The exact Taylor formula of the implied volatility

Stefano Pagliarani[1] **· Andrea Pascucci**[2]

Abstract In a model driven by a multidimensional local diffusion, we study the behavior of the implied volatility σ and its derivatives with respect to log-strike k and maturity T near expiry and at the money. We recover explicit limits of the derivatives $\partial_T^q \partial_k^m \sigma$ for $(T, x-k)$ approaching the origin within the parabolic region $|x-k| \le \lambda\sqrt{T}$, with x denoting the spot log-price of the underlying asset and where λ is a positive and arbitrarily large constant. Such limits yield the exact Taylor formula for the implied volatility within the parabola $|x-k| \le \lambda\sqrt{T}$. In order to include important models of interest in mathematical finance, e.g. Heston, CEV, SABR, the analysis is carried out under the weak assumption that the infinitesimal generator of the diffusion is only locally elliptic.

Keywords Implied volatility · Local-stochastic volatility · Local diffusions · Feller process

Mathematics Subject Classification (2010) 60J60 · 60J70 · 91G20

JEL Classification C02 · C60

S. Pagliarani's work supported by the Chair *Financial Risks* of the *Risk Foundation*.

✉ A. Pascucci
andrea.pascucci@unibo.it

S. Pagliarani
pagliarani@cmap.polytechnique.fr

1 DEAMS, Università di Trieste, Trieste, Italy

2 Dipartimento di Matematica, