R_{1,\sigma^2}(X_t)\|_p + \sup_{t\le T}\|V_{1,t}R_{0,\sigma^2}(X_t)\|_p + |\sigma|^2_\infty \sup_{t\le T}\|R^V_{1,t}\|_p + |\sigma|_\infty\mathcal{M}_1(\sigma)\sup_{t\le T}\|R^X_{0,t}\|_p\}.
\end{aligned}
$$

Then, plugging in the above upper bound the estimates (5.22)-(5.23)-(5.26)-(5.27)-(5.35)-(5.36) $i = 0$, we complete the proof of (5.36) for $i = 1$. Finally for $R^X_{2,t}$, starting from (5.34), we readily have:

$$
\begin{aligned}
\|R^X_{2,t}\|_p \le_c \sqrt{T}\{&\sup_{t\le T}\|R_{2,\sigma}(X_t)\|_p + \sup_{t\le T}\|\Lambda_{1,t}R_{1,\sigma}(X_t)\|_p + \sup_{t\le T}\|\Lambda_{2,t}R_{0,\sigma}(X_t)\|_p + |\sigma|_\infty \sup_{t\le T}\|R^\Lambda_{2,t}\|_p \\
&+ \mathcal{M}_1(\sigma)\sup_{t\le T}\|R^X_{1,t}\|_p + \mathcal{M}_1(\sigma)\sup_{t\le T}\|R^X_{0,t}(X_t + X_t^P)\|_p + \mathcal{M}_1(\sigma)\sup_{t\le T}\|\Lambda_{1,t}R^X_{0,t}\|_p\}
\end{aligned}
$$

$$+T\{\sup_{t\leq T}\|R_{2,\sigma^2}(X_t)\|_p+\sup_{t\leq T}\|V_{1,t}R_{1,\sigma^2}(X_t)\|_p+\sup_{t\leq T}\|V_{2,t}R_{0,\sigma^2}(X_t)\|_p+|\sigma|_\infty^2\sup_{t\leq T}\|R^V_{2,t}\|_p$$
$$+|\sigma|_\infty\mathcal{M}_1(\sigma)\sup_{t\leq T}\|R^X_{1,t}\|_p+\mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)\sup_{t\leq T}\|R^X_{0,t}(X_t+X^P_t)\|_p+|\sigma|_\infty\mathcal{M}_1(\sigma)\sup_{t\leq T}\|V_{1,t}R^X_{0,t}\|_p\}.$$

The proof is completed as before using (5.22)-(5.23)-(5.26)-(5.27)-(5.35)-(5.36) $i=0$ and 1. □

5.3.2 Regularization of the function h by adding a small noise perturbation

To overcome some problems of degeneracy in the Malliavin sense and to compensate the lack of smoothness of the payoff function h, we introduce an extra scalar Brownian motion $\overline{W}$ independent of W and B even if it means enlarge the initial filtration and the initial sigma field. Then we define:

$$h_\delta(x)=\mathbb{E}[h(x+\delta\overline{W}_T)]. \tag{5.37}$$

for a small parameter $\delta>0$. Clearly the function h_δ is of class $C^\infty(\mathbb{R})$ thanks to the smoothness of the Gaussian density and remarkably we can notice that using a conditioning:

$$h_\delta(x)=\mathbb{E}[h_{\delta/\sqrt{2}}(x+\delta\overline{W}_{\frac{T}{2}})]. \tag{5.38}$$

In addition $h\in\mathrm{Lip_b}(\mathbb{R})\Rightarrow \mathrm{h}_\delta\in\mathrm{Lip_b}(\mathbb{R})$ with $C_{h_\delta}\leq C_h$ and $L_{h_\delta}\leq L_h$. The next Lemma estimates the error in terms of δ induced by considering h_δ instead of h in the calculus of expectations and sensitivities which appear in the Theorem 5.2.2.1.

Lemma 5.3.2.1. *Let $\delta>0$. Assume that $h\in\mathrm{Lip_b}(\mathbb{R})$ and that $(\mathcal{H}^\sigma_{x_0})$ is satisfied. Then we have:*

$$\left|\mathbb{E}[h(X_T)]-\mathbb{E}[h_\delta(X_T)]\right|+\left|\mathbb{E}[h(X^P_T)]-\mathbb{E}[h_\delta(X^P_T)]\right|\leq_c L_h\delta\sqrt{T},$$
$$\left|\partial^i_{x^i}\mathbb{E}[h(X^P_T+x)]|_{x=0}-\partial^i_{x^i}\mathbb{E}[h_\delta(X^P_T+x)]|_{x=0}\right|\leq_c L_h\frac{\delta\sqrt{T}}{(\int_0^T\sigma_t^2v_t\mathrm{d}t)^{i/2}},\ \forall i\geq 1.$$

Proof. The first estimate is obvious using the lipschitzianity of h and classical estimates for the auxiliary Brownian motion. For the second write:

$$\mathbb{E}[h_\delta(X^P_T+x)]=\int_{\mathbb{R}}\mathbb{E}[h(y-\frac{\int_0^T\sigma_t^2v_t\mathrm{d}t}{2}+\delta\overline{W}_T)]\frac{e^{-\frac{(y-x)^2}{2\int_0^T\sigma_t^2v_t\mathrm{d}t}}}{\sqrt{2\pi\int_0^T\sigma_t^2v_t\mathrm{d}t}}\mathrm{d}y,$$

to obtain:

$$\partial^i_{x^i}\mathbb{E}[h(X^P_T+x)]|_{x=0}-\partial^i_{x^i}\mathbb{E}[h_\delta(X^P_T+x)]|_{x=0}$$
$$=\int_{\mathbb{R}}\mathbb{E}[h(y-\frac{\int_0^T\sigma_t^2v_t\mathrm{d}t}{2}+\delta\overline{W}_T)-h(y-\frac{\int_0^T\sigma_t^2v_t\mathrm{d}t}{2})]\partial^i_{x^i}\Big\{\frac{e^{-\frac{(y-x)^2}{2\int_0^T\sigma_t^2v_t\mathrm{d}t}}}{\sqrt{2\pi\int_0^T\sigma_t^2v_t\mathrm{d}t}}\Big\}|_{x=0}\mathrm{d}y.$$

Then we complete the proof using again the lipschitzianity of h and standard upper bounds for the derivatives of the Gaussian density. □

In view of the magnitude of the coefficients $C^l_{i,T}$, $C^s_{i,T}$ and $C^{ls}_{i,T}$ defined in Theorem 5.2.2.1, applying Lemma 5.3.2.1, we readily obtain :

$$\left|\mathrm{Error}_{3,\mathrm{h}}\right|=\left|\mathbb{E}[h(X_T)]-\mathbb{E}[h(X^P_T)]-\sum_{i=1}^{6}\eta_{i,T}\mathcal{G}^h_i(X^P_T)\right|$$

$$
\begin{aligned}
&\leq \left|\mathbb{E}[h(X_T)] - \mathbb{E}[h_\delta(X_T)]\right| + \left|\mathbb{E}[h_\delta(X_T^P)] - \mathbb{E}[h(X_T^P)]\right| \\
&\quad + \sum_{i=1}^{6} |\eta_{i,T}| \left|\mathcal{G}_i^{h_\delta}(X_T^P) - \mathcal{G}_i^{h_\delta}(X_T^P)\right| + \left|\mathrm{Error}_{3,\mathrm{h}_\delta}\right| \\
&\leq_c L_h\, \delta \sqrt{T} + \left|\mathrm{Error}_{3,\mathrm{h}_\delta}\right|.
\end{aligned}
$$

Assume now without loss of generality that $\mathcal{M}_1(\sigma) + \xi_{\sup} \neq 0$. We prove the estimate (5.10) if we choose as value for δ:

$$
\delta = |\sigma|_\infty [\xi_{\sup}^3 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2]\, T^{\frac{3}{2}} \tag{5.39}
$$

and establish that:

$$
|\mathrm{Error}_{3,\mathrm{h}_\delta}| \leq_c L_h |\sigma|_\infty [\xi_{\sup}^3 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2] T^2, \tag{5.40}
$$

This is the purpose of the next Subsection.

5.3.3 Malliavin integration by parts formula and proof of estimate (5.40)

We write $W_t = \int_0^t \rho_s \mathrm{d}B_s + \int_0^t \sqrt{1-\rho_s^2} \mathrm{d}B_s^\perp$ where $(B_t^\perp)_{0 \leq t \leq T}$ is a Brownian motion independent of B_t and we consider the calculus of stochastic variations w.r.t. the three-dimensional Brownian motion $(B, B^\perp, \overline{W})$, the Malliavin derivative operator w.r.t. $B, B^\perp$ and $\overline{W}$ being respectively denoted by $(D_t^1(.))_{t\in[0,T]}, (D_t^2(.))_{t\in[0,T]}$ and $(D_t^3(.))_{t\in[0,T]}$. For the second derivatives, we use the obvious notation $(D^{i,j}(.))_{s,t\in[0,T]}$ for $i, j \in \{1,2,3\}$ and so on for the higher derivatives. We freely adopt the notations of [Nualart 2006] for the Sobolev space $\mathbb{D}^{k,p}$ associated to the norm $\|.\|_{k,p}$.
In the following Lemma, we provide estimates of the Malliavin derivatives of $(X_t^P)_{t\in[0,T]}$, $(X_{1,t})_{t\in[0,T]}$ and $(X_{2,t})_{t\in[0,T]}$.

Lemma 5.3.3.1. *Assume that $(\mathcal{H}_{x_0}^{\sigma})$ and (P) hold. Then, $\forall t \in [0,T]$, $X_t^P, X_{1,t}$, $X_{2,t}$, $V_{1,t}$ and $V_{2,t} \in \mathbb{D}^{3,\infty}$. Moreover, we have the following estimates, $\forall p \geq 1$, uniformly in $q, r, s, t \in [0,T]$:*

$$
\begin{aligned}
&\|D_s^1 X_t^P\|_p + \|D_s^2 X_t^P\|_p \leq_c |\sigma|_\infty, && (5.41)\\
&\|D_s^1 \Lambda_{n,t}\|_p + \|D_s^1 V_{n,t}\|_p \leq_c \xi_{\sup}^n T^{\frac{n-1}{2}},\ \forall n \in \{1,2\}, && (5.42)\\
&\|D_s^1 X_{n,t}\|_p + \|D_s^2 X_{n,t}\|_p \leq_c |\sigma|_\infty [\xi_{\sup}^n + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^{n-1}] T^{\frac{n}{2}},\ \forall n \in \{1,2\} && (5.43)\\
&\|D_{r,s}^{1,1} \Lambda_{2,t}\|_p + \|D_{r,s}^{1,1} V_{2,t}\|_p \leq_c \xi_{\sup}^2, && (5.44)\\
&\sum_{i,j\in\{1,2\}} \|D_{r,s}^{i,j} X_{n,t}\|_p \leq_c |\sigma|_\infty [\xi_{\sup}^n + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^{n-1}] T^{\frac{n-1}{2}},\ \forall n \in \{1,2\}, && (5.45)\\
&\sum_{i,j,k\in\{1,2\}} \|D_{q,r,s}^{i,j,k} X_{2,t}\|_p \leq_c |\sigma|_\infty [\xi_{\sup}^2 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})]. && (5.46)
\end{aligned}
$$

Proof. It is obvious that all the variables are in $\mathbb{D}^{3,\infty}$ because by construction we take face to multiple Wiener integrals and temporal integrals of multiple Wiener integrals (see (5.4)-(5.11)-(5.14)-(5.12)-(5.15)). Then the calculus of the derivatives and the L^p-estimates are classical so we skip details. In particular, all the derivatives w.r.t. the third Brownian motion $\overline{W}$ are null as well as the derivatives of $V_{1,t}$, $\Lambda_{1,t}$, $V_{2,t}$ and $\Lambda_{2,t}$ w.r.t. $B^\perp$. □

We now state the result related to integration by parts formulas which is proven later:

Lemma 5.3.3.2. *Assume* $(\mathcal{H}_{x_0}^{\sigma})$ *and* (P)*. For any* $\eta \in [0,1]$*, we define the random variable* $G_{\delta}^{\eta} = X_T^P + \eta(X_{1,T} + \frac{X_{2,T}}{2}) + \delta\overline{W}_{T/2}$*. Then for any* Y *in* $\mathbb{D}^{1,\infty}$*, there exist random variables* $Y_{2,\eta}$ *and* $Y_{3,\eta}$ *in* $\cap_{p\geq 1} L^p$ *such that* $\forall i \in \{2,3\}$*:*

$$\mathbb{E}[Y h_{\delta/\sqrt{2}}^{(i)}(G_{\delta}^{\eta})] = \mathbb{E}[Y_{i,\eta} h_{\delta/\sqrt{2}}^{(1)}(G_{\delta}^{\eta})], \tag{5.47}$$

where for any $p \geq 1$ *and any* $i \in \{2,3\}$*:*

$$\sup_{\eta\in[0,1]} \|Y_{i,\eta}\|_p \leq_c \|Y\|_{i-1,p+\frac{1}{2}} (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-\frac{(i-1)}{2}}. \tag{5.48}$$

We are now in position to achieve the proof of (5.40). Consider $\mathrm{Error}_{3,\mathrm{h}_\delta}$ explicitly written in (5.20). The first term of (5.20) is handled easily using (5.36) $i = 3$. For the second term of (5.20), using (5.38), applying the Lemma 5.3.3.2 with $Y = X_{1,T}X_{2,T} + \frac{X_{2,T}^2}{4}$ and using (5.31)-(5.43), we obtain:

$$\begin{aligned}
&\left|\mathbb{E}[h_\delta^{(2)}(X_T^P)(X_{1,T}X_{2,T} + \frac{X_{2,T}^2}{4})]\right| = \left|\mathbb{E}[h_{\delta/\sqrt{2}}^{(2)}(G_\delta^0)(X_{1,T}X_{2,T} + \frac{X_{2,T}^2}{4})]\right| = \left|\mathbb{E}[h_{\delta/\sqrt{2}}^{(1)}(G_\delta^0)Y_{1,0}]\right| \\
\leq_c & L_h \|Y\|_{1,2} (\int_0^T v_t\sigma_t^2 \mathrm{d}t)^{-1/2} \leq_c L_h |\sigma|_\infty [\xi_{\sup}^3 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2] T^2.
\end{aligned}$$

The last term of (5.20) is handled similarly; apply the Lemma 5.3.3.2 with $Y = (X_{1,T} + \frac{X_{2,T}}{2})^3$ and use (5.31)-(5.43)-(5.45)-(5.46) to obtain the announced result. To complete the proof, it remains to establish the Lemma 5.3.3.2. This is done in the following Subsection.

5.3.4 Proof of Lemma 5.3.3.2

X_T^P is a non-degenerate random variable with Malliavin covariance matrix equal to $\int_0^T \sigma_t^2 v_t \mathrm{d}t > 0$ thanks to $(\mathcal{H}_{x_0}^{\sigma})$ but $G^\eta = X_T^P + \eta(X_{1,T} + \frac{X_{2,T}}{2})$ is degenerate for $\eta > 0$ and this is the second reason to have introduced the small perturbation $\delta\overline{W}_{T/2}$. Consider the random variable $G_\delta^\eta = G^\eta + \delta\overline{W}_{T/2}$ defined in Lemma 5.3.3.2: clearly it belongs to $\mathbb{D}^{3,\infty}$ with Malliavin covariance matrix obviously invertible:

$$\gamma_{G_\delta^\eta} = \sum_{i=1}^{2} \int_0^T (D_t^i G^\eta)^2 \mathrm{d}t + \delta^2 \frac{T}{2} = \gamma_{G^\eta} + \delta^2\frac{T}{2} > \delta^2\frac{T}{2} > 0.$$

Then with (5.39)-(5.41)-(5.43)-(5.45)-(5.46) it readily comes for any $i \in \{1,2\}$ and any $p \geq 1$:

$$\|(D^1 G_\delta^\eta, D^2 G_\delta^\eta, D^3 G_\delta^\eta)\|_{i,p} \leq_c |\sigma|_\infty \sqrt{T}. \tag{5.49}$$

Hence, applying [Nualart 2006, Proposition 1.5.6 and Proposition 2.1.4] and using (5.49) we get the existence of $Y_{2,\eta}$ and $Y_{3,\eta}$ such that for any $i \in \{2,3\}$ and any $p \geq 1$:

$$\begin{aligned}
\|Y_{i,\eta}\|_p \leq_c & \|Y\|_{i-1,p+\frac{1}{2}} \|DG_\delta^\eta\|_{i-1,2^{i-1}p(2p+1)}^{i-1} \|\gamma_{G_\delta^\eta}^{-1}\|_{i-1,2^{i-1}p(2p+1)}^{i-1} \\
\leq_c & \|Y\|_{i-1,p+\frac{1}{2}} (|\sigma|_\infty \sqrt{T})^{i-1} \|\gamma_{G_\delta^\eta}^{-1}\|_{i-1,2^{i-1}p(2p+1)}^{i-1}.
\end{aligned} \tag{5.50}$$

It remains to finely estimate the norms related to the inverse of the Malliavin covariance matrix $\gamma_{G_\delta^\eta}$. First notice that using the definitions of G^η and γ_{G^η} we have:

$$\gamma_{G^\eta} = \gamma_{X_T^P} + \gamma_{\eta(X_{1,T} + \frac{X_{2,T}}{2})} + 2\eta \sum_{i=1}^{2} \int_0^T D_t^i(X_T^P) D_t^i(X_{1,T} + \frac{X_{2,T}}{2}) \mathrm{d}t$$

$$= \int_0^T \sigma_t^2 v_t \mathrm{d}t + \eta^2 \sum_{i=1}^2 \int_0^T [D_t^i(X_{1,T} + \frac{X_{2,T}}{2})]^2 \mathrm{d}t + 2\eta \sum_{i=1}^2 \int_0^T \sigma_t v_t D_t^i(X_{1,T} + \frac{X_{2,T}}{2}) \mathrm{d}t$$

and hence estimates (5.41)-(5.43), $(\mathcal{H}_{x_0}^{\sigma})$ and (P) easily yield to:

$$\sup_{\eta \in [0,1]} \left\|\gamma_{G^\eta} - \int_0^T \sigma_t^2 v_t \mathrm{d}t\right\|_p \leq_c |\sigma|_\infty^2 (\xi_{\sup} + \mathcal{M}_1(\sigma)) T^{\frac{3}{2}}, \tag{5.51}$$

for any $p \geq 1$. This intermediate estimate allows to prove the next Lemma:

Lemma 5.3.4.1. *Assume $(\mathcal{H}_{x_0}^{\sigma})$ and (P). Then $(\gamma_{G_\delta^\eta})^{-1} \in \mathbb{D}^{2,\infty}$ and we have for any $p \geq 1$:*

$$\sup_{\eta \in [0,1]} \|(\gamma_{G_\delta^\eta})^{-1}\|_p \leq_c (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-1}, \tag{5.52}$$

$$\sup_{t \in [0,T],\ \eta \in [0,1]} \sum_{i \in \{1,2\}} \|D_t^i(\gamma_{G_\delta^\eta})^{-1}\|_p \leq_c (\mathcal{M}_1(\sigma) + \xi_{\sup})(\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-1}, \tag{5.53}$$

$$\sup_{s,t \in [0,T],\ \eta \in [0,1]} \sum_{i,j \in \{1,2\}} \|D_{s,t}^{i,j}(\gamma_{G_\delta^\eta})^{-1}\|_p \leq_c [\xi_{\sup}^2 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})](\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-1}. \tag{5.54}$$

Proof. For the sake of brevity, we only prove (5.52) and (5.53) because there is no extra difficulties for (5.54). For (5.52), we have for any $p \geq 1$ and $q \geq 1$:

$$\begin{aligned}
\mathbb{E}[(\gamma_{G_\delta^\eta})^{-p}] &= \mathbb{E}[(\gamma_{G_\delta^\eta})^{-p} \mathbb{1}_{\gamma_{G^\eta} \leq \frac{1}{2}\int_0^T \sigma_t^2 v_t \mathrm{d}t}] + \mathbb{E}[(\gamma_{G_\delta^\eta})^{-p} \mathbb{1}_{\gamma_{G^\eta} > \frac{1}{2}\int_0^T \sigma_t^2 v_t \mathrm{d}t}] \\
&\leq (\delta^2 \frac{T}{2})^{-p} \mathbb{P}(\int_0^T \sigma_t^2 v_t \mathrm{d}t - \gamma_{G^\eta} \geq \frac{\int_0^T \sigma_t^2 v_t \mathrm{d}t}{2}) + (\frac{1}{2}\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-p} \\
&\leq_c (\delta^2 T)^{-p} (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-q} \left\|\gamma_{G^\eta} - \int_0^T \sigma_t^2 v_t \mathrm{d}t\right\|_q^q + (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-p} \\
&\leq_c (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-p} \Big[(\delta^2 T)^{-p} (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-q+p} \left\|\gamma_{G^\eta} - \int_0^T \sigma_t^2 v_t \mathrm{d}t\right\|_q^q + 1\Big],
\end{aligned}$$

where we have used the Markov inequality at the second inequality. Then choosing $q = 6p$ and using (5.39)-(5.51), we readily obtain:

$$\begin{aligned}
\|(\gamma_{G_\delta^\eta})^{-1}\|_p &\leq_c (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-1} \Big[(|\sigma|_\infty [\xi_{\sup}^3 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2] T^2)^{-2} (T|\sigma|_\infty^2)^{-5} (|\sigma|_\infty^2 [\xi_{\sup} + \mathcal{M}_1(\sigma)] T^{\frac{3}{2}})^6 + 1\Big] \\
&\leq_c (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-1}.
\end{aligned}$$

(5.53) is a straightforward application of [Nualart 2006, Lemma 2.1.6]; we have $\forall t \in [0,T]$, $\forall i \in \{1,2\}$:

$$D_t^i(\gamma_{G_\delta^\eta})^{-1} = -\frac{D_t^i \gamma_{G_\delta^\eta}}{\gamma_{G_\delta^\eta}^2} = -2\frac{\int_0^T [D_u^1 G^\eta D_{t,u}^{i,1} G^\eta + D_u^2 G^\eta D_{t,u}^{i,2} G^\eta] \mathrm{d}u.}{\gamma_{G_\delta^\eta}^2}$$

Then using (5.41)-(5.43)-(5.45)-(5.52) we get readily $\forall p \geq 1$:

$$\sup_{t \in [0,T]} \|D_t^i(\gamma_{G_\delta^\eta})^{-1}\|_p \leq_c T|\sigma|_\infty^2 (\mathcal{M}_1(\sigma) + \xi_{\sup})(\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-2},$$

which leads to the announced result. □

Now plug (5.52)-(5.53)-(5.54) in (5.50) to complete the proof.

5.4 Expansion formulas for the implied volatility

In this section we apply our third approximation formula to the particular payoff function $h(x) = (e^x - K)_+$, i.e. a Call payoff function with strike K for which the expansion remains valid (see Subsection 5.2.3). The risk-free rate and the dividend yield[2] are set to 0. In order to obtain more tractable and accurate formulas, we aim at extracting implied volatility expansions from the price approximation formula. It has been shown in Chapter 2 that, in addition to their simplicity, direct implied volatility expansions are more accurate than the corresponding price formulas.

Notations.

▷ **Call options.** We denote by $\mathrm{Call}(S_0, T, K)$ the price at time 0 of a Call option with spot $S_0 = e^{x_0}$, maturity T and strike K, written on the asset $S = e^X$ that is $\mathrm{Call}(S_0, T, K) = \mathbb{E}(e^{X_T} - K)_+$. As usual, ATM (At The Money) Call refers to $S_0 \approx K$, ITM (In The Money) to $S_0 \gg K$, OTM (Out The Money) to $S_0 \ll K$.

▷ **Black-Scholes Call price function.** For the sake of completeness, we give the *Black-Scholes Call price* function depending on log-spot x, total variance y and log-strike k:

$$\mathrm{Call}^{\mathrm{BS}}(x, y, k) = e^x \mathcal{N}(d_1(x, y, k)) - e^k \mathcal{N}(d_2(x, y, k)) \tag{5.55}$$

where:

$$\mathcal{N}(x) = \int_{-\infty}^{x} \mathcal{N}'(u)\mathrm{d}u, \quad \mathcal{N}'(u) = \frac{e^{-u^2/2}}{\sqrt{2\pi}}, \quad d_1(x, y, k) = \frac{x - k}{\sqrt{y}} + \frac{1}{2}\sqrt{y}, \quad d_2(x, y, k) = d_1(x, y, k) - \sqrt{y}.$$

In the following, $x_0 = \log(S_0)$ will represent the log-spot, $k = \log(K)$ the log-strike, $x_{avg} = (x_0 + k)/2 = \log(\sqrt{S_0 K})$ the mid-point between the log-spot and the log-strike, $m = x_0 - k = \log(S_0/K)$ the log-moneyness. The value $\mathrm{Call}^{\mathrm{BS}}(x_0, \int_0^T \sigma_t^2 v_t \mathrm{d}t, k) = \mathbb{E}[(e^{X_T^P} - e^k)_+]$ equals $\mathrm{Call}(S_0, T, K) = \mathbb{E}[(e^{X_T} - K)_+]$ when $\mathcal{M}_1(\sigma) = \xi_{\sup} = 0$. For (x, T, k) given, the *implied Black-Scholes volatility* of a price $\mathrm{Call}(e^x, T, e^k)$ is the unique non-negative volatility parameter $\sigma_{\mathrm{I}}(x, T, k)$ such that:

$$\mathrm{Call}^{\mathrm{BS}}(x, \sigma_{\mathrm{I}}^2(x, T, k)T, k) = \mathrm{Call}(e^x, T, e^k). \tag{5.56}$$

The reader can find in Proposition 5.6.3.2 the definition of $\mathrm{Vega}^{\mathrm{BS}}$ and $\mathrm{Vomma}^{\mathrm{BS}}$ which are the first two derivatives of $\mathrm{Call}^{\mathrm{BS}}$ w.r.t. the volatility parameter.

▷**Quadratic mean of the volatility on** $[0, T]$**.** For any spatial point $z \in \mathbb{R}$, we denote by $\overline{\sigma}_z$ the quadratic mean on $[0, T]$ of $(\sigma_t(z)\sqrt{v_t})_{t \in [0,T]}$ defined by:

$$\overline{\sigma}_z = \sqrt{\frac{1}{T}\int_0^T \sigma_t^2(z) v_t \mathrm{d}t}. \tag{5.57}$$

This notation is frequently used for the points x_0 and x_{avg}. When applied in x_0, we simply write $\bar{\sigma}$ if unambiguous.

[2] Adaptation of the results for non-zero but deterministic risk-free rate and dividend yield is straightforward by considering the discounted asset.

5.4.1 Implied volatility expansion at spot

We introduce new corrective coefficients useful for the implied volatility expansions:

Definition 5.4.1.1. *Assume $(\mathcal{H}^{\sigma}_{x_0})$. We define the following corrective coefficients:*

$$
\begin{aligned}
\gamma_{0a,T} =&\bar\sigma + \frac{C^s_{1,T}}{4\bar\sigma T},\\
\gamma_{0b,T} =&\frac{C^l_{9,T}}{2\bar\sigma^3T^2} - \frac{C^l_{10,T}}{4\bar\sigma^3T^2} - \frac{3C^l_{10,T}}{\bar\sigma^5T^3} + \frac{C^s_{2,T}}{8\bar\sigma T} - \frac{C^s_{2,T}}{2\bar\sigma^3T^2} - \frac{C^s_{3,T}}{16\bar\sigma T} - \frac{C^s_{3,T}}{4\bar\sigma^3T^2} + \frac{3(C^s_{1,T})^2}{8\bar\sigma^5T^3}\\
&+ \frac{C^{ls}_{8,T}}{\bar\sigma^3T^2} - \frac{C^{ls}_{4,T}}{8\bar\sigma T} - \frac{3C^{ls}_{4,T}}{2\bar\sigma^3T^2} - \frac{C^{ls}_{5,T}}{8\bar\sigma T} - \frac{C^{ls}_{5,T}}{2\bar\sigma^3T^2} - \frac{C^{ls}_{6,T}}{2\bar\sigma^3T^2} + \frac{(C^l_{1,T}C^s_{1,T})}{8\bar\sigma^3T^2} + \frac{3(C^l_{1,T}C^s_{1,T})}{2\bar\sigma^5T^3},\\
\gamma_{1a,T} =&-\frac{C^l_{1,T}}{\bar\sigma^3T^2} - \frac{C^s_{1,T}}{2\bar\sigma^3T^2},\\
\gamma_{1b,T} =&-\frac{C^s_{2,T}}{2\bar\sigma^3T^2} + \frac{3(C^s_{1,T})^2}{8\bar\sigma^5T^3} - \frac{C^{ls}_{2,T}+C^{ls}_{3,T}}{2\bar\sigma^3T^2} - \frac{C^{ls}_{4,T}}{2\bar\sigma^3T^2} - \frac{C^{ls}_{6,T}}{4\bar\sigma^3T^2} + \frac{3(C^l_{1,T}C^s_{1,T})}{4\bar\sigma^5T^3},\\
\gamma_{2,T} =&\frac{C^l_{3,T}}{\bar\sigma^5T^3} - \frac{3C^l_{4,T}}{\bar\sigma^5T^3} + 6\frac{C^l_{10,T}}{\bar\sigma^7T^4} + \frac{C^s_{2,T}}{2\bar\sigma^5T^3} + \frac{C^s_{3,T}}{4\bar\sigma^5T^3} - \frac{3(C^s_{1,T})^2}{4\bar\sigma^7T^4} + \frac{C^{ls}_{2,T}+C^{ls}_{3,T}}{\bar\sigma^5T^3}\\
&+ \frac{3C^{ls}_{4,T}}{2\bar\sigma^5T^3} + \frac{C^{ls}_{5,T}}{2\bar\sigma^5T^3} + \frac{C^{ls}_{6,T}}{2\bar\sigma^5T^3} - \frac{3(C^l_{1,T}C^s_{1,T})}{\bar\sigma^7T^4},
\end{aligned}
$$

where $(C^l_{i,T})_{1\le i\le 4}$-$(C^s_{i,T})_{1\le i\le 3}$-$(C^{ls}_{i,T})_{1\le i\le 6}$ are defined in Theorem 5.2.2.1 and $C^l_{9,T}$-$C^l_{10,T}$-$C^{ls}_{8,T}$ are defined by:

$$
\begin{aligned}
C^l_{9,T} =&\omega(\sigma^2v,((\sigma^{(1)})^2+\sigma\sigma^{(2)})v,\sigma^2v)^T_0, \qquad C^l_{10,T} =\omega(\sigma^2v,\sigma\sigma^{(1)}v,\sigma\sigma^{(1)}v,\sigma^2v)^T_0,\\
C^{ls}_{8,T} =&\omega(\rho\xi\sigma v,\sigma\sigma^{(1)},\sigma^2v)^T_0.
\end{aligned}
$$

To obtain an implied volatility expansion, we use the relations between the Greeks w.r.t. the log-spot which naturally appear when applying the expansion of Theorem 5.2.2.1 and the sensitivities w.r.t. the volatility parameter. These relations are available on Appendix 5.6.3. Applying Proposition 5.6.3.3, the third order approximation formula (5.9) can be transformed into:

$$
\begin{aligned}
&\mathrm{Call}(S_0,T,K) \qquad (5.58)\\
=&\mathrm{Call}^{\mathrm{BS}}(x_0,\bar\sigma^2T,k) + \mathrm{Vega}^{\mathrm{BS}}(x_0,\bar\sigma^2T,k)\Big[-\frac{C^l_{1,T}}{\bar\sigma^3T^2}m + \frac{C^l_{2,T}}{2\bar\sigma T} - \frac{C^l_{3,T}}{\bar\sigma^3T^2} + \frac{C^l_{3,T}}{\bar\sigma^5T^3}m^2 - \frac{C^l_{4,T}}{4\bar\sigma T} - \frac{3C^l_{4,T}}{\bar\sigma^3T^2}\\
&+ \frac{3C^l_{4,T}}{\bar\sigma^5T^3}m^2 + \frac{(C^l_{1,T})^2}{8\bar\sigma^3T^2} + \frac{3(C^l_{1,T})^2}{2\bar\sigma^5T^3} - \frac{3(C^l_{1,T})^2}{\bar\sigma^7T^4}m^2 + \frac{C^s_{1,T}}{4\bar\sigma T} - \frac{C^s_{1,T}}{2\bar\sigma^3T^2}m + \frac{C^s_{2,T}}{8\bar\sigma T} - \frac{C^s_{2,T}}{2\bar\sigma^3T^2} - \frac{C^s_{2,T}}{2\bar\sigma^3T^2}m\\
&+ \frac{C^s_{2,T}}{2\bar\sigma^5T^3}m^2 - \frac{C^s_{3,T}}{16\bar\sigma T} - \frac{C^s_{3,T}}{4\bar\sigma^3T^2} + \frac{C^s_{3,T}}{4\bar\sigma^5T^3}m^2 + \frac{3(C^s_{1,T})^2}{8\bar\sigma^5T^3} + \frac{3(C^s_{1,T})^2}{8\bar\sigma^5T^3}m - \frac{3(C^s_{1,T})^2}{4\bar\sigma^7T^4}m^2 + \frac{C^{ls}_{1,T}}{\bar\sigma T}\\
&- \frac{(C^{ls}_{2,T}+C^{ls}_{3,T})}{\bar\sigma^3T^2} - \frac{(C^{ls}_{2,T}+C^{ls}_{3,T})}{2\bar\sigma^3T^2}m + \frac{(C^{ls}_{2,T}+C^{ls}_{3,T})}{\bar\sigma^5T^3}m^2 - \frac{C^{ls}_{4,T}}{8\bar\sigma T} - \frac{3C^{ls}_{4,T}}{2\bar\sigma^3T^2} - \frac{C^{ls}_{4,T}}{2\bar\sigma^3T^2}m + \frac{3C^{ls}_{4,T}}{2\bar\sigma^5T^3}m^2\\
&- \frac{C^{ls}_{5,T}}{8\bar\sigma T} - \frac{C^{ls}_{5,T}}{2\bar\sigma^3T^2} + \frac{C^{ls}_{5,T}}{2\bar\sigma^5T^3}m^2 - \frac{C^{ls}_{6,T}}{2\bar\sigma^3T^2} - \frac{C^{ls}_{6,T}}{4\bar\sigma^3T^2}m + \frac{C^{ls}_{6,T}}{2\bar\sigma^5T^3}m^2 + \frac{(C^l_{1,T}C^s_{1,T})}{8\bar\sigma^3T^2} + \frac{3(C^l_{1,T}C^s_{1,T})}{2\bar\sigma^5T^3}\\
&+ \frac{3(C^l_{1,T}C^s_{1,T})}{4\bar\sigma^5T^3}m - \frac{3(C^l_{1,T}C^s_{1,T})}{\bar\sigma^7T^4}m^2\Big]\\
&+ \frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x_0,\bar\sigma^2T,k)\Big[\Big(\frac{C^l_{1,T}}{\bar\sigma^3T^2}\Big)^2m^2 + (C^s_{1,T})^2\Big(-\frac{m}{2\bar\sigma^3T^2} + \frac{1}{4\bar\sigma T}\Big)^2 + (C^l_{1,T}C^s_{1,T})\Big(\frac{m^2}{\bar\sigma^6T^4} - \frac{m}{2\bar\sigma^4T^3}\Big)\Big]
\end{aligned}
$$

$$+O(K\,|\sigma|_\infty[\xi_{\sup}^3+\mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma)+\xi_{\sup})^2]\,T^2).$$

Using the following relations easy to establish:

$$C^{ls}_{1,T}\omega(\sigma^2 v)_0^T=C^{ls}_{2,T}+C^{ls}_{3,T}+C^{ls}_{8,T},\quad C^{l}_{2,T}\omega(\sigma^2 v)_0^T=C^{l}_{9,T}+2C^{l}_{3,T},\quad C^{l}_{4,T}\omega(\sigma^2 v)_0^T=C^{l}_{10,T}+\frac{(C^{l}_{1,T})^2}{2},$$

we can write that:

$$\frac{C^{ls}_{1,T}}{\bar\sigma T}-\frac{C^{ls}_{2,T}+C^{ls}_{3,T}}{\bar\sigma^3T^2}=\frac{C^{ls}_{8,T}}{\bar\sigma^3T^2},\qquad \frac{C^{l}_{2,T}}{2\bar\sigma T}-\frac{C^{l}_{3,T}}{\bar\sigma^3T^2}=\frac{C^{l}_{9,T}}{2\bar\sigma^3T^2},$$

$$-\frac{C^{l}_{4,T}}{4\bar\sigma T}+\frac{[C^{l}_{1,T}]^2}{8\bar\sigma^3T^2}-\frac{3C^{l}_{4,T}}{\bar\sigma^3T^2}+\frac{3[C^{l}_{1,T}]^2}{2\bar\sigma^5T^3}=-\frac{C^{l}_{10,T}}{4\bar\sigma^3T^2}-\frac{3C^{l}_{10,T}}{\bar\sigma^5T^3},\quad 3\frac{C^{l}_{4,T}}{\bar\sigma^5T^3}-\frac{3[C^{l}_{1,T}]^2}{\bar\sigma^7T^4}=-\frac{3C^{l}_{4,T}}{\bar\sigma^5T^3}+6\frac{C^{l}_{10,T}}{\bar\sigma^7T^4},$$

and finally obtain for (5.58):

$$\begin{aligned}
&\mathrm{Call}(S_0,T,K)\\
=&\mathrm{Call}^{\mathrm{BS}}(x_0,\bar\sigma^2T,k)+\mathrm{Vega}^{\mathrm{BS}}(x_0,\bar\sigma^2T,k)[(\gamma_{0a,T}+\gamma_{0b,T})-\bar\sigma+(\gamma_{1a,T}+\gamma_{1b,T})m+\gamma_{2,T}m^2]\\
&+\frac12\mathrm{Vomma}^{\mathrm{BS}}(x_0,\bar\sigma^2T,k)[\gamma_{0a,T}-\bar\sigma+\gamma_{1a,T}m]^2+O(K\,|\sigma|_\infty[\xi_{\sup}^3+\mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma)+\xi_{\sup})^2]\,T^2)\\
\approx&\mathrm{Call}^{\mathrm{BS}}(x_0,[(\gamma_{0a,T}+\gamma_{0b,T})+(\gamma_{1a,T}+\gamma_{1b,T})m+\gamma_{2,T}m^2]^2T,k),
\end{aligned}$$

This reads as an expansion of the implied volatility and proves the next Theorem:

Theorem 5.4.1.1. *(3rd order expansion of the implied volatility). Assume* ($\mathcal{H}^{\sigma}_{x_0}$) *and* ($P$)*. We have:*

$$\sigma_{\mathrm{I}}(x_0,T,k)=\gamma_{0a,T}+\gamma_{0b,T}+(\gamma_{1a,T}+\gamma_{1b,T})m+\gamma_{2,T}m^2+\mathrm{Error}^{\mathrm{I}}_{3,x_0}. \tag{5.59}$$

At fixed maturity T, the implied volatility approximation is written as a quadratic function w.r.t. the log-moneyness m with the coefficients γ defined in Definition 5.4.1.1.

Corollaries of Theorem 5.4.1.1.

▷**Estimates of** $\mathrm{Error}^{\mathrm{I}}_{3,x_0}$**.** In addition to the above implied volatility expansion, one can under additional technical assumptions upper bound the residual terms. Assume that $|m|\le\mathcal{K}|\sigma|_\infty\sqrt{T}$ for a given $\mathcal{K}>0$ and that $\mathcal{M}_1(\sigma)$, $\mathcal{M}_0(\sigma)$, $\xi_{\sup}$ and T are globally small enough to ensure that $\gamma=\gamma_{0a,T}+\gamma_{0b,T}+(\gamma_{1a,T}+\gamma_{1b,T})m+\gamma_{2,T}m^2>0$. Under these assumptions, one can prove that:

$$|\mathrm{Error}^{\mathrm{I}}_{3,x_0}|=O(|\sigma|_\infty[\xi_{\sup}^3+\mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma)+\xi_{\sup})^2]T^{\frac32}),$$

where the generic constant depends in an increasing way on $\mathcal{K}$, what justifies the label of third order expansion. Although tedious to write, the proof does not contain huge mathematical difficulties and is performed in Chapter 2 Subsection 2.1.4 in the case of pure local volatility models for the second order. For the sake of brevity we leave it as an exercise to the reader giving only the outline: perform a second order expansion of $\sigma\mapsto\mathrm{Call}^{\mathrm{BS}}(x_0,\sigma^2T,k)$ at γ around $\bar\sigma$, a zero order expansion at $\sigma_{\mathrm{I}}(x_0,T,k)$ around γ, apply Theorem 5.2.2.1, use classical estimates for $\partial^n_{\sigma^n}\mathrm{Call}^{\mathrm{BS}}(x_0,\sigma^2T,k)$ for $n=1,2,3$ (see Chapter 2 Corollary 2.6.1.1) and carefully gather terms and evaluate their magnitude.

▷**Short maturity skew and smile behaviours.** We analyse the behaviour of the approximation formula (5.59) at the money (i.e. $m\approx0$) and for short maturity (i.e. $T\ll1$). In view of (5.59) and the various coefficients C^l, C^s, C^{ls} and γ (see Definition 5.4.1.1 and Theorem 5.2.2.1), assuming that σ_t, $\sigma_t^{(1)}$ and $\sigma_t^{(2)}$ are continuous at $t=0$, we obtain for the level, the slope and the curvature ATM:

$$[\sigma_{\mathrm{I}}(x_0,T,k)]|_{k=x_0} \approx \gamma_{0a,T}+\gamma_{0b,T} \approx \sigma_0\sqrt{v_0},$$

$$\partial_k[\sigma_{\mathrm{I}}(x_0,T,k)]|_{k=x_0} \approx -\gamma_{1a,T}-\gamma_{1b,T} \approx -\gamma_{1a,T} \approx \frac{\sigma_0^{(1)}\sqrt{v_0}}{2}+\frac{\rho_0\xi_0}{4\sqrt{v_0}}, \tag{5.60}$$

$$\partial^2_{k^2}[\sigma_{\mathrm{I}}(x_0,T,k)]|_{k=x_0} \approx 2\gamma_{2,T} = \frac{\sigma_0^{(2)}\sqrt{v_0}}{3}-\frac{[\sigma_0^{(1)}]^2\sqrt{v_0}}{6\sigma_0}-\frac{5\rho_0^2\xi_0^2}{24\sigma_0 v_0^{\frac{3}{2}}}+\frac{\xi_0^2}{12\sigma_0 v_0^{\frac{3}{2}}}, \tag{5.61}$$

where we have used the following estimate:

$$|\gamma_{0a,T}+\gamma_{0b,T}-\bar{\sigma}|+|\gamma_{1b,T}| \leq_c T,$$

and consequently neglected these terms considered as maturity bias. We observe that:
1) In case of null correlation, our approximation coincides with [Forde 2012a, Theorem 4.1]. Otherwise, we notice that the terms C^{ls} involving simultaneously ρ_0, ξ_0 and $\sigma_0^{(1)}$ vanish and that the slope of the implied volatility is modified. The correlation is therefore interpreted as a skew parameter and there might be a competition between $\sigma_0^{(1)}$ and ρ in the calibration procedures.
2) For pure local volatility models (i.e. $\xi_{\sup}=0$), we retrieve the results of Theorem 2.3.3.1 of Chapter 2.
3) For pure Heston models (i.e. $\mathcal{M}_1(\sigma)=0$), we recover the expansion given in [Forde 2009, Theorem 2.5]. In the case of zero correlation, the approximation formula (5.59) becomes for short maturity:

$$\sigma_{\mathrm{I}}(x_0,T,k) \approx \bar{\sigma}-\frac{C^s_{3,T}}{16\bar{\sigma}T}-\frac{C^s_{3,T}}{4\bar{\sigma}^3T^2}+\frac{C^s_{3,T}}{4\bar{\sigma}^5T^3}m^2 \approx \bar{\sigma}-\frac{\xi_0^2\sigma_0T}{24\sqrt{v_0}}[\frac{\sigma_0^2v_0T}{4}+1]+\frac{\xi_0^2}{24\sigma_0v_0^{\frac{3}{2}}}m^2.$$

We have retrieved that an uncorrelated Heston model induces a symmetric smile w.r.t. the moneyness. The implied volatility ATM is slightly smaller than the local volatility function ATM and becomes larger ITM or OTM, the smile increasing with the volatility of volatility ξ_0. If we consider a negative correlation, in view of (5.60) (the slope becomes negative and increases in absolute value with $|\rho_0|$) and (5.61) (the curvature is decreasing until reaching zero for $|\rho_0|=\sqrt{2/5}\approx 0.63$), the center of the short maturity smile is shifted to the right and the smile changes from a symmetric shape to a negative skew. The converse is realised for a positive correlation.

▷**Calibration issues for time independent parameters.** Generally the local volatility function is completely determined by a level and a slope parameters identified respectively with the local volatility and its first derivative ATM. This is for instance the case of the CEV model (see (5.71)). For general local and stochastic volatility models, the level of the volatility can be fixed throughout the local volatility function whereas the stochastic variance process can be normalised with an initial value v_0 equal to 1. We have seen that for an uncorrelated local and stochastic volatility model:
i) The level parameter of the local volatility is linked to the short time implied volatility ATM,
ii) The skew parameter of the local volatility is connected to the short time slope of the implied volatility ATM,
iii) Once the local volatility function is identified, the volatility of volatility parameter is linked to the short time curvature of the implied volatility ATM.
These features allow us to suggest good surrogates for these three parameters in view of a calibration procedure by simply estimate the market implied volatility curve for short maturity. But we have observed that the correlation modify the short term skew and it is well known that the mean reversion parameter of a CIR process plays a similar role than the volatility of volatility but in the inverse way.

Thus we can find models having different parameters but reproducing the same smile for one maturity: like for the Heston model, the calibration with a single maturity is an ill-posed problem.

5.4.2 Implied volatility expansion at mid-point

It has been empirically proven in Chapter 2 Section 2.5 throughout exhaustive numerical experiments that for the pure local volatility case, expansions with local volatility function frozen at mid-point $x_{avg} = (x_0 + k)/2 = \log(\sqrt{S_0 K})$ give better results. First we introduce new notations and definitions.

Notations.

▷ **Corrective coefficients frozen at mid-point.** The coefficients C^l, C^s, C^{ls} and γ was naturally defined in Theorem 5.2.2.1 and Definition 5.4.1.1 for the local volatility function σ at log-spot x_0. To consider the same coefficients but with local volatility function frozen at point z where z is generally equal to x_{avg} or x_0, we use the notations $C^l_{i,T}(z)$, $C^s_{i,T}(z)$,$C^{ls}_{i,T}(z)$ and $\gamma_{i,T}(z)$.

▷ **New ellipticity assumption at x_{avg}.** We define similarly $(\mathcal{H}^{\sigma}_{x_{avg}})$ and $(\mathcal{H}^{\sigma}_{x_0})$ by replacing x_0 by x_{avg}. The generic constant in the further estimates will depend in an increasing way on $\frac{|\sigma|^2_\infty T}{\int_0^T \sigma_t^2(x_{avg}) v_t \mathrm{d}t}$.

▷ **Time reversal.** For the coefficients $C^l_{i,T}(x_{avg})$, we introduce the notation $\widetilde{C}^l_{i,T}(x_{avg})$ which means that we have inverted the order of integration of the integrands. For example $\widetilde{C}^l_{1,T}(x_{avg}) = \omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v, \sigma^2(x_{avg})v)_0^T$ instead of $C^l_{1,T}(x_{avg}) = \omega(\sigma^2(x_{avg})v, \sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T$.

Definition 5.4.2.1. *Assume $(\mathcal{H}^{\sigma}_{x_{avg}})$. We define the following corrective coefficients:*

$$\begin{aligned}
\pi_{0a,T}(x_{avg}) =&\gamma_{0a,T}(x_{avg}),\\
\pi_{0b,T}(x_{avg}) =&\gamma_{0b,T}(x_{avg}),\\
\pi_{1a,T}(x_{avg}) =&\frac{\widetilde{C}^l_{1,T}(x_{avg}) - C^l_{1,T}(x_{avg})}{2\bar\sigma^3_{x_{avg}} T^2} - \frac{C^s_{1,T}(x_{avg})}{2\bar\sigma^3_{x_{avg}} T^2},\\
\pi_{1b,T}(x_{avg}) =&\gamma_{1b,T}(x_{avg}) + \frac{C^{ls}_{1,T}(x_{avg})}{4\bar\sigma_{x_{avg}} T} + \frac{C^{ls}_{9,T}(x_{avg})}{8\bar\sigma_{x_{avg}} T} - \frac{C^{ls}_{4,T}(x_{avg}) + C^{ls}_{5,T}(x_{avg}) + C^{ls}_{10,T}(x_{avg})}{8\bar\sigma^3_{x_{avg}} T^2},\\
\pi_{2,T}(x_{avg}) =&\frac{\widetilde{C}^l_{3,T}(x_{avg}) + C^l_{3,T}(x_{avg})}{2\bar\sigma^5_{x_{avg}} T^3} - \frac{3(\widetilde{C}^l_{4,T}(x_{avg}) + C^l_{4,T}(x_{avg}))}{2\bar\sigma^5_{x_{avg}} T^3} - \frac{C^l_{5,T}(x_{avg})}{8\bar\sigma_{x_{avg}} T} + \frac{C^l_{6,T}(x_{avg})}{4\bar\sigma^3_{x_{avg}} T^2}\\
&+6\frac{C^l_{10,T}(x_{avg})}{\bar\sigma^7_{x_{avg}} T^4} + \frac{C^s_{2,T}(x_{avg})}{2\bar\sigma^5_{x_{avg}} T^3} + \frac{C^s_{3,T}(x_{avg})}{4\bar\sigma^5_{x_{avg}} T^3} - \frac{3(C^s_{1,T}(x_{avg}))^2}{4\bar\sigma^7_{x_{avg}} T^4} + \frac{C^{ls}_{2,T}(x_{avg}) + C^{ls}_{3,T}(x_{avg})}{\bar\sigma^5_{x_{avg}} T^3}\\
&+\frac{3C^{ls}_{4,T}(x_{avg})}{2\bar\sigma^5_{x_{avg}} T^3} + \frac{C^{ls}_{5,T}(x_{avg})}{2\bar\sigma^5_{x_{avg}} T^3} + \frac{C^{ls}_{6,T}(x_{avg})}{2\bar\sigma^5_{x_{avg}} T^3} - \frac{3(C^l_{1,T} C^s_{1,T})(x_{avg})}{\bar\sigma^7_{x_{avg}} T^4} - \frac{C^{ls}_{1,T}(x_{avg})}{2\bar\sigma^3_{x_{avg}} T^2} - \frac{C^{ls}_{9,T}(x_{avg})}{4\bar\sigma^3_{x_{avg}} T^2}\\
&+3\frac{C^{ls}_{4,T}(x_{avg}) + C^{ls}_{5,T}(x_{avg}) + C^{ls}_{10,T}(x_{avg})}{4\bar\sigma^5_{x_{avg}} T^3},
\end{aligned}$$

where $C^l_{5,T}(x_{avg})$, $C^l_{6,T}(x_{avg})$, $C^{ls}_{9,T}(x_{avg})$ and $C^{ls}_{10,T}(x_{avg})$ are defined by:

$$\begin{aligned}
C^l_{5,T}(x_{avg}) =&\omega(((\sigma^{(1)})^2 + \sigma\sigma^{(2)})(x_{avg})v)_0^T, \quad C^l_{6,T}(x_{avg}) =\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v, \sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T,\\
C^{ls}_{9,T}(x_{avg}) =&\omega(\rho\xi\sigma^{(1)}(x_{avg})v, \sigma^2(x_{avg}))_0^T, \quad C^{ls}_{10,T}(x_{avg}) =\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v, \rho\xi\sigma(x_{avg})v, \sigma^2(x_{avg}))_0^T.
\end{aligned}$$

To obtain a new implied volatility approximation, we consider the formula 5.59 and we perform a Taylor expansion around the mid-point. First we analyse the leading term $\bar{\sigma}_{x_0}$ of 5.59:

Lemma 5.4.2.1. *Assume* $(\mathcal{H}^{\sigma}_{x_0})$-$(\mathcal{H}^{\sigma}_{x_{avg}})$ *and suppose that* $|m| \le \mathcal{K}|\sigma|_\infty \sqrt{T}$ *for a given* $\mathcal{K} > 0$. *Then we have:*

$$\bar{\sigma}_{x_0} = \bar{\sigma}_{x_{avg}} + \frac{\omega((\sigma\sigma^{(1)})(x_{avg})v)_0^T}{2\bar{\sigma}_{x_{avg}} T} m + \frac{C^l_{5,T}(x_{avg})}{8\bar{\sigma}_{x_{avg}} T} m^2 - \frac{C^l_{6,T}(x_{avg})}{4\bar{\sigma}^3_{x_{avg}} T^2} m^2 + O(|\sigma|_\infty \mathcal{M}_0(\sigma)^2 \mathcal{M}_1(\sigma) T^{\frac{3}{2}}), \quad (5.62)$$

where $C^l_{5,T}$ *and* $C^l_{6,T}$ *are defined in Definition 5.4.2.1.*

Proof. First notice that $(\mathcal{H}^{\sigma}_{x_0})$-$(\mathcal{H}^{\sigma}_{x_{avg}})$ implies the strict positivity of $u\overline{\sigma}^2_{x_0} + (1-u)\overline{\sigma}^2_{x_{avg}}$ for any $u \in [0,1]$. Then apply the Taylor formula twice: firstly for the function $y \mapsto \sqrt{y}$ at $y = \frac{\omega(\sigma^2(x_0)v)_0^T}{T}$ around $y = \frac{\omega(\sigma^2(x_{avg})v)_0^T}{T}$ and secondly for the function $x \mapsto \sigma^2_t(x)$ at $x = x_0$ around $x = x_{avg}$, for any $t \in [0,T]$. It gives:

$$\begin{aligned}
\bar{\sigma}_{x_0} &= \sqrt{\frac{\omega(\sigma^2(x_0)v)_0^T}{T}} = \bar{\sigma}_{x_{avg}} + \frac{\omega(\sigma^2(x_0)v)_0^T - \omega(\sigma^2(x_{avg})v)_0^T}{2\bar{\sigma}_{x_{avg}} T} - \frac{[\omega(\sigma^2(x_0)v)_0^T - \omega(\sigma^2(x_{avg})v)_0^T]^2}{8\bar{\sigma}^3_{x_{avg}} T^2} + R_1 \\
&= \bar{\sigma}_{x_{avg}} + \frac{\omega((\sigma\sigma^{(1)})(x_{avg})v)_0^T}{2\bar{\sigma}_{x_{avg}} T} m + \frac{\omega(((\sigma^{(1)})^2 + \sigma\sigma^{(2)})(x_{avg})v)_0^T}{8\bar{\sigma}_{x_{avg}} T} m^2 - \frac{[\omega((\sigma\sigma^{(1)})(x_{avg})v)_0^T]^2}{8\bar{\sigma}^3_{x_{avg}} T^2} m^2 \\
&\quad + R_1 + R_2 + R_3,
\end{aligned}$$

where:

$$\begin{aligned}
R_1 &= (\omega(\sigma^2(x_0)v)_0^T - \omega(\sigma^2(x_{avg})v)_0^T)^3 \int_0^1 \frac{3}{8T^3(u\overline{\sigma}^2_{x_0} + (1-u)\overline{\sigma}^2_{x_{avg}})^{\frac{5}{2}}} \frac{(1-u)^2}{2} \mathrm{d}u, \\
R_2 &= \frac{m^3}{16\bar{\sigma}_{x_{avg}} T} \int_0^1 \partial^3_{x^3}(\omega(\sigma^2(x)v)_0^T)|_{x=ux_0+(1-u)x_{avg}} \frac{(1-u)^2}{2} \mathrm{d}u, \\
R_3 &= -\frac{(\omega(\sigma^2(x_0)v)_0^T - \omega(\sigma^2(x_{avg})v)_0^T + m\omega((\sigma\sigma^{(1)})(x_{avg})v)_0^T)}{8\bar{\sigma}^3_{x_{avg}} T^2} \frac{m^2}{4} \\
&\quad \times \int_0^1 \partial^2_{x^2}(\omega(\sigma^2(x)v)_0^T)|_{x=ux_0+(1-u)x_{avg}} (1-u) \mathrm{d}u.
\end{aligned}$$

Next remark that $\omega(((\sigma^{(1)})^2 + \sigma\sigma^{(2)})(x_{avg})v)_0^T = C^l_{5,T}(x_{avg})$ and that $[\omega((\sigma\sigma^{(1)})(x_{avg})v)_0^T]^2 = 2C^l_{6,T}(x_{avg})$. Then we readily obtain with the assumption on m and $(\mathcal{H}^{\sigma}_{x_0})$-$(\mathcal{H}^{\sigma}_{x_{avg}})$ that $R_1 + R_2 + R_3 = O(|\sigma|_\infty \mathcal{M}_0(\sigma)^2 \mathcal{M}_1(\sigma) T^{\frac{3}{2}})$. □

Second we analyse the corrective terms:

Lemma 5.4.2.2. *Assume* $(\mathcal{H}^{\sigma}_{x_0})$-$(\mathcal{H}^{\sigma}_{x_{avg}})$ *and suppose that* $|m| \le \mathcal{K}|\sigma|_\infty \sqrt{T}$ *for a given* $\mathcal{K} > 0$. *Then we have:*

$$\begin{aligned}
&\gamma_{0a,T}(x_0) - \bar{\sigma}_{x_0} + \gamma_{1a,T}(x_0) m \qquad (5.63) \\
&= \gamma_{0a,T}(x_{avg}) - \bar{\sigma}_{x_{avg}} + \frac{C^{ls}_{9,T}(x_{avg}) + 2C^{ls}_{1,T}(x_{avg})}{8\bar{\sigma}_{x_{avg}} T} m - \frac{\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T C^s_{1,T}(x_{avg})}{8\bar{\sigma}^3_{x_{avg}} T^2} m + \gamma_{1a,T}(x_{avg}) m \\
&- \frac{C^l_{2,T}(x_{avg}) + 2C^l_{6,T}(x_{avg})}{2\bar{\sigma}^3_{x_{avg}} T^2} m^2 + 3\frac{\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T C^l_{1,T}(x_{avg})}{2\bar{\sigma}^5_{x_{avg}} T^3} m^2 - \frac{C^{ls}_{9,T}(x_{avg}) + 2C^{ls}_{1,T}(x_{avg})}{4\bar{\sigma}^3_{x_{avg}} T^2} m^2
\end{aligned}$$

$$+3\frac{\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T C^s_{1,T}(x_{avg})}{4\bar\sigma^5_{x_{avg}}T^3}m^2+O(|\sigma|_\infty[\xi^3_{\sup}+\mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma)+\xi_{\sup})^2]T^{\frac{3}{2}}),$$

$$\begin{aligned}&\gamma_{0b,T}(x_0)+\gamma_{1b,T}(x_0)m+\gamma_{2,T}(x_0)m^2\\ =&\gamma_{0b,T}(x_{avg})+\gamma_{1b,T}(x_{avg})m+\gamma_{2,T}(x_{avg})m^2+O(|\sigma|_\infty[\xi^3_{\sup}+\mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma)+\xi_{\sup})^2]T^{\frac{3}{2}}).\end{aligned} \tag{5.64}$$

Proof. The above expansions can be proved similarly than the expansion of Lemma 5.4.2.1 with long and tedious computations. Since there is no extra difficulty, we skip further details. □

Lemmas 5.4.2.1 and 5.4.2.2 lead to the next Theorem:

Theorem 5.4.2.1. ***(3rd order expansion of the implied volatility at mid-point)****. Assume $(\mathcal{H}^\sigma_{x_0})$-$(\mathcal{H}^\sigma_{x_{avg}})$ and (P). We have:*

$$\sigma_{\mathrm{I}}(x_0,T,k)=\pi_{0a,T}(x_{avg})+\pi_{0b,T}(x_{avg})+(\pi_{1a,T}(x_{avg})+\pi_{1b,T}(x_{avg}))m+\pi_{2,T}(x_{avg})m^2+\mathrm{Error}^{\mathrm{I}}_{3,\mathrm{x_{avg}}}, \tag{5.65}$$

where the corrective coefficients π are defined in Definition 5.4.2.1.

Proof. We gather terms coming from Lemmas 5.4.2.1 and 5.4.2.2. First notice that:

$$\omega((\sigma\sigma^{(1)})(x_{avg})v)_0^T\omega(\sigma^2(x_{avg})v)_0^T=C^l_{1,T}(x_{avg})+\widetilde{C}^l_{1,T}(x_{avg}).$$

to get:

$$\frac{\omega((\sigma\sigma^{(1)})(x_{avg})v)_0^T}{2\bar\sigma_{x_{avg}}T}m-\frac{C^l_{1,T}(x_{avg})}{\bar\sigma^3_{x_{avg}}T^2}m=\frac{\widetilde{C}^l_{1,T}(x_{avg})-C^l_{1,T}(x_{avg})}{2\bar\sigma^3_{x_{avg}}T^2}m. \tag{5.66}$$

Second remark that:

$$\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T C^s_{1,T}(x_{avg})=C^{ls}_{4,T}(x_{avg})+C^{ls}_{5,T}(x_{avg})+C^{ls}_{10,T}(x_{avg}), \tag{5.67}$$

$$\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T C^l_{1,T}(x_{avg})=2C^l_{4,T}(x_{avg})+\omega((\sigma\sigma^{(1)})(x_{avg})v,\sigma^2(x_{avg})v,(\sigma\sigma^{(1)})(x_{avg})v)_0^T. \tag{5.68}$$

Then use the following relation easy to verify:

$$C^l_{3,T}(x_{avg})-\frac{C^l_{2,T}(x_{avg})\omega(\sigma^2(x_{avg})v)_0^T}{2}+\frac{C^l_{5,T}(x_{avg})[\omega(\sigma^2(x_{avg})v)_0^T]^2}{4}=\frac{C^l_{3,T}(x_{avg})+\widetilde{C}^l_{3,T}(x_{avg})}{2},$$

to write that:

$$\frac{C^l_{3,T}(x_{avg})}{\bar\sigma^5_{x_{avg}}T^3}m^2-\frac{C^l_{2,T}(x_{avg})}{2\bar\sigma^3_{x_{avg}}T^2}m^2+\frac{C^l_{5,T}(x_{avg})}{8\bar\sigma_{x_{avg}}T}m^2=\frac{C^l_{3,T}(x_{avg})+\widetilde{C}^l_{3,T}(x_{avg})}{2\bar\sigma^5_{x_{avg}}T^3}m^2-\frac{C^l_{5,T}(x_{avg})}{8\bar\sigma_{x_{avg}}T}m^2. \tag{5.69}$$

Next, take advantage of the identity:

$$\omega((\sigma\sigma^{(1)})(x_{avg})v,\sigma^2(x_{avg})v,(\sigma\sigma^{(1)})(x_{avg})v)_0^T-C^l_{6,T}(x_{avg})\omega(\sigma^2(x_{avg})v)_0^T=-[C^l_{4,T}(x_{avg})+\widetilde{C}^l_{4,T}(x_{avg})],$$

to obtain:

$$\begin{aligned}&-3\frac{C^l_{4,T}(x_{avg})}{\bar\sigma^5_{x_{avg}}T^3}+\frac{3}{2}\frac{[2C^l_{4,T}(x_{avg})+\omega((\sigma\sigma^{(1)})(x_{avg})v,\sigma^2(x_{avg})v,(\sigma\sigma^{(1)})(x_{avg})v)_0^T]}{\bar\sigma^5_{x_{avg}}T^3}\\ &-\frac{C^l_{6,T}(x_{avg})}{\bar\sigma^3_{x_{avg}}T^2}-\frac{C^l_{6,T}(4x_{avg})}{\bar\sigma^3_{x_{avg}}T^2}\\ =&-3\frac{[C^l_{4,T}(x_{avg})+\widetilde{C}^l_{4,T}(x_{avg})]}{2\bar\sigma^5_{x_{avg}}T^3}+\frac{C^l_{6,T}(x_{avg})}{4\bar\sigma^3_{x_{avg}}T^2}.\end{aligned} \tag{5.70}$$

Finally sum the relations (5.62)-(5.63)-(5.64) and take into account the mathematical reductions (5.66)-(5.67)-(5.68)-(5.69)-(5.70) to obtain the announced result. □

5.5 Numerical experiments

▷**Model and benchmark.** Here we give numerical examples of the accuracy of our implied volatility approximation formula with local volatility at mid-point (see (5.65) in Theorem 5.4.2.1). We consider a time-independent CEV-Heston model:

$$\begin{aligned} \mathrm{d}S_t =& \mu S_t^\beta \sqrt{Y_t}\mathrm{d}W_t,\ S_0 = e^{x_0}, \\ \mathrm{d}Y_t =& \kappa(\theta - Y_t)\mathrm{d}t + \sqrt{Y_t}\xi \mathrm{d}B_t,\ Y_0 = v_0, \\ \mathrm{d}\langle W, B\rangle_t =& \rho \mathrm{d}t. \end{aligned} \tag{5.71}$$

This model is applied directly to the asset price and we apply our various expansion results by considering a fictive log-asset with local volatility function $\sigma(x) = \mu e^{(\beta-1)x}$. Using Proposition 5.6.4.2 in Appendix 5.6.4, the implied volatility formula (5.65) writes explicitly:

$$\begin{aligned} \sigma_{\mathrm{I}}(x_0,T,k) \approx& \mu(S_0K)^{\frac{\beta-1}{2}}\sqrt{\overline{v}}\Big\{1 + \frac{\rho\xi\mu(S_0K)^{\frac{\beta-1}{2}}R_1^sT}{8\overline{v}} + \frac{(\beta-1)^2\mu^2(S_0K)^{\beta-1}\overline{v}T}{24}(1 - \frac{\mu^2(S_0K)^{\beta-1}\overline{v}T}{4}) \\ &+ \frac{\rho^2\xi^2T}{\overline{v}^2}[\frac{3(R_1^s)^2}{32\overline{v}} + R_2^s(\frac{\mu^2(S_0K)^{\beta-1}\overline{v}T}{48} - \frac{1}{12})] - \frac{\xi^2TR_3^s}{\overline{v}^2}[\frac{1}{24} + \frac{\mu^2(S_0K)^{\beta-1}\overline{v}T}{96}] \\ &+ \frac{\rho\xi(\beta-1)\mu(S_0K)^{\frac{\beta-1}{2}}T}{\overline{v}}[\frac{R_1^s}{8} + \mu^2(S_0K)^{\beta-1}\overline{v}T(\frac{R_2^{ls}}{48\overline{v}} - \frac{R_1^s}{32})] \\ &+ \frac{\rho\xi}{\mu(S_0K)^{\frac{\beta-1}{2}}\overline{v}^2}[-\frac{R_1^s}{4} + (\beta-1)\mu^2(S_0K)^{\beta-1}\overline{v}T(\frac{R_1^s}{16} - \frac{R_2^{ls}}{24\overline{v}})]\log(\frac{S_0}{K}) \\ &+ \frac{\rho^2\xi^2T}{\overline{v}^2}[\frac{3(R_1^s)^2}{32\overline{v}} - \frac{R_2^s}{12}]\log(\frac{S_0}{K}) - \frac{(\beta-1)^2}{24}\log^2(\frac{S_0}{K}) \\ &+ \frac{\rho^2\xi^2}{\mu^2(S_0K)^{\beta-1}\overline{v}^3}[\frac{R_2^s}{12} - \frac{3(R_1^s)^2}{16\overline{v}}]\log^2(\frac{S_0}{K}) + \frac{\xi^2R_3^s}{24\mu^2(S_0K)^{\beta-1}\overline{v}^3}\log^2(\frac{S_0}{K})\Big\}, \end{aligned} \tag{5.72}$$

where the coefficients $\overline{v}$, R_1^s, R_2^s, R_3^s and R_2^{ls} are defined in Proposition 5.6.4.1 of Appendix 5.6.4. Note that if the correlation is equal to zero, many terms vanish and the formula becomes very simple.
As a benchmark, we use Monte Carlo methods with a variance reduction technique. The simulated random variable is $(S_T - K)_+$ using an Euler scheme (see [Glasserman 2004, Section 3.4]) and in order to reduce the statistical error, we use the Heston control variate $(S_T^{\mathrm{H}} - K)_+ - \mathbb{E}[(S_T^{\mathrm{H}} - K)_+]$ where $(S_t^H)_{t\in[0,T]}$ follows (5.71) with β fixed at 1. The latter expectation is computed using the Lewis formula [Lewis 2000]. In [Benhamou 2010b], the authors have studied the numerical accuracy of price approximations w.r.t. κ, θ, ξ and ρ in the context of Heston models whereas the influence of the parameters β and μ has been considered in details in Chapter 2 of the thesis in the case of pure local volatility models. This is the reason why we decide to freeze at realistic values the set of model parameters (with an important negative skew) and allow the maturity and the strike to vary in order to see the global accuracy. In all the tests we use the values:

$$S_0 = 1,\quad \mu = 0.25,\quad \beta = 0.5,\quad v_0 = 1,\quad \theta = 1.2,\quad \kappa = 3,\quad \xi = 1.5,\quad \rho = -70\%, \tag{5.73}$$

and we execute the Monte Carlo simulations with 10^7 sample paths and a time discretization of 300 steps by year. Using the Heston control variate, this number of simulations allows to obtain confidence intervals with width reduced to a few bps[3] for a large range of strikes and maturities. All the following computations are performed using C++ on a Intel(R) Core(TM) i5 CPU@2.40GHz with 4 GB of ram.

[3] 1 bp (basis point) is equal to 0.01%.

▷**Accuracy of the implied volatility formula** (5.65)**.** In Tables 5.1, 5.2, 5.3, 5.4, 5.5 and 5.6 (corresponding to the maturities 6M, $1Y$, $2Y$, $3Y$, $5Y$ and $10Y$) we give for various strikes the Black-Scholes implied volatilities estimated by Monte Carlo (`MC`), the bounds of the 95%-confidence interval of the Monte Carlo estimator (`MC-` and `MC+`) and the implied volatilities given by the approximation formula (5.72) ($\mathtt{AF}(\mathtt{x}_{\mathtt{avg}})$). We use the parameters as in (5.73) and the strikes are chosen to be approximately equal to $S_0 e^{q\mu\sqrt{\theta T}}$ where q takes the value of various quantiles of the standard Gaussian law (1%-5%-10%-20%-30%-40%-50%-60%-70%-80%-90%-95%-99%) which allows to cover far ITM and far OTM options. For the sake of completeness, we indicate the computational time to perform the Monte Carlo simulations.

Regarding the results, we see that our approximation formula (5.72) is very accurate, giving errors on implied volatilities smaller than 20 bps for a large range of strikes and maturities. The results for ATM options are truly excellent but we nevertheless observe inaccuracies for extreme strikes, especially for OTM options (however for such strikes the accuracy of the Monte Carlo estimates is less good) and for short maturity. This asymmetry in the errors is probably due to the important correlation. Higher errors for short maturities is a counterintuitive fact with our error estimate (5.10) which was already observed in [Benhamou 2010b] for Heston models. This could be explained by the convergence of the stochastic variance to its stationary regime for long maturities whereas the skew is very important for short maturities owing to the correlation. Thus we observe a maximal error for the whole range of strikes and maturities of approximately 150 bps in Table 5.1 realized for the maturity $6M$ and the extreme strike 0.65. For long maturities ($3Y$, $5Y$ and $10Y$), errors on implied volatility are smaller than 15 bps if we except the largest strike for which the Monte Carlo estimate is questionable because of the very large confidence interval. For instance we sometimes report ND in the tabulars corresponding to the maturities $5Y$ and $10Y$ meaning that the corresponding prices are outside the non-arbitrage bounds.

Last but not least, regarding the computational cost, we observe that we need approximately 2m30s per month of the maturity for the Monte Carlo simulations (4h54m27s for the maturity $10Y$!), whereas the whole set of implied volatilities is computed in less than 1 ms with the implied volatility approximation formula. This is a very significant advantage allowing real-time calibration procedures.

As a conclusion our implied volatility approximation provides very good accuracy with a computational cost close to real-time and is able to deal naturally with general time-dependent local volatility functions.

Table 5.1: Implied Black-Scholes volatilities (%) for the Monte Carlo simulations (execution time: 17m02s) and the approximation AF(x_{avg}) expressed as a function of strikes for $T = 6M$.

Strikes	0.65	0.75	0.80	0.85	0.90	0.95	1	1.05	1.10	1.20	1.25	1.35	1.50
`MC`	34.86	31.86	30.49	29.18	27.94	26.74	25.61	24.52	23.50	21.64	20.82	19.45	18.01
`MC-`	34.85	31.86	30.49	29.18	27.93	26.74	25.61	24.52	23.50	21.64	20.82	19.44	17.95
`MC+`	34.87	31.87	30.49	29.18	27.94	26.75	25.61	24.53	23.50	21.64	20.83	19.46	18.07
$\mathtt{AF}(\mathtt{x}_{\mathtt{avg}})$	35.04	31.93	30.52	29.19	27.93	26.74	25.60	24.52	23.48	21.53	20.61	18.86	16.45

Table 5.2: Implied Black-Scholes volatilities (%) for the Monte Carlo simulations (execution time: 31m33s) and the approximation AF(x_{avg}) expressed as a function of strikes for $T = 1Y$.

