Price expansion formulas for model combining local and stochastic volatility

Romain Bompis, Emmanuel Gobet

Price expansion formulas for model combining local and stochastic volatility

R. Bompis* and E. Gobet†
CMAP, Ecole Polytechnique
Route de Saclay, 91128 Palaiseau cedex, France

First version: July 17th 2012.
This version: December 12th 2012.
Working paper in progress.

Abstract

This paper 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.

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 and al in [HKLW02] or more recently the CEV-Heston model studied notably by Forde et al [FP12] (see [Cox75] for the CEV model and [Hes93] for the Heston model). If the interest of such models is increasing, their use is still challenging because of the lack of closed formulas. The price to be paid for the more realistic dynamic is 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 [HKLW02], Berestycki

*Email: romain.bompis@polytechnique.edu
†Email: emmanuel.gobet@polytechnique.edu

and al [BBF04], Henry Labordère [Lab05]-[Lab08] and Lewis [Lew07] 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 [JT11] who utilise the WKB or ray method (see [Kel78]) 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 [FJ12]: using small noise expansions inspired by [FW98] and large deviation argument, 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 [FP12] the large-time asymptotic of SABR and CEV-Heston models in different regime of strikes using 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 [CFK10] 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 [PP12] which provides an expansion of the characteristic function in a general local and stochastic volatility model (possibly incorporating also Levy 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 perturbation in the work of Hagan et al [HW99] 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.

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 [BGM09] and [BGM10a] to perform non-asymptotic expansion with the help of a Proxy process.

▷**Comparison with previous works and contribution of the paper.** 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 [BGM10b] 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 [BGM10a] or [BGM09], with a suitable paremetrization 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 [BGM10a] for pure local volatility models and of [BGM10b] 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 [AE08] 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):

$$\mathrm{d}X_t = \sigma(t, X_t)\sqrt{V_t}dW_t - \frac{1}{2}\sigma^2(t, X_t)V_t\mathrm{d}t,\ X_0 = x_0, \tag{1.1}$$

$$\mathrm{d}V_t = \alpha_t\mathrm{d}t + \xi_t\sqrt{V_t}\mathrm{d}B_t,\ V_0 = v_0, \tag{1.2}$$

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

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 (1.2) for the sake of clarity in the calculus, but the results developed throughout the paper 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 $y \to e^{\int_0^t \kappa_s\mathrm{d}s}y$ allows us namely to retrieve the formulation (1.2). See details in Appendix A.1.
Our aim is to give an accurate analytical approximation of any European option price of the form:

$$\mathbb{E}[h(X_T)], \tag{1.3}$$

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 (1.1)-(1.2) by the following Gaussian process:

$$\mathrm{d}X_t^P = \sigma(t, x_0)\sqrt{v_t}\mathrm{d}W_t - \frac{1}{2}\sigma^2(t, x_0)v_t\mathrm{d}t, \tag{1.4}$$

$$v_t = v_0 + \int_0^t \alpha_s\mathrm{d}s. \tag{1.5}$$

Such an approximation can be justified if the volatility of volatility is quite small leading to the approximation $V_t \approx v_t$ (which is realistic in practice) **and** if one of the two following situations holds: i) 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 $V_t \approx v_0$, $t \in [0, T]$).
To rely the initial process (1.1)-(1.2) to the proxy process (1.4)-(1.5), we introduce a two-dimensional parameterized process given by:

$$\mathrm{d}X_t^\eta = \sigma(t, \eta X_t^\eta + (1-\eta)x_0)\sqrt{V_t^\eta}\mathrm{d}W_t - \frac{1}{2}\sigma^2(t, \eta X_t^\eta + (1-\eta)x_0)V_t^\eta\mathrm{d}t,\ \ X_0^\eta = x_0, \tag{1.6}$$

$$\mathrm{d}V_t^\eta = \alpha_t \mathrm{d}t + \eta \xi_t \sqrt{V_t^\eta}\mathrm{d}B_t,\ v_0, \tag{1.7}$$

where η is an interpolation parameter lying in the range $[0,1]$, so 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 rely X and the proxy model and to derive successive corrective processes in order to obtain a tractable approximation formula.

▷**Outline of the paper.** The paper is organised as follows. In Section 2 we present a third order price approximation formula in Theorem 2.2.1 which is the main result of the paper. 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 drawn the reader attention to the main difficulties. The explicit calculus of the expansion coefficients is postponed to Appendix A.2. Section 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 paper. In the Section 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 4.1.1 and 4.2.1. Section 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 A, we give intermediate and complementary results.

2 Main Result

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 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:

$$\omega(l_1,\cdots,l_n)_t^T = \omega(l_1\omega(l_2,\cdots,l_n)_\cdot^T)_t^T,$$

for $t \in [0,T]$.

Definition 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 $(V_t)_{t\leq T}$.**

- $(\mathcal{H}_{x_0})$: 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

$$\mathcal{M}_1(\sigma) = \max_{1\leq i\leq 3} |\partial^i_{x^i}\sigma(t,x)|_\infty \text{ and } \mathcal{M}_0(\sigma) = \max_{0\leq i\leq 3} |\partial^i_{x^i}\sigma(t,x)|_\infty.$$

  In addition, we assume the following ellipticity condition: $\int_0^T \sigma_t^2 v_t \mathrm{d}t > 0$.

- (P): α and ξ are measurable, bounded on $[0,T]$ and positive. In addition $\xi_{\inf} > 0$ and $2(\frac{\alpha}{\xi^2})_{\inf} \geq 1$.

Remark 2.1.1. *Because there exists a unique process $(V_t)_{t\leq T}$ satisfying the SDE* (1.2)*, $(\mathcal{H}_{x_0})$ guarantees the existence and the uniqueness of a solution for* (1.1)*, considering generalized stochastic integration w.r.t. semi-martingales (see [Pro04, 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 [BGM10b, 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 2.1.3. $\Lambda_t^\eta = \sqrt{V_t^\eta}$, $\forall t\in[0,T]$, $\forall\eta\in[0,1]$.

In addition, we introduce $(\lambda_t)_{t\in[0,T]}$ defined for any $t\in[0,T]$ by:

$$\lambda_t = \Lambda_t^{\eta=0} = \sqrt{V_t^0} = \sqrt{v_t} = \sqrt{v_0 + \int_0^t \alpha_s \mathrm{d}s}. \tag{2.1}$$

▷**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:

$$\begin{cases} |h(x)| \leq C_h & \forall x\in\mathbb{R}, \\ |\frac{h(y)-h(x)}{y-x}| \leq L_h & \forall (x,y)\in\mathbb{R}^2, x\neq y, \end{cases}$$

This space includes the classical Put payoff function $x\to(K-e^x)_+$ 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 [AP06]).

▷**Generic constants and upper bounds.** We keep the same notation C for all non-negative constants depending on: universal constants, on a number $p\geq 1$ arising in L^p estimates, in a non decreasing way on $\xi_{\sup}$, $\mathcal{M}_0(\sigma)$, $\mathcal{M}_1(\sigma)$, T, $\frac{|\sigma|_\infty^2 T}{\int_0^T \sigma_t^2 v_t \mathrm{d}t}$, v_0, $1/v_0$, $\alpha_{\sup}$,and $|\rho|_\infty$. A generic constant **does not depend** on x_0. We frequently use the short notation $A \leq_c B$ for positive A which means that $A\leq CB$ for a generic constant C. Similarly "$A = O(B)$" means that $|A|\leq CB$ for a generic constant C.

▷**Miscellaneous.** The L^p norm of a random variable is denoted, as usual, by $\|.\|_p$

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

2.2 Third order approximation price formula

We state the main result of the paper:

Theorem 2.2.1. *(3rd order approximation price formula.) Assume* $(\mathcal{H}_{x_0})$ *and* (P)*. Then for any* $h \in \mathrm{Lip_b}(\mathbb{R})$*, we have:*

$$\mathbb{E}[h(X_T)] = \mathbb{E}[h(X_T^P)] + \sum_{i=1}^{6} \eta_{i,T} \mathcal{G}_i^h(X_T^P) + \mathrm{Error}_{3,h}, \tag{2.2}$$

where:

$$\begin{aligned}
\eta_{1,T} =& \frac{C^l_{1,T}}{2} - \frac{C^l_{2,T}}{2} - \frac{C^l_{3,T}}{4} - \frac{C^l_{4,T}}{2} - C^{ls}_{1,T},\\
\eta_{2,T} =& -\frac{3C^l_{1,T}}{2} + \frac{C^l_{2,T}}{2} + \frac{5C^l_{3,T}}{4} + \frac{7C^l_{4,T}}{2} + \frac{(C^l_{1,T})^2}{8} - \frac{C^s_{1,T}}{2}\\
&+ \frac{C^s_{3,T}}{4} + C^{ls}_{1,T} + \frac{C^{ls}_{2,T}}{2} + \frac{C^{ls}_{3,T}}{2} + C^{ls}_{4,T} + \frac{C^{ls}_{5,T}}{2} + \frac{C^{ls}_{6,T}}{4},\\
\eta_{3,T} =& C^l_{1,T} - 2C^l_{3,T} - 6C^l_{4,T} - \frac{3(C^l_{1,T})^2}{4} + \frac{C^s_{1,T}}{2} - \frac{C^s_{2,T}}{2} - \frac{C^s_{3,T}}{2}\\
&- \frac{3C^{ls}_{2,T}}{2} - \frac{3C^{ls}_{3,T}}{2} - \frac{5C^{ls}_{4,T}}{2} - C^{ls}_{5,T} - \frac{3C^{ls}_{6,T}}{4} - \frac{C^l_{1,T}C^s_{1,T}}{4},\\
\eta_{4,T} =& C^l_{3,T} + 3C^l_{4,T} + \frac{13(C^l_{1,T})^2}{8} + \frac{C^s_{2,T}}{2} + \frac{C^s_{3,T}}{4} + C^{ls}_{2,T}\\
&+ C^{ls}_{3,T} + \frac{3C^{ls}_{4,T}}{2} + \frac{C^{ls}_{5,T}}{2} + \frac{C^{ls}_{6,T}}{2} + \frac{(C^s_{1,T})^2}{8} + C^l_{1,T}C^s_{1,T},\\
\eta_{5,T} =& -\frac{3(C^l_{1,T})^2}{2} - \frac{(C^s_{1,T})^2}{4} - \frac{5C^l_{1,T}C^s_{1,T}}{4},\\
\eta_{6,T} =& \frac{(C^l_{1,T})^2}{2} + \frac{(C^s_{1,T})^2}{8} + \frac{C^l_{1,T}C^s_{1,T}}{2},
\end{aligned}$$

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:

$$\mathrm{Error}_{3,h} = O(L_h|\sigma|_\infty[\xi^3_{\sup} + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2]T^2). \tag{2.3}$$

Remark 2.2.1. *Under* $(\mathcal{H}_{x_0})$*,* X_T^P *is a non degenerate normal random variable and consequently, whatever is the regularity of h, the Greeks* $\mathcal{G}_h^i(X_T^P)$ *introduced in* (2.1.2) *exist and are well defined for any integer i. Note also that on the contrary to [BGM10b, Theorem 2.2], we do not assume anymore that the correlation is bounded away to* -1 *and* 1*.*

Remark 2.2.2. *The magnitude of* $\text{Error}_{3,\text{h}}$ *provided in* (2.3) *justifies the label of third order approximation formula because using the notation* $M = \max(\mathcal{M}_0(\sigma), \xi_{\text{sup}})$*, 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* $\max(\mathcal{M}_1(\sigma), \xi_{\text{sup}}) = 0$ *or* $T = 0$*, the approximation formula* (2.2) *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 2.2.3. *If one prefers to restrict to a second order approximation, it simply writes:*

$$\begin{aligned}\mathbb{E}[h(X_T)] =& \mathbb{E}[h(X_T^P)] + C_{1,T}^l[\frac{1}{2}\mathcal{G}_1^h(X_T^P) - \frac{3}{2}\mathcal{G}_2^h(X_T^P) + \mathcal{G}_3^h(X_T^P)] + \frac{C_{1,T}^s}{2}[-\mathcal{G}_2^h(X_T^P) + \mathcal{G}_3^h(X_T^P)] \\ &+ O(L_h|\sigma|_\infty[\xi_{\text{sup}}^2 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\text{sup}})]T^{\frac{3}{2}}).\end{aligned}$$

We let the reader verify that the additional corrective terms of the expansion (2.2) *are bounded up to generic constants by* $L_h|\sigma|_\infty[\xi_{\text{sup}}^2 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\text{sup}})]T^{\frac{3}{2}}$ *using standard upper bounds for the derivatives of the Gausssian density and the magnitude of the additional coefficients C.*

2.3 Corollaries and outline of the proof

▷**Particular cases of pure local volatility model and pure stochastic volatility model.**
a) Observe that if ξ_{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 [BGM10a] (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 [BGM10b] (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 [BGM10b, 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 $C_{3,T}^s$.

▷**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 $h \notin \text{Lip}_{\text{b}}(\mathbb{R})$ (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 (1.6)-(1.7). We recall that the Gaussian proxy process $(X_t^P)_{t\in[0,T]}$ defined in (1.4) is obtained by setting $\eta = 0$. The below 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 (1.6)-(1.7) w.r.t. η and by taking $\eta = 0$ thereafter. For the first corrective processes, we obtain:

$$\begin{aligned}\text{d}X_{1,t} =& [(X_t^P - x_0)\sigma_t^{(1)}\lambda_t + \Lambda_{1,t}\sigma_t](\text{d}W_t - \sigma_t\lambda_t\text{d}t), \qquad X_{2,0} = 0, \qquad (2.4)\\ V_{1,t} =& \int_0^t \xi_s\lambda_s\text{d}B_s, \qquad (2.5)\end{aligned}$$

$$\Lambda_{1,t} = \frac{V_{1,t}}{2\lambda_t}. \tag{2.6}$$

The second corrective processes are:

$$\begin{aligned} \mathrm{d}X_{2,t} =& \{\lambda_t[(X_t^P - x_0)^2\sigma_t^{(2)} + 2X_{1,t}\sigma_t^{(1)}] + 2(X_t^P - x_0)\Lambda_{1,t}\sigma_t^{(1)}\}(\mathrm{d}W_t - \sigma_t\lambda_t\mathrm{d}t) \\ &+ \{\Lambda_{2,t}\sigma_t\mathrm{d}W_t - [(X_t^P - x_0)V_{1,t}\sigma_t^{(1)}\sigma_t + (X_t^P - x_0)^2(\sigma_t^{(1)})^2 v_t + \frac{V_{2,t}}{2}\sigma_t^2]\mathrm{d}t\},\ X_{2,0} = 0 \end{aligned} \tag{2.7}$$

$$V_{2,t} = \int_0^t \xi_s \frac{V_{1,s}}{\lambda_t}\mathrm{d}B_s, \tag{2.8}$$

$$\Lambda_{2,t} = \frac{V_{2,t}}{2\lambda_t} - \frac{V_{1,t}^2}{4(\lambda_t)^3}. \tag{2.9}$$

The reader will notice that under $(\mathcal{H}_{x_0})$, 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 $x = X_T^P$:

$$\begin{aligned} \mathbb{E}[h(X_T)] =& \mathbb{E}[h(X_T^P)] + \mathbb{E}[h^{(1)}(X_T^P)(X_T - X_T^P)] + \frac{1}{2}\mathbb{E}[h^{(2)}(X_T^P)(X_T - X_T^P)^2] \\ &+ \mathbb{E}[(X_T - X_T^P)^3 \int_0^1 h^{(3)}(X_T^P + \eta(X_T - X_T^P))\frac{(1-\eta)^2}{2}\mathrm{d}\eta] \\ =& \mathbb{E}[h(X_T^P)] + \mathbb{E}[h^{(1)}(X_T^P)X_{1,T}] + \mathbb{E}[h^{(1)}(X_T^P)\frac{X_{2,T}}{2}] + \frac{1}{2}\mathbb{E}[h^{(2)}(X_T^P)X_{1,T}^2] + \mathrm{Error}_{3,\mathrm{h}}, \\ \mathrm{Error}_{3,h} =& \mathbb{E}[h^{(1)}(X_T^P)(X_T - \sum_{j=0}^{2}\frac{X_{j,T}}{j!})] + \frac{1}{2}\mathbb{E}[h^{(2)}(X_T^P)(X_T - X_T^P - X_{1,T})(X_T - X_T^P + X_{1,T})] \\ &+ \mathbb{E}[(X_T - X_T^P)^3 \int_0^1 h^{(3)}(X_T^P + \eta(X_T - X_T^P))\frac{(1-\eta)^2}{2}\mathrm{d}\eta], \end{aligned} \tag{2.10}$$

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 A.2:

Lemma 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:*

$$\begin{aligned} &\mathbb{E}\left(\varphi(\int_0^T f_t\mathrm{d}W_t)\int_0^T l_{N,t_N}\int_0^{t_N} l_{N-1,t_N-1}\dots\int_0^{t_2} l_{1,t_1}\mathrm{d}W_{t_1}^{I_1}\dots\mathrm{d}W_{t_{N-1}}^{I_{N-1}}\mathrm{d}W_{t_N}^{I_N}\right) \\ &= \omega(\widehat{l_1},\dots,\widehat{l_N})\mathcal{G}^{\varphi}_{\#\{k:I_k\neq 0\}}(\int_0^T f_t\mathrm{d}W_t), \end{aligned} \tag{2.11}$$

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 (2.2) is given in Appendix A.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 (2.10), 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 formulas like in [BGM10b]. But a straightforward application of this methodology using the representation (2.10) 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 do not satisfy the standard assumptions. Malliavin differentiability is studied by hand in [AE08] 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 [AP06].