Strikes	0.55	0.65	0.75	0.80	0.90	0.95	1	1.05	1.15	1.25	1.40	1.50	1.80
MC	36.36	33.49	31.01	29.89	27.85	26.91	26.02	25.17	23.61	22.22	20.43	19.44	17.32
MC-	36.34	33.48	31.01	29.89	27.84	26.90	26.01	25.17	23.61	22.22	20.43	19.43	17.16
MC+	36.37	33.49	31.02	29.90	27.85	26.91	26.02	25.17	23.62	22.23	20.44	19.45	17.47
AF(x_avg)	36.56	33.58	31.05	29.92	27.85	26.90	26.00	25.15	23.57	22.12	20.16	18.97	15.83

Table 5.3: Implied Black-Scholes volatilities (%) for the Monte Carlo simulations (execution time: 1h4m11s) and the approximation AF(x_{avg}) expressed as a function of strikes for $T = 2Y$.

Strikes	0.45	0.55	0.65	0.75	0.85	0.90	1	1.10	1.20	1.35	1.55	1.80	2.30
MC	37.22	34.46	32.19	30.26	28.59	27.83	26.44	25.20	24.07	22.58	20.89	19.16	16.72
MC-	37.20	34.45	32.18	30.26	28.59	27.83	26.44	25.19	24.07	22.58	20.88	19.14	16.26
MC+	37.24	34.47	32.20	30.27	28.60	27.84	26.45	25.20	24.08	22.59	20.89	19.18	17.05
AF(x_avg)	37.32	34.52	32.22	30.28	28.59	27.83	26.43	25.18	24.04	22.52	20.76	18.87	15.84

5.6 Appendix

5.6.1 Change of model

In this section, we justify why we work without loss of generality with the model (5.1)-(5.2). If we consider a general time-dependent CIR process, the formulation becomes:

$$\begin{aligned} \mathrm{d}X_t =& \Sigma(t,X_t)\sqrt{Y_t}(\mathrm{d}W_t - \frac{\Sigma(t,X_t)\sqrt{Y_t}}{2}\mathrm{d}t),\ X_0 = x_0,\\ \mathrm{d}Y_t =& \kappa_t(\theta_t - Y_t)\mathrm{d}t + \gamma_t\sqrt{Y_t}\mathrm{d}B_t,\ Y_0 = v_0 > 0, \end{aligned}$$

with a correlation $(\rho_t)_{t\in[0,T]}$ between W and B. We assume the next hypothesis:
(P)': κ, θ and γ are positive, measurable and bounded on $[0,T]$ with $\gamma_{inf} > 0$, and $2(\frac{\kappa\theta}{\gamma^2})_{\inf} \geq 1$. Now set $V_t = e^{\int_0^t \kappa_s \mathrm{d}s} Y_t$. A direct application of the Itô's formula leads to:

$$\mathrm{d}V_t = (e^{\int_0^t \kappa_s \mathrm{d}s}\kappa_t\theta_t)\mathrm{d}t + (e^{\frac{1}{2}\int_0^t \kappa_s \mathrm{d}s}\xi_t)\sqrt{V_t}\mathrm{d}B_t,\ V_0 = v_0 > 0,$$

while the dynamic of X becomes:

$$\mathrm{d}X_t = \Sigma(t,X_t)e^{-\frac{1}{2}\int_0^t \kappa_s \mathrm{d}s}\sqrt{V_t}(\mathrm{d}W_t - \frac{\Sigma(t,X_t)e^{-\frac{1}{2}\int_0^t \kappa_s \mathrm{d}s}\sqrt{V_t}}{2}\mathrm{d}t),\ X_0 = x_0.$$

Setting $\alpha_t = e^{\int_0^t \kappa_s \mathrm{d}s}\kappa_t\theta_t$ and $\xi_t = e^{\frac{1}{2}\int_0^t \kappa_s \mathrm{d}s}\gamma_t$ for any $t \in [0,T]$, $\sigma(t,x) = \Sigma(t,x)e^{-\frac{1}{2}\int_0^t \kappa_s \mathrm{d}s}$ for any $(t,x) \in [0,T]\times\mathbb{R}$, we obtain a formulation equivalent to (5.1)-(5.2). Observe that $(P') \Longleftrightarrow (P)$ and that the local volatility functions σ and Σ have the same space regularity.

5.6.2 Explicit computation of the corrective terms of Theorem 5.2.2.1

We give the full derivation of the corrective terms in the approximation (5.9) of Theorem 5.2.2.1. We begin with the proof of Lemma 5.2.3.1 and next we give the details of the computation of the corrective terms.

Table 5.4: Implied Black-Scholes volatilities (%) for the Monte Carlo simulations (execution time: 1h31m44s) and the approximation AF(x_{avg}) expressed as a function of strikes for $T = 3Y$.

Strikes	0.35	0.50	0.55	0.70	0.80	0.90	1	1.10	1.25	1.45	1.75	2.05	2.70
MC	39.08	34.73	33.59	30.74	29.19	27.84	26.65	25.57	24.16	22.55	20.57	18.97	16.32
MC-	39.04	34.71	33.58	30.74	29.19	27.84	26.64	25.57	24.15	22.54	20.55	18.94	15.58
MC+	39.11	34.74	33.60	30.75	29.20	27.85	26.65	25.58	24.16	22.55	20.58	19.00	16.79
AF(x_{avg})	39.13	34.76	33.62	30.76	29.20	27.85	26.65	25.58	24.15	22.52	20.49	18.82	15.97

Table 5.5: Implied Black-Scholes volatilities (%) for the Monte Carlo simulations (execution time: 2h29m18s) and the approximation AF(x_{avg}) expressed as a function of strikes for $T = 5Y$.

Strikes	0.25	0.40	0.50	0.60	0.75	0.85	1	1.15	1.35	1.60	2.05	2.50	3.60
MC	41.27	36.15	33.81	31.93	29.68	28.44	26.86	25.52	24.01	22.45	20.25	18.53	15.59
MC-	41.21	36.12	33.79	31.91	29.67	28.43	26.85	25.51	24.01	22.45	20.23	18.48	ND
MC+	41.33	36.18	33.82	31.94	29.69	28.45	26.86	25.53	24.02	22.46	20.27	18.57	16.76
AF(x_{avg})	41.27	36.16	33.82	31.94	29.69	28.45	26.87	25.53	24.03	22.46	20.24	18.51	15.44

5.6.2.1 Proof of Lemma 5.2.3.1

We proceed by induction. One needs the following technical result:

Lemma 5.6.2.1. *Let $(M_t)_{t\in[0,T]}$ be a square integrable and predictable process, $(f_t)_{t\in[0,T]}$ be a measurable and bounded deterministic function and $\varphi \in C_0^\infty(\mathbb{R})$. Then, we have:*

$$\mathbb{E}\Big(\varphi(\int_0^T f_t \mathrm{d}W_t)\int_0^T M_t \mathrm{d}W_t\Big) = \mathbb{E}\Big(\varphi^{(1)}(\int_0^T f_t \mathrm{d}W_t)\int_0^T f_t M_t \mathrm{d}t\Big),$$
$$\mathbb{E}\Big(\varphi(\int_0^T f_t \mathrm{d}W_t)\int_0^T M_t \mathrm{d}B_t\Big) = \mathbb{E}\Big(\varphi^{(1)}(\int_0^T f_t \mathrm{d}W_t)\int_0^T \rho_t f_t M_t \mathrm{d}t\Big)$$

Proof. These results directly come from the duality relationship of the Malliavin calculus (see Lemma 1.2.1 in [Nualart 2006]). □

If $N = 1$ and $I_N = 0$, there is nothing to prove. If $N = 1$ and $I_N \in \{1,2\}$, Lemma 5.6.2.1 is a particular case of Lemma 5.2.3.1 noting that $\forall i \in \mathbb{N}$, $\mathbb{E}\Big(\varphi^{(i)}(\int_0^T f_t \mathrm{d}W_t)\Big) = \mathcal{G}_i^{\varphi}(\int_0^T f_t \mathrm{d}W_t)$, thanks to the regularity of φ. Suppose that the formula (5.18) is true for $N \geq 2$. Then apply Lemma 5.6.2.1 if $I_{N+1} \in \{1,2\}$ to obtain:

$$\begin{aligned}
&\mathbb{E}\Big(\varphi(\int_0^T f_t \mathrm{d}W_t)\int_0^T l_{N+1,t_{N+1}}\int_0^{t_{N+1}} l_{N,t_N}\dots\int_0^{t_2} l_{1,t_1}\mathrm{d}W_{t_1}^{I_1}\dots\mathrm{d}W_{t_N}^{I_N}\mathrm{d}W_{t_{N+1}}^{I_{N+1}}\Big)\\
=&\mathbb{E}\Big(\varphi^{(\mathbb{1}_{I_{N+1}\neq 0})}(\int_0^T f_t \mathrm{d}W_t)\int_0^T \widehat{l}_{N+1,t_{N+1}}\int_0^{t_{N+1}} l_{N,t_N}\int_0^{t_N}\dots\int_0^{t_2} l_{1,t_1}\mathrm{d}W_{t_1}^{I_1}\dots\mathrm{d}W_{t_N}^{I_N}\mathrm{d}t_{N+1}\Big)\\
=&\mathbb{E}\Big(\varphi^{(\mathbb{1}_{I_{N+1}\neq 0})}(\int_0^T f_t \mathrm{d}W_t)\int_0^T (\int_{t_N}^T \widehat{l}_{N+1,s}\mathrm{d}s) l_{N,t_N}\int_0^{t_N}\dots\int_0^{t_2} l_{1,t_1}\mathrm{d}W_{t_1}^{I_1}\dots\mathrm{d}W_{t_N}^{I_N}\Big),
\end{aligned}$$

where at the last equality we have used the fact that $\int_0^T g_t Z_t \mathrm{d}t = \int_0^T (\int_t^T g_s \mathrm{d}s)\mathrm{d}Z_t$ for any continuous semi-martingale Z starting from 0 and any measurable and bounded deterministic function g (apply the Itô's formula to the product $(\int_t^T g_s \mathrm{d}s)Z_t$). We conclude without difficulty with the induction hypothesis and leave the details to the reader.

Table 5.6: Implied Black-Scholes volatilities (%) for the Monte Carlo simulations (execution time: 4h54m27s) and the approximation AF(x_{avg}) expressed as a function of strikes for $T = 10Y$.

Strikes	0.15	0.25	0.35	0.50	0.65	0.80	1	1.20	1.50	1.95	2.75	3.65	6.30
MC	44.71	39.62	36.40	33.11	30.77	28.97	27.09	25.59	23.81	21.79	19.28	17.33	15.31
MC-	44.60	39.56	36.36	33.08	30.75	28.95	27.07	25.58	23.80	21.78	19.25	17.22	ND
MC+	44.83	39.69	36.44	33.14	30.79	28.98	27.10	25.60	23.82	21.81	19.31	17.43	16.28
AF($\mathtt{x_{avg}}$)	44.69	39.63	36.41	33.12	30.78	28.98	27.10	25.61	23.83	21.81	19.28	17.30	13.70

5.6.2.2 Calculus of the corrective terms

We recall our order 3 approximation:

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

We compute each correction term separately, and pay attention to the different nature contributions in these corrections (pure local volatility part, pure stochastic volatility part and both local and stochastic part).

$\triangleright$**Step 1: contribution with $X_{1,T}$.** Apply the Lemma 5.2.3.1 to $\varphi(\cdot) = h^{(1)}(x_0 - \frac{1}{2}\int_0^T \sigma_t^2 v_t \mathrm{d}t + \cdot)$, $f_t = \sigma_t\lambda_t$ and:

$$X_{1,T} = \int_0^T \left(\sigma_{t_2}^{(1)}\lambda_{t_2}\int_0^{t_2}\sigma_{t_1}\lambda_{t_1}(\mathrm{d}W_{t_1} - \frac{\sigma_{t_1}\lambda_{t_1}}{2}\mathrm{d}t_1) + \frac{\sigma_{t_2}}{2\lambda_{t_2}}\int_0^{t_2}\xi_{t_1}\lambda_{t_1}\mathrm{d}B_{t_1}\right)(\mathrm{d}W_{t_2} - \sigma_{t_2}\lambda_{t_2}\mathrm{d}t_2), \tag{5.74}$$

to get:

$$\mathbb{E}[h^{(1)}(X_T^P)X_{1,T}] = C_{1,T}^l[\mathcal{G}_3^h(X_T^P) - \frac{3}{2}\mathcal{G}_2^h(X_T^P) + \frac{1}{2}\mathcal{G}_1^h(X_T^P)] + \frac{C_{1,T}^s}{2}[\mathcal{G}_3^h(X_T^P) - \mathcal{G}_2^h(X_T^P)], \tag{5.75}$$

where:

$$C_{1,T}^s = \omega(\rho\xi\sigma v, \sigma^2)_0^T, \qquad C_{1,T}^l = \omega(\sigma^2 v, \sigma\sigma^{(1)}v)_0^T.$$

$\triangleright$**Step 2: contribution with $X_{2,T}$.** In view of (5.14)-(5.12)-(5.13)-(5.15)-(5.16), we have:

$$\begin{aligned}\frac{X_{2,T}}{2} =& \frac{1}{2}\int_0^T \lambda_t[(X_t^P - x_0)^2\sigma_t^{(2)} + 2X_{1,t}\sigma_t^{(1)}](\mathrm{d}W_t - \sigma_t\lambda_t\mathrm{d}t) + \frac{1}{2}\int_0^T (X_t^P - x_0)V_{1,t}\frac{\sigma_t^{(1)}}{\lambda_t}(\mathrm{d}W_t - 2\sigma_t\lambda_t\mathrm{d}t)\\ &- \frac{1}{2}\int_0^T (X_t^P - x_0)^2(\sigma_t^{(1)})^2 v_t\mathrm{d}t + \frac{1}{4}\int_0^T \frac{V_{2,t}}{\lambda_t}\sigma_t(\mathrm{d}W_t - \sigma_t\lambda_t\mathrm{d}t) - \frac{1}{8}\int_0^T \frac{V_{1,t}^2}{\lambda_t^3}\sigma_t\mathrm{d}W_t,\end{aligned}$$

where applying the Itô's formula:

$$\frac{(X_t^P - x_0)^2}{2} = \int_0^t (\int_0^{t_2}\sigma_{t_1}\lambda_{t_1}(\mathrm{d}W_{t_1} - \frac{\sigma_{t_1}\lambda_{t_1}}{2}\mathrm{d}t_1))\sigma_{t_2}\lambda_{t_2}(\mathrm{d}W_{t_2} - \frac{\sigma_{t_2}\lambda_{t_2}}{2}\mathrm{d}t_2) + \frac{1}{2}\int_0^t \sigma_{t_1}^2 v_{t_1}\mathrm{d}t_1, \tag{5.76}$$

$$\begin{aligned}(X_t^P - x_0)V_{1,t} =& \int_0^t (\int_0^{t_2}\xi_{s_1}\lambda_{s_1}\mathrm{d}B_{s_1})\sigma_{t_2}\lambda_{t_2}(\mathrm{d}W_{t_2} - \frac{\sigma_{t_2}\lambda_{t_2}}{2}\mathrm{d}t_2)\\ &+ \int_0^t (\int_0^{t_2}\sigma_{t_1}\lambda_{t_1}(\mathrm{d}W_{t_1} - \frac{\sigma_{t_1}\lambda_{t_1}}{2}\mathrm{d}t_1))\xi_{s_2}\lambda_{s_2}\mathrm{d}B_{s_2} + \int_0^t \xi_{t_1}\rho_{t_1}\sigma_{t_1}v_{t_1}\mathrm{d}t_1,\end{aligned} \tag{5.77}$$

$$V_{2,t} = \int_0^t (\int_0^{t_2}\xi_{s_1}\lambda_{s_1}\mathrm{d}B_{s_1})\frac{\xi_{s_2}}{\lambda_{s_2}}\mathrm{d}B_{s_2}. \tag{5.78}$$

(5.74)-(5.76)-(5.77)-(5.78) and applications of Lemma 5.2.3.1 allow to obtain:

$$
\begin{aligned}
&\mathbb{E}[h^{(1)}(X_T^P)\frac{X_{2,T}}{2}]+\frac{1}{8}\mathbb{E}[h^{(2)}(X_T^P)\int_0^T\frac{V_{1,t}^2}{v_t}\sigma_t^2\mathrm{d}t] \qquad (5.79)\\
=&C^l_{3b,T}[\mathcal{G}_4^h(X_T^P)-2\mathcal{G}_3^h(X_T^P)+\frac{5\mathcal{G}_2^h(X_T^P)}{4}-\frac{\mathcal{G}_1^h(X_T^P)}{4}]+\frac{C^l_{2b,T}}{2}[\mathcal{G}_2^h(X_T^P)-\mathcal{G}_1^h(X_T^P)]\\
&+C^l_{4,T}[\mathcal{G}_4^h(X_T^P)-\frac{5\mathcal{G}_3^h(X_T^P)}{2}+2\mathcal{G}_2^h(X_T^P)-\frac{\mathcal{G}_1^h(X_T^P)}{2}]+C^{ls}_{4,T}[\frac{\mathcal{G}_4^h(X_T^P)}{2}-\mathcal{G}_3^h(X_T^P)+\frac{\mathcal{G}_2^h(X_T^P)}{2}]\\
&+(C^{ls}_{2,T}+C^{ls}_{3,T})[\frac{\mathcal{G}_4^h(X_T^P)}{2}-\frac{5\mathcal{G}_3^h(X_T^P)}{4}+\frac{\mathcal{G}_2^h(X_T^P)}{2}]+C^{ls}_{1,T}[\frac{\mathcal{G}_2^h(X_T^P)}{2}-\mathcal{G}_1^h(X_T^P)]\\
&+C^l_{3a,T}[-\mathcal{G}_3^h(X_T^P)+\mathcal{G}_2^h(X_T^P)-\frac{\mathcal{G}_1^h(X_T^P)}{4}]-\frac{C^l_{2a,T}}{2}\mathcal{G}_1^h(X_T^P)+\frac{C^s_{2,T}}{4}[\mathcal{G}_4^h(X_T^P)-\mathcal{G}_3^h(X_T^P)],
\end{aligned}
$$

where:

$$
\begin{aligned}
&C^l_{2a,T}=\omega(\sigma^2v,(\sigma^{(1)})^2v)_0^T, && C^l_{2b,T}=\omega(\sigma^2v,\sigma\sigma^{(2)}v)_0^T, && C^l_{3a,T}=\omega(\sigma^2v,\sigma^2v,(\sigma^{(1)})^2v)_0^T,\\
&C^l_{3b,T}=\omega(\sigma^2v,\sigma^2v,\sigma\sigma^{(2)}v)_0^T, && C^l_{4,T}=\omega(\sigma^2v,\sigma\sigma^{(1)}v,\sigma\sigma^{(1)}v)_0^T, && C^s_{2,T}=\omega(\rho\xi\sigma v,\rho\xi\sigma,\sigma^2)_0^T,\\
&C^{ls}_{1,T}=\omega(\rho\xi\sigma v,\sigma\sigma^{(1)})_0^T, && C^{ls}_{2,T}=\omega(\rho\xi\sigma v,\sigma^2v,\sigma\sigma^{(1)})_0^T, && C^{ls}_{3,T}=\omega(\sigma^2v,\rho\xi\sigma v,\sigma\sigma^{(1)})_0^T,\\
&C^{ls}_{4,T}=\omega(\rho\xi\sigma v,\sigma^2,\sigma\sigma^{(1)}v)_0^T.
\end{aligned}
$$

▷**Step 3: contribution with $X_{1,T}^2$.** Starting from (5.14) and applying the Itô's formula we have:

$$
\begin{aligned}
\frac{1}{2}X_{1,T}^2=&\int_0^T X_{1,t}((X_t^P-x_0)\sigma_t^{(1)}\lambda_t+\frac{V_{1,t}\sigma_t}{2\lambda_t})(\mathrm{d}W_t-\sigma_t\lambda_t\mathrm{d}t) \qquad (5.80)\\
&+\frac{1}{2}\int_0^T((X_t^P-x_0)^2(\sigma_t^{(1)})^2v_t+(X_t^P-x_0)V_{1,t}\sigma_t\sigma_t^{(1)})\mathrm{d}t+\frac{1}{8}\int_0^T\frac{V_{1,t}^2}{v_t}\sigma_t^2\mathrm{d}t,
\end{aligned}
$$

where:

$$
\begin{aligned}
X_{1,t}(X_t^P-x_0)=&\int_0^t[(X_{t_1}^P-x_0)^2\sigma_{t_1}^{(1)}\lambda_{t_1}+\frac{(X_{t_1}^P-x_0)V_{1,t_1}\sigma_{t_1}}{2\lambda_{t_1}}](\mathrm{d}W_{t_1}-\sigma_{t_1}\lambda_{t_1}\mathrm{d}t_1)\\
&+\int_0^t X_{1,t_1}\sigma_{t_1}\lambda_{t_1}(\mathrm{d}W_{t_1}-\frac{\sigma_{t_1}\lambda_{t_1}}{2}\mathrm{d}t_1)+\int_0^t[(X_{t_1}^P-x_0)\sigma_{t_1}\sigma_{t_1}^{(1)}v_{t_1}+\frac{V_{1,t_1}}{2}\sigma_{t_1}^2]\mathrm{d}t_1, \qquad (5.81)\\
X_{1,t}V_{1,t}=&\int_0^T[(X_{t_1}^P-x_0)V_{1,t_1}\sigma_{t_1}^{(1)}\lambda_{t_1}+\frac{V_{1,t_1}^2\sigma_{t_1}}{2\lambda_{t_1}}](\mathrm{d}W_{t_1}-\sigma_{t_1}\lambda_{t_1}\mathrm{d}t_1)+\int_0^t X_{1,t_1}\xi_{t_1}\lambda_{t_1}\mathrm{d}B_{t_1}\\
&+\int_0^t\rho_{t_1}\xi_{t_1}[(X_{t_1}^P-x_0)\sigma_{t_1}^{(1)}v_{t_1}+\frac{V_{1,t_1}\sigma_{t_1}}{2}]\mathrm{d}t_1, \qquad (5.82)\\
V_{1,t}^2=&2\int_0^t(\int_0^t\xi_{t_1}\lambda_{t_1}\mathrm{d}B_{t_1})\xi_{t_2}\lambda_{t_2}\mathrm{d}B_{t_2}+\int_0^t\xi_t^2v_t\mathrm{d}t \qquad (5.83)
\end{aligned}
$$

From Lemma 5.2.3.1 and (5.81)-(5.76)-(5.74)-(5.77) it follows that:

$$
\begin{aligned}
&\mathbb{E}[h^{(2)}(X_T^P)\int_0^T X_{1,t}(X_t^P-x_0)\sigma_t^{(1)}\lambda_t(\mathrm{d}W_t-\sigma_t\lambda_t\mathrm{d}t)] \qquad (5.84)\\
=&C^l_{11a,T}[2\mathcal{G}_6^h(X_T^P)-6\mathcal{G}_5^h(X_T^P)+\frac{13\mathcal{G}_4^h(X_T^P)}{2}-3\mathcal{G}_3^h(X_T^P)+\frac{\mathcal{G}_2^h(X_T^P)}{2}]+C^l_{4,T}[\mathcal{G}_4^h(X_T^P)-2\mathcal{G}_3^h(X_T^P)+\mathcal{G}_2^h(X_T^P)]\\
&+(C^{ls}_{7a,T}+C^{ls}_{7b,T})[\frac{\mathcal{G}_6^h(X_T^P)}{2}-\frac{5\mathcal{G}_5^h(X_T^P)}{4}+\mathcal{G}_4^h(X_T^P)-\frac{\mathcal{G}_3^h(X_T^P)}{4}]+C^{ls}_{4,T}[\frac{\mathcal{G}_4^h(X_T^P)}{2}-\mathcal{G}_3^h(X_T^P)+\frac{\mathcal{G}_2^h(X_T^P)}{2}]
\end{aligned}
$$

$$+C^l_{11b,T}[\mathcal{G}^h_6(X^P_T)-3\mathcal{G}^h_5(X^P_T)+\frac{13\mathcal{G}^h_4(X^P_T)}{4}-\frac{3\mathcal{G}^h_3(X^P_T)}{2}+\frac{\mathcal{G}^h_2(X^P_T)}{4}]+C^{ls}_{7c,T}[\frac{\mathcal{G}^h_6(X^P_T)}{2}-\frac{5\mathcal{G}^h_5(X^P_T)}{4}+\mathcal{G}^h_4(X^P_T)$$
$$-\frac{\mathcal{G}^h_3(X^P_T)}{4}]+C^l_{4,T}[\mathcal{G}^h_4(X^P_T)-\frac{3\mathcal{G}^h_3(X^P_T)}{2}+\frac{\mathcal{G}^h_2(X^P_T)}{2}]+\frac{C^{ls}_{4,T}}{2}[\mathcal{G}^h_4(X^P_T)-\mathcal{G}^h_3(X^P_T)].$$

where:

$$C^l_{11a,T}=\omega(\sigma^2 v,\sigma^2 v,(\sigma\sigma^{(1)})v,(\sigma\sigma^{(1)})v)^T_0,\qquad C^l_{11b,T}=\omega(\sigma^2 v,(\sigma\sigma^{(1)})v,\sigma^2 v,(\sigma\sigma^{(1)})v)^T_0,$$
$$C^{ls}_{7a,T}=\omega(\rho\xi\sigma v,\sigma^2 v,\sigma^2,\sigma\sigma^{(1)}v)^T_0,\qquad C^{ls}_{7b,T}=\omega(\sigma^2 v,\rho\xi\sigma v,\sigma^2,\sigma\sigma^{(1)}v)^T_0,$$
$$C^{ls}_{7c,T}=\omega(\rho\xi\sigma v,\sigma^2,\sigma^2 v,\sigma\sigma^{(1)}v)^T_0.$$

Similarly, using Lemma 5.2.3.1 and (5.82)-(5.77)-(5.83)-(5.74), we have:

$$\frac{1}{2}\mathbb{E}[h^{(2)}(X^P_T)\int_0^T\frac{X_{1,t}V_{1,t}\sigma_t}{\lambda_t}(\mathrm{d}W_t-\sigma_t\lambda_t\mathrm{d}t)] \tag{5.85}$$
$$=(C^{ls}_{7d,T}+C^{ls}_{7e,T})[\frac{\mathcal{G}^h_6(X^P_T)}{2}-\frac{5\mathcal{G}^h_5(X^P_T)}{4}+\mathcal{G}^h_4(X^P_T)-\frac{\mathcal{G}^h_3(X^P_T)}{4}]+C^{ls}_{5,T}[\frac{\mathcal{G}^h_4(X^P_T)}{2}-\mathcal{G}^h_3(X^P_T)+\frac{\mathcal{G}^h_2(X^P_T)}{2}]$$
$$+\frac{C^s_{4a,T}}{2}[\mathcal{G}^h_6(X^P_T)-2\mathcal{G}^h_5(X^P_T)+\mathcal{G}^h_4(X^P_T)]+\frac{C^s_{3,T}}{4}[\mathcal{G}^h_4(X^P_T)-2\mathcal{G}^h_3(X^P_T)+\mathcal{G}^h_2(X^P_T)]$$
$$+C^{ls}_{7f,T}[\frac{\mathcal{G}^h_6(X^P_T)}{2}-\frac{5\mathcal{G}^h_5(X^P_T)}{4}+\mathcal{G}^h_4(X^P_T)-\frac{\mathcal{G}^h_3(X^P_T)}{4}]+\frac{C^s_{4b,T}}{4}[\mathcal{G}^h_6(X^P_T)-2\mathcal{G}^h_5(X^P_T)+\mathcal{G}^h_4(X^P_T)]$$
$$+C^{ls}_{6,T}[\frac{\mathcal{G}^h_4(X^P_T)}{2}-\frac{3\mathcal{G}^h_3(X^P_T)}{4}+\frac{\mathcal{G}^h_2(X^P_T)}{4}]+\frac{C^s_{2,T}}{4}[\mathcal{G}^h_4(X^P_T)-\mathcal{G}^h_3(X^P_T)].$$

where:

$$C^s_{2,T}=\omega(\rho\xi\sigma v,\rho\xi\sigma,\sigma^2)^T_0,\qquad C^s_{3,T}=\omega(\xi^2 v,\sigma^2,\sigma^2)^T_0,$$
$$C^s_{4a,T}=\omega(\rho\xi\sigma v,\xi\rho v\sigma,\sigma^2,\sigma^2)^T_0,\qquad C^s_{4b,T}=\omega(\rho\xi v\sigma,\sigma^2,\rho\xi\sigma v,\sigma^2)^T_0,$$
$$C^{ls}_{5,T}=\omega(\rho\xi\sigma v,\sigma\sigma^{(1)}v,\sigma^2)^T_0,\qquad C^{ls}_{6,T}=\omega(\sigma^2 v,\rho\xi\sigma^{(1)}v,\sigma^2)^T_0,$$
$$C^{ls}_{7d,T}=\omega(\rho\xi\sigma v,\sigma^2 v,\sigma\sigma^{(1)}v,\sigma^2)^T_0,\qquad C^{ls}_{7e,T}=\omega(\sigma^2 v,\rho\xi\sigma v,\sigma\sigma^{(1)}v,\sigma^2)^T_0,$$
$$C^{ls}_{7f,T}=\omega(\sigma^2 v,\sigma\sigma^{(1)}v,\rho\xi\sigma v,\sigma^2)^T_0.$$

Then using again Lemma 5.2.3.1 and (5.76)-(5.77) it comes:

$$\frac{1}{2}\mathbb{E}[h^{(2)}(X^P_T)\int_0^T((X^P_t-x_0)^2(\sigma^{(1)}_t)^2 v_t+(X^P_t-x_0)V_{1,t}\sigma_t\sigma^{(1)}_t)\mathrm{d}t] \tag{5.86}$$
$$=C^l_{3a,T}[\mathcal{G}^h_4(X^P_T)-\mathcal{G}^h_3(X^P_T)+\frac{\mathcal{G}^h_2(X^P_T)}{4}]+\frac{C^l_{2a,T}}{2}\mathcal{G}^h_2(X^P_T)+(C^{ls}_{2,T}+C^{ls}_{3,T})[\frac{\mathcal{G}^h_4(X^P_T)}{2}-\frac{\mathcal{G}^h_3(X^P_T)}{4}]+\frac{C^{ls}_{1,T}}{2}\mathcal{G}^h_2(X^P_T).$$

Finally, we sum the contributions (5.84)-(5.85)-(5.86) to obtain in view of (5.80):

$$\frac{1}{2}\mathbb{E}[h^{(2)}(X^P_T)X^2_{1,T}]-\frac{1}{8}\mathbb{E}[h^{(2)}(X^P_T)\int_0^T\frac{V^2_{1,t}}{v_t}\sigma^2_t\mathrm{d}t] \tag{5.87}$$
$$=(2C^l_{11a,T}+C^l_{11b,T})[\mathcal{G}^h_6(X^P_T)-3\mathcal{G}^h_5(X^P_T)+\frac{13\mathcal{G}^h_4(X^P_T)}{4}-\frac{3\mathcal{G}^h_3(X^P_T)}{2}+\frac{\mathcal{G}^h_2(X^P_T)}{4}]$$
$$+C^l_{4,T}[2\mathcal{G}^h_4(X^P_T)-\frac{7\mathcal{G}^h_3(X^P_T)}{2}+\frac{3\mathcal{G}^h_2(X^P_T)}{2}]+C^{ls}_{4,T}[\mathcal{G}^h_4(X^P_T)-\frac{3\mathcal{G}^h_3(X^P_T)}{2}+\frac{\mathcal{G}^h_2(X^P_T)}{2}]$$
$$+(C^{ls}_{7a,T}+C^{ls}_{7b,T}+C^{ls}_{7c,T}+C^{ls}_{7d,T}+C^{ls}_{7e,T}+C^{ls}_{7f,T})[\frac{\mathcal{G}^h_6(X^P_T)}{2}-\frac{5\mathcal{G}^h_5(X^P_T)}{4}+\mathcal{G}^h_4(X^P_T)-\frac{\mathcal{G}^h_3(X^P_T)}{4}]$$

$$+C^{ls}_{5,T}[\frac{\mathcal{G}^h_4(X^P_T)}{2}-\mathcal{G}^h_3(X^P_T)+\frac{\mathcal{G}^h_2(X^P_T)}{2}]+\frac{C^s_{3,T}}{4}[\mathcal{G}^h_4(X^P_T)-2\mathcal{G}^h_3(X^P_T)+\mathcal{G}^h_2(X^P_T)]+\frac{C^s_{2,T}}{4}[\mathcal{G}^h_4(X^P_T)-\mathcal{G}^h_3(X^P_T)]$$
$$+(2C^s_{4a,T}+C^s_{4b,T})[\frac{\mathcal{G}^h_6(X^P_T)}{4}-\frac{\mathcal{G}^h_5(X^P_T)}{2}+\frac{\mathcal{G}^h_4(X^P_T)}{4}]+C^{ls}_{6,T}[\frac{\mathcal{G}^h_4(X^P_T)}{2}-\frac{3\mathcal{G}^h_3(X^P_T)}{4}+\frac{\mathcal{G}^h_2(X^P_T)}{4}]$$
$$+C^l_{3a,T}[\mathcal{G}^h_4(X^P_T)-\mathcal{G}^h_3(X^P_T)+\frac{\mathcal{G}^h_2(X^P_T)}{4}]+\frac{C^l_{2a,T}}{2}\mathcal{G}^h_2(X^P_T)+(C^{ls}_{2,T}+C^{ls}_{3,T})[\frac{\mathcal{G}^h_4(X^P_T)}{2}-\frac{\mathcal{G}^h_3(X^P_T)}{4}]+\frac{C^{ls}_{1,T}}{2}\mathcal{G}^h_2(X^P_T).$$

▷**Step 4: some mathematical reductions.** There are some relations between the expansion coefficients.The reader can easily verify that:

$$\frac{(C^s_{1,T})^2}{2}=2C^s_{4a,T}+C^s_{4b,T}, \qquad \frac{(C^l_{1,T})^2}{2}=2C^l_{11a,T}+C^l_{11b,T},$$
$$C^l_{1,T}C^s_{1,T}=C^{ls}_{7a,T}+C^{ls}_{7b,T}+C^{ls}_{7c,T}+C^{ls}_{7d,T}+C^{ls}_{7e,T}+C^{ls}_{7f,T}.$$

The first identity is proved in [Benhamou 2010b, section 5.4] and the others are similar. In addition we set $C^l_{2,T}=C^l_{2a,T}+C^l_{2b,T}$ and $C^l_{3,T}=C^l_{3a,T}+C^l_{3b,T}$.

▷**Final step.** Taking advantage of the above simplifications and gathering the different contributions (5.75)-(5.79)-(5.87) of steps $1-2-3$, we obtain the announced formula (5.9), putting together the corrective terms according to the order of the Greeks.

5.6.3 Computations of derivatives of $\mathrm{Call}^{\mathrm{BS}}$ w.r.t the log spot and the volatility

In the following Proposition, we give the derivatives at any order of $\mathrm{Call}^{\mathrm{BS}}$ w.r.t. x:

Proposition 5.6.3.1. *Let $x,k\in\mathbb{R}$ and $y>0$. For any integer $n\geq 1$, we have:*

$$\partial^n_{x^n}\mathrm{Call}^{\mathrm{BS}}(x,y,k)=e^x\mathcal{N}(d_1(x,y,k))+\mathbb{1}_{n\geq 2}\, e^x\mathcal{N}'(d_1(x,y,k))\sum_{j=1}^{n-1}\binom{n-1}{j}(-1)^{j-1}\frac{H_{j-1}(d_1(x,y,k))}{y^{\frac{j}{2}}},$$

where $(H_j)_{j\in\mathbb{N}}$ are the Hermite polynomials defined by $H_j(x)=(-1)^je^{x^2/2}\partial^j_{x^j}(e^{-x^2/2})$ $\forall(j,x)\in\mathbb{N}\times\mathbb{R}$.

In the next Proposition, we provide the formulas of the $\mathrm{Vega}^{\mathrm{BS}}$ and the $\mathrm{Vomma}^{\mathrm{BS}}$:

Proposition 5.6.3.2. *Let $x,k\in\mathbb{R}$, $\sigma>0$ and $T>0$. We have:*

$$\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)=\partial_\sigma\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2T,k)=e^x\sqrt{T}\mathcal{N}'(d_1(x,\sigma^2T,k)),$$
$$\mathrm{Vomma}^{\mathrm{BS}}(x,\sigma^2T,k)=\partial_\sigma\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)=\frac{\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)}{\sigma}d_1(x,\sigma^2T,k)d_2(x,\sigma^2T,k)$$
$$=\frac{\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)}{\sigma}[\frac{(x-k)^2}{\sigma^2T}-\frac{\sigma^2T}{4}].$$

We finally state relations (obtained with Mathematica) between the derivatives of $\mathrm{Call}^{\mathrm{BS}}$ w.r.t. x and the $\mathrm{Vega}^{\mathrm{BS}}$ and the $\mathrm{Vomma}^{\mathrm{BS}}$:

Proposition 5.6.3.3. *Let $x,k\in\mathbb{R}$, $\sigma>0$ and $T>0$. We have:*

$$(\partial^2_{x^2}-\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2T,k)=\frac{\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)}{\sigma T},$$
$$(\partial^3_{x^3}-\frac{3}{2}\partial^2_{x^2}+\frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2T,k)=-\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)\frac{(x-k)}{\sigma^3T^2},$$
$$(\partial^4_{x^4}-2\partial^3_{x^3}+\frac{5}{4}\partial^2_{x^2}-\frac{1}{4}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2T,k)=\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)[\frac{(x-k)^2}{\sigma^5T^3}-\frac{1}{\sigma^3T^2}],$$

$$
\begin{aligned}
&(3\partial^4_{x^4} - 6\partial^3_{x^3} + \frac{7}{2}\partial^2_{x^2} - \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2 T,k) = \mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2 T,k)[3\frac{(x-k)^2}{\sigma^5 T^3} - \frac{3}{\sigma^3 T^2} - \frac{1}{4\sigma T}],\\
&(\frac{1}{2}\partial^6_{x^6} - \frac{3}{2}\partial^5_{x^5} + \frac{13}{8}\partial^4_{x^4} - \frac{3}{4}\partial^3_{x^3} + \frac{1}{8}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2 T,k)\\
=&\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2 T,k)[-3\frac{(x-k)^2}{\sigma^7 T^4} + \frac{1}{8\sigma^3 T^2} + \frac{3}{2\sigma^5 T^3}] + \frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x,\sigma^2 T,k)\frac{(x-k)^2}{\sigma^6 T^4},\\
&(\frac{1}{2}\partial^3_{x^3} - \frac{1}{2}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2 T,k) = \mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2 T,k)[-\frac{(x-k)}{2\sigma^3 T^2} + \frac{1}{4\sigma T}],\\
&(\frac{1}{2}\partial^4_{x^4} - \frac{1}{2}\partial^3_{x^3})\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2 T,k) = \mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2 T,k)[\frac{(x-k)^2}{2\sigma^5 T^3} + \frac{1}{8\sigma T} - \frac{1}{2\sigma^3 T^2} - \frac{(x-k)}{2\sigma^3 T^2}],\\
&(\frac{1}{4}\partial^4_{x^4} - \frac{1}{2}\partial^3_{x^3} + \frac{1}{4}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2 T,k) = \mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2 T,k)[\frac{(x-k)^2}{4\sigma^5 T^3} - \frac{1}{16\sigma T} - \frac{1}{4\sigma^3 T^2}],\\
&(\partial^4_{x^4} - \frac{3}{2}\partial^3_{x^3} + \frac{1}{2}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2 T,k) = \mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2 T,k)[\frac{(x-k)^2}{\sigma^5 T^3} - \frac{1}{\sigma^3 T^2} - \frac{(x-k)}{2\sigma^3 T^2}],\\
&(\frac{3}{2}\partial^4_{x^4} - \frac{5}{2}\partial^3_{x^3} + \partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2 T,k) = \mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2 T,k)[\frac{3(x-k)^2}{2\sigma^5 T^3} - \frac{3}{2\sigma^3 T^2} - \frac{(x-k)}{2\sigma^3 T^2} - \frac{1}{8\sigma T}],\\
&(\frac{1}{2}\partial^6_{x^6} - \frac{5}{4}\partial^5_{x^5} + \partial^4_{x^4} - \frac{1}{4}\partial^3_{x^3})\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2 T,k)\\
=&\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2 T,k)[-3\frac{(x-k)^2}{\sigma^7 T^4} + \frac{3}{2\sigma^5 T^3} + \frac{3(x-k)}{4\sigma^5 T^3} + \frac{1}{8\sigma^3 T^2}]\\
&+\frac{\mathrm{Vomma}^{\mathrm{BS}}(x,\sigma^2 T,k)}{2}[\frac{(x-k)^2}{\sigma^6 T^4} - \frac{(x-k)}{2\sigma^4 T^3}],\\
&(\frac{1}{8}\partial^6_{x^6} - \frac{1}{4}\partial^5_{x^5} + \frac{1}{8}\partial^4_{x^4})\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2 T,k)\\
=&\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2 T,k)[-3\frac{(x-k)^2}{4\sigma^7 T^4} + \frac{3}{8\sigma^5 T^3} + \frac{3(x-k)}{8\sigma^5 T^3}] + \frac{\mathrm{Vomma}^{\mathrm{BS}}(x,\sigma^2 T,k)}{2}[-\frac{(x-k)}{2\sigma^3 T^2} + \frac{1}{4\sigma T}]^2.
\end{aligned}
$$

5.6.4 Applications of the implied volatility expansion at mid-point for time-independent local and stochastic volatility models with CIR-type variance

We specify in this section the form of the implied volatility approximation at mid point when considering the time-independent local and stochastic volatility model with CIR-type variance:

$$
\begin{aligned}
\mathrm{d}X_t &= \sigma(X_t)\sqrt{Y_t}[\mathrm{d}W_t - \frac{\sigma(X_t)\sqrt{Y_t}}{2}\mathrm{d}t], \quad X_0 = x_0,\\
\mathrm{d}Y_t &= \kappa(\theta - Y_t)\mathrm{d}t + \xi\sqrt{Y_t}\mathrm{d}B_t, \quad Y_0 = v_0,\\
\mathrm{d}\langle W,B\rangle_t &= \rho\mathrm{d}t.
\end{aligned}
$$

In view of 5.6.1, we can apply our different price and implied volatility expansion theorems by considering in the various corrective coefficients C (defined in Theorem 5.2.2.1 and Definitions 5.4.1.1 and 5.4.2.1) the time dependent volatility function $\sigma(t,x) = \sigma(x)e^{-\frac{\kappa t}{2}}$, the time dependent deterministic variance function $v_t = v_0 + \theta(e^{\kappa t} - 1)$ and the time-dependent volatility of volatility function $\xi_t = \xi e^{\frac{\kappa t}{2}}$. Thus the coefficients are obtained by simple iterated integrations of exponential functions. Using Mathematica, we derive the following explicit expressions:

Proposition 5.6.4.1. *For $\sigma(t,x) = \sigma(x)e^{-\frac{\kappa t}{2}}$, $v_t = v_0 + \theta(e^{\kappa t} - 1)$, $\xi_t = \xi e^{\frac{\kappa t}{2}}$ and $\rho_t = \rho$, one has:*

$$
\int_0^T v_t e^{-\kappa t}\mathrm{d}t = \bar{v}T, \qquad C^l_{1,T}(x) = \frac{\sigma^3(x)\sigma^{(1)}(x)\bar{v}^2 T^2}{2},
$$

$$
\begin{aligned}
C^l_{2,T}(x) =&\frac{\sigma^2(x)[(\sigma^{(1)})^2+\sigma\sigma^{(2)}](x)\bar{v}^2T^2}{2}, \qquad C^l_{3,T}(x) =C^l_{9,T}(x) = \frac{\sigma^4(x)[(\sigma^{(1)})^2+\sigma\sigma^{(2)}](x)\bar{v}^3T^3}{6},\\
C^l_{4,T}(x) =&\frac{\sigma^4(x)(\sigma^{(1)})^2(x)\bar{v}^3T^3}{6}, \qquad C^l_{10,T}(x) =\frac{\sigma^6(x)(\sigma^{(1)})^2(x)\bar{v}^4T^4}{24},\\
C^l_{5,T}(x) =&[(\sigma^{(1)})^2+\sigma\sigma^{(2)}](x)\bar{v}T, \qquad C^l_{6,T}(x) =\frac{[\sigma\sigma^{(1)}]^2(x)\bar{v}^2T^2}{2},\\
C^s_{1,T}(x) =&\frac{\rho\xi\sigma^3(x)R^s_1T^2}{2}, \qquad C^s_{2,T}(x) =\frac{\rho^2\xi^2\sigma^4(x)R^s_2T^3}{6},\\
C^s_{3,T}(x) =&\frac{\xi^2\sigma^4(x)R^s_3T^3}{6}, \qquad C^{ls}_{1,T}(x) =C^{ls}_{9,T}(x) = \frac{\rho\xi\sigma^2(x)\sigma^{(1)}(x)R^s_1T^2}{2},\\
C^{ls}_{2,T}(x) =&C^{ls}_{5,T}(x) = \frac{\rho\xi\sigma^4(x)\sigma^{(1)}(x)R^{ls}_1T^3}{6}, \qquad C^{ls}_{3,T}(x) =C^{ls}_{6,T}(x) = C^{ls}_{10,T}(x) = \frac{\rho\xi\sigma^4(x)\sigma^{(1)}(x)R^{ls}_2T^3}{6},\\
C^{ls}_{4,T}(x) =&C^{ls}_{8,T}(x) = \frac{\rho\xi\sigma^4(x)\sigma^{(1)}(x)R^{ls}_3T^3}{6},
\end{aligned}
$$

where:

$$
\begin{aligned}
\bar{v} =&(v_0-\theta)\frac{e^{-\kappa T}\left(-1+e^{\kappa T}\right)}{\kappa T}+\theta,\\
R^s_1 =&(v_0-\theta)\frac{e^{-\kappa T}\left(-2\kappa T+2e^{\kappa T}-2\right)}{\kappa^2T^2}+\theta\frac{e^{-\kappa T}\left(2\kappa Te^{\kappa T}-2e^{\kappa T}+2\right)}{\kappa^2T^2},\\
R^s_2 =&(v_0-\theta)\frac{e^{-\kappa T}\left(-3\kappa T(\kappa T+2)+6e^{\kappa T}-6\right)}{\kappa^3T^3}+\theta\frac{e^{-\kappa T}\left(6e^{\kappa T}(\kappa T-2)+6\kappa T+12\right)}{\kappa^3T^3},\\
R^s_3 =&(v_0-\theta)\frac{e^{-2\kappa T}\left(-6e^{\kappa T}\kappa T+3e^{2\kappa T}-3\right)}{\kappa^3T^3}+\theta\frac{e^{-2\kappa T}\left(12e^{\kappa T}+3e^{2\kappa T}(2\kappa T-3)-3\right)}{2\kappa^3T^3},\\
R^{ls}_1 =&\frac{3}{2\kappa^3T^3}\{e^{-2\kappa T}(v_0-\theta)(v_0(3+2\kappa T)-\theta(5+2\kappa T))+2e^{-\kappa T}(\theta^2(4+\kappa T(6+\kappa T))\\
&-\theta v_0(-2+\kappa T(4+\kappa T))-2v_0^2)+\theta^2(4\kappa T-13)+4\theta v_0+v_0^2\},\\
R^{ls}_2 =&\frac{3}{\kappa^3T^3}\{-e^{-2\kappa T}(v_0-\theta)^2+e^{-\kappa T}(\theta^2(-4+\kappa T(-2+\kappa T))-\theta v_0(-2+\kappa T(-4+\kappa T))-2\kappa Tv_0^2)\\
&+\theta^2(5+\kappa T(-4+\kappa T))+2\theta v_0(-2+\kappa T)+v_0^2\},\\
R^{ls}_3 =&\frac{3}{2\kappa^3T^3}\{e^{-2\kappa T}(v_0-\theta)(v_0(3+2\kappa T)-\theta(5+2\kappa T))-4e^{-\kappa T}(v_0-2\theta)^2\\
&+\theta^2(11+2\kappa T(-4+\kappa T))+4\theta v_0(-2+\kappa T)+v_0^2\}.
\end{aligned}
$$

We have in addition the relation:

$$
\bar{v}T\frac{R^s_1T^2}{2}=\frac{(R^{ls}_1+R^{ls}_2+R^{ls}_3)T^3}{6}. \tag{5.88}
$$

Using the relation (5.88), one gets without difficulty:

$$
\begin{aligned}
\frac{C^{ls}_{8,T}(x)}{\bar{\sigma}^3_xT^2}-\frac{3C^{ls}_{4,T}(x)}{2\bar{\sigma}^3_xT^2}-\frac{C^{ls}_{5,T}(x)}{2\bar{\sigma}^3_xT^2}-\frac{C^{ls}_{6,T}(x)}{2\bar{\sigma}^3_xT^2}+\frac{3(C^l_{1,T}C^s_{1,T})(x)}{2\bar{\sigma}^5_xT^3} =&\frac{(C^l_{1,T}C^s_{1,T})(x)}{2\bar{\sigma}^5_xT^3},\\
-\frac{C^{ls}_{4,T}(x)}{8\bar{\sigma}_xT}-\frac{C^{ls}_{5,T}(x)}{8\bar{\sigma}_xT}+\frac{(C^l_{1,T}C^s_{1,T})(x)}{8\bar{\sigma}^3_xT^2} =&\frac{C^{ls}_{6,T}(x)}{8\bar{\sigma}_xT}-\frac{(C^l_{1,T}C^s_{1,T})(x)}{8\bar{\sigma}^3_xT^2},\\
-\frac{(C^{ls}_{2,T}+C^{ls}_{3,T})(x)}{2\bar{\sigma}^3_xT^2}-\frac{C^{ls}_{4,T}(x)}{2\bar{\sigma}^3_xT^2}-\frac{C^{ls}_{6,T}(x)}{4\bar{\sigma}^3_xT^2}+\frac{3(C^l_{1,T}C^s_{1,T})(x)}{4\bar{\sigma}^5_xT^3} =&-\frac{C^{ls}_{6,T}(x)}{4\bar{\sigma}^3_xT^2}-\frac{(C^l_{1,T}C^s_{1,T})(x)}{4\bar{\sigma}^5_xT^3},
\end{aligned}
$$

$$\frac{(C^{ls}_{2,T}+C^{ls}_{3,T})(x)}{\bar{\sigma}^5_x T^3}+\frac{3C^{ls}_{4,T}(x)}{2\bar{\sigma}^5_x T^3}+\frac{C^{ls}_{5,T}(x)}{2\bar{\sigma}^5_x T^3}+\frac{C^{ls}_{6,T}(x)}{2\bar{\sigma}^5_x T^3}-\frac{3(C^{l}_{1,T}C^{s}_{1,T})(x)}{\bar{\sigma}^7_x T^4}=0,$$

$$\frac{C^{ls}_{1,T}(x)}{4\bar{\sigma}_x T}+\frac{C^{ls}_{9,T}(x)}{8\bar{\sigma}_x T}-\frac{C^{ls}_{4,T}(x)+C^{ls}_{5,T}(x)+C^{ls}_{10,T}(x)}{8\bar{\sigma}^3_x T^2}=\frac{(C^{l}_{1,T}C^{s}_{1,T})(x)}{2\bar{\sigma}^5_x T^3},$$

$$-\frac{C^{ls}_{1,T}(x)}{2\bar{\sigma}^3_x T^2}-\frac{C^{ls}_{9,T}(x)}{4\bar{\sigma}^3_x T^2}+3\frac{C^{ls}_{4,T}(x)+C^{ls}_{5,T}(x)+C^{ls}_{10,T}(x)}{4\bar{\sigma}^5_x T^3}=0.$$

Then the above mathematical reductions allow to obtain the following expressions for the coefficients γ and π defined in Definitions 5.4.1.1 and 5.4.2.1:

Proposition 5.6.4.2. *For $\sigma(t,x)=\sigma(x)e^{-\frac{\kappa t}{2}}$, $v_t=v_0+\theta(e^{\kappa t}-1)$, $\xi_t=\xi e^{\frac{\kappa t}{2}}$ and $\rho_t=\rho$, one has:*

$$\begin{aligned}
\gamma_{0a,T}(x_0)=&\sigma(x_0)\sqrt{\bar{v}}\{1+\frac{\rho\xi\sigma(x_0)R^s_1 T}{8\bar{v}}\},\\
\gamma_{1a,T}(x_0)=&\sigma(x_0)\sqrt{\bar{v}}\{-\frac{\sigma^{(1)}(x_0)}{2\sigma(x_0)}-\frac{\rho\xi R^s_1}{4\sigma(x_0)\bar{v}^2}\},\\
\gamma_{0b,T}(x_0)=&\sigma(x_0)\sqrt{\bar{v}}\{\bar{v}T[\frac{\sigma(x_0)\sigma^{(2)}(x_0)}{12}-(\sigma^{(1)})^2(x_0)(\frac{1}{24}+\frac{\sigma^2(x_0)\bar{v}T}{96})]\\
&+\frac{\rho^2\xi^2 T}{\bar{v}^2}[\frac{3(R^s_1)^2}{32\bar{v}}+R^s_2(\frac{\sigma^2(x_0)\bar{v}T}{48}-\frac{1}{12})]-\frac{\xi^2 T R^s_3}{\bar{v}^2}[\frac{1}{24}+\frac{\sigma^2(x_0)\bar{v}T}{96}]\\
&+\frac{\rho\xi\sigma^{(1)}(x_0)T}{\bar{v}}[\frac{R^s_1}{8}+\sigma^2(x_0)\bar{v}T(\frac{R^{ls}_2}{48\bar{v}}-\frac{R^s_1}{32})]\},\\
\gamma_{1b,T}(x_0)=&\sigma(x_0)\sqrt{\bar{v}}\{\frac{\rho^2\xi^2 T}{\bar{v}^2}[\frac{3(R^s_1)^2}{32\bar{v}}-\frac{R^s_2}{12}]-\frac{\rho\xi\sigma^{(1)}(x_0)T}{\bar{v}}[\frac{R^s_1}{16}+\frac{R^{ls}_2}{24\bar{v}}]\},\\
\gamma_{2,T}(x_0)=&\sigma(x_0)\sqrt{\bar{v}}\{\frac{\sigma^{(2)}(x_0)}{6\sigma(x_0)}-\frac{(\sigma^{(1)})^2(x_0)}{12\sigma^2(x_0)}+\frac{\rho^2\xi^2}{\sigma^2(x_0)\bar{v}^3}[\frac{R^s_2}{12}-\frac{3(R^s_1)^2}{16\bar{v}}]+\frac{\xi^2R^s_3}{24\sigma^2(x_0)\bar{v}^3}\},\\
\pi_{1a,T}(x_{avg})=&-\sigma(x_{avg})\sqrt{\bar{v}}\frac{\rho\xi R^s_1}{4\sigma(x_{avg})\bar{v}^2},\\
\pi_{1b,T}(x_{avg})=&\sigma(x_{avg})\sqrt{\bar{v}}\{\frac{\rho^2\xi^2 T}{\bar{v}^2}[\frac{3(R^s_1)^2}{32\bar{v}}-\frac{R^s_2}{12}]+\frac{\rho\xi\sigma^{(1)}(x_{avg})T}{\bar{v}}[\frac{R^s_1}{16}-\frac{R^{ls}_2}{24\bar{v}}]\},\\
\pi_{2,T}(x_{avg})=&\sigma(x_{avg})\sqrt{\bar{v}}\{\frac{\sigma^{(2)}(x_{avg})}{24\sigma(x_{avg})}-\frac{(\sigma^{(1)})^2(x_{avg})}{12\sigma^2(x_{avg})}+\frac{\rho^2\xi^2}{\sigma^2(x_{avg})\bar{v}^3}[\frac{R^s_2}{12}-\frac{3(R^s_1)^2}{16\bar{v}}]+\frac{\xi^2R^s_3}{24\sigma^2(x_{avg})\bar{v}^3}\}.
\end{aligned}$$

CHAPTER 6

Smile and Skew behaviors for the CEV-Heston model

The aim of this Chapter is to study the impact of the CEV-Heston model parameters on the smile and the skew behaviors w.r.t. the maturity and the strike in order to illustrate numerically the discussion of the previous Chapter 5 Section 5.4. This is done for time-independent parameters using the third order implied volatility approximation formula at mid-point (see Chapter 5 Theorem 5.4.2.1 and Equation 5.72). We use our approximation formula for the sake of brevity instead of performing various time costing Monte Carlo simulations and we are confident with the fact that although not perfectly equal to the true implied volatilities, the estimations provided by our expansion formula are enough accurate to give a good overview of the influence of the model parameters.
We mainly focus on the volatility of volatility ξ, the skew parameter β and the correlation ρ. We allow these parameters to vary (independently or simultaneously) and we fix during all the tests the values of x_0, μ, v_0, θ (connected to the level of the long term variance at spot) and κ (which plays a similar role that the volatility of volatility but in the inverse way). We choose the following values:

$$x_0 = 0, \qquad \mu = 25\%, \qquad v_0 = 1, \qquad \theta = 1.2, \qquad \kappa = 3.$$

The maturity varies from $T = 3M$ to $T = 10Y$ and the log-moneyness to $m = -0.8$ to $m = 0.8$. We start from the symmetrical situation given with the choice of the values $\xi = 1.5$, $\beta = 1$ and $\rho = 0$ and then we study the impact of modifications of these values on the smile and the skew. The approximation of the implied volatility surface in this symmetrical situation is given in Figure 6.1 and we notice a marked smile symmetric w.r.t. the log-moneyness for short maturity (this confirms the property proved in [Renault 1996]). The smile flattens for long maturity and the long term volatility converges to the value $\mu\sqrt{\theta} \approx 27.40\%$ as it is proved in [Lewis 2000, Chapter 6].

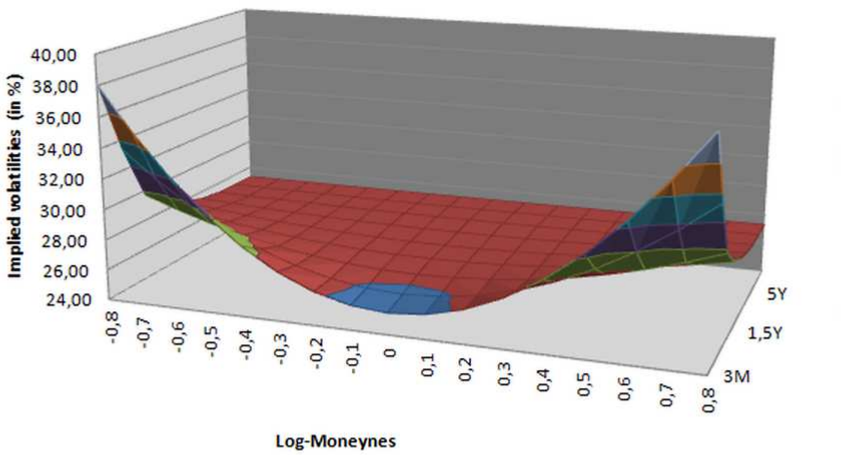

Figure 6.1: Implied Black-Scholes volatilities written as function of log-moneyness and maturities. Parameters: $\xi = 1.5, \beta = 1$ and $\rho = 0$.

▷**Impact of the skew parameter β.** We remark from the figure 6.2 that:

- When β decreases, the center of the short maturity smile is shifted to the right and a negative skew progressively appears.
- For long maturities and small values of β, we observe an important negative skew whereas the curvature decreases. That justifies the name of "skew parameter" because β induces a negative skew at both short and long maturity.

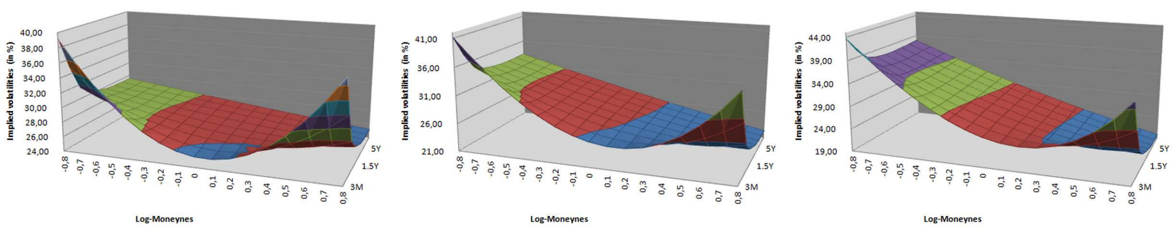

Figure 6.2: Implied Black-Scholes volatilities written as function of log-moneyness and maturities (on the left for $\beta = 0.8$, at the center for $\beta = 0.5$ and on the right for $\beta = 0.2$). Other parameters: $\xi = 1.5$ and $\rho = 0$.

▷**Impact of the correlation ρ.** We notice on figure 6.3 that:

- For negative values of ρ and short maturity, the impact of the correlation is close to the impact of the skew parameter β. We nevertheless mention that when ρ increases in absolute value, the curvature quickly decreases and for ρ close to -1, the curve seems to become concave.

- For positive values of ρ, the center of the short maturity smile is shifted to the left. More ρ is close to 1, more the smile shape changes from a symmetric smile to an inverted skew slightly concave.

- For all the correlations, we observe for long maturity an almost flat curve approximately equal to $\mu\sqrt{\theta} \approx 27.40\%$. We therefore interpret ρ as a short term skew parameter.

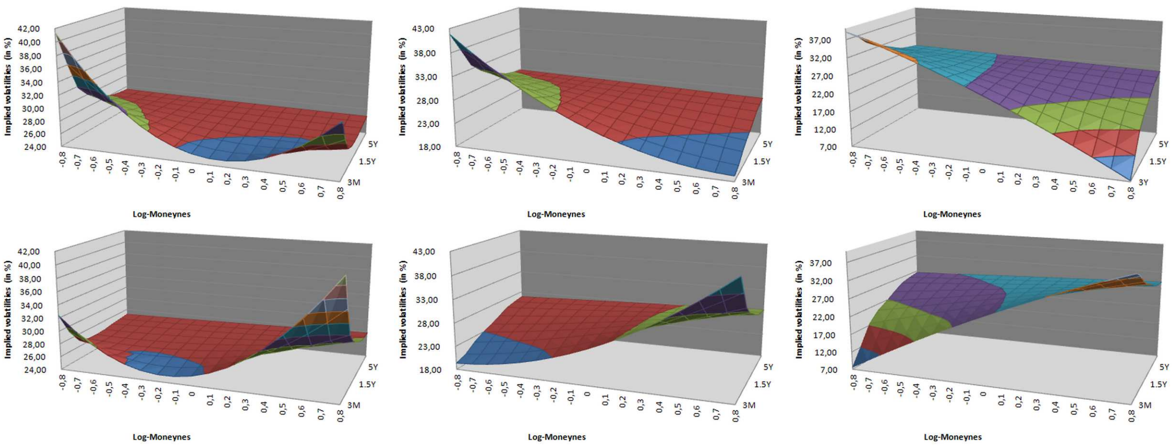

Figure 6.3: Implied Black-Scholes volatilities written as function of log-moneyness and maturities (from left to right and from top to bottom: $\rho = -20\%$, $\rho = -50\%$, $\rho = -70\%$, $\rho = 20\%$, $\rho = 50\%$, $\rho = 70\%$). Other parameters: $\xi = 1.5$ and $\beta = 1$.

▷**Impact of the volatility of volatility ξ.** We notice on figure 6.4 that:

- For small values of ξ, the short term implied volatility is not far from a flat surface. When ξ increases, the smile for short maturity progressively appears and we notice a U shape.

- For long maturities , the implied volatility flattens to the value $\mu\sqrt{\theta} \approx 27.40\%$ whatever is the volatility of volatility. This confirms that ξ is a short term smile or curvature parameter.

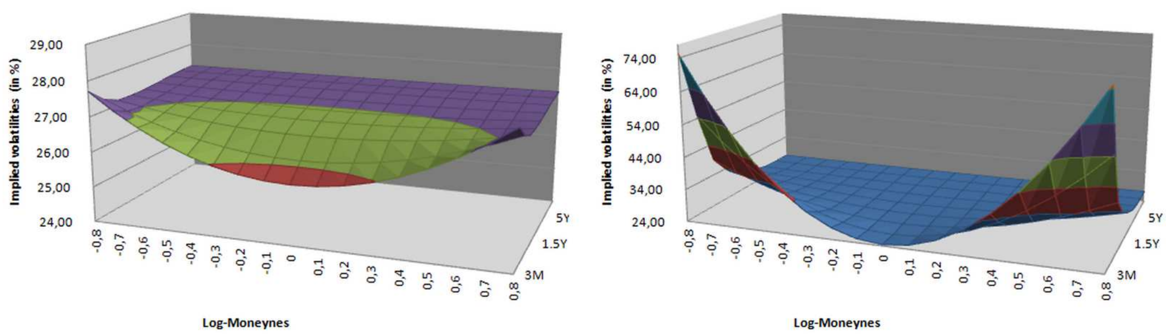

Figure 6.4: Implied Black-Scholes volatilities written as function of log-moneyness and maturities (on the left for $\xi = 0.6$ and on the right for $\xi = 3$). Other parameters: $\beta = 1$ and $\rho = 0$.

▷**Joint impact of ξ and β.** We notice on figure 6.5 that:

- For small values of ξ and β, we observe an important negative skew at both short and long maturity. The surface is close to an inclined plane and we retrieve the features of pure local volatiliy models.
- For large ξ and small β, we observe for short maturity and negative log-moneyness a negative skew (the curvature seems less emphasized in this left side) whereas the smile is more noticeable on the right side: the curve is pulled up with a strong convexity. For long maturity only the influence of β is visible and we observe a quite important negative skew.

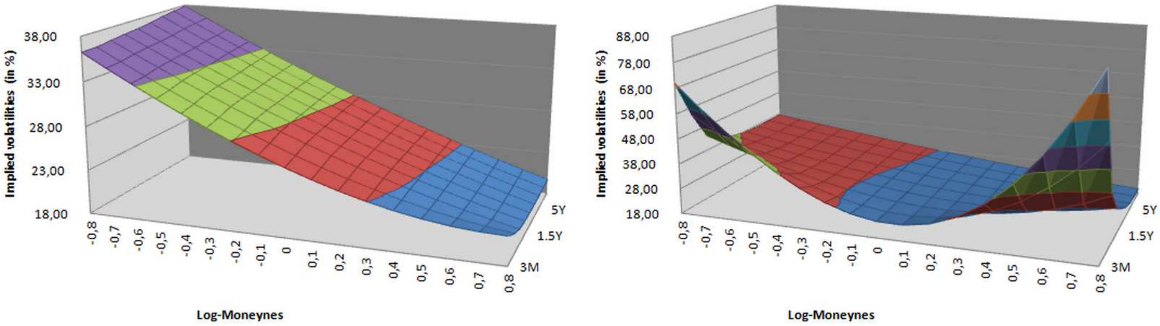

Figure 6.5: Implied Black-Scholes volatilities written as function of log-moneyness and maturities (on the left for $\xi = 0.6$ and on the right for $\xi = 3$). Other parameters: $\beta = 0.2$ and $\rho = 0$.

▷**Joint impact of ξ and ρ.** We notice on figure 6.6 that:

- For ξ close to 0 and ρ close to -1, there is an important negative skew for short maturity, skew which vanishes for long maturity.
- When ξ increases, the negative skew is more emphasized for short maturity what seems to indicate that, when ρ is large enough (in absolute value), its impact prevails over the influence of ξ and that increasing ξ emphasizes the skew. As expected the implied volatility curve flattens for long maturity.

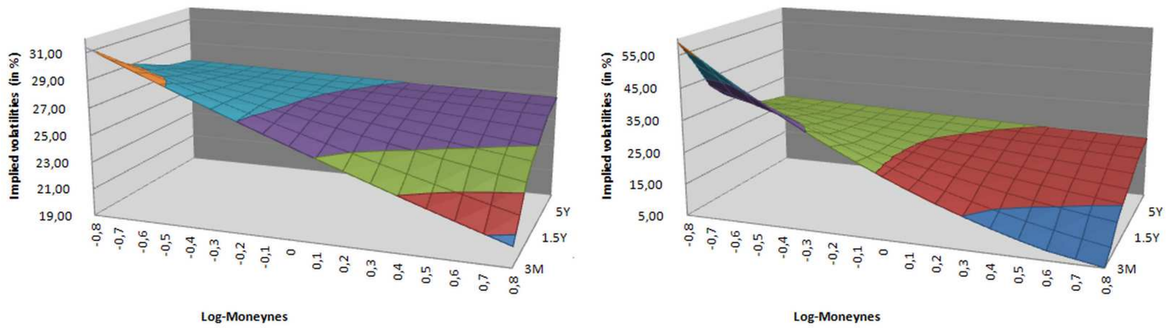

Figure 6.6: Implied Black-Scholes volatilities written as function of log-moneyness and maturities (on the left for $\xi = 0.6$ and on the right for $\xi = 3$). Other parameters: $\beta = 1$ and $\rho = -0.7$.

▷**Joint impact of ρ and β.** We notice on figure 6.7 that:

- For small β and ρ close to -1 we observe for short maturity a very important negative skew as if the influence of β and ρ were added. At long maturity, it remains only the influence of β and the skew is less emphasized.

- For ρ close to zero, the influences of β and ξ prevail. We are close to the behaviour observed on figure 6.5 right-side.

- For ρ close to 1, we notice for short maturity a positive skew du to the important correlation which the influence dominates the impact of β and at long maturity an important negative skew due to β. We thus observe a twisted implied volatility surface and we do not know if this behaviour may correspond to something observable in financial markets.

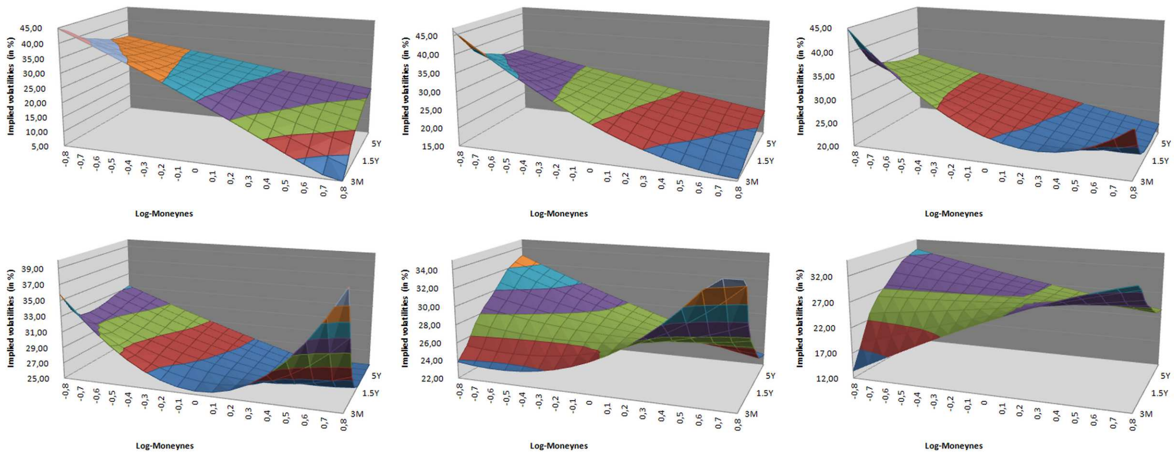

Figure 6.7: Implied Black-Scholes volatilities written as function of log-moneyness and maturities (from left to right and from top to bottom: $\rho = -70\%$, $\rho = -50\%$, $\rho = -20\%$, $\rho = 20\%$, $\rho = 50\%$, $\rho = 70\%$). Other parameters: $\xi = 1.5$ and $\beta = 0.5$.

▷**Joint impact of ξ, ρ and β.** We notice on figures 6.8 and 6.9 that:

- When ξ is close to 0, ρ close to 1 and $\beta = 0.5$, we observe an almost flat curve for short maturity as if the impact of ρ (reduced by the small value of ξ) was of the same magnitude of the impact of β but in the opposite way, so that the both influences are cancelled. For long maturities, a negative skew appears due to β.
- When we fix ρ to 0.7, the more ξ is large and the more β is close to 1, the more the positive skew is important for short maturity. For long maturity the negative skew seems to depend only on β.
- When we fix ξ to 3 and β to 0.5 and allow ρ to vary from -0.6 to -0.3 (10% per 10%), the short-term negative skew progressively becomes less emphasized and a smile with center slightly centred on the right progressively appears. It remains a negative skew for long maturities only due to β.

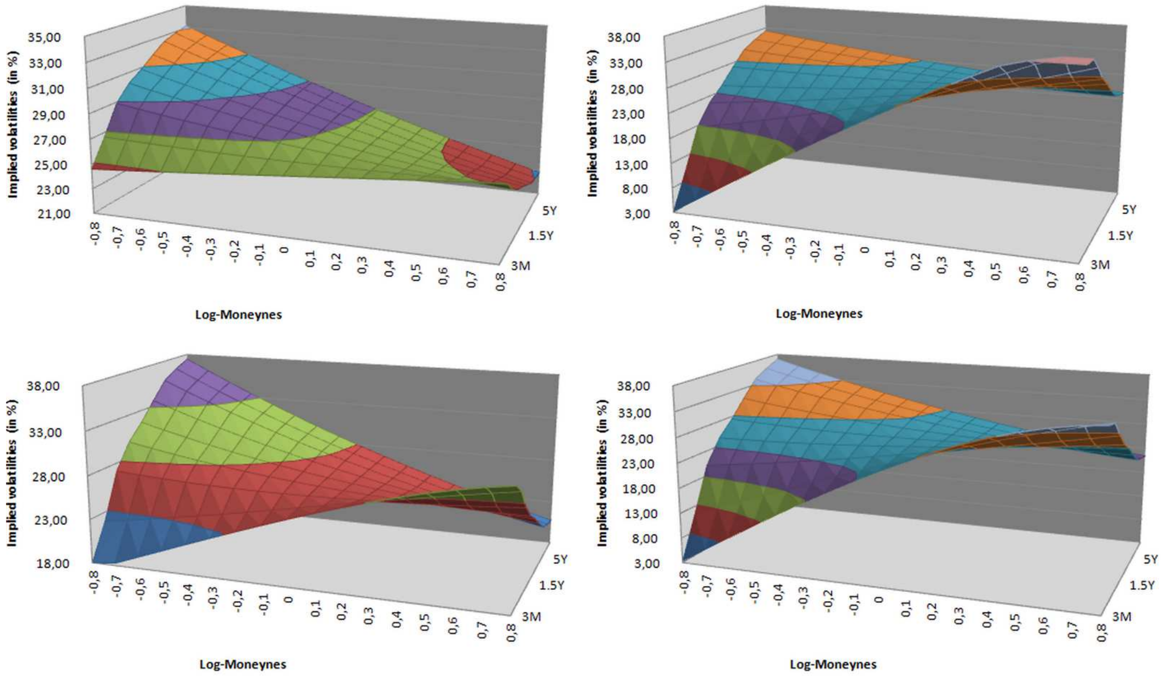

Figure 6.8: Implied Black-Scholes volatilities written as function of log-moneyness and maturities. From left to right and from top to bottom: ($\xi = 0.6, \rho = 0.7, \beta = 0.5$), ($\xi = 2.5, \rho = 0.7, \beta = 0.5$), ($\xi = 1.5$, $\rho = 0.7, \beta = 0.2$), ($\xi = 2.5, \rho = 0.7, \beta = 0.2$).

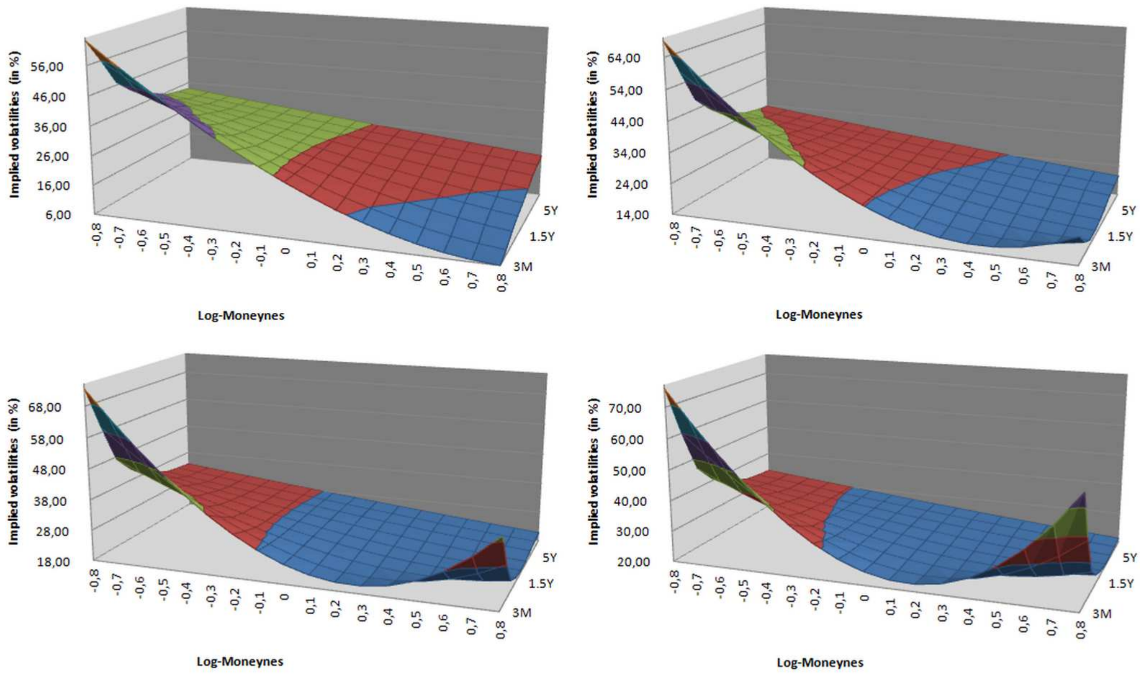

Figure 6.9: Implied Black-Scholes volatilities written as function of log-moneyness and maturities (from left to right and from top to bottom: $\rho = -60\%$, $\rho = -50\%$, $\rho = -40\%$, $\rho = -30\%$). Other parameters: $\xi = 3$ and $\beta = 0.5$.

Part III

Price approximation formulas for barrier options

Chapter 7

Price expansions for regular down barrier options

Contents

We introduce in this Chapter option price expansions for regular barrier options focusing on the down and out case and deducing from it the down and in case. In the framework of time-dependent local volatility models, we derive new formulas using a method mixing Itô calculus and PDE approach. We choose a Gaussian proxy model and express the difference with the local volatility model using the PDE associated to the Gaussian proxy process. Then we smartly combine expansion of the local volatility function, Itô calculus, key relations (involving martingales, convolution simplifications) and PDE arguments to obtain the approximation formulas with tight error estimates using the derivatives of the Gaussian proxy kernel. The presence of the hitting times complicates the analysis and this framework has not been very studied in the literature even with pure PDE point of view (because only the joint use of stochastic analysis and PDE arguments may presumably lead to the key reductions). In order to simplify this already quite delicate analysis, we assume that the risk-free rate and the dividend yield are equal to 0 and consider martingale assets. The excellent accuracy of our formulas is illustrated throughout various numerical experiments.

7.1 Introduction

Barrier options belong to the most popular path-dependent derivatives. For example a down and out option with level of barrier equal to B, written on the underlying S starting from an initial value greater than B and with terminal payoff h at the maturity T pays to its owner $h(S_T)$ at time T if the barrier has not be reached by the underlying by that time. If the asset price hits the barrier, the option expires worthless. On the contrary, a down and in option expires worthless unless the underlying reaches the barrier before expiry. Analogously we can define the upper barrier options. Note that in this work we only consider the case of regular barrier option, i.e. with payoff function equal to zero beyond the level of the barrier. This is for example the case of a down and out/in Call option with strike greater than the barrier level. These options present the particularity to have a bounded delta.

▷**Comparison with the literature.** The literature on pricing and hedging barrier options is very profuse and it is obvious that we will not pay tribute to it in just few references. Explicit formulas are available for the Black-Scholes or Bachelier models because the joint law of the asset price and its running minimum (or maximum) is known. We refer to Reiner et al. [Reiner 1991] for straightforward computations or to the works of Carr et al. [Carr 1998a] and El Karoui et al. [El Karoui 1999] where the authors use the reflection principle, which can be seen as a symmetry property of the Brownian motion, to deduce prices and hedges of barrier options.
In general no closed-form formulas are available for time-dependent local volatility models. If the joint law of the asset price and its running minimum is not known or if, in a dual point of view, one can not solve the associated PDE with boundary conditions, one needs to use numerical methods like finite difference approaches (see Boyle et al. [Boyle 1998]) or sophisticated Monte Carlo simulations (see Pham [Pham 2010] or Gobet [Gobet 2009] for general reviews). In addition to not being real-time methods, their application in the context of barrier option is quite delicate in comparison to the plain-vanilla framework. For example, it is well known that the usual Euler scheme in Monte Carlo simulations yields to approximations overestimating the exact value because there is no control of the diffusion path between two successive discrete dates. As a consequence, if we consider a n-time discretization with the step $\frac{T}{n}$, we obtain an error between the real barrier option price and the price obtained by the Euler scheme of order $\sqrt{\frac{T}{n}}$ instead of the classical order $\frac{T}{n}$ for standard European options (see [Gobet 2010]). This is the reason why we find in the literature advanced Monte Carlo schemes (as in the previous cited references) to improve the discretization order of convergence like Brownian bridge techniques to take into account eventual exits of the diffusion outside the domain between two consecutive discretization dates (see [Baldi 1995] or [Gobet 2001]) or methods of shifting the barrier inside the activation zone of the option to compensate the overestimation bias (see Costantini et al. [Costantini 2006], Gobet and Menozzi [Gobet 2010] or the works of Broadie et al. [Broadie 1997, Broadie 1999] where on the contrary the authors use a continuity correction to price discrete barrier options with continuous barrier formulas).
An other point of view consists of deriving explicit analytical approximations of continuous-time barriers and this is the purpose of the Chapter. We adapt the Proxy principle developed in [Benhamou 2009] and [Benhamou 2010a] (which consists roughly speaking in performing a non-asymptotic expansion of the quantity of interest around a Proxy model) in order to avoid the use of the stochastic analysis which seems inadequate to handle the hitting times. As a result we provide explicit and accurate ana-

lytical formulas with tight error estimates (written in terms of the magnitude of the model parameters and the maturity) for the pricing of regular down barrier options written on an asset following a time-inhomogeneous diffusion. We firstly concentrate on the down and out case and then easily deduce the down and in case. By symmetry, all the results can be transposed for regular up barrier options. To the best of our knowledge, no other analytical and tractable approximation of barrier options prices allowing to deal with general time-dependent local volatility models is available in the literature. We mention the very recent work of Kato et al. [Kato 2012] who perform an asymptotic expansion of the solution of the PDE with Cauchy-Dirichlet boundary condition associated to the expectation to approximate. But as mentioned before, results are available only for time-homogeneous diffusions and as discussed in Chapter 2 of the thesis, a PDE error analysis (which is far from straightforward) gives error estimates only in power of the perturbation parameter ϵ considering that the other model parameters (magnitude of the volatility, skew of the local volatility function…) and the maturity have no important asymptotic whereas the influence of the regularity of the payoff function is not taken into account.

An other approach, very developed in the last decade, is the dual problem of the research of an exact or at least approximate static hedge which consists to compute or approximate the price of a barrier option with an European-type contingent claim of the same maturity. See for instance Andersen et al. [Andersen 2002], Giese et al. [Giese 2007], Maruhn et al. [Maruhn 2009] or Ilhan et al. [Ilhan 2009]. However it has been shown in Bardos et al. [Bardos 2002] that exact static hedging strategies may not be available although the authors show the existence of an approximate static hedge for any smooth enough diffusion model, approximate static hedge which has to be computed numerically as in many quoted references. More recently, we cite Carr et al. [Carr 2011] in which the authors find for time-homogeneous diffusions an explicit analytical expression for the payoff function allowing the static hedge under some regularity assumptions. Some limitations of these methods are: 1)once the payoff function is found, it remains for pricing issues to effectively compute the price of the vanilla option, which is possible only if the law of the asset price is known and thus this generally leads to untractable formulas; 2)results are oftenly available only for time-independent coefficients due to the use of the Laplace transform and Sturm-Liouville equations (we refer to [Davydov 2001] or [Davydov 2003]).

▷**Formulation of the problem.** In this Chapter, we consider financial products written on a single asset which the log price at time t is denoted by X_t. We model the dynamic of X through a linear Brownian motion defined on a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{0\le t\le T},\mathbb{P})$, where $T>0$ is a fixed terminal time and $(\mathcal{F}_t)_{0\le t\le T}$ is the completion of the natural filtration of W. We consider that the risk-free rate and the dividend yield are equal to 0 [1] and we directly model the log asset by setting:

$$X_t = x_0 + \int_0^t \sigma(s,X_s)(\mathrm{d}W_s - \frac{\sigma(s,X_s)}{2}\mathrm{d}s). \tag{7.1}$$

Our main objective is to give an accurate analytic approximation of a down and out regular barrier option, written as:

$$\mathbb{E}[h(X_T)\mathbb{1}_{\inf\limits_{t\in[0,T]} X_t > b}] = \mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}], \tag{7.2}$$

where $\mathbb{E}$ stands for the standard expectation operator under the risk neutral probability, $b<x_0$ is the level of the barrier, h is a locally Lipschitz payoff function such that $h(x)=0$, $\forall x\le b$ and $\tau_b=\inf\{t>0:\ X_t=b\}$ is the first hitting time of the level b for the process X.

Following the proxy approach developed in [Benhamou 2009] [Benhamou 2010b], the idea is to use a Gaussian proxy process $(X_t^P)_{t\in[0,T]}$ obtained by freezing the space variable in the coefficient σ:

$$X_t^P = x_0 + \int_0^t \sigma(s,x_0)(\mathrm{d}W_s - \frac{\sigma(s,x_0)}{2}\mathrm{d}s). \tag{7.3}$$

[1] The case of non trivial cost of carry seems not to be a straightforward extension and this is left for further research.

Similarly, we introduce the first hitting time of the level b for the proxy process X^P: $\tau_b^P = \inf\{t > 0 : X_t^P = b\}$. The practical interest of this proxy is that the joint law of $(X_T^P, \inf_{t\in[0,T]} X_t^P)$ is known. Heuristically, this law is expected to be close to the law of $(X_T, \inf_{t\in[0,T]} X_t)$ if the spatial derivatives of the local volatility function σ are small, or if $|\sigma|_\infty$ is small, or if the maturity T is short. But we do not only replace $(X_T, \inf_{t\in[0,T]} X_t)$ by $(X_T^P, \inf_{t\in[0,T]} X_t^P)$, we also provide correction terms in order to achieve a higher accuracy (see Theorems 7.2.2.1 and 7.2.3.1).

▷**Comparison with previous works and contribution of the Chapter.** As a difference with previous works using the proxy approach [Benhamou 2009], [Benhamou 2010b], [Benhamou 2010a], [Benhamou 2012] or Parts I and II of the thesis, we do not use a parameterization to link the initial model and the proxy model and we do not employ the stochastic analysis to justify our expansions. There are two difficulties: firstly the hitting times are not smooth w.r.t. regular perturbations and secondly one can not directly apply the Malliavin calculus on them because of the indicator function on the minimum (see [Nualart 2006]) whereas even the minimum of a Brownian motion is only once Malliavin differentiable (see [Nualart 2006]).

To overcome these difficulties, we follow an approach presented in Chapter 2 Subsection 2.2.2 mixing Itô calculus and PDE. The idea is to represent the error $\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}] - \mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P>T}]$ using the PDE associated to the proxy:

$$v_{\underline{o},T}^{P,h}(t,x) = \mathbb{E}[h(X_T^P)\mathbb{1}_{\inf_{s\in[t,T]} X_s^P > b}|X_t^P = x].$$

Then we smartly combine expansion of the local volatility function, Itô calculus, martingale properties and PDEs to obtain approximation formulas. The calculus of the corrective terms is not anymore performed using the Malliavin integration by parts formula, which can be view as a static operation focusing only on the law of a random variables (typically X_T^P), but we follow an approach involving all the dynamic structure of the processes $(X_t)_{t\in[0,T]}$ and $(X_t^P)_{t\in[0,T]}$. If the tools remain more standard, the explicit derivation of the expansion coefficients is a rather tricky and one has to carefully combine martingale properties and Itô calculus (this is somewhat related to Bismut-Elworthy-Li formula like in [Thalmaier 1998] and [Delarue 2003]).

We provide explicit second and third order formulas which coincide with [Benhamou 2010a, Theorems 2.1 and 2.2] if the level of the barrier b tends to $-\infty$. In case of non trivial barrier level, the corrective terms are composed of new terms of the form $\mathbb{E}[\mathbb{1}_{\tau_b^P\le T}\vec{\omega}(m_1,..,m_n)_{\tau_b^P}^T\partial_{x^i}^i v_{\underline{o},T}^{P,h}(\tau_b^P,b)]$, where $\vec{\omega}$ is an integral operator defined in the following Subsection and $(m_1(t),..,m_n(t))_{t\in[0,T]}$ are functions depending on σ^2 and its spatial derivatives computed at $x = x_0$.

▷**Outline of the Chapter.** The Chapter is organised as follows. In Section 7.2, we state our main approximation results: in the general time-dependent local volatility framework, we provide order 2 and order 3 approximations of regular down and out barrier options prices (see Theorems 7.2.2.1-7.2.3.1) with an estimation of the error justifying the order. Section 7.3 is devoted to the derivation of the expansion and the justification of the error magnitude. We give in Section 7.4 Corollaries on the pricing of regular down and in barrier options prices (see Theorems 7.4.0.1-7.4.0.2). Then in Section 7.5 we apply our approximation formulas to the particular case of regular down barrier Call options and derive new expansions with the local volatility frozen at mid-point between the log-spot and the log-strike. In addition, we show that in the time-homogeneous framework, our expansion formulas reduced to totally explicit and very simple expressions and we finally gather numerical experiments illustrating the high-performance of our approximation formulas. In Appendix 7.6.1 we give some additional results concerning the Gaussian density and the Gaussian hitting times density (relations between their partial derivatives and convolution properties). Appendices 7.6.2 and 7.6.3 are devoted to the proof of technical results.

7.2 Derivation of the expansion

7.2.1 Notations and definitions

The following notations and definitions are repeatedly used in this Chapter.

▹**Differentiation.** For any measurable function f of $(t,x) \in [0,T] \times \mathbb{R}$, we write if these derivatives have a meaning: $f_t(x) = f(t,x)$ and $f_t^{(i)}(x) = \partial^i_{x^i} f(t,x)$. When considering the spatial point x_0, we omit if unambiguous the dependence w.r.t. the spatial component and write $f_t = f(t,x_0)$ and $f_t^{(i)} = \partial^i_{x^i} f(t,x_0)$. For instance we have $\sigma_t := \sigma(t,x_0)$ and $\sigma_t^{(i)} := \partial^i_{x^i}\sigma(t,x_0)$.

▹**Total variance.** We define the local variance function by $\Sigma_t(x) = \sigma_t^2(x)$. Then we define the total variance for the process X^P on the period $[t,T]$ for any $t \in [0,T]$ by $\mathcal{V}_t^T = \int_t^T \Sigma_s \mathrm{d}s$.

▹ **Integral Operator.** We define the integral operator $\vec{\omega}$ as follows: for any measurable and bounded function l of $t \in [0,T]$, for any $s \leq t \leq T$, we set:

$$\vec{\omega}(l)_s^t = \int_s^t l_u \mathrm{d}u$$

For any measurable and bounded functions $(l_1, \cdots, l_n)$, we define its n-times iteration by:

$$\vec{\omega}(l_1, \cdots, l_n)_s^t = \vec{\omega}(l_1 \vec{\omega}(l_2, \cdots, l_n)_s^{\cdot})_s^t, \ \forall s \leq t \leq T.$$

The reader should pay attention to the fact that the Definition of the operator $\vec{\omega}$ is different from the previous Chapters. We adopt here a forward convention instead of a backward convention due to the new methodology employed in the Chapter. For instance, $\vec{\omega}(l_1,l_2)_0^T = \int_0^T l_1(t)(\int_0^t l_2(s)\mathrm{d}s)\mathrm{d}t$ instead of the old Definition $\omega(l_1,l_2)_0^T = \int_0^T l_1(t)(\int_t^T l_2(s)\mathrm{d}s)\mathrm{d}t$.

▹ **Assumption ($\widetilde{\mathcal{H}}^{\sigma}_{x_0}$) on σ.** σ is a bounded measurable function of $(t,x) \in [0,T] \times \mathbb{R}$, and three times continuously differentiable in x with bounded derivatives. Using the notation $|\sigma|_\infty = \sup_{(t,x)\in[0,T]\times\mathbb{R}} |\sigma(t,x)|$, we set:

$$\mathcal{M}_1(\sigma) = \max_{1\leq i\leq 3} \sup_{(t,x)\in[0,T]\times\mathbb{R}} |\partial^i_{x^i}\sigma(t,x)| \text{ and } \mathcal{M}_0(\sigma) = \max(|\sigma|_\infty, \mathcal{M}_1(\sigma)).$$

In addition, there exists a constant $C_\sigma > 0$ such that $\sigma_{\inf}(x_0) = \inf_{t\in[0,T]} \sigma_t(x_0) \geq C_\sigma$.
The reader can notice the stronger ellipticity assumption assumed for the proxy process in comparison to the previous Chapters where we supposed ($\mathcal{H}^{\sigma}_{x_0}$): $\int_0^T \sigma^2(x_0)\mathrm{d}t > 0$ (which implies that X_T^P is a non-degenerate Gaussian variable). Our dynamic approach leverages all the trajectory of the proxy process and our assumption ensures that $\forall t \in [0,T]$, X_t^P defined in (7.3) is a non-degenerate Gaussian variable.

▹ **Assumptions on the payoff function h.** h belongs to $\underline{\mathrm{Lip}}(\mathbb{R},b)$ the space of real valued locally Lipschitz functions such that $h(x) = 0$ for any $x \leq b$. There exists a constant $C_h \geq 0$ such that:

$$|h(x)| \leq C_h e^{C_h|x|}, \qquad |\frac{h(y)-h(x)}{y-x}| \leq C_h e^{\frac{C_h}{2}(|x|+|y|)} \quad (\forall y \neq x).$$

In particular, h is a.e. differentiable with $|h^{(1)}(x)| \leq C_h e^{C_h|x|}$. This space includes the Call payoff function $h(x) = (e^x - e^k)_+$ with log-strike $k \geq b$.

Remark 7.2.1.1. *By symmetry, all the results of the Chapter are naturally extended to regular up barrier options, i.e. for any payoff function $h \in \overline{\mathrm{Lip}}(\mathbb{R},b)$ the space of real valued functions locally Lipschitz null on $[b,+\infty]$. This space notably includes the Put payoff function $h(x) = (e^k - e^x)_+$ with log-strike $k \leq b$.*

▷ **Constants and error estimates.** We use the following notations to state our error estimates throughout the Chapter:

- "$A = O(B)$" means that $|A| \leq CB$ where C stands for a generic constant that is a non-negative increasing function of T, $\mathcal{M}_0(\sigma)$, $\mathcal{M}_1(\sigma)$ and of the oscillation ratio $\frac{|\sigma|_\infty}{C_\sigma}$.
- Similarly, if A is positive, $A \leq_c B$ means that $A \leq CB$ for a generic constant C.
- The L^p-norm of a random variable Z is denoted as usual by $\|Z\|_p = \mathbb{E}[|Z|^p]^{\frac{1}{p}}$.

▷**Differential operators.** Some specific differential operators are frequently utilized in this Chapter. For convenience we introduce notations to denote them:

Definition 7.2.1.1.

$$\mathcal{L}_1^x = \partial_x - \frac{1}{2}\mathcal{I}, \qquad \mathcal{L}_2^x = \partial_{x^2}^2 - \partial_x,$$

$$\mathcal{L}_3^x = \mathcal{L}_2^x \circ \mathcal{L}_1^x = \partial_{x^3}^3 - \frac{3}{2}\partial_{x^2}^2 + \frac{1}{2}x, \qquad \mathcal{L}_4^x = \mathcal{L}_2^x \circ \mathcal{L}_2^x = \partial_{x^4}^4 - 2\partial_{x^3}^3 + \partial_{x^2}^2,$$

$$\mathcal{L}_5^x = \mathcal{L}_2^x \circ \mathcal{L}_3^x = \partial_{x^5}^5 - \frac{5}{2}\partial_{x^4}^4 + 2\partial_{x^3}^3 - \frac{1}{2}\partial_{x^2}^2, \qquad \mathcal{L}_6^x = \mathcal{L}_2^x \circ \mathcal{L}_4^x = \partial_{x^6}^6 - 3\partial_{x^5}^5 + 3\partial_{x^4}^4 - \partial_{x^3}^3,$$

where $\mathcal{I}$ denotes the identity operator. We have in addition the following relations easy to obtain:

$$\mathcal{L}_1^x \circ \mathcal{L}_1^x = \mathcal{L}_2^x + \frac{1}{4}\mathcal{I}, \qquad \mathcal{L}_1^x \circ \mathcal{L}_3^x = \mathcal{L}_4^x + \frac{1}{4}\mathcal{L}_2^x,$$

$$\mathcal{L}_1^x \circ \mathcal{L}_4^x = \mathcal{L}_3^x \circ \mathcal{L}_2^x = \mathcal{L}_5^x, \qquad \mathcal{L}_1^x \circ \mathcal{L}_5^x = \mathcal{L}_6^x + \frac{1}{4}\mathcal{L}_4^x.$$

▷ **Density of the Proxy and its hitting times.** The Gaussian Proxy process at time $t \in [0, T]$ $X_t^P = x_0 - \frac{1}{2}\mathcal{V}_0^t + \int_0^t \sigma_s \mathrm{d}W_s$ has the explicit density:

$$\mathcal{D}^P(0, t, x_0, y) := \mathcal{D}^P(0, t, y - x_0) = \frac{e^{-\frac{(y - x_0 + \frac{1}{2}\mathcal{V}_0^t)^2}{2\mathcal{V}_0^t}}}{\sqrt{2\pi\mathcal{V}_0^t}}.$$

With the conditional information that $X_s^P = x$ for $0 \leq s < t \leq T$, the conditional law of X_t^P is given by the density:

$$\mathcal{D}^P(s, t, x, y) := \mathcal{D}^P(s, t, y - x) = \frac{e^{-\frac{(y - x + \frac{1}{2}\mathcal{V}_s^t)^2}{2\mathcal{V}_s^t}}}{\sqrt{2\pi\mathcal{V}_s^t}}. \tag{7.4}$$

We use the notation $\mathcal{N}(x) = \int_{-\infty}^x \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\mathrm{d}y$ for the cumulative Gaussian function at point x and we denote by $\overline{\mathcal{N}}(x) = 1 - \mathcal{N}(x) = \int_x^\infty \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\mathrm{d}y$ its associated complementary function. The derivative of $\mathcal{N}$ which is the standard Gaussian density is naturally denoted by $\mathcal{N}'(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$.
We define the first hitting time of b for X^P after time $s \in [0, T[$ by:

$$\tau_{s,b}^P = \inf\{t > s : \ X_t^P = b\}, \tag{7.5}$$

with the convention $\tau^P_{0,b} = \tau^P_b$. The density of the hitting times for a drifted Brownian with constant volatility is well known (see [Karatzas 1991]). Using a regular time-change thanks to $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$ (see [Revuz 1999]) on can deduce the time-inhomogeneous case. The first hitting time τ^P_b for the level $b < x_0$ for the Gaussian proxy process X^P has the density $\mathcal{J}^P$:

$$\mathcal{J}^P(0,t,x_0,b) := \mathcal{J}^P(0,t,b-x_0) = \Sigma_t \frac{|b-x_0|}{\mathcal{V}^t_0} \frac{e^{-\frac{(b-x_0+\frac{1}{2}\mathcal{V}^t_0)^2}{2\mathcal{V}^t_0}}}{\sqrt{2\pi\mathcal{V}^t_0}} = \Sigma_t \frac{(x_0-b)}{\mathcal{V}^t_0}\mathcal{D}^P(0,t,b-x_0),$$

for any $t \geq 0$. Consequently with the conditional information that $X^P_s = x > b$ for $s < T$, the density of the first hitting time $\tau^P_{s,b}$ of the level b after the time s is given by:

$$\mathcal{J}^P(s,t,x,b) := \mathcal{J}^P(s,t,b-x) = \Sigma_t \frac{(x-b)}{\mathcal{V}^t_s}\mathcal{D}^P(s,t,b-x)\mathbb{1}_{t\geq s}, \tag{7.6}$$

for any $t \in [s,T]$. We summarize in Appendix 7.6.1 some useful properties of $\mathcal{D}^P$, $\mathcal{J}^P$ and $\mathcal{N}$.

▷**Proxy pricing kernel.** We introduce the proxy pricing kernel for the down and out option defined for any $x \geq b$ and any $t \leq T$ by:

$$v^{P,h}_{\underline{o},T}(t,x) = \mathbb{E}[h(X^P_T)\mathbb{1}_{\tau^P_{t,b}>T}|X^P_t = x]. \tag{7.7}$$

We denote similarly by $v^{P,h}_T(t,x) = \mathbb{E}[h(X^P_T)|X^P_t = x]$ the proxy pricing kernel for the plain vanilla option. We prefer work firstly with the down and out case because $v^{P,h}_{\underline{o},T}(t,x)$ is naturally connected to a Cauchy-Dirichlet problem. Under the ellipticity assumption $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$,whatever is the regularity of the payoff function h, the proxy pricing kernel is $C^\infty([0,T[\times[b,+\infty[)$ and satisfies the following parabolic PDE with Dirichlet boundary condition and with Cauchy terminal condition (see [Cattiaux 1986]-[Cattiaux 1991]):

$$\begin{cases} \partial_t v^{P,h}_{\underline{o},T}(t,x) + \frac{1}{2}\Sigma_t \mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t,x) = 0, \quad (t,x) \in [0,T[\times[b,+\infty[, \\ v^{P,h}_{\underline{o},T}(t,b) = 0, \quad t \in [0,T], \\ v^{P,h}_{\underline{o},T}(T,x) = h(x), \quad x \in]b,+\infty[. \end{cases} \tag{7.8}$$

The solution in closed-form is given $\forall(t,x) \in [0,T]\times[b,+\infty[$ by:

$$\begin{aligned} v^{P,h}_{\underline{o},T}(t,x) &= \int_b^\infty h(y)(1-e^{-\frac{2(x-b)(y-b)}{\mathcal{V}^T_t}})\mathcal{D}^P(t,T,y-x)\mathrm{d}y \\ &= \int_b^\infty h(y)\mathcal{D}^P(t,T,y-x)\mathrm{d}y - e^{(x-b)}\int_b^\infty h(y)\mathcal{D}^P(t,T,y+x-2b)\mathrm{d}y, \end{aligned} \tag{7.9}$$

where $\mathcal{D}^P$ is the density of the Proxy defined in (7.4). If $h \in \underline{\mathrm{Lip}}(\mathbb{R},b)$, we can consider the integration on the whole real axis (because $h(x) = 0$, $\forall x \leq b$) and the first term in the r.h.s. of (7.9) is exactly equal to $v^{P,h}_T(t,x)$ classical solution of the system (7.8) without the boundary condition.

Our further calculations are based on two key Lemmas of these kernels of pricing. The first Lemma gives relations for the sensitivities w.r.t. x computed at the barrier:

Lemma 7.2.1.1. *Assume $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$. For any $t \in [0,T[$, we have:*

$$\mathcal{L}^x_n v^{P,h}_{\underline{o},T}(t,b) = 0, \quad \forall n \in \{2,4,6\} \tag{7.10}$$

$$\mathcal{L}^x_n v^{P,h}_{\underline{o},T}(t,b) = 2\mathcal{L}^x_n v^{P,h}_T(t,b), \quad \forall n \in \{1,3,5\}, \tag{7.11}$$

where we recall that the differential operators $\mathcal{L}^x$ are defined in Definition 7.2.1.1.

Proof. The PDE solved by $v^{P,h}_{\underline{o},T}(t,x)$ (see (7.8)) and the Definition 7.2.1.1 allow to interpret $\mathcal{L}^x_n v^{P,h}_{\underline{o},T}(t,b)$, $\forall n\in\{2,4,6\}$, as iterated derivatives of $v^{P,h}_{\underline{o},T}(t,x)$ w.r.t. t at $x=b$. Besides we have $v^{P,h}_{\underline{o},T}(t,b)=0$, $\forall t\in[0,T[$ (see the boundary condition in (7.8)). For (7.11), it is sufficient in view of the Definition 7.2.1.1 to show that $\mathcal{L}^x_1 v^{P,h}_{\underline{o},T}(t,b)=\partial_x v^{P,h}_{\underline{o},T}(t,b)=2\mathcal{L}^x_1 v^{P,h}_{T}(t,b)$ which can be proven by a straightforward calculus. □

The second Lemma is a Martingale property which can be proven using standard arguments involving Itô's formula and PDE simplifications (see [Thalmaier 1998] and [Delarue 2003]):

Lemma 7.2.1.2. *Assume* $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$. $\forall 0\le s\le t<T$, $\forall x\ge b$ *and for any* $n\in\mathbb{N}$, *we have:*

$$\mathbb{E}[\partial^n_{x^n} v^{P,h}_{\underline{o},T}(t\wedge\tau^P_{s,b}, X^P_{t\wedge\tau^P_{s,b}})|X^P_s=x]=\partial^n_{x^n} v^{P,h}_{\underline{o},T}(s,x).$$

The above Lemmas are oftenly used to decompose sensitivities computed before τ^P_b in a martingale part and a part (eventually null) with sensitivities computed at (τ^P_b,b). For example we have, $\forall t\in[0,T[$, $\partial_x v^{P,h}_{\underline{o},T}(t,X^P_t)\mathbb{1}_{\tau^P_b>t}=\partial_x v^{P,h}_{\underline{o},T}(t\wedge\tau^P_b, X^P_{t\wedge\tau^P_b})-\mathbb{1}_{\tau^P_b\le t}\partial_x v^{P,h}_{T}(\tau^P_b,b)$ and $\mathbb{E}[\partial_x v^{P,h}_{\underline{o},T}(t\wedge\tau^P_b, X^P_{t\wedge\tau^P_b})]=\partial_x v^{P,h}_{\underline{o},T}(0,x_0)$.

7.2.2 Second order expansion

Applying the Itô's formula to $v^{P,h}_{\underline{o},T}(t,X_t)$ between 0 and $T\wedge\tau_b$ where $v^{P,h}_{\underline{o},T}$ is the proxy pricing kernel defined in (7.7) and $(X_t)_{t\in[0,T]}$ defined in (7.1) we obtain:

$$\begin{aligned}
&v^{P,h}_{\underline{o},T}(T\wedge\tau_b, X_{T\wedge\tau_b}) \qquad (7.12)\\
=&v^{P,h}_{\underline{o},T}(0,x_0)+\int_0^{T\wedge\tau_b}\sigma_t(X_t)\partial_x v^{P,h}_{\underline{o},T}(t,X_t)\mathrm{d}W_t+\int_0^{T\wedge\tau_b}(\partial_t+\frac{1}{2}\Sigma_t(X_t)\mathcal{L}^x_2)v^{P,h}_{\underline{o},T}(t,X_t)\mathrm{d}t\\
=&\mathbb{E}[h(X^P_T)\mathbb{1}_{\tau^P_b>T}]+\int_0^{T\wedge\tau_b}\sigma_t(X_t)\partial_x v^{P,h}_{\underline{o},T}(t,X_t)\mathrm{d}W_t+\frac{1}{2}\int_0^{T\wedge\tau_b}(\Sigma_t(X_t)-\Sigma_t)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t,X_t)\mathrm{d}t,
\end{aligned}$$

where we used the fact at the second equality that $v^{P,h}_{\underline{o},T}$ follows the PDE (7.8) on $]0,T[\times]b,+\infty[$. Then notice that:

$$v^{P,h}_{\underline{o},T}(T\wedge\tau_b, X_{T\wedge\tau_b})=v^{P,h}_{\underline{o},T}(T,X_T)\mathbb{1}_{\tau_b>T}+\underbrace{v^{P,h}_{\underline{o},T}(\tau_b,b)\mathbb{1}_{\tau_b\le T}}_{=0}=v^{P,h}_{\underline{o},T}(T,X_T)\mathbb{1}_{\tau_b>T}=h(X_T)\mathbb{1}_{\tau_b>T},$$

and that $(\int_0^{t\wedge\tau_b}\sigma_s(X_s)\partial_x v^{P,h}_{\underline{o},T}(s,X_s)\mathrm{d}W_s)_{t\in[0,T]}$ is a true $(\mathcal{F}_t)_{t\in[0,T]}$-martingale owing to $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$ and to Lemma 7.3.1.3 postponed to Section 7.3. Thus taking the expectation in (7.12), we finally get:

Proposition 7.2.2.1. ***(Robustness-type formulation).*** *Assume* $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$. *For any* $h\in\underline{\mathrm{Lip}}(\mathbb{R},b)$, *we have:*

$$\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}]=\mathbb{E}[h(X^P_T)\mathbb{1}_{\tau^P_b>T}]+\frac{1}{2}\mathbb{E}[\int_0^{T\wedge\tau_b}(\Sigma_t(X_t)-\Sigma_t)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t,X_t)\mathrm{d}t]. \tag{7.13}$$

Remark 7.2.2.1. *We have obtained some kind of Black-Scholes robustness formula w.r.t. the log-asset like in [El Karoui 1998]. This expresses the difference between the barrier prices in the local volatility model and in the proxy model, or equivalently the expectation of the tracking error when we try to hedge with the proxy model a barrier option with real underlying following the SDE (7.1). The result indicates that the error is the expectation of a temporal integral of sensitivities (involving the operator* $\mathcal{L}^x_2$*) of the proxy pricing kernel computed along the path of the initial process X, sensitivities which are weighted with the difference of the instantaneous variances in the two models.*

Employing a first order expansion for the function $\Sigma_t(.)$ at $x = X_t$ around $x = x_0$ gives:

$$\begin{aligned}\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}] =&\mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P>T}] + \frac{1}{2}\mathbb{E}[\int_0^T (X_t - x_0)\Sigma_t^{(1)}\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t,X_t)\mathbb{1}_{\tau_b>t}\mathrm{d}t]\\ &+\frac{1}{2}\mathbb{E}[\int_0^{T\wedge\tau_b}(X_t-x_0)^2\{\int_0^1(1-\alpha)\Sigma_t^{(2)}(x_0+\alpha(X_t-x_0))\mathrm{d}\alpha\}\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t,X_t)\mathrm{d}t].\end{aligned} \tag{7.14}$$

We consider the last term as an error which will be analysed later. The second term can not be computed in closed-form because of the presence of (X_t,τ_b) which law is not known. Observe that we can switch the expectation and the temporal integral owing to the Lipschitz regularity of h which gives rise to singular terms of the form $(T-t)^{-\frac{1}{2}}$ but integrable at T (see Lemma 7.3.1.3) and to the integrability of X thanks to $(\widetilde{\mathcal{H}}_{x_0}^{\sigma})$ (see Lemma 7.3.1.1).
To approximate $\mathbb{E}\Big[(X_t-x_0)\Sigma_t^{(1)}\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t,X_t)\mathbb{1}_{\tau_b>t}\Big]$ for any $t\in[0,T[$, we use similar arguments and Proposition 7.2.2.1 by replacing T by t and $h(x)$ by $\phi_t(x) = (x-x_0)\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t,x)$. We introduce:

$$v_{\underline{o},t}^{P,\phi_t}(s,x) = \mathbb{E}[\phi_t(X_t^P)\mathbb{1}_{\tau_{s,b}^P>t}|X_s^P = x], \tag{7.15}$$

for any $s\in[0,t]$ and any $x\geq b$ the solution of the system (7.8) on $]0,t[\times]b,\infty[$ but with terminal condition ϕ_t. We interpret $v_{\underline{o},t}^{P,\phi_t}$ as the price function of a new down and out barrier option with maturity t and the regular payoff ϕ_t (thanks to Lemma 7.2.1.1 equation (7.10), $\phi_t(b) = 0$).
The methodology previously employed leads to:

$$\begin{aligned}&\frac{1}{2}\mathbb{E}[\int_0^T(X_t-x_0)\Sigma_t^{(1)}\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t,X_t)\mathbb{1}_{\tau_b>t}\mathrm{d}t]\\ =&\frac{1}{2}\int_0^T\Sigma_t^{(1)}\mathbb{E}[\phi_t(X_t)\mathbb{1}_{\tau_b>t}]\mathrm{d}t = \frac{1}{2}\int_0^T\Sigma_t^{(1)}\mathbb{E}[v_{\underline{o},t}^{P,\phi_t}(t,X_t)\mathbb{1}_{\tau_b>t}]\mathrm{d}t = \frac{1}{2}\int_0^T\Sigma_t^{(1)}\mathbb{E}[v_{\underline{o},t}^{P,\phi_t}(t\wedge\tau_b,X_{t\wedge\tau_b})]\mathrm{d}t\\ =&\frac{1}{2}\int_0^T\Sigma_t^{(1)}v_{\underline{o},t}^{P,\phi_t}(0,x_0)\mathrm{d}t+\frac{1}{4}\int_0^T\Sigma_t^{(1)}\mathbb{E}[\int_0^{t\wedge\tau_b}(\Sigma_s(X_s)-\Sigma_s)\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,X_s)\mathrm{d}s]\mathrm{d}t.\end{aligned} \tag{7.16}$$

The last term is again neglected and the explicit calculus of the first term is given in the Lemma 7.3.2.1. We have paved the way to the next Theorem which proof of the error estimate is postponed to Section 7.3.3:

Theorem 7.2.2.1. ***(2nd order approximation price formula for down and out regular barrier options)***. *Assuming $(\widetilde{\mathcal{H}}_{x_0}^{\sigma})$, we have for any $h\in\underline{\mathrm{Lip}}(\mathbb{R},b)$ and for any $x_0\geq b$:*

$$\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}] = \mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P>T}]+Cor_{1,\underline{o}}+Cor_{2,\underline{o}}+O(|\sigma|_\infty\mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)T^{\frac{3}{2}}), \tag{7.17}$$

where the differential operator $\mathcal{L}_3^x$ is defined in Definition 7.2.1.1 and where:

$$Cor_{1,\underline{o}} = \frac{1}{2}\vec{\omega}(\Sigma^{(1)},\Sigma)_0^T\mathcal{L}_3^x v_{\underline{o},T}^{P,h}(0,x_0), \qquad Cor_{2,\underline{o}} = -\frac{1}{2}\mathbb{E}[\mathbb{1}_{\tau_b^P\leq T}\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^P}^T\mathcal{L}_3^x v_{\underline{o},T}^{P,h}(\tau_b^P,b)].$$

Remark 7.2.2.2. *The above price approximation is the sum of three terms:*

1. $\mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P>T}]$*: The leading order which is the price of a regular down and out barrier option in the Black-Scholes model which corresponds to the case of space-independent local volatility function σ. There is a closed-form formula for Call options and one can apply a numerical integration for general payoff functions.*

2. $Cor_{1,\underline{o}}$*: This term is a weighted sum of sensitivities w.r.t. the log-spot x_0 at the initial date. We interpret it as a correction due to the spatial dependence of the local volatility σ. This corrective term is similar to the volatility correction terms of [Benhamou 2010a, Theorem 2.1].*

3. *$Cor_{2,\underline{\varrho}}$: This correction term reads as a barrier correction and we can interpret it as a sensitivity w.r.t. the barrier level. For the Call payoff, we compute it easily with a numerical integration on the segment $[0,T]$, the density of the hitting times being known (see (7.6)) and the sensitivity $\mathcal{L}_3^x v_{\underline{\varrho},T}^{P,h}(t,x))|_{x=b}$ being available in closed-form for any $t\in[0,T[$. Owing to $(\widetilde{\mathcal{H}}_{x_0}^{\sigma})$, the weighted sensitivities $\mathbb{1}_{\tau_b^P\le T}\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^P}^T\mathcal{L}_3^x v_{\underline{\varrho},T}^{P,h}(\tau_b^P,b)$ are well defined and this term does not explode when $\tau_b^P\to T$ (see below).*

Remark 7.2.2.3. *The reader can notice that if b tends to $-\infty$, the correction term $Cor_{2,\underline{\varrho}}$ vanishes. We have indeed applying the Lemma 7.3.1.3 postponed to Section 7.3 owing to $(\widetilde{\mathcal{H}}_{x_0}^{\sigma})$:*

$$|Cor_{2,\underline{\varrho}}|\le_c|\sigma|_\infty^2\mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)\mathbb{E}[\mathbb{1}_{\tau_b^P\le T}(T-\tau_b^P)^2(\mathcal{V}_{\tau_b^P}^T)^{-1}]\le_c\mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)T\mathbb{P}(\tau_b^P\le T),$$

the above probability tending to 0 as $b\to-\infty$. We have in addition convergence of the leading term $\mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P>T}]$ and the sensitivities $\mathcal{L}_3^x v_{\underline{\varrho},T}^{P,h}(0,x_0)$ in $Cor_{1,\underline{\varrho}}$ towards $v_T^{P,h}(0,x_0)$ and $\mathcal{L}_3^x v_T^{P,h}(0,x_0)$. As a conclusion, we exactly retrieve the expansion of [Benhamou 2010a, Theorem 2.1] if $b\to-\infty$.
If $b=x_0$, we have $\tau_b^P=0$ a.s. and thus $\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}]=\mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P>T}]=0$ whereas $Cor_{1,\underline{\varrho}}+Cor_{2,\underline{\varrho}}=0$. Our price approximation is coherent with the fact that the price becomes equal to zero if $b=x_0$.

7.2.3 Third order expansion

We start from the robustness formula (7.13) which we recall the expression:

$$\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}]=\mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P>T}]+\frac{1}{2}\mathbb{E}[\int_0^{T\wedge\tau_b}(\Sigma_t(X_t)-\Sigma_t)\mathcal{L}_2^x v_{\underline{\varrho},T}^{P,h}(t,X_t)\mathrm{d}t].$$

Perform a second order expansion for the function Σ at $x=X_T$ around $x=x_0$:

$$\begin{aligned}\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}]=&\mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P>T}]+\frac{1}{2}\mathbb{E}[\int_0^{T\wedge\tau_b}\{(X_t-x_0)\Sigma_t^{(1)}+\frac{1}{2}(X_t-x_0)^2\Sigma_t^{(2)}\}\mathcal{L}_2^x v_{\underline{\varrho},T}^{P,h}(t,X_t)\mathrm{d}t]\\&+\frac{1}{2}\mathbb{E}[\int_0^{T\wedge\tau_b}(X_t-x_0)^3\{\int_0^1\frac{(1-\alpha)^2}{2}\Sigma_t^{(3)}((1-\alpha)x_0+\alpha X_t)\mathrm{d}\alpha\}\mathcal{L}_2^x v_{\underline{\varrho},T}^{P,h}(t,X_t)\mathrm{d}t].\end{aligned}\quad(7.18)$$

The last term is considered like an error. For the second term we introduce $\forall 0\le t<T$ the payoff function $\psi_t(x)=(x-x_0)^2\mathcal{L}_2^x v_{\underline{\varrho},T}^{P,h}(t,x)$. Then following the arguments previously employed in the Section 7.2.2, we obtain that the second term of the r.h.s. of (7.18) is equal to:

$$\begin{aligned}&\frac{1}{2}\int_0^T\{\Sigma_t^{(1)}\mathbb{E}[\phi_t(X_t)\mathbb{1}_{\tau_b>t}]+\frac{\Sigma_t^{(2)}}{2}\mathbb{E}[\psi_t(X_t)\mathbb{1}_{\tau_b>t}]\}\mathrm{d}t\\=&\frac{1}{2}\int_0^T\{\Sigma_t^{(1)}\mathbb{E}[v_{\underline{\varrho},t}^{P,\phi_t}(t,X_t)\mathbb{1}_{\tau_b>t}]+\frac{\Sigma_t^{(2)}}{2}\mathbb{E}[v_{\underline{\varrho},t}^{P,\psi_t}(t,X_t)\mathbb{1}_{\tau_b>t}]\mathrm{d}t\\=&\frac{1}{2}\int_0^T(\Sigma_t^{(1)}v_{\underline{\varrho},t}^{P,\phi_t}(0,x_0)+\frac{1}{2}\Sigma_t^{(2)}v_{\underline{\varrho},t}^{P,\psi_t}(0,x_0))\mathrm{d}t+\frac{1}{4}\int_0^T\Sigma_t^{(1)}\mathbb{E}[\int_0^{t\wedge\tau_b}(\Sigma_s(X_s)-\Sigma_s)\mathcal{L}_2^x v_{\underline{\varrho},t}^{P,\phi_t}(s,X_s)\mathrm{d}s]\mathrm{d}t\\&+\frac{1}{8}\int_0^T\Sigma_t^{(2)}\mathbb{E}[\int_0^{t\wedge\tau_b}(\Sigma_s(X_s)-\Sigma_s)\mathcal{L}_2^x v_{\underline{\varrho},t}^{P,\psi_t}(s,X_s)\mathrm{d}s]\mathrm{d}t,\end{aligned}\quad(7.19)$$

where for any $0\le s\le t<T$ and any $x\ge b$:

$$v_{\underline{\varrho},t}^{P,\psi_t}(s,x)=\mathbb{E}[\psi_t(X_t^P)\mathbb{1}_{\tau_{s,b}^P>t}|X_s^P=x],\quad(7.20)$$

is the solution of the system (7.8) on $]0,t[\times]b,\infty[$ but having as terminal condition the regular payoff function ψ_t. The magnitude of the last term will be analysed later. To obtain a global error of amplitude

4 w.r.t. the interest parameters, we need to approximate the second term. We write performing a first order Taylor expansion:

$$
\begin{aligned}
&\frac{1}{4}\int_0^T \Sigma_t^{(1)}\mathbb{E}[\int_0^{t\wedge\tau_b}(\Sigma_s(X_s)-\Sigma_s)\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,X_s)\mathrm{d}s]\mathrm{d}t\\
=&\frac{1}{4}\int_0^T \Sigma_t^{(1)}\mathbb{E}[\int_0^{t\wedge\tau_b}(X_s-x_0)\Sigma_s^{(1)}\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,X_s)\mathrm{d}s]\mathrm{d}t\\
+&\frac{1}{4}\int_0^T \Sigma_t^{(1)}\mathbb{E}[\int_0^{t\wedge\tau_b}\{\int_0^1\Sigma_s^{(2)}((1-\alpha)x_0+\alpha X_s)(1-\alpha)\mathrm{d}\alpha\}(X_s-x_0)^2\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,X_s)\mathrm{d}s]\mathrm{d}t.
\end{aligned}
\tag{7.21}
$$

Then we neglect for the moment the last term and we refine the second term. Using the same methodology we introduce $\forall 0\le s<t<T$ the payoff functions $\rho_{s,t}(x)=(x-x_0)\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,x)$. Note that $\rho_{s,t}(b)=0$ (see (7.39) in Lemma 7.3.2.3) and as previously, for any $0\le u\le s<t<T$ and $x\ge b$ we denote by:

$$
v_{\underline{o},s}^{P,\rho_{s,t}}(u,x)=\mathbb{E}[\rho_{s,t}(X_s^P)\mathbb{1}_{\tau_{u,b}^P>s}|X_u^P=x],
\tag{7.22}
$$

the solution of (7.8) on $]0,s[\times]b,\infty[$ associated to the regular payoff function $\rho_{s,t}$. Thus we get:

$$
\begin{aligned}
&\frac{1}{4}\int_0^T \Sigma_t^{(1)}\mathbb{E}[\int_0^{t\wedge\tau_b}(X_s-x_0)\Sigma_s^{(1)}\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,X_s)\mathrm{d}s]\mathrm{d}t\\
=&\frac{1}{4}\int_0^T \Sigma_t^{(1)}(\int_0^t\Sigma_s^{(1)}v_{\underline{o},s}^{P,\rho_{s,t}}(0,x_0)\mathrm{d}s)\mathrm{d}t+\frac{1}{8}\int_0^T\Sigma_t^{(1)}(\int_0^t\Sigma_s^{(1)}\mathbb{E}[\int_0^{s\wedge\tau_b}(\Sigma_u(X_u)-\Sigma_u)\mathcal{L}_2^x v_{\underline{o},s}^{P,\rho_{s,t}}(u,X_u)\mathrm{d}u]\mathrm{d}s)\mathrm{d}t.
\end{aligned}
\tag{7.23}
$$

The explicit calculus of the new corrective terms coming from the r.h.s. of (7.19)-(7.23) is given in Lemma 7.3.4.4 whereas the error analysis is performed in Section 7.3.5. This leads to the following Theorem:

Theorem 7.2.3.1. *(**3rd order approximation price formula for down and out regular barrier options**). Assuming $(\widetilde{\mathcal{H}}_{x_0}^{\sigma})$, we have for any $h\in\underline{\mathrm{Lip}}(\mathbb{R},b)$ and any $x_0\ge b$:*

$$
\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b>T}]=\mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P>T}]+\sum_{n=1}^{5}Cor_{n,\underline{o}}+O(|\sigma|_\infty\mathcal{M}_1(\sigma)[\mathcal{M}_0(\sigma)]^2T^2),
\tag{7.24}
$$

where the corrective terms $Cor_{1,\underline{o}}$ and $Cor_{2,\underline{o}}$ are defined in Theorem 7.2.2.1, where:

$$
\begin{aligned}
Cor_{3,\underline{o}}=&\frac{1}{4}\vec{\omega}(\Sigma^{(2)},\Sigma)_0^T\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(0,x_0)+\frac{1}{2}\vec{\omega}(\Sigma^{(2)},\Sigma,\Sigma)_0^T(\mathcal{L}_4^x+\frac{1}{4}\mathcal{L}_2^x)v_{\underline{o},T}^{P,h}(0,x_0)\\
&+\frac{1}{4}\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)},\Sigma)_0^T(3\mathcal{L}_4^x+\frac{1}{2}\mathcal{L}_2^x)v_{\underline{o},T}^{P,h}(0,x_0)+\frac{1}{8}\vec{\omega}^2(\Sigma^{(1)},\Sigma)_0^T(\mathcal{L}_6^x+\frac{1}{4}\mathcal{L}_4^x)v_{\underline{o},T}^{P,h}(0,x_0),\\
Cor_{4,\underline{o}}=&\frac{1}{4}(x_0-b)\mathbb{E}\Big[\mathbb{1}_{\tau_b^P\le T}\{2[\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)})_{\tau_b^P}^T+\vec{\omega}(\Sigma^{(2)},\Sigma)_{\tau_b^P}^T+\vec{\omega}(\Sigma^{(1)})_{\tau_b^P}^T\vec{\omega}(\Sigma^{(1)},\Sigma)_0^{\tau_b^P}(\mathcal{V}_0^{\tau_b^P})^{-1}]\mathcal{L}_3^x\\
&+[2\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)},\Sigma)_{\tau_b^P}^T+\vec{\omega}(\Sigma^{(1)},\Sigma,\Sigma^{(1)})_{\tau_b^P}^T+\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^P}^T\vec{\omega}(\Sigma^{(1)},\Sigma)_0^{\tau_b^P}(\mathcal{V}_0^{\tau_b^P})^{-1}]\mathcal{L}_5^x\}v_{\underline{o},T}^{P,h}(\tau_b^P,b)\Big],\\
Cor_{5,\underline{o}}=&\frac{1}{4}\int_0^T\Sigma_r(\mathcal{V}_0^r)^{-1}\{2\vec{\omega}(\Sigma^{(1)})_r^T\mathcal{L}_3^x+\vec{\omega}(\Sigma^{(1)},\Sigma)_r^T\mathcal{L}_5^x\}v_{\underline{o},T}^{P,h}(r,b)\\
&\times\mathbb{E}\Big[\mathbb{1}_{\tau_b^P\le r}\{\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^P}^r\mathcal{V}_0^{\tau_b^P}(\mathcal{V}_{\tau_b^P}^r)^{-1}-\vec{\omega}(\Sigma^{(1)},\Sigma)_0^{\tau_b^P}-\mathcal{V}_0^{\tau_b^P}\vec{\omega}(\Sigma^{(1)})_{\tau_b^P}^r\}\mathcal{D}^P(\tau_b^P,r,0)\Big]\mathrm{d}r,
\end{aligned}
$$

and where the differential operators $\mathcal{L}^x$ are defined in Definition 7.2.1.1.

Remark 7.2.3.1. *The reader will notice that if $b\to-\infty$, then $Cor_{2,\underline{o}}+Cor_{4,\underline{o}}+Cor_{5,\underline{o}}\to 0$ and we retrieve the results of [Benhamou 2010a, Theorem 2.2], whereas if $x_0=b$, according to Lemma 7.2.1.1 equation (7.10), all the corrective terms vanish.*

7.3 Calculus of the corrective terms and error analysis

7.3.1 Preliminary results

We begin with classical but useful estimates easily obtained with standard inequalities:

Lemma 7.3.1.1. *Assume* $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$. *For any* $p \geq 1$, *we have:*

$$\sup_{t\in[0,T]} \|X_t - x_0\|_p \leq_c |\sigma|_\infty \sqrt{T}, \qquad \sup_{t\in[0,T]} \|e^{|X_t|}\|_p \leq e^{x_0} e^{C_p|\sigma|^2_\infty T},$$

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}} |\Sigma^{(1)}_t(x)| \leq_c |\sigma|_\infty \mathcal{M}_1(\sigma), \qquad \sup_{(t,x,i)\in[0,T]\times\mathbb{R}\times\{2,3\}} |\Sigma^{(i)}_t(x)| \leq_c \mathcal{M}_0(\sigma)\mathcal{M}_1(\sigma),$$

for a constant C_p *depending only on* p.

We recall well known properties of the Gaussian cumulative and density functions:

Lemma 7.3.1.2. *In this Lemma,* $\mathcal{P}$ *denotes an arbitrary polynomial function and* C *an arbitrary positive constant.*

- $x \mapsto \mathcal{P}(x)e^{Cx}\mathcal{N}'(x)$ *is a bounded function.*
- *For any* $x < 0$ *(respectively* $x > 0$*), we have* $\mathcal{N}(x) \leq \frac{\mathcal{N}'(x)}{|x|}$ *(respectively* $\overline{\mathcal{N}}(x) \leq \frac{\mathcal{N}'(x)}{x}$*) and consequently* $x \mapsto |x|\mathcal{N}(x)$ *and* $x \mapsto |x|e^{C|x|}\mathcal{N}(x)$ *(respectively* $x \mapsto x\overline{\mathcal{N}}(x)$ *and* $x \mapsto xe^{Cx}\overline{\mathcal{N}}(x)$*) are bounded functions on* $]-\infty, 0]$ *(respectively* $[0; +\infty[$*).*

We now announce a Lemma related to the spatial derivatives of $v^{P,h}_{\underline{o},T}(t,x)$ and their estimates.

Lemma 7.3.1.3. *Assume* $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$ *and that* $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$. $\forall (t,x) \in [0,T[\times[b,+\infty[$ *and* $\forall n \geq 1$, *we have:*

$$|\partial^n_{x^n} v^{P,h}_{\underline{o},T}(t,x)| \leq_c e^{C_h|x|}(\mathcal{V}^T_t)^{-\frac{n-1}{2}}, \tag{7.25}$$

for a generic constant independent of x. *We easily deduce that* $\forall t \in [0,T[$ *and for any* $p \geq 1$:

$$\|v^{P,h}_{\underline{o},T}(t,X_t)\mathbb{1}_{\tau_b>t}\|_p \leq_c C^{-(n-1)}_{\sigma}(T-t)^{-\frac{n-1}{2}}. \tag{7.26}$$

Proof. Starting from (7.9), a straightforward calculus leads to the following expressions for the spatial derivatives of $v^{P,h}_{\underline{o},T}(t,x)$, for any $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$, $\forall n \geq 1$, $\forall (t,x) \in [0,T[\times[b,+\infty[$:

$$\begin{aligned}
&\partial^n_{x^n} v^{P,h}_{\underline{o},T}(t,x)\\
=&\partial^{n-1}_{x^{n-1}}\{\int_{\mathbb{R}} h^{(1)}(y+x)\mathcal{D}^P(t,T,y)\mathrm{d}y - e^{(x-b)}\int_{\mathbb{R}} h(y)\mathcal{D}^P(t,T,y+x-2b)\mathrm{d}y\\
&+ e^{(x-b)}\int_{\mathbb{R}} h^{(1)}(y-x)\mathcal{D}^P(t,T,y-2b)\mathrm{d}y\}\\
=&(\mathcal{V}^T_t)^{-\frac{n-1}{2}}\int_{\mathbb{R}} h^{(1)}(y)H_{n-1}(\frac{(y-x+\frac{1}{2}\mathcal{V}^T_t)}{\sqrt{\mathcal{V}^T_t}})\mathcal{D}^P(t,T,y-x)\mathrm{d}y - e^{(x-b)}\int_{\mathbb{R}} h(y)\mathcal{D}^P(t,T,y+x-2b)\mathrm{d}y\\
&+ e^{(x-b)}\sum_{k=1}^{n}\binom{n}{k}\frac{(-1)^{k-1}}{(\mathcal{V}^T_t)^{\frac{k-1}{2}}}\int_{\mathbb{R}} h^{(1)}(y)H_{k-1}(\frac{(y+x-2b+\frac{1}{2}\mathcal{V}^T_t)}{\sqrt{\mathcal{V}^T_t}})\mathcal{D}^P(t,T,y+x-2b)\mathrm{d}y\\
=&(\mathcal{V}^T_t)^{-\frac{n-1}{2}}\int_{\frac{b-x+\frac{1}{2}\mathcal{V}^T_t}{\sqrt{\mathcal{V}^T_t}}}^{\infty} h^{(1)}(y+x-\frac{1}{2}\mathcal{V}^T_t)H_{n-1}(y)\mathcal{N}'(y)\mathrm{d}y - e^{(x-b)}\int_{\frac{-b+x+\frac{1}{2}\mathcal{V}^T_t}{\sqrt{\mathcal{V}^T_t}}}^{\infty} h(y-x+2b-\frac{1}{2}\mathcal{V}^T_t)\mathcal{N}'(y)\mathrm{d}y
\end{aligned}$$

$$+e^{(x-b)}\sum_{k=1}^{n}\binom{n}{k}\frac{(-1)^{k-1}}{(\mathcal{V}_t^T)^{\frac{k-1}{2}}}\int_{\frac{-b+x+\frac{1}{2}\mathcal{V}_t^T}{\sqrt{\mathcal{V}_t^T}}}^{\infty} h^{(1)}(y-x+2b-\frac{1}{2}\mathcal{V}_t^T)H_{k-1}(y)\mathcal{N}'(y)\mathrm{d}y, \tag{7.27}$$

where H_n denote the n^{th} Hermite polynomial defined for any $x\in\mathbb{R}$ by:

$$H_n(x) = -(1)^n e^{x^2/2}\partial^n_{x^n}(e^{-x^2/2}). \tag{7.28}$$

Then using the assumption on h and the fact that $x \geq b$, one obtains for any $y \geq \frac{-b+x+\frac{\mathcal{V}_t^T}{2}}{\sqrt{\mathcal{V}_t^T}} \geq 0$:

$$\left|e^{(x-b)}h^{(1)}(y-x+2b-\frac{\mathcal{V}_t^T}{2})\right|+\left|e^{(x-b)}h(y-x+2b-\frac{\mathcal{V}_t^T}{2})\right| \leq_c e^{C_h|x|}e^{(1+2C_h)(x-b)}e^{C_h y}.$$

Then, making the change of variable $z = x-b$, we show that for any $k\in\{1,\dots,n\}$ $z\mapsto e^{(1+2C_h)z}\int_{\frac{z+\frac{1}{2}\mathcal{V}_t^T}{\sqrt{\mathcal{V}_t^T}}}^{\infty} e^{C_h y}|H_{k-1}(y)|\mathcal{N}'(y)\mathrm{d}y$ is a bounded function on $\mathbb{R}_+$ tending to 0 as z tends to ∞ (i.e. b tends to $-\infty$). For z enough large we indeed have $|H_{k-1}(y)| \leq e^{C_h y}$ for any $y \geq \frac{z+\frac{1}{2}\mathcal{V}_t^T}{\sqrt{\mathcal{V}_t^T}}$ and then we can write $e^{(1+2C_h)z}\int_{\frac{z+\frac{1}{2}\mathcal{V}_t^T}{\sqrt{\mathcal{V}_t^T}}}^{\infty} e^{2C_h y}\mathcal{N}'(y)\mathrm{d}y = e^{(1+2C_h)z}e^{2C_h^2}\overline{\mathcal{N}}(\frac{z+\frac{1}{2}\mathcal{V}_t^T}{\sqrt{\mathcal{V}_t^T}}-2C_h)$. We conclude with Lemma 7.3.1.2. The result (7.25) follows without difficulty from these observations. Using the Lemma 7.3.1.1, (7.26) is a straightforward consequence of (7.25). □

Remark 7.3.1.1. *The above Lemma shows that the spatial regularity of the payoff function h allows to obtain a first spatial derivative of* $v^{P,h}_{\underline{o},T}$ *bounded at maturity. Then the next derivatives explode at maturity with the speed* $(T-t)^{\frac{n-1}{2}}$. *If h is not a regular payoff (*$h(b)\neq 0$*) but still remains a.e. differentiable, the first spatial derivative of* $v^{P,h}_{\underline{o},T}$ *becomes equal for any* $t<T$ *and any* $x\geq b$ *to:*

$$\begin{aligned}\partial_x v^{P,h}_{\underline{o},T}(t,x) = &\int_{b-x}^{\infty} h^{(1)}(y+x)\mathcal{D}^P(t,T,y)\mathrm{d}y - e^{(x-b)}\int_b^{\infty} h(y)\mathcal{D}^P(t,T,y+x-2b)\mathrm{d}y\\ &+e^{(x-b)}\int_{b+x}^{\infty} h^{(1)}(y-x)\mathcal{D}^P(t,T,y-2b)\mathrm{d}y + h(b)\{\mathcal{D}^P(t,T,b-x)+e^{x-b}\mathcal{D}^P(t,T,x-b)\}.\end{aligned}$$

If h is not anymore a.e. differentiable, the derivative writes for any $t<T$ *and any* $x\geq b$*:*

$$\begin{aligned}\partial_x v^{P,h}_{\underline{o},T}(t,x) =&(\mathcal{V}_t^T)^{-\frac{1}{2}}\int_b^{\infty} h(y)\frac{(y-x+\frac{1}{2}\mathcal{V}_t^T)}{\sqrt{\mathcal{V}_t^T}}\mathcal{D}^P(t,T,y-x)\mathrm{d}y - e^{(x-b)}\int_b^{\infty} h(y)\mathcal{D}^P(t,T,y+x-2b)\mathrm{d}y\\ &+(\mathcal{V}_t^T)^{-\frac{1}{2}}e^{(x-b)}\int_b^{\infty} h(y)\frac{(y+x-2b+\frac{1}{2}\mathcal{V}_t^T)}{\sqrt{\mathcal{V}_t^T}}\mathcal{D}^P(t,T,y+x-2b)\mathrm{d}y,\end{aligned}$$

Consequently, whatever is the regularity of h, $v^{P,h}_{\underline{o},T}(t,x)$ *explodes at maturity and at the barrier with the speed* $\sqrt{T-t}$. *For instance, we could replace* $(\mathcal{V}_t^T)^{-\frac{n-1}{2}}$ *by* $(\mathcal{V}_t^T)^{-\frac{n}{2}}$ *in (7.25) for a non regular payoff.*

7.3.2 Calculus of $v^{P,\phi_t}_{o,t}$ and estimate of its spatial derivatives

The first step is to make explicit $v^{P,\phi_t}_{\underline{o},t}$ defined in (7.15) and to link it to the derivatives of the Proxy kernel pricing $v^{P,h}_{\underline{o},T}$. This is the purpose of the next Lemma:

Lemma 7.3.2.1. *Assume* ($\widetilde{\mathcal{H}}^{\sigma}_{x_0}$) *and that* $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$. $\forall 0 \le s \le t < T$ *and* $\forall x \ge b$, *we have:*

$$v^{P,\phi_t}_{\underline{o},t}(s,x) = (x - x_0)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(s,x) + \mathcal{V}^t_s \mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(s,x) - \mathbb{E}[\mathbb{1}_{\tau^P_{s,b} \le t} \mathcal{V}^t_{\tau^P_{s,b}} \mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(\tau^P_{s,b}, b) | X^P_s = x]. \tag{7.29}$$

We deduce that for any $x_0 \ge b$, *we have:*

$$\frac{1}{2}\int_0^T \Sigma^{(1)}_t v^{P,\phi_t}_{\underline{o},t}(0,x_0)\mathrm{d}t = \frac{1}{2}\vec{\omega}(\Sigma^{(1)},\Sigma)^T_0 \mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(0,x_0) - \frac{1}{2}\mathbb{E}[\mathbb{1}_{\tau^P_b \le T}\vec{\omega}(\Sigma^{(1)},\Sigma)^T_{\tau^P_b}\mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(\tau^P_b, b)].$$

Remark 7.3.2.1. *In the expression* (7.29), *we retrieve the fact that* $\forall 0 \le s \le t < T$, $v^{P,\phi_t}_{\underline{o},t}(s,b) = 0$. *By Lemma* 7.2.1.1 *equation* (7.10), *we indeed have* $\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(s,b) = 0$ *and the fact that* $\tau^P_{s,b} = s$ *a.s. if* $X^P_s = b$ *allows to cancel all the remaining terms of the r.h.s. of* (7.29).

Proof. By definition, we have, $\forall x > b$ (we trivially obtain 0 if $x = b$):

$$\begin{aligned}
v^{P,\phi_t}_{\underline{o},t}(s,x) =& \mathbb{E}[\phi_t(X^P_t)\mathbb{1}_{\tau^P_{s,b} > t}|X^P_s = x] = \mathbb{E}[(X^P_t - x_0)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t, X^P_t)\mathbb{1}_{\tau^P_{s,b} > t}|X^P_s = x]\\
=& \mathbb{E}[(X^P_{t\wedge\tau^P_{s,b}} - x_0)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t\wedge\tau^P_{s,b}, X^P_{t\wedge\tau^P_{s,b}})|X^P_s = x] - (b - x_0)\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\le t}\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(\tau^P_{s,b}, b)|X^P_s = x]\\
=& \mathbb{E}[(X^P_{t\wedge\tau^P_{s,b}} - x_0)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t\wedge\tau^P_{s,b}, X^P_{t\wedge\tau^P_{s,b}})|X^P_s = x],
\end{aligned}$$

where we used at the last line the fact that $\phi_t(b) = 0$ (see (7.10)). Then apply the Itô's Lemma for the product $(X^P_{t\wedge\tau^P_{s,b}} - x_0)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t\wedge\tau^P_{s,b}, X^P_{t\wedge\tau^P_{s,b}})$ to obtain:

$$\begin{aligned}
&(X^P_{t\wedge\tau^P_{s,b}} - x_0)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t\wedge\tau^P_{s,b}, X^P_{t\wedge\tau^P_{s,b}})\\
=&(X^P_s - x_0)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(s, X^P_s) + \int_s^{t\wedge\tau^P_{s,b}} (X^P_u - x_0)(\partial_x \circ \mathcal{L}^x_2) v^{P,h}_{\underline{o},T}(u, X^P_u)\sigma_u \mathrm{d}W_u\\
&+ \int_s^{t\wedge\tau^P_{s,b}} (X^P_u - x_0)(\partial_t + \frac{1}{2}\Sigma_u \mathcal{L}^x_2)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(u, X^P_u)\mathrm{d}u\\
&+ \int_s^{t\wedge\tau^P_{s,b}} \mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(u, X^P_u)\sigma_u(\mathrm{d}W_u - \frac{\sigma_u}{2}\mathrm{d}u) + \int_s^{t\wedge\tau^P_{s,b}} (\partial_x \circ \mathcal{L}^x_2) v^{P,h}_{\underline{o},T}(u, X^P_u)\Sigma_u \mathrm{d}u.
\end{aligned}$$

Then taking the expectation with the conditional information that $X^P_s = x > b$, the simplifications coming from the PDE solved by $v^{P,h}_{\underline{o},T}$ (7.8) on the domain $]0,T[\times]b,\infty[$ and the fact that $(\partial_x - \frac{1}{2}\mathcal{I}) \circ \mathcal{L}^x_2 = \mathcal{L}^x_1 \circ \mathcal{L}^x_2 = \mathcal{L}^x_3$ (see Definition 7.2.1.1) yield:

$$\mathbb{E}[(X^P_{t\wedge\tau^P_{s,b}} - x_0)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t\wedge\tau^P_{s,b}, X^P_{t\wedge\tau^P_{s,b}})|X^P_s = x] = (x - x_0)\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(s,x) + \mathbb{E}[\int_s^{t\wedge\tau^P_{s,b}} \mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(u, X^P_u)\Sigma_u \mathrm{d}u \big| X^P_s = x].$$

For the second term we use again a decomposition to get:

$$\begin{aligned}
&\mathbb{E}[\int_s^{t\wedge\tau^P_{s,b}} \mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(u, X^P_u)\Sigma_u \mathrm{d}u \big| X^P_s = x] = \mathbb{E}[\int_s^t \mathbb{1}_{\tau^P_{s,b} > u}\mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(u, X^P_u)\Sigma_u \mathrm{d}u \big| X^P_s = x]\\
=&\mathbb{E}[\int_s^t \mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(u\wedge\tau^P_{s,b}, X^P_{u\wedge\tau^P_{s,b}})\Sigma_u \mathrm{d}u \big| X^P_s = x] - \mathbb{E}[\int_s^t \mathbb{1}_{\tau^P_{s,b}\le u}\mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(\tau^P_{s,b}, b)\Sigma_u \mathrm{d}u \big| X^P_s = x]\\
=&\mathcal{V}^t_s \mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(s,x) - \mathbb{E}[\mathbb{1}_{\tau^P_{s,b} < t}\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(\tau^P_{s,b}, b)|X^P_s = x],
\end{aligned}$$

where we applied at the last equality the Lemma 7.2.1.2. That achieves the proof of the first statement of the Lemma. The second statement easily follows. □

The next Lemma provides the estimate of $\mathcal{L}_2^x\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\le t}\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x]$:

Lemma 7.3.2.2. *Assume $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$ and that $h\in \underline{\mathrm{Lip}}(\mathbb{R},b)$. $\forall 0\le s<t<T$, $\forall x\ge b$, we have:*

$$\mathcal{L}_2^x\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\le t}\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x] =\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\le t}(2\mathcal{L}_3^x+\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x]. \tag{7.30}$$

In addition, we have the following estimates $\forall 0\le s<t<T$, $\forall x\ge b$:

$$\left|\mathcal{L}_2^x\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\le t}\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x]\right| \le_c e^{C_h|x|}\{(\mathcal{V}_s^{\frac{s+t}{2}})^{-\frac12}(\mathcal{V}^T_t)^{-\frac12}+(\mathcal{V}^T_{\frac{s+t}{2}})^{-1}\}, \tag{7.31}$$

$$(x-b)\left|\mathcal{L}_2^x\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\le t}\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x]\right| \le_c e^{C_h|x|}\{(\mathcal{V}^T_t)^{-\frac12}+(\mathcal{V}^T_0)^{\frac12}(\mathcal{V}^T_{\frac{s+t}{2}})^{-1}\}, \tag{7.32}$$

where the generic constants are independent of x.