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 (3.17)), we can write:

$$\begin{aligned}
\mathbb{E}[h_\delta(X_T)] =& \mathbb{E}[h_\delta(X_T^P + X_{1,T} + \frac{X_{2,T}}{2})] + \mathbb{E}[(X_T - \sum_{j=0}^{2} \frac{X_{j,T}}{j!}) \int_0^1 h_\delta^{(1)}((1-\eta) \sum_{j=0}^{2} \frac{X_{j,T}}{j!} + \eta X_T) \mathrm{d}\eta] \\
=& \mathbb{E}[h_\delta(X_T^P)] + \mathbb{E}[h_\delta^{(1)}(X_T^P)(X_{1,T} + \frac{X_{2,T}}{2})] + \frac{1}{2}\mathbb{E}[h_\delta^{(2)}(X_T^P)X_{1,T}^2] + \mathrm{Error}_{3,h_\delta}, \qquad (2.12) \\
\mathrm{Error}_{3,h_\delta} =& \mathbb{E}\big[(X_T - \sum_{j=0}^{2} \frac{X_{j,T}}{j!}) \int_0^1 h_\delta^{(1)}((1-\eta) \sum_{j=0}^{2} \frac{X_{j,T}}{j!} + \eta X_T) \mathrm{d}\eta\big] + \frac{1}{2}\mathbb{E}[h_\delta^{(2)}(X_T^P)(X_{1,T}X_{2,T} + \frac{X_{2,T}^2}{4})] \\
&+ \mathbb{E}\big[(X_{1,T} + \frac{X_{2,T}}{2})^3 \int_0^1 \frac{(1-\eta)^2}{2} h_\delta^{(3)}(X_T^P + \eta(X_{1,T} + \frac{X_{2,T}}{2})) \mathrm{d}\eta\big]. \qquad (2.13)
\end{aligned}$$

As h is supposed Lipschitz bounded, the first term of (2.13) which involves only the first derivative of h_δ can be handled without Malliavin calculus. The two last terms of (2.13) 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 [GM05] or [GM11]).
The complete analyse of the error is given in the below subsection.

3 Error analysis

We establish the estimate (2.3) in several steps:

1. L^p norms estimates of the residuals processes,
2. small noise perturbation to smooth the function h,
3. careful use of Malliavin integration by parts formulas to achieve the proof.

3.1 Approximation of X, V, Λ and error estimates

Approximation of V, Λ and error estimates

Definition 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* $\left((\Lambda_{j,t})_{t\in[0,T]}\right)_{j\in\{1,2\}}$ *are defined in* (2.6)-(2.9)*. Replacing* Λ *by* V*, we define similarly the* V*-residual processes using the notation* R^V*.*

Proposition 3.1.1. *Assume* (P)*. Then for any* $p \geq 1$*, we have:*

$$\sqrt{v_0} \leq \lambda_{\inf} \leq \lambda_{\sup} \leq \sqrt{v_0 + T\alpha_{\sup}}, \tag{3.1}$$

$$\sup_{t\in[0,T]} \|\Lambda_{i,t}\|_p \leq_c (\xi_{\sup}\sqrt{T})^i, \ \forall i \in \{1,2\}, \tag{3.2}$$

$$\sup_{t\in[0,T]} \|R^{\Lambda}_{i,t}\|_p \leq_c (\xi_{\sup}\sqrt{T})^{i+1}, \ \forall i \in \{0,1,2\}. \tag{3.3}$$

Proof. (3.1) is obvious in view of (2.1). The proofs of (3.2) and (3.3) can be found in [BGM10b, Propositions 4.6, 4.7 and 4.8] replacing in the original paper κ by zero and $\kappa\theta_t$ by α_t. □

Corollary 3.1.1. *Assume* (P)*. Then one has for any* $p \geq 1$*:*

$$v_0 \leq v_{\inf} \leq v_{\sup} \leq v_0 + T\alpha_{\sup}, \tag{3.4}$$

$$\sup_{t\in[0,T]} \|V_t\|_p \leq_c 1 + v_0, \tag{3.5}$$

$$\sup_{t\in[0,T]} \|V_{i,t}\|_p \leq_c (\xi_{\sup}\sqrt{T})^i, \ \forall i \in \{1,2\}, \tag{3.6}$$

$$\sup_{t\in[0,T]} \|R^{V}_{i,t}\|_p \leq_c (\xi_{\sup}\sqrt{T})^{i+1}, \ \forall i \in \{0,1,2\}. \tag{3.7}$$

Proof. The proof of (3.4) and (3.5) are easy. (3.6) are obtained readily with (2.5), (2.8) and (3.2). Proofs of (3.7) are available in [BGM10b, 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 3.1.2. *Assume* $(\mathcal{H}_{x_0})$*. 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* (2.4)-(2.7)*. 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:*

$$R_{n,\sigma}(X_t) = \sigma_t(X_t) - \sum_{i=0}^{n} \frac{(X_t - x_0)^i}{i!}\sigma^{(i)}_t. \tag{3.8}$$

Replacing σ *by* σ^2*, we use the similar notation* $R_{n,\sigma^2}(X_t)$*.*

Standard computations involving Burkholder-Davis-Gundy and Holder inequalities yield:

$$\begin{aligned}\|X_t - x_0\|_p^p \leq_c & t^{\frac{p}{2}-1} \int_0^t \|\sigma_s(X_s)\sqrt{V_s}\|_p^p \mathrm{d}s + t^{p-1} \int_0^t \|\sigma_s^2(X_s)V_s\|_p^p \mathrm{d}s \\ \leq_c & t^{\frac{p}{2}-1}|\sigma|_\infty^p \int_0^t \mathbb{E}[V_s^{p/2}]\mathrm{d}s + t^{p-1}|\sigma|_\infty^{2p} \int_0^t \mathbb{E}[V_s^p]\mathrm{d}s \leq_c (|\sigma|_\infty \sqrt{T})^p, \end{aligned} \tag{3.9}$$

for any $p \geq 2$, where we have applied the estimate (3.5) at the last line. We now intend to handle X-residual processes, and the next results are intermediate steps. In the below Lemma, we provide L^p estimates of $X_t^P - x_0$, $X_{1,t}$ and $X_{2,t}$.

Lemma 3.1.1. *Assume* $(\mathcal{H}_{x_0})$ *and* (P)*. For any* $p \geq 1$*:*

$$\sup_{t\in[0,T]} \|X_t^P - x_0\|_p \leq_c |\sigma|_\infty \sqrt{T}, \tag{3.10}$$

$$\sup_{t\in[0,T]} \|X_{i,t}\|_p \leq_c |\sigma|_\infty [\xi_{\sup}^i + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^{i-1}] T^{\frac{i+1}{2}}, \ \forall i \in \{1,2\}. \tag{3.11}$$

Proof. (3.10) is similar to (3.9). For (3.11) $i = 1$: starting from (2.4), the same computations as before give:

$$\|X_{1,t}\|_p \leq_c \mathcal{M}_1(\sigma)\sqrt{T}(1 + \mathcal{M}_0(\sigma)\sqrt{T}) \sup_{t\in[0,T]} \|X_t^P - x_0\|_p + |\sigma|_\infty \sqrt{T}(1 + \mathcal{M}_0(\sigma)\sqrt{T}) \sup_{t\in[0,T]} \|V_{1,t}\|_p.$$

We conclude using (3.10) and (3.6). For (3.11) $i = 2$, one has from (2.7):

$$\begin{aligned}\|X_{2,t}\|_p \leq_c & \mathcal{M}_1(\sigma)\sqrt{T}(1 + \mathcal{M}_0(\sigma)\sqrt{T})(\sup_{t\in[0,T]} \|(X_t^P - x_0)^2\|_p + \sup_{t\in[0,T]} \|X_{1,t}\|_p) + |\sigma|_\infty \sqrt{T}(1 + \mathcal{M}_0(\sigma)\sqrt{T}) \sup_{t\in[0,T]} \|V_{2,t}\|_p \\ & + |\sigma|_\infty \sqrt{T} \sup_{t\in[0,T]} \|V_{1,t}^2\|_p + \mathcal{M}_1(\sigma)\sqrt{T}(1 + \mathcal{M}_0(\sigma)\sqrt{T}) \sup_{t\in[0,T]} \|X_t^P - x_0\|_{2p} \sup_{t\in[0,T]} \|V_{1,t}\|_{2p}.\end{aligned}$$

We conclude using (3.10), (3.11) $i = 1$ and (3.6). □

We give in the following Lemma the explicit equations solved by the X-residual processes:

Lemma 3.1.2. *Assume* $(\mathcal{H}_{x_0})$ *and* (P)*. One has:*

$$\mathrm{d}R_{0,t}^X = [\lambda_t R_{0,\sigma}(X_t) + \sigma_t(X_t)R_{0,t}^\Lambda]\mathrm{d}W_t - \frac{1}{2}[v_t R_{0,\sigma^2}(X_t) + \sigma_t^2(X_t)R_{0,t}^V]\mathrm{d}t, \ R_{0,0}^X = 0, \tag{3.12}$$

$$\begin{aligned}\mathrm{d}R_{1,t}^X = & [\lambda_t R_{1,\sigma}(X_t) + \Lambda_{1,t}R_{0,\sigma}(X_t) + \sigma_t(X_t)R_{1,t}^\Lambda + \lambda_t\sigma_t^{(1)}R_{0,t}^X]\mathrm{d}W_t \\ & - \frac{1}{2}[v_t R_{1,\sigma^2}(X_t) + V_{1,t}R_{0,\sigma^2}(X_t) + \sigma_t^2(X_t)R_{1,t}^V + 2v_t\sigma_t\sigma_t^{(1)}R_{0,t}^X]\mathrm{d}t, \ R_{1,0}^X = 0,\end{aligned} \tag{3.13}$$

$$\begin{aligned}\mathrm{d}R_{2,t}^X = & [\lambda_t R_{2,\sigma}(X_t) + \Lambda_{1,t}R_{1,\sigma}(X_t) + \frac{\Lambda_{2,t}}{2}R_{0,\sigma}(X_t) + \sigma_t(X_t)R_{2,t}^\Lambda + \lambda_t\sigma_t^{(1)}R_{1,t}^X + \lambda_t\frac{\sigma_t^{(2)}}{2}R_{0,t}^X(X_t + X_t^P) \\ & + \Lambda_{1,t}\sigma_t^{(1)}R_{0,t}^X]\mathrm{d}W_t - \frac{1}{2}[v_t R_{2,\sigma^2}(X_t) + V_{1,t}R_{1,\sigma^2}(X_t) + \frac{V_{2,t}}{2}R_{0,\sigma^2}(X_t) + \sigma_t^2(X_t)R_{2,t}^V + 2v_t\sigma_t\sigma_t^{(1)}R_{1,t}^X \\ & + v_t((\sigma_t^{(1)})^2 + \sigma_t^{(2)}\sigma_t)R_{0,t}^X(X_t + X_t^P) + 2V_{1,t}\sigma_t\sigma_t^{(1)}R_{0,t}^X]\mathrm{d}t, \ R_{2,0}^X = 0.\end{aligned} \tag{3.14}$$

Proof. The verification of these identities is tedious but without mathematical difficulties. For convenience, we detail some computations. To obtain (3.12), start from (1.1) and (1.4) and write:

$$\begin{aligned}
\mathrm{d}R^X_{0,t} =&[\sigma_t(X_t)\Lambda_t - \sigma_t\lambda_t]\mathrm{d}W_t - \frac{1}{2}[\sigma_t^2(X_t)V_t - \sigma_t^2 v_t]\mathrm{d}t\\
=&[\lambda_t(\sigma_t(X_t) - \sigma_t) + \sigma_t(X_t)(\Lambda_t - \lambda_t)]\mathrm{d}W_t - \frac{1}{2}[v_t(\sigma_t^2(X_t) - \sigma_t^2) + \sigma_t^2(X_t)(V_t - v_t)]\mathrm{d}t\\
=&[\lambda_t R_{0,\sigma}(X_t) + \sigma_t(X_t)R^{\Lambda}_{0,t}]\mathrm{d}W_t - \frac{1}{2}[v_t R_{0,\sigma^2}(X_t) + \sigma_t^2(X_t)R^V_{0,t}]\mathrm{d}t.
\end{aligned}$$

Similarly for (3.13), using (3.12) and (2.4), we get:

$$\begin{aligned}
\mathrm{d}R^X_{1,t} =&\mathrm{d}R^X_{0,t} - \mathrm{d}X_{1,t}\\
=&[\lambda_t R_{0,\sigma}(X_t) + \sigma_t(X_t)R^{\Lambda}_{0,t}]\mathrm{d}W_t - \frac{1}{2}[v_t R_{0,\sigma^2}(X_t) + \sigma_t^2(X_t)R^V_{0,t}]\mathrm{d}t\\
&- [(X^P_t - x_0)\sigma^{(1)}_t\lambda_t + \Lambda_{1,t}\sigma_t](\mathrm{d}W_t - \sigma_t\lambda_t\mathrm{d}t)\\
=&[\lambda_t R_{0,\sigma}(X_t) + \Lambda_{1,t}R_{0,\sigma}(X_t) + \sigma_t(X_t)R^{\Lambda}_{1,t}]\mathrm{d}W_t - \frac{1}{2}[v_t R_{0,\sigma^2}(X_t) + V_{1,t}R_{0,\sigma^2}(X_t) + \sigma_t^2(X_t)R^V_{1,t}]\mathrm{d}t\\
&- (X^P_t - x_0)\sigma^{(1)}_t\lambda_t(\mathrm{d}W_t - \sigma_t\lambda_t\mathrm{d}t)\\
=&[\lambda_t R_{1,\sigma}(X_t) + \Lambda_{1,t}R_{0,\sigma}(X_t) + \sigma_t(X_t)R^{\Lambda}_{1,t} + \lambda_t\sigma^{(1)}_t R^X_{0,t}]\mathrm{d}W_t\\
&- \frac{1}{2}[v_t R_{1,\sigma^2}(X_t) + V_{1,t}R_{0,\sigma^2}(X_t) + \sigma_t^2(X_t)R^V_{1,t} + 2v_t\sigma_t\sigma^{(1)}_t R^X_{0,t}]\mathrm{d}t.
\end{aligned}$$

Now consider (3.14). Start from (3.13)-(2.7) 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_t^2(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)^2 v_t + \frac{V_{2,t}}{2}\sigma_t^2]\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\\
&- \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^{(1)}_t R^X_{1,t}]\mathrm{d}t\\
&- \frac{1}{2}\{\lambda_t(X^P_t - x_0)^2\sigma^{(2)}_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}[(X^P_t - x_0)V_{1,t}\sigma^{(1)}_t\sigma_t + (X^P_t - x_0)^2(\sigma^{(1)}_t)^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^{(1)}_t R^X_{1,t} + \lambda_t\frac{\sigma^{(2)}_t}{2}R^X_{0,t}(X_t + X^P_t)]\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^{(1)}_t)^2 + \sigma^{(2)}_t\sigma_t)R^X_{0,t}(X_t + X^P_t)\\
&+ 2v_t\sigma_t\sigma^{(1)}_t R^X_{1,t}]\mathrm{d}t - (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}(X^P_t - x_0)V_{1,t}\sigma^{(1)}_t\sigma_t\mathrm{d}t
\end{aligned}$$

$$
\begin{aligned}
=&[\lambda_t R_{2,\sigma}(X_t) + \Lambda_{1,t} R_{1,\sigma}(X_t) + \frac{\Lambda_{2,t}}{2} R_{0,\sigma}(X_t) + \sigma_t(X_t) R^{\Lambda}_{2,t} + \lambda_t \sigma^{(1)}_t R^X_{1,t} + \lambda_t \frac{\sigma^{(2)}_t}{2} R^X_{0,t}(X_t + X^P_t) \\
&+ \Lambda_{1,t} \sigma^{(1)}_t R^X_{0,t}]\mathrm{d}W_t - \frac{1}{2}[v_t R_{2,\sigma^2}(X_t) + V_{1,t} R_{1,\sigma^2}(X_t) + \frac{V_{2,t}}{2} R_{0,\sigma^2}(X_t) + \sigma^2_t(X_t) R^V_{2,t} + 2 v_t \sigma_t \sigma^{(1)}_t R^X_{1,t} \\
&+ v_t((\sigma^{(1)}_t)^2 + \sigma^{(2)}_t \sigma_t) R^X_{0,t}(X_t + X^P_t) + 2 V_{1,t} \sigma_t \sigma^{(1)}_t R^X_{0,t}]\mathrm{d}t.
\end{aligned}
$$

□

An intermediate result is the estimates of $R_{n,\sigma}(X_t)$ and $R_{n,\sigma^2}(X_t)$. Assuming $(\mathcal{H}_{x_0})$, from the Taylor-Lagrange inequality, we have $|R_{n,\sigma}(X_t)| \leq_c |X_t - x_0|^{n+1} \mathcal{M}_1(\sigma)$ and $|R_{n,\sigma^2}(X_t)| \leq_c |X_t - x_0|^{n+1} \mathcal{M}_0(\sigma)\mathcal{M}_1(\sigma)$. Combined with (3.9), this readily gives $\forall p \geq 2$ and $\forall j \in \{0, .., 2\}$:

$$
\sup_{t\in[0,T]} \|R_{j,\sigma}(X_t)\|_p \leq_c (|\sigma|_\infty \sqrt{T})^{j+1} \mathcal{M}_1(\sigma), \ \sup_{t\in[0,T]} \|R_{j,\sigma^2}(X_t\|_p \leq_c (|\sigma|_\infty \sqrt{T})^{j+1} \mathcal{M}_0(\sigma)\mathcal{M}_1(\sigma). \tag{3.15}
$$