Proof. For the first part, using the explicit form of density of the hitting times (7.6) and the relation (7.62) of Proposition 7.6.1.1, a straightforward calculus leads to, $\forall 0\le s<t<T$ and $\forall x>b$:

$$\begin{aligned}\mathcal{L}_2^x\mathbb{E}[\mathcal{V}^t_{\tau^P_{s,b}}\mathbb{1}_{\tau^P_{s,b}\le t}\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x] &= \mathcal{L}_2^x\{\int_s^t \mathcal{J}^P(s,r,x,b)\mathcal{V}^t_r\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\}\\ =\int_s^t \mathcal{L}_2^x\{\mathcal{J}^P(s,r,x,b)\}\mathcal{V}^t_r\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r \quad &= 2\int_s^t \partial_r\{\Sigma_r^{-1}\mathcal{J}^P(s,r,x,b)\}\mathcal{V}^t_r\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r.\end{aligned}$$

Then we perform an integration by parts, using the fact that $\mathcal{J}^P(s,r,x,b)$ (respectively $\mathcal{V}^t_r\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(r,b)$) tends to 0 as r tends to s (respectively to t), to obtain:

$$\begin{aligned}\mathcal{L}_2^x\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\le t}\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x] &= -2\int_s^t \Sigma_r^{-1}\mathcal{J}^P(s,r,x,b)\partial_r\{\mathcal{V}^t_r\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(r,b)\}\mathrm{d}r\\ =\int_s^t \mathcal{J}^P(s,r,x,b)(2\mathcal{I}-\frac{2}{\Sigma_r}\mathcal{V}^t_r\partial_r)\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r \quad &= \int_s^t \mathcal{J}^P(s,r,x,b)(2\mathcal{I}+\mathcal{V}^t_r\mathcal{L}_2^x)\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(r,x)\mathrm{d}r,\\ =\int_s^t \mathcal{J}^P(s,r,x,b)\{2\mathcal{L}_3^x+\mathcal{V}^t_r\mathcal{L}_5^x\}v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r,&\end{aligned}$$

using the PDE (7.8) solved by $v^{P,h}_{\underline{o},T}$ and the Definition of $\mathcal{L}_5^x=\mathcal{L}_3^x\circ\mathcal{L}_2^x$ (see Definition 7.2.1.1). That achieves the proof of (7.30). Note that up to a passing to the limit, the formula (7.30) remains valid if x tends to b: one has to remove the expectations and replace $\tau^P_{s,b}$ by s.
For (7.31), we suppose that $x>b$, otherwise a straightforward application of Lemma 7.3.1.3 gives:

$$\begin{aligned}\left|\mathcal{L}_2^x\{\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\le t}\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x]\}|_{x=b}\right| &=\left|2\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(s,b)+\mathcal{V}^t_s\mathcal{L}_5^x v^{P,h}_{\underline{o},T}(s,b)\right|\\ &\le_c e^{C_h|x|}\{(\mathcal{V}^T_s)^{-1}+\mathcal{V}^t_s(\mathcal{V}^T_s)^{-2}\}\le_c e^{C_h|x|}(\mathcal{V}^T_s)^{-1}.\end{aligned}$$

Now, assuming that $x>b$, we only show the estimate for $\int_s^t\mathcal{J}^P(s,r,x,b)\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r$, the treatment of the term with $\mathcal{L}_5^x$ being similar. We could use the estimate (7.25) to directly get:

$$\left|\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\le t}\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x]\right|\le_c \mathbb{P}(\tau^P_{s,b}\le t|X^P_s=x)e^{C_h|x|}(\mathcal{V}^T_t)^{-1}\le_c e^{C_h|x|}(\mathcal{V}^T_t)^{-1},$$

but problems will arise in the error estimate of Theorem 7.2.2.1 because
$\int_0^T\int_0^t(T-t)^{-1}\mathrm{d}s\mathrm{d}t=\int_0^T\frac{t}{T-t}\mathrm{d}t=\infty$. To overcome this difficulty we split the domain of integration $[s,t]$ by writing $[s,t]=[s,\frac{s+t}{2}]\cup[\frac{s+t}{2},t]$. For the first part, we apply Lemma 7.3.1.3 to obtain:

$$\left|\int_s^{\frac{s+t}{2}}\mathcal{J}^P(s,r,x,b)\mathcal{L}_3^x v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\right|\le_c\mathbb{P}(\tau^P_{s,b}\le\frac{t+s}{2}|X^P_s=x)e^{C_h|x|}(\mathcal{V}^T_{\frac{s+t}{2}})^{-1}\le_c e^{C_h|x|}(\mathcal{V}^T_{\frac{s+t}{2}})^{-1}. \tag{7.33}$$

For the second one, use the estimate (7.65) in Proposition 7.6.1.2, the estimate (7.25) and the Cauchy Schwarz inequality to get:

$$\Big|\int_{\frac{s+t}{2}}^{t} \mathcal{J}^P(s,r,x,b)\mathcal{L}_3^x v_{\underline{o},T}^{P,h}(r,b)\mathrm{d}r\Big| \tag{7.34}$$

$$\leq_c e^{C_h|x|}\int_{\frac{s+t}{2}}^{t}\Sigma_r(\mathcal{V}_s^r)^{-1}(\mathcal{V}_r^T)^{-1}\mathrm{d}r \leq_c e^{C_h|x|}\sqrt{\int_{\frac{s+t}{2}}^{t}\Sigma_r(\mathcal{V}_s^r)^{-2}\mathrm{d}r}\sqrt{\int_{\frac{s+t}{2}}^{t}\Sigma_r(\mathcal{V}_r^T)^{-2}\mathrm{d}r} \leq_c e^{C_h|x|}(\mathcal{V}_s^{\frac{s+t}{2}})^{-\frac{1}{2}}(\mathcal{V}_t^T)^{-\frac{1}{2}},$$

what allows us to conclude. It remains to show the bound (7.32). First we have similarly to (7.33) using (7.65):

$$\Big|\int_{s}^{\frac{s+t}{2}} (x-b)\mathcal{J}^P(s,r,x,b)\mathcal{L}_3^x v_{\underline{o},T}^{P,h}(r,b)\mathrm{d}r\Big| \leq_c e^{C_h|x|}(\mathcal{V}_{\frac{s+t}{2}}^T)^{-1}\int_s^{\frac{s+t}{2}}\Sigma_r(\mathcal{V}_s^r)^{-\frac{1}{2}}\mathrm{d}r \leq_c e^{C_h|x|}(\mathcal{V}_{\frac{s+t}{2}}^T)^{-1}(\mathcal{V}_s^{\frac{s+t}{2}})^{\frac{1}{2}}. \tag{7.35}$$

Then we refine the estimate (7.34) to achieve the proof:

$$\Big|\int_{\frac{s+t}{2}}^{t}(x-b)\mathcal{J}^P(s,r,x,b)\mathcal{L}_3^x v_{\underline{o},T}^{P,h}(r,b)\mathrm{d}r\Big| \leq_c e^{C_h|x|}\int_{\frac{s+t}{2}}^{t}\Sigma_r(\mathcal{V}_s^r)^{-\frac{1}{2}}(\mathcal{V}_r^T)^{-1}\mathrm{d}r \leq_c e^{C_h|x|}(\mathcal{V}_t^T)^{-\frac{1}{2}}. \tag{7.36}$$

□

We now state in the following Lemma the estimate of $\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,x)$:

Lemma 7.3.2.3. *Assume* $(\widetilde{\mathcal{H}}_{x_0}^{\sigma})$ *and that* $h \in \underline{\mathrm{Lip}}(\mathbb{R},b)$. $\forall 0 \leq s < t < T$ *and* $\forall x \geq b$, *we have:*

$$\left|\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,x)\right| \leq_c e^{C_h|x|}\{|x-x_0|(\mathcal{V}_s^T)^{-\frac{3}{2}} + (\mathcal{V}_s^T)^{-1} + (\mathcal{V}_s^{\frac{s+t}{2}})^{-\frac{1}{2}}(\mathcal{V}_t^T)^{-\frac{1}{2}} + (\mathcal{V}_{\frac{s+t}{2}}^T)^{-1}\}, \tag{7.37}$$

where the generic constant is independent of x. We deduce that for any $0 \leq s < t < T$ *and for any* $p \geq 1$*:*

$$\|\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,X_s)\mathbb{1}_{\tau_b > s}\|_p \leq_c C_\sigma^{-2}\{\frac{\sqrt{T}}{(T-s)^{\frac{3}{2}}} + \frac{1}{T-s} + \frac{1}{\sqrt{(t-s)}\sqrt{T-t}} + \frac{1}{T-\frac{(s+t)}{2}}\} \tag{7.38}$$

In addition, we have:

$$\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,b) = 0. \tag{7.39}$$

Proof. Combining Lemmas 7.3.2.1 and 7.3.2.2, we easily get $\forall 0 \leq s < t < T$ and $\forall x \geq b$:

$$\begin{aligned}\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s,x) =& 2(\mathcal{L}_1^x \circ \mathcal{L}_2^x)v_{\underline{o},T}^{P,h}(s,x) + (x-x_0)(\mathcal{L}_2^x \circ \mathcal{L}_2^x)v_{\underline{o},T}^{P,h}(s,x) + \mathcal{V}_s^t(\mathcal{L}_2^x \circ \mathcal{L}_3^x)v_{\underline{o},T}^{P,h}(s,x)\\ &- \mathbb{E}[\mathbb{1}_{\tau_{s,b}^P \leq t}(2\mathcal{L}_3^x + \mathcal{V}_{\tau_{s,b}^P}^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(\tau_{s,b}^P,b)|X_s^P = x]\\ =& 2\mathcal{L}_3^x v_{\underline{o},T}^{P,h}(s,x) + (x-x_0)\mathcal{L}_4^x v_{\underline{o},T}^{P,h}(s,x) + \mathcal{V}_s^t\mathcal{L}_5^x v_{\underline{o},T}^{P,h}(s,x)\\ &- \mathbb{E}[\mathbb{1}_{\tau_{s,b}^P \leq t}(2\mathcal{L}_3^x + \mathcal{V}_{\tau_{s,b}^P}^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(\tau_{s,b}^P,b)|X_s^P = x],\end{aligned} \tag{7.40}$$

using the Definition 7.2.1.1 of the differential operators $\mathcal{L}^x$. Then the estimate (7.37) is easily obtained with the estimates (7.31) of Lemma 7.3.2.2 and (7.25) of Lemma 7.3.1.3. The estimate (7.38) directly follows from (7.37) using the Lemma 7.3.1.1. For (7.39), use the Lemma 7.2.1.1 equation (7.10) to write that $\mathcal{L}_4^x v_{\underline{o},T}^{P,h}(s,b) = 0$ and the result becomes obvious. □

7.3.3 Proof of the error estimate in Theorem 7.2.2.1

Assuming $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$ and that $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$, we obtain in view of (7.14) and (7.16) that the error term in Theorem 7.2.2.1 is equal to:

$$\begin{aligned}\mathrm{Error}_{2,\mathrm{h}} =& \frac{1}{2}\mathbb{E}[\int_0^{T\wedge\tau_b^P}(X_t - x_0)^2\{\int_0^1(1-\alpha)\Sigma_t^{(2)}((1-\alpha)x_0 + \alpha X_t)\mathrm{d}\alpha\}\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t, X_t)\mathrm{d}t] \\ &+ \frac{1}{4}\int_0^T \Sigma_t^{(1)}\mathbb{E}[\int_0^{t\wedge\tau_b^P}(\Sigma_s(X_s) - \Sigma_s)\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s, X_s)\mathrm{d}s]\mathrm{d}t. \end{aligned} \tag{7.41}$$

Then use Lemmas 7.3.1.3 and 7.3.1.1 and standard inequalities to get for first term of the r.h.s. of (7.41):

$$\begin{aligned} &\left|\mathbb{E}[\int_0^{T\wedge\tau_b^P}(X_t - x_0)^2\{\int_0^1(1-\alpha)\Sigma_t^{(2)}((1-\alpha)x_0 + \alpha X_t)\mathrm{d}\alpha\}\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t, X_t)\mathrm{d}t]\right| \\ \le_c& \mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)\sup_{t\in[0,T]}\|X_t - x_0\|_4^2\int_0^T\|\mathbb{1}_{\tau_b^P>t}\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t, X_t)\|_2\mathrm{d}t \\ \le_c& \mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)|\sigma|_\infty^2 T C_\sigma^{-1}\int_0^T\frac{\mathrm{d}t}{\sqrt{T-t}} \le_c |\sigma|_\infty\mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)T^{\frac{3}{2}}. \end{aligned}$$

Regarding the second term of the r.h.s. of (7.41), utilize the Lemma 7.3.2.3 equation (7.38) to get:

$$\begin{aligned} &\left|\int_0^T\Sigma_t^{(1)}\mathbb{E}[\int_0^{t\wedge\tau_b^P}(\Sigma_s(X_s) - \Sigma_s)\mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s, X_s)\mathrm{d}s]\mathrm{d}t\right| \\ \le_c& \mathcal{M}_1(\sigma)|\sigma|_\infty\sup_{s\in[0,T]}\|\Sigma_s(X_s) - \Sigma_s\|_2 C_\sigma^{-2}\int_0^T\int_0^t\{\frac{\sqrt{T}}{(T-s)^{\frac{3}{2}}} + \frac{1}{T-s} + \frac{1}{\sqrt{(t-s)}\sqrt{T-t}} + \frac{1}{T-\frac{(s+t)}{2}}\}\mathrm{d}s\mathrm{d}t \\ \le_c& [\mathcal{M}_1(\sigma)]^2\frac{|\sigma|_\infty^3}{C_\sigma^2}T^{\frac{3}{2}} \le_c |\sigma|_\infty[\mathcal{M}_1(\sigma)]^2 T^{\frac{3}{2}}. \end{aligned}$$

We have finished the proof.

Remark 7.3.3.1. *We would like to point out that the singular term* $\frac{\sqrt{T}}{(T-s)^{\frac{3}{2}}} + \frac{1}{T-s} + \frac{1}{\sqrt{(t-s)}\sqrt{T-t}} + \frac{1}{T-\frac{(s+t)}{2}}$ *appearing in the above double integral remains fortunately integrable. For non regular payoff functions (Call payoffs with strike lower than the barrier for instance) or for digital options, the first spatial derivative of* $v_{\underline{o},T}^{P,h}$ *may explode at maturity and the singularities arising in iterated integrals are not anymore integrable. Our approach seems inappropriate and this is the reason why we restrict ourself to regular and a.e. once time differentiable payoff function.*
In case of trivial barrier level (b tends to $-\infty$*) and assuming the strong ellipticity condition* $\inf_{(t,x)\in[0,T]\times\mathbb{R}}\sigma(t,x) > 0$ $(\mathcal{H}^\sigma)$*, the reader familiar with the Malliavin calculus will notice that Malliavin integration by parts could be performed to handle terms of the form* $\mathbb{E}[(X_t - x_0)^k\Sigma_t^{(j)}(X_t)\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t, X_t)]$ *by reducing the order (and thus the irregularities) of the derivatives applied to* $v_{\underline{o},T}^{P,h}$ *and thus to treat the case of binary payoff functions and retrieve the results of [Benhamou 2010a, Theorem 2.1 and 4.3].*

7.3.4 Calculus of $v_{o,t}^{P,\psi_t}$, $v_{o,s}^{P,\rho_{s,t}}$ and estimates of their derivatives

We begin with the calculus of $v_{\underline{o},t}^{P,\psi_t}$ defined in (7.20):

Lemma 7.3.4.1. *Assume* $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$ *and that* $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$*.* $\forall 0 \le s \le t < T$ *and* $\forall x \ge b$*, we have:*

$$v_{\underline{o},t}^{P,\psi_t}(s,x) = \{(x-x_0)^2\mathcal{L}_2^x + \mathcal{V}_s^t\mathcal{L}_2^x + 2(x-x_0)\mathcal{V}_s^t\mathcal{L}_3^x + \vec{\omega}(\Sigma,\Sigma)_s^t(2\mathcal{L}_4^x + \frac{1}{2}\mathcal{L}_2^x)\}v_{\underline{o},T}^{P,h}(s,x)$$

$$+2(x_0-b)\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\leq t}\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x]. \tag{7.42}$$

Proof. We follow the proof of Lemma 7.3.2.1. By definition of $v^{P,\psi_t}_{\underline{o},t}$, an application of Lemma 7.2.1.1 equation (7.10) gives for any $0\leq s\leq t<T$ and $\forall x\geq b$:

$$v^{P,\psi_t}_{\underline{o},t}(s,x)=\mathbb{E}[(X^P_{t\wedge\tau^P_{s,b}}-x_0)^2\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t\wedge\tau^P_{s,b},X^P_{t\wedge\tau^P_{s,b}})|X^P_s=x].$$

The application of the Itô's formula for the product $(X^P_{t\wedge\tau^P_{s,b}}-x_0)^2\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t\wedge\tau^P_{s,b},X^P_{t\wedge\tau^P_{s,b}})$ and simplifications coming from PDE (7.8) yield:

$$\begin{aligned}
&\mathbb{E}[(X^P_{t\wedge\tau^P_{s,b}}-x_0)^2\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(t\wedge\tau^P_{s,b},X^P_{t\wedge\tau^P_{s,b}})|X^P_s=x]\\
=&(x-x_0)^2\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(s,x)+2\mathbb{E}[\int_s^{t\wedge\tau^P_{s,b}}(X^P_u-x_0)(\mathcal{L}^x_1\circ\mathcal{L}^x_2)v^{P,h}_{\underline{o},T}(u,X^P_u)\Sigma_u\mathrm{d}u\big|X^P_s=x]\\
&+\mathbb{E}[\int_s^{t\wedge\tau^P_{s,b}}\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(u,X^P_u)\Sigma_u\mathrm{d}u\big|X^P_s=x].
\end{aligned}$$

Using the fact that $\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(u,b)=0$ and the Lemma 7.2.1.2, the last term of the r.h.s. is obviously equal to $\mathcal{V}^t_s\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(s,x)$. For the second term, use the Definition 7.2.1.1 to write $\mathcal{L}^x_1\circ\mathcal{L}^x_2=\mathcal{L}^x_3$, then the classical decomposition, the Itô's Lemma, the PDE (7.8) and the Lemma 7.2.1.2 give:

$$\begin{aligned}
&\mathbb{E}[\int_s^{t\wedge\tau^P_{s,b}}(X^P_u-x_0)\mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(u,X^P_u)\Sigma_u\mathrm{d}u\big|X^P_s=x]\\
=&(x_0-b)\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\leq t}\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x]+(x-x_0)\mathcal{V}^t_s\mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(s,x)\\
&+\mathbb{E}[\int_s^t\Sigma_u(\int_s^{u\wedge\tau^P_{s,b}}(\mathcal{L}^x_1\circ\mathcal{L}^x_3)v^{P,h}_{\underline{o},T}(r,X^P_r)\Sigma_r\mathrm{d}r)\mathrm{d}u\big|X^P_s=x].
\end{aligned}$$

Then the relations $\mathcal{L}^x_1\circ\mathcal{L}^x_3=\mathcal{L}^x_4+\frac{1}{4}\mathcal{L}^x_2$ (see Definition 7.2.1.1) and $(\mathcal{L}^x_4+\frac{1}{4}\mathcal{L}^x_2)v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)=0$ (see Lemma 7.2.1.1 equation (7.10)) give that the last term of the above r.h.s. is equal to $\vec{\omega}(\Sigma,\Sigma)^t_s(\mathcal{L}^x_4+\frac{1}{4}\mathcal{L}^x_2)v^{P,h}_{\underline{o},T}(s,x)$. Combining all the contributions achieves the proof. □

Lemma 7.3.4.2. *Assume $(\widetilde{\mathcal{H}^\sigma_{x_0}})$ and that $h\in\underline{\mathrm{Lip}}(\mathbb{R},b)$. $\forall 0\leq s<t<T$ and $\forall x\geq b$, we have:*

$$\begin{aligned}
\left|\mathcal{L}^x_2 v^{P,\psi_t}_{\underline{o},t}(s,x)\right|\leq_c e^{C_h|x|}\{&|x-x_0|^2(\mathcal{V}^T_s)^{-\frac{3}{2}}+|x-x_0|(\mathcal{V}^T_s)^{-1}+(\mathcal{V}^T_t)^{-\frac{1}{2}}+|x-x_0|(\mathcal{V}^{\frac{s+t}{2}}_s)^{-\frac{1}{2}}(\mathcal{V}^T_t)^{-\frac{1}{2}}\\
&+[(\mathcal{V}^T_0)^{\frac{1}{2}}+|x-x_0|](\mathcal{V}^T_{\frac{s+t}{2}})^{-1}\},
\end{aligned}\tag{7.43}$$

for a generic constant independent of x. We deduce that $\forall 0\leq s<t<T$, $\forall x\geq b$ and for any $p\geq 1$:

$$\|\mathcal{L}^x_2 v^{P,\psi_t}_{\underline{o},t}(s,X_s)\mathbb{1}_{\tau_b>s}\|_p\leq_c C^{-1}_\sigma\{\frac{T}{(T-s)^{\frac{3}{2}}}+\frac{\sqrt{T}}{T-s}+\frac{1}{\sqrt{T-t}}+\frac{\sqrt{T}}{\sqrt{t-s}\sqrt{T-t}}+\frac{\sqrt{T}}{T-\frac{(s+t)}{2}}\}. \tag{7.44}$$

Proof. Using Lemmas 7.3.4.1 and 7.3.2.2, we easily obtain $\forall 0\leq s<t<T$, $\forall x\geq b$:

$$\begin{aligned}
\mathcal{L}^x_2 v^{P,\psi_t}_{\underline{o},t}(s,x)=&\{(x-x_0)^2(\mathcal{L}^x_2\circ\mathcal{L}^x_2)+4(x-x_0)(\mathcal{L}^x_1\circ\mathcal{L}^x_2)+2\mathcal{L}^x_2+\mathcal{V}^t_s(\mathcal{L}^x_2\circ\mathcal{L}^x_2)+2(x-x_0)\mathcal{V}^t_s(\mathcal{L}_2\circ\mathcal{L}^x_3)\\
&+4\mathcal{V}^t_s(\mathcal{L}^x_1\circ\mathcal{L}^x_3)+\vec{\omega}(\Sigma,\Sigma)^t_s[2(\mathcal{L}^x_2\circ\mathcal{L}^x_4)+\frac{1}{2}(\mathcal{L}^x_2\circ\mathcal{L}^x_2)]\}v^{P,h}_{\underline{o},T}(s,x)\\
&+2(x_0-b)\mathbb{E}[\mathbb{1}_{\tau^P_{s,b}\leq t}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_{s,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{s,b},b)|X^P_s=x].
\end{aligned}$$

Then writing $x_0-b=x-b+x_0-x$, the announced estimates (7.43) and (7.44) directly follow from Lemmas 7.3.2.2, 7.3.1.3 and 7.3.1.1. We skip further details. □

In the next Lemma, we give a useful Martingale property of $v^{P,\phi_t}_{\underline{o},t}$:

Lemma 7.3.4.3. *For any $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$, for any $0 \le u \le s < t < T$, for any $x \ge b$ and for any $n \in \mathbb{N}$, we have:*

$$\mathbb{E}[\partial^n_{x^n} v^{P,\phi_t}_{\underline{o},t}(s \wedge \tau^P_{u,b}, X^P_{s\wedge\tau^P_{u,b}})|X^P_u = x] = \partial^n_{x^n} v^{P,\phi_t}_{\underline{o},t}(u, x).$$

Proof. The proof is similar to the proof of Lemma 7.2.1.2, so we skip it. □

We now give the explicit calculus of $v^{P,\rho_{s,t}}_{\underline{o},s}$ defined in (7.22) in the following Lemma which proof, a little bit tedious to write, is postponed to Appendix 7.6.2:

Lemma 7.3.4.4. *Assume $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$ and that $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$. $\forall 0 \le u \le s < t < T$, $\forall x \ge b$, we have:*

$$\begin{aligned}
&v^{P,\rho_{s,t}}_{\underline{o},s}(u,x)\\
=&(x-x_0)\{(2\mathcal{L}^x_3 + (x-x_0)\mathcal{L}^x_4 + \mathcal{V}^t_u\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(u,x) - \mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le t}(2\mathcal{L}^x_3 + \mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{u,b}, b)|X^P_u = x]\}\\
&+\mathcal{V}^s_u\{3\mathcal{L}^x_4 + \frac{1}{2}\mathcal{L}^x_2 + (x-x_0)\mathcal{L}^x_5 + \mathcal{V}^t_u(\mathcal{L}^x_6 + \frac{1}{4}\mathcal{L}^x_4)\}v^{P,h}_{\underline{o},T}(u,x)\\
&+(x_0-b)\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\mathcal{L}^x_5 v^{P,h}_{\underline{o},T}(\tau^P_{u,b}, b)|X^P_u = x] + (x-b)\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}(2\mathcal{L}^x_3 + \mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{u,b}, b)|X^P_u = x]\\
&-\mathcal{V}^s_u\mathcal{L}^x_1\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\in[s,t]}(2\mathcal{L}^x_3 + \mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{u,b}, b)|X^P_u = x]\\
&+\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\int_s^t \frac{\Sigma_r}{\mathcal{V}^r_{\tau^P_{u,b}}}\mathcal{D}^P(\tau^P_{u,b}, r, 0)(2\mathcal{L}^x_3 + \mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\big|X^P_u = x],
\end{aligned}\tag{7.45}$$

where using the relation (7.64) in Proposition 7.6.1.1,
$\mathcal{V}^s_u\mathcal{L}^x_1\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\in[s,t]}(2\mathcal{L}^x_3 + \mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{u,b}, b)|X^P_u = x]$ is equal to:

$$\mathcal{V}^s_u\int_s^t \Sigma_r(\frac{1}{\mathcal{V}^r_u} - \frac{(x-b)^2}{(\mathcal{V}^r_u)^2})\mathcal{D}^P(u, r, b-x)(2\mathcal{L}^x_3 + \mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r.$$

We deduce that for any $x_0 \ge b$, using the Definition of the corrective terms $Cor_{k,\underline{o}}$ in Theorem 7.2.3.1:

$$\frac{1}{4}\int_0^T \Sigma^{(2)}_t v^{P,\psi_t}_{\underline{o},t}(0,x_0)\mathrm{d}t + \frac{1}{4}\int_0^T \Sigma^{(1)}_t(\int_0^t \Sigma^{(1)}_s v^{P,\rho_{s,t}}_{\underline{o},s}(0,x_0)\mathrm{d}s)\mathrm{d}t = \sum_{n=3}^{5} Cor_{n,\underline{o}}.\tag{7.46}$$

We finally provide an estimate of $\mathcal{L}^x_2 v^{P,\rho_{s,t}}_{\underline{o},s}(u,x)$, the proof being performed in Appendix 7.6.2.

Lemma 7.3.4.5. *Assume $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$ and that $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$. $\forall 0 \le u < s < t < T$ and $\forall x \ge b$, we have the following estimate:*

$$\begin{aligned}
&\left|\mathcal{L}^x_2 v^{P,\rho_{s,t}}_{\underline{o},s}(u,x)\right| \qquad (7.47)\\
\le_c & e^{C_h|x|}\{|x-x_0|^2(\mathcal{V}^T_u)^{-\frac{5}{2}} + |x-x_0|[(\mathcal{V}^T_{\frac{u+t}{2}})^{-2} + (\mathcal{V}^{\frac{u+t}{2}}_u)^{-\frac{3}{2}}(\mathcal{V}^T_t)^{-\frac{1}{2}} + (\mathcal{V}^{\frac{u+t}{2}}_u)^{-1}(\mathcal{V}^T_{\frac{s+t}{2}})^{-1}\\
&+(\mathcal{V}^{\frac{u+s}{2}}_u)^{-\frac{1}{2}}(\mathcal{V}^T_{\frac{s+t}{2}})^{-\frac{3}{2}}] + (\mathcal{V}^T_{\frac{s+t}{2}})^{-\frac{3}{2}} + (\mathcal{V}^{\frac{u+t}{2}}_u)^{-1}(\mathcal{V}^T_t)^{-\frac{1}{2}} + (\mathcal{V}^{\frac{u+s}{2}}_u)^{-\frac{1}{2}}(\mathcal{V}^T_{\frac{s+t}{2}})^{-1} + (\mathcal{V}^T_0)^{\frac{1}{2}}(\mathcal{V}^T_{\frac{u+t}{2}})^{-2}\\
&+(\mathcal{V}^s_{\frac{u+s}{2}})^{-\frac{3}{4}}(\mathcal{V}^T_t)^{-\frac{3}{4}} + (\mathcal{V}^T_0)^{\frac{1}{20}}(\mathcal{V}^{\frac{u+s}{2}}_u)^{-\frac{3}{4}}(\mathcal{V}^T_t)^{-\frac{4}{5}}\},
\end{aligned}$$

where the generic constant is independent of x. We deduce that $\forall 0 \le u < s < t < T$, $\forall p \ge 1$:

$$\|\mathcal{L}^x_2 v^{P,\rho_{s,t}}_{\underline{o},s}(u, X_u)\mathbb{1}_{\tau_b > u}\|_p \tag{7.48}$$

$$
\begin{aligned}
\leq_c C_\sigma^{-3}\{ & \frac{T}{(T-u)^{\frac{5}{2}}} + \frac{\sqrt{T}}{(T-\frac{u+t}{2})^2} + \frac{\sqrt{T}}{(t-u)^{\frac{3}{2}}\sqrt{T-t}} + \frac{\sqrt{T}}{(t-u)(T-\frac{s+t}{2})} + \frac{\sqrt{T}}{\sqrt{s-u}(T-\frac{s+t}{2})^{\frac{3}{2}}} \\
& + \frac{1}{(T-\frac{s+t}{2})^{\frac{3}{2}}} + \frac{1}{(t-u)\sqrt{T-t}} + \frac{1}{\sqrt{s-u}(T-\frac{s+t}{2})} + \frac{1}{(s-u)^{\frac{3}{4}}(T-t)^{\frac{3}{4}}} + \frac{T^{\frac{1}{20}}}{(s-u)^{\frac{3}{4}}(T-t)^{\frac{4}{5}}}\}.
\end{aligned}
$$

7.3.5 Proof of the error estimate in Theorem 7.2.2.1

Assuming $(\widetilde{\mathcal{H}}_{x_0}^{\sigma})$ and that $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$, we obtain in view of (7.18)-(7.19)-(7.21)-(7.23) that the error term in Theorem 7.2.2.1 is equal to:

$$
\begin{aligned}
\mathrm{Error}_{3,\mathrm{h}} = & \frac{1}{2}\mathbb{E}[\int_0^{T\wedge\tau_b} (X_t - x_0)^3 \{\int_0^1 \frac{(1-\alpha)^2}{2}\Sigma_t^{(3)}((1-\alpha)x_0 + \alpha X_t)\mathrm{d}\alpha\}\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t, X_t)\mathrm{d}t] \\
& + \frac{1}{8}\int_0^T \Sigma_t^{(2)}\mathbb{E}[\int_0^{t\wedge\tau_b} (\Sigma_s(X_s) - \Sigma_s)\mathcal{L}_2^x v_{\underline{o},t}^{P,\psi_t}(s, X_s)\mathrm{d}s]\mathrm{d}t \\
& + \frac{1}{4}\int_0^T \Sigma_t^{(1)}\mathbb{E}[\int_0^{t\wedge\tau_b} \{\int_0^1 \Sigma_s^{(2)}((1-\alpha)x_0 + \alpha X_s)(1-\alpha)\mathrm{d}\alpha\}(X_s - x_0)^2 \mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s, X_s)\mathrm{d}s]\mathrm{d}t \\
& + \frac{1}{8}\int_0^T \Sigma_t^{(1)}(\int_0^t \Sigma_s^{(1)}\mathbb{E}[\int_0^{s\wedge\tau_b} (\Sigma_u(X_u) - \Sigma_u)\mathcal{L}_2^x v_{\underline{o},s}^{P,\rho_{s,t}}(u, X_u)\mathrm{d}u]\mathrm{d}s)\mathrm{d}t.
\end{aligned} \tag{7.49}
$$

We easily bound the first term of the r.h.s. of (7.49) using Lemmas 7.3.1.3 and 7.3.1.1 and $(\widetilde{\mathcal{H}}_{x_0}^{\sigma})$:

$$
\begin{aligned}
& \left|\mathbb{E}[\int_0^{T\wedge\tau_b} (X_t - x_0)^3 \{\int_0^1 \frac{(1-\alpha)^2}{2}\Sigma_t^{(3)}((1-\alpha)x_0 + \alpha X_t)\mathrm{d}\alpha\}\mathcal{L}_2^x v_{\underline{o},T}^{P,h}(t, X_t)\mathrm{d}t]\right| \\
\leq_c & \mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)|\sigma|_\infty^3 T^{\frac{3}{2}} C_\sigma^{-1} \int_0^T \frac{\mathrm{d}t}{\sqrt{T-t}} \leq_c \mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)|\sigma|_\infty^2 T^2.
\end{aligned}
$$

For the second term of (7.49), we utilize the Lemma 7.3.4.2 to obtain:

$$
\begin{aligned}
& \left|\int_0^T \Sigma_t^{(2)}\mathbb{E}[\int_0^{t\wedge\tau_b} (\Sigma_s(X_s) - \Sigma_s)\mathcal{L}_2^x v_{\underline{o},t}^{P,\psi_t}(s, X_s)\mathrm{d}s]\mathrm{d}t\right| \\
\leq_c & [\mathcal{M}_1(\sigma)]^2\mathcal{M}_0(\sigma)|\sigma|_\infty^2 \sqrt{T} C_\sigma^{-1} \\
& \times \int_0^T (\int_0^t \{\frac{T}{(T-s)^{\frac{3}{2}}} + \frac{\sqrt{T}}{T-s} + \frac{1}{\sqrt{T-t}} + \frac{\sqrt{T}}{\sqrt{t-s}\sqrt{T-t}} + \frac{\sqrt{T}}{T-\frac{(s+t)}{2}}\}\mathrm{d}s)\mathrm{d}t \\
\leq_c & [\mathcal{M}_1(\sigma)]^2\mathcal{M}_0(\sigma)|\sigma|_\infty T^2.
\end{aligned}
$$

We now pass to the third term of (7.49) and to treat it, we use the Lemma 7.3.2.3 to get:

$$
\begin{aligned}
& \left|\int_0^T \Sigma_t^{(1)}\mathbb{E}[\int_0^{t\wedge\tau_b} \{\int_0^1 \Sigma_s^{(2)}((1-\alpha)x_0 + \alpha X_s)(1-\alpha)\mathrm{d}\alpha\}(X_s - x_0)^2 \mathcal{L}_2^x v_{\underline{o},t}^{P,\phi_t}(s, X_s)\mathrm{d}s]\mathrm{d}t\right| \\
\leq_c & [\mathcal{M}_1(\sigma)]^2\mathcal{M}_0(\sigma)|\sigma|_\infty^3 T C_\sigma^{-2} \int_0^T (\int_0^t \{\frac{\sqrt{T}}{(T-s)^{\frac{3}{2}}} + \frac{1}{T-s} + \frac{1}{\sqrt{(t-s)}\sqrt{T-t}} + \frac{1}{T-\frac{(s+t)}{2}}\}\mathrm{d}s)\mathrm{d}t \\
\leq_c & [\mathcal{M}_1(\sigma)]^2\mathcal{M}_0(\sigma)|\sigma|_\infty T^2.
\end{aligned}
$$

We finish with the last term of (7.49). An application of Lemma 7.3.4.5 yields:

$$
\left|\int_0^T \Sigma_t^{(1)}(\int_0^t \Sigma_s^{(1)}\mathbb{E}[\int_0^{s\wedge\tau_b} (\Sigma_u(X_u) - \Sigma_u)\mathcal{L}_2^x v_{\underline{o},s}^{P,\rho_{s,t}}(u, X_u)\mathrm{d}u]\mathrm{d}s)\mathrm{d}t\right|
$$

$$
\begin{aligned}
\leq_c & [\mathcal{M}_1(\sigma)]^3 |\sigma|_\infty^4 \sqrt{T} C_\sigma^{-3} \int_0^T \Big(\int_0^t \Big(\int_0^s \Big\{ \frac{T}{(T-u)^{\frac{5}{2}}} + \frac{\sqrt{T}}{(T-\frac{u+t}{2})^2} + \frac{\sqrt{T}}{(t-u)^{\frac{3}{2}} \sqrt{T-t}} \\
& + \frac{\sqrt{T}}{(t-u)(T-\frac{s+t}{2})} + \frac{\sqrt{T}}{\sqrt{s-u}(T-\frac{s+t}{2})^{\frac{3}{2}}} + \frac{1}{(T-\frac{s+t}{2})^{\frac{3}{2}}} + \frac{1}{(t-u)\sqrt{T-t}} \\
& + \frac{1}{\sqrt{s-u}(T-\frac{s+t}{2})} + \frac{1}{(s-u)^{\frac{3}{4}}(T-t)^{\frac{3}{4}}} + \frac{T^{\frac{1}{20}}}{(s-u)^{\frac{3}{4}}(T-t)^{\frac{4}{5}}} \Big\} \mathrm{d}u \Big) \mathrm{d}s \Big) \mathrm{d}t \\
\leq_c & [\mathcal{M}_1(\sigma)]^3 |\sigma|_\infty T^2.
\end{aligned}
$$

That achieves the proof.

7.4 Applications to the pricing of down and in barrier options

We denote by $v_{\underline{i},T}^{P,h}(t,x) = \mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P \leq T} | X_t^P = x]$,$\forall (t,x) \in [0,T] \times [b,+\infty[$, the proxy pricing kernel for the down and in option and the obvious relation:

$$
v_{\underline{o},T}^{P,h}(t,x) + v_{\underline{i},T}^{P,h}(t,x) = v_T^{P,h}(t,x), \tag{7.50}
$$

allows us to deduce results for the down and in case from the down and out case, applying the results of [Benhamou 2010a] for the plain-vanilla part. For $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$,
$v_{\underline{i},T}^{P,h}(t,x) = \int_{\mathbb{R}} h(y) e^{-\frac{2(x-b)(y-b)}{v_t^T}} \mathcal{D}^P(t,T,y-x)\mathrm{d}y$ which is the second term of the r.h.s of (7.9). We easily deduce from the relation (7.50) the next Lemma:

Lemma 7.4.0.1. *Assume* $(\widetilde{\mathcal{H}}_{x_0}^\sigma)$ *and that* $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$. *For any* $t \in [0,T[$, *we have:*

$$
\mathcal{L}_n^x v_{\underline{i},T}^{P,h}(t,b) = \mathcal{L}_n^x v_T^{P,h}(t,b), \quad \forall n \in \{2,4,6\}, \tag{7.51}
$$

$$
\mathcal{L}_n^x v_{\underline{i},T}^{P,h}(t,b) = -\mathcal{L}_n^x v_T^{P,h}(t,b) = -\frac{1}{2}\mathcal{L}_n^x v_{\underline{o},T}^{P,h}(t,b), \quad \forall n \in \{1,3,5\}, \tag{7.52}
$$

We now announce directly the main results:

Theorem 7.4.0.1. ***(2nd order approximation price formula for down and in regular barrier options)****.* *Assume* $(\widetilde{\mathcal{H}}_{x_0}^\sigma)$. *Then for any* $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$, *we have for any* $x_0 \geq b$:

$$
\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b \leq T}] = \mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P \leq T}] + Cor_{1,\underline{i}} + Cor_{2,\underline{i}} + O(|\sigma|_\infty \mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)T^{\frac{3}{2}}), \tag{7.53}
$$

where:

$$
Cor_{1,\underline{i}} = \frac{1}{2}\vec{\omega}(\Sigma^{(1)},\Sigma)_0^T \mathcal{L}_3^x v_{\underline{i},T}^{P,h}(0,x_0), \qquad Cor_{2,\underline{i}} = -\mathbb{E}[\mathbb{1}_{\tau_b^P \leq T}\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^P}^T \mathcal{L}_3^x v_{\underline{i},T}^{P,h}(\tau_b^P, b)].
$$

Proof. We write $\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b \leq T}] = \mathbb{E}[h(X_T)] - \mathbb{E}[h(X_T)\mathbb{1}_{\tau_b > T}]$ and we apply [Benhamou 2010a, Theorem 2.1] for the plain vanilla part and the Theorem 7.2.2.1 for the down and out option to readily obtain:

$$
\begin{aligned}
\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b \leq T}] = & \mathbb{E}[h(X_T^P)] - \mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P > T}] + \frac{1}{2}\vec{\omega}(\Sigma^{(1)},\Sigma)_0^T \mathcal{L}_3^x \underbrace{(v_T^{P,h}(0,x) - v_{\underline{o},T}^{P,h}(0,x))}_{v_{\underline{i},T}^{P,h}(0,x)}|_{x=x_0} \\
& + \frac{1}{2}\mathbb{E}[\mathbb{1}_{\tau_b^P \leq T}\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^P}^T \mathcal{L}_3^x v_{\underline{o},T}^{P,h}(\tau_b^P, b)] + O(|\sigma|_\infty \mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)T^{\frac{3}{2}}).
\end{aligned}
$$

Then by Lemma 7.4.0.1 equation (7.52), we have $\mathcal{L}_3^x v_{\underline{o},T}^{P,h}(\tau_b^P, b) = 2\mathcal{L}_3^x v_T^{P,h}(\tau_b^P, b) = -2\mathcal{L}_3^x v_{\underline{i},T}^{P,h}(\tau_b^P, b)$. □

Remark 7.4.0.1. *Observe that if $x_0 = b$, we take face to a plain vanilla option and owing to Lemma 7.4.0.1 we are coherent with [Benhamou 2010a, Theorem 2.1], whereas if $b \to -\infty$, all the term of the expansion vanish.*

Theorem 7.4.0.2. *(**3rd order approximation price formula for down and in regular barrier options**). Assuming $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$, we have for any $h \in \underline{\mathrm{Lip}}(\mathbb{R}, b)$ and any $x_0 \geq b$:*

$$\mathbb{E}[h(X_T)\mathbb{1}_{\tau_b \leq T}] = \mathbb{E}[h(X_T^P)\mathbb{1}_{\tau_b^P \leq T}] + \sum_{n=1}^{5} Cor_{n,\underline{i}} + O(|\sigma|_\infty \mathcal{M}_1(\sigma)[\mathcal{M}_0(\sigma)]^2 T^2), \tag{7.54}$$

where the corrective terms $Cor_{n,\underline{i}}$ are obtained by replacing $v^{P,h}_{\underline{o},T}(0, x_0)$ by $v^{P,h}_{\underline{i},T}(0, x_0)$ and $v^{P,h}_{\underline{o},T}(r, b) = 2v^{P,h}_T(r, b)$ by $2v^{P,h}_{\underline{i},T}(r, b) = -2v^{P,h}_T(r, b)$ $\forall r \in [0, T]$ in the various corrective terms $Cor_{k,\underline{o}}$ defined in Theorem 7.2.3.1.

Remark 7.4.0.2. *If $x_0 = b$, $Cor_{4,\underline{i}} = Cor_{5,\underline{i}} = 0$ and owing to Lemma 7.4.0.1 equation (7.51), $\mathcal{L}^x_n v^{P,h}_{\underline{i},T}(0, x_0) = \mathcal{L}^x_n v^{P,h}_T(0, x_0)$, $\forall n \in \{2, 4, 6\}$. We are consistent with [Benhamou 2010a, Theorem 2.2]. If $b \to -\infty$, all the corrective terms vanish.*

7.5 Applications to regular down barrier Call options

In this section, we apply our various price approximations to the particular case of regular down Call options. The payoff function is now equal to $h(x) = (e^x - e^k)_+$ with $\min(x_0, k) \geq b$. In order to obtain more accurate approximations (see Chapter 2), we derive new expansions with the local volatility frozen at mid-point $x_{avg} = \frac{x_0+k}{2}$. Then we show that if the local volatility function is time-homogeneous, our prices expansions reduce to totally closed-form formulas with a numerical cost close to zero.

7.5.1 Notations

▷ **Barrier Call options.** We denote by $\mathrm{Call}(S_0, T, K)$ the price at time 0 of a Call option with spot S_0, maturity T and strike K, written on the asset $S = e^X$ that is $\mathrm{Call}(S_0, T, K) = \mathbb{E}(e^{X_T} - K)_+$. We use the notation $\mathrm{DoCall}(S_0, T, K, B)$ (respectively $\mathrm{DiCall}(S_0, T, K, B)$) for the price at time 0 of a regular down and out (respectively in) barrier Call option with barrier level $B = e^b \leq \min(K, S_0)$ that is $\mathrm{DoCall}(S_0, T, K, B) = \mathbb{E}[(e^{X_T} - K)_+\mathbb{1}_{\tau_b > T}]$ (respectively $\mathrm{DiCall}(S_0, T, K, B) = \mathbb{E}[(e^{X_T} - K)_+\mathbb{1}_{\tau_b \leq T}]$).
As usual, ATM (At The Money) Call refers to $x_0 \approx k$, ITM (In The Money) to $x_0 \gg k$, OTM (Out The Money) to $x_0 \ll k$.

▷ **Barrier Black-Scholes Call price function.** For the sake of completeness, we give the Black-Scholes Call price function depending on log-spot x, total variance $y > 0$ and log-strike k:

$$\mathrm{Call}^{\mathrm{BS}}(x, y, k) = e^x \mathcal{N}(d_1(x, y, k)) - e^k \mathcal{N}(d_2(x, y, k)) \tag{7.55}$$

where:

$$d_1(x, y, k) = \frac{x - k}{\sqrt{y}} + \frac{1}{2}\sqrt{y}, \qquad d_2(x, y, k) = d_1(x, y, k) - \sqrt{y}.$$

We recall that we have the simple following relation:

$$\partial_y \mathrm{Call}^{\mathrm{BS}}(x, y, k) = \frac{1}{2}\mathcal{L}^x_2 \mathrm{Call}^{\mathrm{BS}}(x, y, k) = \frac{1}{2}\mathcal{L}^k_2 \mathrm{Call}^{\mathrm{BS}}(x, y, k). \tag{7.56}$$

We use the notation $\mathrm{DoCall}^{\mathrm{BS}}(x,y,k,b)$ (respectively $\mathrm{DiCall}^{\mathrm{BS}}(x,y,k,b)$) for the price of a regular down and out (respectively in) barrier Call option in the Black-Scholes model with log-barrier equal to b. For convenience, we recall that we have the following formulas:

$$\mathrm{DiCall}^{\mathrm{BS}}(x,y,k,b)=\mathrm{Call}^{\mathrm{BS}}(b,y,x+k-b),\quad \mathrm{DoCall}^{\mathrm{BS}}(x,y,k,b)=\mathrm{Call}^{\mathrm{BS}}(x,y,k)-\mathrm{DiCall}^{\mathrm{BS}}(x,y,k,b). \tag{7.57}$$

Notice that $\partial_x\mathrm{DiCall}^{\mathrm{BS}}(x,y,k,b)=\partial_k\mathrm{Call}^{\mathrm{BS}}(b,y,k)|_{k=x+k-b}$ and the facts that $\mathcal{L}_n^x\mathrm{Call}^{\mathrm{BS}}(x,y,k)=\mathcal{L}_n^k\mathrm{Call}^{\mathrm{BS}}(x,y,k)$ for $n\in\{2,4,6\}$ whereas $\mathcal{L}_n^x\mathrm{Call}^{\mathrm{BS}}(x,y,k)=-\mathcal{L}_n^k\mathrm{Call}^{\mathrm{BS}}(x,y,k)$ for $n\in\{3,5\}$ (see Chapter 2 Proposition 2.6.1.3) allow to retrieve the results of Lemmas 7.2.1.1 and 7.4.0.1.

The expansions of Theorems 7.2.2.1-7.2.3.1 and Corollary 7.4.0.1 remain valid for the regular Call payoff replacing in the various corrective coefficients Cor the sensitivities by the corresponding derivatives w.r.t. x of the functions $\mathrm{DoCall}^{\mathrm{BS}}(x,y,k,b)$ or $\mathrm{DiCall}^{\mathrm{BS}}(x,y,k,b)$.
In the following, $x_0=\log(S_0)$ will represent the log-spot, $k=\log(K)$ the log-strike, $b=\log(B)\le k$ the log-barrier and $x_{avg}=(x_0+k)/2=\log(\sqrt{S_0K})$ the mid-point between the log-spot and the log-strike.

▷ **Total volatility at x_{avg}, $\tau_b^{P,x_{avg}}$ and $(\widetilde{\mathcal{H}}^\sigma_{x_{avg}})$.** When freezing the local volatility and local variance functions σ and Σ in x_{avg}, we denote by $\mathcal{V}_t^T(x_{avg})=\int_t^T\Sigma_s(x_{avg})\mathrm{d}s$ the total variance at point x_{avg} on the period $[t,T]$. We extend this notation for the integral operator $\vec{\omega}$ acting on Σ and its derivatives computed at x_{avg} and we introduce the notation $Cor_{k,\underline{\varrho}}(x_{avg})$ to denote the same corrective terms introduced in Theorem 7.2.3.1 but with the local volatility frozen at x_{avg}. For instance we have $\vec{\omega}(\Sigma^{(1)},\Sigma)_0^T(x_{avg})=\vec{\omega}(\Sigma^{(1)}(x_{avg}),\Sigma(x_{avg}))_0^T$ and $Cor_{1,\underline{\varrho}}(x_{avg})=\frac{1}{2}\vec{\omega}(\Sigma^{(1)},\Sigma)_0^T(x_{avg})\mathcal{L}_3^x\mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b)$.
Similarly we introduce the new proxy process $(X_t^{P,x_{avg}})_{t\in[0,T]}$ which is a Gaussian process defined like in (7.3) but with the local volatility frozen at x_{avg}. For this process we introduce the first hitting time of the level b: $\tau_b^{P,x_{avg}}=\inf\{t\ge0:X_t^{P,x_{avg}}=b\}$. We denote by $\mathcal{D}^{P,x_{avg}}$ the density of $X_t^{P,x_{avg}}$ and by $\mathcal{J}^{P,x_{avg}}$ the density associated to the hitting times of $(X_t^{P,x_{avg}})_{t\in[0,T]}$.
Finally we define the assumption $(\widetilde{\mathcal{H}}^\sigma_{x_{avg}})$ similarly to $(\widetilde{\mathcal{H}}^\sigma_{x_0})$ replacing x_0 by x_{avg} in $(\widetilde{\mathcal{H}}^\sigma_{x_0})$.

7.5.2 Regular down barrier Call option approximations with the local volatility at mid-point.

To obtain new expansions with the local volatility frozen at x_{avg}, we perform an expansion of the local volatility in the approximation formulas given in Theorems 7.2.2.1-7.2.3.1 and Corollary 7.4.0.1. The results are summarised in the following Proposition proven in Appendix 7.6.3.

Proposition 7.5.2.1. *Assume $(\widetilde{\mathcal{H}}^\sigma_{x_0})$-$(\widetilde{\mathcal{H}}^\sigma_{x_{avg}})$. We have for any $b\le\min(x_0,k)$:*

$$\begin{aligned}\mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T,k,b)=&\mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b)+Cor_{6,\underline{\varrho}}(x_{avg})+O(|\sigma|_\infty\mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)T^{\frac{3}{2}})\\=&\mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b)+(Cor_{6,\underline{\varrho}}+Cor_{7,\underline{\varrho}})(x_{avg})+O(|\sigma|_\infty\mathcal{M}_1(\sigma)[\mathcal{M}_0(\sigma)]^2T^2),\end{aligned}$$

$$\begin{aligned}Cor_{1,\underline{\varrho}}=&Cor_{1,\underline{\varrho}}(x_{avg})+O(|\sigma|_\infty\mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)T^{\frac{3}{2}})=(Cor_{1,\underline{\varrho}}+Cor_{8,\underline{\varrho}})(x_{avg})+O(|\sigma|_\infty\mathcal{M}_1(\sigma)[\mathcal{M}_0(\sigma)]^2T^2),\\Cor_{2,\underline{\varrho}}=&Cor_{2,\underline{\varrho}}(x_{avg})+O(|\sigma|_\infty\mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma)T^{\frac{3}{2}})=(Cor_{2,\underline{\varrho}}+Cor_{9,\underline{\varrho}})(x_{avg})+O(|\sigma|_\infty\mathcal{M}_1(\sigma)[\mathcal{M}_0(\sigma)]^2T^2),\end{aligned}$$

$$\sum_{n=3}^5Cor_{n,\underline{\varrho}}=\sum_{n=3}^5Cor_{n,\underline{\varrho}}(x_{avg})+O(|\sigma|_\infty\mathcal{M}_1(\sigma)[\mathcal{M}_0(\sigma)]^2T^2),$$

where:

$$Cor_{6,\underline{\varrho}}(x_{avg})=\frac{1}{4}(x_0-k)\vec{\omega}(\Sigma^{(1)}(x_{avg}))_0^T\mathcal{L}_2^x\mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b),$$

$$Cor_{7,\underline{o}}(x_{avg}) = \frac{1}{16}(x_0-k)^2\{\vec{\omega}(\Sigma^{(2)}(x_{avg}))_0^T\mathcal{L}_2^x + \vec{\omega}(\Sigma^{(1)},\Sigma^{(1)})_0^T(x_{avg})\mathcal{L}_4^x\}\mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b),$$

$$Cor_{8,\underline{o}}(x_{avg}) = \frac{1}{4}(x_0-k)\{[\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)})_0^T + \vec{\omega}(\Sigma^{(2)},\Sigma)_0^T](x_{avg})\mathcal{L}_3^x$$
$$+ [\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)},\Sigma)_0^T + \frac{1}{2}\vec{\omega}(\Sigma^{(1)},\Sigma,\Sigma^{(1)})_0^T](x_{avg})\mathcal{L}_5^x\}\mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b),$$

$$Cor_{9,\underline{o}}(x_{avg}) = -\frac{1}{4}(x_0-k)\mathbb{E}[\mathbb{1}_{\tau_b^{P,x_{avg}}<T}\{[\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)})^T_{\tau_b^{P,x_{avg}}} + \vec{\omega}(\Sigma^{(2)},\Sigma)^T_{\tau_b^{P,x_{avg}}}$$
$$+ \vec{\omega}(\Sigma^{(1)})_0^{\tau_b^{P,x_{avg}}}\vec{\omega}(\Sigma^{(1)})^T_{\tau_b^{P,x_{avg}}}](x_{avg})\mathcal{L}_3^x + [\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)},\Sigma)^T_{\tau_b^{P,x_{avg}}} + \frac{1}{2}\vec{\omega}(\Sigma^{(1)},\Sigma,\Sigma^{(1)})^T_{\tau_b^{P,x_{avg}}}$$
$$+ \frac{1}{2}\vec{\omega}(\Sigma^{(1)})_0^{\tau_b^{P,x_{avg}}}\vec{\omega}(\Sigma^{(1)},\Sigma)^T_{\tau_b^{P,x_{avg}}}](x_{avg})\mathcal{L}_5^x\}\mathrm{DoCall}^{\mathrm{BS}}(b,\mathcal{V}^T_{\tau_b^{P,x_{avg}}}(x_{avg}),k,b)].$$

We easily deduce from Proposition 7.5.2.1 the next Theorem:

Theorem 7.5.2.1. ***(2nd and 3rd order approximations for regular down barrier Call options with local volatility at x_{avg}).***
Assume $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$-$(\widetilde{\mathcal{H}}^{\sigma}_{x_{avg}})$. Then for any $b \le \min(x_0,k)$, we have:

$$\begin{aligned}\mathrm{DoCall}(S_0,T,K,B) =& \mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b) + \sum_{n\in\{1,2,6\}} Cor_{n,\underline{o}}(x_{avg}) + O(|\sigma|_\infty\mathcal{M}_0(\sigma)\mathcal{M}_1(\sigma)T^{\frac{3}{2}})\\ =& \mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b) + \sum_{n=1}^{9} Cor_{n,\underline{o}}(x_{avg}) + O(|\sigma|_\infty[\mathcal{M}_0(\sigma)]^2\mathcal{M}_1(\sigma)T^2),\end{aligned}$$

where the corrective coefficients $Cor_{n,\underline{o}}$ are defined in Theorems 7.2.2.1-7.2.3.1 and in Proposition 7.5.2.1. Under the same hypotheses, one has:

$$\begin{aligned}\mathrm{DiCall}(S_0,T,K,B) =& \mathrm{DiCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b) + \sum_{n\in\{1,2,6\}} Cor_{n,\underline{i}}(x_{avg}) + O(|\sigma|_\infty\mathcal{M}_0(\sigma)\mathcal{M}_1(\sigma)T^{\frac{3}{2}})\\ =& \mathrm{DiCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b) + \sum_{n=1}^{9} Cor_{n,\underline{i}}(x_{avg}) + O(|\sigma|_\infty[\mathcal{M}_0(\sigma)]^2\mathcal{M}_1(\sigma)T^2),\end{aligned}$$

where we define $Cor_{n,\underline{i}}(x_{avg})$, $n \in \{1,..,9\}$ by replacing in $Cor_{n,\underline{o}}(x_{avg})$ $\mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b)$ by $\mathrm{DiCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T(x_{avg}),k,b)$, $n \in \{1,3,6,7,8\}$ and by setting them equal to $-Cor_{n,\underline{o}}(x_{avg})$, $n \in \{2,4,5,9\}$.

7.5.3 Reductions in the time-homogeneous framework

The various formulas of Theorems 7.2.2.1-7.2.3.1-7.5.2.1 and Corollary 7.4.0.1 are explicit up to a numerical integration of the terms expressed as an expectation involving the first hitting time of the barrier. Remarkably, these formulas are available in closed-forms through an expression containing only Gaussian functions if the local volatility is time-independent (extension to local volatility functions with separable time and space variables is straightforward). To see this, we leverage convolution properties of the Gaussian hitting times density postponed to Proposition 7.6.1.4. That leads to the following Proposition proven in Appendix 7.6.3.

Proposition 7.5.3.1. *We suppose that the local volatility function is time-homogeneous and we assume $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$-$(\widetilde{\mathcal{H}}^{\sigma}_{x_{avg}})$. Then for any $b \le \min(x_0,k)$, we have for any $z \in \{x_0,x_{avg}\}$:*

$Cor_{2,\underline{o}}(z)$

$$=\frac{1}{2}e^{b}(k-b)\frac{\Sigma^{(1)}(z)}{\Sigma(z)}\{(x_0-b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}-\mathcal{V}_0^T(z)\mathcal{D}^{P,z}(0,T,2b-x_0-k)\},$$

$Cor_{4,\underline{o}}(z)$

$$=\frac{1}{2}e^{b}(x_0-b)(k-b)\frac{\Sigma^{(2)}(z)}{\Sigma(z)}\{\mathcal{V}_0^T(z)\mathcal{D}^{P,z}(0,T,2b-x_0-k)-(x_0-b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}\}$$

$$+\frac{1}{8}e^{b}(x_0-b)(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\{(x_0-b)[3+\frac{5}{8}\mathcal{V}_0^T(z)+\frac{1}{8}(x_0-k)(x_0+k-2b)]\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}$$

$$+[2(k-b)^2+(k-b)(x_0-b)-\mathcal{V}_0^T(z)(2+\frac{1}{8}(x_0-b)(x_0-k)+\frac{1}{2}\mathcal{V}_0^T(z))]\mathcal{D}^{P,z}(0,T,2b-x_0-k)\},$$

$Cor_{5,\underline{o}}(z)$

$$=\frac{1}{8}e^{b}(x_0-b)(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\{(x_0-b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}-\mathcal{V}_0^T(z)\mathcal{D}^{P,z}(0,T,2b-x_0-k)$$

$$+\frac{1}{8}(\mathcal{V}_0^T(z))^{\frac{3}{2}}e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}$$

$$\times(-\frac{1}{6}H_3(x)\mathcal{N}(x)-\frac{1}{6}H_2(x)\mathcal{N}'(x)+\frac{1}{2}\mathcal{N}'(x)-\frac{(x_0-k)}{2\sqrt{\mathcal{V}_0^T(z)}}[(H_2(x)+2)\mathcal{N}(x)+x\mathcal{N}'(x)])\Big|_{x=\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}}}\},$$

$Cor_{9,\underline{o}}(z)$

$$=\frac{1}{4}e^{b}(x_0-k)(k-b)\frac{\Sigma^{(2)}(z)}{\Sigma(z)}\{(x_0-b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}-\mathcal{V}_0^T(z)\mathcal{D}^{P,z}(0,T,2b-x_0-k)\}$$

$$+\frac{1}{4}e^{b}(x_0-k)(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\{[\frac{1}{2}\mathcal{V}_0^T(z)+\frac{1}{2}(2b-x_0-k)(k-b)+\frac{1}{8}(\mathcal{V}_0^T(z))^2]\mathcal{D}^{P,z}(0,T,2b-x_0-k)$$

$$-(x_0-b)[1+\frac{1}{8}\mathcal{V}_0^T(z)]\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}\}.$$

Remark 7.5.3.1. *We remarkably obtain very simple second order formulas with the local volatility at mid-point. Combining the explicit forms of $Cor_{1,\underline{o}}$, $Cor_{2,\underline{o}}$ and $Cor_{6,\underline{o}}$ (see Propositions 7.5.2.1-7.5.3.1), we easily get:*

$$\mathrm{DoCall}(S_0,T,K,B)=\mathrm{DoCall}^{\mathrm{BS}}(x_0,\Sigma(x_{avg})T,k,b)+Cor_{10,\underline{o}}+O(|\sigma|_\infty\mathcal{M}_0(\sigma)\mathcal{M}_1(\sigma)T^{\frac{3}{2}}),\tag{7.58}$$

$$\mathrm{DiCall}(S_0,T,K,B)=\mathrm{DiCall}^{\mathrm{BS}}(x_0,\Sigma(x_{avg})T,k,b)-Cor_{10,\underline{o}}+O(|\sigma|_\infty\mathcal{M}_0(\sigma)\mathcal{M}_1(\sigma)T^{\frac{3}{2}}),\tag{7.59}$$

$$Cor_{10,\underline{o}}=\frac{1}{2}\frac{\Sigma^{(1)}(x_{avg})}{\Sigma(x_{avg})}(k-b)(x_0-b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(x_{avg})}})e^{x_{avg}-\frac{\mathcal{V}_0^T(x_{avg})}{8}}.$$

The expansions formulas (7.59) and (7.58) reduce to a suitable regular down barrier Call option price with the local volatility frozen at x_{avg} plus or minus a simple and explicit term $Cor_{10,\underline{o}}$ symmetric w.r.t. the variables x_0 and k. Notice that if $b=x_0$ or $b=k$, this additional term vanishes. In particular if $b=k$, owing to (7.57), the expansion formula (7.59) is simply a Black-Scholes Put price with log-spot x_0, log-strike b and total variance equal to $\Sigma(x_{avg})T$, whereas the approximation formula (7.58) is the difference of a Black-Scholes Call price and a Black-Scholes Put price with the same features, i.e. $e^{x_0}-e^{b}$. We let the reader verify that the third order formulas also give these very simple approximations if $b=k$.

7.5.4 Numerical experiments

▷**Model and benchmark.** Here we give numerical examples of the accuracy of our third order approximation with local volatility at x_{avg} (see Theorem 7.5.2.1), denoted by `App(3,`x_{avg}`)`, for regular down out barrier Call options in the time-homogeneous framework because of the high simplicity of our formulas owing to Proposition 7.5.3.1. As a benchmark model, we consider a time-independent CEV model:

$$dS_t = \nu S_0^{1-\beta} S_t^{\beta} dW_t,\ S_0 = e^{x_0}, \tag{7.60}$$

with a spot spot value $S_0 = 100$, a level of volatility $\nu = 25\%$ and a skew parameter $\beta = 0.5$. This model is applied directly to the asset price and we apply our results by considering a fictive log-asset with local volatility function $\sigma(x) = \nu e^{(\beta-1)(x-x_0)}$. Although the local volatility function as well its derivatives are not bounded, the following tests nevertheless prove the excellent accuracy of our formula. An advantage of this model widely used by the practitioners in the industry of finance is that Call options prices (see [Schroder 1989]) are available in closed-form as well as barrier Call options prices. Davydov et al. derived in [Davydov 2001] closed-form formulas for the Laplace transforms of barrier options prices under the CEV diffusion. Then they developed in [Davydov 2003] eigenfunctions expansions to invert the Laplace transform. Alternatively, and this is the methodology chosen in this work, the Laplace transform inversion can be performed with the Abate and Whitt algorithm (see [Abate 1995]).

▷**Set of parameters.** We study the accuracy of `App(3,`x_{avg}`)` for various maturities, strikes and levels of barrier. We report in Table 7.1 the maturities and the strikes evolving approximately as $S_0 \exp(c\nu\sqrt{T})$ where c takes the value of various quantiles of the standard Gaussian law (1%-5%-10%-20%-30%-40%-50%-60%-70%-80%-90%-95%-99%) to cover both far ITM and far OTM options. We report in Table 7.2 the maturities and the corresponding levels of barrier evolving similarly to the strikes (but with quantiles lower than 50% because the level of barrier has to be smaller than the spot 100). For the sake of completeness we also consider the case $B = 0$ (pure vanilla case) associated to the quantile 0. Then we report in Tables 7.3 and 7.4 the values of the down out barrier Call option prices for the whole set of maturities, strikes and barriers (keeping only the strikes greater than the barriers to consider regular options) obtained with the closed-form formula and in Tables 7.5 and 7.6, the errors on prices using `App(3,`x_{avg}`)`.

▷**Analysis of the results.** The Tables 7.5 and 7.6 show that `App(3,`x_{avg}`)` is extremely accurate for all the maturities, strikes and barriers. The maximum error in absolute value for the whole set of parameters is about 10^{-2} obtained for the largest maturities. For small maturities (see Table 7.5), the errors are generally of magnitude $10^{-4} - 10^{-5}$ (with smallest errors of magnitude 10^{-7} !) and for large maturities (see Table 7.6), of magnitude $10^{-3} - 10^{-4}$ (with smallest errors of magnitude 10^{-6} !). The errors are increasing w.r.t. T and at fixed maturity, are globally of the same magnitude according to K and B. In particular we notice that the barrier prices approximations are as accurate than those of the vanilla prices. This reinforces our belief that the accuracy of the expansion strongly depends on the regularity of the involved functionals which are comparable in the case of regular barrier options and plain vanilla options. We nevertheless remark that the errors are slightly smaller for very OTM options (certainly because the price is close to 0) and are very small for ITM options with $K = B$ (at most 10^{-5}). As noticed in Remark 7.5.3.1, $K = B$ is a particular situation with a price approximation given by $S_0 - K$ what is very close to the real price as seen in Tables 7.3 and 7.4.

Table 7.1: Set of maturities and strikes for the numerical experiments

$T\backslash K$	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
3M	70	75	80	85	90	95	100	105	110	115	125	130	135
6M	65	75	80	85	90	95	100	105	110	120	125	135	150
1Y	55	65	75	80	90	95	100	105	115	125	140	150	180
1.5Y	50	60	70	75	85	95	100	110	115	130	150	165	200
2Y	45	55	65	75	85	90	100	110	120	135	155	180	230
3Y	35	50	55	70	80	90	100	110	125	145	175	205	270
5Y	25	40	50	60	75	85	100	115	135	160	205	250	260
10Y	15	25	35	50	65	80	100	120	150	195	275	365	630

Table 7.2: Set of maturities and barriers for the numerical experiments

$T\backslash B$	0%	1%	5%	10%	20%	30%	40%	45%	49%
3M	0	70	75	80	85	90	95	97.5	99.5
6M	0	65	75	80	85	90	95	97.5	99.5
1Y	0	55	65	75	80	90	95	97.5	99.5
1.5Y	0	50	60	70	75	85	95	97.5	99
2Y	0	45	55	65	75	85	90	97.5	99
3Y	0	35	50	55	70	80	90	95	99
5Y	0	25	40	50	60	75	85	90	98.5
10Y	0	15	25	35	50	65	80	90	98

7.6 Appendix

7.6.1 Properties of the Gaussian density, the Gaussian cumulative function and the Gaussian hitting times density

We give in the following Propositions some simple relations between the partial derivatives of $\mathcal{D}^P$ and $\mathcal{J}^P$, some estimates of their derivatives and integration results for $\mathcal{N}$:

Proposition 7.6.1.1. *Assume $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$. For any $0 \le s < t \le T$, for any $x,y \in \mathbb{R}$ and any $b < x$, we have:*

$$\partial_s \mathcal{D}^P(s,t,x,y) = -\frac{\Sigma_s}{2}\mathcal{L}^x_2\mathcal{D}^P(s,t,x,y), \quad \partial_t \mathcal{D}^P(s,t,x,y) = \frac{\Sigma_t}{2}\mathcal{L}^x_2\mathcal{D}^P(s,t,x,y), \tag{7.61}$$

$$\partial_s \mathcal{J}^P(s,t,x,b) = -\frac{\Sigma_s}{2}\mathcal{L}^x_2\mathcal{J}^P(s,t,x,b), \quad \partial_t\Big(\frac{\mathcal{J}^P(s,t,x,b)}{\Sigma_t}\Big) = \frac{1}{2}\mathcal{L}^x_2\mathcal{J}^P(s,t,x,b), \tag{7.62}$$

$$\mathcal{J}^P(s,t,x,b) = -\Sigma_t\mathcal{L}^x_1\mathcal{D}^P(s,t,x,b), \tag{7.63}$$

$$\begin{aligned}\mathcal{L}^x_1\mathcal{J}^P(s,t,x,b) =& \Sigma_t\Big(\frac{1}{\mathcal{V}^t_s} - \frac{(x-b)^2}{(\mathcal{V}^t_s)^2}\Big)\mathcal{D}^P(s,t,x,b) = \frac{\Sigma_t}{\mathcal{V}^t_s}\mathcal{D}^P(s,t,x,b) - \frac{(x-b)}{\mathcal{V}^t_s}\mathcal{J}^P(s,t,x,b)\\ &= -\Sigma_t\Big(\mathcal{L}^x_2 + \frac{1}{4}\mathcal{I}\Big)\mathcal{D}^P(s,t,x,b) = -\Big(2\partial_t + \frac{1}{4}\Sigma_t\mathcal{I}\Big)\mathcal{D}^P(s,t,x,b).\end{aligned} \tag{7.64}$$

Proposition 7.6.1.2. *Assume $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$. Using the fact that for any polynomial function $\mathcal{P}$, $x \mapsto \mathcal{P}(x)\mathcal{N}'(x)$ is a bounded function, we have for any $0 \le s < t \le T$, for any $(x,y) \in]-\infty,0]\times\mathbb{R}$ and any integers n and m:*

$$|y|^m\left|\partial^n_{y^n}\mathcal{D}^P(s,t,y)\right| \le_c (\mathcal{V}^t_s)^{\frac{m-n-1}{2}}, \qquad |x|^m\left|\partial^n_{x^n}\mathcal{J}^P(s,t,x)\right| \le_c \Sigma_t(\mathcal{V}^t_s)^{\frac{m-n-2}{2}}. \tag{7.65}$$

The results remain valid if we assume $(\widetilde{\mathcal{H}}^{\sigma}_{x_{avg}})$ and consider $\mathcal{D}^{P,x_{avg}}$, $\mathcal{J}^{P,x_{avg}}$, $\Sigma_t(x_{avg})$ and $\mathcal{V}^t_s(x_{avg})$.

Table 7.3: Down out barrier Call options prices in the CEV model ($\beta = 0.5, \nu = 0.25$) obtained with the closed-form formula for the maturities $T = 3M$, $T = 6M$, $T = 1Y$ and $T = 1.5Y$.