We now state the result related to the estimates of the residuals processes:

Proposition 3.1.2. *Assume that* $(\mathcal{H}_{x_0})$ *and* (P) *hold. Then for any* $p \geq 1$*, we have:*

$$
\sup_{t\in[0,T]} \|R^X_{j,t}\|_p \leq_c |\sigma|_\infty \{\xi^{j+1}_{\sup} + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^j\} T^{\frac{j}{2}+1}, \ \forall j \in \{0, 1, 2\}. \tag{3.16}
$$

Proof. We leverage the explicit equation solved by the residuals $(R^X_{j,t})_{t\in[0,T]}$ (see Lemma 3.1.2). We begin with $R^X_{0,t}$. Starting from (3.12) and using standard inequalities, it readily follows:

$$
\|R^X_{0,t}\|_p \leq_c \sqrt{T}[\sup_{t\leq T} \|R_{0,\sigma}(X_t)\|_p + |\sigma|_\infty \sup_{t\leq T} \|R^{\Lambda}_{0,t}\|_p] + T[\sup_{t\leq T} \|R_{0,\sigma^2}(X_t)\|_p + |\sigma|^2_\infty \sup_{t\leq T} \|R^V_{0,t}\|_p].
$$

We conclude using (3.15)-(3.3)-(3.7). Similarly for $R^X_{1,t}$ given in (3.13), we obtain:

$$
\begin{aligned}
\|R^X_{1,t}\|_p \leq_c \sqrt{T}\{&\sup_{t\leq T} \|R_{1,\sigma}(X_t)\|_p + \sup_{t\leq T} \|\Lambda_{1,t} R_{0,\sigma}(X_t)\|_p + |\sigma|_\infty \sup_{t\leq T} \|R^{\Lambda}_{1,t}\|_p + \mathcal{M}_1(\sigma) \sup_{t\leq T} \|R^X_{0,t}\|_p\} \\
&+ T\{\sup_{t\leq T} \|R_{1,\sigma^2}(X_t)\|_p + \sup_{t\leq T} \|V_{1,t} R_{0,\sigma^2}(X_t)\|_p + |\sigma|^2_\infty \sup_{t\leq T} \|R^V_{1,t}\|_p + |\sigma|_\infty \mathcal{M}_1(\sigma) \sup_{t\leq T} \|R^X_{0,t}\|_p\}.
\end{aligned}
$$

Then, plugging in the above upper bound the estimates (3.2)-(3.3)-(3.6)-(3.7)-(3.15)-(3.16) $i = 0$, we complete the proof of (3.16) for $i = 1$. Finally for $R^X_{2,t}$, starting from (3.14), we readily have:

$$
\begin{aligned}
\|R^X_{2,t}\|_p \leq_c \sqrt{T}\{&\sup_{t\leq T} \|R_{2,\sigma}(X_t)\|_p + \sup_{t\leq T} \|\Lambda_{1,t} R_{1,\sigma}(X_t)\|_p + \sup_{t\leq T} \|\Lambda_{2,t} R_{0,\sigma}(X_t)\|_p + |\sigma|_\infty \sup_{t\leq T} \|R^{\Lambda}_{2,t}\|_p \\
&+ \mathcal{M}_1(\sigma) \sup_{t\leq T} \|R^X_{1,t}\|_p + \mathcal{M}_1(\sigma) \sup_{t\leq T} \|R^X_{0,t}(X_t + X^P_t)\|_p + \mathcal{M}_1(\sigma) \sup_{t\leq T} \|\Lambda_{1,t} R^X_{0,t}\|_p\} \\
&+ T\{\sup_{t\leq T} \|R_{2,\sigma^2}(X_t)\|_p + \sup_{t\leq T} \|V_{1,t} R_{1,\sigma^2}(X_t)\|_p + \sup_{t\leq T} \|V_{2,t} R_{0,\sigma^2}(X_t)\|_p + |\sigma|^2_\infty \sup_{t\leq T} \|R^V_{2,t}\|_p \\
&+ |\sigma|_\infty \mathcal{M}_1(\sigma) \sup_{t\leq T} \|R^X_{1,t}\|_p + \mathcal{M}_1(\sigma)\mathcal{M}_0(\sigma) \sup_{t\leq T} \|R^X_{0,t}(X_t + X^P_t)\|_p + |\sigma|_\infty \mathcal{M}_1(\sigma) \sup_{t\leq T} \|V_{1,t} R^X_{0,t}\|_p\}.
\end{aligned}
$$

The proof is completed as before using (3.2)-(3.3)-(3.6)-(3.7)-(3.15)-(3.16) $i = 0$ and 1. □

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's mean enlarge the initial filtration and sigma field. Then we define:

$$h_\delta(x) = \mathbb{E}[h(x + \delta \overline{W}_T)]. \tag{3.17}$$

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:

$$h_\delta(x) = \mathbb{E}[h_{\delta/\sqrt{2}}(x + \delta \overline{W}_{\frac{T}{2}})]. \tag{3.18}$$

In addition $h \in \mathrm{Lip_b}(\mathbb{R}) \Rightarrow \mathrm{h}_\delta \in \mathrm{Lip_b}(\mathbb{R})$ with $C_{h_\delta} = C_h$ and $L_{h_\delta} = L_h$. The next Lemma estimate 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 2.2.1.

Lemma 3.2.1. *Let $\delta > 0$. Assume that $h \in \mathrm{Lip_b}(\mathbb{R})$ and that $(\mathcal{H}_{x_0})$ is satisfied. Then we have:*

$$\Big|\mathbb{E}[h(X_T)] - \mathbb{E}[h_\delta(X_T)]\Big| + \Big|\mathbb{E}[h(X_T^P)] - \mathbb{E}[h_\delta(X_T^P)]\Big| \leq_c L_h \delta \sqrt{T},$$

$$\Big|\partial^i_{x^i}\mathbb{E}[h(X_T^P + x)]|_{x=0} - \partial^i_{x^i}\mathbb{E}[h_\delta(X_T^P + x)]|_{x=0}\Big| \leq_c L_h \frac{\delta\sqrt{T}}{(\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{i/2}}, \ \forall i \geq 1.$$

Proof. The first estimate is obvious using the lipschitzianity of h and classical estimates for the auxiliary Brownian motion. For the second write:

$$\mathbb{E}[h_\delta(X_T^P + x)] = \int_{\mathbb{R}} \mathbb{E}[h(y - \frac{\int_0^T \sigma_t^2 v_t \mathrm{d}t}{2} + \delta\overline{W}_T)] \frac{e^{-\frac{(y-x)^2}{2\int_0^T \sigma_t^2 v_t \mathrm{d}t}}}{\sqrt{2\pi \int_0^T \sigma_t^2 v_t \mathrm{d}t}} \mathrm{d}y,$$

to obtain:

$$\begin{aligned}
&\partial^i_{x^i}\mathbb{E}[h(X_T^P + x)]|_{x=0} - \partial^i_{x^i}\mathbb{E}[h_\delta(X_T^P + x)]|_{x=0} \\
&= \int_{\mathbb{R}} \mathbb{E}[h(y - \frac{\int_0^T \sigma_t^2 v_t \mathrm{d}t}{2} + \delta\overline{W}_T) - h(y - \frac{\int_0^T \sigma_t^2 v_t \mathrm{d}t}{2})] \partial^i_{x^i}\{\frac{e^{-\frac{(y-x)^2}{2\int_0^T \sigma_t^2 v_t \mathrm{d}t}}}{\sqrt{2\pi \int_0^T \sigma_t^2 v_t \mathrm{d}t}}\}|_{x=0} \mathrm{d}y.
\end{aligned}$$

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 2.2.1, applying Lemma 3.2.1, we readily obtain :

$$\begin{aligned}
\Big|\mathrm{Error}_{3,h}\Big| =& \Big|\mathbb{E}[h(X_T)] - \mathbb{E}[h(X_T^P)] - \sum_{i=1}^6 \eta_{i,T}\mathcal{G}_i^h(X_T^P)\Big| \\
\leq& \Big|\mathbb{E}[h(X_T)] - \mathbb{E}[h_\delta(X_T)]\Big| + \Big|\mathbb{E}[h_\delta(X_T^P)] - \mathbb{E}[h(X_T^P)]\Big| \\
&+ \sum_{i=1}^6 |\eta_{i,T}|\Big|\mathcal{G}_i^{h_\delta}(X_T^P) - \mathcal{G}_i^{h_\delta}(X_T^P)\Big| + \Big|\mathrm{Error}_{3,h_\delta}\Big|
\end{aligned}$$

$$\leq_c L_h\,\delta\,\sqrt{T} + \left|\text{Error}_{3,h_\delta}\right|.$$

Assume now without loss of generality that $\mathcal{M}_1(\sigma) + \xi_{\sup} \neq 0$. We prove the estimate (2.3) if we choose as value for δ:

$$\delta = |\sigma|_\infty[\xi_{\sup}^3 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2]\; T^{\frac{3}{2}} \tag{3.19}$$

and establish that:

$$|\text{Error}_{3,h_\delta}| \leq_c L_h|\sigma|_\infty[\xi_{\sup}^3 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2]T^2, \tag{3.20}$$

This is the purpose of the next subsection.

3.3 Malliavin integration by part formula and proof of estimate (3.20)

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 $D(.) = (D_t^1(.))_{t\in[0,T]}, (D_t^2(.))_{t\in[0,T]}, (D_t^3(.))_{t\in[0,T]}$. 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 [Nua06] 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 3.3.1. *Assume that $(\mathcal{H}_{x_0})$ 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 $q, r, s, t \in [0,T]$:*

$$\|D_s^1 X_t^P\|_p + \|D_s^2 X_t^P\|_p \leq_c |\sigma|_\infty, \tag{3.21}$$

$$\|D_s^1 \Lambda_{n,t}\|_p + \|D_s^1 V_{n,t}\|_p \leq_c \xi_{\sup}^n T^{\frac{n-1}{2}},\ \forall n \in \{1,2\}, \tag{3.22}$$

$$\|D_s^1 X_{n,t}\|_p + \|D_s^2 X_{n,t}\|_p \leq_c |\sigma|_\infty[\xi_{\sup}^n + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^{n-1}]T^{\frac{n}{2}},\ \forall n \in \{1,2\} \tag{3.23}$$

$$\|D_{r,s}^{1,1} \Lambda_{2,t}\|_p + \|D_{r,s}^{1,1} V_{2,t}\|_p \leq_c \xi_{\sup}^2, \tag{3.24}$$

$$\sum_{i,j\in\{1,2\}} \|D_{r,s}^{i,j} X_{n,t}\|_p \leq_c |\sigma|_\infty[\xi_{\sup}^n + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^{n-1}]T^{\frac{n-1}{2}},\ \forall n \in \{1,2\}, \tag{3.25}$$

$$\sum_{i,j,k\in\{1,2\}} \|D_{q,r,s}^{i,j,k} X_{2,t}\|_p \leq_c |\sigma|_\infty[\xi_{\sup}^2 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})]. \tag{3.26}$$

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 (1.4)-(2.4)-(2.7)-(2.5)-(2.8)). Then the calculus of the derivatives and the L^p-estimates is 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 3.3.2. *Assume $(\mathcal{H}_{x_0})$ 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 $\forall i \in \{2,3\}$:*

$$\mathbb{E}[Y h_{\delta/\sqrt{2}}^{(i)}(G_\delta^\eta)] = \mathbb{E}[Y_{i,\eta} h_{\delta/\sqrt{2}}^{(1)}(G_\delta^\eta)], \tag{3.27}$$

where for any $p \geq 1$ and any $i \in \{2,3\}$:

$$\sup_{\eta\in[0,1]} \|Y_{i,\eta}\|_p \leq_c \|Y\|_{i-1,p+\frac{1}{2}} \Big(\int_0^T \sigma_t^2 v_t \mathrm{d}t\Big)^{-\frac{(i-1)}{2}}. \tag{3.28}$$

We are now in position to achieve the proof of (3.20). Consider $\mathrm{Error}_{3,h_\delta}$ explicitly written in (2.13). The first term of (2.13) is handled easily using (3.16) $i = 3$. For the second term of (2.13), using (3.18), applying the Lemma 3.3.2 with $Y = X_{1,T}X_{2,T} + \frac{X_{2,T}^2}{4}$ and using (3.11)-(3.23), we obtain:

$$\Big|\mathbb{E}[h_\delta^{(2)}(X_T^P)(X_{1,T}X_{2,T} + \frac{X_{2,T}^2}{4})]\Big| = \Big|\mathbb{E}[h^{(2)}_{\delta/\sqrt{2}}(G_{0,\delta})(X_{1,T}X_{2,T} + \frac{X_{2,T}^2}{4})]\Big| = \Big|\mathbb{E}[h^{(1)}_{\delta/\sqrt{2}}(G_{0,\delta})Y_{1,0}]\Big|$$
$$\leq_c L_h\|Y\|_{1,2}\Big(\int_0^T v_t\sigma_t^2\mathrm{d}t\Big)^{-1/2} \leq_c L_h|\sigma|_\infty[\xi_{\sup}^3 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2]T^2.$$

The last term of (2.13) is handled similarly; apply the Lemma 3.3.2 with $Y = (X_{1,T} + \frac{X_{2,T}}{2})^3$ and use (3.11)-(3.23)-(3.25)-(3.26) to obtain the announced result. To complete the proof, it remains to establish the Lemma 3.3.2. This is done in the below subsection.

3.4 Proof of Lemma 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 $(\mathcal{H}_{x_0})$ 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 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 (3.19)-(3.21)-(3.23)-(3.25)-(3.26) it readily comes for any $i \in \{1,2\}$ and any $p \geq 1$:

$$\|(D^1G_\delta^\eta, D^2G_\delta^\eta, D^3G_\delta^\eta)\|_{i,p} \leq_c |\sigma|_\infty\sqrt{T}. \tag{3.29}$$

Hence, applying [Nua06, Proposition 1.5.6 and Proposition 2.1.4] and using (3.29) we get the existence of $Y_{2,\eta}$ and $Y_{3,\eta}$ such that for any $i \in \{2,3\}$ and any $p \geq 1$:

$$\begin{aligned}\|Y_{i,\eta}\|_p \leq_c &\|Y\|_{i-1,p+\frac{1}{2}}\|DG_\delta^\eta\|^{i-1}_{i-1,2^{i-1}p(2p+1)}\|\gamma^{-1}_{G_\delta^\eta}\|^{i-1}_{i-1,2^{i-1}p(2p+1)} \\ \leq_c &\|Y\|_{i-1,p+\frac{1}{2}}(|\sigma|_\infty\sqrt{T})^{i-1}\|\gamma^{-1}_{G_\delta^\eta}\|^{i-1}_{i-1,2^{i-1}p(2p+1)}.\end{aligned} \tag{3.30}$$

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:

$$\begin{aligned}\gamma_{G^\eta} =&\gamma_{X_T^P} + \gamma_{\eta(X_{1,T}+\frac{X_{2,T}}{2})} + 2\eta\sum_{i=1}^2\int_0^T D_t^i(X_T^P)D_t^i(X_{1,T} + \frac{X_{2,T}}{2})\mathrm{d}t \\ =&\int_0^T \sigma_t^2 v_t\mathrm{d}t + \eta^2\sum_{i=1}^2\int_0^T [D_t^i(X_{1,T} + \frac{X_{2,T}}{2})]^2\mathrm{d}t + 2\eta\sum_{i=1}^2\int_0^T \sigma_t v_t D_t^i(X_{1,T} + \frac{X_{2,T}}{2})\mathrm{d}t\end{aligned}$$

and hence estimates (3.21)-(3.23), $(\mathcal{H}_{x_0})$ and (P) easily yield to:

$$\sup_{\eta\in[0,1]} \Big\|\gamma_{G^\eta} - \int_0^T \sigma_t^2 v_t \mathrm{d}t\Big\|_p \leq_c |\sigma|_\infty^2(\xi_{\sup} + \mathcal{M}_1(\sigma))T^{\frac{3}{2}}, \tag{3.31}$$

for any $p \geq 1$. This intermediate estimate allows to prove the below Lemma:

Lemma 3.4.1. *Assume $(\mathcal{H}_{x_0})$ and (P). Then $(\gamma_{G_\delta^\eta})^{-1} \in \mathbb{D}^{2,\infty}$ and we have for any $p \geq 1$:*

$$\sup_{\eta\in[0,1]} \|(\gamma_{G_\delta^\eta})^{-1}\|_p \leq_c (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-1}, \tag{3.32}$$

$$\sup_{t\in[0,T],\ \eta\in[0,1]} \sum_{i\in\{1,2\}} \|D_t^i(\gamma_{G_\delta^\eta})^{-1}\|_p \leq_c (\mathcal{M}_1(\sigma) + \xi_{\sup})(\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-1}, \tag{3.33}$$

$$\sup_{s,t\in[0,T],\ \eta\in[0,1]} \sum_{i,j\in\{1,2\}} \|D_{s,t}^{i,j}(\gamma_{G_\delta^\eta})^{-1}\|_p \leq_c [\xi_{\sup}^2 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})](\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-1}. \tag{3.34}$$

Proof. For the sake of brevity, we only prove (3.32) and (3.33) because there is no extra difficulties for (3.34). For (3.32), we have for any $p \geq 1$ and $q \geq 1$:

$$\begin{aligned}
\mathbb{E}[(\gamma_{G_\delta^\eta})^{-p}] &= \mathbb{E}[(\gamma_{G_\delta^\eta})^{-p}\mathbb{1}_{\gamma_{G^\eta} \leq \frac{1}{2}\int_0^T \sigma_t^2 v_t \mathrm{d}t}] + \mathbb{E}[(\gamma_{G_\delta^\eta})^{-p}\mathbb{1}_{\gamma_{G^\eta} > \frac{1}{2}\int_0^T \sigma_t^2 v_t \mathrm{d}t}] \\
&\leq (\delta^2 \frac{T}{2})^{-p}\mathbb{P}(\int_0^T \sigma_t^2 v_t \mathrm{d}t - \gamma_{G^\eta} \geq \frac{\int_0^T \sigma_t^2 v_t \mathrm{d}t}{2}) + (\frac{1}{2}\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-p} \\
&\leq_c (\delta^2 T)^{-p}(\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-q}\Big\|\gamma_{G^\eta} - \int_0^T \sigma_t^2 v_t \mathrm{d}t\Big\|_q^q + (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-p} \\
&\leq_c (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-p}[(\delta^2 T)^{-p}(\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-q+p}\Big\|\gamma_{G^\eta} - \int_0^T \sigma_t^2 v_t \mathrm{d}t\Big\|_q^q] + 1],
\end{aligned}$$

where we have used the Markov inequality at the second inequality. Then choosing $q = 6p$ and using (3.19)-(3.31), we readily obtain:

$$\begin{aligned}
\|(\gamma_{G_\delta^\eta})^{-1}\|_p &\leq_c (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-1} \\
&\quad \times [(|\sigma|_\infty[\xi_{\sup}^3 + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2]T^2)^{-2}(T|\sigma|_\infty^2)^{-5}(|\sigma|_\infty^2[\xi_{\sup} + \mathcal{M}_1(\sigma)]T^{\frac{3}{2}})^6 + 1] \\
&\leq_c (\int_0^T \sigma_t^2 v_t \mathrm{d}t)^{-1}.
\end{aligned}$$

(3.33) is a straightforward application of [Nua06, 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 (3.21)-(3.23)-(3.25)-(3.32) 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 (3.32)-(3.33)-(3.34) in (3.30) to complete the proof.

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 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 [BG12] that, in addition to their simplicity, direct implied volatility expansions are more accurate than the corresponding price formulas.

Notations.

▷ **Call options.** We denote by $\text{Call}(S_0, T, K)$ 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 $\text{Call}(S_0, T, K) = \mathbb{E}(e^{X_T} - K)_+$. 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:

$$\text{Call}^{\text{BS}}(x, y, k) = e^x \mathcal{N}(d_1(x, y, k)) - e^k \mathcal{N}(d_2(x, y, k)) \tag{4.1}$$

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 $\text{Call}^{\text{BS}}(x_0, \int_0^T \sigma_t^2 v_t \mathrm{d}t, k) = \mathbb{E}[(e^{X_T^P} - e^k)_+]$ equals $\text{Call}(S_0, T, K) = \mathbb{E}[(e^{X_T} - K)_+]$ when $\mathcal{M}_1(\sigma) = \xi_{\text{sup}} = 0$. For (x, T, k) given, the *implied Black-Scholes volatility* of a price $\text{Call}(e^x, T, e^k)$ is the unique non-negative volatility parameter $\sigma_{\text{I}}(x, T, k)$ such that:

$$\text{Call}^{\text{BS}}(x, \sigma_{\text{I}}^2(x, T, k)T, k) = \text{Call}(e^x, T, e^k). \tag{4.2}$$

The reader can find in Proposition A.3.2 the definition of Vega^{BS} and Vomma^{BS} which are the first two derivatives of Call^{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:

$$\overline{\sigma}_z = \sqrt{\frac{1}{T}\int_0^T \sigma_t^2(z) v_t \mathrm{d}t}. \tag{4.3}$$

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.

4.1 Implied volatility expansions at spot

We introduce new corrective coefficients useful for the implied volatility expansions:

Definition 4.1.1. *Assume* $(\mathcal{H}_{x_0})$. *We define the following corrective coefficients:*

$$\begin{aligned}
\gamma_{0a,T} =& \bar{\sigma} + \frac{C^s_{1,T}}{4\bar{\sigma}T},\\
\gamma_{0b,T} =& \frac{C^l_{9,T}}{2\bar{\sigma}^3T^2} - \frac{C^l_{10,T}}{4\bar{\sigma}^3T^2} - \frac{3C^l_{10,T}}{\bar{\sigma}^5T^3} + \frac{C^s_{2,T}}{8\bar{\sigma}T} - \frac{C^s_{2,T}}{2\bar{\sigma}^3T^2} - \frac{C^s_{3,T}}{16\bar{\sigma}T} - \frac{C^s_{3,T}}{4\bar{\sigma}^3T^2} + \frac{3(C^s_{1,T})^2}{8\bar{\sigma}^5T^3}\\
&+ \frac{C^{ls}_{8,T}}{\bar{\sigma}^3T^2} - \frac{C^{ls}_{4,T}}{8\bar{\sigma}T} - \frac{3C^{ls}_{4,T}}{2\bar{\sigma}^3T^2} - \frac{C^{ls}_{5,T}}{8\bar{\sigma}T} - \frac{C^{ls}_{5,T}}{2\bar{\sigma}^3T^2} - \frac{C^{ls}_{6,T}}{2\bar{\sigma}^3T^2} + \frac{(C^l_{1,T}C^s_{1,T})}{8\bar{\sigma}^3T^2} + \frac{3(C^l_{1,T}C^s_{1,T})}{2\bar{\sigma}^5T^3},
\end{aligned}$$

$$\begin{aligned}
\gamma_{1a,T} =& -\frac{C^l_{1,T}}{\bar{\sigma}^3T^2} - \frac{C^s_{1,T}}{2\bar{\sigma}^3T^2},\\
\gamma_{1b,T} =& -\frac{C^s_{2,T}}{2\bar{\sigma}^3T^2} + \frac{3(C^s_{1,T})^2}{8\bar{\sigma}^5T^3} - \frac{C^{ls}_{2,T} + C^{ls}_{3,T}}{2\bar{\sigma}^3T^2} - \frac{C^{ls}_{4,T}}{2\bar{\sigma}^3T^2} - \frac{C^{ls}_{6,T}}{4\bar{\sigma}^3T^2} + \frac{3(C^l_{1,T}C^s_{1,T})}{4\bar{\sigma}^5T^3},\\
\gamma_{2,T} =& \frac{C^l_{3,T}}{\bar{\sigma}^5T^3} - \frac{3C^l_{4,T}}{\bar{\sigma}^5T^3} + 6\frac{C^l_{10,T}}{\bar{\sigma}^7T^4} + \frac{C^s_{2,T}}{2\bar{\sigma}^5T^3} + \frac{C^s_{3,T}}{4\bar{\sigma}^5T^3} - \frac{3(C^s_{1,T})^2}{4\bar{\sigma}^7T^4} + \frac{C^{ls}_{2,T} + C^{ls}_{3,T}}{\bar{\sigma}^5T^3}\\
&+ \frac{3C^{ls}_{4,T}}{2\bar{\sigma}^5T^3} + \frac{C^{ls}_{5,T}}{2\bar{\sigma}^5T^3} + \frac{C^{ls}_{6,T}}{2\bar{\sigma}^5T^3} - \frac{3(C^l_{1,T}C^s_{1,T})}{\bar{\sigma}^7T^4},
\end{aligned}$$

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 2.2.1 and* $C^l_{9,T}$-$C^l_{10,T}$-$C^{ls}_{8,T}$ *are defined by:*

$$\begin{aligned}
C^l_{9,T} =& \omega(\sigma^2 v, ((\sigma^{(1)})^2 + \sigma\sigma^{(2)})v, \sigma^2 v)^T_0, \qquad & C^l_{10,T} =& \omega(\sigma^2 v, \sigma\sigma^{(1)}v, \sigma\sigma^{(1)}v, \sigma^2 v)^T_0,\\
C^{ls}_{8,T} =& \omega(\rho\xi\sigma v, \sigma\sigma^{(1)}, \sigma^2 v)^T_0.
\end{aligned}$$

To obtain implied volatility expansions, we use the relations between the Greeks w.r.t. the log-spot which naturally appear when applying the expansion of Theorem 2.2.1 and the sensitivities w.r.t. the volatility parameter. These relations are available on Appendix A.3. Applying Proposition A.3.3, the third order approximation formula (2.2) can be transformed into:

$$\begin{aligned}
&\mathrm{Call}(S_0, T, K) \qquad (4.4)\\
=&\mathrm{Call}^{\mathrm{BS}}(x_0, \bar{\sigma}^2T, k) + \mathrm{Vega}^{\mathrm{BS}}(x_0, \bar{\sigma}^2T, k)\Big[- \frac{C^l_{1,T}}{\bar{\sigma}^3T^2}m + \frac{C^l_{2,T}}{2\bar{\sigma}T} - \frac{C^l_{3,T}}{\bar{\sigma}^3T^2} + \frac{C^l_{3,T}}{\bar{\sigma}^5T^3}m^2 - \frac{C^l_{4,T}}{4\bar{\sigma}T} - \frac{3C^l_{4,T}}{\bar{\sigma}^3T^2}\\
&+ \frac{3C^l_{4,T}}{\bar{\sigma}^5T^3}m^2 + \frac{(C^l_{1,T})^2}{8\bar{\sigma}^3T^2} + \frac{3(C^l_{1,T})^2}{2\bar{\sigma}^5T^3} - \frac{3(C^l_{1,T})^2}{\bar{\sigma}^7T^4}m^2 + \frac{C^s_{1,T}}{4\bar{\sigma}T} - \frac{C^s_{1,T}}{2\bar{\sigma}^3T^2}m + \frac{C^s_{2,T}}{8\bar{\sigma}T} - \frac{C^s_{2,T}}{2\bar{\sigma}^3T^2} - \frac{C^s_{2,T}}{2\bar{\sigma}^3T^2}m\\
&+ \frac{C^s_{2,T}}{2\bar{\sigma}^5T^3}m^2 - \frac{C^s_{3,T}}{16\bar{\sigma}T} - \frac{C^s_{3,T}}{4\bar{\sigma}^3T^2} + \frac{C^s_{3,T}}{4\bar{\sigma}^5T^3}m^2 + \frac{3(C^s_{1,T})^2}{8\bar{\sigma}^5T^3} + \frac{3(C^s_{1,T})^2}{8\bar{\sigma}^5T^3}m - \frac{3(C^s_{1,T})^2}{4\bar{\sigma}^7T^4}m^2 + \frac{C^{ls}_{1,T}}{\bar{\sigma}T}\\
&- \frac{(C^{ls}_{2,T} + C^{ls}_{3,T})}{\bar{\sigma}^3T^2} - \frac{(C^{ls}_{2,T} + C^{ls}_{3,T})}{2\bar{\sigma}^3T^2}m + \frac{(C^{ls}_{2,T} + C^{ls}_{3,T})}{\bar{\sigma}^5T^3}m^2 - \frac{C^{ls}_{4,T}}{8\bar{\sigma}T} - \frac{3C^{ls}_{4,T}}{2\bar{\sigma}^3T^2} - \frac{C^{ls}_{4,T}}{2\bar{\sigma}^3T^2}m + \frac{3C^{ls}_{4,T}}{2\bar{\sigma}^5T^3}m^2\\
&- \frac{C^{ls}_{5,T}}{8\bar{\sigma}T} - \frac{C^{ls}_{5,T}}{2\bar{\sigma}^3T^2} + \frac{C^{ls}_{5,T}}{2\bar{\sigma}^5T^3}m^2 - \frac{C^{ls}_{6,T}}{2\bar{\sigma}^3T^2} - \frac{C^{ls}_{6,T}}{4\bar{\sigma}^3T^2}m + \frac{C^{ls}_{6,T}}{2\bar{\sigma}^5T^3}m^2 + \frac{(C^l_{1,T}C^s_{1,T})}{8\bar{\sigma}^3T^2} + \frac{3(C^l_{1,T}C^s_{1,T})}{2\bar{\sigma}^5T^3}
\end{aligned}$$

$$+\frac{3(C^l_{1,T}C^s_{1,T})}{4\bar\sigma^5T^3}m-\frac{3(C^l_{1,T}C^s_{1,T})}{\bar\sigma^7T^4}m^2\Big]$$
$$+\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x_0,\bar\sigma^2T,k)\Big[(\frac{C^l_{1,T}}{\bar\sigma^3T^2})^2m^2+(C^s_{1,T})^2(-\frac{m}{2\bar\sigma^3T^2}+\frac{1}{4\bar\sigma T})^2+(C^l_{1,T}C^s_{1,T})(\frac{m^2}{\bar\sigma^6T^4}-\frac{m}{2\bar\sigma^4T^3})\Big]$$
$$+O(K\,|\sigma|_\infty[\xi^3_{\sup}+\mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma)+\xi_{\sup})^2]\,T^2).$$

Using the below relations easy to establish:

$$C^{ls}_{1,T}\omega(\sigma^2v)^T_0=C^{ls}_{2,T}+C^{ls}_{3,T}+C^{ls}_{8,T},\quad C^l_{2,T}\omega(\sigma^2v)^T_0=C^l_{9,T}+2C^l_{3,T},\quad C^l_{4,T}\omega(\sigma^2v)^T_0=C^l_{10,T}+\frac{(C^l_{1,T})^2}{2},$$

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 (4.4):

$$\begin{aligned}&\mathrm{Call}(S_0,T,K)\\ =&\mathrm{Call}^{\mathrm{BS}}(x_0,\bar\sigma^2T,k)+\mathrm{Vega}^{\mathrm{BS}}(x_0,\bar\sigma^2T,k)[(\gamma_{0a,T}+\gamma_{0b,T})-\bar\sigma+(\gamma_{1a,T}+\gamma_{1b,T})m+\gamma_{2,T}m^2]\\ &+\frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x_0,\bar\sigma^2T,k)[\gamma_{0a,T}-\bar\sigma+\gamma_{1a,T}m]^2+O(K\,|\sigma|_\infty[\xi^3_{\sup}+\mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma)+\xi_{\sup})^2]\,T^2)\\ \approx&\mathrm{Call}^{\mathrm{BS}}(x_0,[(\gamma_{0a,T}+\gamma_{0b,T})+(\gamma_{1a,T}+\gamma_{1b,T})m+\gamma_{2,T}m^2]^2T,k),\end{aligned}$$

This reads as an expansion of the implied volatility and proves the below Theorem:

Theorem 4.1.1. ***(3rd order expansion of the implied volatility)****. Assume* $(\mathcal{H}_{x_0})$ *and* (P)*. We have:*

$$\sigma_{\mathrm{I}}(x_0,T,k)=\gamma_{0a,T}+\gamma_{0b,T}+(\gamma_{1a,T}+\gamma_{1b,T})m+\gamma_{2,T}m^2+\mathrm{Error}^{\mathrm{I}}_{3,x_0}. \tag{4.5}$$

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 4.1.1.*

Corollaries of Theorem 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 $|m|\le C_m|\sigma|_\infty\sqrt{T}$ for a given $C_m>0$ 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:

$$|\mathrm{Error}^{\mathrm{I}}_{3,x_0}|=O(|\sigma|_\infty[\xi^3_{\sup}+\mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma)+\xi_{\sup})^2]T^{\frac{3}{2}}),$$

where the generic constant depends in an increasing way on C_m, what justifies the label of third order expansion. Although tedious to write, the proof does not contain huge mathematical difficulties and is performed in [BG12, Subsection 3.4] 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 $\sigma \to \text{Call}^{\text{BS}}(x_0, \sigma^2 T, k)$ at γ around $\bar{\sigma}$, a zero order expansion at $\sigma_1(x_0, T, k)$ around γ, apply Theorem 2.2.1, use classical estimates for $\partial^n_{\sigma^n}\text{Call}^{\text{BS}}(x_0, \sigma^2 T, k)$ for $n = 1, 2, 3$ (see [BG12, Corollary 33]) and carefully gather terms and evaluate their magnitude.

▷**Short maturity skew and smile behaviours.** We analyse the behaviour of the approximation formula (4.5) at the money (i.e. $m \approx 0$) and for short maturity (i.e. $T \ll 1$). In view of (4.5) and the various coefficients C^l, C^s, C^{ls} and γ (see Definition 4.1.1 and Theorem 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:

$$
\begin{aligned}
[\sigma_1(x_0, T, k)]|_{k=x_0} \approx & \gamma_{0a,T} + \gamma_{0b,T} \approx \sigma_0 \sqrt{v_0}, \\
\partial_k[\sigma_1(x_0, T, k)]|_{k=x_0} \approx & -\gamma_{1a,T} - \gamma_{1b,T} \approx -\gamma_{1a,T} \approx \frac{\sigma_0^{(1)}\sqrt{v_0}}{2} + \frac{\rho_0 \xi_0}{4\sqrt{v_0}}, \qquad (4.6) \\
\partial^2_{k^2}[\sigma_1(x_0, T, k)]|_{k=x_0} \approx & 2\gamma_{2,T} = \frac{\sigma_0^{(2)}\sqrt{v_0}}{3} - \frac{[\sigma_0^{(1)}]^2\sqrt{v_0}}{6\sigma_0} - \frac{5\rho_0^2\xi_0^2}{24\sigma_0 v_0^{\frac{3}{2}}} + \frac{\xi_0^2}{12\sigma_0 v_0^{\frac{3}{2}}}, \qquad (4.7)
\end{aligned}
$$

where we have used the below 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 [FJ12, Theorem 4.1]. Otherwise, we notice that the mixt terms 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. $\xi_{\text{sup}} = 0$), we retrieve the results of [BG12, Theorem 22].
3) For pure Heston models (i.e. $\mathcal{M}_1(\sigma) = 0$), we recover the expansion given in [FJ09, Theorem 2.5]. In the case of zero correlation, the approximation formula (4.5) becomes for short maturity:

$$
\sigma_1(x_0, T, k) \approx \bar{\sigma} - \frac{C^s_{3,T}}{16\bar{\sigma}T} - \frac{C^s_{3,T}}{4\bar{\sigma}^3 T^2} + \frac{C^s_{3,T}}{4\bar{\sigma}^5 T^3} m^2 \approx \bar{\sigma} - \frac{\xi_0^2 \sigma_0 T}{24\sqrt{v_0}}\left[\frac{\sigma_0^2 v_0 T}{4} + 1\right] + \frac{\xi_0^2}{24\sigma_0 v_0^{\frac{3}{2}}} m^2.
$$

We have retrieved that an uncorrelated Heston model induces 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 (4.6) (the slope becomes negative and increases in absolute value with $|\rho_0|$) and (4.7) (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 (5.1)). 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 relied 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.