$B \backslash K$	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
0%	30.02	25.07	20.22	15.61	11.41	7.82	4.98	2.94	1.59	0.79	0.15	0.06	0.02
1%	30.00	25.06	20.22	15.61	11.41	7.82	4.98	2.94	1.59	0.79	0.15	0.06	0.02
5%		25.00	20.20	15.60	11.41	7.82	4.98	2.94	1.59	0.79	0.15	0.06	0.02
10%			20.00	15.54	11.39	7.81	4.98	2.93	1.59	0.79	0.15	0.06	0.02
20%				15.00	11.17	7.74	4.96	2.93	1.59	0.79	0.15	0.06	0.02
30%					10.00	7.18	4.72	2.84	1.56	0.78	0.15	0.06	0.02
40%						5.00	3.51	2.23	1.29	0.67	0.14	0.05	0.02
45%							1.75	1.18	0.72	0.39	0.09	0.04	0.01
49%							0.48	0.34	0.21	0.12	0.03	0.01	0.00
0%	35.08	25.46	20.95	16.76	12.99	9.74	7.05	4.91	3.29	1.32	0.79	0.25	0.03
1%	35.00	25.45	20.94	16.75	12.99	9.74	7.05	4.91	3.29	1.32	0.79	0.25	0.03
5%		25.00	20.73	16.66	12.95	9.72	7.04	4.91	3.29	1.32	0.79	0.25	0.03
10%			20.00	16.27	12.75	9.63	7.00	4.89	3.28	1.32	0.79	0.25	0.03
20%				15.00	12.00	9.21	6.77	4.78	3.23	1.31	0.78	0.25	0.03
30%					10.00	7.91	5.97	4.31	2.97	1.24	0.75	0.24	0.03
40%						5.00	3.94	2.95	2.11	0.94	0.59	0.20	0.03
45%							1.83	1.41	1.04	0.49	0.31	0.11	0.02
49%							0.49	0.38	0.29	0.14	0.09	0.03	0.01
0%	45.15	35.58	26.68	22.62	15.51	12.53	9.95	7.77	4.50	2.44	0.86	0.40	0.03
1%	45.00	35.55	26.67	22.61	15.51	12.53	9.95	7.77	4.50	2.44	0.86	0.40	0.03
5%		35.00	26.50	22.52	15.48	12.52	9.95	7.77	4.50	2.44	0.86	0.40	0.03
10%			25.00	21.54	15.10	12.29	9.81	7.69	4.48	2.43	0.86	0.40	0.03
20%				20.00	14.35	11.79	9.49	7.49	4.40	2.40	0.85	0.39	0.03
30%					10.00	8.50	7.07	5.75	3.56	2.03	0.76	0.36	0.03
40%						5.00	4.25	3.54	2.29	1.36	0.54	0.26	0.02
45%							1.88	1.59	1.06	0.65	0.27	0.13	0.01
49%							0.49	0.42	0.28	0.18	0.07	0.04	0.00
0%	50.27	40.83	32.04	27.98	20.70	14.69	12.18	8.11	6.52	3.18	1.06	0.43	0.04
1%	50.00	40.75	32.01	27.96	20.70	14.69	12.18	8.11	6.51	3.18	1.06	0.42	0.04
5%		40.00	31.71	27.78	20.63	14.66	12.17	8.11	6.51	3.17	1.06	0.42	0.04
10%			30.00	26.57	20.06	14.41	12.00	8.04	6.47	3.16	1.06	0.42	0.04
20%				25.00	19.19	13.97	11.69	7.89	6.37	3.14	1.06	0.42	0.04
30%					15.00	11.41	9.74	6.80	5.58	2.85	0.99	0.40	0.04
40%						5.00	4.40	3.26	2.75	1.52	0.58	0.25	0.03
45%							1.90	1.44	1.23	0.70	0.28	0.12	0.01
50%							0.98	0.74	0.64	0.37	0.15	0.07	0.01

Proposition 7.6.1.3. *Using standard properties of the Hermite polynomials defined in* (7.28)*, we have* $\forall n \in \mathbb{N}$ *and* $\forall x \in \mathbb{R}$*:*

$$\int_{-\infty}^{x} H_n(y)\mathcal{N}(y)\mathrm{d}y = \frac{1}{n+1}(H_{n+1}(x)\mathcal{N}(x) + H_n(x)\mathcal{N}'(x)). \tag{7.66}$$

We summarize in the next Proposition some useful convolution properties of $\mathcal{D}^P$ and $\mathcal{J}^P$:

Proposition 7.6.1.4. *Assume* $(\widetilde{\mathcal{H}}^{\sigma}_{x_0})$*. One has for any* $0 \le r < t \le T$ *and for any* $a, b < 0$*:*

$$\int_r^t \mathcal{J}^P(r,s,a)\mathcal{J}^P(s,t,b)\mathrm{d}s = \mathcal{J}^P(r,t,a+b), \tag{7.67}$$

$$\int_r^t \mathcal{J}^P(r,s,a)\mathcal{D}^P(s,t,b)\mathrm{d}s = \mathcal{D}^P(r,t,a+b), \tag{7.68}$$

Table 7.4: Down out barrier Call options prices in the CEV model ($\beta = 0.5, \nu = 0.25$) obtained with the closed-form formula for the maturities $T = 2Y$, $T = 3Y$, $T = 5Y$ and $T = 10Y$.

$B\backslash K$	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
0%	55.35	45.96	37.17	29.22	22.30	19.26	14.05	9.94	6.84	3.69	1.48	0.41	0.02
1%	55.00	45.84	37.13	29.21	22.30	19.26	14.05	9.94	6.84	3.69	1.48	0.41	0.02
5%		45.00	36.76	29.05	22.23	19.22	14.03	9.94	6.83	3.69	1.48	0.41	0.02
10%			35.00	28.12	21.76	18.89	13.87	9.87	6.80	3.68	1.48	0.41	0.02
20%				25.00	19.87	17.45	13.06	9.42	6.57	3.60	1.46	0.41	0.02
30%					15.00	13.43	10.41	7.75	5.56	3.15	1.33	0.38	0.02
40%						10.00	7.94	6.04	4.42	2.58	1.12	0.33	0.02
45%							4.48	3.50	2.62	1.58	0.72	0.22	0.01
49%							0.98	0.78	0.60	0.37	0.18	0.06	0.00
0%	65.37	51.45	47.09	35.09	28.17	22.20	17.17	13.05	8.36	4.37	1.47	0.44	0.02
1%	65.00	51.36	47.03	35.08	28.17	22.20	17.17	13.05	8.36	4.37	1.47	0.44	0.02
5%		50.00	46.04	34.72	27.99	22.11	17.13	13.03	8.36	4.37	1.47	0.44	0.02
10%			45.00	34.25	27.73	21.97	17.05	12.99	8.34	4.36	1.47	0.44	0.02
20%				30.00	24.94	20.19	15.95	12.31	8.03	4.26	1.46	0.44	0.02
30%					20.00	16.62	13.44	10.59	7.10	3.87	1.37	0.42	0.02
40%						10.00	8.34	6.77	4.73	2.71	1.02	0.33	0.02
45%							4.59	3.79	2.71	1.60	0.63	0.21	0.01
49%							0.98	0.82	0.60	0.36	0.15	0.05	0.00
0%	75.55	61.90	53.47	45.71	35.44	29.54	22.08	16.17	10.37	5.70	1.74	0.47	0.01
1%	75.00	61.72	53.39	45.68	35.43	29.53	22.08	16.17	10.37	5.70	1.74	0.47	0.01
5%		60.00	52.38	45.08	35.16	29.38	22.01	16.14	10.36	5.69	1.74	0.47	0.01
10%			50.00	43.49	34.31	28.83	21.73	16.00	10.30	5.68	1.74	0.47	0.01
20%				40.00	32.19	27.33	20.85	15.50	10.08	5.59	1.73	0.47	0.01
30%					25.00	21.75	17.15	13.11	8.78	5.03	1.61	0.45	0.01
40%						15.00	12.16	9.54	6.60	3.91	1.32	0.38	0.01
45%							8.75	6.96	4.91	2.97	1.04	0.31	0.01
49%							1.47	1.20	0.88	0.55	0.21	0.07	0.00
0%	86.12	77.50	69.43	58.41	48.71	40.28	30.90	23.41	15.12	7.53	1.97	0.39	0.00
1%	85.00	76.87	69.06	58.23	48.62	40.24	30.88	23.41	15.12	7.53	1.97	0.39	0.00
5%		75.00	67.80	57.53	48.22	40.01	30.77	23.35	15.10	7.52	1.97	0.39	0.00
10%			65.00	55.75	47.09	39.29	30.38	23.14	15.01	7.50	1.97	0.39	0.00
20%				50.00	43.03	36.44	28.63	22.07	14.51	7.34	1.95	0.38	0.00
30%					35.00	30.28	24.37	19.18	12.93	6.74	1.85	0.37	0.00
40%						20.00	16.58	13.40	9.37	5.10	1.49	0.32	0.00
45%							9.16	7.55	5.43	3.07	0.95	0.21	0.00
49%							1.97	1.65	1.21	0.71	0.23	0.05	0.00

$$\int_r^t \Sigma_s \mathcal{D}^P(r,s,a)\mathcal{D}^P(s,t,b)\mathrm{d}s = \mathcal{N}\Big(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}\Big)e^{-\frac{(a+b)}{2}-\frac{1}{8}\mathcal{V}_r^t} = (\Sigma_t)^{-1}\int_r^t \Sigma_s \mathcal{N}\Big(\frac{a}{\sqrt{\mathcal{V}_r^s}}\Big)e^{-\frac{a}{2}-\frac{1}{8}\mathcal{V}_r^s}\mathcal{J}^P(s,t,b)\mathrm{d}s, \tag{7.69}$$

$$\int_r^t \Sigma_s \mathcal{N}\Big(\frac{a}{\sqrt{\mathcal{V}_r^s}}\Big)e^{-\frac{a}{2}-\frac{1}{8}\mathcal{V}_r^s}\mathcal{D}^P(s,t,b)\mathrm{d}s = (a+b)\mathcal{N}\Big(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}\Big)e^{-\frac{(a+b)}{2}-\frac{1}{8}\mathcal{V}_r^t} + \mathcal{V}_r^t\mathcal{D}^P(r,t,a+b), \tag{7.70}$$

$$\int_r^t \mathcal{J}^P(r,s,a)\mathcal{V}_s^t\mathcal{J}^P(s,t,b)\mathrm{d}s = -\Sigma_t b\mathcal{D}^P(r,t,a+b), \tag{7.71}$$

$$\int_r^t \mathcal{J}^P(r,s,a)\mathcal{V}_s^t\mathcal{D}^P(s,t,b)\mathrm{d}s = \mathcal{V}_r^t\mathcal{D}^P(r,t,a+b) + a\mathcal{N}\Big(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}\Big)e^{-\frac{(a+b)}{2}-\frac{1}{8}\mathcal{V}_r^t}, \tag{7.72}$$

Table 7.5: Errors on down out barrier Call options prices in the CEV model ($\beta = 0.5, \nu = 0.25$) using `App(3,`x_{avg}`)` for the maturities $T = 3M$, $T = 6M$, $T = 1Y$ and $T = 1.5Y$.

$B\backslash K$	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
0%	−8E−7	−6E−7	−4E−7	−3E−7	−6E−7	−1E−6	−1E−6	−9E−7	−6E−7	−3E−7	−3E−7	−4E−7	−5E−7
1%	−5E−7	−5E−6	−5E−6	8E−7	−6E−6	7E−6	−2E−7	−6E−7	−6E−7	−3E−7	−3E−7	−4E−7	−5E−7
5%		−3E−6	−2E−5	1E−5	1E−5	6E−7	2E−6	3E−7	−4E−7	−3E−7	−3E−7	−4E−7	−5E−7
10%			−2E−5	−2E−5	2E−5	3E−5	3E−5	1E−5	4E−6	1E−6	−3E−7	−4E−7	−5E−7
20%				−1E−5	−1E−5	5E−6	4E−5	4E−5	3E−5	1E−5	1E−6	−6E−8	−4E−7
30%					−1E−5	−4E−6	4E−6	1E−5	2E−5	2E−5	6E−6	2E−6	4E−7
40%						−1E−5	1E−5	7E−6	−3E−6	−9E−6	−8E−6	−5E−6	−3E−6
45%							2E−6	8E−6	4E−6	−2E−6	−1E−5	−8E−6	−6E−6
49%							−6E−6	2E−6	2E−6	7E−7	−2E−6	−3E−6	−2E−6
0%	−4E−6	−2E−6	−2E−6	−3E−6	−5E−6	−6E−6	−6E−6	−5E−6	−4E−6	−2E−6	−1E−6	−2E−6	−2E−6
1%	−6E−7	4E−5	9E−5	8E−5	4E−5	2E−5	1E−6	−3E−6	−3E−6	−2E−6	−1E−6	−2E−6	−2E−6
5%		−1E−6	−4E−5	2E−5	1E−4	2E−4	2E−4	1E−4	7E−5	1E−5	5E−6	−7E−7	−2E−6
10%			8E−7	−5E−6	4E−6	8E−5	1E−4	2E−4	1E−4	6E−5	3E−5	6E−6	−2E−6
20%				3E−6	2E−5	2E−5	3E−5	6E−5	9E−5	8E−5	6E−5	2E−5	4E−7
30%					−1E−8	4E−5	4E−5	2E−5	1E−6	−1E−6	2E−6	4E−6	−3E−7
40%						8E−7	3E−5	4E−5	3E−5	−2E−5	−4E−5	−5E−5	−2E−5
45%							5E−6	2E−5	2E−5	3E−6	−1E−5	−3E−5	−2E−5
49%							−2E−6	4E−6	6E−6	3E−6	−3E−7	−6E−6	−6E−6
0%	−2E−5	−1E−5	−1E−5	−2E−5	−3E−5	−3E−5	−3E−5	−3E−5	−2E−5	−1E−5	−6E−6	−8E−6	−1E−5
1%	−9E−6	−5E−5	4E−4	4E−4	1E−4	6E−5	2E−5	−6E−6	−2E−5	−1E−5	−6E−6	−8E−6	−1E−5
5%		−2E−5	−1E−4	2E−4	7E−4	7E−4	6E−4	4E−4	2E−4	5E−5	2E−6	−6E−6	−1E−5
10%			−2E−5	1E−5	4E−5	2E−4	4E−4	6E−4	6E−4	4E−4	1E−4	4E−5	−1E−5
20%				−1E−5	7E−5	5E−5	9E−5	2E−4	3E−4	3E−4	2E−4	9E−5	−6E−6
30%					−2E−5	1E−4	2E−4	2E−4	2E−5	−1E−4	−1E−4	−1E−4	−3E−5
40%						−1E−5	8E−5	1E−4	1E−4	3E−6	−1E−4	−2E−4	−7E−5
45%							1E−5	4E−5	6E−5	4E−5	−4E−5	−7E−5	−5E−5
49%							−4E−6	7E−6	2E−5	1E−5	−6E−6	−2E−5	−1E−5
0%	−6E−5	−3E−5	−4E−5	−6E−5	−8E−5	−9E−5	−9E−5	−8E−5	−7E−5	−3E−5	−2E−5	−2E−5	−3E−5
1%	3E−5	−3E−4	1E−3	1E−3	8E−4	3E−4	2E−4	3E−6	−2E−5	−3E−5	−2E−5	−2E−5	−3E−5
5%		3E−5	−4E−4	1E−4	1E−3	2E−3	1E−3	9E−4	6E−4	2E−4	7E−6	−2E−5	−3E−5
10%			3E−5	9E−5	1E−5	6E−4	1E−3	1E−3	1E−3	9E−4	2E−4	6E−5	−3E−5
20%				2E−5	2E−4	1E−4	3E−4	7E−4	8E−4	8E−4	4E−4	1E−4	−2E−5
30%					1E−5	4E−4	4E−4	1E−4	3E−5	−1E−4	−9E−5	−5E−5	−3E−5
40%						1E−5	1E−4	2E−4	2E−4	6E−5	−2E−4	−3E−4	−2E−4
45%							2E−5	1E−4	1E−4	7E−5	−7E−5	−1E−4	−1E−4
49%							5E−6	5E−5	6E−5	5E−5	−2E−5	−6E−5	−5E−5

$$\int_r^t \mathcal{J}^P(r,s,a)(\mathcal{V}_s^t)^2 \mathcal{J}^P(s,t,b)\mathrm{d}s = -\Sigma_t b \mathcal{V}_r^t \mathcal{D}^P(r,t,a+b) - \Sigma_t ab \mathcal{N}\left(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}\right) e^{-\frac{(a+b)}{2} - \frac{1}{8}\mathcal{V}_r^t}, \tag{7.73}$$

$$\begin{aligned} &\int_r^t \Sigma_s \mathcal{D}^P(r,s,a)\mathcal{V}_s^t \mathcal{D}^P(s,t,b)\mathrm{d}s \\ =&\frac{1}{2}\mathcal{V}_r^t e^{-\frac{(a+b)}{2} - \frac{1}{8}\mathcal{V}_r^t}\{\mathcal{N}(x) + \frac{(a-b)}{\sqrt{\mathcal{V}_r^t}}[x\mathcal{N}(x) + \mathcal{N}'(x)]\}\Big|_{x=\frac{a+b}{\sqrt{\mathcal{V}_r^t}}}, \end{aligned} \tag{7.74}$$

$$e^{\frac{(a+b)}{2} + \frac{1}{8}\mathcal{V}_r^t} \int_r^t \Sigma_s \mathcal{N}\left(\frac{a}{\sqrt{\mathcal{V}_r^s}}\right) e^{-\frac{a}{2} - \frac{1}{8}\mathcal{V}_r^s} \mathcal{V}_s^t \mathcal{D}^P(s,t,b)\mathrm{d}s \tag{7.75}$$

Table 7.6: Errors on down out barrier Call options prices in the CEV model ($\beta = 0.5, \nu = 0.25$) using `App(3,` x_{avg}`)` for the maturities $T = 2Y$, $T = 3Y$, $T = 5Y$ and $T = 10Y$.

B\K	1%	5%	10%	20%	30%	40%	50%	60%	70%	80%	90%	95%	99%
0%	−1E−4	−7E−5	−8E−5	−1E−4	−2E−4	−2E−4	−2E−4	−2E−4	−1E−4	−7E−5	−3E−5	−5E−5	−5E−5
1%	−2E−6	−9E−4	2E−3	2E−3	1E−3	1E−3	3E−4	4E−5	−6E−5	−6E−5	−3E−5	−5E−5	−5E−5
5%		8E−6	−8E−4	9E−4	3E−3	3E−3	2E−3	2E−3	8E−4	2E−4	2E−5	−4E−5	−5E−5
10%			3E−6	−1E−4	2E−4	8E−4	2E−3	2E−3	2E−3	1E−3	4E−4	3E−5	−5E−5
20%				6E−6	5E−4	4E−4	2E−4	4E−4	8E−4	1E−3	7E−4	2E−4	−4E−5
30%					7E−6	4E−4	7E−4	4E−4	4E−5	−4E−4	−4E−4	−2E−4	−6E−5
40%						5E−6	5E−4	6E−4	3E−4	−2E−4	−6E−4	−6E−4	−1E−4
45%							2E−4	3E−4	4E−4	1E−4	−3E−4	−5E−4	−2E−4
49%							3E−6	6E−5	9E−5	7E−5	−2E−5	−1E−4	−6E−5
0%	−4E−4	−2E−4	−2E−4	−4E−4	−4E−4	−5E−4	−5E−4	−4E−4	−3E−4	−2E−4	−8E−5	−1E−4	−1E−4
1%	5E−5	4E−4	3E−3	4E−3	2E−3	1E−3	2E−4	−1E−4	−2E−4	−2E−4	−8E−5	−1E−4	−1E−4
5%		1E−5	−8E−4	4E−4	4E−3	6E−3	6E−3	5E−3	2E−3	7E−4	4E−5	−1E−4	−1E−4
10%			2E−5	−8E−4	1E−3	4E−3	6E−3	6E−3	4E−3	2E−3	3E−4	−7E−5	−1E−4
20%				1E−5	1E−3	6E−4	4E−4	8E−4	2E−3	2E−3	1E−3	4E−4	−1E−4
30%					1E−5	1E−3	1E−3	8E−4	−1E−4	−6E−4	−4E−4	−2E−4	−1E−4
40%						3E−6	8E−4	1E−3	8E−4	−2E−4	−1E−3	−1E−3	−3E−4
45%							2E−4	6E−4	7E−4	2E−4	−6E−4	−9E−4	−3E−4
49%							9E−6	1E−4	2E−4	1E−4	−7E−5	−2E−4	−1E−4
0%	−1E−3	−7E−4	−8E−4	−1E−3	−1E−3	−2E−3	−2E−3	−1E−3	−9E−4	−5E−4	−3E−4	−4E−4	−3E−4
1%	6E−6	−1E−3	1E−2	1E−2	8E−3	4E−3	6E−4	−5E−4	−7E−4	−5E−4	−3E−4	−4E−4	−3E−4
5%		2E−5	−4E−3	−2E−3	1E−2	2E−2	2E−2	1E−2	5E−3	2E−3	−6E−5	−4E−4	−3E−4
10%			2E−5	−2E−4	−4E−4	4E−3	1E−2	1E−2	1E−2	7E−3	1E−3	−2E−4	−3E−4
20%				1E−5	2E−3	7E−4	1E−3	5E−3	8E−3	8E−3	3E−3	4E−4	−3E−4
30%					6E−6	3E−3	3E−3	2E−3	−8E−4	−2E−3	−8E−4	−4E−4	−2E−4
40%						2E−6	2E−3	3E−3	1E−3	−2E−3	−4E−3	−3E−3	−5E−4
45%							1E−3	2E−3	2E−3	−9E−5	−3E−3	−3E−3	−6E−4
49%							3E−5	3E−4	4E−4	3E−4	−3E−4	−6E−4	−2E−4
0%	7E−5	1E−3	−3E−4	−3E−3	−5E−3	−6E−3	−6E−3	−5E−3	−4E−3	−2E−3	−1E−3	−2E−3	−4E−4
1%	−4E−6	−4E−2	1E−2	7E−2	6E−2	3E−2	1E−2	2E−3	−1E−3	−1E−3	−1E−3	−2E−3	−4E−4
5%		1E−5	−2E−2	−3E−3	5E−2	7E−2	6E−2	4E−2	1E−2	2E−3	−9E−4	−2E−3	−4E−4
10%			1E−5	−8E−3	−6E−3	2E−2	5E−2	6E−2	4E−2	2E−2	8E−4	−2E−3	−4E−4
20%				−5E−7	7E−3	2E−3	3E−4	9E−3	2E−2	3E−2	8E−3	−8E−4	−4E−4
30%					3E−6	9E−3	9E−3	3E−3	−5E−3	−4E−3	4E−4	−1E−3	−4E−4
40%						4E−6	7E−3	9E−3	4E−3	−8E−3	−1E−2	−8E−3	−5E−4
45%							2E−3	5E−3	5E−3	2E−5	−9E−3	−8E−3	−6E−4
49%							1E−4	8E−4	1E−3	8E−4	−1E−3	−2E−3	−2E−4

$$=\frac{1}{2}(\mathcal{V}_r^t)^{\frac{3}{2}}\{-\frac{1}{6}H_3(x)\mathcal{N}(x)-\frac{1}{6}H_2(x)\mathcal{N}'(x)+\frac{1}{2}\mathcal{N}'(x)+\frac{(a-b)}{2\sqrt{\mathcal{V}_r^t}}[(H_2(x)+2)\mathcal{N}(x)+x\mathcal{N}'(x)]\}\Big|_{x=\frac{a+b}{\sqrt{\mathcal{V}_r^t}}},$$

where H_n denote the n^{th} Hermite polynomial defined in (7.28). The results remain valid if we assume $(\widetilde{\mathcal{H}}^{\sigma}_{x_{avg}})$ and consider $\mathcal{D}^{P,x_{avg}}$, $\mathcal{J}^{P,x_{avg}}$ and the local variance function frozen at x_{avg}.

Remark 7.6.1.1. *There are some natural intuitions of the equalities (7.67)-(7.68) if we consider the case of the standard Brownian motion. By independence of the increments we have that $\tau_a + \tau_b = \tau_{a+b}$ in law. Similarly we have $\mathbb{P}(W_t < a+b) = \mathbb{E}[\mathbb{1}_{\tau_a<t}\mathbb{P}(W_{t-\tau_a} < b)]$.*

Proof. We introduce the auxiliary process $(Z_t = \int_0^t \sigma_s \mathrm{d}W_s)_{t\in[0,T]}$ and for $y < 0$ and $0 \le r < t \le T$, we set

$\tau^Z_{r,y} = \inf\{s > r : Z_s - Z_r = y\}$. $Z_t - Z_r$ and $\tau^Z_{r,y}$ admit the densities:

$$\mathcal{D}(r,t,z) = \frac{e^{-\frac{z^2}{2\mathcal{V}_r^t}}}{\sqrt{2\pi\mathcal{V}_r^t}}, \quad \mathcal{J}(r,s,y) = -\frac{\Sigma_s y}{\mathcal{V}_r^s}\mathcal{D}(r,s,y)\mathbb{1}_{s\geq r} = \Sigma_s\partial_y\widetilde{\mathcal{D}}(r,s,y)\mathbb{1}_{s\geq r}. \tag{7.76}$$

We aim at showing that $\mathcal{D}$ and $\mathcal{J}$ satisfy equivalent convolution properties and then we will conclude using the fact that $-\frac{(a+\frac{1}{2}\mathcal{V}_r^s)^2}{2\mathcal{V}_r^s} - \frac{(b+\frac{1}{2}\mathcal{V}_s^t)^2}{2\mathcal{V}_s^t} = -\frac{a^2}{2\mathcal{V}_r^s} - \frac{b^2}{2\mathcal{V}_s^t} - \frac{(a+b)}{2} - \frac{1}{8}\mathcal{V}_r^t$.
First for (7.67) we have using the definition of $\mathcal{J}$ and the independent increments of Z:

$$\begin{aligned}\int_r^t \mathcal{J}(r,s,a+b)\mathrm{d}s =& \mathbb{P}(\tau^Z_{r,a+b} \leq t) = \mathbb{E}[\mathbb{1}_{\tau^Z_{r,a}\leq t}\mathbb{P}(\tau^Z_{\tau^Z_{r,a},b} \leq t)]\\ =& \int_r^t \mathbb{P}(\tau^Z_{s,b} \leq t)\mathcal{J}(r,s,a)\mathrm{d}s = \int_r^t (\int_s^t \mathcal{J}(s,u,b)\mathrm{d}u)\mathcal{J}(r,s,a)\mathrm{d}s.\end{aligned}$$

Then derive w.r.t. t to get $\int_r^t \mathcal{J}(r,s,a)\mathcal{J}(s,t,b)\mathrm{d}s = \mathcal{J}(r,t,a+b)$.
Then for (7.68) utilize (7.76) and (7.67) to get by integration:

$$\begin{aligned}\int_r^t \mathcal{J}(r,s,a)\mathcal{D}(s,t,b)\mathrm{d}s =& \int_r^t \mathcal{J}(r,s,a)(\frac{1}{\Sigma_t}\int_{-\infty}^b \mathcal{J}(s,t,x)\mathrm{d}x)\mathrm{d}s\\ =& \frac{1}{\Sigma_t}\int_{-\infty}^b (\int_r^t \mathcal{J}(r,s,a)\mathcal{J}(s,t,x)\mathrm{d}s)\mathrm{d}x = \frac{1}{\Sigma_t}\int_{-\infty}^b \mathcal{J}(r,t,a+x)\mathrm{d}x = \mathcal{D}(r,t,a+b)\end{aligned}$$

Similarly we deduce (7.69) from (7.68) using again an integration. The proof of (7.70) is analogous using (7.66) and (7.69). We skip details. For (7.71), we use (7.68) and write

$$\int_r^t \mathcal{J}(r,s,a)\mathcal{V}_s^t\mathcal{J}(s,t,b)\mathrm{d}s = -\Sigma_t b\int_r^t \mathcal{J}(r,s,a)\mathcal{D}(s,t,b)\mathrm{d}s = -\Sigma_t b\mathcal{D}(r,t,a+b).$$

Then by integration it comes:

$$\begin{aligned}&\int_r^t \mathcal{J}(r,s,a)\mathcal{V}_s^t\mathcal{D}(s,t,b)\mathrm{d}s = -\int_{-\infty}^b x\mathcal{D}(r,t,a+x)\mathrm{d}x\\ =& -\int_{-\infty}^b (a+x)\mathcal{D}(r,t,a+x)\mathrm{d}x + a\int_{-\infty}^b \mathcal{D}(r,t,a+x)\mathrm{d}x = \mathcal{V}_r^t\mathcal{D}(r,t,a+b) + a\mathcal{N}(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}).\end{aligned}$$

We now pass to (7.73). Use (7.72) to get:

$$\int_r^t \mathcal{J}(r,s,a)(\mathcal{V}_s^t)^2\mathcal{J}(s,t,b)\mathrm{d}s = -\Sigma_t b\int_r^t \mathcal{J}(r,s,a)\mathcal{V}_s^t\mathcal{D}(s,t,b)\mathrm{d}s = -\Sigma_t b\mathcal{V}_r^t\mathcal{D}(r,t,a+b) - \Sigma_t ab\mathcal{N}(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}).$$

Then by integration, starting from (7.72) and using (7.66), one obtains:

$$\begin{aligned}&\int_r^t \Sigma_s\mathcal{D}(r,s,a)\mathcal{V}_s^t\mathcal{D}(s,t,b)\mathrm{d}s = \int_{-\infty}^a (\mathcal{V}_r^t\mathcal{D}(r,t,x+b) + x\mathcal{N}(\frac{x+b}{\sqrt{\mathcal{V}_r^t}}))\mathrm{d}x\\ =& \mathcal{V}_r^t\{\mathcal{N}(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}) + \int_{-\infty}^{\frac{a+b}{\sqrt{\mathcal{V}_r^t}}} (x - \frac{b}{\sqrt{\mathcal{V}_r^t}})\mathcal{N}(x)\mathrm{d}x\}\\ =& \mathcal{V}_r^t\{\mathcal{N}(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}) + \frac{1}{2}[H_2(\frac{a+b}{\sqrt{\mathcal{V}_r^t}})\mathcal{N}(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}) + H_1(\frac{a+b}{\sqrt{\mathcal{V}_r^t}})\mathcal{N}'(\frac{a+b}{\sqrt{\mathcal{V}_r^t}})] - \frac{b}{\sqrt{\mathcal{V}_r^t}}[\frac{(a+b)}{\sqrt{\mathcal{V}_r^t}}\mathcal{N}(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}) + \mathcal{N}'(\frac{a+b}{\sqrt{\mathcal{V}_r^t}})]\}\\ =& \frac{1}{2}\mathcal{V}_r^t\{\mathcal{N}(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}) + \frac{(a-b)}{\sqrt{\mathcal{V}_r^t}}[\frac{(a+b)}{\sqrt{\mathcal{V}_r^t}}\mathcal{N}(\frac{a+b}{\sqrt{\mathcal{V}_r^t}}) + \mathcal{N}'(\frac{a+b}{\sqrt{\mathcal{V}_r^t}})]\}.\end{aligned}$$

Finally, we get using (7.74) and (7.66):

$$\int_r^t \Sigma_s \mathcal{N}(\frac{a}{\sqrt{\mathcal{V}_r^s}})\mathcal{V}_s^t \mathcal{D}(s,t,b)\mathrm{d}s = \int_{-\infty}^a \frac{1}{2}\mathcal{V}_r^t\{\mathcal{N}(\frac{y+b}{\sqrt{\mathcal{V}_r^t}}) + \frac{(y-b)}{\sqrt{\mathcal{V}_r^t}}[\frac{(y+b)}{\sqrt{\mathcal{V}_r^t}}\mathcal{N}(\frac{y+b}{\sqrt{\mathcal{V}_r^t}}) + \mathcal{N}'(\frac{y+b}{\sqrt{\mathcal{V}_r^t}})]\}\mathrm{d}y$$

$$=\frac{1}{2}(\mathcal{V}_r^t)^{\frac{3}{2}}\{x\mathcal{N}(x)+\mathcal{N}'(x)+\int_{-\infty}^x(y-\frac{2b}{\sqrt{\mathcal{V}_r^t}})[y\mathcal{N}(y)+\mathcal{N}'(y)]\mathrm{d}y\}\Big|_{x=\frac{a+b}{\sqrt{\mathcal{V}_r^t}}}$$

$$=\frac{1}{2}(\mathcal{V}_r^t)^{\frac{3}{2}}\{x\mathcal{N}(x)+\mathcal{N}'(x)+\frac{(a-b)}{2\sqrt{\mathcal{V}_r^t}}[H_2(x)\mathcal{N}(x)+x\mathcal{N}'(x)+2\mathcal{N}(x)]$$
$$-\frac{1}{2}\int_{-\infty}^x[H_2(y)\mathcal{N}(y)+y\mathcal{N}'(y)+2\mathcal{N}(y)]\mathrm{d}y\}\Big|_{x=\frac{a+b}{\sqrt{\mathcal{V}_r^t}}}$$

$$=\frac{1}{2}(\mathcal{V}_r^t)^{\frac{3}{2}}\{x\mathcal{N}(x)+\mathcal{N}'(x)+\frac{(a-b)}{2\sqrt{\mathcal{V}_r^t}}[H_2(x)\mathcal{N}(x)+x\mathcal{N}'(x)+2\mathcal{N}(x)]$$
$$-\frac{1}{6}[H_3(x)\mathcal{N}(x)+H_2(x)\mathcal{N}'(x)-3\mathcal{N}'(x)+6x\mathcal{N}(x)+6\mathcal{N}'(x)]\}\Big|_{x=\frac{a+b}{\sqrt{\mathcal{V}_r^t}}}$$

$$=\frac{1}{2}(\mathcal{V}_r^t)^{\frac{3}{2}}\{\frac{1}{6}[-x^3+3x]\mathcal{N}(x)+\frac{1}{6}[-x^2+4]\mathcal{N}'(x)+\frac{(a-b)}{2\sqrt{\mathcal{V}_r^t}}[H_2(x)\mathcal{N}(x)+x\mathcal{N}'(x)+2\mathcal{N}(x)]\}\Big|_{x=\frac{a+b}{\sqrt{\mathcal{V}_r^t}}}.$$

□

7.6.2 Proof of Lemmas 7.3.4.4-7.3.4.5.

▷**Proof of Lemma 7.3.4.4.** As $\mathcal{L}_2^x v_{\underline{O},T}^{P,\phi_t}(u,b) = 0$ for any u $\in [0,T[$ (see Lemma 7.3.2.3 equation (7.11)), we have following the proof of Lemma 7.3.2.1 and using Lemma 7.3.4.3 that for any $0 \le u \le s < t < T$ and any $x \ge b$:

$$\begin{aligned} v_{\underline{O},s}^{P,\rho_{s,t}}(u,x) =&\mathbb{E}[(X^P_{s\wedge\tau^P_{u,b}}-x_0)\mathcal{L}_2^x v_{\underline{O},t}^{P,\phi_t}(s\wedge\tau^P_{u,b},X^P_{s\wedge\tau^P_{u,b}})|X_u^P=x] \\ =&(x-x_0)\mathcal{L}_2^x v_{\underline{O},t}^{P,\phi_t}(u,x)+\mathbb{E}[\int_u^{s\wedge\tau^P_{u,b}}\mathcal{L}_3^x v_{\underline{O},t}^{P,\phi_t}(r,X_r^P)\Sigma_r\mathrm{d}r|X_u^P=x] \\ =&(x-x_0)\mathcal{L}_2^x v_{\underline{O},t}^{P,\phi_t}(u,x)+\mathcal{V}_u^s\mathcal{L}_3^x v_{\underline{O},t}^{P,\phi_t}(u,x)-\mathbb{E}[\mathcal{V}^s_{\tau^P_{u,b}}\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{L}_3^x v_{\underline{O},t}^{P,\phi_t}(\tau^P_{u,b},b)|X_u^P=x]. \end{aligned} \tag{7.77}$$

We now intend to express $v_s^{P,\rho_{s,t}}(u,x)$ in term of derivatives of $v_{\underline{O},T}^{P,h}$. First, we have with (7.40):

$$\begin{aligned} (x-x_0)\mathcal{L}_2^x v_{\underline{O},t}^{P,\phi_t}(u,x) =&(x-x_0)\{2\mathcal{L}_3^x+(x-x_0)\mathcal{L}_4^x+\mathcal{V}_u^t\mathcal{L}_5^x\}v_{\underline{O},T}^{P,h}(u,x) \\ &-(x-x_0)\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le t}(2\mathcal{L}_3^x+\mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}_5^x)v_{\underline{O},T}^{P,h}(\tau^P_{u,b},b)|X_u^P=x]. \end{aligned} \tag{7.78}$$

Then write $\mathcal{L}_3^x v_{\underline{O},t}^{P,\phi_t}(u,x) = (\mathcal{L}_1^x \circ \mathcal{L}_2^x)v_{\underline{O},t}^{P,\phi_t}(u,x)$ and use Definition 7.2.1.1 to obtain on the one hand:

$$\begin{aligned} \mathcal{V}_u^s\mathcal{L}_3^x v_{\underline{O},t}^{P,\phi_t}(u,x) =&\mathcal{V}_u^s\{\mathcal{L}_1^x\circ(2\mathcal{L}_3^x+(x-x_0)\mathcal{L}_4^x+\mathcal{V}_u^t\mathcal{L}_5^x)\}v_{\underline{O},T}^{P,h}(u,x) \\ &-\mathcal{V}_u^s\mathcal{L}_1^x\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le t}(2\mathcal{L}_3^x+\mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}_5^x)v_{\underline{O},T}^{P,h}(\tau^P_{u,b},b)|X_u^P=x] \\ =&\mathcal{V}_u^s\{3\mathcal{L}_4^x+\frac{1}{2}\mathcal{L}_2^x+(x-x_0)\mathcal{L}_5^x+\mathcal{V}_u^t(\mathcal{L}_6^x+\frac{1}{4}\mathcal{L}_4^x)\}v_{\underline{O},T}^{P,h}(u,x) \\ &-\mathcal{V}_u^s\mathcal{L}_1^x\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le t}(2\mathcal{L}_3^x+\mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}_5^x)v_{\underline{O},T}^{P,h}(\tau^P_{u,b},b)|X_u^P=x], \end{aligned} \tag{7.79}$$

and on the other hand with Lemma 7.2.1.1:

$$\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\mathcal{L}_3^x v_{\underline{O},t}^{P,\phi_t}(\tau^P_{u,b},b)|X_u^P=x]$$

$$
\begin{aligned}
=&(b-x_0)\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\mathcal{L}^x_5 v^{P,h}_{\underline{o},T}(\tau^P_{u,b},b)|X^P_u=x]\\
&-\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\lim_{(r,y)\to(\tau^P_{u,b},b)}\mathcal{L}^y_1\{\mathbb{E}[\mathbb{1}_{\tau^P_{r,b}\le t}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_{r,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{r,b},b)|X^P_r=y]\}\big|X^P_u=x].
\end{aligned}\tag{7.80}
$$

Then using the relation (7.64) of Proposition 7.6.1.1, we obtain:

$$
\begin{aligned}
&\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\lim_{(r,y)\to(\tau^P_{u,b},b)}\mathcal{L}^y_1\{\mathbb{E}[\mathbb{1}_{\tau^P_{r,b}\le t}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_{r,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{r,b},b)|X^P_r=y]\}\big|X^P_u=x]\\
=&\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\lim_{\epsilon\uparrow 0}\int_{\tau^P_{u,b}}^t\Sigma_r(\frac{1}{\mathcal{V}^r_{\tau^P_{u,b}}}-\frac{\epsilon^2}{(\mathcal{V}^r_{\tau^P_{u,b}})^2})\mathcal{D}^P(\tau^P_{u,b},r,\epsilon)(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\big|X^P_u=x]\\
=&\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\int_s^t\frac{\Sigma_r}{\mathcal{V}^r_{\tau^P_{u,b}}}\mathcal{D}^P(\tau^P_{u,b},r,0)(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\big|X^P_u=x]\\
&+\lim_{\epsilon\uparrow 0}\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\int_{\tau^P_{u,b}}^s\Sigma_r(\frac{1}{\mathcal{V}^r_{\tau^P_{u,b}}}-\frac{\epsilon^2}{(\mathcal{V}^r_{\tau^P_{u,b}})^2})\mathcal{D}^P(\tau^P_{u,b},r,\epsilon)(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\big|X^P_u=x].
\end{aligned}\tag{7.81}
$$

We now treat the second term of the above r.h.s.. Write that $\mathcal{V}^s_{\tau^P_{u,b}}=\mathcal{V}^r_{\tau^P_{u,b}}+\mathcal{V}^s_r$, use again the Proposition 7.6.1.1 and utilize the convolution results postponed to Proposition 7.6.1.4 to get:

$$
\begin{aligned}
&\lim_{\epsilon\uparrow 0}\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\int_{\tau^P_{u,b}}^s\Sigma_r(\frac{1}{\mathcal{V}^r_{\tau^P_{u,b}}}-\frac{\epsilon^2}{(\mathcal{V}^r_{\tau^P_{u,b}})^2})\mathcal{D}^P(\tau^P_{u,b},r,\epsilon)(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\big|X^P_u=x]\\
=&\lim_{\epsilon\uparrow 0}\int_u^s\mathcal{J}^P(u,\theta,b-x)\mathcal{V}^s_\theta(\int_\theta^s\Sigma_r(\frac{1}{\mathcal{V}^r_\theta}-\frac{\epsilon^2}{(\mathcal{V}^r_\theta)^2})\mathcal{D}^P(\theta,r,\epsilon)(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r)\mathrm{d}\theta\\
=&\lim_{\epsilon\uparrow 0}\int_u^s(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)(\int_u^r\mathcal{J}^P(u,\theta,b-x)\Sigma_r(1-\frac{\epsilon^2}{\mathcal{V}^r_\theta})\mathcal{D}^P(\theta,r,\epsilon)\mathrm{d}\theta)\mathrm{d}r\\
&+\lim_{\epsilon\uparrow 0}\int_u^s\mathcal{V}^s_r(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)((-\partial_\epsilon-\frac{1}{2}\mathcal{I})\int_u^r\mathcal{J}^P(u,\theta,b-x)\mathcal{J}^P(\theta,r,\epsilon)\mathrm{d}\theta)\mathrm{d}r\\
=&\lim_{\epsilon\uparrow 0}\int_u^s(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\{\Sigma_r\mathcal{D}^P(u,r,b-x+\epsilon)+\epsilon\mathcal{J}^P(u,r,b-x+\epsilon)\}\mathrm{d}r\\
&+\lim_{\epsilon\uparrow 0}\int_u^s\mathcal{V}^s_r(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)(-\partial_\epsilon-\frac{1}{2}\mathcal{I})\mathcal{J}^P(u,r,b-x+\epsilon)\mathrm{d}r\\
=&\lim_{\epsilon\uparrow 0}\int_u^s(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\Big[\Sigma_r\mathcal{D}^P(u,r,b-x+\epsilon)\{1+\mathcal{V}^s_r(\frac{1}{\mathcal{V}^r_u}-\frac{(x-b+\epsilon)^2}{(\mathcal{V}^r_u)^2})\}+\epsilon\mathcal{J}^P(u,r,b-x+\epsilon)\Big]\mathrm{d}r\\
=&\int_u^s(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\Sigma_r\mathcal{D}^P(u,r,b-x)\{1+\mathcal{V}^s_r(\frac{1}{\mathcal{V}^r_u}-\frac{(x-b)^2}{(\mathcal{V}^r_u)^2})\}\mathrm{d}r\\
=&\int_u^s\Sigma_r(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathcal{D}^P(u,r,b-x)\{\frac{(x-b)^2}{\mathcal{V}^r_u}+\mathcal{V}^s_u(\frac{1}{\mathcal{V}^r_u}-\frac{(x-b)^2}{(\mathcal{V}^r_u)^2})\}\mathrm{d}r\\
=&\{(x-b)\mathcal{I}+\mathcal{V}^s_u\mathcal{L}^x_1\}\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{u,b},b)|X^P_u=x].
\end{aligned}\tag{7.82}
$$

Regrouping the intermediate results (7.81)-(7.82) leads to:

$$
\begin{aligned}
&\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\lim_{(r,y)\to(\tau^P_{u,b},b)}\mathcal{L}^y_1\{\mathbb{E}[\mathbb{1}_{\tau^P_{r,b}\le t}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_{r,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{r,b},b)|X^P_r=y]\}\big|X^P_u=x]\\
=&\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\int_s^t\frac{\Sigma_r}{\mathcal{V}^r_{\tau^P_{u,b}}}\mathcal{D}^P(\tau^P_{u,b},r,0)(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\big|X^P_u=x]\\
&+\{(x-b)\mathcal{I}+\mathcal{V}^s_u\mathcal{L}^x_1\}\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{u,b},b)|X^P_u=x].
\end{aligned}\tag{7.83}
$$

Combine (7.77)-(7.78)-(7.79)-(7.80)-(7.83) to achieve the proof of (7.45). We now pass to the proof of (7.46). First use Lemma 7.3.4.1 to get for the contributions coming from $v^{P,\psi_t}_{\underline{o},t}$:

$$\begin{aligned}\frac{1}{4}\int_0^T \Sigma^{(2)}_t v^{P,\psi_t}_{\underline{o},t}(0,x_0)\mathrm{d}t =&\frac{1}{4}\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(0,x_0)\int_0^T \Sigma^{(2)}_t \mathcal{V}^t_s \mathrm{d}t + \frac{1}{2}(x_0-b)\int_0^T \Sigma^{(2)}_t \mathbb{E}[\mathbb{1}_{\tau^P_b\le t}\mathcal{V}^t_{\tau^P_b}\mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(\tau^P_b,b)]\mathrm{d}t\\ &+\frac{1}{2}(\mathcal{L}^x_4+\frac{1}{4}\mathcal{L}^x_2)v^{P,h}_{\underline{o},T}(0,x_0)\int_0^T \Sigma^{(2)}_t\vec{\omega}(\Sigma,\Sigma)^t_0\mathrm{d}t\\ =&\frac{1}{4}\vec{\omega}(\Sigma^{(2)},\Sigma)^T_0\mathcal{L}^x_2 v^{P,h}_{\underline{o},T}(0,x_0)+\frac{1}{2}(x_0-b)\mathbb{E}[\mathbb{1}_{\tau^P_b\le T}\vec{\omega}(\Sigma^{(2)},\Sigma)^T_{\tau^P_b}\mathcal{L}^x_3 v^{P,h}_{\underline{o},T}(\tau^P_b,b)]\\ &+\frac{1}{2}\vec{\omega}(\Sigma^{(2)},\Sigma,\Sigma)^T_0(\mathcal{L}^x_4+\frac{1}{4}\mathcal{L}^x_2)v^{P,h}_{\underline{o},T}(0,x_0).\end{aligned}\tag{7.84}$$

For $v^{P,\rho_{s,t}}_{\underline{o},s}$, we have:

$$\frac{1}{4}\int_0^T \Sigma^{(1)}_t(\int_0^t \Sigma^{(1)}_s v^{P,\rho_{s,t}}_{\underline{o},s}(0,x_0)\mathrm{d}s)\mathrm{d}t = \frac{1}{4}\sum_{n=1}^5 T_n(v^{P,\rho_{s,t}}_{\underline{o},s}),\tag{7.85}$$

where $T_1(v^{P,\rho_{s,t}}_{\underline{o},s})$ is equal to:

$$\begin{aligned}&(3\mathcal{L}^x_4+\frac{1}{2}\mathcal{L}^x_2)v^{P,h}_{\underline{o},T}(0,x_0)\int_0^T \Sigma^{(1)}_t(\int_0^t \Sigma^{(1)}_s\mathcal{V}^s_0\mathrm{d}s)\mathrm{d}t+(\mathcal{L}^x_6+\frac{1}{4}\mathcal{L}^x_4)v^{P,h}_{\underline{o},T}(0,x_0)\int_0^T \Sigma^{(1)}_t\mathcal{V}^t_0(\int_0^t \Sigma^{(1)}_s\mathcal{V}^s_0\mathrm{d}s)\mathrm{d}t\\ =&\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)},\Sigma)^T_0(3\mathcal{L}^x_4+\frac{1}{2}\mathcal{L}^x_2)v^{P,h}_{\underline{o},T}(0,x_0)+\frac{1}{2}\vec{\omega}^2(\Sigma^{(1)},\Sigma)^T_0(\mathcal{L}^x_6+\frac{1}{4}\mathcal{L}^x_4)v^{P,h}_{\underline{o},T}(0,x_0),\end{aligned}$$

where $T_2(v^{P,\rho_{s,t}}_{\underline{o},s})$ is equal to:

$$(x_0-b)\int_0^T \Sigma^{(1)}_t\int_0^t \Sigma^{(1)}_s\mathbb{E}[\mathbb{1}_{\tau^P_b\le s}\mathcal{V}^s_{\tau^P_b}\mathcal{L}^x_5 v^{P,h}_{\underline{o},T}(\tau^P_b,x)]\mathrm{d}s\mathrm{d}t=(x_0-b)\mathbb{E}[\mathbb{1}_{\tau^P_b\le T}\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)},\Sigma)^T_{\tau^P_b}\mathcal{L}^x_5 v^{P,h}_{\underline{o},T}(\tau^P_b,b)],$$

where $T_3(v^{P,\rho_{s,t}}_{\underline{o},s})$ is equal to:

$$\begin{aligned}&(x_0-b)\int_0^T \Sigma^{(1)}_t(\int_0^t \Sigma^{(1)}_s\mathbb{E}[\mathbb{1}_{\tau^P_b\le s}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_b}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_b,b)]\mathrm{d}s)\mathrm{d}t\\ =&(x_0-b)\mathbb{E}[\mathbb{1}_{\tau^P_b\le T}\{2\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)})^T_{\tau^P_b}\mathcal{L}^x_3+(\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)},\Sigma)^T_{\tau^P_b}+\vec{\omega}(\Sigma^{(1)},\Sigma,\Sigma^{(1)})^T_{\tau^P_b})\mathcal{L}^x_5\}v^{P,h}_{\underline{o},T}(\tau^P_b,b)],\end{aligned}$$

where $T_4(v^{P,\rho_{s,t}}_{\underline{o},s})$ is equal to:

$$\begin{aligned}&-\int_0^T \Sigma^{(1)}_t(\int_0^t \Sigma^{(1)}_s\mathcal{V}^s_0(\int_s^t \Sigma_r(\frac{1}{\mathcal{V}^r_0}-\frac{(x_0-b)^2}{(\mathcal{V}^r_0)^2})\mathcal{D}^P(0,r,b-x_0)(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r)\mathrm{d}s)\mathrm{d}t\\ =&-\int_0^T \frac{\Sigma_r}{\mathcal{V}^r_0}\vec{\omega}(\Sigma^{(1)},\Sigma)^r_0\{2\vec{\omega}(\Sigma^{(1)})^T_r\mathcal{L}^x_3+\vec{\omega}(\Sigma^{(1)},\Sigma)^T_r\mathcal{L}^x_5\}v^{P,h}_{\underline{o},T}(r,b)\mathcal{D}^P(0,r,b-x_0)\mathrm{d}r\\ &+(x_0-b)\mathbb{E}[\mathbb{1}_{\tau^P_b\le T}\vec{\omega}(\Sigma^{(1)},\Sigma)^{\tau^P_b}_0(\mathcal{V}^{\tau^P_b}_0)^{-1}\{2\vec{\omega}(\Sigma^{(1)})^T_{\tau^P_b}\mathcal{L}^x_3+\vec{\omega}(\Sigma^{(1)},\Sigma)^T_{\tau^P_b}\mathcal{L}^x_5\}v^{P,h}_{\underline{o},T}(\tau^P_b,b)],\end{aligned}$$

and where $T_5(v^{P,\rho_{s,t}}_{\underline{o},s})$ is equal to:

$$\begin{aligned}&\int_0^T \Sigma^{(1)}_t(\int_0^t \Sigma^{(1)}_s\mathbb{E}[\mathbb{1}_{\tau^P_b\le s}\mathcal{V}^s_{\tau^P_b}\int_s^t \Sigma_r(\mathcal{V}^r_{\tau^P_b})^{-1}\mathcal{D}^P(\tau^P_b,r,0)(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r]\mathrm{d}s)\mathrm{d}t\\ =&\int_0^T \Sigma_r\{2\vec{\omega}(\Sigma^{(1)})^T_r\mathcal{L}^x_3+\vec{\omega}(\Sigma^{(1)},\Sigma)^T_r\mathcal{L}^x_5\}v^{P,h}_{\underline{o},T}(r,b)\mathbb{E}[\mathbb{1}_{\tau^P_b\le r}\vec{\omega}(\Sigma^{(1)},\Sigma)^r_{\tau^P_b}(\mathcal{V}^r_{\tau^P_b})^{-1}\mathcal{D}^P(\tau^P_b,r,0)]\mathrm{d}r,\end{aligned}$$

Then we write that:

$$\begin{aligned}
\vec{\omega}(\Sigma^{(1)},\Sigma)^r_{\tau^P_b}(\mathcal{V}^r_{\tau^P_b})^{-1} =&\vec{\omega}(\Sigma^{(1)},\Sigma)^r_{\tau^P_b}(\mathcal{V}^r_0)^{-1}+\vec{\omega}(\Sigma^{(1)},\Sigma)^r_{\tau^P_b}\frac{\mathcal{V}^{\tau^P_b}_0}{\mathcal{V}^r_0\mathcal{V}^r_{\tau^P_b}}\\
=&\vec{\omega}(\Sigma^{(1)},\Sigma)^r_0(\mathcal{V}^r_0)^{-1}-\int_0^{\tau^P_b}\Sigma_t(\int_t^r\Sigma^{(1)}_s\mathrm{d}s)\mathrm{d}t(\mathcal{V}^r_0)^{-1}+\vec{\omega}(\Sigma^{(1)},\Sigma)^r_{\tau^P_b}\frac{\mathcal{V}^{\tau^P_b}_0}{\mathcal{V}^r_0\mathcal{V}^r_{\tau^P_b}}\\
=&\vec{\omega}(\Sigma^{(1)},\Sigma)^r_0(\mathcal{V}^r_0)^{-1}-[\vec{\omega}(\Sigma^{(1)},\Sigma)^{\tau^P_b}_0+\mathcal{V}^{\tau^P_b}_0\vec{\omega}(\Sigma^{(1)})^r_{\tau^P_b}](\mathcal{V}^r_0)^{-1}+\vec{\omega}(\Sigma^{(1)},\Sigma)^r_{\tau^P_b}\frac{\mathcal{V}^{\tau^P_b}_0}{\mathcal{V}^r_0\mathcal{V}^r_{\tau^P_b}}.
\end{aligned}$$

In addition, up to a passing to the limit, an application of Proposition 7.6.1.4 relation (7.68) yields:

$$\mathbb{E}[\mathbb{1}_{\tau^P_b\le r}\mathcal{D}^P(\tau^P_b,r,0)]=\int_0^r\mathcal{J}^P(0,\theta,b-x_0)\mathcal{D}^P(\theta,r,0)\mathrm{d}\theta=\mathcal{D}^P(0,r,b-x_0).$$

The two above mathematical reductions allow us to obtain for the sum $T_4(v^{P,\rho_{s,t}}_{\underline{o},s})+T_5(v^{P,\rho_{s,t}}_{\underline{o},s})$:

$$\begin{aligned}
&(x_0-b)\mathbb{E}[\mathbb{1}_{\tau^P_b\le T}\vec{\omega}(\Sigma^{(1)},\Sigma)^{\tau^P_b}_0(\mathcal{V}^{\tau^P_b}_0)^{-1}\{2\vec{\omega}(\Sigma^{(1)})^T_{\tau^P_b}\mathcal{L}^x_3+\vec{\omega}(\Sigma^{(1)},\Sigma)^T_{\tau^P_b}\mathcal{L}^x_5\}v^{P,h}_{\underline{o},T}(\tau^P_b,b)]\\
&+\int_0^T\Sigma_r(\mathcal{V}^r_0)^{-1}\{2\vec{\omega}(\Sigma^{(1)})^T_r\mathcal{L}^x_3+\vec{\omega}(\Sigma^{(1)},\Sigma)^T_r\mathcal{L}^x_5\}v^{P,h}_{\underline{o},T}(r,b)\mathbb{E}[\mathbb{1}_{\tau^P_b\le r}\{\vec{\omega}(\Sigma^{(1)},\Sigma)^r_{\tau^P_b}\mathcal{V}^{\tau^P_b}_0(\mathcal{V}^r_{\tau^P_b})^{-1}\\
&-\vec{\omega}(\Sigma^{(1)},\Sigma)^{\tau^P_b}_0-\mathcal{V}^{\tau^P_b}_0\vec{\omega}(\Sigma^{(1)})^r_{\tau^P_b}\}\mathcal{D}^P(\tau^P_b,r,0)]\mathrm{d}r
\end{aligned}\tag{7.86}$$

Combine (7.84)-(7.85)-(7.86) and the expressions of $T_n(v^{P,\rho_{s,t}}_{\underline{o},s})$ for $n=1,2,3$ to finish the proof.

▷**Proof of Lemma 7.3.4.5.** Let $0\le u<s<t<T$ and $x>b$. In view of (7.45), we have $\mathcal{L}^x_2 v^{P,\rho_{s,t}}_{\underline{o},s}(u,x)=\sum_{n=1}^{6}\Theta_n(u,s,t,x)$ where:

$$\begin{aligned}
\Theta_1(u,s,t,x)=&\mathcal{L}^x_2\{(x-x_0)[2\mathcal{L}^x_3+(x-x_0)\mathcal{L}^x_4+\mathcal{V}^t_u\mathcal{L}^x_5]\\
&+\mathcal{V}^s_u[3\mathcal{L}^x_4+\frac{1}{2}\mathcal{L}^x_2+(x-x_0)\mathcal{L}^x_5+\mathcal{V}^t_u(\mathcal{L}^x_6+\frac{1}{4}\mathcal{L}^x_4)]\}v^{P,h}_{\underline{o},T}(u,x),
\end{aligned}\tag{7.87}$$

$$\Theta_2(u,s,t,x)=-\mathcal{L}^x_2\{(x-x_0)\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le t}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{u,b},b)|X^P_u=x]\},\tag{7.88}$$

$$\Theta_3(u,s,t,x)=(x_0-b)\mathcal{L}^x_2\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\mathcal{L}^x_5 v^{P,h}_{\underline{o},T}(\tau^P_{u,b},b)|X^P_u=x],\tag{7.89}$$

$$\Theta_4(u,s,t,x)=\mathcal{L}^x_2\{(x-b)\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{u,b},b)|X^P_u=x]\},\tag{7.90}$$

$$\Theta_5(u,s,t,x)=-\mathcal{V}^s_u(\mathcal{L}^x_2\circ\mathcal{L}^x_1)\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\in[s,t]}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{u,b},b)\big|X^P_u=x],\tag{7.91}$$

$$\Theta_6(u,s,t,x)=-\mathcal{L}^x_2\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le s}\mathcal{V}^s_{\tau^P_{u,b}}\int_s^t\frac{\Sigma_r}{\mathcal{V}^r_{\tau^P_{u,b}}}\mathcal{D}^P(\tau^P_{u,b},r,0)(2\mathcal{L}^x_3+\mathcal{V}^t_r\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\big|X^P_u=x],\tag{7.92}$$

With Lemma 7.3.1.3, we easily obtain for (7.87):

$$|\Theta_1(u,s,t,x)|\le_c e^{C_h|x|}\{|x-x_0|^2(\mathcal{V}^T_u)^{-\frac{5}{2}}+|x-x_0|(\mathcal{V}^T_u)^{-2}+(\mathcal{V}^T_u)^{-\frac{3}{2}}\}\tag{7.93}$$

We now pass to (7.88). We have $\Theta_2(u,s,t,x)=\Theta_{2a}(u,s,t,x)+2\Theta_{2b}(u,s,t,x)$ with:

$$\Theta_{2a}(u,s,t,x)=-(x-x_0)\mathcal{L}^x_2\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\le t}(2\mathcal{L}^x_3+\mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}^x_5)v^{P,h}_{\underline{o},T}(\tau^P_{u,b},b)|X^P_u=x],$$

$$\Theta_{2b}(u,s,t,x) = -\mathcal{L}_1^x \mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\leq t}(2\mathcal{L}_3^x + \mathcal{V}^t_{\tau^P_{u,b}}\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(\tau^P_{u,b},b)|X^P_u = x]$$

For $\Theta_{2a}(u,s,t,x)$, following the proof of Lemma 7.3.2.2, split the domain of integration by writing $[u,t] = [u,\frac{u+t}{2}]\cup[\frac{u+t}{2},t]$ to get using Proposition 7.6.1.1 relation (7.62), Lemma 7.3.1.3, an integration by parts, the Cauchy-Schwarz inequality and estimates (7.47)-(7.65):

$$\begin{aligned}
&|\Theta_{2a}(u,s,t,x)|\\
=&|x-x_0|\Big|\int_u^{\frac{u+t}{2}}\partial_r\{\Sigma_r^{-1}\mathcal{J}^P(u,r,x,b)\}(2\mathcal{L}_3^x+\mathcal{V}^t_r\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r+\int_{\frac{u+t}{2}}^t\mathcal{L}_2^x\mathcal{J}^P(u,r,x,b)(2\mathcal{L}_3^x+\mathcal{V}^t_r\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\Big|\\
\leq_c&|x-x_0|\{\Big|\mathcal{J}^P(u,\frac{u+t}{2},x,b)\Sigma^{-1}_{\frac{u+t}{2}}(2\mathcal{L}_3^x+\mathcal{V}^t_{\frac{u+t}{2}}\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(\frac{u+t}{2},b)\Big|\\
&+\Big|\int_u^{\frac{u+t}{2}}\Sigma_r^{-1}\mathcal{J}^P(u,r,x,b)\partial_r\{(2\mathcal{L}_3^x+\mathcal{V}^t_r\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(r,b)\}\mathrm{d}r\Big|+e^{C_h|x|}\int_{\frac{u+t}{2}}^t\Sigma_r(\mathcal{V}^r_u)^{-2}(\mathcal{V}^T_r)^{-1}\mathrm{d}r\}\\
\leq_c&|x-x_0|e^{C_h|x|}\{(\mathcal{V}_u^{\frac{u+t}{2}})^{-1}(\mathcal{V}^T_{\frac{u+t}{2}})^{-1}+\int_u^{\frac{u+t}{2}}\mathcal{J}^P(u,r,x,b)(\mathcal{V}^T_r)^{-2}\mathrm{d}r+(\mathcal{V}_u^{\frac{u+t}{2}})^{-\frac{3}{2}}(\mathcal{V}^T_t)^{-\frac{1}{2}}\}\\
\leq_c&|x-x_0|e^{C_h|x|}\{(\mathcal{V}_u^{\frac{u+t}{2}})^{-1}(\mathcal{V}^T_{\frac{u+t}{2}})^{-1}+(\mathcal{V}^T_{\frac{u+t}{2}})^{-2}+(\mathcal{V}_u^{\frac{u+t}{2}})^{-\frac{3}{2}}(\mathcal{V}^T_t)^{-\frac{1}{2}}\}.
\end{aligned}$$

For $\Theta_{2b}(u,s,t,x)$, use Proposition 7.6.1.1 relation (7.64) to obtain with the same arguments:

$$\begin{aligned}
&|\Theta_{2b}(u,s,t,x)|\\
=&\Big|\int_u^{\frac{u+t}{2}}(2\partial_r+\frac{\Sigma_r}{4}\mathcal{I})\mathcal{D}^P(u,r,x,b)(2\mathcal{L}_3^x+\mathcal{V}^t_r\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r+\int_{\frac{u+t}{2}}^t\mathcal{L}_1^x\mathcal{J}^P(u,r,x,b)(2\mathcal{L}_3^x+\mathcal{V}^t_r\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\Big|\\
\leq_c&\Big|\mathcal{D}^P(u,\frac{u+t}{2},x,b)(2\mathcal{L}_3^x+\mathcal{V}^t_{\frac{u+t}{2}}\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(\frac{u+t}{2},b)\Big|\\
&+\Big|\int_u^{\frac{u+t}{2}}\mathcal{D}^P(u,r,x,b)(\partial_r+\Sigma_r\mathcal{I})\{(2\mathcal{L}_3^x+\mathcal{V}^t_r\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(r,b)\}\mathrm{d}r\Big|+e^{C_h|x|}\int_{\frac{u+t}{2}}^t\Sigma_r(\mathcal{V}^r_u)^{-\frac{3}{2}}(\mathcal{V}^T_r)^{-1}\mathrm{d}r\\
\leq_c&e^{C_h|x|}\{(\mathcal{V}_u^{\frac{u+t}{2}})^{-\frac{1}{2}}(\mathcal{V}^T_{\frac{u+t}{2}})^{-1}+\int_u^{\frac{u+t}{2}}\Sigma_r\mathcal{D}^P(u,r,x,b)(\mathcal{V}^T_r)^{-2}\mathrm{d}r+(\mathcal{V}_u^{\frac{u+t}{2}})^{-1}(\mathcal{V}^T_t)^{-\frac{1}{2}}\}\\
\leq_c&e^{C_h|x|}\{(\mathcal{V}_u^{\frac{u+t}{2}})^{-\frac{1}{2}}(\mathcal{V}^T_{\frac{u+t}{2}})^{-1}+(\mathcal{V}_u^{\frac{u+t}{2}})^{\frac{1}{2}}(\mathcal{V}^T_{\frac{u+t}{2}})^{-2}+(\mathcal{V}_u^{\frac{u+t}{2}})^{-1}(\mathcal{V}^T_t)^{-\frac{1}{2}}\}.
\end{aligned}$$

Thus with the above intermediate results, one deduces:

$$\begin{aligned}
|\Theta_2(u,s,t,x)|\leq_c e^{C_h|x|}\{&|x-x_0|(\mathcal{V}_u^{\frac{u+t}{2}})^{-1}(\mathcal{V}^T_{\frac{u+t}{2}})^{-1}+|x-x_0|(\mathcal{V}^T_{\frac{u+t}{2}})^{-2}+|x-x_0|(\mathcal{V}_u^{\frac{u+t}{2}})^{-\frac{3}{2}}(\mathcal{V}^T_t)^{-\frac{1}{2}}\\
&+(\mathcal{V}_u^{\frac{u+t}{2}})^{-\frac{1}{2}}(\mathcal{V}^T_{\frac{u+t}{2}})^{-1}+(\mathcal{V}_u^{\frac{u+t}{2}})^{\frac{1}{2}}(\mathcal{V}^T_{\frac{u+t}{2}})^{-2}+(\mathcal{V}_u^{\frac{u+t}{2}})^{-1}(\mathcal{V}^T_t)^{-\frac{1}{2}}\}. \qquad (7.94)
\end{aligned}$$

For $\Theta_3(u,s,t,x)$, following the proof of Lemma 7.3.2.2, we easily obtain with an integration by parts:

$$\Theta_3(u,s,t,x) = (x_0-b)\mathbb{E}[\mathbb{1}_{\tau^P_{u,b}\leq s}(2\mathcal{L}_5^x+\mathcal{V}^s_{\tau^P_{u,b}}\mathcal{L}_2^x\circ\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(\tau^P_{u,b},b)|X^P_u = x].$$

Then we write $x_0-b = x-b-(x-x_0)$ and we readily obtain using Lemma 7.3.1.3, the Cauchy-Schwarz inequality and (7.47)-(7.65):

$$\begin{aligned}
&|\Theta_3(u,s,t,x)| \qquad (7.95)\\
\leq_c&(|x-b|+|x-x_0|)\Big|\int_u^{\frac{u+s}{2}}\mathcal{J}^P(u,r,x,b)(2\mathcal{L}_5^x+\mathcal{V}^s_r\mathcal{L}_2^x\circ\mathcal{L}_5^x)v^{P,h}_{\underline{o},T}(r,b)\mathrm{d}r\Big|
\end{aligned}$$

$$
\begin{aligned}
&+(|x-b|+|x-x_0|)\Big|\int_{\frac{u+s}{2}}^{s}\mathcal{J}^P(u,r,x,b)(2\mathcal{L}_5^x+\mathcal{V}_r^s\mathcal{L}_2^x\circ\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(r,b)\mathrm{d}r\Big|\\
\le_c & e^{C_h|x|}\{(\mathcal{V}_{\frac{u+s}{2}}^T)^{-2}\int_u^{\frac{u+s}{2}}[(\mathcal{V}_u^r)^{-\frac12}+|x-x_0|\mathcal{J}^P(u,r,x,b)]\mathrm{d}r+\int_{\frac{u+s}{2}}^{s}\Sigma_r[(\mathcal{V}_u^r)^{-\frac12}+|x-x_0|(\mathcal{V}_u^r)^{-1}](\mathcal{V}_r^T)^{-2}\mathrm{d}r\}\\
\le_c & e^{C_h|x|}\{(\mathcal{V}_{\frac{u+s}{2}}^T)^{-2}(\mathcal{V}_u^{\frac{u+s}{2}})^{\frac12}+(\mathcal{V}_{\frac{u+s}{2}}^T)^{-2}|x-x_0|+(\mathcal{V}_s^T)^{-\frac32}+|x-x_0|(\mathcal{V}_u^{\frac{u+s}{2}})^{-\frac12}(\mathcal{V}_s^T)^{-\frac32}\}.
\end{aligned}
$$

Then we focus on (7.90). We have $\Theta_4(u,s,t,x)=\Theta_{4a}(u,s,t,x)+\Theta_{4b}(u,s,t,x)$ with:

$$
\begin{aligned}
\Theta_{4a}(u,s,t,x)&=(x-b)\mathcal{L}_2^x\mathbb{E}[\mathbb{1}_{\tau_{u,b}^P\le s}(2\mathcal{L}_3^x+\mathcal{V}_{\tau_{u,b}^P}^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(\tau_{u,b}^P,b)|X_u^P=x],\\
\Theta_{4b}(u,s,t,x)&=2\mathcal{L}_1^x\mathbb{E}[\mathbb{1}_{\tau_{u,b}^P\le s}(2\mathcal{L}_3^x+\mathcal{V}_{\tau_{u,b}^P}^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(\tau_{u,b}^P,b)|X_u^P=x].
\end{aligned}
$$

Using integrations by parts, one gets for $\Theta_{4a}(u,s,t,x)$:

$$
\begin{aligned}
\Theta_{4a}(u,s,t,x)=&(x-b)\Sigma_s^{-1}\mathcal{J}^P(u,s,x,b)(2\mathcal{L}_3^xv_{\underline{o},T}^{P,h}(s,b)+\mathcal{V}_s^t\mathcal{L}_5^xv_{\underline{o},T}^{P,h}(s,b))\\
&+(x-b)\mathbb{E}[\mathbb{1}_{\tau_{u,b}^P\le s}(4\mathcal{L}_5^x+\mathcal{V}_{\tau_{u,b}^P}^t\mathcal{L}_2^x\circ\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(\tau_{u,b}^P,b)|X_u^P=x],
\end{aligned}
$$

and thus we easily obtain using similar arguments employed for $\Theta_3(u,s,t,x)$:

$$
|\Theta_{4a}(u,s,t,x)|\le_c e^{C_h|x|}\{(\mathcal{V}_u^{\frac{u+s}{2}})^{\frac12}(\mathcal{V}_{\frac{u+s}{2}}^T)^{-2}+(\mathcal{V}_s^T)^{-\frac32}+(\mathcal{V}_u^s)^{-\frac12}(\mathcal{V}_s^T)^{-1}\}.
$$

Then using notably Proposition 7.6.1.1 relation (7.64), one obtains for $\Theta_{4b}(u,s,t,x)$:

$$
\begin{aligned}
&|\Theta_{4b}(u,s,t,x)|\\
\le_c & \Big|\int_u^{\frac{u+s}{2}}(2\partial_r+\frac14\Sigma_r\mathcal{I})\mathcal{D}^P(u,r,x,b)(2\mathcal{L}_3^x+\mathcal{V}_r^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(r,b)\mathrm{d}r\Big|+\Big|\int_{\frac{u+s}{2}}^{s}\mathcal{L}_1^x\mathcal{J}^P(u,r,x,b)(2\mathcal{L}_3^x+\mathcal{V}_r^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(r,b)\mathrm{d}r\Big|\\
\le_c & \mathcal{D}^P(u,\frac{u+s}{2},x,b)\Big|(2\mathcal{L}_3^x+\mathcal{V}_{\frac{u+s}{2}}^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(\frac{u+s}{2},b)\Big|\\
&+\Big|\int_u^{\frac{u+s}{2}}\mathcal{D}^P(u,r,x,b)(\partial_r+\Sigma_r)\{(2\mathcal{L}_3^x+\mathcal{V}_r^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(r,b)\}\mathrm{d}r\Big|+e^{C_h|x|}\int_{\frac{u+s}{2}}^{s}\Sigma_r(\mathcal{V}_u^r)^{-\frac32}(\mathcal{V}_r^T)^{-1}\mathrm{d}r\\
\le_c & e^{C_h|x|}\{(\mathcal{V}_u^{\frac{u+s}{2}})^{-\frac12}(\mathcal{V}_{\frac{u+s}{2}}^T)^{-1}+\int_u^{\frac{u+s}{2}}\Sigma_r\mathcal{D}^P(u,r,x,b)(\mathcal{V}_r^T)^{-2}\mathrm{d}r+(\mathcal{V}_u^{\frac{u+s}{2}})^{-\frac12}(\mathcal{V}_s^T)^{-1}\}\\
\le_c & e^{C_h|x|}\{(\mathcal{V}_u^{\frac{u+s}{2}})^{-\frac12}(\mathcal{V}_{\frac{u+s}{2}}^T)^{-1}+(\mathcal{V}_u^{\frac{u+s}{2}})^{\frac12}(\mathcal{V}_{\frac{u+s}{2}}^T)^{-2}+(\mathcal{V}_u^{\frac{u+s}{2}})^{-\frac12}(\mathcal{V}_s^T)^{-1}\}.
\end{aligned}
$$

Combining the estimates for $\Theta_{4a}(u,s,t,x)$ and $\Theta_{4b}(u,s,t,x)$ leads to:

$$
\Theta_4(u,s,t,x)\le_c e^{C_h|x|}\{(\mathcal{V}_u^{\frac{u+s}{2}})^{-\frac12}(\mathcal{V}_s^T)^{-1}+(\mathcal{V}_u^{\frac{u+s}{2}})^{\frac12}(\mathcal{V}_{\frac{u+s}{2}}^T)^{-2}+(\mathcal{V}_s^T)^{-\frac32}\}\tag{7.96}
$$

Next we treat (7.91). Again the same tools previously employed yield:

$$
\begin{aligned}
&|\Theta_5(u,s,t,x)| &(7.97)\\
=&\mathcal{V}_u^s\Big|\int_s^{\frac{s+t}{2}}(\mathcal{L}_2^x\circ\mathcal{L}_1^x)\mathcal{J}^P(u,r,x,b)(2\mathcal{L}_3^x+\mathcal{V}_r^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(r,b)\mathrm{d}r\\
&+\int_{\frac{s+t}{2}}^{t}(\mathcal{L}_2^x\circ\mathcal{L}_1^x)\mathcal{J}^P(u,r,x,b)(2\mathcal{L}_3^x+\mathcal{V}_r^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(r,b)\mathrm{d}r\Big|\\
\le_c&\mathcal{V}_u^s\{\Sigma_s^{-1}\Big|\mathcal{L}_1^x\mathcal{J}^P(u,s,x,b)(2\mathcal{L}_3^x+\mathcal{V}_s^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(s,b)\Big|+\Sigma_{\frac{s+t}{2}}^{-1}\Big|\mathcal{L}_1^x\mathcal{J}^P(u,\frac{s+t}{2},x,b)(2\mathcal{L}_3^x+\mathcal{V}_{\frac{s+t}{2}}^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(\frac{s+t}{2},b)\Big|\}
\end{aligned}
$$

$$+\mathcal{V}_u^s\Big|\int_s^{\frac{s+t}{2}}\Sigma_r^{-1}\mathcal{L}_1^x\mathcal{J}^P(u,r,x,b)\partial_r\{(2\mathcal{L}_3^x+\mathcal{V}_r^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(r,b)\}\mathrm{d}r\Big|+e^{C_h|x|}\mathcal{V}_u^s\int_{\frac{s+t}{2}}^t\Sigma_r(\mathcal{V}_u^r)^{-\frac{5}{2}}(\mathcal{V}_r^T)^{-1}\mathrm{d}r$$
$$\leq_c e^{C_h|x|}\{\mathcal{V}_u^s(\mathcal{V}_u^s)^{-\frac{3}{2}}(\mathcal{V}_s^T)^{-1}+\mathcal{V}_u^s(\mathcal{V}_u^{\frac{s+t}{2}})^{-\frac{3}{2}}(\mathcal{V}_{\frac{s+t}{2}}^T)^{-1}+\mathcal{V}_u^s\int_s^{\frac{s+t}{2}}\Sigma_r(\mathcal{V}_u^r)^{-\frac{3}{2}}(\mathcal{V}_r^T)^{-2}\mathrm{d}r+\mathcal{V}_u^s(\mathcal{V}_u^{\frac{s+t}{2}})^{-2}(\mathcal{V}_t^T)^{-\frac{1}{2}}\}$$
$$\leq_c e^{C_h|x|}\{(\mathcal{V}_u^s)^{-\frac{1}{2}}(\mathcal{V}_{\frac{s+t}{2}}^T)^{-1}+(\mathcal{V}_{\frac{s+t}{2}}^T)^{-\frac{3}{2}}+(\mathcal{V}_u^{\frac{s+t}{2}})^{-1}(\mathcal{V}_t^T)^{-\frac{1}{2}}\}.$$

We finally estimate $\Theta_6(u,s,t,x)$ defined in (7.92). Perform an integration by parts and use Hölder inequalities to obtain :

$$\begin{aligned}
&|\Theta_6(u,s,t,x)|\\
=&\Big|\int_s^t\Sigma_r(2\mathcal{L}_3^x+\mathcal{V}_r^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(r,b)\{\int_u^s\mathcal{L}_2^x\mathcal{J}^P(u,l,x,b)\frac{\mathcal{V}_l^s}{\mathcal{V}_l^r}\mathcal{D}^P(l,r,0)\mathrm{d}l\}\mathrm{d}r\Big|\\
\leq_c&\Big|\int_s^t\Sigma_r(2\mathcal{L}_3^x+\mathcal{V}_r^t\mathcal{L}_5^x)v_{\underline{o},T}^{P,h}(r,b)\{\int_u^s\mathcal{J}^P(u,l,x,b)(\Sigma_l)^{-1}\partial_l\{\frac{\mathcal{V}_l^s}{\mathcal{V}_l^r}\mathcal{D}^P(l,r,0)\}\mathrm{d}l\}\mathrm{d}r\Big|\\
\leq_c&e^{C_h|x|}\int_s^t\Sigma_r(\mathcal{V}_r^T)^{-1}\{\int_u^s\mathcal{J}^P(u,l,x,b)(\mathcal{V}_l^r)^{-\frac{3}{2}}\mathrm{d}l\}\mathrm{d}r\\
\leq_c&e^{C_h|x|}\{\int_s^t\Sigma_r(\mathcal{V}_r^T)^{-1}(\mathcal{V}_{\frac{u+s}{2}}^r)^{-\frac{3}{2}}\mathrm{d}r+\int_s^t\Sigma_r(\mathcal{V}_r^T)^{-1}\{\int_{\frac{u+s}{2}}^s\Sigma_l(\mathcal{V}_u^l)^{-1}(\mathcal{V}_l^r)^{-\frac{3}{2}}\mathrm{d}l\}\mathrm{d}r\}\\
\leq_c&e^{C_h|x|}\{(\int_s^t\Sigma_r(\mathcal{V}_r^T)^{-4}\mathrm{d}r)^{\frac{1}{4}}(\int_s^t\Sigma_r(\mathcal{V}_{\frac{u+s}{2}}^r)^{-2}\mathrm{d}r)^{\frac{3}{4}}+\int_s^t\Sigma_r(\mathcal{V}_r^T)^{-1}(\int_{\frac{u+s}{2}}^s\Sigma_l(\mathcal{V}_u^l)^{-4}\mathrm{d}l)^{\frac{1}{4}}(\int_{\frac{u+s}{2}}^s\Sigma_l(\mathcal{V}_l^r)^{-2}\mathrm{d}l)^{\frac{3}{4}}\mathrm{d}r\}\\
\leq_c&e^{C_h|x|}\{(\mathcal{V}_t^T)^{-\frac{3}{4}}(\mathcal{V}_{\frac{u+s}{2}}^s)^{-\frac{3}{4}}+(\mathcal{V}_u^{\frac{u+s}{2}})^{-\frac{3}{4}}\int_s^t\Sigma_r(\mathcal{V}_r^T)^{-1}(\mathcal{V}_s^r)^{-\frac{3}{4}}\mathrm{d}r\}\\
\leq_c&e^{C_h|x|}\{(\mathcal{V}_t^T)^{-\frac{3}{4}}(\mathcal{V}_{\frac{u+s}{2}}^s)^{-\frac{3}{4}}+(\mathcal{V}_u^{\frac{u+s}{2}})^{-\frac{3}{4}}(\int_s^t\Sigma_r(\mathcal{V}_r^T)^{-5}\mathrm{d}r)^{\frac{1}{5}}(\int_s^t\Sigma_r(\mathcal{V}_s^r)^{-\frac{15}{16}}\mathrm{d}r)^{\frac{4}{5}}\mathrm{d}r\}\\
\leq_c&e^{C_h|x|}\{(\mathcal{V}_t^T)^{-\frac{3}{4}}(\mathcal{V}_{\frac{u+s}{2}}^s)^{-\frac{3}{4}}+(\mathcal{V}_u^{\frac{u+s}{2}})^{-\frac{3}{4}}(\mathcal{V}_t^T)^{-\frac{4}{5}}(\mathcal{V}_s^t)^{\frac{1}{20}}\}.
\end{aligned}\tag{7.98}$$

We conclude combining the estimates (7.93)-(7.94)-(7.95)-(7.96)-(7.97)-(7.98) which remain valid if x tends to b with a passing to the limit.