4.2 Implied volatility expansions at mid-point

It has been empirically proven in [BG12] 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 2.2.1 and Definition 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 $(\mathcal{H}_{x_{avg}})$ and $(\mathcal{H}_{x_0})$ 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 $\widetilde{C}^l_{1,T}(x_{avg}) = \omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v, \sigma^2(x_{avg})v)^T_0$ instead of $C^l_{1,T}(x_{avg}) = \omega(\sigma^2(x_{avg})v, \sigma(x_{avg})\sigma^{(1)}(x_{avg})v)^T_0$.

Definition 4.2.1. *Assume $(\mathcal{H}_{x_{avg}})$. We define the following corrective coefficients:*

$$
\begin{aligned}
\pi_{0a,T}(x_{avg}) =& \gamma_{0a,T}(x_{avg}),\\
\pi_{0b,T}(x_{avg}) =& \gamma_{0b,T}(x_{avg}),\\
\pi_{1a,T}(x_{avg}) =& \frac{\widetilde{C}^l_{1,T}(x_{avg}) - C^l_{1,T}(x_{avg})}{2\bar{\sigma}^3_{x_{avg}}T^2} - \frac{C^s_{1,T}(x_{avg})}{2\bar{\sigma}^3_{x_{avg}}T^2},\\
\pi_{1b,T}(x_{avg}) =& \gamma_{1b,T}(x_{avg}) + \frac{C^{ls}_{1,T}(x_{avg})}{4\bar{\sigma}_{x_{avg}}T} + \frac{C^{ls}_{9,T}(x_{avg})}{8\bar{\sigma}_{x_{avg}}T} - \frac{C^{ls}_{4,T}(x_{avg}) + C^{ls}_{5,T}(x_{avg}) + C^{ls}_{10,T}(x_{avg})}{8\bar{\sigma}^3_{x_{avg}}T^2},
\end{aligned}
$$

$$
\begin{aligned}
\pi_{2,T}(x_{avg}) =& \frac{\widetilde{C}^l_{3,T}(x_{avg}) + C^l_{3,T}(x_{avg})}{2\bar{\sigma}^5_{x_{avg}}T^3} - \frac{3(\widetilde{C}^l_{4,T}(x_{avg}) + C^l_{4,T}(x_{avg}))}{2\bar{\sigma}^5_{x_{avg}}T^3} - \frac{C^l_{5,T}(x_{avg})}{8\bar{\sigma}_{x_{avg}}T} + \frac{C^l_{6,T}(x_{avg})}{4\bar{\sigma}^3_{x_{avg}}T^2}\\
&+ 6\frac{C^l_{10,T}(x_{avg})}{\bar{\sigma}^7_{x_{avg}}T^4} + \frac{C^s_{2,T}(x_{avg})}{2\bar{\sigma}^5_{x_{avg}}T^3} + \frac{C^s_{3,T}(x_{avg})}{4\bar{\sigma}^5_{x_{avg}}T^3} - \frac{3(C^s_{1,T}(x_{avg}))^2}{4\bar{\sigma}^7_{x_{avg}}T^4} + \frac{C^{ls}_{2,T}(x_{avg}) + C^{ls}_{3,T}(x_{avg})}{\bar{\sigma}^5_{x_{avg}}T^3}
\end{aligned}
$$

$$+\frac{3C^{ls}_{4,T}(x_{avg})}{2\bar\sigma^5_{x_{avg}}T^3}+\frac{C^{ls}_{5,T}(x_{avg})}{2\bar\sigma^5_{x_{avg}}T^3}+\frac{C^{ls}_{6,T}(x_{avg})}{2\bar\sigma^5_{x_{avg}}T^3}-\frac{3(C^{l}_{1,T}C^{s}_{1,T})(x_{avg})}{\bar\sigma^7_{x_{avg}}T^4}-\frac{C^{ls}_{1,T}(x_{avg})}{2\bar\sigma^3_{x_{avg}}T^2}-\frac{C^{ls}_{9,T}(x_{avg})}{4\bar\sigma^3_{x_{avg}}T^2}$$
$$+3\frac{C^{ls}_{4,T}(x_{avg})+C^{ls}_{5,T}(x_{avg})+C^{ls}_{10,T}(x_{avg})}{4\bar\sigma^5_{x_{avg}}T^3},$$

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:*

$$C^{l}_{5,T}(x_{avg})=\omega(((\sigma^{(1)})^2+\sigma\sigma^{(2)})(x_{avg})v)^T_0,\quad C^{l}_{6,T}(x_{avg})=\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v,\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)^T_0,$$
$$C^{ls}_{9,T}(x_{avg})=\omega(\rho\xi\sigma^{(1)}(x_{avg})v,\sigma^2(x_{avg}))^T_0,\quad C^{ls}_{10,T}(x_{avg})=\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v,\rho\xi\sigma(x_{avg})v,\sigma^2(x_{avg}))^T_0.$$

To obtain a new implied volatility approximation, we consider the formula 4.5 and we perform a Taylor expansion around the mid-point. First we analyse the leading term $\bar\sigma_{x_0}$ of 4.5:

Lemma 4.2.1. *Assume* $(\mathcal{H}_{x_0})$-$(\mathcal{H}_{x_{avg}})$ *and suppose that* $|m|\le C_m|\sigma|_\infty\sqrt{T}$ *for a given* $C_m>0$. *Then we have:*

$$\bar\sigma_{x_0}=\bar\sigma_{x_{avg}}+\frac{\omega((\sigma\sigma^{(1)})(x_{avg})v)^T_0}{2\bar\sigma_{x_{avg}}T}m+\frac{C^{l}_{5,T}(x_{avg})}{8\bar\sigma_{x_{avg}}T}m^2-\frac{C^{l}_{6,T}(x_{avg})}{4\bar\sigma^3_{x_{avg}}T^2}m^2+O(|\sigma|_\infty\mathcal{M}_0(\sigma)^2\mathcal{M}_1(\sigma)T^{\frac{3}{2}}),\quad(4.8)$$

where $C^{l}_{5,T}$ *and* $C^{l}_{6,T}$ *are defined in Definition 4.2.1.*

Proof. First notice that $(\mathcal{H}_{x_0})$-$(\mathcal{H}_{x_{avg}})$ 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 $y\to\sqrt{y}$ at $y=\frac{\omega(\sigma^2(x_0)v)^T_0}{T}$ around $y=\frac{\omega(\sigma^2(x_{avg})v)^T_0}{T}$ and secondly for the function $x\to\sigma^2_t(x)$ at $x=x_0$ around $x=x_{avg}$, for any $t\in[0,T]$. It gives:

$$\begin{aligned}\bar\sigma_{x_0}&=\sqrt{\frac{\omega(\sigma^2(x_0)v)^T_0}{T}}=\bar\sigma_{x_{avg}}+\frac{\omega(\sigma^2(x_0)v)^T_0-\omega(\sigma^2(x_{avg})v)^T_0}{2\bar\sigma_{x_{avg}}T}-\frac{[\omega(\sigma^2(x_0)v)^T_0-\omega(\sigma^2(x_{avg})v)^T_0]^2}{8\bar\sigma^3_{x_{avg}}T^2}+R_1\\&=\bar\sigma_{x_{avg}}+\frac{\omega((\sigma\sigma^{(1)})(x_{avg})v)^T_0}{2\bar\sigma_{x_{avg}}T}m+\frac{\omega(((\sigma^{(1)})^2+\sigma\sigma^{(2)})(x_{avg})v)^T_0}{8\bar\sigma_{x_{avg}}T}m^2-\frac{[\omega((\sigma\sigma^{(1)})(x_{avg})v)^T_0]^2}{8\bar\sigma^3_{x_{avg}}T^2}m^2\\&\quad+R_1+R_2+R_3,\end{aligned}$$

where:

$$\begin{aligned}R_1&=(\omega(\sigma^2(x_0)v)^T_0-\omega(\sigma^2(x_{avg})v)^T_0)^3\int_0^1\frac{3}{8T^3(u\overline{\sigma}^2_{x_0}+(1-u)\overline{\sigma}^2_{x_{avg}})^{\frac{5}{2}}}\frac{(1-u)^2}{2}\mathrm{d}u,\\R_2&=\frac{m^3}{16\bar\sigma_{x_{avg}}T}\int_0^1\partial^3_{x^3}(\omega(\sigma^2(x)v)^T_0)|_{x=ux_0+(1-u)x_{avg}}\frac{(1-u)^2}{2}\mathrm{d}u,\\R_3&=-\frac{(\omega(\sigma^2(x_0)v)^T_0-\omega(\sigma^2(x_{avg})v)^T_0+m\omega((\sigma\sigma^{(1)})(x_{avg})v)^T_0)}{8\bar\sigma^3_{x_{avg}}T^2}\frac{m^2}{4}\\&\quad\times\int_0^1\partial^2_{x^2}(\omega(\sigma^2(x)v)^T_0)|_{x=ux_0+(1-u)x_{avg}}(1-u)\mathrm{d}u.\end{aligned}$$

Next remark that $\omega(((\sigma^{(1)})^2+\sigma\sigma^{(2)})(x_{avg})v)^T_0=C^{l}_{5,T}(x_{avg})$ and that $[\omega((\sigma\sigma^{(1)})(x_{avg})v)^T_0]^2=2C^{l}_{6,T}(x_{avg})$. Then we readily obtain with the assumption on m and $(\mathcal{H}_{x_0})$-$(\mathcal{H}_{x_{avg}})$ 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 4.2.2. *Assume* $(\mathcal{H}_{x_0})$-$(\mathcal{H}_{x_{avg}})$ *and suppose that* $|m| \leq C_m|\sigma|_\infty \sqrt{T}$ *for a given* $C_m > 0$. *Then we have:*

$$\begin{aligned}
&\gamma_{0a,T}(x_0) - \bar{\sigma}_{x_0} + \gamma_{1a,T}(x_0)m \qquad (4.9)\\
=&\gamma_{0a,T}(x_{avg}) - \bar{\sigma}_{x_{avg}} + \frac{C^{ls}_{9,T}(x_{avg}) + 2C^{ls}_{1,T}(x_{avg})}{8\bar{\sigma}_{x_{avg}}T}m - \frac{\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T C^s_{1,T}(x_{avg})}{8\bar{\sigma}^3_{x_{avg}}T^2}m + \gamma_{1a,T}(x_{avg})m\\
&- \frac{C^l_{2,T}(x_{avg}) + 2C^l_{6,T}(x_{avg})}{2\bar{\sigma}^3_{x_{avg}}T^2}m^2 + 3\frac{\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T C^l_{1,T}(x_{avg})}{2\bar{\sigma}^5_{x_{avg}}T^3}m^2 - \frac{C^{ls}_{9,T}(x_{avg}) + 2C^{ls}_{1,T}(x_{avg})}{4\bar{\sigma}^3_{x_{avg}}T^2}m^2\\
&+ 3\frac{\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T C^s_{1,T}(x_{avg})}{4\bar{\sigma}^5_{x_{avg}}T^3}m^2 + O(|\sigma|_\infty[\xi^3_{\sup} + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2]T^{\frac{3}{2}}),\\
&\gamma_{0b,T}(x_0) + \gamma_{1b,T}(x_0)m + \gamma_{2,T}(x_0)m^2 \qquad (4.10)\\
=&\gamma_{0b,T}(x_{avg}) + \gamma_{1b,T}(x_{avg})m + \gamma_{2,T}(x_{avg})m^2 + O(|\sigma|_\infty[\xi^3_{\sup} + \mathcal{M}_1(\sigma)(\mathcal{M}_0(\sigma) + \xi_{\sup})^2]T^{\frac{3}{2}}).
\end{aligned}$$

Proof. The above expansions can be proved similarly than the expansion of Lemma 4.2.1 with long and tedious computations. Since there is no extra difficulty, we skip further details. □

Lemmas 4.2.1 and 4.2.2 lead to the below Theorem:

Theorem 4.2.1. ***(3rd order expansion of the implied volatility at mid-point).*** *Assume* $(\mathcal{H}_{x_0})$ *and* (P). *We have:*

$$\sigma_{\mathrm{I}}(x_0, T, k) = \pi_{0a,T}(x_{avg}) + \pi_{0b,T}(x_{avg}) + (\pi_{1a,T}(x_{avg}) + \pi_{1b,T}(x_{avg}))m + \pi_{2,T}(x_{avg})m^2 + \mathrm{Error}^{\mathrm{I}}_{3,x_{avg}}, \qquad (4.11)$$

where the corrective coefficients π *are defined in Definition 4.2.1.*

Proof. We gather terms coming from Lemmas 4.2.1 and 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:

$$\frac{\omega((\sigma\sigma^{(1)})(x_{avg})v)_0^T}{2\bar{\sigma}_{x_{avg}}T}m - \frac{C^l_{1,T}(x_{avg})}{\bar{\sigma}^3_{x_{avg}}T^2}m = \frac{\widetilde{C}^l_{1,T}(x_{avg}) - C^l_{1,T}(x_{avg})}{2\bar{\sigma}^3_{x_{avg}}T^2}m. \qquad (4.12)$$

Second remark that:

$$\begin{aligned}
\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T C^s_{1,T}(x_{avg}) =& C^{ls}_{4,T}(x_{avg}) + C^{ls}_{5,T}(x_{avg}) + C^{ls}_{10,T}(x_{avg}), \qquad (4.13)\\
\omega(\sigma(x_{avg})\sigma^{(1)}(x_{avg})v)_0^T C^l_{1,T}(x_{avg}) =& 2C^l_{4,T}(x_{avg}) + \omega((\sigma\sigma^{(1)})(x_{avg})v, \sigma^2(x_{avg})v, (\sigma\sigma^{(1)})(x_{avg})v)_0^T. \qquad (4.14)
\end{aligned}$$

Then use the below 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:

$$\frac{C^l_{3,T}(x_{avg})}{\bar\sigma^5_{x_{avg}}T^3}m^2-\frac{C^l_{2,T}(x_{avg})}{2\bar\sigma^3_{x_{avg}}T^2}m^2+\frac{C^l_{5,T}(x_{avg})}{8\bar\sigma_{x_{avg}}T}m^2=\frac{C^l_{3,T}(x_{avg})+\widetilde{C}^l_{3,T}(x_{avg})}{2\bar\sigma^5_{x_{avg}}T^3}m^2-\frac{C^l_{5,T}(x_{avg})}{8\bar\sigma_{x_{avg}}T}m^2. \quad (4.15)$$

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:

$$\begin{aligned}
&-3\frac{C^l_{4,T}(x_{avg})}{\bar\sigma^5_{x_{avg}}T^3}+\frac{3}{2}\frac{[2C^l_{4,T}(x_{avg})+\omega((\sigma\sigma^{(1)})(x_{avg})v,\sigma^2(x_{avg})v,(\sigma\sigma^{(1)})(x_{avg})v)_0^T]}{\bar\sigma^5_{x_{avg}}T^3}\\
&-\frac{C^l_{6,T}(x_{avg})}{\bar\sigma^3_{x_{avg}}T^2}-\frac{C^l_{6,T}(4x_{avg})}{\bar\sigma^3_{x_{avg}}T^2}\\
=&-3\frac{[C^l_{4,T}(x_{avg})+\widetilde{C}^l_{4,T}(x_{avg})]}{2\bar\sigma^5_{x_{avg}}T^3}+\frac{C^l_{6,T}(x_{avg})}{4\bar\sigma^3_{x_{avg}}T^2}.
\end{aligned} \quad (4.16)$$

Finally sum the relations (4.8)-(4.9)-(4.10) and take into account the mathematical reductions (4.12)-(4.13)-(4.14)-(4.15)-(4.16) to obtain the announced result. □

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 (4.11) in Theorem 4.2.1). We consider a time-independent CEV-Heston model:

$$\begin{aligned}
\mathrm{d}S_t&=\mu S_t^\beta\sqrt{Y_t}\mathrm{d}W_t,\ S_0=e^{x_0},\\
\mathrm{d}Y_t&=\kappa(\theta-Y_t)\mathrm{d}t+\sqrt{Y_t}\xi\mathrm{d}B_t,\ Y_0=v_0,\\
\mathrm{d}\langle W,B\rangle_t&=\rho\mathrm{d}t.
\end{aligned} \quad (5.1)$$