7.6.3 Proof of Propositions 7.5.2.1-7.5.3.1.

▷**Proof of Proposition 7.5.2.1.** We skip the proof for the leading term $\mathrm{DoCall}^{\mathrm{BS}}(x_0,\mathcal{V}_0^T,k,b)$ and for the corrective terms $Cor_{1,\underline{o}}$ and $Cor_{3,\underline{o}}$, the proof being very similar to the proof of Lemmas 2.1.3.1 and 2.1.3.1 and Theorem 2.3.2.1 of Chapter 2. For instance for the part $\mathrm{Call}^{\mathrm{BS}}(b,y,x_0+k-b)$ in $\mathrm{DoCall}^{\mathrm{BS}}(x_0,y,k,b)$, we have $|x_0-k|\leq x_0+k-2b$ for $b\leq\min(x_0,k)$ and an application of Corollary 2.6.1.1 of Chapter 2 allows to easily estimate the residuals in the expansions.
We now prove the expansion for $Cor_{2,\underline{o}}$. Let $y>0$. By Lemma 7.2.1.1, we have $\mathcal{L}_3^x\mathrm{DoCall}^{\mathrm{BS}}(x,y,k,b)|_{x=b}=2\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,y,k)$ and by Proposition 2.6.1.3 of Chapter 2, we have $\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,y,k)=e^b\frac{(k-b)}{y^{\frac{3}{2}}}\mathcal{N}'(\frac{b-k+\frac{1}{2}y}{\sqrt{y}})$. Consequently in view of the density of the hitting times (see (7.6)), we can write $\mathcal{J}^{P,x}(0,\tau_b^{P,x},b-x_0)=e^{-b}\Sigma_t(x)\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(x,\mathcal{V}_0^{\tau_b^{P,x}}(x),x_0)|_{x=b}$ to get:

$$\begin{aligned}
Cor_{2,\underline{o}}=&-\mathbb{E}[\mathbb{1}_{\tau_b^P<T}\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^P}^T\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_{\tau_b^P}^T,k,b)]\\
=&-e^{-b}\int_0^T\Sigma_t\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_0^t,x_0)\vec{\omega}(\Sigma^{(1)},\Sigma)_t^T\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_t^T,k)\mathrm{d}t.
\end{aligned}$$

Perform a Taylor expansion of $\Lambda_t : z \mapsto -e^{-b}\Sigma_t(z)\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b, \mathcal{V}_0^t(z), x_0)\vec{\omega}(\Sigma^{(1)}, \Sigma)_t^T(z)\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b, \mathcal{V}_t^T(z), k)$ at $z = x_0$ around $z = x_{avg}$ to obtain with (7.56):

$$
\begin{aligned}
&\mathbb{E}[\mathbb{1}_{\tau_b^P<T}\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^P}^T\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_{\tau_b^P}^T,k)]\\
=&\mathbb{E}[\mathbb{1}_{\tau_b^{P,x_{avg}}<T}\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^{P,x_{avg}}}^T(x_{avg})\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_{\tau_b^{P,x_{avg}}}^T(x_{avg}),k,b)]\\
&+\frac{1}{2}(x_0-k)\mathbb{E}[\mathbb{1}_{\tau_b^{P,x_{avg}}<T}\{[\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)})_{\tau_b^{P,x_{avg}}}^T+\vec{\omega}(\Sigma^{(2)},\Sigma)_{\tau_b^{P,x_{avg}}}^T](x_{avg})\mathcal{L}_3^x\\
&+[\vec{\omega}(\Sigma^{(1)},\Sigma^{(1)},\Sigma)_{\tau_b^{P,x_{avg}}}^T+\frac{1}{2}\vec{\omega}(\Sigma^{(1)},\Sigma,\Sigma^{(1)})_{\tau_b^{P,x_{avg}}}^T](x_{avg})\mathcal{L}_5^x\}\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_{\tau_b^{P,x_{avg}}}^T(x_{avg}),k)]\\
&+\frac{e^{-b}}{2}(x_0-k)\int_0^T\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_0^t(x_{avg}),x_0)[\Sigma_t^{(1)}\vec{\omega}(\Sigma^{(1)},\Sigma)_t^T](x_{avg})\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_t^T(x_{avg}),k)\mathrm{d}t\\
&+\frac{e^{-b}}{2}(x_0-k)\int_0^T\partial_y\{\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(x,y,x_0)|_{x=b}\}|_{y=\mathcal{V}_0^t(x_{avg})}[\Sigma_t\vec{\omega}(\Sigma^{(1)})_0^t\vec{\omega}(\Sigma^{(1)},\Sigma)_t^T](x_{avg})\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_t^T(x_{avg}),k)\mathrm{d}t\\
&+R,
\end{aligned}
$$

where $R = \frac{1}{4}(x_0-k)^2\int_0^T\{\int_0^1(1-\alpha)\Lambda_t^{(2)}(\alpha x_{avg}+(1-\alpha)x_0)\mathrm{d}\alpha\}\mathrm{d}t$. Using the decomposition $(x_0-k) = x_0-b+b-k$, the hypotheses $(\widetilde{\mathcal{H}}_{x_0}^{\sigma})$-$(\widetilde{\mathcal{H}}_{x_{avg}}^{\sigma})$, standard inequalities and Corollary 2.6.1.1 of Chapter 2, one obtains $R = O(|\sigma|_\infty\mathcal{M}_1(\sigma)[\mathcal{M}_0(\sigma)]^2T^2)$. Besides we easily show that:

$$
\begin{aligned}
&\mathbb{E}[\mathbb{1}_{\tau_b^P<T}\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^P}^T\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_{\tau_b^P}^T,k)]\\
=&\mathbb{E}[\mathbb{1}_{\tau_b^{P,x_{avg}}<T}\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^{P,x_{avg}}}^T(x_{avg})\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_{\tau_b^{P,x_{avg}}}^T(x_{avg}),k)]+O(|\sigma|_\infty\mathcal{M}_1(\sigma)[\mathcal{M}_0(\sigma)]^2T^{\frac{3}{2}}).
\end{aligned}
$$

Then write $\Sigma_t(x_{avg})\partial_y\{\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(x,y,x_0)|_{x=b}\}|_{y=\mathcal{V}_0^t(x_{avg})} = \partial_t\{\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(x,\mathcal{V}_0^t(x_{avg}),x_0)|_{x=b}\}$ and perform an integration by parts to get with (7.56) and Lemma 7.2.1.1:

$$
\begin{aligned}
&e^{-b}\int_0^T\partial_y\{\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(x,y,x_0)|_{x=b}\}|_{y=\mathcal{V}_0^t(x_{avg})}[\Sigma_t\vec{\omega}(\Sigma^{(1)})_0^t\vec{\omega}(\Sigma^{(1)},\Sigma)_t^T](x_{avg})\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_t^T(x_{avg}),k)\mathrm{d}t\\
=&-e^{-b}\int_0^T\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_0^t(x_{avg}),x_0)[\Sigma_t^{(1)}\vec{\omega}(\Sigma^{(1)},\Sigma)_t^T](x_{avg})\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_t^T(x_{avg}),k)\mathrm{d}t\\
&+\mathbb{E}[\mathbb{1}_{\tau_b^{P,x_{avg}}<T}\vec{\omega}(\Sigma^{(1)})_0^{\tau_b^{P,x_{avg}}}(x_{avg})\{\vec{\omega}(\Sigma^{(1)})_{\tau_b^{P,x_{avg}}}^T(x_{avg})\mathcal{L}_3^x+\frac{1}{2}\vec{\omega}(\Sigma^{(1)},\Sigma)_{\tau_b^{P,x_{avg}}}^T(x_{avg})\mathcal{L}_5^x\}\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_{\tau_b^{P,x_{avg}}}^T(x_{avg}),k)].
\end{aligned}
$$

Combine the intermediate results to achieve the proof. Although tedious to write, the proofs for $Cor_{3,\underline{\varrho}}$, $Cor_{4,\underline{\varrho}}$ and $Cor_{5,\underline{\varrho}}$ are very similar and do not contain huge mathematical difficulty so we leave them as an exercise to the reader.

▷**Proof of Proposition 7.5.3.1.** We begin with $Cor_{2,\underline{\varrho}}(z)$. First using Proposition 2.6.1.3 of Chapter 2, we can write $\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(x,\mathcal{V}_{\tau_b^{P,z}}^T(z),k)|_{x=b} = \frac{e^b}{\Sigma(z)}\mathcal{J}^{P,z}(\tau_b^{P,z},T,b-k)$ and then by definition (see Theorem 7.2.2.1), one has $Cor_{2,\underline{\varrho}}(z) = -\frac{1}{2}\frac{e^b}{\Sigma(z)}\Sigma(z)\Sigma^{(1)}(z)\int_0^T\mathcal{J}^{P,z}(0,t,b-x_0)(T-t)^2\mathcal{J}^{P,z}(t,T,b-k)\mathrm{d}t$. One concludes with Proposition 7.6.1.4 Relation (7.73).

We now pass to $Cor_{4,\underline{\varrho}}(z)$. By definition we have (see Theorem 7.2.3.1):

$$
Cor_{4,\underline{\varrho}}(z) = (x_0-b)[\frac{1}{2}Cor_{4a,\underline{\varrho}}(z)+\frac{1}{8}Cor_{4b,\underline{\varrho}}(z)], \tag{7.99}
$$

where:

$$
Cor_{4a,\underline{\varrho}}(z) = \mathbb{E}[\mathbb{1}_{\tau_b^{P,z}\le T}\{\Sigma(z)\Sigma^{(2)}(z)(T-\tau_b^{P,z})^2+(\Sigma^{(1)})^2(z)(T-\tau_b^{P,z})T\}\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_{\tau_b^{P,z}}^T(z),k)],
$$

$$Cor_{4b,\underline{\varrho}}(z)=\mathbb{E}[\mathbb{1}_{\tau_b^{P,z}\le T}\Sigma(z)(\Sigma^{(1)})^2(z)\{(T-\tau_b^{P,z})^3+(T-\tau_b^{P,z})^2T\}\mathcal{L}_5^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}^T_{\tau_b^{P,z}}(z),k)].$$

Using Proposition 7.6.1.4 relations (7.73) and (7.71), we easily get:

$$\begin{aligned}Cor_{4a,\underline{\varrho}}(z)=&e^b(k-b)\frac{\Sigma^{(2)}(z)}{\Sigma(z)}\{\mathcal{V}_0^T(z)\mathcal{D}^{P,z}(0,T,2b-x_0-k)-(x_0-b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}}e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}})\}\\&+e^b(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\mathcal{V}_0^T(z)\mathcal{D}^{P,z}(0,T,2b-x_0-k).\end{aligned}\tag{7.100}$$

Then a straightforward calculus yields:

$$\begin{aligned}\mathcal{L}_5^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}^T_{\tau_b^{P,z}}(z)),k)=&(\mathcal{L}_2^x\circ\mathcal{L}_3^x)\mathrm{Call}^{\mathrm{BS}}(x,\mathcal{V}^T_{\tau_b^{P,z}}(z),k)|_{x=b}\\=&e^b\{\frac{(k-b)^3}{(\mathcal{V}^T_{\tau_b^{P,z}}(z))^3}-3\frac{(k-b)}{(\mathcal{V}^T_{\tau_b^{P,z}}(z))^2}-\frac{1}{4}\frac{(k-b)}{\mathcal{V}^T_{\tau_b^{P,z}}(z)}\}\mathcal{D}^{P,z}(\tau_b^{P,z},T,b-k).\end{aligned}\tag{7.101}$$

Thus it comes with Proposition 7.6.1.4 relations (7.67)-(7.68)-(7.72)-(7.74):

$$\begin{aligned}&\mathbb{E}[\mathbb{1}_{\tau_b^{P,z}\le T}\Sigma(z)(\Sigma^{(1)})^2(z)\{(T-\tau_b^{P,z})^3+(T-\tau_b^{P,z})^2T\}\frac{(k-b)^3}{(\mathcal{V}^T_{\tau_b^{P,z}})^3(z)}\mathcal{D}^{P,z}(\tau_b^{P,z},T,b-k)]\\=&(k-b)^2\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\int_0^T\mathcal{J}^{P,z}(0,t,b-x_0)\{(k-b)\mathcal{D}^{P,z}(t,T,b-k)+T\mathcal{J}^{P,z}(t,T,b-k)\}\mathrm{d}t\\=&(k-b)^2\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\{(k-b)\mathcal{D}^{P,z}(0,T,2b-x_0-k)+T\mathcal{J}^{P,z}(0,T,2b-x_0-k)\}\\=&(k-b)^2\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\mathcal{D}^{P,z}(0,T,2b-x_0-k)[2(k-b)+(x_0-b)];\\&\mathbb{E}[\mathbb{1}_{\tau_b^{P,z}\le T}\Sigma(z)(\Sigma^{(1)})^2(z)\{(T-\tau_b^{P,z})^3+(T-\tau_b^{P,z})^2T\}\frac{(k-b)}{(\mathcal{V}^T_{\tau_b^{P,z}})^2(z)}\mathcal{D}^{P,z}(\tau_b^{P,z},T,b-k)]\\=&(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\int_0^T\mathcal{J}^{P,z}(0,t,b-x_0)\{\mathcal{V}_t^T(z)+\mathcal{V}_0^T(z)\}\mathcal{D}^{P,z}(t,T,b-k)\mathrm{d}t\\=&(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\{2\mathcal{V}_0^T(z)\mathcal{D}^{P,z}(0,T,2b-x_0-k)-(x_0-b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}\};\\&\mathbb{E}[\mathbb{1}_{\tau_b^{P,z}\le T}\Sigma(z)(\Sigma^{(1)})^2(z)\{(T-\tau_b^{P,z})^3+(T-\tau_b^{P,z})^2T\}\frac{(k-b)}{\mathcal{V}^T_{\tau_b^{P,z}}(z)}\mathcal{D}^{P,z}(\tau_b^{P,z},T,b-k)]\\=&(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\int_0^T\mathcal{J}^{P,z}(0,t,b-x_0)\mathcal{V}_t^T(z)\{2\mathcal{V}_0^T(z)-\mathcal{V}_0^t(z)\}\mathcal{D}^{P,z}(t,T,b-k)\mathrm{d}t\\=&(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\int_0^T\{2\mathcal{V}_0^T(z)\mathcal{J}^{P,z}(0,t,b-x_0)-\Sigma(z)(x_0-b)\mathcal{D}^{P,z}(0,t,b-x_0)\}\mathcal{V}_t^T(z)\mathcal{D}^{P,z}(t,T,b-k)\mathrm{d}t\\=&(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\{\mathcal{V}_0^T(z)\mathcal{D}^{P,z}(0,T,2b-x_0-k)[\frac{1}{2}(x_0-b)(x_0-k)+2\mathcal{V}_0^T(z)]\\&+\frac{1}{2}(x_0-b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}[(x_0-k)(2b-x_0-k)-5\mathcal{V}_0^T(z)]\}.\end{aligned}$$

Combining the intermediate results gives with (7.101):

$$Cor_{4b,\underline{\varrho}}(z)$$

$$=e^{b}(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\{(x_0-b)[3+\frac{5}{8}\mathcal{V}_0^T(z)+\frac{1}{8}(x_0-k)(x_0+k-2b)]\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}$$
$$+[2(k-b)^2+(k-b)(x_0-b)-\mathcal{V}_0^T(z)(6+\frac{1}{8}(x_0-b)(x_0-k)+\frac{1}{2}\mathcal{V}_0^T(z))]\mathcal{D}^{P,z}(0,T,2b-x_0-k)\},$$

what leads to the announced result with (7.99)-(7.100).
Next we treat $Cor_{5,\underline{\varrho}}(z)$. We have by definition (see Theorem 7.2.3.1):

$$Cor_{5,\underline{\varrho}}(z)=-\frac{1}{2}Cor_{5a,\underline{\varrho}}(z)-\frac{1}{8}Cor_{5b,\underline{\varrho}}(z), \tag{7.102}$$

where:

$$Cor_{5a,\underline{\varrho}}(z)=\int_0^T\Sigma(z)(\Sigma^{(1)})^2(z)(T-r)\mathcal{L}_3^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_r^T(z),k)\mathbb{E}[\mathbb{1}_{\tau_b^{P,z}\le r}\tau_b^{P,z}\mathcal{D}^{P,z}(\tau_b^{P,z},r,0)]\mathrm{d}r,$$
$$Cor_{5b,\underline{\varrho}}(z)=\int_0^T\Sigma^2(z)(\Sigma^{(1)})^2(z)(T-r)^2\mathcal{L}_5^x\mathrm{Call}^{\mathrm{BS}}(b,\mathcal{V}_r^T(z),k)\mathbb{E}[\mathbb{1}_{\tau_b^{P,z}\le r}\tau_b^{P,z}\mathcal{D}^{P,z}(\tau_b^{P,z},r,0)]\mathrm{d}r.$$

Then, applying the Proposition 7.6.1.1 relation (7.69), it comes up to a passing to the limit:

$$\Sigma(z)\mathbb{E}[\mathbb{1}_{\tau_b^{P,z}\le r}\tau_b^{P,z}\mathcal{D}^{P,z}(\tau_b^{P,z},r,0)]=(x_0-b)\Sigma(z)\int_0^r\mathcal{D}^{P,z}(0,\theta,b-x_0)\mathcal{D}^{P,z}(\theta,r,0)\mathrm{d}\theta$$
$$=(x_0-b)\mathcal{N}(\frac{b-x_0}{\sqrt{\mathcal{V}_0^r(z)}})e^{\frac{x_0-b}{2}-\frac{1}{8}\mathcal{V}_0^r(z)}.$$

Thus we easily get on the one hand using relation (7.70):

$$Cor_{5a,\underline{\varrho}}(z)$$
$$=e^{b}(x_0-b)(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\Sigma(z)\int_0^T\mathcal{N}(\frac{b-x_0}{\sqrt{\mathcal{V}_0^r(z)}})e^{\frac{x_0-b}{2}-\frac{1}{8}\mathcal{V}_0^r(z)}\mathcal{D}^{P,z}(r,T,b-k)\mathrm{d}r$$
$$=e^{b}(x_0-b)(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\{\mathcal{V}_0^T(z)\mathcal{D}^{P,z}(0,T,2b-x_0-k)-(x_0+k-2b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}\},$$

and on the other hand using relations (7.69)-(7.70)-(7.75) and equation (7.101):

$$Cor_{5b,\underline{\varrho}}(z)$$
$$=e^{b}(x_0-b)(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\int_0^T\mathcal{N}(\frac{b-x_0}{\sqrt{\mathcal{V}_0^r(z)}})e^{\frac{x_0-b}{2}-\frac{1}{8}\mathcal{V}_0^r(z)}$$
$$\times\{(k-b)\mathcal{J}^{P,z}(r,T,b-k)-3\Sigma(z)\mathcal{D}^{P,z}(r,T,b-k)-\frac{1}{4}\Sigma(z)\mathcal{V}_r^T(z)\mathcal{D}^{P,z}(r,T,b-k)\}\mathrm{d}r$$
$$=e^{b}(x_0-b)(k-b)\frac{(\Sigma^{(1)})^2(z)}{\Sigma^2(z)}\{(k-b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}-3\mathcal{V}_0^T(z)\mathcal{D}^{P,z}(0,T,2b-x_0-k)$$
$$+3(x_0+k-2b)\mathcal{N}(\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}})e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}-\frac{1}{8}(\mathcal{V}_0^T(z))^{\frac{3}{2}}(-\frac{1}{6}H_3(x)\mathcal{N}(x)-\frac{1}{6}H_2(x)\mathcal{N}'(x)+\frac{1}{2}\mathcal{N}'(x)$$
$$-\frac{(x_0-k)}{2\sqrt{\mathcal{V}_0^T(z)}}[(H_2(x)+2)\mathcal{N}(x)+x\mathcal{N}'(x)])\Big|_{x=\frac{2b-x_0-k}{\sqrt{\mathcal{V}_0^T(z)}}}e^{\frac{x_0+k}{2}-b-\frac{\mathcal{V}_0^T(z)}{8}}\}.$$

That achieves the proof. We leave the proof for $Cor_{9,\underline{\varrho}}(z)$ defined in Proposition 7.5.2.1 to the reader which is very similar to that of $Cor_{4,\underline{\varrho}}(z)$.

Part IV

Efficient weak approximations in multidimensional diffusions

Chapter 8

Stochastic Approximation Finite Element method: analytical formulas for multidimensional diffusion process

Submitted in *SIAM Journal of Numerical Analysis*.

Contents

We derive an analytical weak approximation of a multidimensional diffusion as coefficients or time are small. Our methodology combines the use of Gaussian proxys to approximate the law of the diffusion and a Finite Element interpolation of the terminal function applied to the diffusion. We call this method Stochastic Approximation Finite Element (SAFE for short) method. We provide error bounds of our global approximation depending on the diffusion process coefficients, the time horizon and the regularity of the terminal function. Then we give estimates of the computational cost of our algorithm. This shows an improved efficiency compared to Monte-Carlo methods in small and medium dimensions (up to 10), which is confirmed by numerical experiments. For high dimensions (greater than 10) we can perform

Monte-Carlo simulations on the proxy and this shows a speed gain by a factor 100 in comparison to Monte-Carlo methods applied on the diffusion owing to the exact simulation without discretization of the proxy.

8.1 Introduction

Motivation and contribution of the Chapter. We consider for $d \geq 1$ a d-dimensional stochastic differential equation (SDE) defined by:

$$X_t = x_0 + \sum_{j=1}^{q} \int_0^t \sigma_j(s, X_s) \mathrm{d}W_s^j + \int_0^t b(s, X_s) \mathrm{d}s,$$

where $(W_t)_{t\geq 0}$ is a standard Brownian motion in $\mathbb{R}^q$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with the usual assumptions on the filtration $(\mathcal{F}_t)_{t\geq 0}$. Here, σ is a $d \times q$ matrix and b is a d-dimensional vector, their entries being regular and bounded functions. We are interested in deriving analytical approximations of

$$\mathbb{E}[h(X_T)] \tag{8.1}$$

for a given function h, at least Lipschitz continuous, and a fixed time horizon $T > 0$. The explicit calculus of (8.1) is most of the time impossible because the marginal law of the diffusion X is not known and because of the general form of the function h. Hence it is usual to perform a numerical method. For low dimension (say $d \leq 3$), we may use PDE schemes since $(x_0, T) \mapsto \mathbb{E}[h(X_T)]$ solves a linear parabolic PDE but the complexity is increasing very quickly with the dimension d. For higher dimension, Monte-Carlo methods are preferred, but although almost insensitive to the dimension, they only evaluate the above expectation for a single (x_0, T). The aim of this work is to provide an alternative numerical method, based on analytical approximation, and we highlight an approach suiting well to general functions h without specific form (under reasonable conditions) and to rather general diffusion models. The quick and efficient approximation of SDE distributions is fundamental, it is widely used as a cornerstone of probabilistic algorithms related to the dynamic programming problems which necessitate the evaluation of many nested conditional expectations (for instance, see [Bally 2003] for optimal stopping problems and [Lemor 2006] for Backward SDEs).
Our subsequent numerical method (called Stochastic Approximation Finite Element, SAFE for short) relies both on the weak approximation of the marginal law of X_T and on the approximation of the function h. Firstly, to approximate the law of X, we consider the Gaussian proxy process obtained by freezing at $x = x_0$ the diffusion coefficients:

$$X_t^P = x_0 + \sum_{j=1}^{q} \int_0^t \sigma_j(s, x_0) \mathrm{d}W_s^j + \int_0^t b(s, x_0) \mathrm{d}s. \tag{8.2}$$

Using the Proxy principle of [Gobet 2012a] , we derive a weak approximation in the form (see Theorem 8.2.1.1):

$$\mathbb{E}[h(X_T)] \approx \mathbb{E}[h(X_T^P)] + \sum_{|\alpha|\leq 3} w_{\alpha,T} \partial^{|\alpha|}_{\epsilon_{\alpha_1}\dots\epsilon_{\alpha_{|\alpha|}}} (\mathbb{E}[h(X_T^P + \epsilon)])\Big|_{\epsilon=0}, \tag{8.3}$$

where $\alpha \in \{1, \dots, d\}^{|\alpha|}$ is a multi-index, $w_{\alpha,T}$ are weights depending explicitly on the SDE coefficients and where the sensitivities $\partial^{|\alpha|}_{\epsilon_{\alpha_1}\dots\epsilon_{\alpha_{|\alpha|}}} (\mathbb{E}[h(X_T^P + \epsilon)])\Big|_{\epsilon=0}$ are well defined as soon as the law of X_T^P is non degenerate. Apart from few specific cases of functions h (for example if h has separable variables combined with the independence of the X_T^P components), the representation (8.3) can not be directly computed in closed forms: however, it can be rewritten in a simple expectation form suitable for simple and direct Monte-Carlo simulations (see Theorem 8.2.1.2). To obtain fully analytical formulas, another ingredient is needed. The second step is to approximate the function h by a local interpolation based on suitable shape functions of Finite Element Methods (see Theorems 8.2.2.1-8.2.2.2-8.2.4.1). Denoting by $\hat{h}$ the resulting interpolation of h, the final structure of approximation becomes

$$\mathbb{E}[h(X_T)] \approx \mathbb{E}[\hat{h}(X_T^P)] + \sum_{|\alpha|\leq 3} w_{\alpha,T}\partial^{|\alpha|}_{\epsilon_{\alpha_1}\cdots\epsilon_{\alpha_{|\alpha|}}}(\mathbb{E}[\hat{h}(X_T^P+\epsilon)])\Big|_{\epsilon=0},$$

which accuracy and complexity are given in Theorems 8.2.3.1-8.2.4.1 and Corollaries 8.2.3.1-8.2.4.1. The convergence holds as b, σ or T go to 0 in a suitable sense. The key feature in this methodology is that the interpolation procedure is done in such a way that the calculus of the above expectations is fully explicit and reduces to computations involving the c.d.f. of a one-dimensional Gaussian r.v. and its derivatives (see Subsection 8.2.2). The flexibility and the accuracy of our formulas allow their use as it stands or alternatively it could serve as a control variates tool to improve Monte-Carlo methods.

Background results. We briefly describe the main known approaches to approximate the distribution of a SDE. Time discretization schemes are broadly described in [Kloeden 2010]: they consist in replacing X by an approximation $\hat{X}$ easier to simulate, the evaluation of $\mathbb{E}(h(\hat{X}_T))$ is then made using Monte-Carlo simulations. The balance between discretization and integration errors is described in [Duffie 1995].

Alternatively, the cubature on Wiener space by Kusuoka-Lyons-Victoir [Kusuoka 2004, Lyons 2004] is a well-established theory. It is based on a smart discrete approximation of the Wiener measure, which leads to solving ODEs in order to approximate X. The splitting method by Ninomiya-Victoir [Ninomiya 2008] also reduces to solving ODEs. Clearly, these approaches are different from ours.

The quantization method [Graf 2000] is aimed at approximating the distribution of X_T with a fixed number of points, optimally w.r.t. a L^p-norm; for applications to stochastic processes, see for instance [Bally 2003]. This differs from the current work.

The use of asymptotic methods has been much developed during the recent years, mostly in the fields of mathematical finance. As opposed to our setting, the related works deal mainly with one-dimensional processes and specific h. Mathematical approaches are numerous, see [Fouque 2011, Lorig 2013a] and the Chapter 2 of the thesis among others. We nevertheless count some studies devoted to the multidimensional case in the framework of averaged diffusions. Among them we cite the work of Pascucci et al. [Foschi 2013] for the pricing of Asian options in local volatility models, the work of Tankov and Gulisashvili [Gulisashvili 2013] concerning the asymptotic approximation for sums of log-normal r.v. with application to basket option approximations and the work of Avellaneda et al. [Avellaneda 2003] in which is provided an asymptotic formula of the implied volatility in the framework of basket options assuming that each stock follows a one-factor risk-neutral process (with eventually a correlation between the assets).

Here, we address the multi-dimensional case with general functions h and general diffusions X, extending much the setting of previous references.

Organization of the Chapter. In the following, we introduce notations and assumptions that are used throughout the paper. We state in Section 8.2 the main results of the Chapter:

- We first provide in Theorem 8.2.1.1 a second order weak approximation of $\mathbb{E}[h(X_T)]$ using the Gaussian proxy, the magnitude of the error being estimated w.r.t. the SDE coefficients and the time horizon T. The previous approximation, involving correction terms as expectation sensitivities, has an interesting representation as a simple expectation, much suitable for direct Monte-Carlo simulations, see Theorem 8.2.1.2.

- We then perform a suitable multilinear interpolation of the function h in Theorem 8.2.2.1, the accuracy results being given in Theorem 8.2.2.2 according to the regularity of h. The resulting formulas are fully explicit.
- We finally establish a final approximation combining both the weak expansion and the interpolation of h in Theorem 8.2.3.1, providing tight error estimates as well a complexity analysis (Corollary 8.2.3.1).
- Results are extended in Theorem 8.2.4.1 and Corollary 8.2.4.1 considering multiquadratic finite elements.

The proof of the error estimates of Theorems 8.2.1.1 and 8.2.2.2 are respectively given in Sections 8.3 and 8.4. Numerical experiments illustrating the performance of our algorithm in comparison to Monte-Carlo methods are presented in Section 8.5. Appendix 8.6.1 is devoted to the explicit derivation of the corrective terms of the weak expansion provided in Theorem 8.2.1.1.

Notations, definitions and assumptions.

▹**Linear algebra.** The j-th column of a matrix A will be denoted by A_j (or $A_{j,t}$ if A is a time-dependent matrix) and its i-th row by A^i. A^* denotes the transpose of A and if it is a squared matrix, $\det(A)$ stands for its determinant. $\mathcal{I}_m$ denotes the m-dimensional identity matrix, $\langle \cdot, \cdot \rangle$ the inner product on $\mathbb{R}^m$ and $|\cdot|$ is the Euclidean norm on $\mathbb{R}^m$.

▹ **Functions.** As usual, $C^k(O_1, O_2)$ stands for the set of functions $g : O_1 \to O_2$ that are k-times continuously differentiable, where O_1, O_2 are some subsets of Euclidean spaces. Let p_1, p_2 be in $\mathbb{N}\backslash\{0\}$. For any function $g = (g_1, \dots, g_{p_2})^* : [0,T] \times \mathbb{R}^{p_1} \to \mathbb{R}^{p_2}$, we denote $|g|_\infty = \sup_{(t,x)\in[0,T]\times\mathbb{R}^{p_1}} |g(t,x)|$. If g is sufficiently differentiable w.r.t. the variable x, its gradient which takes values in $\mathbb{R}^{p_2} \otimes \mathbb{R}^{p_1}$ is simply denoted $\nabla g(t,x) = (\partial_{x_1} g(t,x), \dots, \partial_{x_{p_1}} g(t,x))$; when $p_2 = 1$, we denote its Hessian matrix by $H(g)(t,x) = (\partial^2_{x_i,x_j} g(t,x))_{i,j\in\{1,\dots,p_1\}}$. Furthermore, we often use the short notation $\partial^\alpha g(t,x)$ for $\partial^{|\alpha|}_{x_{\alpha_1}\dots x_{\alpha_{|\alpha|}}} g(t,x)$, i.e. the partial derivative of g w.r.t. a multi-index α according to the space variable. We denote by $\mathrm{Lip_b}(\mathbb{R}^\mathrm{d}, \mathbb{R})$ the space of Lipschitz functions $h : \mathbb{R}^d \to \mathbb{R}$ satisfying $\mathrm{C}_{\mathrm{Lip},h} := \sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|h(x)-h(y)|}{|x-y|} < +\infty$.

▹ **About the Gaussian proxy.** Whenever unambiguous, we use the notations $\sigma_t := \sigma(t, x_0)$ and $b_t := b(t, x_0)$ for any $t \in [0,T]$ and we denote by $\Sigma_t := \sigma_t \sigma_t^*$ the d-dimensional non-negative definite covariance matrix at time t associated to the Gaussian process X^P defined in (8.2). We start with an easy result, which notations are used throughout the work.

Proposition 8.1.0.1.

1. *The distribution of X^P_T is normal with mean $m^P_T = x_0 + \int_0^T b_t \mathrm{d}t$ and covariance matrix $\mathcal{V}^P_T = \int_0^T \Sigma_t \mathrm{d}t$.*
2. *There is a d-dimensional orthogonal matrix $\mathcal{U}_\mathcal{V}$ such that $\mathcal{V}^P_T = \mathcal{U}_\mathcal{V} \mathcal{D}^P_T \mathcal{U}^{-1}_\mathcal{V}$ where $\mathcal{D}^P_T := \mathrm{diag}(\lambda_1^2 T, \dots, \lambda_d^2 T)$ is a d-dimensional diagonal matrix containing the eigenvalues of $\mathcal{V}^P_T$.*

▷ **Assumption ($\mathcal{H}^{\sigma,b}_{x_0}$) on σ and b.**

($\mathcal{H}^{\sigma,b}_{x_0}$)-i) σ and b are bounded measurable functions from $[0,T]\times\mathbb{R}^d$ to $\mathbb{R}^{d\times q}$ and $\mathbb{R}^d$ respectively, they are twice continuously differentiable w.r.t. x, with uniformly bounded derivatives, and their second derivatives are locally $\alpha\in(0,1]$-Hölder continuous w.r.t. x. We set:

$$\mathcal{M}_1(\sigma,b)=\sum_{\alpha:1\le|\alpha|\le 2}(|\partial^\alpha\sigma|_\infty+|\partial^\alpha b|_\infty)\quad\text{and}\quad\mathcal{M}_0(\sigma,b)=\max(|\sigma|_\infty,|b|_\infty,\mathcal{M}_1(\sigma,b)).$$

To avoid uninteresting situations, we assume $\mathcal{M}_0(\sigma,b)>0$.

($\mathcal{H}^{\sigma,b}_{x_0}$)-ii) There is a constant $\overline{C}_{\mathcal{V}}\ge 1$ such that

$$\overline{C}_{\mathcal{V}}\mathcal{M}_0(\sigma,b)\ge\max_{i\in\{1,\dots,d\}}\lambda_i\ge\min_{i\in\{1,\dots,d\}}\lambda_i\ge(\overline{C}_{\mathcal{V}})^{-1}\mathcal{M}_0(\sigma,b).$$

In particular, the matrix $\mathcal{V}^P_T$ is positive definite.

From ($\mathcal{H}^{\sigma,b}_{x_0}$) and Property 8.1.0.1 we easily deduce

Proposition 8.1.0.2.

1. *The distribution of X^P_T has a density $f^P(x)=\frac{e^{-\frac{1}{2}(x-m^P_T)^*(\mathcal{V}^P_T)^{-1}(x-m^P_T)}}{(2\pi)^{\frac{d}{2}}\sqrt{\det(\mathcal{V}^P_T)}}$, such that for any multi-index α*

$$\left|\partial^\alpha f^P(x)\right|\le C_{\alpha,d}(\mathcal{M}_0(\sigma,b)\sqrt{T})^{-(d+|\alpha|)}\exp\Big(-\frac{|x-m^P_T|^2}{C_{\alpha,d}[\mathcal{M}_0(\sigma,b)]^2T}\Big),\tag{8.4}$$

for a constant $C_{\alpha,d}>0$ that depends in a non-decreasing way on T, $\mathcal{M}_0(\sigma,b)$ and $\overline{C}_{\mathcal{V}}$.

2. *For any measurable function $\phi:\mathbb{R}^d\to\mathbb{R}$ exponentially bounded, define $\overline{\phi}^P:\epsilon\in\mathbb{R}^d\mapsto\overline{\phi}^P(\epsilon)=\mathbb{E}[\phi(X^P_T+\epsilon)]$. Then $\overline{\phi}^P$ is of class C^∞ and all the derivatives $\partial^{|\alpha|}_{\epsilon_{\alpha_1}\dots\epsilon_{\alpha_{|\alpha|}}}\overline{\phi}^P(0):=\partial^{|\alpha|}_{\epsilon_{\alpha_1}\dots\epsilon_{\alpha_{|\alpha|}}}\overline{\phi}^P(\epsilon)|_{\epsilon=0}$ exist for any multi-index $\alpha\in\{1,\dots,d\}^{|\alpha|}$.*

▷ **Miscellaneous.** We use the following notations to state our error estimates throughout the Chapter:

- "$A=O(B)$" means that $|A|\le CB$ where C stands for a generic constant that is a non-negative non-decreasing function of the parameters d, T, $\mathcal{M}_0(\sigma,b)$, $\mathcal{M}_1(\sigma,b)$ and $\overline{C}_{\mathcal{V}}$. Unless made explicit, a generic constant may depend on the test function h.
- Similarly, if A is non-negative, $A\le_c B$ means that $A\le CB$ for a generic constant C.

Lastly, for a r.v. $Y\in\mathbb{R}^m$ ($m\ge 1$) and for $p\ge 1$, $\|Y\|_p=(\mathbb{E}|Y|^p)^{\frac{1}{p}}$ stands for its L^p-norm.

8.2 Main results

8.2.1 Second order weak approximation and Monte Carlo simulations on the Proxy

The model proxy has the advantage to have an explicit Gaussian law and the accuracy of the approximations $\sigma(t,X_t)\approx\sigma(t,x_0)$ and $b(t,X_t)\approx b(t,x_0)$ can be justified if $\mathcal{M}_1(\sigma,b)$, $\mathcal{M}_0(\sigma,b)$ and T are globally small enough (see Lemma 8.3.1.1). Nevertheless, we can not reasonably expect

$\mathbb{E}[h(X_T)] \approx \overline{h}^P(0) = \mathbb{E}[h(X_T^P)]$ to be solely accurate enough and we provide correction terms. To derive them, we make an intensive use of the next interpolated process:

$$X_t^\eta = x_0 + \sum_{j=1}^q \int_0^t \sigma_j(s, \eta X_s^\eta + (1-\eta)x_0)\mathrm{d}W_s^j + \int_0^t b(s, \eta X_s^\eta + (1-\eta)x_0)\mathrm{d}s, \quad \eta \in [0,1], \tag{8.5}$$

so that $X^{\eta=1} = X$ and $X^{\eta=0} = X^P$. Under $(\mathcal{H}_{x_0}^{\sigma,b})$-i), almost surely for any t, $\eta \mapsto X_t^\eta$ is $C^2([0,1],\mathbb{R}^d)$, see [Kunita 1997]. The dynamics of the two first derivatives $(\dot{X}_t^\eta := \partial_\eta X_t^\eta)_{t\geq 0}$ and $(\ddot{X}_t^\eta := \partial^2_{\eta^2} X_t^\eta)_{t\geq 0}$ are obtained by a straight differentiation of the SDE satisfied by X^η:

$$\begin{aligned}
\dot{X}_t^\eta =& \sum_{j=1}^q \int_0^t \nabla\sigma_j(s, x_0 + \eta(X_s^\eta - x_0))(X_s^\eta - x_0 + \eta\dot{X}_s^\eta)\mathrm{d}W_s^j \\
&+ \int_0^t \nabla b(s, x_0 + \eta(X_s^\eta - x_0))(X_s^\eta - x_0 + \eta\dot{X}_s^\eta)\mathrm{d}s, \qquad (8.6)\\
(\ddot{X}_t^\eta)^i =& \sum_{j=1}^q \int_0^t [(X_s^\eta - x_0 + \eta\dot{X}_s^\eta)^* H(\sigma_j^i)(s, x_0 + \eta(X_s^\eta - x_0))(X_s^\eta - x_0 + \eta\dot{X}_s^\eta) \\
&\qquad + \nabla\sigma_j^i(s, x_0 + \eta(X_s^\eta - x_0))(2\dot{X}_s^\eta + \eta\ddot{X}_s^\eta)]\mathrm{d}W_s^j \\
&+ \int_0^t [(X_s^\eta - x_0 + \eta\dot{X}_s^\eta)^* H(b^i)(s, x_0 + \eta(X_s^\eta - x_0))(X_s^\eta - x_0 + \eta\dot{X}_s^\eta) \\
&\qquad + \nabla b^i(s, x_0 + \eta(X_s^\eta - x_0))(2\dot{X}_s^\eta + \eta\ddot{X}_s^\eta)]\mathrm{d}s, \qquad \forall i \in \{1,\dots,d\}. \qquad (8.7)
\end{aligned}$$

Setting $\sigma'_{j,t} := \nabla\sigma_j(t,x_0)$, $\Sigma'_{j,t} := \nabla\Sigma_j(t,x_0)$ and $b'_t := \nabla b(t,x_0)$, $\dot{X} := \dot{X}^{\eta=0}$ is solution of the SDE:

$$\dot{X}_t = \sum_{j=1}^q \int_0^t \sigma'_{j,s}\,(X_s^P - x_0)\mathrm{d}W_s^j + \int_0^t b'_s\,(X_s^P - x_0)\mathrm{d}s. \tag{8.8}$$

Then combining Taylor expansions for the interpolated process X^η and the function h (here assumed to be smooth enough for the sake of brevity), we propose the following weak stochastic approximation:

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] + \mathbb{E}[\nabla h(X_T^P)\dot{X}_T] + \mathrm{Error}_{2,h}^{\mathrm{SA}}, \tag{8.9}$$

where the explicit calculus of the corrective term $\mathbb{E}[\nabla h(X_T^P)\dot{X}_T]$ is performed in Proposition 8.6.1.1 whereas the estimate of the error $\mathrm{Error}_{2,h}^{\mathrm{SA}}$ is postponed to Section 8.3. This leads to the following Theorem (stated for only Lipschitz function h).

Theorem 8.2.1.1. ***(Second order weak approximation using the Gaussian proxy).*** *Assume $(\mathcal{H}_{x_0}^{\sigma,b})$ and suppose that $h \in \mathrm{Lip_b}(\mathbb{R}^\mathrm{d},\mathbb{R})$. Then we have:*

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] + \mathrm{Cor}_{2,h} + \mathrm{Error}_{2,h}^{\mathrm{SA}}, \tag{8.10}$$

where:

$$\begin{aligned}
\mathrm{Cor}_{2,h} =& \nabla\overline{h}^P(0)\int_0^T b'_t\Big(\int_0^t b_s\mathrm{d}s\Big)\mathrm{d}t + \sum_{i,j=1}^d \partial^2_{\epsilon_i,\epsilon_j}\overline{h}^P(0)\Big[\int_0^T (b_t^i)'\Big(\int_0^t \Sigma_{j,s}\mathrm{d}s\Big)\mathrm{d}t + \frac{1}{2}\int_0^T (\Sigma_{j,t}^i)'\Big(\int_0^t b_s\mathrm{d}s\Big)\mathrm{d}t\Big] \\
&+ \frac{1}{2}\sum_{i,j,k=1}^d \partial^3_{\epsilon_i,\epsilon_j,\epsilon_k}\overline{h}^P(0)\int_0^T (\Sigma_{j,t}^i)'\Big(\int_0^t \Sigma_{k,s}\mathrm{d}s\Big)\mathrm{d}t, \qquad (8.11)
\end{aligned}$$

recalling $b'_t = \nabla b(t,x_0)$ and $(\Sigma_{j,t}^i)' = \nabla[\sigma\sigma^*]_j^i(t,x_0)$. The stochastic approximation error term is estimated as follows:

$$|\mathrm{Error}_{2,h}^{\mathrm{SA}}| \leq_c \mathrm{C}_{\mathrm{Lip},h}\mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T^{\frac{3}{2}}. \tag{8.12}$$

Remind that $\mathrm{Cor}_{2,h}$ is well defined whatever the smoothness of h (see Property 8.1.0.2).

Remark 8.2.1.1. *The weak approximation is constituted by a leading order $\overline{h}^P(0)$ plus a sum of weighted sensitivities, i.e. derivatives of $\overline{h}^P$ at zero, up to the third order. The error is of order 3 w.r.t. the standard deviation $\mathcal{M}_0(\sigma,b)\sqrt{T}$ and is null if $\mathcal{M}_1(\sigma,b)=0$ or if $\mathrm{C}_{\mathrm{Lip},h}=0$ (i.e. h is constant). That justifies the label of second order weak approximation. When $h(x)=\phi(\sum_{i=1}^d \eta_i x_i)$ with $\eta_i\geq 0$, the above expansion coincides with that of [Gobet 2012a, Theorem 2.1]related to averaged diffusions.*

Although the density of the Gaussian proxy is known, the approximation formula (8.10) does not reduce to fully explicit calculations, due to the general form of h. Nevertheless, we can derive another representation as an expectation of $h(X_T^P)$ modified by an explicit weight: this is easily obtained by transferring the ε-differentiation of the expectation $\overline{h}^P(\varepsilon)$ (associated to correction terms) into a differentiation of the proxy Gaussian density f^P. This is a somewhat standard argument, in particular regarding the Malliavin calculus applications [Nualart 2006, Section 6.2], we skip details of the derivation. The advantage of this representation as an expectation is to make possible its evaluation by standard Monte-Carlo methods involving only simulations of the Gaussian proxy X_T^P (exact and without discretization), see our subsequent numerical experiments.

Theorem 8.2.1.2. *Under the notations and assumptions of Theorem 8.2.1.1, the main terms of the stochastic approximation are*

$$\mathbb{E}[h(X_T^P)]+\mathrm{Cor}_{2,h}=\mathbb{E}\Big[h(X_T^P)\Big\{1+\mathcal{W}[\Sigma,b;x_0]_0^T([\mathcal{V}_0^T]^{-1}(X_T^P-m_T^P))\Big\}\Big], \tag{8.13}$$

where we set, for $\boldsymbol{Y}\in\mathbb{R}^d$,

$$\begin{aligned}\mathcal{W}[\Sigma,b;x_0]_0^T(\boldsymbol{Y}) =&< \boldsymbol{Y}, \int_0^T b_t'(\int_0^t b_s \mathrm{d}s)\mathrm{d}t >\\
&+\sum_{i,j=1}^d \Big\{\boldsymbol{Y}^i\boldsymbol{Y}^j-([\mathcal{V}_0^T]^{-1})^i_j\Big\}\Big[\int_0^T (b_t^i)'(\int_0^t \Sigma_{j,s}\mathrm{d}s)\mathrm{d}t+\frac{1}{2}\int_0^T(\Sigma^i_{j,t})'(\int_0^t b_s\mathrm{d}s)\mathrm{d}t\Big]\\
&+\frac{1}{2}\sum_{i,j,k=1}^d\Big\{\boldsymbol{Y}^i\boldsymbol{Y}^j\boldsymbol{Y}^k-\boldsymbol{Y}^k([\mathcal{V}_0^T]^{-1})^i_j-\boldsymbol{Y}^j([\mathcal{V}_0^T]^{-1})^i_k-\boldsymbol{Y}^i([\mathcal{V}_0^T]^{-1})^j_k\Big\}\int_0^T(\Sigma^i_{j,t})'(\int_0^t \Sigma_{k,s}\mathrm{d}s)\mathrm{d}t.\end{aligned}$$

As an alternative to a Monte-Carlo evaluation based on (8.13), we provide in the following Subsection a new numerical method to approximate the expansion formula (8.10) taking advantage of a multilinear interpolation with hat functions, which theoretical accuracy is given according to the h-smoothness. The extension to multiquadratic interpolation is presented afterwards.

8.2.2 An efficient algorithm using multilinear finite elements

We define the hat function Λ_z^μ with center $z\in\mathbb{R}$ and size parameter $\mu>0$ by:

$$\Lambda_z^\mu(y)=\frac{y-(z-\mu)}{\mu}\mathbb{1}_{y\in[z-\mu,z[}+\frac{z+\mu-y}{\mu}\mathbb{1}_{y\in[z,z+\mu]}. \tag{8.14}$$

Observe that $\mathbb{E}(\Lambda_z^\mu(G_1))$ is known in explicit form when G_1 is a scalar Gaussian r.v. (like the proxy): therefore, replacing h by a linear interpolation $\hat{h}$ (using the Λ_z^μ-function) leads to a fully explicit formula for (8.10). To extend to the d-dimensional case, we wish to use tensor products of such a function basis in all directions to provide an interpolation of h. However remark that for any (G_1,G_2) Gaussian vector and any z_1,z_2,μ_1,μ_2, the calculus of $\mathbb{E}[\Lambda_{z_1}^{\mu_1}(G_1)\Lambda_{z_2}^{\mu_2}(G_2)]$ is not tractable, except in the case of *zero correlation*; thus an additional ingredient is necessary to maintain explicit formulas. In order to be placed in a situation of uncorrelated Gaussian r.v., we introduce an affine transformation $\mathcal{A}$ of the space, composed of a rotation using the d-dimensional diagonal matrix $\mathcal{U}_{\mathcal{V}}$ (involved in the diagonal decomposition of $\mathcal{V}_T^P$) and a translation of vector m_T^P (the expectation of X_T^P). The following presentation is aimed at providing the construction of the right grid (nodes, directions, size).

▷ **Description of the methodology.** We consider a finite product grid in $\mathbb{R}^d$ defined by:

$$\mathcal{Y} = (y_i^j)_{(i,j)\in\{1,\dots,d\}\times\{0,\dots,N\}}, \quad y_i^j = -\mathrm{R}_i + j\delta_i, \quad \mathrm{R}_i = \mathrm{R}\lambda_i\sqrt{T}, \quad \delta_i = \delta\lambda_i\sqrt{T}, \quad \delta = \frac{2\mathrm{R}}{N}, \tag{8.15}$$

where we recall that $\lambda_i^2 T$ are the eigenvalues of the covariance matrix $\mathcal{V}_T^P$ and where the grid parameters R and δ are to be specified according to the final approximation accuracy desired. We assume $N \in \mathbb{N}^*$. The grid $\mathcal{Y}$ contains N^d small hypercubes and their vertices are the nodes with coordinates $(y_1^{j_1},\dots,y_d^{j_d})^*$ for any $j_1,\dots,j_d \in \{0,\dots,N\}$. Then we define a new grid $\mathcal{X} = (x^{j_1,\dots,j_d})_{j_1,\dots,j_d\in\{0,\dots,N\}}$ image of $\mathcal{Y}$ by the transformation $\mathcal{A} : x \mapsto \mathcal{A}x = m_T^P + \mathcal{U}_{\mathcal{V}}x$:

$$x^{j_1,\dots,j_d} = (x_1^{j_1,\dots,j_d},\dots,x_d^{j_1,\dots,j_d})^* := \mathcal{A}(y_1^{j_1},\dots,y_d^{j_d})^*.$$

The convex hull of $\mathcal{Y}$ is the hypercube D^P, and let us introduce its image $\widetilde{D}^P$ by $\mathcal{A}$:

$$D^P = [-\mathrm{R}_1,\mathrm{R}_1]\times\cdots\times[-\mathrm{R}_d,\mathrm{R}_d], \qquad \widetilde{D}^P := \mathcal{A}(D^P). \tag{8.16}$$

Then we define the multilinear interpolation of h based on the grid $\mathcal{X}$ by setting, for any $x \in \mathbb{R}^d$,

$$h(x) \approx \hat{h}(x) := \sum_{j_1,\dots,j_d\in\{0,\dots,N\}} h(x^{j_1,\dots,j_d}) \prod_{i=1}^d \Lambda_{y_i^{j_i}}^{\delta_i}((\mathcal{U}_{\mathcal{V}}^{-1}(x-m_T^P))^i). \tag{8.17}$$

Notice that $\hat{h}$ is continuous, vanishes outside the domain $\mathcal{A}([-\mathrm{R}_1-\delta_1,\mathrm{R}_1+\delta_1]\times\cdots\times[-\mathrm{R}_d-\delta_d,\mathrm{R}_d+\delta_d])$, and the restriction of $\hat{h}$ to D^P is Lipschitz continuous with a Lipschitz constant at most equal to $\mathrm{C}_{\mathrm{Lip},h}$. The above construction is very similar to multilinear Lagrange finite elements on d-parallelotope, see [Brenner 2008] for a general reference.

▷ **Explicit approximation.** Using (8.17) and taking the expectation, we get for the leading order of the expansion (8.10):

$$\mathbb{E}[h(X_T^P)] \approx \mathbb{E}[\hat{h}(X_T^P)] = \sum_{j_1,\dots,j_d\in\{0,\dots,N\}} h(x^{j_1,\dots,j_d})\mathbb{E}[\prod_{i=1}^d \Lambda_{y_i^{j_i}}^{\delta_i}((\mathcal{U}_{\mathcal{V}}^{-1}(X_T^P-m_T^P))^i)], \tag{8.18}$$

where $\mathcal{U}_{\mathcal{V}}^{-1}(X_T^P - m_T^P)$ has the centered Gaussian law with independent components. Thus combining independence and scaling argument, setting

$$y_0 = (y_0^j)_{j\in\{-1,\dots,N+1\}} := (-\mathrm{R}+j\delta)_{j\in\{-1,\dots,N+1\}} \tag{8.19}$$

and using (8.14)-(8.15)-(8.19), a straightforward calculus leads to

$$\mathbb{E}[\prod_{i=1}^d \Lambda_{y_i^{j_i}}^{\delta_i}((\mathcal{U}_{\mathcal{V}}^{-1}(X_T^P-m_T^P))^i)] = \prod_{i=1}^d \mathbb{E}[\Lambda_{y_i^{j_i}}^{\delta_i}(\lambda_i\sqrt{T}W_1^1)] = \prod_{i=1}^d \mathbb{E}[\Lambda_{y_0^{j_i}}^{\delta}(W_1^1)] = \prod_{i=1}^d \beta_{j_i}^{\delta}(y_0), \tag{8.20}$$

where

$$\beta_j^\delta(y_0) := \frac{\beta(y_0^{j+1}) - 2\beta(y_0^j) + \beta(y_0^{j-1})}{\delta}, \quad \beta(x) := x\mathcal{N}(x) + \mathcal{N}'(x), \quad \mathcal{N}(x) := \int_{-\infty}^x \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \mathrm{d}y = \beta'(x). \tag{8.21}$$

Next for the corrective terms (8.11), we similarly replace h by $\hat{h}$, which gives $\mathrm{Cor}_{2,h} \approx \mathrm{Cor}_{2,\hat{h}}$. For any multi-index $\alpha \in \{1,\ldots,d\}^{|\alpha|}$, we get for the derivatives of $\overline{\hat{h}}^P$ using (8.18)-(8.20):

$$\partial^{|\alpha|}_{\epsilon_{\alpha_1},\ldots,\epsilon_{\alpha_{|\alpha|}}} \overline{\hat{h}}^P(\epsilon) = \sum_{j_1,\ldots,j_d \in \{0,\ldots,N\}} h(x^{j_1,\ldots,j_d}) \partial^{|\alpha|}_{\epsilon_{\alpha_1},\ldots,\epsilon_{\alpha_{|\alpha|}}} \prod_{i=1}^d \mathbb{E}[\Lambda^{\delta_i}_{y_i^{j_i}}(\lambda_i \sqrt{T} W_1^1 + (\mathcal{U}_{\mathcal{V}}^{-1}\epsilon)^i)],$$

$$\mathbb{E}[\Lambda^{\delta_i}_{y_i^{j_i}}(\lambda_i \sqrt{T} W_1^1 + (\mathcal{U}_{\mathcal{V}}^{-1}\epsilon)^i)] = \mathbb{E}[\Lambda^{\delta}_{y_0^{j_i} - \frac{(\mathcal{U}_{\mathcal{V}}^{-1}\epsilon)^i}{\lambda_i\sqrt{T}}}(W_1^1)] = \beta^\delta_{j_i}\Big(y_0 - \frac{(\mathcal{U}_{\mathcal{V}}^{-1}\epsilon)^i}{\lambda_i \sqrt{T}}\Big).$$

Thus it is sufficient to compute the perturbed coefficients β^δ as in (8.21) according to the new translated grid $y_0 - \frac{(\mathcal{U}_{\mathcal{V}}^{-1}\epsilon)^i}{\lambda_i\sqrt{T}}$ and then to differentiate w.r.t. ϵ at $\epsilon = 0$, which leads to explicit calculations. The next result summarizes the previous analysis, in combination with Theorem 8.2.1.1.

Theorem 8.2.2.1. ***(SAFE method with multilinear finite elements)****.*
Assume $(\mathcal{H}^{\sigma,b}_{x_0})$ *and suppose that* $h \in \mathrm{Lip_b}(\mathbb{R}^\mathrm{d},\mathbb{R})$*. Define*

$$\overline{\hat{h}}^P(\epsilon) := \sum_{j_1,\ldots,j_d \in \{0,\ldots,N\}} h(x^{j_1,\ldots,j_d}) \prod_{i=1}^d \beta^\delta_{j_i}\Big(y_0 - \frac{(\mathcal{U}_{\mathcal{V}}^{-1}\epsilon)^i}{\lambda_i\sqrt{T}}\Big).$$

where the weight functions $\beta^\delta_{j_i}$ *and the grid* y_0 *are respectively defined in* (8.21) *and* (8.19)*. Then we have*

$$\mathbb{E}[h(X_T)] = \overline{\hat{h}}^P(0) + \mathrm{Cor}_{2,\hat{h}} + \mathrm{Error}^{\mathrm{SA}}_{2,h} + \mathrm{Error}^{\mathrm{FEL}}_h, \tag{8.22}$$

where

$$\begin{aligned}\mathrm{Cor}_{2,\hat{h}} =& \nabla\overline{\hat{h}}^P(0)\int_0^T b_t'(\int_0^t b_s \mathrm{d}s)\mathrm{d}t + \sum_{i,j=1}^d \partial^2_{\epsilon_i,\epsilon_j}\overline{\hat{h}}^P(0)\Big[\int_0^T (b_t^i)'(\int_0^t \Sigma_{j,s}\mathrm{d}s)\mathrm{d}t + \frac{1}{2}\int_0^T (\Sigma^i_{j,t})'(\int_0^t b_s\mathrm{d}s)\mathrm{d}t\Big]\\ &+ \frac{1}{2}\sum_{i,j,k=1}^d \partial^3_{\epsilon_i,\epsilon_j,\epsilon_k}\overline{\hat{h}}^P(0)\int_0^T (\Sigma^i_{j,t})'(\int_0^t \Sigma_{k,s}\mathrm{d}s)\mathrm{d}t,\end{aligned}$$

and where the error using the multilinear finite elements approximation is defined by:

$$\mathrm{Error}^{\mathrm{FEL}}_h := \overline{h}^P(0) - \overline{\hat{h}}^P(0) + \mathrm{Cor}_{2,h} - \mathrm{Cor}_{2,\hat{h}}. \tag{8.23}$$

▷ **Accuracy results.** The accuracy of multilinear interpolation depends on the smoothness of the function h to approximate. Our goal is not to be exhaustive in this respect but rather to give few settings relevant for the practical applications that we have in mind. For a detailed exposure on Finite Elements accuracy, see [Brenner 2008]. We distinguish three kinds of increasingly strong assumptions:

- **(H1)** : $h \in \mathrm{Lip_b}(\mathbb{R}^\mathrm{d},\mathbb{R})$.
- **(H2)** : $h \in \mathrm{Lip_b}(\mathbb{R}^\mathrm{d},\mathbb{R})$, piecewise C^2, in the sense that there are an integer $N_h \in \mathbb{N}^*$, N_h domains (non empty open connected sets of $\mathbb{R}^d$) $(D^i)_{i\in\{1,\ldots,N_h\}}$, such that:

1. $\forall i \in \{1,\dots,N_h\}$, either the domain D^i has a compact boundary ∂D^i of class C^2, or D^i is a half-space,
2. $\mathbb{R}^d = \bigcup_{i=1}^{i=N_h} \overline{D^i}$,
3. $\forall i \in \{1,\dots,N_h\}$, the restriction of h to D^i, which is denoted by h_i, is a $C^2(\overline{D^i},\mathbb{R}^d)$-function with bounded derivatives.

(H3) : $h \in C^2(\mathbb{R}^d,\mathbb{R})$ with bounded derivatives.

We state the accuracy results in the following Theorem, which proof is postponed to Section 8.4.

Theorem 8.2.2.2. *(Accuracy of SAFE method with multilinear finite elements). Assume $(\mathcal{H}_{x_0}^{\sigma,b})$ and suppose that h satisfies at least **(H1)**. Recalling the density f^P of X_T^P in (8.4) and the domain $\widetilde{D}^P$ in (8.16), for any multi-index α set*

$$\mathcal{G}_h^{\alpha,\mathrm{T}} = \int_{\mathbb{R}^d} \mathbb{1}_{y\notin\widetilde{D}^P} h(y)\partial^\alpha (f^P(y-\epsilon))\Big|_{\epsilon=0}\mathrm{d}y, \qquad \mathcal{G}_h^{\alpha,\mathrm{I}} = \int_{\mathbb{R}^d} \mathbb{1}_{y\in\widetilde{D}^P} h(y)\partial^\alpha (f^P(y-\epsilon))\Big|_{\epsilon=0}\mathrm{d}y. \tag{8.24}$$

Then, define $\mathrm{Cor}_{2,h}^{\mathrm{T}}$ (respectively $\mathrm{Cor}_{2,h}^{\mathrm{I}}$) replacing in $\mathrm{Cor}_{2,h}$ the sensitivities $\partial^\alpha \overline{h}^P(0)$ by $\mathcal{G}_h^{\alpha,\mathrm{T}}$ (respectively $\mathcal{G}_h^{\alpha,\mathrm{I}}$) so that $\mathrm{Cor}_{2,h} = \mathrm{Cor}_{2,h}^{\mathrm{T}} + \mathrm{Cor}_{2,h}^{\mathrm{I}}$; proceed similarly with $\mathcal{G}_{\hat h}^{\alpha,\mathrm{T}}$, $\mathcal{G}_{\hat h}^{\alpha,\mathrm{I}}$, $\mathrm{Cor}_{2,\hat h}^{\mathrm{T}}$ and $\mathrm{Cor}_{2,\hat h}^{\mathrm{I}}$. Then the multilinear finite elements error (8.23) is decomposed as

$$\mathrm{Error}_h^{\mathrm{FEL}} = \mathrm{Error}_h^{\mathrm{FEL,T}} + \mathrm{Error}_h^{\mathrm{FEL,I}},$$

where the Truncation Error

$$\mathrm{Error}_h^{\mathrm{FEL,T}} := \mathbb{E}[(h(X_T^P) - \hat h(X_T^P))\mathbb{1}_{X_T^P\notin\widetilde{D}^P}] + \mathrm{Cor}_{2,h}^{\mathrm{T}} - \mathrm{Cor}_{2,\hat h}^{\mathrm{T}}$$

strongly depends on the size parameter R *introduced in* (8.15)*, and where the Interpolation Error on $\widetilde{D}^P$*

$$\mathrm{Error}_h^{\mathrm{FEL,I}} := \mathbb{E}[(h(X_T^P) - \hat h(X_T^P))\mathbb{1}_{X_T^P\in\widetilde{D}^P}] + \mathrm{Cor}_{2,h}^{\mathrm{I}} - \mathrm{Cor}_{2,\hat h}^{\mathrm{I}}$$

depends on the grid mesh δ. On the one hand, for $h \in \mathrm{Lip_b}(\mathbb{R}^\mathrm{d},\mathbb{R})$, the truncation error is such

$$|\mathrm{Error}_h^{\mathrm{FEL,T}}| \leq_c (|h(m_T^P)| + \mathrm{C}_{\mathrm{Lip},h})\exp(-\mathrm{R}^2/4). \tag{8.25}$$

On the other hand, the Interpolation Error is estimated as follows, according to the regularity of h:

$$\mathrm{Error}_h^{\mathrm{FEL,I}} \leq c \begin{cases} \mathrm{C}_{\mathrm{Lip},h}\delta\mathcal{M}_0(\sigma,b)\sqrt{T} & \textit{under } \textbf{(H1)}, \\ \left\{\mathrm{C}_{\mathrm{Lip},h} + \max\limits_{i\le N_h,\ \alpha:|\alpha|=2} |\partial^\alpha h_i|_\infty\right\}\delta\mathcal{M}_0(\sigma,b)\sqrt{T}\left[\delta + \mathcal{M}_0(\sigma,b)\sqrt{T}\right] & \textit{under } \textbf{(H2)}, \\ \sup\limits_{\alpha:|\alpha|=2} |\partial^\alpha h|_\infty\, \delta^2[\mathcal{M}_0(\sigma,b)\sqrt{T}]^2 & \textit{under } \textbf{(H3)}, \end{cases} \tag{8.26}$$

*where the generic constant c in case **(H2)** depends on the domains.*

8.2.3 Final approximation and complexity of the SAFE algorithm based on multilinear finite elements

Combining Theorems 8.2.1.1 (weak approximation with the Gaussian proxy) and 8.2.2.2 (suitable interpolation of h using the hat functions), we derive a final approximation of $\mathbb{E}[h(X_T)]$ with a choice of parameters R and δ (see (8.15)) allowing to obtain a global error of order at most equal to

$$\mathcal{E} = [\mathcal{M}_0(\sigma,b)\sqrt{T}]^3.$$

The proof of the following Theorem is left to the reader.

Theorem 8.2.3.1. *Assume $(\mathcal{H}_{x_0}^{\sigma,b})$ and suppose that h satisfies at least* ***(H1)*** *. Consider the local approximation $\hat{h}$ of h defined in* (8.17) *with parameters* R *and* δ *set as follows:*

$$\mathrm{R} := 2\sqrt{\log(1/\mathcal{E})}, \qquad \delta := c\begin{cases} [\max_i \lambda_i \sqrt{T}]^2 & \textit{under } \textbf{(H1)}, \\ \max_i \lambda_i \sqrt{T} & \textit{under } \textbf{(H2)}, \\ [\max_i \lambda_i \sqrt{T}]^{\frac{1}{2}} & \textit{under } \textbf{(H3)}, \end{cases}$$

for an arbitrary fixed constant c. Then, the global error is of order 3 w.r.t. $\mathcal{M}_0(\sigma,b)\sqrt{T}$:

$$\mathbb{E}[h(X_T)] = \mathbb{E}[\hat{h}(X_T^P)] + \mathrm{Cor}_{2,\hat{h}} + O([\mathcal{M}_0(\sigma,b)\sqrt{T}]^3).$$

Now, let us analyze the algorithm complexity w.r.t. the target error $\mathcal{E}$, according to the regularity of h.