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 A.4.2 in Appendix A.4, the implied volatility formula (4.11) writes explicitly:

$$\begin{aligned}
\sigma_{\mathrm{I}}(x_0,T,k)\approx&\mu(S_0K)^{\frac{\beta-1}{2}}\sqrt{\overline{v}}\Big\{1+\frac{\rho\xi\mu(S_0K)^{\frac{\beta-1}{2}}R_1^sT}{8\overline{v}}+\frac{(\beta-1)^2\mu^2(S_0K)^{\beta-1}\overline{v}T}{24}(1-\frac{\mu^2(S_0K)^{\beta-1}\overline{v}T}{4}) \quad (5.2)\\
&+\frac{\rho^2\xi^2T}{\overline{v}^2}[\frac{3(R_1^s)^2}{32\overline{v}}+R_2^s(\frac{\mu^2(S_0K)^{\beta-1}\overline{v}T}{48}-\frac{1}{12})]-\frac{\xi^2TR_3^s}{\overline{v}^2}[\frac{1}{24}+\frac{\mu^2(S_0K)^{\beta-1}\overline{v}T}{96}]\\
&+\frac{\rho\xi(\beta-1)\mu(S_0K)^{\frac{\beta-1}{2}}T}{\overline{v}}[\frac{R_1^s}{8}+\mu^2(S_0K)^{\beta-1}\overline{v}T(\frac{R_2^{ls}}{48\overline{v}}-\frac{R_1^s}{32})]\\
&+\frac{\rho\xi}{\mu(S_0K)^{\frac{\beta-1}{2}}\overline{v}^2}[-\frac{R_1^s}{4}+(\beta-1)\mu^2(S_0K)^{\beta-1}\overline{v}T(\frac{R_1^s}{16}-\frac{R_2^{ls}}{24\overline{v}})]\log(\frac{S_0}{K})\\
&+\frac{\rho^2\xi^2T}{\overline{v}^2}[\frac{3(R_1^s)^2}{32\overline{v}}-\frac{R_2^s}{12}]\log(\frac{S_0}{K})-\frac{(\beta-1)^2}{24}\log^2(\frac{S_0}{K})
\end{aligned}$$

$$+\frac{\rho^2\xi^2}{\mu^2(S_0K)^{\beta-1}\overline{v}^3}\Big[\frac{R_2^s}{12}-\frac{3(R_1^s)^2}{16\overline{v}}\Big]\log^2\Big(\frac{S_0}{K}\Big)+\frac{\xi^2R_3^s}{24\mu^2(S_0K)^{\beta-1}\overline{v}^3}\log^2\Big(\frac{S_0}{K}\Big)\Big\},$$

where the coefficients $\overline{v}$, R_1^s, R_2^s, R_3^s and R_2^{ls} are defined in Proposition A.4.1 of Appendix A.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 [Gla04, 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 (5.1) with β fixed at 1. The latter expectation is computed using the Lewis formula [Lew00]. In [BGM10b], 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 [BG12] 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:

$$S_0=1,\quad \mu=0.25,\quad \beta=0.5,\quad v_0=1,\quad \theta=1.2,\quad \kappa=3,\quad \xi=1.5,\quad \rho=-70\%, \tag{5.3}$$

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.

▷**Accuracy of the implied volatility formula** (4.11)**.** In Tables 1, 2, 3, 4, 5 and 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 (5.2) (AF(x_{avg})). We use the parameters as in (5.3) and the strikes are chosen to be approximately equal to $S_0e^{q\mu\sqrt{\theta T}}$ 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 (5.2) 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 (2.3) which was already observed in [BGM10b] 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 1 realized for the maturity $3M$ 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 report ND in the tabulars corresponding to the maturities $5Y$ and $10Y$ meaning that the corresponding prices are outside the 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

[3] 1 bp (basis point) is equal to 0.01%.

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 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
AF(x_{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 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 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

A Appendix

A.1 Change of model

In this section, we justify why we work without loss of generality with the model (1.1)-(1.2). If we consider a general time-dependent CIR process, the formulation becomes:

$$\mathrm{d}X_t = \Sigma(t, X_t)\sqrt{Y_t}(dW_t - \frac{\Sigma(t, X_t)\sqrt{Y_t}}{2}\mathrm{d}t),\ X_0 = x_0,$$

Table 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: 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

$$\mathrm{d}Y_t = \kappa_t(\theta_t - Y_t)\mathrm{d}t + \gamma_t\sqrt{Y_t}\mathrm{d}B_t,\ Y_0 = v_0 > 0,$$

with a correlation $(\rho_t)_{t\in[0,T]}$ between W and B. We assume the below 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 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 (1.1)-(1.2). Observe that $(P') \Longleftrightarrow (P)$ and that the local volatility functions σ and Σ have the same space regularity.

A.2 Explicit computation of the corrective terms of Theorem 2.2.1

We give the full derivation of the corrective terms in the approximation (2.2) of Theorem 2.2.1. We begin with the proof of Lemma 2.3.1 and next we give the details of the computation of the corrective terms.

A.2.1 Proof of Lemma 2.3.1

We proceed by induction. One needs the following technical result:

Table 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(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

Lemma A.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 [Nua06]). □

If $N = 1$ and $I_N = 0$, there is nothing to prove. If $N = 1$ and $I_N \in \{1, 2\}$, Lemma A.2.1 is a particular case of Lemma 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 (2.11) is true for $N \geq 2$. Then apply Lemma A.2.1 if $I_{N+1} \in \{1, 2\}$ to obtain:

$$\begin{aligned}
&\mathbb{E}\Big(\varphi(\int_0^T f_t \mathrm{d}W_t)\int_0^T l_{N+1,t_{N+1}}\int_0^{t_{N+1}} l_{N,t_N}\ldots\int_0^{t_2} l_{1,t_1}\mathrm{d}W_{t_1}^{I_1}\ldots\mathrm{d}W_{t_N}^{I_N}\mathrm{d}W_{t_{N+1}}^{I_{N+1}}\Big)\\
=&\mathbb{E}\Big(\varphi^{(\mathbb{1}_{I_{N+1}\neq 0})}(\int_0^T f_t \mathrm{d}W_t)\int_0^T \widehat{l}_{N+1,t_{N+1}}\int_0^{t_{N+1}} l_{N,t_N}\int_0^{t_N}\ldots\int_0^{t_2} l_{1,t_1}\mathrm{d}W_{t_1}^{I_1}\ldots\mathrm{d}W_{t_N}^{I_N}\mathrm{d}t_{N+1}\Big)\\
=&\mathbb{E}\Big(\varphi^{(\mathbb{1}_{I_{N+1}\neq 0})}(\int_0^T f_t \mathrm{d}W_t)\int_0^T (l_{N,t_N}\int_{t_N}^T \widehat{l}_{N+1,s}\mathrm{d}s)\int_0^{t_N}\ldots\int_0^{t_2} l_{1,t_1}\mathrm{d}W_{t_1}^{I_1}\ldots\mathrm{d}W_{t_N}^{I_N}\Big),
\end{aligned}$$

where at the last equality we have used the fact that $\int_0^T g_t Z_t \mathrm{d}t = \int_0^T (\int_t^T g_s \mathrm{d}s) dZ_t$ for any continuous semi-martingale Z starting from 0 and any measurable and bounded deterministic function g (apply the Itô formula to the product $(\int_t^T g_s ds)Z_t$). We conclude without difficulty with the induction hypothesis and leave the details to the reader.

A.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).

▷**Step 1: contribution with $X_{1,T}$.** Apply the Lemma 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:

$$X_{1,T} = \int_0^T \left(\sigma_{t_2}^{(1)}\lambda_{t_2}\int_0^{t_2}\sigma_{t_1}\lambda_{t_1}(\mathrm{d}W_{t_1} - \frac{\sigma_{t_1}\lambda_{t_1}}{2}\mathrm{d}t_1) + \frac{\sigma_{t_2}}{2\lambda_{t_2}}\int_0^{t_2}\xi_{t_1}\lambda_{t_1}\mathrm{d}B_{t_1}\right)(\mathrm{d}W_{t_2} - \sigma_{t_2}\lambda_{t_2}\mathrm{d}t_2), \tag{A.1}$$

to get:

$$\mathbb{E}[h^{(1)}(X_T^P)X_{1,T}] = C_{1,T}^l[\mathcal{G}_3^h(X_T^P) - \frac{3}{2}\mathcal{G}_2^h(X_T^P) + \frac{1}{2}\mathcal{G}_1^h(X_T^P)] + \frac{C_{1,T}^s}{2}[\mathcal{G}_3^h(X_T^P) - \mathcal{G}_2^h(X_T^P)], \tag{A.2}$$

where:

$$C_{1,T}^s = \omega(\rho\xi\sigma v, \sigma^2)_0^T, \qquad C_{1,T}^l = \omega(\sigma^2 v, \sigma\sigma^{(1)}v)_0^T.$$

▷**Step 2: contribution with $X_{2,T}$.** In view of (2.7)-(2.5)-(2.6)-(2.8)-(2.9), 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 formula:

$$\frac{(X_t^P - x_0)^2}{2} = \int_0^t (\int_0^{t_2}\sigma_{t_1}\lambda_{t_1}(\mathrm{d}W_{t_1} - \frac{\sigma_{t_1}\lambda_{t_1}}{2}\mathrm{d}t_1))\sigma_{t_2}\lambda_{t_2}(\mathrm{d}W_{t_2} - \frac{\sigma_{t_2}\lambda_{t_2}}{2}\mathrm{d}t_2) + \frac{1}{2}\int_0^t \sigma_{t_1}^2 v_{t_1}\mathrm{d}t_1, \tag{A.3}$$

$$\begin{aligned}(X_t^P - x_0)V_{1,t} =& \int_0^t (\int_0^{t_2}\xi_{s_1}\lambda_{s_1}\mathrm{d}B_{s_1})\sigma_{t_2}\lambda_{t_2}(\mathrm{d}W_{t_2} - \frac{\sigma_{t_2}\lambda_{t_2}}{2}\mathrm{d}t_2)\\ &+ \int_0^t (\int_0^{t_2}\sigma_{t_1}\lambda_{t_1}(\mathrm{d}W_{t_1} - \frac{\sigma_{t_1}\lambda_{t_1}}{2}\mathrm{d}t_1))\xi_{s_2}\lambda_{s_2}\mathrm{d}B_{s_2} + \int_0^t \xi_{t_1}\rho_{t_1}\sigma_{t_1}v_{t_1}\mathrm{d}t_1,\end{aligned} \tag{A.4}$$

$$V_{2,t} = \int_0^t (\int_0^{t_2}\xi_{s_1}\lambda_{s_1}\mathrm{d}B_{s_1})\frac{\xi_{s_2}}{\lambda_{s_2}}\mathrm{d}B_{s_2}. \tag{A.5}$$

(A.1)-(A.3)-(A.4)-(A.5) and applications of Lemma 2.3.1 allow to obtain:

$$\begin{aligned}&\mathbb{E}[h^{(1)}(X_T^P)\frac{X_{2,T}}{2}] + \frac{1}{8}\mathbb{E}[h^{(2)}(X_T^P)\int_0^T \frac{V_{1,t}^2}{v_t}\sigma_t^2\mathrm{d}t] \qquad (A.6)\\ =&C_{3b,T}^l[\mathcal{G}_4^h(X_T^P) - 2\mathcal{G}_3^h(X_T^P) + \frac{5\mathcal{G}_2^h(X_T^P)}{4} - \frac{\mathcal{G}_1^h(X_T^P)}{4}] + \frac{C_{2b,T}^l}{2}[\mathcal{G}_2^h(X_T^P) - \mathcal{G}_1^h(X_T^P)]\\ &+ C_{4,T}^l[\mathcal{G}_4^h(X_T^P) - \frac{5\mathcal{G}_3^h(X_T^P)}{2} + 2\mathcal{G}_2^h(X_T^P) - \frac{\mathcal{G}_1^h(X_T^P)}{2}] + C_{4,T}^{ls}[\frac{\mathcal{G}_4^h(X_T^P)}{2} - \mathcal{G}_3^h(X_T^P) + \frac{\mathcal{G}_2^h(X_T^P)}{2}]\\ &+ (C_{2,T}^{ls} + C_{3,T}^{ls})[\frac{\mathcal{G}_4^h(X_T^P)}{2} - \frac{5\mathcal{G}_3^h(X_T^P)}{4} + \frac{\mathcal{G}_2^h(X_T^P)}{2}] + C_{1,T}^{ls}[\frac{\mathcal{G}_2^h(X_T^P)}{2} - \mathcal{G}_1^h(X_T^P)]\\ &+ C_{3a,T}^l[-\mathcal{G}_3^h(X_T^P) + \mathcal{G}_2^h(X_T^P) - \frac{\mathcal{G}_1^h(X_T^P)}{4}] - \frac{C_{2a,T}^l}{2}\mathcal{G}_1^h(X_T^P) + \frac{C_{2,T}^s}{4}[\mathcal{G}_4^h(X_T^P) - \mathcal{G}_3^h(X_T^P)],\end{aligned}$$

where:

$$C^l_{2a,T} = \omega(\sigma^2 v, (\sigma^{(1)})^2 v)^T_0, \qquad C^l_{2b,T} = \omega(\sigma^2 v, \sigma\sigma^{(2)} v)^T_0, \qquad C^l_{3a,T} = \omega(\sigma^2 v, \sigma^2 v, (\sigma^{(1)})^2 v)^T_0,$$
$$C^l_{3b,T} = \omega(\sigma^2 v, \sigma^2 v, \sigma\sigma^{(2)} v)^T_0, \quad C^l_{4,T} = \omega(\sigma^2 v, \sigma\sigma^{(1)} v, \sigma\sigma^{(1)} v)^T_0, \quad C^s_{2,T} = \omega(\rho\xi\sigma v, \rho\xi\sigma, \sigma^2)^T_0,$$
$$C^{ls}_{1,T} = \omega(\rho\xi\sigma v, \sigma\sigma^{(1)})^T_0, \qquad C^{ls}_{2,T} = \omega(\rho\xi\sigma v, \sigma^2 v, \sigma\sigma^{(1)})^T_0, \qquad C^{ls}_{3,T} = \omega(\sigma^2 v, \rho\xi\sigma v, \sigma\sigma^{(1)})^T_0,$$
$$C^{ls}_{4,T} = \omega(\rho\xi\sigma v, \sigma^2, \sigma\sigma^{(1)} v)^T_0.$$

▷**Step 3: contribution with $X^2_{1,T}$.** Starting from (2.7) and applying the Itô formula we have:

$$\frac{1}{2}X^2_{1,T} = \int_0^T X_{1,t}((X^P_t - x_0)\sigma^{(1)}_t \lambda_t + \frac{V_{1,t}\sigma_t}{2\lambda_t})(\mathrm{d}W_t - \sigma_t\lambda_t \mathrm{d}t) \tag{A.7}$$
$$+ \frac{1}{2}\int_0^T ((X^P_t - x_0)^2(\sigma^{(1)}_t)^2 v_t + (X^P_t - x_0)V_{1,t}\sigma_t\sigma^{(1)}_t)\mathrm{d}t + \frac{1}{8}\int_0^T \frac{V^2_{1,t}}{v_t}\sigma^2_t \mathrm{d}t,$$

where:

$$X_{1,t}(X^P_t - x_0) = \int_0^t [(X^P_{t_1} - x_0)^2\sigma^{(1)}_{t_1}\lambda_{t_1} + \frac{(X^P_{t_1} - x_0)V_{1,t_1}\sigma_{t_1}}{2\lambda_{t_1}}](\mathrm{d}W_{t_1} - \sigma_{t_1}\lambda_{t_1}\mathrm{d}t_1)$$
$$+ \int_0^t X_{1,t_1}\sigma_{t_1}\lambda_{t_1}(\mathrm{d}W_{t_1} - \frac{\sigma_{t_1}\lambda_{t_1}}{2}\mathrm{d}t_1) + \int_0^t [(X^P_{t_1} - x_0)\sigma_{t_1}\sigma^{(1)}_{t_1} v_{t_1} + \frac{V_{1,t_1}}{2}\sigma^2_{t_1}]\mathrm{d}t_1, \tag{A.8}$$
$$X_{1,t}V_{1,t} = \int_0^T [(X^P_{t_1} - x_0)V_{1,t_1}\sigma^{(1)}_{t_1}\lambda_{t_1} + \frac{V^2_{1,t_1}\sigma_{t_1}}{2\lambda_{t_1}}](\mathrm{d}W_{t_1} - \sigma_{t_1}\lambda_{t_1}\mathrm{d}t_1) + \int_0^t X_{1,t_1}\xi_{t_1}\lambda_{t_1}\mathrm{d}B_{t_1}$$
$$+ \int_0^t \rho_{t_1}\xi_{t_1}[(X^P_{t_1} - x_0)\sigma^{(1)}_{t_1} v_{t_1} + \frac{V_{1,t_1}\sigma_{t_1}}{2}]\mathrm{d}t_1, \tag{A.9}$$
$$V^2_{1,t} = 2\int_0^t (\int_0^t \xi_{t_1}\lambda_{t_1}\mathrm{d}B_{t_1})\xi_{t_2}\lambda_{t_2}\mathrm{d}B_{t_2} + \int_0^t \xi^2_t v_t \mathrm{d}t \tag{A.10}$$