Denote by $C(d)$ the computational cost for the elementary operations at each node $x^{j_1,\dots,j_d}$ of the local approximation: apart from the evaluation of $h(x^{j_1,\dots,j_d})$, computations are mainly dedicated to the calculus of the β weights defined in (8.21) and their derivatives, which is simple and can even be made off-line. Therefore, the total computational cost of the algorithm is $C^{\mathrm{FEL}}_{calculus} = O\Big(C(d)(N+1)^d\Big)$. Since $N = 2\mathrm{R}/\delta$, the complexity of the algorithm to reach the target error $\mathcal{E} = [\mathcal{M}_0(\sigma,b)\sqrt{T}]^3$ can be evaluated in the following manner.

Corollary 8.2.3.1. *With the previous notations and assumptions, as $\mathcal{E} \to 0$ we have*

$$C^{\mathrm{FEL}}_{calculus} = \begin{cases} O\Big([\log(1/\mathcal{E})]^{d/2}\mathcal{E}^{-\frac{2d}{3}}\Big) & \textit{under } \textbf{(H1)}, \\ O\Big([\log(1/\mathcal{E})]^{d/2}\mathcal{E}^{-\frac{d}{3}}\Big) & \textit{under } \textbf{(H2)}, \\ O\Big([\log(1/\mathcal{E})]^{d/2}\mathcal{E}^{-\frac{d}{6}}\Big) & \textit{under } \textbf{(H3)}. \end{cases} \tag{8.27}$$

Let us briefly discuss the theoretical efficiency of our algorithm in comparison to a direct Monte-Carlo method. If we perform M simulations of the diffusion X_T via an Euler scheme $X_T^{\Delta t}$ with time step Δt, the total computational cost is of order $M \times (T/\Delta t)$, whereas the mean square error is (see [Duffie 1995])

$$\mathbb{V}ar[h(X_T^{\Delta t})]M^{-1} + \Big(\mathbb{E}[h(X_T)] - \mathbb{E}[h(X_T^{\Delta t})]\Big)^2.$$

The first term (statistical error) is approximately equal to

$$\mathbb{V}ar[h(X_T)]M^{-1} = \mathbb{V}ar[h(X_T) - h(x_0)]M^{-1} = O([\mathrm{C}_{\mathrm{Lip},h}\mathcal{M}_0(\sigma,b)\sqrt{T}]^2 M^{-1}).$$

The second error term (discretization error) is a bit delicate to analyze under our assumptions: Kebaier shows in [Kebaier 2005, Proposition 2.2] that for any $\alpha \in (\frac{1}{2}, 1]$, there is a Lipschitz function h so that the discretization error is $O((\Delta t)^\alpha)$. In [Talay 1990, Bally 1996], the order $\alpha = 1$ is established for smooth h or for uniform (hypo)-elliptic σ, but these assumptions are not fulfilled in our setting. To encompass general results, we rather use strong convergence estimates, and we specialize them to our setting:

$$\Big|\mathbb{E}[h(X_T)] - \mathbb{E}[h(X_T^{\Delta t})\Big| \le \mathrm{C}_{\mathrm{Lip},h}\mathbb{E}|X_T - X_T^{\Delta t}| \le_c \mathrm{C}_{\mathrm{Lip},h}[\mathcal{M}_0(\sigma,b)]^2\sqrt{T}\sqrt{\Delta t}.$$

We now tune the parameters to achieve a L^2-error of the same order as the SAFE method when $\mathcal{E} \to 0$, i.e. to have a mean square error $\mathcal{E}^2 = [\mathcal{M}_0(\sigma,b)\sqrt{T}]^6$: this is achieved by taking $M^{-1} \sim [\mathcal{M}_0(\sigma,b)\sqrt{T}]^4 = \mathcal{E}^{\frac{4}{3}}$ and $\Delta t \sim [\mathcal{M}_0(\sigma,b)]^2 T^2 = \mathcal{E}^{\frac{2}{3}} T$. Therefore, the computational cost is $C^{\mathrm{MC}}_{calculus} = O(\mathcal{E}^{-2})$, independently of the dimension. Thus, in view of (8.27) and neglecting logarithmic factors, the SAFE method with multilinear finite elements is (in the sense of this theoretical comparison) more competitive than Monte-Carlo methods up to the dimension

$$d = 3 \text{ under } \textbf{(H1)}\,, \qquad d = 6 \text{ under } \textbf{(H2)}\,, \qquad d = 12 \text{ under } \textbf{(H3)}\,.$$

We now analyse the efficiency of the SAFE methodology with multilinear finite elements in comparison to Monte-Carlo simulations on the Gaussian proxy X_T^P (see (8.13)). The simulation of X_T^P is exact and without discretization, and thus similar arguments to those previously employed lead to a computational cost $C_{calculus}^{\text{MC Proxy}} = O(M) = O(\mathcal{E}^{-\frac{4}{3}})$. Consequently, in view of (8.27), the SAFE method with multilinear finite elements presents a better theoretical efficiency than Monte-Carlo simulations on the proxy up to the dimension:

$$d = 2 \text{ under } \textbf{(H1)}\,, \qquad d = 4 \text{ under } \textbf{(H2)}\,, \qquad d = 8 \text{ under } \textbf{(H3)}\,.$$

Subsequent experiments are coherent with these rules.

8.2.4 SAFE with multiquadratic finite elements

For smooth functions at least of class $C^3(\mathbb{R}^d, \mathbb{R})$, we propose an extension using multiquadratic finite elements [Brenner 2008, Section 3.5] allowing to reach a given accuracy with a lower computational cost. We define two basis functions of center and size parameters $z \in \mathbb{R}$ and $\mu > 0$:

$$\begin{aligned}
\zeta_z^\mu(y) &= \frac{y-(z-\mu)}{\mu}\Big(2\Big(\frac{y-(z-\mu)}{\mu}\Big)-1\Big)\mathbb{1}_{y\in[z-\mu,z[} + \frac{z+\mu-y}{\mu}\Big(2\Big(\frac{z+\mu-y}{\mu}\Big)-1\Big)\mathbb{1}_{y\in[z,z+\mu]},\\
\Xi_z^\mu(y) &= \frac{(y-(z-\mu))}{\mu}\frac{(z+\mu-y)}{\mu}\mathbb{1}_{y\in[z-\mu,z+\mu]}.
\end{aligned}$$

To obtain a multiquadratic interpolation of h, for any $i \in \{1,\dots,d\}$ and any $j \in \{0,\dots,N-1\}$ we denote by $y_0^{j+\frac{1}{2}} = -\mathrm{R} + (j+\frac{1}{2})\delta$ and $y_i^{j+\frac{1}{2}} = -\mathrm{R}_i + (j+\frac{1}{2})\delta_i$ the middles of respectively the segments $[y_0^j, y_0^{j+1}]$ and $[y_i^j, y_i^{j+1}]$. Then we naturally extend
$\mathcal{X} = (x^{j_1,\dots,j_d})_{j_1,\dots,j_d\in\{0,\frac{1}{2},\dots,N-\frac{1}{2},N\}} = (\mathcal{A}(y_1^{j_1},\dots,y_d^{j_d})^*)_{j_1,\dots,j_d\in\{0,\frac{1}{2},\dots,N-\frac{1}{2},N\}}$ and the multiquadratic interpolation of h is defined by:

$$Qh(x) := \sum_{j_1,\dots,j_d\in\{0,\frac{1}{2},\dots,N-\frac{1}{2},N\}} h(x^{j_1,\dots,j_d}) \prod_{i=1}^d \big[\zeta_{y_i^{j_i}}^{\delta_i}\big(((\mathcal{U}_\mathcal{V}^{-1}(x-m_T^P))^i\big)\mathbb{1}_{2j_i\equiv 0[2]} + \Xi_{y_i^{j_i}}^{\delta_i/2}\big((\mathcal{U}_\mathcal{V}^{-1}(x-m_T^P))^i\big)\mathbb{1}_{2j_i\equiv 1[2]}\big],$$

which is a continuous approximation, piecewise quadratic in all directions on $\widetilde{D}^P$ and vanishing outside the domain $\mathcal{A}([-\mathrm{R}_1-\delta_1, \mathrm{R}_1+\delta_1]\times\cdots\times[-\mathrm{R}_d-\delta_d, \mathrm{R}_d+\delta_d])$. A direct computation (a little bit tedious but without mathematical difficulty) yields:

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\frac{(\mathcal{U}_\mathcal{V}^{-1}(X_T^P-m_T^P))^i - y_i^{j_i-1}}{\delta_i}\Big)^2\mathbb{1}_{(\mathcal{U}_\mathcal{V}^{-1}(X_T^P-m_T^P))^i\in[y_i^{j_i-1},y_i^{j_i}[}\\
&\quad+\Big(\frac{y_i^{j_i+1}-(\mathcal{U}_\mathcal{V}^{-1}(X_T^P-m_T^P))^i}{\delta_i}\Big)^2\mathbb{1}_{(\mathcal{U}_\mathcal{V}^{-1}(X_T^P-m_T^P))^i\in[y_i^{j_i},y_i^{j_i+1}]}\Big] = 2\mathcal{B}_{j_i}^\delta(y_0),\\
&\mathbb{E}[\zeta_{y_i^{j_i}}^{\delta_i}((\mathcal{U}_\mathcal{V}^{-1}(X_T^P-m_T^P))^i)] = 4\mathcal{B}_{j_i}^\delta(y_0) - \beta_{j_i}^\delta(y_0),\\
&\mathbb{E}[\Xi_{y_i^{j_i}}^{\delta_i/2}((\mathcal{U}_\mathcal{V}^{-1}(X_T^P-m_T^P))^i)] = -2\mathcal{B}_{j_i}^{\delta/2}(y_0) + 2\beta_{j_i}^{\delta/2}(y_0),
\end{aligned}$$

where the function β is defined in (8.21) and where:

$$\mathcal{B}^{\delta}_{j}(y_0) := \frac{\mathcal{B}(y_0^{j+1}) - \mathcal{B}(y_0^{j-1}) - 2\delta\mathcal{B}'(y_0^{j})}{\delta^2}, \qquad \mathcal{B}(x) := \int_{-\infty}^{x} \beta(u)\mathrm{d}u = \frac{(x^2+1)\mathcal{N}(x) + x\mathcal{N}'(x)}{2}, \qquad (8.28)$$
$$\beta^{\delta/2}_{j}(y_0) := \frac{\beta(y_0^{j+1/2}) - 2\beta(y_0^{j}) + \beta(y_0^{j-1/2})}{\delta/2}, \qquad \mathcal{B}^{\delta/2}_{j}(y_0) := \frac{\mathcal{B}(y_0^{j+1/2}) - \mathcal{B}(y_0^{j-1/2}) - 2(\delta/2)\mathcal{B}'(y_0^{j})}{(\delta/2)^2}.$$

We are now in a position to announce the following Theorem, which proof is very similar to those of Theorems 8.2.2.1-8.2.2.2 and is thus left to the reader.

Theorem 8.2.4.1. *(SAFE using multiquadratic finite elements).*
Assume $(\mathcal{H}^{\sigma,b}_{x_0})$ *and suppose that* $h \in C^3(\mathbb{R}^d, \mathbb{R})$ *with bounded derivatives. Define*

$$\overline{Qh}^{P}(\epsilon) := \sum_{j_1,\dots,j_d \in \{0,\frac{1}{2},\dots,N-\frac{1}{2},N\}} h(x^{j_1,\dots,j_d}) \prod_{i=1}^{d} \Big[\Big(4\mathcal{B}^{\delta}_{j_i}\Big(y_0 - \frac{(\mathcal{U}_{\mathcal{V}}^{-1}\epsilon)^i}{\lambda_i\sqrt{T}}\Big) - \beta^{\delta}_{j_i}\Big(y_0 - \frac{(\mathcal{U}_{\mathcal{V}}^{-1}\epsilon)^i}{\lambda_i\sqrt{T}}\Big)\Big)\mathbb{1}_{2j_i \equiv 0[2]}$$
$$+ \Big(-2\mathcal{B}^{\delta/2}_{j_i}\Big(y_0 - \frac{(\mathcal{U}_{\mathcal{V}}^{-1}\epsilon)^i}{\lambda_i\sqrt{T}}\Big) + 2\beta^{\delta/2}_{j_i}\Big(y_0 - \frac{(\mathcal{U}_{\mathcal{V}}^{-1}\epsilon)^i}{\lambda_i\sqrt{T}}\Big)\Big)\mathbb{1}_{2j_i \equiv 1[2]}\Big],$$

where the weight functions $\mathcal{B}^{\delta}_{j_i}$, $\mathcal{B}^{\delta/2}_{j_i}$, $\beta^{\delta}_{j_i}$ *and* $\beta^{\delta/2}_{j_i}$ *and the grid* y_0 *are defined in* (8.28), (8.21) *and* (8.19). *Then*

$$\mathbb{E}[h(X_T)] = \overline{Qh}^{P}(0) + \mathrm{Cor}_{2,Qh} + O(\mathrm{C}_{\mathrm{Lip},h}\mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T^{\frac{3}{2}}) + \mathrm{Error}_h^{\mathrm{FEQ,T}} + \mathrm{Error}_h^{\mathrm{FEQ,I}},$$

where the truncation error $\mathrm{Error}_h^{\mathrm{FEQ,T}}$ *is estimated as in* (8.25) *and where the interpolation error using the multiquadratic interpolation is estimated as follows:* $\mathrm{Error}_h^{\mathrm{FEQ,I}} = O(\delta^3[\mathcal{M}_0(\sigma,b)\sqrt{T}]^3)$.

In order to achieve a global accuracy $\mathcal{E} = [\mathcal{M}_0(\sigma,b)\sqrt{T}]^3$, we conclude as for Corollary 8.2.3.1: namely, take $\mathrm{R} = 2\sqrt{\log(1/\mathcal{E})}$ and $\delta = c$ for an arbitrary fixed constant $c > 0$. It leads to the following complexity result, which demonstrates a better asymptotic theoretical efficiency compared to Monte-Carlo methods in any dimension.

Corollary 8.2.4.1. *With the previous notations and assumptions, as* $\mathcal{E} \to 0$ *we have*
$C^{\mathrm{FEQ}}_{calculus} = O\big([\log(1/\mathcal{E})]^{d/2}\big)$.

For functions mixing various local regularity properties, we have better to use different shape functions and different mesh sizes, and possibly sparse grids [Bungartz 2004]. It can be done easily according to the examples to handle, we do not develop further this adaptive viewpoint.

8.3 Proof of the error estimate in Theorem 8.2.1.1

The estimate of $\mathrm{Error}_{2,h}^{\mathrm{SA}}$ provided in (8.12) is proved in three steps:

1. L^p-estimates of the interpolated process X^η and of its derivatives;
2. Gaussian regularization of h with a small Brownian perturbation;
3. Malliavin integration by parts formula.

We may stress that the step 2 with Gaussian regularization is a crucial ingredient for a secured use of Malliavin calculus integration by parts under the sole pointwise ellipticity $(\mathcal{H}^{\sigma,b}_{x_0})$.

8.3.1 L^p-norm estimates of $X^\eta - x_0$, $\dot{X}^\eta$ and $\ddot{X}^\eta$

Lemma 8.3.1.1. *Assume* $(\mathcal{H}_{x_0}^{\sigma,b})$. *We have the following estimates* $\forall p \geq 1$*:*

$$\sup_{t\in[0,T],\eta\in[0,1]} \|X_t^\eta - x_0\|_p \leq_c \mathcal{M}_0(\sigma,b)\sqrt{T}, \tag{8.29}$$

$$\sup_{t\in[0,T],\eta\in[0,1]} \|\dot{X}_t^\eta\|_p \leq_c \mathcal{M}_1(\sigma,b)\mathcal{M}_0(\sigma,b)T, \qquad \sup_{t\in[0,T],\eta\in[0,1]} \|\ddot{X}_t^\eta\|_p \leq_c \mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T^{\frac{3}{2}}. \tag{8.30}$$

Proof. W.l.o.g. we can assume $p \geq 2$. The estimate (8.29) is standard using classic inequalities and the upper bounds on b and σ in $(\mathcal{H}_{x_0}^{\sigma,b})$. Regarding the estimate of $\dot{X}^\eta$, start from (8.6): usual computations based on Burkholder-Davis-Gundy inequalities and the Gronwall lemma lead to

$$\mathbb{E}|\dot{X}_t^\eta|^p \leq_c (\mathcal{M}_1(\sigma,b)\sqrt{T})^p \sup_{t\in[0,T]} \mathbb{E}|X_t^\eta - x_0|^p,$$

with some generic constant c uniform in $t \in [0,T]$ and $\eta \in [0,1]$. Then plug the estimate (8.29) into this inequality to conclude. Finally for the last estimate of (8.30), start from (8.7) and apply the same inequalities to obtain:

$$\mathbb{E}|\ddot{X}_t^\eta|^p \leq_c (\mathcal{M}_1(\sigma,b)\sqrt{T})^p \Big\{ \sup_{t\in[0,T]} \mathbb{E}|X_t^\eta - x_0|^{2p} + \sup_{t\in[0,T]} \mathbb{E}|\dot{X}_t^\eta|^{2p} + \sup_{t\in[0,T]} \mathbb{E}|\dot{X}_t^\eta|^p \Big\};$$

using the previous estimates easily achieves the proof. □

8.3.2 Regularization of h with a small noise perturbation

Let $\overline{W}$ be an extra independent d-dimensional Brownian motion defined on the same probability space, suppose that $\mathcal{M}_1(\sigma,b) \neq 0$ w.l.o.g. and define the small parameter $\xi = \mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T > 0$ and the $C^\infty(\mathbb{R}^d,\mathbb{R})$-function :

$$h_\xi(x) = \mathbb{E}[h(x + \xi\overline{W}_T)] = \mathbb{E}[h_{\xi/\sqrt{2}}(x + \xi\overline{W}_{\frac{T}{2}})]. \tag{8.31}$$

Replacing h by h_ξ in our expansion analysis (8.10) induces extra errors quantified below.

Lemma 8.3.2.1. *Assume* $(\mathcal{H}_{x_0}^{\sigma,b})$. *For any* $h \in \mathrm{Lip_b}(\mathbb{R}^d,\mathbb{R})$ *and any multi-index* α, *we have:*

$$|h - h_\xi|_\infty \leq_c \mathrm{C}_{\mathrm{Lip},h}\mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T^{\frac{3}{2}},$$

$$\left|\partial^\alpha \overline{h}^P(0) - \partial^\alpha \overline{h_\xi}^P(0)\right| \leq_c \mathrm{C}_{\mathrm{Lip},h}\mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T^{\frac{3}{2}} \Big(\mathcal{M}_0(\sigma,b)\sqrt{T}\Big)^{-|\alpha|}.$$

Proof. The first estimate readily follows from (8.31) and $h \in \mathrm{Lip_b}(\mathbb{R}^d,\mathbb{R})$. About the second estimate, using the density f^P of X_T^P (see (8.4)), write

$$\partial^\alpha \overline{h}^P(0) - \partial^\alpha \overline{h_\xi}^P(0) = \int_{\mathbb{R}^d} [h(y) - h_\xi(y)]\partial^\alpha (f^P(y-\epsilon))\big|_{\epsilon=0} \mathrm{d}y,$$

and complete the proof by combining the first estimate with standard upper bounds for the derivatives of f^P (see Property 8.1.0.2). □

An important consequence of the above lemma is a nice control of correction and error terms w.r.t. h and h_ξ. The proof is straightforward and we skip it.

Corollary 8.3.2.1. *Under the assumptions of Lemma 8.3.2.1, we have*

$$\begin{aligned}
\left|\mathrm{Cor}_{2,h} - \mathrm{Cor}_{2,h_\xi}\right| &\le_c [\mathcal{M}_1(\sigma,b)]^2[\mathcal{M}_0(\sigma,b)]^2 T^2,\\
\left|\mathrm{Error}^{\mathrm{SA}}_{2,h} - \mathrm{Error}^{\mathrm{SA}}_{2,h_\xi}\right| &:= \left|\mathbb{E}[h(X_T)] - \mathbb{E}[h(X^P_T)] - \mathrm{Cor}_{2,h} - (\mathbb{E}[h_\xi(X_T)] - \mathbb{E}[h_\xi(X^P_T)] - \mathrm{Cor}_{2,h_\xi})\right|\\
&\le_c \mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T^{\frac{3}{2}}.
\end{aligned}$$

Hence, proving the error estimate (8.12) is reduced to prove the following Proposition:

Proposition 8.3.2.1. *Under previous assumptions, we have* $|\mathrm{Error}^{\mathrm{SA}}_{2,h_\xi}| \le_c \mathrm{C}_{\mathrm{Lip},h}\mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T^{\frac{3}{2}}$.

8.3.3 Malliavin integration by parts formula and proof of Proposition (8.3.2.1)

This step is aimed at deriving error estimates related to h_ξ, depending only on $\mathrm{C}_{\mathrm{Lip},h_\xi} \le \mathrm{C}_{\mathrm{Lip},h}$ and not on higher derivatives of h_ξ, thanks to Malliavin integration by parts formula: we make it possible because of the extra independent noise $\xi\overline{W}_{T/2}$ in (8.31) which plays an important role on the degeneracy event (see Lemma 8.3.4.1). This ingredient is crucial to make our error analysis available.

Let us detail the argumentation. We consider the Malliavin calculus w.r.t the $(q+d)$-dimensional Brownian motion $(W,\overline{W})$, the Malliavin derivatives associated to W (respectively $\overline{W}$) being denoted by D (respectively $\overline{D}$). We refer to [Nualart 2006] for the related theory and the notations. First we extend the estimates provided in Lemma 8.3.1.1 to the norms $\|\cdot\|_{k,p}$ related to the Sobolev space $\mathbb{D}^{k,p}$, $k=1,2$.

Lemma 8.3.3.1. *Under $(\mathcal{H}^{\sigma,b}_{x_0})$, for any $\eta\in[0,1]$, $i\in\{1,\ldots,d\}$ and $t\in[0,T]$, $(X^\eta_t)^i\in\mathbb{D}^{2,\infty}$ and $(\dot X^\eta_t)^i\in\mathbb{D}^{1,\infty}$. In addition we have the following estimates for any $p\ge 1$:*

$$\|D_r(X^\eta_t)^i\|_p \le_c \mathcal{M}_0(\sigma,b), \qquad \|D^2_{r,s}(X^\eta_t)^i\|_p \le_c \mathcal{M}_0(\sigma,b)\mathcal{M}_1(\sigma,b), \tag{8.32}$$

$$\|D_r(\dot X^\eta_t)^i\|_p \le_c \mathcal{M}_1(\sigma,b)\mathcal{M}_0(\sigma,b)\sqrt{T}, \tag{8.33}$$

uniformly in $r,s,t\in[0,T]$, $i\in\{1,\ldots,d\}$ and $\eta\in[0,1]$.

Proof. The inclusions in $\mathbb{D}^{1,\infty}$ and $\mathbb{D}^{2,\infty}$ are standard to justify under our assumptions, we skip it and we focus on the L^p-estimates. W.l.o.g. we assume $p\ge 2$. We only detail the computations for $\dot X^\eta_t$. Start from (8.6) to get $\forall t\in[0,T]$, $\forall r\in[0,t]$, $\forall(i,k)\in\{1,\ldots,d\}\times\{1,\ldots,q\}$:

$$\begin{aligned}
(D_r\dot X^\eta_t)^i_k =& \nabla\sigma^i_k(r,x_0+\eta(X^\eta_r-x_0))(X^\eta_r-x_0+\eta\dot X^\eta_r)\\
&+\sum_{j=1}^q\int_r^t (X^\eta_u-x_0+\eta\dot X^\eta_u)^* H(\sigma^i_j)(u,x_0+\eta(X^\eta_u-x_0))(D_rX^\eta_u)_k\mathrm{d}W^j_u\\
&+\sum_{j=1}^q\int_r^t \nabla\sigma^i_j(u,x_0+\eta(X^\eta_u-x_0))(D_rX^\eta_u+\eta D_r\dot X^\eta_u)_k\mathrm{d}W^j_u\\
&+\int_r^t (X^\eta_u-x_0+\eta\dot X^\eta_u)^* H(b^i)(u,x_0+\eta(X^\eta_u-x_0))(D_rX^\eta_u)_k\mathrm{d}u\\
&+\int_r^t \nabla b^i(u,x_0+\eta(X^\eta_u-x_0))(D_rX^\eta_u+\eta D_r\dot X^\eta_u)_k\mathrm{d}u.
\end{aligned}$$

Then apply the Young, Burkholder-Davis-Gundy and Hölder inequalities combined to the Gronwall lemma: it gives

$$\mathbb{E}|(D_r\dot X^\eta_t)_k|^p \le_c [\mathcal{M}_1(\sigma,b)]^p\Big\{\mathbb{E}|X^\eta_r-x_0|^p + \mathbb{E}|\dot X^\eta_r|^p + \sup_{u\in[0,T]}\sqrt{\mathbb{E}|(D_rX^\eta_u)_k|^{2p}}\, T^{p/2}$$

$$\times\Big(\sup_{u\in[0,T]}\sqrt{\mathbb{E}|X_u^\eta-x_0|^{2p}}+\sqrt{\sup_{u\in[0,T]}\mathbb{E}|\dot X_u^\eta|^{2p}}\Big)+\sup_{u\in[0,T]}\mathbb{E}|(D_rX_u^\eta)_k|^p\,T^{p/2}\Big\}.$$

In view of (8.29), (8.30) and (8.32), the announced result (8.33) is proved. □

We now state the key result to establish Proposition (8.3.2.1).

Proposition 8.3.3.1. *Assume $(\mathcal{H}_{x_0}^{\sigma,b})$. For $\eta\in[0,1]$, we consider the d-dimensional random variable:*

$$F^\eta=X_T^P+\eta(X_T-X_T^P)+\xi\overline{W}_{\frac{T}{2}}. \tag{8.34}$$

Then for any $Y\in\mathbb{D}^{1,\infty}$ and any $i,j\in\{1,\dots,d\}$, there exists a random variable $Y_{j,\eta}\in\cap_{p\geq1}L^p$ such that

$$\mathbb{E}[\partial^2_{x_i,x_j}h_{\xi/\sqrt{2}}(F_\eta)Y]=\mathbb{E}[\partial_{x_i}h_{\xi/\sqrt{2}}(F_\eta)Y_{j,\eta}],$$

where $\|Y_{j,\eta}\|_p\leq_c\|Y\|_{1,3p}(\mathcal{M}_0(\sigma,b)\sqrt{T})^{-1}$, uniformly in $\eta\in[0,1]$, for any $p\geq1$.

Before proving it, let us complete the proof of Proposition 8.3.2.1. Define the residual processes

$$R_T^{0,X}:=X_T-X_T^P=\int_0^1\dot X_T^\eta\mathrm{d}\eta,\qquad R_T^{1,X}:=X_T-X_T^P-\dot X_T=\int_0^1(1-\eta)\ddot X_T^\eta\mathrm{d}\eta, \tag{8.35}$$

which enable us to represent the error (8.9) as follows:

$$\mathrm{Error}^{\mathrm{SA}}_{2,h_\xi}=\mathbb{E}[\nabla h_\xi(X_T^P)R_T^{1,X}]+\mathbb{E}\Big(\int_0^1(1-\eta)[R_T^{0,X}]^*H(h_\xi)(X_T^P+\eta(X_T-X_T^P))R_T^{0,X}\mathrm{d}\eta\Big). \tag{8.36}$$

The first error term is easily handled combining (8.35), Lemma 8.3.1.1 and $\mathrm{C}_{\mathrm{Lip},h_\xi}\leq\mathrm{C}_{\mathrm{Lip},h}$:

$$\mathbb{E}[\nabla h_\xi(X_T^P)R_T^{1,X}]=\mathcal{O}(\mathrm{C}_{\mathrm{Lip},h}(\mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2T^{\frac32}).$$

For the second error term of (8.36), use (8.31), apply for any $i,j\in\{1,\dots,d\}$ the above Proposition 8.3.3.1 with $Y=(R_{0,t}^X)_i(R_{0,t}^X)_j$ and use Lemmas 8.3.1.1 and 8.3.3.1 combined with (8.35) to get:

$$\begin{aligned}\mathbb{E}\Big[\int_0^1(1-\eta)[R_T^{0,X}]^*H(h_\xi)(X_T^P+\eta(X_T-X_T^P))R_T^{0,X}\mathrm{d}\eta\Big]&=\int_0^1(1-\eta)\mathbb{E}\Big[[R_T^{0,X}]^*H(h_{\xi/\sqrt{2}})(F^\eta)R_T^{0,X}\Big]\mathrm{d}\eta\\&=\mathcal{O}(\mathrm{C}_{\mathrm{Lip},h}[\mathcal{M}_1(\sigma,b)]^2\mathcal{M}_0(\sigma,b)T^{\frac32}).\end{aligned}$$

Proposition 8.3.2.1 is proved.

8.3.4 Proof of Proposition 8.3.3.1

It is clear that under $(\mathcal{H}_{x_0}^{\sigma,b})$, F^η defined in (8.34) is in $\mathbb{D}^{2,\infty}$ and is non degenerate since its Malliavin covariance matrix γ_{F^η} is such that

$$\gamma_{F^\eta}=\int_0^T D_t(X_T^P+\eta(X_T-X_T^P))\,[D_t(X_T^P+\eta(X_T-X_T^P))]^*\mathrm{d}t+\xi^2\frac{T}{2}\mathcal{I}_d\geq\xi^2\frac{T}{2}\mathcal{I}_d, \tag{8.37}$$

$$\gamma_{F^0}=\mathcal{V}_T^P+\xi^2\frac{T}{2}I_d\geq(\overline{C}_{\mathcal{V}})^{-2}[\mathcal{M}_0(\sigma,b)]^2T\mathcal{I}_d. \tag{8.38}$$

Then [Nualart 2006, Propositions 2.1.4 and 1.5.6] proves the existence of $Y_{j,\eta}$ such that for any $p\geq1$,

$$\|Y_{j,\eta}\|_p\leq_c\|Y\|_{1,3p}\|\gamma_{F^\eta}^{-1}\|_{1,3p}\|(DF^\eta,\overline{D}F^\eta)\|_{1,3p}\leq_c\|Y\|_{1,3p}\|\gamma_{F^\eta}^{-1}\|_{1,3p}\|\mathcal{M}_0(\sigma,b)\sqrt{T},$$

where we have used (8.32) and the value of ξ. To complete the proof of Proposition 8.3.3.1, apply the following estimates related to $\gamma_{F^\eta}^{-1}$:

Lemma 8.3.4.1. *Assume* $(\mathcal{H}^{\sigma,b}_{x_0})$*. For any* $p \geq 1$ *and any* $\eta \in [0,1]$*, we have:*

$$\|\det{}^{-1}(\gamma_{F^\eta})\|_p \leq_c (\mathcal{M}_0(\sigma,b)\sqrt{T})^{-2d}, \tag{8.39}$$

$$\|(\gamma^{-1}_{F^\eta})^i_j\|_p \leq_c (\mathcal{M}_0(\sigma,b)\sqrt{T})^{-2}, \qquad \sup_{t\in[0,T]} \|D_t(\gamma^{-1}_{F^\eta})^i_j\|_p \leq_c \mathcal{M}_1(\sigma,b)(\mathcal{M}_0(\sigma,b)\sqrt{T})^{-2}, \tag{8.40}$$

uniformly in $\eta \in [0,1]$.

Proof. Here we rely much on the assumption $(\mathcal{H}^{\sigma,b}_{x_0})$ about the oscillation of eigenvalues of $\mathcal{V}^P_T$. All the next generic constants are uniform in $\eta \in [0,1]$. Using the definition of γ_{F^η} (8.37), write

$$\gamma_{F^\eta} - \gamma_{F^0} = \eta \int_0^T \Big\{ D_t X^P_T [D_t(X_T - X^P_T)]^* + D_t(X_T - X^P_T)[D_t X^P_T + \eta D_t(X_T - X^P_T)]^* \Big\} \mathrm{d}t.$$

Then, combining (8.35), (8.32) and (8.33), it readily follows ($\forall p \geq 1$)

$$\|\gamma_{F^\eta} - \gamma_{F^0}\|_p \leq_c \mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T^{\frac{3}{2}}, \qquad \|\gamma_{F^\eta}\|_p \leq_c (\mathcal{M}_0(\sigma,b)\sqrt{T})^2. \tag{8.41}$$

Thus, it is an easy exercice (using the Leibniz formula for determinant) to deduce

$$\frac{\|\det(\gamma_{F^\eta}) - \det(\gamma_{F^0})\|_p}{\det(\gamma_{F^0})} \leq_c \frac{\mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T^{\frac{3}{2}} (\mathcal{M}_0(\sigma,b)\sqrt{T})^{2(d-1)}}{(\mathcal{M}_0(\sigma,b)\sqrt{T})^{2d}} \leq \mathcal{M}_1(\sigma,b)\sqrt{T}. \tag{8.42}$$

We are in a position to prove (8.39). For any given $p \geq 1$ and $m \geq 1$,

$$\begin{aligned}
\mathbb{E}[\det{}^{-p}(\gamma_{F^\eta})] =& \mathbb{E}[\det{}^{-p}(\gamma_{F^\eta}) \mathbb{1}_{\det(\gamma_{F^\eta}) \leq \frac{1}{2}\det(\gamma_{F^0})}] + \mathbb{E}[\det{}^{-p}(\gamma_{F^\eta}) \mathbb{1}_{\det(\gamma_{F^\eta}) > \frac{1}{2}\det(\gamma_{F^0})}] \\
\leq_c& (\xi^2 T)^{-dp} \mathbb{P}\Big(\det(\gamma_{F^0}) - \det(\gamma_{F^\eta}) \geq \frac{\det(\gamma_{F^0})}{2}\Big) + \det{}^{-p}(\gamma_{F^0}) \\
\leq_c& ([\mathcal{M}_1(\sigma,b)]^2 [\mathcal{M}_0(\sigma,b)]^4 T^3)^{-dp} \det{}^{-m}(\gamma_{F^0}) \mathbb{E}[|\det(\gamma_{F^\eta}) - \det(\gamma_{F^0})|^m] + (\mathcal{M}_0(\sigma,b)\sqrt{T})^{-2pd} \\
\leq_c& ([\mathcal{M}_1(\sigma,b)]^2 [\mathcal{M}_0(\sigma,b)]^4 T^3)^{-dp} (\mathcal{M}_1(\sigma,b)\sqrt{T})^m + (\mathcal{M}_0(\sigma,b)\sqrt{T})^{-2pd},
\end{aligned}$$

where we have notably used (8.37) at the first inequality, the Markov inequality, the definition of ξ and (8.38) at the second one and (8.42) at the last one. Then choose $m = 4pd$ to readily obtain:

$$\mathbb{E}[\det{}^{-p}(\gamma_{F^\eta})] \leq_c (\mathcal{M}_0(\sigma,b)\sqrt{T})^{-2pd}.$$

The first estimate of (8.40) readily follows from (8.39)-(8.41). Regarding the second one, it is a consequence of $D_t(\gamma^{-1}_{F^\eta})^i_j = -\sum_{m,l=1}^d (\gamma^{-1}_{F^\eta})^i_m (\gamma^{-1}_{F^\eta})^l_j D_t(\gamma_{F^\eta})^m_l$ (see [Nualart 2006, Lemma 2.1.6]) and of the estimate $\|D_t(\gamma_{F^\eta})^i_j\|_p \leq_c \mathcal{M}_1(\sigma,b)[\mathcal{M}_0(\sigma,b)]^2 T$ which comes from (8.32). □

8.4 Proof of Theorem 8.2.2.2

8.4.1 Truncation Error $\mathrm{Error}^{\mathrm{FEL,T}}_h$

▷**Estimate of** $\mathbb{E}[(h - \hat{h})(X^P_T)\mathbb{1}_{X^P_T \notin \widetilde{D}^P}]$. By construction of $\hat{h}$, we have
$|\hat{h}(x)| \leq \sup_{|x| \in \widetilde{D}^P} |h(x)| \leq |h(m^P_T)| + \mathrm{C}_{\mathrm{Lip},h} \sqrt{d} \max_i \mathrm{R}_i$, which becomes
$|\hat{h}(x)| \leq |h(m^P_T)| + \sqrt{d}\frac{\max_i \mathrm{R}_i}{\min_i \mathrm{R}_i} \mathrm{C}_{\mathrm{Lip},h} |x - m^P_T|$ for $x \notin \widetilde{D}^P$. Besides, $|h(x)| \leq |h(m^P_T)| + \mathrm{C}_{\mathrm{Lip},h}|x - m^P_T|$. Therefore

$$|h(x) - \hat{h}(x)| \leq 2|h(m^P_T)| + \Big(1 + \sqrt{d}\frac{\max_i \mathrm{R}_i}{\min_i \mathrm{R}_i}\Big) \mathrm{C}_{\mathrm{Lip},h} |x - m^P_T|, \qquad \forall x \notin \widetilde{D}^P. \tag{8.43}$$

Then using the Cauchy-Schwarz inequality and Lemma 8.3.1.1, we get:

$$|\mathbb{E}[(h-\hat{h})(X_T^P)\mathbb{1}_{X_T^P\notin\widetilde{D}^P}]| \le_c |h(m_T^P)|\mathbb{P}(X_T^P\notin\widetilde{D}^P)+\mathrm{C}_{\mathrm{Lip},h}\mathcal{M}_0(\sigma,b)\sqrt{T}\Big[\mathbb{P}(X_T^P\notin\widetilde{D}^P)\Big]^{\frac{1}{2}}.$$

By the definition (8.16) of the domains D^P and $\widetilde{D}^P$, we have

$$\mathbb{P}(X_T^P\notin\widetilde{D}^P)=\mathbb{P}(W_1\notin[-\mathrm{R},\mathrm{R}]^d)\le d\,\mathbb{P}(|W_1^1|>\mathrm{R})\le 2de^{-\frac{\mathrm{R}^2}{2}} \tag{8.44}$$

using a standard Gaussian concentration inequality. Finally, we have proved

$$|\mathbb{E}[(h-\hat{h})(X_T^P)\mathbb{1}_{X_T^P\notin\widetilde{D}^P}]| \le_c (|h(m_T^P)|+\mathrm{C}_{\mathrm{Lip},h})e^{-\frac{\mathrm{R}^2}{4}}. \tag{8.45}$$

▷**Estimate of** $\mathrm{Cor}^{\mathrm{T}}_{2,h}-\mathrm{Cor}^{\mathrm{T}}_{2,\hat{h}}$. Starting from the definition (8.24) and using the inequality (8.43), the estimates (8.29)-(8.44) and the Hölder inequality, we derive that $\left|\mathcal{G}_h^{\alpha,\mathrm{T}}-\mathcal{G}_{\hat{h}}^{\alpha,\mathrm{T}}\right|$ is bounded by

$$\Big\{\int_{\mathbb{R}^d}\mathbb{1}_{y\notin\widetilde{D}^P}\left|h(y)-\hat{h}(y)\right|^4 f^P(y)\mathrm{d}y\Big\}^{\frac14}\Big\{\int_{\mathbb{R}^d}\left|(f^P(y))^{-1}\partial^\alpha(f^P(y-\epsilon))\right|_{\epsilon=0}\Big|^4 f^P(y)\mathrm{d}y\Big\}^{\frac14}\Big\{\int_{\mathbb{R}^d}\mathbb{1}_{y\notin\widetilde{D}^P}f^P(y)\mathrm{d}y\Big\}^{\frac12}$$
$$\le_c(|h(m_T^P)|+\mathrm{C}_{\mathrm{Lip},h})\,(\mathcal{M}_0(\sigma,b)\sqrt{T})^{-|\alpha|}e^{-\frac{\mathrm{R}^2}{4}},$$

which easily leads to $|\mathrm{Cor}^{\mathrm{T}}_{2,h}-\mathrm{Cor}^{\mathrm{T}}_{2,\hat{h}}|\le_c(|h(m_T^P))|+\mathrm{C}_{\mathrm{Lip},h})e^{-\frac{\mathrm{R}^2}{4}}$. Combining this last estimate with (8.45) achieves the proof of (8.25).

8.4.2 Interpolation Error $\mathrm{Error}_h^{\mathrm{FEL,I}}$

For any $j_1,\dots,j_d\in\{0,\dots,N-1\}$ we denote by $D^{P,j_1,..j_d}$ the hyperrectangle $[y_1^{j_1},y_1^{j_1+1}]\times\cdots\times[y_d^{j_d},y_d^{j_d+1}]$ and we set $\widetilde{D}^{P,j_1,..j_d}=\mathcal{A}(D^{P,j_1,..j_d})$.

▷**Estimate of** $\mathbb{E}[(h-\hat{h})(X_T^P)\mathbb{1}_{X_T^P\in\widetilde{D}^P}]$**, cases (H1) and (H3) .** Our local approximation consists in using a tensor product finite elements of order 1 on d-parallelotope. We have for any $j_1,\dots,j_d\in\{0,\dots,N-1\}$

$$\mathrm{diam}(\widetilde{D}^{P,j_1,..j_d}):=\max_{x,x'\in\widetilde{D}^{P,j_1,..j_d}}|x-x'|=\mathrm{diam}(D^{P,j_1,..j_d})$$
$$=\Big(\sum_{i=1}^d\delta_i^2\Big)^{\frac12}=\delta\Big(\sum_{i=1}^d\lambda_i^2T\Big)^{\frac12}\le\mu:=\overline{C}_{\mathcal{V}}\sqrt{d}\delta\mathcal{M}_0(\sigma,b)\sqrt{T}. \tag{8.46}$$

Hence an application of [Brenner 2008, Theorem 4.6.14] yields that

$$\sup_{j_1,\dots,j_d}\sup_{x\in\widetilde{D}^{P,j_1..j_d}}\left|h(x)-\hat{h}(x)\right|\le c_0\begin{cases}\mathrm{C}_{\mathrm{Lip},h}\,\mu, & \text{case } \textbf{(H1)}\,,\\ \big(\sum_{\alpha:|\alpha|=2}|\partial^\alpha h|_\infty\big)\,\mu^2, & \text{case } \textbf{(H3)}\,,\end{cases} \tag{8.47}$$

where c_0 is a universal constant. This obviously leads to :

$$\left|\mathbb{E}[(h-\hat{h})(X_T^P)\mathbb{1}_{X_T^P\in\widetilde{D}^P}]\right|\le_c\begin{cases}\mathrm{C}_{\mathrm{Lip},h}\delta\mathcal{M}_0(\sigma,b)\sqrt{T}, & \text{case } \textbf{(H1)}\,,\\ \big(\sum_{\alpha:|\alpha|=2}|\partial^\alpha h|_\infty\big)\delta^2[\mathcal{M}_0(\sigma,b)\sqrt{T}]^2, & \text{case } \textbf{(H3)}\,.\end{cases} \tag{8.48}$$

▷**Estimate of** $\mathbb{E}[(h-\hat{h})(X_T^P)\mathbb{1}_{X_T^P\in\widetilde{D}^P}]$**, case (H2) .** We denote by $\mathcal{J}$ the set of integers $\mathbf{j}=(j_1,\dots,j_d)\in\{0,\dots,N\}^d$ such that $\widetilde{D}^{P,j_1,..j_d}$ does not intersect $\bigcup_{i=1}^{N_h}\partial D^i$: for $\mathbf{j}\in\mathcal{J}$, the restriction of h

to $\widetilde{D}^{P,j_1,..j_d}$ coincides with a C^2-function, to which we can apply the estimate (8.47) in the case **(H3)**. Otherwise, we can nevertheless use the estimate in the case **(H1)**. It gives

$$\left|\mathbb{E}[(h-\hat{h})(X_T^P)\mathbb{1}_{X_T^P\in\widetilde{D}^P}]\right| \leq c_0\sum_{\mathbf{j}\in\mathcal{J}}\mathbb{E}\Big(\mu^2\max_{i\in\{1,\dots,N_h\}}\big(\sum_{\alpha:|\alpha|=2}|\partial^\alpha h_i|_\infty\big)\mathbb{1}_{X_T^P\in\widetilde{D}^{P,j_1,..j_d}}\Big)+c_0\sum_{\mathbf{j}\notin\mathcal{J}}\mathbb{E}\Big(\mu\mathrm{C}_{\mathrm{Lip},h}\mathbb{1}_{X_T^P\in\widetilde{D}^{P,j_1,..j_d}}\Big).$$

Moreover, by the definition of μ (see (8.46)), $x\in\widetilde{D}^{P,j_1,..j_d}$ for some $\mathbf{j}\notin\mathcal{J}$ implies that $\mathrm{d}(x,\partial D^i)\leq\mu$ for some i: therefore, $\sum_{\mathbf{j}\notin\mathcal{J}}\mathbb{P}\Big(X_T^P\in\widetilde{D}^{P,j_1,..j_d}\Big)\leq\sum_{i=1}^{N_h}\mathbb{P}(\mathrm{d}(X_T^P,\partial D^i)\leq\mu)$. An application of the following lemma finally leads to

$$\left|\mathbb{E}[(h-\hat{h})(X_T^P)\mathbb{1}_{X_T^P\in\widetilde{D}^P}]\right| \leq_c \Big(\mathrm{C}_{\mathrm{Lip},h}+\max_{i\in\{1,\dots,N_h\}}\big(\sum_{\alpha:|\alpha|=2}|\partial^\alpha h_i|_\infty\big)\Big)\delta^2\mathcal{M}_0(\sigma,b)\sqrt{T}. \tag{8.49}$$

Lemma 8.4.2.1. *Assume $(\mathcal{H}_{x_0}^{\sigma,b})$ and* **(H2)** *, let $\mu=\overline{C}_{\mathcal{V}}\sqrt{d}\delta\mathcal{M}_0(\sigma,b)\sqrt{T}$. Then $\forall i\in\{1,\dots,N_h\}$, we have:*

$$\mathbb{P}(\mathrm{d}(X_T^P,\partial D^i)\leq\mu)\leq_c\delta.$$

Proof. We give here a proof in the particular case where D^i is a half-space of the form $D^i=\{x\in\mathbb{R}^d, x_1>0\}$. Using the bound (8.4) for the density of X_T^p, we have

$$\mathbb{P}(\mathrm{d}(X_T^P,\partial D^i)\leq\mu)\leq_c\int_{\mathbb{R}^d}\mathbb{1}_{-\mu<x_1\leq\mu}\frac{e^{-\frac{|x-m_T^P|^2}{C_{\emptyset,d}[\mathcal{M}_0(\sigma,b)]^2T}}}{(\mathcal{M}_0(\sigma,b)\sqrt{T})^d}\mathrm{d}x\leq_c\int_{-\mu}^{\mu}\frac{e^{-\frac{|x_1-(m_T^P)_1|^2}{C_{\emptyset,d}[\mathcal{M}_0(\sigma,b)]^2T}}}{(\mathcal{M}_0(\sigma,b)\sqrt{T})}\mathrm{d}x_1,$$

the last integral being obviously bounded by $\frac{2\mu}{\mathcal{M}_0(\sigma,b)\sqrt{T}}\leq_c\delta$. For the general case, the idea is to locally map D^i into a half-space by using local charts, so that we are reduced to the first case. We skip these standard computations and refer for instance to [Gobet 2001]. □

▷**Estimate of** $\mathrm{Cor}^{\mathrm{I}}_{2,h}-\mathrm{Cor}^{\mathrm{I}}_{2,\hat{h}}$**.** Using the definition (8.24), the estimates (8.47) of $|\mathbb{1}_{y\in\widetilde{D}^P}(h-\hat{h})(y)|$, the upper bounds (8.4) on $\partial^\alpha f^P$, we easily get:

$$\left|\mathcal{G}_h^{\alpha,\mathrm{I}}-\mathcal{G}_{\hat{h}}^{\alpha,\mathrm{I}}\right|\leq_c\begin{cases}\mathrm{C}_{\mathrm{Lip},h}\delta(\mathcal{M}_0(\sigma,b)\sqrt{T})^{1-|\alpha|}, & \text{cases } \textbf{(H1)} \text{ and } \textbf{(H2)}\,,\\ \big\{\sum_{\alpha:|\alpha|=2}|\partial^\alpha h|_\infty\big\}\delta^2(\mathcal{M}_0(\sigma,b)\sqrt{T})^{2-|\alpha|}, & \text{case } \textbf{(H3)}\,.\end{cases}$$

It readily follows that

$$\left|\mathrm{Cor}^{\mathrm{I}}_{2,h}-\mathrm{Cor}^{\mathrm{I}}_{2,\hat{h}}\right|\leq_c\begin{cases}\mathrm{C}_{\mathrm{Lip},h}\delta(\mathcal{M}_0(\sigma,b)\sqrt{T})^{2}, & \text{cases } \textbf{(H1)} \text{ and } \textbf{(H2)}\,,\\ \big\{\sum_{\alpha:|\alpha|=2}|\partial^\alpha h|_\infty\big\}\delta^2(\mathcal{M}_0(\sigma,b)\sqrt{T})^{3}, & \text{case } \textbf{(H3)}\,.\end{cases}$$

Gathering the above estimate with (8.48)-(8.49) yields the announced result (8.26).

8.5 Numerical experiments

8.5.1 Model and set of parameters

For the numerical tests, we consider the case of d independent one-dimensional elliptic diffusions $(X^i)_{i\in\{1,\dots,d\}}$ driven by their own scalar Brownian motion ($d=q$). We choose the same dynamics for all diffusions: for any $i\in\{1,\dots,d\}$, set

$$\sigma(x)=1-\frac{x}{1+x+x^2},\qquad \mathrm{d}X_t^i=(\mu+\frac{1}{2}\nu^2\sigma^{(1)}(X_t^i))\sigma(X_t^i)\mathrm{d}t+\nu\sigma(X_t^i)\mathrm{d}W_t^i,\qquad X_0^i=x_0.$$

We easily check that σ takes values in $[\frac{2}{3}, 2]$ and is of class C^∞ with bounded derivatives, the first derivative vanishing at ± 1. A key feature of this diffusion model is that (owing to the Lamperti transform, see [Karatzas 1991, p. 294-295]) $X_t^i = g(f(x_0) + \mu t + \nu W_t^i)$ where g is the inverse function of

$$f(x) = \int_0^x \frac{\mathrm{d}y}{\sigma(y)} = x + \frac{1}{2}\log(1 + x^2).$$

Thus, using numerical inversions of f to evaluate g, we can exactly simulate X_T without discretization scheme to get our reference values. We consider four types of terminal functions from the most regular case to the less regular one:

$$\begin{aligned}
h_1(x) &= \frac{100 e^{\frac{1}{d}\sum_{i=1}^d x_i}}{1 + e^{\frac{1}{d}\sum_{i=1}^d x_i}} \ (C^\infty \text{ with bounded derivatives, case } \textbf{(H3)}\text{),}\\
h_2(x) &= 100 e^{\frac{1}{d}\sum_{i=1}^d x_i} \ (C^\infty\text{, case } \textbf{(H3)} \text{ with unbounded derivatives),}\\
h_3(x) &= \frac{100}{d}(\sum_{i=1}^d x_i)_+ \ \text{(Lipschitz, case } \textbf{(H2)}\text{),}\\
h_4(x) &= 100 \max(x_1, \dots, x_d) \quad \text{(Lipschitz, case } \textbf{(H1)}\text{).}
\end{aligned}$$

Besides, we investigate three sets of parameters:

($\mathcal{P}_1$): $x_0 = 0$, $\nu = 20\%$, $\mu = 0$ (standard situation);

($\mathcal{P}_2$): $x_0 = 0$, $\nu = \mu = 20\%$ (large drift);

($\mathcal{P}_3$): $x_0 = 1$, $\nu = 20\%$, $\mu = 0$, (small magnitude and variations of σ).

In all the numerical tests, we set $T = 1$ and the purpose is to compare the following approximations:

1. `SAFE Lin` **(H1)** - **(H2)** - **(H3)** : the SAFE method using multilinear finite elements (see Theorem 8.2.2.1) and fixing its parameters as follows (according to Theorems 8.2.2.2-8.2.3.1): R = 4 and $\delta = [\nu\sigma(x_0)]^2$, $\delta = \nu\sigma(x_0)$ and $\delta = [\nu\sigma(x_0)]^{\frac{1}{2}}$ for respectively **(H1)** - **(H2)** - **(H3)** .

2. `SAFE Quad 1-2-3-4-5`: the SAFE method using multiquadratic finite elements (see Theorem 8.2.4.1) using R = 4 and $\delta = 0.5$, $\delta = 1$, $\delta = 2$, $\delta = 4$ and $\delta = 8$ for respectively `1-2-3-4-5`. The methods `SAFE Quad 4-5` using large steps are only investigated in dimension 10.

3. `MC`: the estimation by Monte-Carlo simulations without discretization scheme. Keep in mind that the execution times reported in the test tables should be multiplied by about a factor 100 to take into account the usual time discretization effort. In all the tests we use 10^7 sample paths. The value in parentheses is the half-width of the related 95%-symmetric confidence interval.

4. `MC Proxy`: the estimation by Monte-Carlo simulations of the stochastic approximation formula (8.10) (see Theorem (8.2.1.2)). In all the tests we use 10^7 sample paths. We provide additional tests to show the efficiency of this method (which is always faster than `MC`) in high dimensions (greater than 20 and up to 100) for which the SAFE methods seem too slow to be applied.

8.5.2 Results

We study both the accuracy and the speed of the `SAFE` method in comparison to Monte-Carlo simulations up to dimension 10. For high dimensions (from 20 to 100) we show the accuracy of `MC Proxy` which is

Table 8.1: Estimation of the expectations in dimension 2 with MC, MC proxy, SAFE methods (set of parameters ($\mathcal{P}_1$)) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	49.47	98.99	4.89	9.70	1m2s
	(±2.3E−2)	(±9.0E−3)	(±4.5E−3)	(±9.5E−3)	
`MC Proxy`	49.49	99.01	4.84	9.67	43s
	(±1.1E−2)	(±2.3E−2)	(±5.5E−3)	(±1.1E−2)	
`SAFE Lin` **(H1)**	49.49	99.02	4.85	9.70	$< 10^{-4}$s
`SAFE Lin` **(H2)**	49.49	99.02	4.86	9.71	$< 10^{-4}$s
`SAFE Lin` **(H3)**	49.49	99.05	4.97	9.79	$< 10^{-4}$s
`SAFE Quad 1`	49.49	99.02	4.85	9.66	$< 10^{-4}$s
`SAFE Quad 2`	49.49	99.01	4.87	9.55	$< 10^{-4}$s
`SAFE Quad 3`	49.49	99.01	4.86	9.24	$< 10^{-4}$s

Table 8.2: Estimation of the expectations in dimension 3 with MC, MC proxy, SAFE methods (set of parameters ($\mathcal{P}_1$)) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	49.47	98.62	3.84	15.01	1m40s
	(±1.9E−3)	(±7.4E−3)	(±3.7E−3)	(±8.2E−3)	
`MC Proxy`	49.48	98.64	3.78	14.98	1m12s
	(±1.3E−2)	(±2.7E−2)	(±4.6E−3)	(±1.1E−2)	
`SAFE Lin` **(H1)**	49.48	98.66	3.80	15.02	16s
`SAFE Lin` **(H2)**	49.48	98.67	3.81	15.04	0.1s
`SAFE Lin` **(H3)**	49.49	98.69	3.87	15.17	0.02s
`SAFE Quad 1`	49.48	98.66	3.79	14.96	0.08s
`SAFE Quad 2`	49.48	98.66	3.81	14.78	0.02s
`SAFE Quad 3`	49.49	98.66	3.85	14.18	$< 10^{-4}$s

always faster by a factor 100 than `MC` whatever is the dimension. In Tables 8.1-8.2-8.3-8.4-8.5-8.6-8.7-8.8-8.9-8.10-8.11-8.12-8.13, we report the results in dimensions 2-3-4-5-6-7-8-9-10-20-30-50-100, with execution times, for all methods and all terminal functions using the set of parameters ($\mathcal{P}_1$). In Tables 8.14-8.15 we give the results in dimension 4 using the sets of parameters ($\mathcal{P}_2$) and ($\mathcal{P}_3$). We finally plot in Figure 8.1 the relative errors (compared to the `MC Proxy` estimator) given by `SAFE Lin` and `SAFE Quad` w.r.t. $\log(N)$, using R = 5, in dimension 4, for the four terminal functions. All these computations have been coded in C++ on a Intel(R) Core(TM) i5 CPU@2.40GHz with 4 GB of ram.

▷**Influence of the `SAFE` method.** Regarding the Tables 8.1-8.2-8.3-8.4-8.5, first we notice that the values obtained by `MC` and `MC Proxy` are close to each other (error smaller than 1% in absolute value) which indicates a very good accuracy of the stochastic approximation of Theorems 8.2.1.1 and 8.2.1.2. The mesh size δ is larger for `SAFE Quad` than for `SAFE Lin`; nevertheless we remark that `SAFE Lin` and `Quad` give close results and give good deterministic approximations of the values obtained by `MC Proxy` (the value towards which we expect convergence) and `MC` (the target value). Their relative errors in comparison to `MC Proxy` are often smaller than 1% in absolute value.

Table 8.3: Estimation of the expectations in dimension 4 with MC, MC proxy, SAFE methods (set of parameters ($\mathcal{P}_1$)) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
MC	49.47	98.45	3.22	18.30	2m1s
	(±1.7E−3)	(±6.4E−3)	(±3.2E−3)	(±7.4E−3)	
MC Proxy	49.49	98.50	3.16	18.25	1m23s
	(±1.5E−2)	(±3.1E−2)	(±4.0E−3)	(±1.1E−2)	
SAFE Lin **(H1)**	49.48	98.48	3.17	18.28	1h16m
SAFE Lin **(H2)**	49.48	98.49	3.18	18.31	7s
SAFE Lin **(H3)**	49.48	98.50	3.23	18.47	0.3s
SAFE Quad 1	49.48	98.48	3.17	18.20	3s
SAFE Quad 2	49.48	98.48	3.17	17.98	0.2s
SAFE Quad 3	49.48	98.49	3.22	16.92	0.02s

Table 8.4: Estimation of the expectations in dimension 5 with MC, MC proxy and SAFE methods (set of parameters ($\mathcal{P}_1$)) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
MC	49.47	98.34	2.78	20.63	2m38s
	(±1.5E−3)	(±5.8E−3)	(±2.8E−3)	(±6.9E−3)	
MC Proxy	49.49	98.38	2.74	20.53	1m51s
	(±1.7E−2)	(±3.4E−2)	(±3.6E−3)	(±1.2E−2)	
SAFE Lin **(H2)**	49.47	98.37	2.76	20.60	4m44s
SAFE Lin **(H3)**	49.48	98.39	2.81	20.79	4s
SAFE Quad 1	49.47	98.37	2.74	20.47	2m2s
SAFE Quad 2	49.48	98.37	2.73	20.21	2s
SAFE Quad 3	49.48	98.38	2.78	18.62	0.5s

Table 8.5: Estimation of the expectations in dimension 6 with MC, MC proxy and SAFE methods (set of parameters ($\mathcal{P}_1$)) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
MC	49.47	98.27	2.47	22.42	2m58s
	(±1.3E−3)	(±5.3E−3)	(±2.5E−3)	(±6.6E−3)	
MC Proxy	49.47	98.29	2.43	22.25	2m2s
	(±1.9E−2)	(±3.7E−2)	(±3.3E−3)	(±1.3E−2)	
SAFE Lin **(H2)**	49.47	98.30	2.44	22.33	4h56m
SAFE Lin **(H3)**	49.48	98.31	2.49	22.54	2m7s
SAFE Quad 1	49.47	98.29	2.43	22.19	1h30m
SAFE Quad 2	49.48	98.30	2.43	21.90	1m30s
SAFE Quad 3	49.48	98.31	2.46	19.77	2s

Table 8.6: Estimation of the expectations in dimension 7 with MC, MC proxy and SAFE methods (set of parameters $(\mathcal{P}_1)$) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	49.47	98.22	2.23	23.85	3m38s
	(±1.3E−3)	(±4.9E−3)	(±2.3E−3)	(±6.3E−3)	
`MC Proxy`	49.46	98.22	2.18	23.61	2m36s
	(±2.0E−2)	(±4.0E−2)	(±3.1E−3)	(±1.4E−2)	
`SAFE Lin` **(H3)**	49.47	98.26	2.24	23.93	47m7s
`SAFE Quad 2`	49.47	98.24	2.19	23.24	34m22s
`SAFE Quad 3`	49.48	98.25	2.20	20.62	21s

Table 8.7: Estimation of the expectations in dimension 8 with MC, MC proxy and SAFE methods (set of parameters $(\mathcal{P}_1)$) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	49.47	98.17	2.03	25.04	3m57s
	(±1.2E−3)	(±4.6E−3)	(±2.1E−3)	(±6.1E−3)	
`MC Proxy`	49.46	98.18	1.99	24.74	2m41s
	(±2.2E−2)	(±4.3E−2)	(±2.9E−3)	(±1.5E−2)	
`SAFE Quad 3`	49.48	98.21	2.00	21.29	3m39s

Table 8.8: Estimation of the expectations in dimension 9 with MC, MC proxy and SAFE methods (set of parameters $(\mathcal{P}_1)$) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	49.47	98.14	1.87	26.05	4m40s
	(±2.0E−3)	(±4.3E−3)	(±2.1E−3)	(±5.9E−3)	
`MC Proxy`	49.46	98.15	1.83	25.69	3m6s
	(±2.3E−2)	(±4.6E−2)	(±2.8E−3)	(±1.7E−2)	
`SAFE Quad 3`	49.48	98.18	1.83	21.85	36m25s

▷**Influence of the number of points in `SAFE` methods.** We observe in Tables 8.1-8.2-8.3-8.4-8.5 that as expected, with fewer points (`SAFE Lin` **(H3)** or `Quad 3`), we globally lose accuracy.

▷**Influence of the dimension and the terminal function.** Generally speaking, for h_1,h_2,h_3 the accuracy is very good whatever the dimension is. For h_4, results get worse as d increases. We notice in Tables 8.5-8.6-8.7-8.8-8.9-8.10-8.11-8.12-8.13 that in high dimension, the `SAFE` methods and `MC Proxy` give more accurate results for h_3 (probably because the average of the r.v. induces smaller fluctuations) but less accurate for h_4 (certainly due to the regularity breakdown of the function max in high-dimension).

▷**Speed results.** Regarding the execution times, we notice that in dimension 2, the use of `SAFE` is almost instantaneous versus 1 or 2 minutes for accurate Monte-Carlo simulations. For the dimension 3, all the method take less than 1s except `SAFE Lin` **(H1)** which need 16s what remains very quick.

Table 8.9: Estimation of the expectations in dimension 10 with MC, MC proxy and SAFE methods (set of parameters ($\mathcal{P}_1$)) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	49.47	98.12	1.73	26.93	4m50s
	(±1.0E−3)	(±4.1E−3)	(±1.9E−3)	(±5.8E−3)	
`MC Proxy`	49.49	98.18	1.70	26.52	3m15s
	(±2.4E−2)	(±4.8E−2)	(±2.7E−3)	(±1.8E−2)	
`SAFE Quad 3`	49.47	98.15	1.69	22.35	5h49m
`SAFE Quad 4`	49.48	98.16	1.82	13.32	1m
`SAFE Quad 5`	49.48	98.17	1.60	21.05	0.39s

Table 8.10: Estimation of the expectations in dimension 20 with MC, MC proxy and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	49.47	98.00	1.01	32.18	9m35s
	(±7.4E−4)	(±2.9E−3)	(±1.2E−3)	(±5.0E−3)	
`MC Proxy`	49.50	98.10	0.99	31.28	6m33s
	(±3.4E−2)	(±6.8E−2)	(±1.8E−3)	(±2.8E−2)	

Table 8.11: Estimation of the expectations in dimension 30 with MC, MC proxy and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	49.47	97.97	0.70	34.91	14m38s
	(±6.0E−4)	(±2.3E−3)	(±9.2E−4)	(±4.7E−3)	
`MC Proxy`	49.49	98.03	0.69	33.59	9m52s
	(±4.2E−2)	(±8.3E−2)	(±1.7E−3)	(±3.6E−2)	

Table 8.12: Estimation of the expectations in dimension 50 with MC, MC proxy and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	49.47	97.94	0.42	38.07	23m50s
	(±4.7E−4)	(±1.8E−3)	(±6.3E−4)	(±4.3E−3)	
`MC Proxy`	49.48	98.00	0.41	36.17	16m12s
	(±5.4E−2)	(±1.1E−1)	(±1.3E−3)	(±5.0E−2)	

In dimension 4, it takes 1h16 for `SAFE Lin` **(H1)**, 7s for `SAFE Lin` **(H2)** and less for the other `SAFE` methods. Even if 1h16 seems to be quite important, it is still faster than `MC` (taking into account the discretization effort) but slower than `MC Proxy`. We notice that in dimension 5 `SAFE Lin` **(H2)** and `Quad 1` need 4m44s and 2m2s, a performance which is close to `MC Proxy`. In dimension 6, `SAFE Lin` **(H2)** takes almost 5h which is comparable to `MC`. `SAFE Quad 1` takes 1h30 but `SAFE Lin` **(H3)** (2m7s), `Quad 2` and `3` (1m30s ands 2s) are very competitive. In dimension 7, `SAFE Lin` **(H3)**,

Table 8.13: Estimation of the expectations in dimension 100 with MC, MC proxy and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	49.47	97.92	0.17	41.94	47m37s
	(±3.3E−4)	(±1.3E−3)	(±3.3E−4)	(±3.9E−3)	
`MC Proxy`	49.46	97.94	0.17	39.12	33m10s
	(±7.7E−2)	(±1.5E−1)	(±9.0E−4)	(±7.8E−2)	

Table 8.14: Estimation of the expectations in dimension 4 with MC, MC proxy, SAFE methods (set of parameters ($\mathcal{P}_2$)) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	54.26	119.15	17.27	34.55	2m1s
	(±1.4E−3)	(±6.5E−3)	(±5.3E−3)	(±6.7E−3)	
`MC Proxy`	54.15	118.56	16.86	32.91	1m23s
	(±2.4E−4)	(±5.5E−2)	(±1.3E−2)	(±2.4E−2)	
`SAFE Lin` **(H1)**	54.20	118.70	16.91	33.01	1h16m
`SAFE Lin` **(H2)**	54.19	118.68	16.90	33.04	7s
`SAFE Lin` **(H3)**	54.18	118.68	16.91	33.21	0.3s
`SAFE Quad 1`	54.19	118.68	16.90	32.91	3s
`SAFE Quad 2`	54.19	118.66	16.89	32.63	0.2s
`SAFE Quad 3`	54.18	118.64	16.88	31.62	0.02s

Table 8.15: Estimation of the expectations in dimension 4 with MC, MC proxy, SAFE methods (set of parameters ($\mathcal{P}_3$)) and execution time.

method / function	h_1	h_2	h_3	h_4	exec. time
`MC`	73.08	272.43	100.00	113.76	2m1s
	(±8.1E−4)	(±1.1E−2)	(±4.1E−3)	(±5.8E−3)	
`MC Proxy`	73.09	272.44	100.00	113.75	1m23s
	(±8.3E−4)	(±1.1E−2)	(±4.2E−3)	(±6.0E−3)	
`SAFE Lin` **(H1)**	73.07	272.37	99.97	113.71	7h10m
`SAFE Lin` **(H2)**	73.07	272.38	99.98	113.73	35s
`SAFE Lin` **(H3)**	73.08	272.41	99.99	113.81	0.6s
`SAFE Quad 1`	73.08	272.41	99.99	113.67	3s
`SAFE Quad 2`	73.08	272.40	99.99	113.54	0.2s
`SAFE Quad 3`	73.07	272.39	99.98	113.04	0.02s

`Quad 2` and `3` respectively necessitate 47m, 34m and 21s. In higher dimension (≥ 8), only `SAFE Quad 3` is competitive with very satisfying results except for the max terminal function. It takes 3m39s in dimension 8 (close to `MC Proxy`), about 36m in dimension 9 and almost 6h in dimension 10 (comparable to `MC`). Of course, this huge amount of computations can be easily split among parallel processors: this is a nice feature of the SAFE method. Also, the efficiency may be much improved by using sparse grids

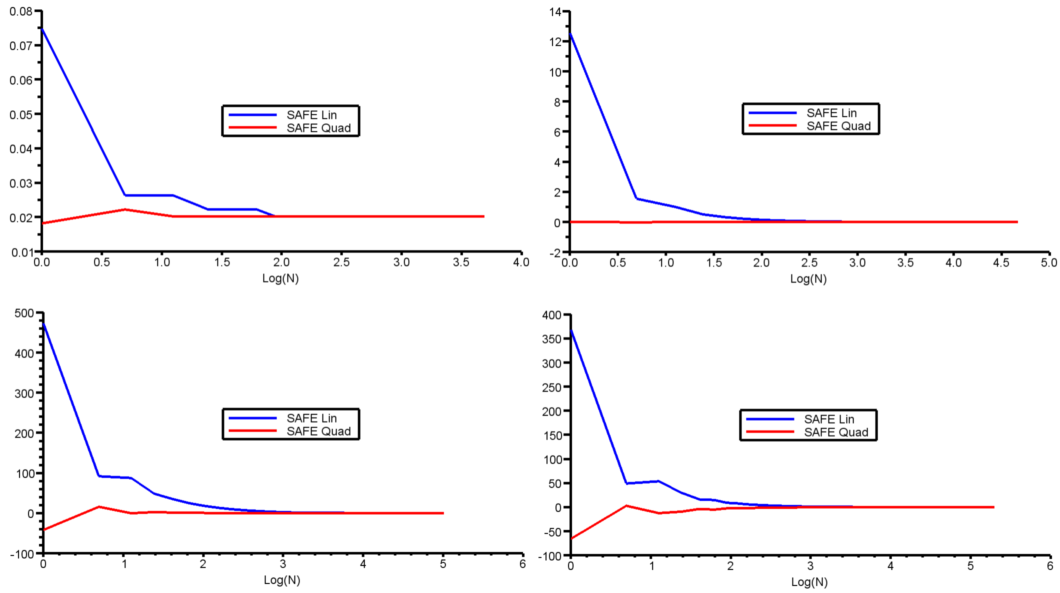

Figure 8.1: Relative errors (in %) in comparison to `MC Proxy` of `SAFE Lin` and `Quad` w.r.t. $\log(N)$ in dimension 4 (from left to right and from top to bottom: terminal functions h_1, h_2, h_3 and h_4).

[Bungartz 2004], this is left to further investigation. With very few points, `SAFE Quad 4-5` give in dimension 10 very good results for h_1 and h_2, good results for h_3 (relative errors of order 5%) and pretty poor results with h_4, for execution times of respectively 1m and less than 1s. This shows that for smooth functions, we may reach very high accuracies with few points.
When the dimension becomes high (≥ 20) the `SAFE` methodology is probably less competitive and we may prefer to use `MC Proxy` which presents a speed gain by a factor 100 in comparison to `MC` due to the lack of discretization. Besides the accuracy is very good for the terminal functions h_1, h_2 and h_3 and remains satisfying for h_4 (relative errors for h_4 are or order -3%,-4%,-5% and -7% for respectively dimensions 20, 30, 50 and 100). The calculations are performed in approximately 6m, 10m, 16m and 33m in respectively dimensions 20, 30, 50 and 100 what is promising.
As a conclusion, for small and medium dimensions (≤ 10) and enough smooth functions the use of `SAFE` is very efficient and for high dimensions (> 10), the utilization of `MC Proxy` is a good alternative which gives rise to very good approximations and to always faster calculations than `MC` whatever is the dimension.

▷**Convergence results of `SAFE`.** We notice from Figure 8.1 that the convergence is quite fast for both `SAFE Lin` and `Quad`. Moreover the more regular the terminal function, the faster the convergence. For large regularity of h, only few points are needed to achieve convergence. On the contrary for the function max, the convergence speeds are similar (taking into account that the computational cost at fixed N is higher for `SAFE Quad`).

▷**Influence of the drift and the initial point.** Table 8.14 (parameters ($\mathcal{P}_2$)) shows that the accuracy is worsened as the drift gets larger, inaccuracies increasing with the irregularity of the terminal function. This transport term probably increases the deviations of the diffusion and worsens the accuracy of the proxy. On the contrary, considering an initial point at 1 which induces for σ small variations and magnitude leads to better results as presented in Table 8.15 (parameters ($\mathcal{P}_3$)), notably for the function h_3. All these observations are coherent with our error estimates in Theorems 8.2.1.1-8.2.3.1-8.2.4.1.

8.6 Appendix

8.6.1 Computation of the correction terms in Theorems 8.2.1.1 and representation as sensitivities

We follow the routine of [Gobet 2012a], with some adaptations due the multidimensional setting. We first provide integration by parts formulas useful for the explicit derivation of the correction terms. In the following, $c_t : [0,T] \to \mathbb{R}^{1\times q}$, $C_t : [0,T] \to \mathbb{R}^{d\times q}$ are square integrable and predictable processes, $A_t : [0,T] \to \mathbb{R}^{d\times q}$ is a square integrable and deterministic process and $\psi, \psi_1, \dots, \psi_d : \mathbb{R}^d \to \mathbb{R}$ are smooth functions with bounded derivatives.

Lemma 8.6.1.1. *One has:*

$$\mathbb{E}[\psi(\int_0^T A_t\mathrm{d}W_t)\int_0^T c_t\mathrm{d}W_t] = \mathbb{E}[\nabla\psi(\int_0^T A_t\mathrm{d}W_t)\int_0^T A_t c_t^*\mathrm{d}t], \tag{8.50}$$

$$\mathbb{E}[(\psi_1(\int_0^T A_t\mathrm{d}W_t),\dots,\psi_\mathrm{d}(\int_0^T A_t\mathrm{d}W_t))\int_0^T C_t\mathrm{d}W_t] = \sum_{i,j=1}^d \mathbb{E}[\partial_{x_j}\psi_i(\int_0^T A_t\mathrm{d}W_t)\int_0^T (A_tC_t^*)_i^j\mathrm{d}t]. \tag{8.51}$$

Proof. We begin with (8.50). The process A being deterministic, the Malliavin derivative $D_s\int_0^T A_t\mathrm{d}W_t$ is equal to $A_s\mathbb{1}_{s\le T}$ and $\psi(\int_0^T A_t\mathrm{d}W_t)\in\mathbb{D}^{1,\infty}$ with $D_s[\psi(\int_0^T A_t\mathrm{d}W_t)] = \psi'(\int_0^T A_t\mathrm{d}W_t)A_s\mathbb{1}_{s\le T}$. Then identify $\int_0^T c_t\mathrm{d}W_t$ with the Skorohod operator and apply the duality relationship [Nualart 2006, Definition 1.3.1 and Proposition 1.3.11]. To derive (8.51), apply (8.50) with $c_t = C_t^i$ for any $i\in\{1,\dots,d\}$:

$$\begin{aligned}\mathbb{E}[(\psi_1(\int_0^T A_t\mathrm{d}W_t),\dots,\psi_\mathrm{d}(\int_0^T A_t\mathrm{d}W_t))\int_0^T C_t\mathrm{d}W_t] &= \sum_{i=1}^d \mathbb{E}[\nabla\psi_i(\int_0^T A_t\mathrm{d}W_t)\int_0^T A_t(C_t^i)^*\mathrm{d}t]\\ &= \sum_{i,j=1}^d \mathbb{E}[\partial_{x_j}\psi_i(\int_0^T A_t\mathrm{d}W_t)\int_0^T (A_tC_t^*)_i^j\mathrm{d}t].\end{aligned}$$

□

We are now in a position to prove:

Proposition 8.6.1.1. *Assume $(\mathcal{H}_{x_0}^{\sigma,b})$. For any smooth function $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ with bounded derivatives, we have:*

$$\mathbb{E}[\nabla\phi(X_T^P)\dot{X}_T] = \mathrm{Cor}_{2,\phi},$$

where $\mathrm{Cor}_{2,\phi}$ is defined in (8.11).

Proof. In the proof, we repeatedly use the identity $\mathbb{E}(\partial^\alpha\phi(X_T^P)) = \partial^\alpha\overline{\phi}^P(0)$ for any multi-index α. In view of (8.9) and (8.8), one has to transform $\mathbb{E}[\nabla\phi(X_T^P)\dot{X}_T] = I_1 + I_2$ with

$$I_1 := \mathbb{E}[\nabla\phi(X_T^P)\int_0^T b_t'(X_t^P - x_0)\mathrm{d}t], \qquad I_2 := \mathbb{E}[\nabla\phi(X_T^P)\int_0^T \sum_{j=1}^d \sigma_{j,t}'(X_t^P - x_0)\mathrm{d}W_t^j].$$

We begin with I_1. Writing $X_t^P - x_0 = \int_0^t \sigma_s\mathrm{d}W_s + \int_0^t b_s\mathrm{d}s$, we obtain $I_1 = I_{1a} + I_{1b}$ with

$$I_{1a} := \mathbb{E}[\nabla\phi(X_T^P)\int_0^T b_t'(\int_0^t b_s\mathrm{d}s)\mathrm{d}t] = \nabla\overline{\phi}^P(0)\int_0^T b_t'(\int_0^t b_s\mathrm{d}s)\mathrm{d}t,$$

$$I_{1b} := \mathbb{E}[\nabla\phi(X_T^P)\int_0^T b_t'(\int_0^t \sigma_s \mathrm{d}W_s)\mathrm{d}t] = \int_0^T \mathbb{E}[\nabla\phi(X_T^P)(\int_0^t b_t'\sigma_s \mathrm{d}W_s)]\mathrm{d}t.$$

The contribution I_{1a} is in its final form. Regarding I_{1b}, for any $t \in [0,T]$ apply formula (8.51) with $(\psi_1,\dots,\psi_d)(x) = \nabla\phi(m_T^P + x)$, $(A_s)_{s\in[0,T]} = (\sigma_s)_{s\in[0,T]}$ and $(C_s)_{s\in[0,T]} = (\mathbb{1}_{s\le t} b_t'\sigma_s)_{s\in[0,T]}$ to directly obtain:

$$I_{1b} = \mathbb{E}[\sum_{i,j=1}^d \partial^2_{x_i,x_j}\phi(X_T^P)\int_0^T (\int_0^t \sigma_s\sigma_s^*(b_t')^*\mathrm{d}s)_i^j \mathrm{d}t] = \sum_{i,j=1}^d \partial^2_{\epsilon_i,\epsilon_j}\overline{\phi}^P(0)\int_0^T (b_t^i)'(\int_0^t \Sigma_{j,s}\mathrm{d}s)\mathrm{d}t.$$

We now handle the term I_2. Let D_t be the d-dimensional square matrix whose j-th column is equal to $\sigma'_{j,t}\,(X_t^P - x_0)$: using again (8.51), we derive

$$\begin{aligned}
I_2 =& \mathbb{E}[\nabla\phi(X_T^P)\int_0^T D_t \mathrm{d}W_t] = \mathbb{E}[\sum_{i,j=1}^d \partial^2_{x_i,x_j}\phi(X_T^P)\int_0^T (\sigma_t D_t^*)_i^j \mathrm{d}t]\\
=& \mathbb{E}[\sum_{i,j,k=1}^d \partial^2_{x_i,x_j}\phi(X_T^P)\int_0^T \sigma^j_{k,t}D^i_{k,t}\mathrm{d}t] = \mathbb{E}[\sum_{i,j,k,l=1}^d \partial^2_{x_i,x_j}\phi(X_T^P)\int_0^T [\sigma^j_k\partial_{x_l}\sigma^i_k](t,x_0)(X_t^P - x_0)^l \mathrm{d}t].
\end{aligned}$$

Thanks to the symmetry of the Hessian matrix $H(\phi)$, we also have

$$\begin{aligned}
I_2 =& \frac{1}{2}\mathbb{E}[\sum_{i,j,k,l=1}^d \partial^2_{x_i,x_j}\phi(X_T^P)\int_0^T [\sigma^i_k\partial_{x_l}\sigma^j_k + \sigma^j_k\partial_{x_l}\sigma^i_k](t,x_0)(X_t^P - x_0)^l \mathrm{d}t]\\
=& \frac{1}{2}\mathbb{E}[\sum_{i,j,l=1}^d \partial^2_{x_i,x_j}\phi(X_T^P)\int_0^T \partial_{x_l}\Sigma^i_j(t,x_0)(X_t^P - x_0)^l \mathrm{d}t] = \frac{1}{2}\mathbb{E}[\sum_{i,j=1}^d \partial^2_{x_i,x_j}\phi(X_T^P)\int_0^T (\Sigma^i_{j,t})'(X_t^P - x_0)\mathrm{d}t].
\end{aligned}$$

From this point, the computations are similar to those for I_1; briefly, writing $X_t^P - x_0 = \int_0^t \sigma_s \mathrm{d}W_s + \int_0^t b_s \mathrm{d}s$, we decompose $I_2 = \frac{1}{2}(I_{2a} + I_{2b})$ with

$$\begin{aligned}
I_{2a} :=& \mathbb{E}[\sum_{i,j=1}^d \partial^2_{x_i,x_j}\phi(X_T^P)\int_0^T (\Sigma^i_{j,t})'(\int_0^t b_s\mathrm{d}s)\mathrm{d}t] = \sum_{i,j=1}^d \partial^2_{\epsilon_i,\epsilon_j}\overline{\phi}^P(0)\int_0^T (\Sigma^i_{j,t})'(\int_0^t b_s\mathrm{d}s)\mathrm{d}t,\\
I_{2b} :=& \mathbb{E}[\sum_{i,j=1}^d \partial^2_{x_i,x_j}\phi(X_T^P)\int_0^T (\Sigma^i_{j,t})'(\int_0^t \sigma_s\mathrm{d}W_s)\mathrm{d}t] = \int_0^T \mathbb{E}[\sum_{i,j=1}^d \partial^2_{x_i,x_j}\phi(X_T^P)(\int_0^t (\Sigma^i_{j,t})'\sigma_s\mathrm{d}W_s)]\mathrm{d}t.
\end{aligned}$$

Then use (8.50) and the symmetry of Σ_t in order to get

$$\begin{aligned}
I_{2b} =& \int_0^T \mathbb{E}[\sum_{i,j=1}^d \nabla(\partial^2_{x_i,x_j}\phi)(X_T^P)(\int_0^t \sigma_s[(\Sigma^i_{j,t})'\sigma_s]^*\mathrm{d}s)]\mathrm{d}t\\
=& \mathbb{E}[\sum_{i,j=1}^d \nabla(\partial^2_{x_i,x_j}\phi)(X_T^P)]\int_0^T (\int_0^t \Sigma_s\mathrm{d}s)((\Sigma^i_{j,t})')^*\mathrm{d}t\\
=& \sum_{i,j,k=1}^d \partial^3_{\epsilon_i,\epsilon_j,\epsilon_k}\overline{\phi}^P(0)[\int_0^T (\int_0^t \Sigma_s\mathrm{d}s)((\Sigma^i_{j,t})')^*\mathrm{d}t]^k = \sum_{i,j,k=1}^d \partial^3_{\epsilon_i,\epsilon_j,\epsilon_k}\overline{\phi}^P(0)\int_0^T (\Sigma^i_{j,t})'(\int_0^t \Sigma_{k,s}\mathrm{d}s)\mathrm{d}t.
\end{aligned}$$

Gathering all the contributions achieves the proof. □

CHAPTER 9

Price approximations in multidimensional CEV models using the SAFE methods

The purpose of this Chapter is to present additional numerical tests for the pricing of multi-asset products (Basket, Geometrical mean, Worst of and Best of Put options) in multidimensional CEV models using the SAFE methods. We apply the results of the previous Chapter 8.

Model and set of parameters. We consider d independent one-dimensional diffusions $(X^i)_{i\in\{1,\dots,d\}}$ modelling the log-assets defined by the SDEs:

$$dX_t^i = \sigma(X_t^i)[dW_t^i - \frac{1}{2}\sigma(X_t^i)dt], \quad X_0^i = x_0^i = \log(100), \tag{9.1}$$

with $\sigma(x) = 0.2e^{-0.2(x-\log(100))}$. We are interesting by the pricing of multi-asset Put options with payoffs:

$$(K - \frac{1}{d}\sum_{i=1}^{d}\exp(x_i))_+ \ \texttt{(Basket)}, \qquad (K - \exp(\frac{1}{d}\sum_{i=1}^{d} x_i))_+ \ \texttt{(Geo. mean)}$$
$$(K - \min_{i\in\{1,..,d\}}\exp(x_i))_+ \ \texttt{(Worst of)}, \qquad (K - \max_{i\in\{1,..,d\}}\exp(x_i))_+ \ \texttt{(Best of)},$$

where the strike K takes the values 90,95,100,105 and 110. In all the tests, we set $T = 1$ and the aim of the numerical experiments is to compare the following approximations:

1. `SAFE Lin` **(H1) - (H2) - (H3)** : the SAFE method using multilinear finite elements (see Theorem 8.2.2.1) and fixing its parameters as follows (according to Theorems 8.2.2.2-8.2.3.1): R = 4 and $\delta = (0.2)^2$, $\delta = 0.2$ and $\delta = (0.2)^{\frac{1}{2}}$ for respectively **(H1) - (H2) - (H3)** .

2. `SAFE Quad 1-2-3-4-5`: the SAFE method using multiquadratic finite elements (see Theorem 8.2.4.1) using R = 4 and $\delta = 0.5$, $\delta = 1$, $\delta = 2$, $\delta = 4$ and $\delta = 8$ for respectively `1-2-3-4-5`. The methods `SAFE Quad 4-5` using large steps are only investigated in dimension 10.

3. `MC` : the estimation by Monte-Carlo simulations using in all the tests a time discretization of 300 steps by year and 10^7 sample paths. The value in parentheses is the half-width of the related 95%-symmetric confidence interval. Note that in dimension 1, one can use the closed-form formula (see [Schroder 1989]) denoted by `True Price`.

4. `MC Proxy`: the estimation by Monte-Carlo simulations of the stochastic approximation formula (8.10) (see Theorem (8.2.1.2)). In all the tests we use 10^7 sample paths. In dimension 1, the stochastic approximation formula (8.10) is explicit and we denote its calculus by `Proxy Price`.

Table 9.1: Estimation of the Put prices with the closed-form formula, the proxy method, SAFE method and execution time.

payoff	method / strikes	90	95	100	105	110	exec. time
`Put`	`True Price`	3.66	5.56	7.97	10.87	14.22	$< 10^{-4}$s
	`Proxy Price`	3.66	5.56	7.97	10.87	14.22	$< 10^{-4}$s
	`SAFE Lin (H1)`	3.66	5.56	7.96	10.86	14.22	$< 10^{-4}$s
	`SAFE Lin (H2)`	3.66	5.55	7.93	10.85	14.21	$< 10^{-4}$s
	`SAFE Lin (H3)`	3.80	5.65	8.11	10.98	14.34	$< 10^{-4}$s
	`SAFE Quad 1`	3.68	5.56	7.99	10.87	14.25	$< 10^{-4}$s
	`SAFE Quad 2`	3.55	5.67	8.06	10.87	14.06	$< 10^{-4}$s
	`SAFE Quad 3`	4.13	5.88	8.21	11.40	14.59	$< 10^{-4}$s

Results. In Tables 9.1-9.2-9.3-9.4-9.5-9.6-9.7-9.8-9.9-9.10, we report the results in dimensions 1-2-3-4-5-6-7-8-9-10, with execution times, for all methods, all payoffs and all strikes. For all these computations, we used C++ on a Intel(R) Core(TM) i5 CPU@2.40GHz with 4 GB of ram.

▷**Influence of the method and of the number of points.** We notice on Tables 9.2-9.3-9.4-9.5-9.6 that the values obtained by `MC` and `MC Proxy` are very close to each other whatever the strike is (1 or 2 bps of error) and that all the SAFE methods provide results very close to `MC Proxy`. As expected, the methods with more points give better results. It seems that globally the results are a little bit better than those obtained for the elliptic diffusion in Chapter 8.

▷**Influence of the dimension, the payoff and the strikes.** We remark on Table 9.1 that in dimension 1, `Proxy Price` and `True Price` are extremely close with errors smaller than 1 bp. The best SAFE methods (`SAFE Lin (H1)` -`SAFE Lin (H2)` -`SAFE Quad 1`) yield to error reduced to few bps. For the other dimensions, the `Basket` and the `Geo. Mean` payoffs present very good accuracy whatever the strike is. The results for the more irregular payoffs (`Worst of` and `Best of`) get worse as d increases.

▷**Speed results.** We retrieve the same results that those obtained with the elliptic diffusion in Chapter 8 except that all the execution times are a little bit more important because there are 5 strikes to compute per payoff (and for `MC`, we have to use discretization which greatly increases the execution time):

- `SAFE Lin (H1)` is more competitive than `MC Proxy` up to the dimension 3 and than `MC` up to the dimension 4,
- `SAFE Lin (H2)` -`SAFE Quad 1` are close to `MC Proxy` in dimension 5 and are comparable to `MC` in dimension 6,
- `SAFE Lin (H3)` -`SAFE Quad 2` are close to `MC Proxy` in dimension 6 and are comparable to `MC` in dimension 7,
- `SAFE Quad 3` is more competitive than `MC Proxy` up to the dimension 8 and than `MC` up to the dimension 10,
- `SAFE Quad 4`-`SAFE Quad 5` are faster than `MC Proxy` in dimension 10.

Table 9.2: Estimation of the Put prices in dimension 2 with MC, MC proxy, SAFE methods and execution time.

payoff	method / strikes	90	95	100	105	110	exec. time
Basket	MC	1.83	3.39	5.65	8.61	12.18	2h20m
		(2.7E−3)	(3.7E−3)	(4.8E−3)	(5.8E−3)	(6.7E−3)	
	MC Proxy	1.82	3.39	5.64	8.60	12.17	2m40s
		(2.8E−3)	(3.8E−3)	(4.9E−3)	(5.9E−3)	(6.8E−3)	
	SAFE Lin **(H1)**	1.82	3.39	5.64	8.60	12.17	$< 10^{-4}$s
	SAFE Lin **(H2)**	1.83	3.40	5.66	8.61	12.18	$< 10^{-4}$s
	SAFE Lin **(H3)**	1.87	3.46	5.73	8.67	12.22	$< 10^{-4}$s
	SAFE Quad 1	1.82	3.39	5.65	8.60	12.17	$< 10^{-4}$s
	SAFE Quad 2	1.82	3.41	5.65	8.62	12.16	$< 10^{-4}$s
	SAFE Quad 3	1.85	3.55	5.47	8.67	12.08	$< 10^{-4}$s
Geo. Mean	MC	2.05	3.74	6.14	9.23	12.92	
		(2.8E−3)	(3.9E−3)	(5.0E−3)	(6.0E−3)	(6.8E−3)	
	MC Proxy	2.04	3.73	6.13	9.22	12.91	
		(3.0E−3)	(4.0E−3)	(5.1E−3)	(6.1E−3)	(7.0E−3)	
	SAFE Lin **(H1)**	2.04	3.73	6.13	9.22	12.91	
	SAFE Lin **(H2)**	2.05	3.75	6.13	9.23	12.92	
	SAFE Lin **(H3)**	2.11	3.83	6.24	9.32	12.99	
	SAFE Quad 1	2.04	3.73	6.13	9.22	12.91	
	SAFE Quad 2	2.05	3.76	6.15	9.25	12.91	
	SAFE Quad 3	2.05	3.82	6.15	9.35	12.90	
Worst of	MC	6.61	9.69	13.34	17.45	21.91	
		(5.4E−3)	(6.4E−3)	(7.3E−3)	(7.9E−3)	(8.4E−3)	
	MC Proxy	6.60	9.68	13.33	17.45	21.90	
		(5.7E−3)	(6.7E−3)	(7.6E−3)	(8.3E−3)	(8.8E−3)	
	SAFE Lin **(H1)**	6.60	9.68	13.33	17.44	21.90	
	SAFE Lin **(H2)**	6.62	9.70	13.37	17.47	21.91	
	SAFE Lin **(H3)**	6.82	9.80	13.51	17.56	22.03	
	SAFE Quad 1	6.63	9.67	13.34	17.42	21.89	
	SAFE Quad 2	6.41	9.74	13.36	17.40	21.69	
	SAFE Quad 3	7.04	9.93	13.40	17.74	22.09	
Best of	MC	0.71	1.43	2.60	4.28	6.54	
		(1.7E−3)	(2.5E−3)	(3.5E−3)	(4.5E−3)	(5.6E−3)	
	MC Proxy	0.71	1.43	2.59	4.28	6.53	
		(1.7E−3)	(2.5E−3)	(3.5E−3)	(4.5E−3)	(5.6E−3)	
	SAFE Lin **(H1)**	0.71	1.43	2.60	4.28	6.54	
	SAFE Lin **(H2)**	0.71	1.45	2.63	4.31	6.54	
	SAFE Lin **(H3)**	0.78	1.50	2.71	4.40	6.66	
	SAFE Quad 1	0.73	1.46	2.64	4.32	6.60	
	SAFE Quad 2	0.70	1.59	2.76	4.33	6.44	
	SAFE Quad 3	1.22	1.84	3.02	5.05	7.08	

Table 9.3: Estimation of the Put prices in dimension 3 with MC, MC proxy, SAFE methods and execution time.

payoff	method / strikes	90	95	100	105	110	exec. time
Basket	MC	1.11	2.46	4.62	7.63	11.36	3h44m
		(1.9E−3)	(2.9E−3)	(4.0E−3)	(5.0E−3)	(5.8E−3)	
	MC Proxy	1.11	2.45	4.61	7.62	11.36	3m16s
		(2.0E−3)	(3.0E−3)	(4.1E−3)	(5.1E−3)	(5.9E−3)	
	SAFE Lin **(H1)**	1.11	2.45	4.61	7.62	11.36	46s
	SAFE Lin **(H2)**	1.12	2.46	4.62	7.63	11.36	0.4s
	SAFE Lin **(H3)**	1.16	2.51	4.66	7.66	11.37	0.05s
	SAFE Quad 1	1.11	2.45	4.61	7.62	11.36	0.2s
	SAFE Quad 2	1.11	2.45	4.62	7.62	11.36	0.03s
	SAFE Quad 3	1.17	2.51	4.59	7.57	11.36	$< 10^{-4}$s
Geo. Mean	MC	1.36	2.90	5.28	8.51	12.43	
		(2.1E−3)	(3.1E−3)	(4.2E−3)	(5.2E−3)	(5.9E−3)	
	MC Proxy	1.35	2.89	5.28	8.50	12.42	
		(2.2E−3)	(3.3E−3)	(4.3E−3)	(5.3E−3)	(6.1E−3)	
	SAFE Lin **(H1)**	1.35	2.89	5.27	8.50	12.42	
	SAFE Lin **(H2)**	1.36	2.90	5.29	8.51	12.43	
	SAFE Lin **(H3)**	1.41	2.97	5.34	8.57	12.46	
	SAFE Quad 1	1.35	2.89	5.27	8.50	12.41	
	SAFE Quad 2	1.36	2.87	5.29	8.49	12.43	
	SAFE Quad 3	1.37	2.97	5.34	8.41	12.45	
Worst of	MC	9.03	12.82	17.12	21.74	26.56	
		(5.8E−3)	(6.6E−3)	(7.2E−3)	(7.5E−3)	(7.8E−3)	
	MC Proxy	9.01	12.81	17.11	21.74	26.56	
		(6.3E−3)	(7.1E−3)	(7.7E−3)	(8.2E−3)	(8.4E−3)	
	SAFE Lin **(H1)**	9.01	12.81	17.10	21.73	26.55	
	SAFE Lin **(H2)**	9.03	12.84	17.14	21.76	26.57	
	SAFE Lin **(H3)**	9.26	12.94	17.28	21.85	26.67	
	SAFE Quad 1	9.02	12.78	17.09	21.69	26.51	
	SAFE Quad 2	8.74	12.79	17.04	21.62	26.32	
	SAFE Quad 3	9.14	12.77	16.85	21.62	26.38	
Best of	MC	0.17	0.44	1.01	1.99	3.51	
		(7.1E−4)	(1.2E−3)	(2.0E−3)	(2.9E−3)	(3.9E−3)	
	MC Proxy	0.17	0.45	1.01	1.99	3.51	
		(7.2E−4	(1.2E−3)	(2.0E−3)	(2.9E−3)	(3.9E−3)	
	SAFE Lin **(H1)**	0.17	0.45	1.01	1.99	3.51	
	SAFE Lin **(H2)**	0.17	0.45	1.03	2.01	3.51	
	SAFE Lin **(H3)**	0.20	0.49	1.09	2.10	3.63	
	SAFE Quad 1	0.18	0.47	1.04	2.04	3.58	
	SAFE Quad 2	0.18	0.56	1.14	2.02	3.45	
	SAFE Quad 3	0.42	0.64	1.28	2.58	3.88	

Table 9.4: Estimation of the Put prices in dimension 4 with MC, MC proxy, SAFE methods and execution time.

payoff	method / strikes	90	95	100	105	110	exec. time
Basket	MC	0.74	1.92	4.00	7.06	10.94	4h34m
		(1.4E−3)	(2.4E−3)	(3.5E−3)	(4.5E−3)	(5.2E−3)	
	MC Proxy	0.74	1.91	4.00	7.05	10.93	3ms40s
		(1.5E−3)	(2.5E−3)	(3.5E−3)	(4.5E−3)	(5.3E−3)	
	SAFE Lin (H1)	0.74	1.91	4.00	7.05	10.93	2h35m
	SAFE Lin (H2)	0.75	1.92	4.00	7.06	10.93	17s
	SAFE Lin (H3)	0.77	1.96	4.04	7.08	10.93	0.7s
	SAFE Quad 1	0.74	1.91	4.00	7.05	10.93	7s
	SAFE Quad 2	0.74	1.92	4.00	7.06	10.93	0.5s
	SAFE Quad 3	0.74	1.91	3.97	7.02	10.88	0.03s
Geo. Mean	MC	0.98	2.39	4.77	8.10	12.19	
		(1.6E−3)	(2.7E−3)	(3.7E−3)	(4.7E−3)	(5.4E−3)	
	MC Proxy	0.97	2.38	4.76	8.09	12.18	
		(1.7E−3)	(2.8E−3)	(3.9E−3)	(4.8E−3)	(5.5E−3)	
	SAFE Lin (H1)	0.97	2.38	4.75	8.09	12.18	
	SAFE Lin (H2)	0.98	2.39	4.76	8.10	12.19	
	SAFE Lin (H3)	1.02	2.45	4.81	8.13	12.20	
	SAFE Quad 1	0.97	2.38	4.76	8.09	12.18	
	SAFE Quad 2	0.98	2.38	4.76	8.10	12.19	
	SAFE Quad 3	0.99	2.43	4.81	8.13	12.19	
Worst of	MC	11.02	15.27	19.90	24.74	29.68	
		(5.9E−3)	(6.5E−3)	(6.9E−3)	(7.1E−3)	(7.2E−3)	
	MC Proxy	11.01	15.26	19.89	24.74	29.68	
		(6.5E−3)	(7.2E−3)	(7.7E−3)	(8.0E−3)	(8.2E−3)	
	SAFE Lin (H1)	11.00	15.25	19.88	24.72	29.66	
	SAFE Lin (H2)	11.03	15.28	19.91	24.75	29.68	
	SAFE Lin (H3)	11.25	15.38	20.05	24.84	29.79	
	SAFE Quad 1	11.00	15.21	19.85	24.66	29.61	
	SAFE Quad 2	10.69	15.14	19.73	24.55	29.42	
	SAFE Quad 3	10.72	14.83	19.26	24.17	29.09	
Best of	MC	0.04	0.15	0.42	1.00	2.02	
		(3.2E−4)	(6.6E−4)	(1.2E−3)	(1.9E−3)	(2.9E−3)	
	MC Proxy	0.04	0.15	0.42	1.00	2.02	
		(3.3E−4)	(6.5E−4)	(1.2E−3)	(1.9E−3)	(2.8E−3)	
	SAFE Lin (H1)	0.04	0.15	0.42	1.00	2.02	
	SAFE Lin (H2)	0.04	0.16	0.44	1.01	2.02	
	SAFE Lin (H3)	0.06	0.18	0.47	1.10	2.12	
	SAFE Quad 1	0.05	0.16	0.45	1.04	2.09	
	SAFE Quad 2	0.06	0.22	0.51	1.01	2.01	
	SAFE Quad 3	0.15	0.22	0.60	1.42	2.25	

Table 9.5: Estimation of the Put prices in dimension 5 with MC, MC proxy, SAFE methods and execution time.

payoff	method / strikes	90	95	100	105	110	exec. time
Basket	MC	0.52	1.56	3.58	6.68	10.67	5h43m
		(5.7E−3)	(2.0E−3)	(3.1E−3)	(4.1E−3)	(4.8E−3)	
	MC Proxy	0.52	1.56	3.58	6.68	10.67	4m11s
		(1.2E−3)	(2.1E−3)	(3.2E−3)	(4.2E−3)	(4.9E−3)	
	SAFE Lin **(H2)**	0.52	1.56	3.58	6.68	10.66	14m25s
	SAFE Lin **(H3)**	0.54	1.59	3.61	6.69	10.65	15s
	SAFE Quad 1	0.52	1.56	3.58	6.68	10.67	5m10s
	SAFE Quad 2	0.52	1.56	3.58	6.68	10.67	11s
	SAFE Quad 3	0.54	1.55	3.57	6.67	10.66	0.5s
Geo. Mean	MC	0.73	2.04	4.41	7.84	12.06	
		(1.3E−3)	(2.3E−3)	(3.4E−3)	(4.3E−3)	(4.9E−3)	
	MC Proxy	0.73	2.04	4.40	7.83	12.06	
		(1.4E−3)	(2.4E−3)	(3.5E−3)	(4.5E−3)	(5.1E−3)	
	SAFE Lin **(H2)**	0.74	2.04	4.41	7.83	12.05	
	SAFE Lin **(H3)**	0.76	2.08	4.46	7.87	12.07	
	SAFE Quad 1	0.73	2.03	4.39	7.82	12.05	
	SAFE Quad 2	0.73	2.04	4.38	7.83	12.05	
	SAFE Quad 3	0.72	2.05	4.44	7.84	12.04	
Worst of	MC	12.70	17.23	22.04	26.97	31.95	
		(5.9E−3)	(6.4E−3)	(6.6E−3)	(6.7E−3)	(6.7E−3)	
	MC Proxy	12.68	17.21	22.02	26.95	31.93	
		(6.7E−3)	(7.2E−3)	(7.6E−3)	(7.8E−3)	(8.1E−3)	
	SAFE Lin **(H2)**	12.69	17.23	22.03	26.96	31.94	
	SAFE Lin **(H3)**	12.91	17.33	22.17	27.06	32.05	
	SAFE Quad 1	12.65	17.14	21.95	26.87	31.85	
	SAFE Quad 2	12.33	17.01	21.79	26.71	31.65	
	SAFE Quad 3	11.95	16.37	21.01	25.98	30.94	
Best of	MC	0.01	0.05	0.19	0.52	1.21	
		(1.5E−4)	(3.6E−4)	(7.3E−4)	(1.3E−3)	(2.1E−3)	
	MC Proxy	0.01	0.05	0.19	0.52	1.21	
		(1.5E−4)	(3.5E−4)	(7.1E−4)	(1.3E−3)	(2.0E−3)	
	SAFE Lin **(H2)**	0.01	0.06	0.20	0.54	1.21	
	SAFE Lin **(H3)**	0.02	0.07	0.22	0.61	1.30	
	SAFE Quad 1	0.01	0.06	0.20	0.56	1.27	
	SAFE Quad 2	0.02	0.09	0.24	0.52	1.24	
	SAFE Quad 3	0.05	0.08	0.30	0.83	1.36	

Table 9.6: Estimation of the Put prices in dimension 6 with MC, MC proxy, SAFE methods and execution time.

payoff	method / strikes	90	95	100	105	110	exec. time
Basket	MC	0.38	1.31	3.27	6.41	10.50	6h52m
		(9.1E−4)	(1.8E−3)	(2.9E−3)	(3.8E−3)	(4.5E−3)	
	MC Proxy	0.38	1.30	3.27	6.41	10.50	4m36s
		(9.5E−4)	(1.8E−3)	(2.9E−3)	(3.9E−3)	(4.6E−3)	
	SAFE Lin **(H2)**	0.38	1.31	3.27	6.41	10.49	8h2m
	SAFE Lin **(H3)**	0.40	1.33	3.29	6.41	10.47	4m47s
	SAFE Quad 1	0.38	1.30	3.26	6.41	10.50	3h2m
	SAFE Quad 2	0.38	1.30	3.26	6.41	10.50	3m25s
	SAFE Quad 3	0.38	1.32	3.26	6.40	10.47	5s
Geo. Mean	MC	0.57	1.78	4.14	7.65	11.98	
		(1.1E−3)	(2.1E−3)	(3.2E−3)	(4.1E−3)	(4.6E−3)	
	MC Proxy	0.56	1.78	4.13	7.64	11.98	
		(1.2E−3)	(2.2E−3)	(3.3E−3)	(4.2E−3)	(4.8E−3)	
	SAFE Lin **(H2)**	0.57	1.79	4.14	7.65	11.97	
	SAFE Lin **(H3)**	0.60	1.82	4.19	7.68	11.99	
	SAFE Quad 1	0.56	1.78	4.13	7.64	11.97	
	SAFE Quad 2	0.56	1.77	4.13	7.64	11.97	
	SAFE Quad 3	0.58	1.76	4.16	7.62	11.98	
Worst of	MC	14.12	18.82	23.72	28.69	33.68	
		(5.9E−3)	(6.2E−3)	(6.3E−3)	(6.3E−3)	(6.4E−3)	
	MC Proxy	14.09	18.80	23.70	28.67	33.66	
		(6.7E−3)	(7.2E−3)	(7.5E−3)	(7.8E−3)	(8.0E−3)	
	SAFE Lin **(H2)**	14.11	18.81	23.71	28.67	33.66	
	SAFE Lin **(H3)**	14.31	18.93	23.84	28.79	33.78	
	SAFE Quad 1	14.05	18.72	23.62	28.58	33.57	
	SAFE Quad 2	13.73	18.55	23.42	28.39	33.36	
	SAFE Quad 3	12.94	17.56	22.33	27.32	32.31	
Best of	MC	3.2E−3	0.02	0.08	0.28	0.75	
		(7.5E−5)	(2.0E−4)	(4.6E−4)	(9.1E−4)	(1.6E−3)	
	MC Proxy	3.2E−3	0.02	0.08	0.28	0.75	
		(7.5E−5)	(2.0E−4)	(4.5E−4)	(8.8E−4)	(1.5E−3)	
	SAFE Lin **(H2)**	3.3E−3	0.02	0.09	0.29	0.75	
	SAFE Lin **(H3)**	5.1E−3	0.03	0.10	0.35	0.81	
	SAFE Quad 1	3.9E−3	0.02	0.09	0.31	0.80	
	SAFE Quad 2	8.1E−3	0.04	0.12	0.27	0.80	
	SAFE Quad 3	1.8E−2	0.03	0.17	0.50	0.84	

Table 9.7: Estimation of the Put prices in dimension 7 with MC, MC proxy, SAFE methods and execution time.

payoff	method / strikes	90	95	100	105	110	exec. time
Basket	MC	0.28	1.11	3.03	6.21	10.38	7h43m
		(7.5E−4)	(1.6E−3)	(2.6E−3)	(3.6E−3)	(4.3E−3)	
	MC Proxy	0.28	1.11	3.03	6.21	10.38	5m9s
		(7.9E−4)	(1.6E−3)	(2.7E−3)	(3.7E−3)	(4.4E−3)	
	SAFE Lin **(H3)**	0.30	1.14	3.04	6.20	10.34	1h36m
	SAFE Quad 2	0.28	1.11	3.02	6.20	10.38	1h2m
	SAFE Quad 3	0.29	1.12	3.02	6.18	10.36	46s
Geo. Mean	MC	0.45	1.58	3.93	7.52	11.94	
		(9.6E−4)	(1.9E−3)	(3.0E−3)	(3.8E−3)	(4.3E−3)	
	MC Proxy	0.45	1.58	3.93	7.51	11.93	
		(1.0E−3)	(2.0E−3)	(3.1E−3)	(4.0E−3)	(4.6E−3)	
	SAFE Lin **(H3)**	0.47	1.62	3.97	7.54	11.94	
	SAFE Quad 2	0.44	1.58	3.93	7.51	11.93	
	SAFE Quad 3	0.43	1.57	3.94	7.52	11.92	
Worst of	MC	15.32	20.14	25.09	30.07	35.07	
		(5.8E−3)	(6.0E−3)	(6.0E−3)	(6.1E−3)	(6.1E−3)	
	MC Proxy	15.30	20.11	25.06	30.05	35.05	
		(6.8E−3)	(7.2E−3)	(7.5E−3)	(7.8E−3)	(8.1E−3)	
	SAFE Lin **(H3)**	15.51	20.25	25.20	30.18	35.18	
	SAFE Quad 2	14.94	19.83	24.76	29.74	34.73	
	SAFE Quad 3	13.77	18.53	23.38	28.37	33.37	
Best of	MC	8.9E−4	0.01	0.04	0.15	0.47	
		(3.7E−5)	(1.2E−4)	(3.0E−4)	(6.4E−4)	(1.2E−3)	
	MC Proxy	9.3E−4	0.01	0.04	0.16	0.47	
		(3.8E−5)	(1.1E−4)	(2.9E−4)	(6.2E−4)	(1.2E−3)	
	SAFE Lin **(H3)**	1.6E−3	0.01	0.05	0.21	0.52	
	SAFE Quad 2	3.3E−3	0.02	0.06	0.14	0.54	
	SAFE Quad 3	6.3E−3	0.01	0.10	0.31	0.53	

Table 9.8: Estimation of the Put prices in dimension 8 with MC, MC proxy, SAFE methods and execution time.

payoff	method / strikes	90	95	100	105	110	exec. time
`Basket`	`MC`	0.21	0.96	2.83	6.05	10.30	9h11m
		(6.2E−4)	(1.4E−3)	(2.5E−3)	(3.4E−3)	(4.0E−3)	
	`MC Proxy`	0.21	0.96	2.83	6.05	10.29	5m36s
		(6.6E−4)	(1.5E−3)	(2.5E−3)	(3.5E−3)	(4.2E−3)	
	`SAFE Quad 3`	0.22	0.96	2.82	6.03	10.28	6m36
`Geo. Mean`	`MC`	0.36	1.42	3.76	7.41	11.91	
		(8.3E−4)	(1.7E−3)	(2.8E−3)	(3.7E−3)	(4.1E−3)	
	`MC Proxy`	0.36	1.42	3.76	7.41	11.90	
		(8.9E−4)	(1.8E−3)	(2.9E−3)	(3.8E−3)	(4.4E−3)	
	`SAFE Quad 3`	0.36	1.43	3.76	7.41	11.90	
`Worst of`	`MC`	16.37	21.26	26.23	31.22	36.22	
		(5.7E−3)	(5.8E−3)	(5.8E−3)	(5.8E−3)	(5.8E−3)	
	`MC Proxy`	16.35	21.23	26.20	31.20	36.20	
		(6.8E−3)	(7.2E−3)	(7.5E−3)	(7.8E−3)	(8.1E−3)	
	`SAFE Quad 3`	14.50	19.34	24.24	29.24	34.24	
`Best of`	`MC`	2.5E−4	2.7E−3)	0.02	0.09	0.30	
		(1.9E−5)	(6.7E−5)	(1.9E−4)	(4.6E−4)	(9.2E−4)	
	`MC Proxy`	2.8E−4	2.8E−3	0.02	0.09	0.30	
		(1.9E−5)	(6.6E−5)	(1.9E−4)	(4.4E−4)	(8.9E−4)	
	`SAFE Quad 3`	2.2E−3	3.4E−3	0.06	0.19	0.33	

Table 9.9: Estimation of the Put prices in dimension 9 with MC, MC proxy, SAFE methods and execution time.

payoff	method / strikes	90	95	100	105	110	exec. time
`Basket`	`MC`	0.16	0.84	2.67	5.92	10.24	9h47m
		(5.3E−4)	(1.3E−3)	(2.3E−3)	(3.3E−3)	(3.9E−3)	
	`MC Proxy`	0.16	0.84	2.67	5.92	10.23	6m10s
		(5.6E−4)	(1.3E−3)	(2.4E−3)	(3.4E−3)	(4.0E−3)	
	`SAFE Quad 3`	0.17	0.84	2.66	5.91	10.22	1h5m
`Geo. Mean`	`MC`	0.29	1.29	3.63	7.33	11.90	
		(7.2E−4)	(1.6E−3)	(2.7E−3)	(3.5E−3)	(3.9E−3)	
	`MC Proxy`	0.29	1.29	3.62	7.33	11.89	
		(7.8E−4)	(1.7E−3)	(2.8E−3)	(3.7E−3)	(4.2E−3)	
	`SAFE Quad 3`	0.29	1.30	3.61	7.32	11.88	
`Worst of`	`MC`	17.29	22.22	27.20	32.20	37.20	
		(5.5E−3)	(5.6E−3)	(5.7E−3)	(5.7E−3)	(5.7E−3)	
	`MC Proxy`	17.27	22.19	27.18	32.18	37.18	
		(6.8E−3)	(7.2E−3)	(7.5E−3)	(7.9E−3)	(8.2E−3)	
	`SAFE Quad 3`	15.14	20.04	24.98	29.98	34.97	
`Best of`	`MC`	7.2E−5	1.1E−3	0.01	0.05	0.19	
		(8.8E−6)	(4.0E−5)	(1.3E−4)	(3.3E−4)	(7.1E−4)	
	`MC Proxy`	7.6E−4	1.1E−3	0.01	0.05	0.19	
		(9.5E−5)	(3.8E−5)	(1.2E−4)	(3.1E−4)	(6.8E−4)	
	`SAFE Quad 3`	7.8E−3	1.2E−3	0.04	0.12	0.21	

Table 9.10: Estimation of the Put prices in dimension 10 with MC, MC proxy, SAFE methods and execution time.

payoff	method / strikes	90	95	100	105	110	exec. time
`Basket`	`MC`	0.13	0.74	2.53	5.82	10.19	10h1m
		(4.5E−4)	(1.2E−3)	(2.2E−3)	(3.2E−3)	(3.7E−3)	
	`MC Proxy`	0.13	0.74	2.53	5.82	10.19	6m28s
		(4.8E−4)	(1.2E−3)	(2.3E−3)	(3.3E−3)	(3.9E−3)	
	`SAFE Quad 3`	0.13	0.74	2.53	5.80	10.17	10h24m
	`SAFE Quad 4`	0.16	0.81	2.46	5.48	9.76	1m45s
	`SAFE Quad 5`	0.10	0.66	2.73	6.41	10.35	0.7s
`Geo. Mean`	`MC`	0.24	1.18	3.51	7.27	11.89	
		(6.4E−4)	(1.5E−3)	(2.6E−3)	(3.4E−3)	(3.7E−3)	
	`MC Proxy`	0.24	1.18	3.51	7.27	11.88	
		(6.9E−4)	(1.6E−3)	(2.7E−3)	(3.5E−3)	(4.0E−3)	
	`SAFE Quad 3`	0.24	1.19	3.49	7.27	11.88	
	`SAFE Quad 4`	0.21	1.26	3.63	7.28	11.81	
	`SAFE Quad 5`	0.21	1.19	3.39	7.33	11.96	
`Worst of`	`MC`	18.11	23.06	28.05	33.05	38.05	
		(5.4E−3)	(5.5E−3)	(5.5E−3)	(5.5E−3)	(5.5E−3)	
	`MC Proxy`	18.08	23.03	28.02	33.02	38.02	
		(6.8E−3)	(7.2E−3)	(7.6E−3)	(8.0E−3)	(8.3E−3)	
	`SAFE Quad 3`	15.14	20.04	24.98	29.97	34.97	
	`SAFE Quad 4`	14.25	18.80	23.53	28.53	33.53	
	`SAFE Quad 5`	11.55	12.81	15.55	20.55	25.55	
`Best of`	`MC`	3.1E−5	4.5E−4	4.4E−3	0.03	0.13	
		(6.1E−06)	(2.6E−5)	(8.7E−5)	(2.4E−4)	(5.6E−4)	
	`MC Proxy`	2.2E−5	4.0E−4	4.2E−3	0.03	0.13	
		(4.5E−06)	(2.2E−5)	(7.8E−5)	(2.2E−4)	(5.3E−4)	
	`SAFE Quad 3`	0.0	0.0	0.0	0.00	0.00	
	`SAFE Quad 4`	0.0	0.0	0.0	0.00	0.00	
	`SAFE Quad 5`	0.0	0.0	0.0	0.00	0.00	

List of Figures

List of Tables

Bibliography

[Abate 1995] J. Abate and W. whitt. *Numerical inversion of Laplace transforms of probability distributions*. Orsa Journal of Computing, vol. 7, no. 1, pages 36–43, 1995. (Cited on page 192.)

[Achdou 2005] Y. Achdou and O. Pironneau. Computational methods for option pricing. SIAM series, Frontiers in Applied Mathematics, Philadelphia, 2005. (Cited on page 3.)

[Alòs 2008] E. Alòs and C.O. Ewald. *Malliavin differentiability of the Heston volatility and applications to option pricing*. Advances in Applied Probability, vol. 40, no. 1, pages 144–162, 2008. (Cited on pages 12, 21, 123 and 129.)

[Alòs 2012] E. Alòs. *A decomposition formula for option prices in the Heston model and applications to option pricing approximation*. Finance and Stochastics, vol. 16, no. 3, pages 403–422, 2012. (Cited on page 41.)

[Andersen 2000] L. Andersen and J. Andreasen. *Volatility skews and extensions of the Libor market model*. Applied Mathematical Finance, vol. 7, no. 1, pages 1–32, 2000. (Cited on page 65.)

[Andersen 2002] L. Andersen, J. Andreasen and D. Eliezer. *Static replication of barrier options: some general results*. Journal of Computational Finance, vol. 5, no. 4, pages 1–25, 2002. (Cited on page 169.)

[Andersen 2006] L. Andersen and V. Piterbarg. *Moment explosion in stochastic volatility models*. Finance and Stochastics, vol. 11, no. 1, pages 29–50, 2006. (Cited on pages 126 and 129.)

[Avellaneda 2003] M. Avellaneda, D. Boyer-Olson, J. Busca and P. Friz. *Application of large deviation methods to the pricing of index options in finance*. Comptes Rendus Mathematique, vol. 336, no. 3, pages 263 – 266, 2003. (Cited on page 213.)

[Azencott 83] R. Azencott. *Densité des diffusions en temps petit: développements asymptotiques. Partie I*. Séminaire de Probabilités XVIII - Lecture Notes in Mathematics 1059, 1982–83. (Cited on page 11.)

[Bachelier 1900] L. Bachelier. *Théorie de la spéculation*. Annales Scientifiques de l'École Normale Supérieure, vol. 17, pages 21–86, 1900. (Cited on page 3.)

[Baldi 1995] P. Baldi. *Exact Asymptotics for the Probability of Exit from a Domain and Applications to Simulation*. The Annals of Probability, vol. 23, no. 4, pages 1644–1670, 1995. (Cited on page 168.)

[Bally 1996] V. Bally and D. Talay. *The law of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution function*. Probab. Theory Related Fields, vol. 104-1, pages 43–60, 1996. (Cited on page 221.)

[Bally 2003] V. Bally and G. Pagès. *A quantization algorithm for solving multi-dimensional optimal stopping problems*. Bernoulli, vol. 9, no. 6, pages 1003–1049, 2003. (Cited on pages 212 and 213.)

[Bardos 2002] C. Bardos, R. Douady and A. Fursikov. *Static hedging of barrier options with a smile*. ESAIM: Control, Optimization and Calculus of Variations, vol. 8, pages 127–142, 2002. (Cited on page 169.)

[Barndorff-Nielson 2001] O. E. Barndorff-Nielson and N. Shephard. *Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial mathematics*. J. Royal Stat. Soc., vol. 63, pages 167–241, 2001. (Cited on page 82.)

[Benaim 2009] S. Benaim and P. Friz. *Regular variation and smile asymptotics*. Mathematical Finance, vol. 19, no. 1, pages 1–12, 2009. (Cited on page 8.)

[Benhamou 2009] E. Benhamou, E. Gobet and M. Miri. *Smart expansion and fast calibration for jump diffusion*. Finance and Stochastics, vol. 13, no. 4, pages 563–589, 2009. (Cited on pages 2, 40, 41, 84, 123, 168, 169 and 170.)

[Benhamou 2010a] E. Benhamou, E. Gobet and M. Miri. *Expansion formulas for European options in a local volatility model*. International Journal of Theoretical and Applied Finance, vol. 13, no. 4, pages 603–634, 2010. (Cited on pages 2, 19, 21, 40, 45, 46, 53, 84, 123, 127, 168, 170, 175, 176, 177, 183, 187 and 188.)

[Benhamou 2010b] E. Benhamou, E. Gobet and M. Miri. *Time dependent Heston model*. SIAM Journal on Financial Mathematics, vol. 1, pages 289–325, 2010. (Cited on pages 2, 40, 89, 123, 125, 127, 128, 129, 130, 131, 145, 146, 152, 169 and 170.)

[Benhamou 2012] E. Benhamou, E. Gobet and M. Miri. *Analytical formulas for local volatility model with stochastic rates*. Quantitative finance, vol. 2, no. 2, pages 185–198, 2012. (Cited on pages 5, 40, 59 and 170.)

[Berestycki 2004] H. Berestycki, J. Busca and I. Florent. *Computing the implied volatility in stochastic volatility models*. Comm. Pure Appl. Math., vol. 10, pages 1352–1373, 2004. (Cited on page 122.)

[Bergomi 2005] L. Bergomi. *Smile dynamics II*. Risk, pages 67–73, 2005. (Cited on page 83.)

[Bergomi 2008] L. Bergomi. *Smile dynamics III*. Risk, pages 90–96, 2008. (Cited on page 83.)

[Beyer 2008] P. Beyer and J. Kienitz. *Pricing Forward Start Options in Models Based on (Time-Changed) Levy Processes*. Available at `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1319703`, 2008. (Cited on page 82.)

[Black 1973] F. Black and M. Scholes. *The Pricing of Options and Corporate Liabilities*. Journal of Political Economy, vol. 81, pages 637–659, 1973. (Cited on pages 3 and 81.)

[Black 1976] F. Black. *The Pricing of Commodity Contracts*. Journal of Financial Economics, vol. 3, pages 167–179, 1976. (Cited on page 2.)

[Bossy 2007] 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. (Cited on page 89.)

[Boyle 1998] P. Boyle and Y. Tian. *An explicit finite difference approach to the pricing of barrier options*. Applied Mathematical Finance, vol. 5, no. 1, pages 17–43, 1998. (Cited on page 168.)

[Brenner 2008] S.C. Brenner and L.R. Scott. The mathematical theory of finite element methods. Texts in Applied Mathematics 15. New York, NY: Springer., third édition, 2008. (Cited on pages 218, 219, 222 and 228.)

[Brigo 2006] D. Brigo and F. Mercurio. Interest Rate Models Theory and Practice. Springer Finance, 2nd édition, 2006. (Cited on page 82.)

[Britton-Jones 2000] M. Britton-Jones and A. Neuberger. *Option Prices, Implied Processes, and Stochastic Volatility*. The Journal of FINANCE, vol. 55, no. 2, pages 839–866, 2000. (Cited on page 81.)

[Broadie 1997] M. Broadie, P. Glasserman and S. Kou. *A continuity correction for discrete barrier options*. Mathematical Finance, vol. 7, no. 4, pages 325–348, 1997. (Cited on page 168.)

[Broadie 1999] M. Broadie, P. Glasserman and S. Kou. *Connecting discrete and continuous path-dependent options*. Finance and Stochastics, vol. 3, no. 1, pages 55–82, 1999. (Cited on page 168.)

[Bungartz 2004] H-J. Bungartz and M. Griebel. *Sparse grids*. Acta Numerica, vol. 13, pages 147–269, 2004. (Cited on pages 223 and 236.)

[Carr 1998a] P. Carr, K. Ellis and V. Gupta. *Static hedging of path dependent options*. Journal of Finance, vol. 53, pages 1165–1190, 1998. (Cited on page 168.)

[Carr 1998b] P. Carr and D.B. Madan. *Option valuation using the fast Fourier transform*. Journal of Computational Finance, vol. 2, no. 4, pages 61–73, 1998. (Cited on page 4.)

[Carr 2011] P. Carr and S. Nadtochiy. *Static Hedging under Time-Homogeneous Diffusions*. SIAM Journal of Financial Mathematics, pages 794–838, 2011. (Cited on page 169.)

[Cattiaux 1986] P. Cattiaux. *Hypoellipticité et hypoellipticité partielle pour les diffusions avec une condition frontière*. Ann. Inst. H. Poincaré Probab. Statist., vol. 22, no. 1, pages 67–112, 1986. (Cited on page 173.)

[Cattiaux 1991] P. Cattiaux. *Calcul stochastique et opérateurs dégénérés du second ordre - II. Problème de Dirichlet*. Bull. Sc. Math., 2ème série, vol. 115, pages 81–122, 1991. (Cited on page 173.)

[Choi 2010] S.Y. Choi, J.P. Fouque and J.H Kim. *Option pricing under hybrid stochastic and local volatility*. preprint, 2010. (Cited on page 122.)

[Corielli 2010] F. Corielli, P. Foschi and A. Pascucci. *Parametrix Approximation of Diffusion Transition Densities*. SIAM Journal on Financial Mathematics, vol. 1, pages 833–867, 2010. (Cited on page 14.)

[Costantini 2006] C. Costantini, E. Gobet and N. El Karoui. *Boundary sensitivities for diffusion processes in time dependent domains*. Applied Mathematics Optimization, vol. 54, no. 2, pages 159–187, 2006. (Cited on page 168.)

[Cox 1975] 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. (Cited on pages 31 and 122.)

[Davydov 2001] D. Davydov and V. Linetsky. *Pricing and Hedging Path-Dependent Options Under the CEV Process*. Management Science, vol. 47, no. 7, pages 949–965, 2001. (Cited on pages 169 and 192.)

[Davydov 2003] D. Davydov and V. Linetsky. *Pricing Options on Scalar Diffusions: An Eigenfunction Expansion Approach*. Operations Research, vol. 51, no. 2, pages 185–209, 2003. (Cited on pages 169 and 192.)

[De Marco 2011] S. De Marco. *Smoothness and asymptotic estimates of densities for SDEs with locally smooth coefficients and applications to square root-type diffusions*. Ann. Appl. Probab., vol. 21, no. 4, pages 1282–1321, 2011. (Cited on page 21.)

[Delarue 2003] F. Delarue. *Estimates of the solutions of a system of quasi-linear PDEs. A probabilistic scheme*. Séminaire de Probabilités XXXVII - Lecture Notes in Mathematics 1832, pages 290–332, 2003. (Cited on pages 170 and 174.)

[Duffie 1995] D. Duffie and P. Glynn. *Efficient Monte Carlo simulation of security prices*. Ann. Appl. Probab., vol. 5, no. 4, pages 897–905, 1995. (Cited on pages 213 and 221.)

[Duffie 2003] D. Duffie, D. Filipovic and W. Schachermayer. *Affine processes and applications in finance*. The Annals of Applied Probability, vol. 13, no. 3, pages 984–1053, 2003. (Cited on page 9.)

[Dupire 1994] B. Dupire. *Pricing with a smile*. Risk, vol. 7, no. 1, pages 18–20, 1994. (Cited on pages 34 and 81.)

[El Karoui 1998] N. El Karoui, M. Jeanblanc-Picqué and S.E. Shreve. *Robustness of the Black and Scholes formula*. Math. Finance, vol. 8, no. 2, pages 93–126, 1998. (Cited on page 174.)

[El Karoui 1999] N. El Karoui and M. Jeanblanc. *Options exotiques*. Finance, pages 49–67, 1999. (Cited on page 168.)

[Emanuel 1982] D. Emanuel and J. MacBeth. *Further results on the constant elasticity of variance call option pricing model*. J. Financial and Quant. Anal., no. 17, pages 533–554, 1982. (Cited on page 31.)

[Faà di Bruno 1857] C.F. Faà di Bruno. *Note sur une nouvelle formule de calcul différentiel*. Quaterly J. Pure Appl. Math, vol. 1, pages 359–360, 1857. (Cited on page 105.)

[Fleming 1986] W.H. Fleming and P.E. Souganidis. *Asymptotic series and the method of vanishing viscosity*. Indiana Univ. Math. J., vol. 35, pages 425–447, 1986. (Cited on page 11.)

[Forde 2009] M. Forde and A. Jacquier. *Small-time asymptotics for implied volatility under the Heston model*. International Journal of Theoretical and Applied Finance, vol. 12, no. 6, pages 861–876, 2009. (Cited on page 141.)

[Forde 2011] M. Forde and A. Jacquier. *The large-maturity smile for the Heston model*. Finance and Stochastics, vol. 15, no. 4, pages 755–780, 2011. (Cited on page 9.)

[Forde 2012a] M. Forde and A. Jacquier. *Small-time asymptotics for an uncorrelated local-stochastic volatility model*. forthcoming in Applied Mathematical Finance, 2012. (Cited on pages 122 and 141.)

[Forde 2012b] M. Forde and A. Pogudin. *The Large-maturity smile for the SABR and CEV-Heston models*. preprint, 2012. (Cited on page 122.)

[Foschi 2013] P. Foschi, S. Pagliarani and A. Pascucci. *Black-Scholes formulae for Asian options in local volatility models*. J. Comput. Appl. Math., pages 442–459, 2013. (Cited on page 213.)

[Fouque 2000] J.P. Fouque, G. Papanicolaou and R. Sircar. Derivatives in financial markets with stochastic volatility. Cambridge University Press, 2000. (Cited on page 15.)

[Fouque 2003] J.P. Fouque, G. Papanicolaou, R. Sircar and S. Solna. *Singular perturbations in option pricing*. SIAM J. Appl. Math., vol. 63, no. 5, pages 1648–1665, 2003. (Cited on pages 13 and 15.)

[Fouque 2004] J.P. Fouque, G. Papanicolaou, R. Sircar and K. Solna. *Maturity Cycles in Implied Volatility*. Finance and Stochastics, vol. 8, no. 4, pages 451–477, 2004. (Cited on pages 13 and 16.)

[Fouque 2011] J.P. Fouque, G. Papanicolaou, R. Sircar and S. Knut. Multiscale stochastic volatility for equity, interest rate, and credit derivatives. Cambridge University Press, Cambridge, 2011. (Cited on pages 11, 17 and 213.)

[Fournié 1997] E. Fournié, J. Lebuchoux and N. Touzi. *Small noise expansion and importance sampling*. Asymptot. Anal., vol. 14, no. 4, pages 361–376, 1997. (Cited on page 11.)

[Freidlin 1998] M.I. Freidlin and A.D. Wentzell. Random perturbations of dynamical systems. New York, NY: Springer, second édition, 1998. Transl. from the Russian by Joseph Szuecs. (Cited on pages 10 and 122.)

[Gatheral 2003] J. Gatheral. *Case studies in financial modeling lecture notes*. Courant Institute of Mathematical Sciences, New York University, 2003. (Cited on page 83.)

[Gatheral 2004] J. Gatheral. *A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives*. Presentation at Global Derivatives & Risk Management, Madrid, www.math.nyu.edu/fellows_fin_math/gatheral/madrid2004.pdf, May 2004. (Cited on page 9.)

[Gatheral 2006] J. Gatheral. The volatility surface, a practioner's guide. Wiley Finance, 2006. (Cited on page 82.)

[Gatheral 2011] J. Gatheral and A. Jacquier. *Convergence of Heston to SVI*. Quantitative Finance, vol. 11, no. 8, pages 1129–1132, 2011. (Cited on page 9.)

[Gatheral 2012] J. Gatheral, E.P. Hsu, P. Laurence, C. Ouyang and T.H Wang. *Asymptotics of implied volatility in local volatility models*. Methematical Finance, vol. 22, no. 10, 2012. (Cited on page 14.)

[Giese 2007] A. Giese and J. Maruhn. *Cost Optimal static super-replication of barrier options: an opimization approach*. Journal of Computational Finance, vol. 10, no. 3, pages 71–97, 2007. (Cited on page 169.)

[Glasserman 2004] P. Glasserman. Monte carlo methods in financial engineering. Springer Verlag, New York, 2004. (Cited on pages 3 and 145.)

[Glasserman 2011] P. Glasserman and Q. Wu. *Forward and Future Implied Volatility*. International Journal of Theoretical and Applied Finance, vol. 14, no. 3, 2011. (Cited on pages 82 and 83.)

[Gobet 2001] E. Gobet. *Euler schemes and half-space approximation for the simulation of diffusions in a domain*. ESAIM: Probability and Statistics, vol. 5, pages 261–297, 2001. (Cited on pages 168 and 229.)

[Gobet 2005] E. Gobet and R. Munos. *Sensitivity analysis using Itô-Malliavin calculus and martingales. Application to stochastic control problem*. Journal of Control and Optimization, vol. 43, no. 5, pages 1676–1713, 2005. (Cited on page 130.)

[Gobet 2009] E. Gobet. *Advanced Monte Carlo methods for barrier and related exotic options*. In Alain Bensoussan, Qiang Zhang and Philippe G. Ciarlet, editeurs, Special Volume: Mathematical Modeling and Numerical Methods in Finance, volume 15 of *Handbook of Numerical Analysis*, pages 497–528. Elsevier, 2009. (Cited on page 168.)

[Gobet 2010] E. Gobet and S. Menozzi. *Stopped diffusion processes: boundary corrections and overshoot*. Stochastic Processes and Their Applications, vol. 120, no. 2, pages 130–162, 2010. (Cited on page 168.)

[Gobet 2012a] E. Gobet and M. Miri. *Weak approximation of averaged diffusion processes*. Forthcoming in Stochastic Processes and their Applications, 2012. (Cited on pages 3, 33, 40, 41, 130, 212, 217 and 237.)

[Gobet 2012b] E. Gobet and A. Suleiman. *New approximations in local volatility models*. In Y. Kabanov, editeur, Musiela Festschrift. Springer, 2012. (Cited on pages 34, 45 and 105.)

[Graf 2000] S. Graf and H. Luschgy. Foundations of quantization for probability distributions, volume 1730 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000. (Cited on page 213.)

[Gulisashvili 2010] A. Gulisashvili. *Asymptotic formulas with error estimates for call pricing functions and the implied volatility at extreme strikes*. SIAM J. Financial Math., vol. 1, pages 609–641, 2010. (Cited on page 8.)

[Gulisashvili 2012] A. Gulisashvili, editeur. Analytically tractable stochastic stock price models. Forthcoming in Springer Finance, 2012. (Cited on page 8.)

[Gulisashvili 2013] A. Gulisashvili and P. Tankov. *Asymptotics for sums of log-normal random variables and applications to finance*. forthcoming, 2013. (Cited on page 213.)

[Hagan 1999] P. S. Hagan and D. E. Woodward. *Equivalent Black Volatilities*. Applied Mathematical Finance, vol. 6, pages 147–157, 1999. (Cited on pages 13, 14, 19, 30, 57, 67 and 122.)

[Hagan 2002] P. S. Hagan, D. Kumar, A. S. Lesniewski and D. E. Woodward. *Managing smile risk*. Willmott Magazine, pages 84–108, 2002. (Cited on pages 13, 82 and 122.)

[Henry-Labordère 2005] P. Henry-Labordère. *A general asymptotic implied volatility for stochastic volatility models*. arXiv:condmat/04317, 2005. (Cited on page 122.)

[Henry-Labordère 2008] P. Henry-Labordère. Analysis, geometry, and modeling in finance: Advanced methods in option pricing. Chapman and Hall, 2008. (Cited on pages 14, 19, 30, 57 and 122.)

[Heston 1993] S. Heston. *A closed-form solutions for options with stochastic volatility with applications to bond and currency options,*. Review of Financial Studies, vol. 6, pages 327–343, 1993. (Cited on pages 2, 81 and 122.)

[Hong 2004] G. Hong. *Forward Smile and Derivative Pricing*. `http://www.globalriskguard.com/resources/deriv/fwd_3.pdf`, 2004. (Cited on page 82.)

[Hull 1990] J. C. Hull and A. White. *Pricing interest-rate derivative securities*. Rev. Finan. Stud, vol. 3, pages 573–592, 1990. (Cited on page 2.)

[Ilhan 2009] A. Ilhan, M. Jonsson and R. Sircar. *Optimal static dynamic hedges for exotic options under convex risk measures*. Stochastic Processes and Applications, vol. 119, no. 10, pages 3608–3632, 2009. (Cited on page 169.)

[Jacquier 2011] A. Jacquier, M. Keller-Ressel and A. Mijatovic. *Large deviations in stochastic volatility with jumps: asymptotic implied volatility for affine models*. preprint, http://arxiv.org/abs/1108.3998, 2011. (Cited on page 9.)

[Jacquier 2012] A. Jacquier and P. Roome. *Asymptotics of forward implied volatilitys*. Preprint Available at http://arxiv.org/abs/1212.0779, 2012. (Cited on page 83.)

[Jordan 2011] R. Jordan and C. Tier. *Asymptotic Approximations to Deterministic and Stochastic Volatility Models*. SIAM J. Fin. Math., vol. 2, no. 1, pages 935–964, 2011. (Cited on page 122.)

[Karatzas 1991] I. Karatzas and S. Shreve. Brownian motion and stochastic calculus. 1991. (Cited on pages 173 and 230.)

[Kato 2012] T. Kato, A. Takahashi and T. Yamada. *An Asymptotic Expansion for Solutions of Cauchy-Dirichlet Problem for Second Order Parabolic PDEs and its Application to Pricing Barrier Options*. Rapport technique, 2012. (Cited on page 169.)

[Kebaier 2005] A. Kebaier. *Statistical Romberg extrapolation: a new variance reduction method and applications to option pricing*. The Annals of Applied Probability, vol. 15, no. 4, pages 2681–2705, 2005. (Cited on page 221.)

[Keller-Ressel 2008] M. Keller-Ressel and F. Kilin. *Forward-start options in the Barndorff-Nielsen-Shephard Model*. CQPF Working Paper Series at the Frankfurt School of Finance 18, 2008. (Cited on page 82.)

[Keller 1978] J. Keller. *Rays, waves and asymptotics*. Bull. Amer. Math. Soc, vol. 84, no. 78, pages 727–750, 1978. (Cited on page 122.)

[Kevorkian 1985] J. Kevorkian and J.D. Cole. Perturbation methods in applied mathematics. Springer-Verlag, 1985. (Cited on page 17.)

[Kloeden 1995] P.E. Kloeden and E. Platen. Numerical solution of stochastic differential equations. Springer Verlag, 1995. (Cited on page 14.)

[Kloeden 2010] P.E. Kloeden and E. Platen. Numerical solution of stochastic differential equations. 4th corrected printing. Applications of Mathematics 23. Berlin: Springer, 2010. (Cited on page 213.)

[Kruse 2005] S. Kruse and U. Nogel. *On the pricing of forward starting options in Hestons model on stochastic volatility*. Finance and Stochastics, vol. 9, no. 2, pages 233–250, 2005. (Cited on page 82.)

[Kunita 1984] H. Kunita. *Stochastic flows of diffeomorphisms*. École d'été de Probabilités de St-Flour XII. Lecture notes., vol. 1097, 1984. (Cited on pages 31 and 104.)

[Kunita 1997] H. Kunita. Stochastic flows and stochastic differential equations. Cambridge Studies in Advanced Mathematics. 24. Cambridge: Cambridge University Press, 1997. (Cited on page 216.)

[Kunitomo 2001] N. Kunitomo and A. Takahashi. *The Asymptotic Expansion Approach to the Valuation of Interest Rate Contingent Claims*. Mathematical Finance, vol. 11, no. 1, pages 117–151, 2001. (Cited on page 12.)

[Kusuoka 2004] S. Kusuoka. *Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus*. Kusuoka, Shigeo (ed.) et al., Advances in mathematical economics. Vol. 6. Tokyo: Springer. Adv. Math. Econ. 6, 69-83 (2004)., 2004. (Cited on page 213.)

[Lee 2005] R.W. Lee. *Implied Volatility: Statics, Dynamics, and Probabilistic Interpretation*. In Recent Advances in Applied Probability. Springer, 2005. (Cited on page 7.)

[Lemor 2006] J.P. Lemor, E. Gobet and X. Warin. *Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations*. Bernoulli, vol. 12, no. 5, pages 889–916, 2006. (Cited on page 212.)

[Lewis 2000] A.L. Lewis. Option valuation under stochastic volatility with mathematica code. Finance Press, Newport Beach, California, 2000. (Cited on pages 4, 145 and 157.)

[Lewis 2007] A. Lewis. *Geometries and smile asymptotics for a class of Stochastic Volatility models*. www.optioncity.net, 2007. (Cited on page 122.)

[Lorig 2013a] M. Lorig, S. Pagliarani and A. Pascucci. *Implied vol for any local-stochastic vol model*. Preprint SSRN, 2013. (Cited on pages 122 and 213.)

[Lorig 2013b] M. Lorig, S. Pagliarani and A. Pascucci. *A Taylor series approach to pricing and implied vol for LSV models*. Preprint, vol. 10, no. 1, 2013. (Cited on page 14.)

[Lucic 2003] V. Lucic. *Forward-start options in stochastic volatility models*. Wilmott Magazine, 2003. (Cited on page 82.)

[Lyons 2004] T. Lyons and N. Victoir. *Cubature on Wiener space*. Proc. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci., vol. 460, no. 2041, pages 169–198, 2004. (Cited on page 213.)

[Maruhn 2009] J. Maruhn and E. Sachs. *Robust static hedging of barrier options in stochastic volatility models*. Mathematical Methods of Operations Research, vol. 70, no. 13, pages 405–433, 2009. (Cited on page 169.)

[Merton 1976] R. Merton. *Option pricing when underlying stock returns are discontinuous*. Journal of Financial Economics, vol. 3, no. 1-2, pages 125–144, 1976. (Cited on page 2.)

[Musiela 2005] M. Musiela and M. Rutkowski. Martingale methods in financial modelling. Springer Verlag, second édition, 2005. (Cited on page 7.)

[Ninomiya 2008] S. Ninomiya and N. Victoir. *Weak approximation of stochastic differential equations and application to derivative pricing*. Appl. Math. Finance, vol. 15, no. 2, pages 107–121, 2008. (Cited on page 213.)

[Nualart 2006] D. Nualart. Malliavin calculus and related topics. Springer Verlag, Berlin, second édition, 2006. (Cited on pages 41, 65, 135, 136, 137, 148, 170, 217, 225, 226, 227 and 237.)

[Osajima 2007] 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. (Cited on page 12.)

[Pagliarani 2011] S. Pagliarani and A. Pascucci. *Analytical approximation of the transition density in a local volatility model*. Central European Journal of Mathematics, vol. 10, no. 1, pages 250–270, 2011. (Cited on pages 14 and 18.)

[Pagliarani 2013a] S. Pagliarani and A. Pascucci. *Local Stochastic Volatility with Jumps*. to appear in Int. J. Theor. Appl. Finance, 2013. (Cited on pages 14 and 122.)

[Pagliarani 2013b] S. Pagliarani, A. Pascucci and C. Riga. *Adjoint expansions in local Levy models*. SIAM J. Financial Math., vol. 4, no. 1, 2013. (Cited on page 14.)

[Papanicolaou 1999] G. Papanicolaou and R. Sircar. *Stochastic Volatility, Smile and Asymptotics*. Applied Mathematical Finance, vol. 6, no. 2, pages 107–145, 1999. (Cited on page 15.)

[Pham 2010] H. Pham. *Large deviations in mathematical finance*. arXiv, 2010. (Cited on page 168.)

[Protter 2004] P. Protter. Stochastic integration and differential equations. 2004. (Cited on pages 106 and 125.)

[Reiner 1991] E. Reiner and M. Rubinstein. *Breaking down the barriers*. Risk, vol. 4, no. 8, pages 28–35, 1991. (Cited on page 168.)

[Renault 1996] E. Renault and N. Touzi. *Option hedging and implied volatilities in a stochastic volatility model*. Mathematical Finance, vol. 6, pages 279–302, 1996. (Cited on page 157.)

[Revuz 1999] D. Revuz and M. Yor. Continuous martingales and brownian motion. Springer-Verlag, Hardcover, 3rd édition, 1999. (Cited on pages 87 and 173.)

[Rogers 2010] L.C.G. Rogers and M.R. Tehranchi. *Can the implied volatility surface move by parallel shifts?* Finance and Stochastics, vol. 14, pages 235–248, 2010. (Cited on page 8.)

[Schachermayer 2008] W. Schachermayer and J. Teichmann. *How close are the option pricing formulas of Bachelier and Black-Merton-Scholes?* Math. Finance, vol. 18, no. 1, pages 155–170, 2008. (Cited on page 6.)

[Schonbucher 1999] P. Schonbucher. *A market model for stochastic implied volatility*. Philosophical Transactions of the Royal Society, vol. Series A, no. 357, pages 2071–2092, 1999. (Cited on page 83.)

[Schroder 1989] M. Schroder. *Computing the Constant Elasticity of Variance Option Pricing Formula*. Journal of Finance, vol. 44, pages 211–219, 1989. (Cited on pages 55, 94, 192 and 239.)

[Scott 1987] L. Scott. *Option Pricing when the Variance changes randomly : Theory, Estimation, and an Application*. J. Financial and Quantitative Analysis, vol. 22, pages 419–438, 1987. (Cited on page 15.)

[Talay 1990] D. Talay and L.Tubaro. *Expansion of the global error for numerical schemes solving stochastic differential equations*. Stochastic Analysis and Applications, vol. 8-4, pages 94–120, 1990. (Cited on page 221.)

[Tehranchi 2009] M.R. Tehranchi. *Asymptotics of implied volatility far from maturity*. Journal of Applied Probability, vol. 46, no. 3, pages 629–650, 2009. (Cited on page 8.)

[Thalmaier 1998] A. Thalmaier and F.Y. Wang. *Gradient estimates for harmonic functions on regular domains in Riemannian manifolds*. J. Funct. Anal., vol. 155, no. 1, pages 109–124, 1998. (Cited on pages 170 and 174.)

[Uchida 2004] M. Uchida and N. Yoshida. *Asymptotic Expansion for Small Diffusions Applied to Option Pricing*. Statistical Inference for Stochastics Processes, vol. 7, pages 189–223, 2004. (Cited on page 12.)

[Watanabe 1987] S. Watanabe. *Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels*. Annals of Probability, vol. 15, no. 1, pages 1–39, 1987. (Cited on pages 12 and 14.)

[Wu 2010] Q. Wu. *Series Expansion of the SABR Joint Density*. Mathematical Finance, 2010. (Cited on page 82.)

[Yoshida 1992a] N. Yoshida. *Asymptotic Expansion for Statistics Related to Small Diffusions*. J. Japan Statistical Society, vol. 22, no. 2, pages 139–159, 1992. (Cited on page 12.)

[Yoshida 1992b] N. Yoshida. *Asymptotic Expansions of Maximum Likelihood Estimator for Small Diffusion via the Theory of Malliavin-Watanabe*. Probability Theory and Related Fields, vol. 92, pages 275–311, 1992. (Cited on page 12.)