From Lemma 2.3.1 and (A.8)-(A.3)-(A.1)-(A.4) it follows that:

$$\mathbb{E}[h^{(2)}(X^P_T)\int_0^T X_{1,t}(X^P_t - x_0)\sigma^{(1)}_t\lambda_t(\mathrm{d}W_t - \sigma_t\lambda_t\mathrm{d}t)] \tag{A.11}$$
$$= C^l_{11a,T}[2\mathcal{G}^h_6(X^P_T) - 6\mathcal{G}^h_5(X^P_T) + \frac{13\mathcal{G}^h_4(X^P_T)}{2} - 3\mathcal{G}^h_3(X^P_T) + \frac{\mathcal{G}^h_2(X^P_T)}{2}] + C^l_{4,T}[\mathcal{G}^h_4(X^P_T) - 2\mathcal{G}^h_3(X^P_T) + \mathcal{G}^h_2(X^P_T)]$$
$$+ (C^{ls}_{7a,T} + C^{ls}_{7b,T})[\frac{\mathcal{G}^h_6(X^P_T)}{2} - \frac{5\mathcal{G}^h_5(X^P_T)}{4} + \mathcal{G}^h_4(X^P_T) - \frac{\mathcal{G}^h_3(X^P_T)}{4}] + C^{ls}_{4,T}[\frac{\mathcal{G}^h_4(X^P_T)}{2} - \mathcal{G}^h_3(X^P_T) + \frac{\mathcal{G}^h_2(X^P_T)}{2}]$$
$$+ C^l_{11b,T}[\mathcal{G}^h_6(X^P_T) - 3\mathcal{G}^h_5(X^P_T) + \frac{13\mathcal{G}^h_4(X^P_T)}{4} - \frac{3\mathcal{G}^h_3(X^P_T)}{2} + \frac{\mathcal{G}^h_2(X^P_T)}{4}] + C^{ls}_{7c,T}[\frac{\mathcal{G}^h_6(X^P_T)}{2} - \frac{5\mathcal{G}^h_5(X^P_T)}{4} + \mathcal{G}^h_4(X^P_T)$$
$$- \frac{\mathcal{G}^h_3(X^P_T)}{4}] + C^l_{4,T}[\mathcal{G}^h_4(X^P_T) - \frac{3\mathcal{G}^h_3(X^P_T)}{2} + \frac{\mathcal{G}^h_2(X^P_T)}{2}] + \frac{C^{ls}_{4,T}}{2}[\mathcal{G}^h_4(X^P_T) - \mathcal{G}^h_3(X^P_T)].$$

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 2.3.1 and (A.9)-(A.4)-(A.10)-(A.1), we have:

$$\frac{1}{2}\mathbb{E}[h^{(2)}(X_T^P)\int_0^T \frac{X_{1,t}V_{1,t}\sigma_t}{\lambda_t}(\mathrm{d}W_t - \sigma_t\lambda_t\mathrm{d}t)] \tag{A.12}$$
$$\begin{aligned}
=&(C_{7d,T}^{ls} + C_{7e,T}^{ls})[\frac{\mathcal{G}_6^h(X_T^P)}{2} - \frac{5\mathcal{G}_5^h(X_T^P)}{4} + \mathcal{G}_4^h(X_T^P) - \frac{\mathcal{G}_3^h(X_T^P)}{4}] + C_{5,T}^{ls}[\frac{\mathcal{G}_4^h(X_T^P)}{2} - \mathcal{G}_3^h(X_T^P) + \frac{\mathcal{G}_2^h(X_T^P)}{2}]\\
&+ \frac{C_{4a,T}^s}{2}[\mathcal{G}_6^h(X_T^P) - 2\mathcal{G}_5^h(X_T^P) + \mathcal{G}_4^h(X_T^P)] + \frac{C_{3,T}^s}{4}[\mathcal{G}_4^h(X_T^P) - 2\mathcal{G}_3^h(X_T^P) + \mathcal{G}_2^h(X_T^P)]\\
&+ C_{7f,T}^{ls}[\frac{\mathcal{G}_6^h(X_T^P)}{2} - \frac{5\mathcal{G}_5^h(X_T^P)}{4} + \mathcal{G}_4^h(X_T^P) - \frac{\mathcal{G}_3^h(X_T^P)}{4}] + \frac{C_{4b,T}^s}{4}[\mathcal{G}_6^h(X_T^P) - 2\mathcal{G}_5^h(X_T^P) + \mathcal{G}_4^h(X_T^P)]\\
&+ C_{6,T}^{ls}[\frac{\mathcal{G}_4^h(X_T^P)}{2} - \frac{3\mathcal{G}_3^h(X_T^P)}{4} + \frac{\mathcal{G}_2^h(X_T^P)}{4}] + \frac{C_{2,T}^s}{4}[\mathcal{G}_4^h(X_T^P) - \mathcal{G}_3^h(X_T^P)].
\end{aligned}$$

where:

$$\begin{aligned}
&C_{2,T}^s = \omega(\rho\xi\sigma\nu, \rho\xi\sigma, \sigma^2)_0^T, && C_{3,T}^s = \omega(\xi^2\nu, \sigma^2, \sigma^2)_0^T, && C_{4a,T}^s = \omega(\rho\xi\sigma\nu, \xi\rho\nu\sigma, \sigma^2, \sigma^2)_0^T,\\
&C_{4b,T}^s = \omega(\rho\xi\nu\sigma, \sigma^2, \rho\xi\sigma\nu, \sigma^2)_0^T, && C_{5,T}^{ls} = \omega(\rho\xi\sigma\nu, \sigma\sigma^{(1)}\nu, \sigma^2)_0^T, && C_{6,T}^{ls} = \omega(\sigma^2\nu, \rho\xi\sigma^{(1)}\nu, \sigma^2)_0^T,\\
&C_{7d,T}^{ls} = \omega(\rho\xi\sigma\nu, \sigma^2\nu, \sigma\sigma^{(1)}\nu, \sigma^2)_0^T, && C_{7e,T}^{ls} = \omega(\sigma^2\nu, \rho\xi\sigma\nu, \sigma\sigma^{(1)}\nu, \sigma^2)_0^T, && C_{7f,T}^{ls} = \omega(\sigma^2\nu, \sigma\sigma^{(1)}\nu, \rho\xi\sigma\nu, \sigma^2)_0^T.
\end{aligned}$$

Then using again Lemma 2.3.1 and (A.3)-(A.4) it comes:

$$\frac{1}{2}\mathbb{E}[h^{(2)}(X_T^P)\int_0^T ((X_t^P - x_0)^2(\sigma_t^{(1)})^2\nu_t + (X_t^P - x_0)V_{1,t}\sigma_t\sigma_t^{(1)})\mathrm{d}t] \tag{A.13}$$
$$=C_{3a,T}^l[\mathcal{G}_4^h(X_T^P) - \mathcal{G}_3^h(X_T^P) + \frac{\mathcal{G}_2^h(X_T^P)}{4}] + \frac{C_{2a,T}^l}{2}\mathcal{G}_2^h(X_T^P) + (C_{2,T}^{ls} + C_{3,T}^{ls})[\frac{\mathcal{G}_4^h(X_T^P)}{2} - \frac{\mathcal{G}_3^h(X_T^P)}{4}] + \frac{C_{1,T}^{ls}}{2}\mathcal{G}_2^h(X_T^P).$$

Finally, we sum the contributions (A.11)-(A.12)-(A.13) to obtain in view of (A.7):

$$\frac{1}{2}\mathbb{E}[h^{(2)}(X_T^P)X_{1,T}^2] - \frac{1}{8}\mathbb{E}[h^{(2)}(X_T^P)\int_0^T \frac{V_{1,t}^2}{\nu_t}\sigma_t^2\mathrm{d}t] \tag{A.14}$$
$$\begin{aligned}
=&(2C_{11a,T}^l + C_{11b,T}^l)[\mathcal{G}_6^h(X_T^P) - 3\mathcal{G}_5^h(X_T^P) + \frac{13\mathcal{G}_4^h(X_T^P)}{4} - \frac{3\mathcal{G}_3^h(X_T^P)}{2} + \frac{\mathcal{G}_2^h(X_T^P)}{4}]\\
&+ C_{4,T}^l[2\mathcal{G}_4^h(X_T^P) - \frac{7\mathcal{G}_3^h(X_T^P)}{2} + \frac{3\mathcal{G}_2^h(X_T^P)}{2}] + C_{4,T}^{ls}[\mathcal{G}_4^h(X_T^P) - \frac{3\mathcal{G}_3^h(X_T^P)}{2} + \frac{\mathcal{G}_2^h(X_T^P)}{2}]\\
&+ (C_{7a,T}^{ls} + C_{7b,T}^{ls} + C_{7c,T}^{ls} + C_{7d,T}^{ls} + C_{7e,T}^{ls} + C_{7f,T}^{ls})[\frac{\mathcal{G}_6^h(X_T^P)}{2} - \frac{5\mathcal{G}_5^h(X_T^P)}{4} + \mathcal{G}_4^h(X_T^P) - \frac{\mathcal{G}_3^h(X_T^P)}{4}]\\
&+ C_{5,T}^{ls}[\frac{\mathcal{G}_4^h(X_T^P)}{2} - \mathcal{G}_3^h(X_T^P) + \frac{\mathcal{G}_2^h(X_T^P)}{2}] + \frac{C_{3,T}^s}{4}[\mathcal{G}_4^h(X_T^P) - 2\mathcal{G}_3^h(X_T^P) + \mathcal{G}_2^h(X_T^P)] + \frac{C_{2,T}^s}{4}[\mathcal{G}_4^h(X_T^P) - \mathcal{G}_3^h(X_T^P)]\\
&+ (2C_{4a,T}^s + C_{4b,T}^s)[\frac{\mathcal{G}_6^h(X_T^P)}{4} - \frac{\mathcal{G}_5^h(X_T^P)}{2} + \frac{\mathcal{G}_4^h(X_T^P)}{4}] + C_{6,T}^{ls}[\frac{\mathcal{G}_4^h(X_T^P)}{2} - \frac{3\mathcal{G}_3^h(X_T^P)}{4} + \frac{\mathcal{G}_2^h(X_T^P)}{4}]\\
&+ C_{3a,T}^l[\mathcal{G}_4^h(X_T^P) - \mathcal{G}_3^h(X_T^P) + \frac{\mathcal{G}_2^h(X_T^P)}{4}] + \frac{C_{2a,T}^l}{2}\mathcal{G}_2^h(X_T^P) + (C_{2,T}^{ls} + C_{3,T}^{ls})[\frac{\mathcal{G}_4^h(X_T^P)}{2} - \frac{\mathcal{G}_3^h(X_T^P)}{4}] + \frac{C_{1,T}^{ls}}{2}\mathcal{G}_2^h(X_T^P).
\end{aligned}$$

▷**Step 4: some mathematical reductions.** There are some relations between the expansion coefficients. The reader can easily verify that:

$$\frac{(C_{1,T}^s)^2}{2} = 2C_{4a,T}^s + C_{4b,T}^s, \qquad \frac{(C_{1,T}^l)^2}{2} = 2C_{11a,T}^l + C_{11b,T}^l,$$

$$C^l_{1,T}C^s_{1,T} = C^{ls}_{7a,T} + C^{ls}_{7b,T} + C^{ls}_{7c,T} + C^{ls}_{7d,T} + C^{ls}_{7e,T} + C^{ls}_{7f,T}.$$

The first identity is proved in [BGM10b, 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 (A.2)-(A.6)-(A.14) of steps $1-2-3$, we obtain the announced formula (2.2), putting together the corrective terms according to the order of the Greeks.

A.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 A.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 $H_j(x) = (-1)^j e^{x^2/2}\partial^j_{x^j}(e^{-x^2/2})\ \forall (j,x) \in \mathbb{N}\times\mathbb{R}$.

In the next Proposition, we provide the formulas of the $\mathrm{Vega}^{\mathrm{BS}}$ and the $\mathrm{Vomma}^{\mathrm{BS}}$:

Proposition A.3.2. *Let $x, k \in \mathbb{R}$, $\sigma > 0$ and $T > 0$. We have:*

$$\begin{aligned}
\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k) &= \partial_\sigma\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2T,k) = e^x\sqrt{T}\mathcal{N}'(d_1(x,\sigma^2T,k)),\\
\mathrm{Vomma}^{\mathrm{BS}}(x,\sigma^2T,k) &= \partial_\sigma\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k) = \frac{\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)}{\sigma}d_1(x,\sigma^2T,k)d_2(x,\sigma^2T,k)\\
&= \frac{\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)}{\sigma}\Big[\frac{(x-k)^2}{\sigma^2T} - \frac{\sigma^2T}{4}\Big].
\end{aligned}$$

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 A.3.3. *Let $x, k \in \mathbb{R}$, $\sigma > 0$ and $T > 0$. We have:*

$$\begin{aligned}
&(\partial^2_{x^2} - \partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2T,k) = \frac{\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)}{\sigma T},\\
&(\partial^3_{x^3} - \frac{3}{2}\partial^2_{x^2} + \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2T,k) = -\mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)\frac{(x-k)}{\sigma^3T^2},\\
&(\partial^4_{x^4} - 2\partial^3_{x^3} + \frac{5}{4}\partial^2_{x^2} - \frac{1}{4}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2T,k) = \mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)\Big[\frac{(x-k)^2}{\sigma^5T^3} - \frac{1}{\sigma^3T^2}\Big],\\
&(3\partial^4_{x^4} - 6\partial^3_{x^3} + \frac{7}{2}\partial^2_{x^2} - \frac{1}{2}\partial_x)\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2T,k) = \mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)\Big[3\frac{(x-k)^2}{\sigma^5T^3} - \frac{3}{\sigma^3T^2} - \frac{1}{4\sigma T}\Big],
\end{aligned}$$

$$\begin{aligned}
&(\frac{1}{2}\partial^6_{x^6} - \frac{3}{2}\partial^5_{x^5} + \frac{13}{8}\partial^4_{x^4} - \frac{3}{4}\partial^3_{x^3} + \frac{1}{8}\partial^2_{x^2})\mathrm{Call}^{\mathrm{BS}}(x,\sigma^2T,k)\\
&= \mathrm{Vega}^{\mathrm{BS}}(x,\sigma^2T,k)\Big[-3\frac{(x-k)^2}{\sigma^7T^4} + \frac{1}{8\sigma^3T^2} + \frac{3}{2\sigma^5T^3}\Big] + \frac{1}{2}\mathrm{Vomma}^{\mathrm{BS}}(x,\sigma^2T,k)\frac{(x-k)^2}{\sigma^6T^4},
\end{aligned}$$

$$(\frac{1}{2}\partial^3_{x^3} - \frac{1}{2}\partial^2_{x^2})\text{Call}^{\text{BS}}(x,\sigma^2T,k) = \text{Vega}^{\text{BS}}(x,\sigma^2T,k)[-\frac{(x-k)}{2\sigma^3T^2} + \frac{1}{4\sigma T}],$$
$$(\frac{1}{2}\partial^4_{x^4} - \frac{1}{2}\partial^3_{x^3})\text{Call}^{\text{BS}}(x,\sigma^2T,k) = \text{Vega}^{\text{BS}}(x,\sigma^2T,k)[\frac{(x-k)^2}{2\sigma^5T^3} + \frac{1}{8\sigma T} - \frac{1}{2\sigma^3T^2} - \frac{(x-k)}{2\sigma^3T^2}],$$
$$(\frac{1}{4}\partial^4_{x^4} - \frac{1}{2}\partial^3_{x^3} + \frac{1}{4}\partial^2_{x^2})\text{Call}^{\text{BS}}(x,\sigma^2T,k) = \text{Vega}^{\text{BS}}(x,\sigma^2T,k)[\frac{(x-k)^2}{4\sigma^5T^3} - \frac{1}{16\sigma T} - \frac{1}{4\sigma^3T^2}],$$
$$(\partial^4_{x^4} - \frac{3}{2}\partial^3_{x^3} + \frac{1}{2}\partial^2_{x^2})\text{Call}^{\text{BS}}(x,\sigma^2T,k) = \text{Vega}^{\text{BS}}(x,\sigma^2T,k)[\frac{(x-k)^2}{\sigma^5T^3} - \frac{1}{\sigma^3T^2} - \frac{(x-k)}{2\sigma^3T^2}],$$
$$(\frac{3}{2}\partial^4_{x^4} - \frac{5}{2}\partial^3_{x^3} + \partial^2_{x^2})\text{Call}^{\text{BS}}(x,\sigma^2T,k) = \text{Vega}^{\text{BS}}(x,\sigma^2T,k)[\frac{3(x-k)^2}{2\sigma^5T^3} - \frac{3}{2\sigma^3T^2} - \frac{(x-k)}{2\sigma^3T^2} - \frac{1}{8\sigma T}],$$
$$(\frac{1}{2}\partial^6_{x^6} - \frac{5}{4}\partial^5_{x^5} + \partial^4_{x^4} - \frac{1}{4}\partial^3_{x^3})\text{Call}^{\text{BS}}(x,\sigma^2T,k)$$
$$= \text{Vega}^{\text{BS}}(x,\sigma^2T,k)[-3\frac{(x-k)^2}{\sigma^7T^4} + \frac{3}{2\sigma^5T^3} + \frac{3(x-k)}{4\sigma^5T^3} + \frac{1}{8\sigma^3T^2}]$$
$$+ \frac{\text{Vomma}^{\text{BS}}(x,\sigma^2T,k)}{2}[\frac{(x-k)^2}{\sigma^6T^4} - \frac{(x-k)}{2\sigma^4T^3}],$$
$$(\frac{1}{8}\partial^6_{x^6} - \frac{1}{4}\partial^5_{x^5} + \frac{1}{8}\partial^4_{x^4})\text{Call}^{\text{BS}}(x,\sigma^2T,k)$$
$$= \text{Vega}^{\text{BS}}(x,\sigma^2T,k)[-3\frac{(x-k)^2}{4\sigma^7T^4} + \frac{3}{8\sigma^5T^3} + \frac{3(x-k)}{8\sigma^5T^3}] + \frac{\text{Vomma}^{\text{BS}}(x,\sigma^2T,k)}{2}[-\frac{(x-k)}{2\sigma^3T^2} + \frac{1}{4\sigma T}]^2.$$

A.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:

$$\begin{aligned} dX_t &= \sigma(X_t)\sqrt{Y_t}[dW_t - \frac{\sigma(X_t)\sqrt{Y_t}}{2}dt], \quad X_0 = x_0, \\ dY_t &= \kappa(\theta - Y_t)dt + \xi\sqrt{Y_t}dB_t, \quad Y_0 = v_0, \\ d\langle W, B\rangle_t &= \rho dt. \end{aligned}$$

In view of A.1, we can apply our different price and implied volatility expansion theorems by considering in the various corrective coefficients C (defined in Theorem 2.2.1 and Definitions 4.1.1 and 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 A.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:*

$$\int_0^T v_t e^{-\kappa t}dt = \overline{v}T, \qquad C^l_{1,T}(x) = \frac{\sigma^3(x)\sigma^{(1)}(x)\overline{v}^2T^2}{2},$$
$$C^l_{2,T}(x) = \frac{\sigma^2(x)[(\sigma^{(1)})^2 + \sigma\sigma^{(2)}](x)\overline{v}^2T^2}{2}, \qquad C^l_{3,T}(x) = C^l_{9,T}(x) = \frac{\sigma^4(x)[(\sigma^{(1)})^2 + \sigma\sigma^{(2)}](x)\overline{v}^3T^3}{6},$$
$$C^l_{4,T}(x) = \frac{\sigma^4(x)(\sigma^{(1)})^2(x)\overline{v}^3T^3}{6}, \qquad C^l_{10,T}(x) = \frac{\sigma^6(x)(\sigma^{(1)})^2(x)\overline{v}^4T^4}{24},$$

$$
\begin{aligned}
C^{l}_{5,T}(x) =& [(\sigma^{(1)})^2 + \sigma\sigma^{(2)}](x)\bar{v}T, & C^{l}_{6,T}(x) =& \frac{[\sigma\sigma^{(1)}]^2(x)\bar{v}^2T^2}{2},\\
C^{s}_{1,T}(x) =& \frac{\rho\xi\sigma^3(x)R^s_1T^2}{2}, & C^{s}_{2,T}(x) =& \frac{\rho^2\xi^2\sigma^4(x)R^s_2T^3}{6},\\
C^{s}_{3,T}(x) =& \frac{\xi^2\sigma^4(x)R^s_3T^3}{6}, & C^{ls}_{1,T}(x) =& C^{ls}_{9,T}(x) = \frac{\rho\xi\sigma^2(x)\sigma^{(1)}(x)R^s_1T^2}{2},\\
C^{ls}_{2,T}(x) =& C^{ls}_{5,T}(x) = \frac{\rho\xi\sigma^4(x)\sigma^{(1)}(x)R^{ls}_1T^3}{6}, & C^{ls}_{3,T}(x) =& C^{ls}_{6,T}(x) = C^{ls}_{10,T}(x) = \frac{\rho\xi\sigma^4(x)\sigma^{(1)}(x)R^{ls}_2T^3}{6},\\
C^{ls}_{4,T}(x) =& C^{ls}_{8,T}(x) = \frac{\rho\xi\sigma^4(x)\sigma^{(1)}(x)R^{ls}_3T^3}{6},
\end{aligned}
$$

where:

$$
\begin{aligned}
\bar{v} &= (v_0-\theta)\frac{e^{-\kappa T}\left(-1+e^{\kappa T}\right)}{\kappa T} + \theta,\\
R^s_1 &= (v_0-\theta)\frac{e^{-\kappa T}\left(-2\kappa T + 2e^{\kappa T} - 2\right)}{\kappa^2T^2} + \theta\frac{e^{-\kappa T}\left(2\kappa T e^{\kappa T} - 2e^{\kappa T} + 2\right)}{\kappa^2T^2},\\
R^s_2 &= (v_0-\theta)\frac{e^{-\kappa T}\left(-3\kappa T(\kappa T+2) + 6e^{\kappa T} - 6\right)}{\kappa^3T^3} + \theta\frac{e^{-\kappa T}\left(6e^{\kappa T}(\kappa T-2) + 6\kappa T + 12\right)}{\kappa^3T^3},\\
R^s_3 &= (v_0-\theta)\frac{e^{-2\kappa T}\left(-6e^{\kappa T}\kappa T + 3e^{2\kappa T} - 3\right)}{\kappa^3T^3} + \theta\frac{e^{-2\kappa T}\left(12e^{\kappa T} + 3e^{2\kappa T}(2\kappa T-3) - 3\right)}{2\kappa^3T^3},\\
R^{ls}_1 &= \frac{3}{2\kappa^3T^3}\{e^{-2\kappa T}(v_0-\theta)\big(v_0(3+2\kappa T) - \theta(5+2\kappa T)\big) + 2e^{-\kappa T}\big(\theta^2(4+\kappa T(6+\kappa T))\\
&\quad - \theta v_0(-2+\kappa T(4+\kappa T)) - 2v_0^2\big) + \theta^2(4\kappa T - 13) + 4\theta v_0 + v_0^2\},\\
R^{ls}_2 &= \frac{3}{\kappa^3T^3}\{-e^{-2\kappa T}(v_0-\theta)^2 + e^{-\kappa T}\big(\theta^2(-4+\kappa T(-2+\kappa T)) - \theta v_0(-2+\kappa T(-4+\kappa T)) - 2\kappa T v_0^2\big)\\
&\quad + \theta^2(5+\kappa T(-4+\kappa T)) + 2\theta v_0(-2+\kappa T) + v_0^2\},\\
R^{ls}_3 &= \frac{3}{2\kappa^3T^3}\{e^{-2\kappa T}(v_0-\theta)\big(v_0(3+2\kappa T) - \theta(5+2\kappa T)\big) - 4e^{-\kappa T}(v_0-2\theta)^2\\
&\quad + \theta^2(11+2\kappa T(-4+\kappa T)) + 4\theta v_0(-2+\kappa T) + v_0^2\}.
\end{aligned}
$$

We have in addition the relation:

$$
\bar{v}T\frac{R^s_1T^2}{2} = \frac{(R^{ls}_1 + R^{ls}_2 + R^{ls}_3)T^3}{6} \tag{A.15}
$$

Using the relation (A.15), one gets without difficulty :

$$
\begin{aligned}
\frac{C^{ls}_{8,T}(x)}{\bar{\sigma}_x^3T^2} - \frac{3C^{ls}_{4,T}(x)}{2\bar{\sigma}_x^3T^2} - \frac{C^{ls}_{5,T}(x)}{2\bar{\sigma}_x^3T^2} - \frac{C^{ls}_{6,T}(x)}{2\bar{\sigma}_x^3T^2} + \frac{3(C^l_{1,T}C^s_{1,T})(x)}{2\bar{\sigma}_x^5T^3} &= \frac{(C^l_{1,T}C^s_{1,T})(x)}{2\bar{\sigma}_x^5T^3},\\
-\frac{C^{ls}_{4,T}(x)}{8\bar{\sigma}_xT} - \frac{C^{ls}_{5,T}(x)}{8\bar{\sigma}_xT} + \frac{(C^l_{1,T}C^s_{1,T})(x)}{8\bar{\sigma}_x^3T^2} &= \frac{C^{ls}_{6,T}(x)}{8\bar{\sigma}_xT} - \frac{(C^l_{1,T}C^s_{1,T})(x)}{8\bar{\sigma}_x^3T^2},\\
-\frac{(C^{ls}_{2,T} + C^{ls}_{3,T})(x)}{2\bar{\sigma}_x^3T^2} - \frac{C^{ls}_{4,T}(x)}{2\bar{\sigma}_x^3T^2} - \frac{C^{ls}_{6,T}(x)}{4\bar{\sigma}_x^3T^2} + \frac{3(C^l_{1,T}C^s_{1,T})(x)}{4\bar{\sigma}_x^5T^3} &= -\frac{C^{ls}_{6,T}(x)}{4\bar{\sigma}_x^3T^2} - \frac{(C^l_{1,T}C^s_{1,T})(x)}{4\bar{\sigma}_x^5T^3},
\end{aligned}
$$

$$\frac{(C^{ls}_{2,T}+C^{ls}_{3,T})(x)}{\bar\sigma_x^5T^3}+\frac{3C^{ls}_{4,T}(x)}{2\bar\sigma_x^5T^3}+\frac{C^{ls}_{5,T}(x)}{2\bar\sigma_x^5T^3}+\frac{C^{ls}_{6,T}(x)}{2\bar\sigma_x^5T^3}-\frac{3(C^{l}_{1,T}C^{s}_{1,T})(x)}{\bar\sigma_x^7T^4}=0,$$

$$\frac{C^{ls}_{1,T}(x)}{4\bar\sigma_xT}+\frac{C^{ls}_{9,T}(x)}{8\bar\sigma_xT}-\frac{C^{ls}_{4,T}(x)+C^{ls}_{5,T}(x)+C^{ls}_{10,T}(x)}{8\bar\sigma_x^3T^2}=\frac{(C^{l}_{1,T}C^{s}_{1,T})(x)}{2\bar\sigma_x^5T^3},$$

$$-\frac{C^{ls}_{1,T}(x)}{2\bar\sigma_x^3T^2}-\frac{C^{ls}_{9,T}(x)}{4\bar\sigma_x^3T^2}+3\frac{C^{ls}_{4,T}(x)+C^{ls}_{5,T}(x)+C^{ls}_{10,T}(x)}{4\bar\sigma_x^5T^3}=0.$$

Then the above mathematical reductions allow to obtain the following expressions for the coefficients γ and π defined in Definitions 4.1.1 and 4.2.1 :

Proposition A.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:*

$$\begin{aligned}
\gamma_{0a,T}(x_0) =&\sigma(x_0)\sqrt{\bar v}\{1+\frac{\rho\xi\sigma(x_0)R_1^sT}{8\bar v}\},\\
\gamma_{1a,T}(x_0) =&\sigma(x_0)\sqrt{\bar v}\{-\frac{\sigma^{(1)}(x_0)}{2\sigma(x_0)}-\frac{\rho\xi R_1^s}{4\sigma(x_0)\bar v^2}\},\\
\gamma_{0b,T}(x_0) =&\sigma(x_0)\sqrt{\bar v}\{\bar vT[\frac{\sigma(x_0)\sigma^{(2)}(x_0)}{12}-(\sigma^{(1)})^2(x_0)(\frac{1}{24}+\frac{\sigma^2(x_0)\bar vT}{96})]\\
&+\frac{\rho^2\xi^2T}{\bar v^2}[\frac{3(R_1^s)^2}{32\bar v}+R_2^s(\frac{\sigma^2(x_0)\bar vT}{48}-\frac{1}{12})]-\frac{\xi^2TR_3^s}{\bar v^2}[\frac{1}{24}+\frac{\sigma^2(x_0)\bar vT}{96}]\\
&+\frac{\rho\xi\sigma^{(1)}(x_0)T}{\bar v}[\frac{R_1^s}{8}+\sigma^2(x_0)\bar vT(\frac{R_2^{ls}}{48\bar v}-\frac{R_1^s}{32})]\},\\
\gamma_{1b,T}(x_0) =&\sigma(x_0)\sqrt{\bar v}\{\frac{\rho^2\xi^2T}{\bar v^2}[\frac{3(R_1^s)^2}{32\bar v}-\frac{R_2^s}{12}]-\frac{\rho\xi\sigma^{(1)}(x_0)T}{\bar v}[\frac{R_1^s}{16}+\frac{R_2^{ls}}{24\bar v}]\},\\
\gamma_{2,T}(x_0) =&\sigma(x_0)\sqrt{\bar v}\{\frac{\sigma^{(2)}(x_0)}{6\sigma(x_0)}-\frac{(\sigma^{(1)})^2(x_0)}{12\sigma^2(x_0)}+\frac{\rho^2\xi^2}{\sigma^2(x_0)\bar v^3}[\frac{R_2^s}{12}-\frac{3(R_1^s)^2}{16\bar v}]+\frac{\xi^2R_3^s}{24\sigma^2(x_0)\bar v^3}\},\\
\pi_{1a,T}(x_{avg}) =&-\sigma(x_{avg})\sqrt{\bar v}\frac{\rho\xi R_1^s}{4\sigma(x_{avg})\bar v^2},\\
\pi_{1b,T}(x_{avg}) =&\sigma(x_{avg})\sqrt{\bar v}\{\frac{\rho^2\xi^2T}{\bar v^2}[\frac{3(R_1^s)^2}{32\bar v}-\frac{R_2^s}{12}]+\frac{\rho\xi\sigma^{(1)}(x_{avg})T}{\bar v}[\frac{R_1^s}{16}-\frac{R_2^{ls}}{24\bar v}]\},\\
\pi_{2,T}(x_{avg}) =&\sigma(x_{avg})\sqrt{\bar v}\{\frac{\sigma^{(2)}(x_{avg})}{24\sigma(x_{avg})}-\frac{(\sigma^{(1)})^2(x_{avg})}{12\sigma^2(x_{avg})}+\frac{\rho^2\xi^2}{\sigma^2(x_{avg})\bar v^3}[\frac{R_2^s}{12}-\frac{3(R_1^s)^2}{16\bar v}]+\frac{\xi^2R_3^s}{24\sigma^2(x_{avg})\bar v^3}\}.
\end{aligned}$$

References

[AE08] E. Alos and C.O. Ewald. Malliavin differentiability of the heston volatility and applications to option pricing. *Advances in Applied Probability*, 40(1):144–162, 2008.

[AP06] L. Andersen and V. Piterbarg. Moment explosion in stochastic volatility models. *Finance and Stochastics*, 11(1):29–50, 2006.

[BBF04] H. Berestycki, J. Busca, and I. Florent. Computing the implied volatility in stochastic volatility models. *Comm. Pure Appl. Math.*, 10:1352–1373, 2004.

[BG12] R. Bompis and E. Gobet. Asymptotic and non asymptotic approximations for option valuation. *preprint available on HAL and SSRN*, 2012.

[BGM09] E. Benhamou, E. Gobet, and M. Miri. Smart expansion and fast calibration for jump diffusion. *Finance and Stochastics*, 13(4):563–5894, 2009.

[BGM10a] E. Benhamou, E. Gobet, and M. Miri. Expansions formulas of european options in a local volatility model. *International Journal of Theoretical and Applied Finance*, 13(4):603–634, 2010.

[BGM10b] E. Benhamou, E. Gobet, and M. Miri. Time dependent heston model. *SIAM Journal on Financial Mathematics*, 1:289–325, 2010.

[CFK10] S.Y. Choi, J.P. Fouque, and J.H Kim. Option pricing under hybrid stochastic and local volatility. *preprint*, 2010.

[Cox75] 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.

[FJ09] M. Forde and A. Jacquier. Small-time asymptotics for implied volatility under the Heston model. *International Journal of Theoretical and Applied Finance*, 12(6):861–876, 2009.

[FJ12] M. Forde and A. Jacquier. Small-time asymptotics for an uncorrelated local-stochastic volatility model. *forthcoming in Applied Mathematical Finance*, 2012.

[FP12] M. Forde and A. Pogudin. The large-maturity smile for the SABR and CEV-Heston models. *preprint*, 2012.

[FW98] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*. Second ed., Springer-Verlag, New York, 1998.

[Gla04] P. Glasserman. *Monte Carlo methods in financial engineering*. Springer Verlag, New York, 2004.

[GM05] E. Gobet and R. Munos. Sensitivity analysis using itô-malliavin calculus and martingales. application to stochastic control problem. *Journal of Control and Optimization*, 43(5):1676–1713, 2005.

[GM11] E. Gobet and M. Miri. Weak approximation of averaged diffusion processes. *preprint available on HAL*, 2011.

[Hes93] S. Heston. A closed-form solutions for options with stochastic volatility with applications to bond and currency options,. *Review of Financial Studies*, 6:327–343, 1993.

[HKLW02] P. S. Hagan, D. Kumar, A. S. Lesniewski, and D. E. Woodward. Managing smile risk. *Willmott Magazine*, pages 84–108, 2002.

[HW99] P. S. Hagan and D. E. Woodward. Equivalent Black volatilities. *Applied mathematical Finance*, 6:147–157, 1999.

[JT11] Richard Jordan and Charles Tier. Asymptotic approximations to deterministic and stochastic volatility models. *SIAM J. Fin. Math.*, 2(1):935–964, 2011.

[Kel78] J. Keller. Rays, waves and asymptotics. *Bull. Amer. Math. Soc*, 84(78):727–750, 1978.

[Lab05] P. Henry Labordère. A general asymptotic implied volatility for stochastic volatility models. *arXiv:condmat/04317*, 2005.

[Lab08] P. Henry Labordère. *Analysis, Geometry, and Modeling in Finance: Advanced Methods in Option Pricing*. Chapman and Hall, 2008.

[Lew00] A.L. Lewis. *Option Valuation under Stochastic Volatility with Mathematica Code*. Finance Press, Newport Beach, California, 2000.

[Lew07] A. Lewis. Geometries and smile asymptotics for a class of stochastic volatility models. *www.optioncity.net*, 2007.

[Nua06] D. Nualart. *Malliavin calculus and related topics*. Springer-Verlag, Berlin, second edition, 2006.

[PP12] S. Pagliarani and A. Pascucci. Local stochastic volatility with jumps. *Preprint available on SSRN*, 2012.

[Pro04] P. Protter. *Stochastic integration and differential equations*. 2004.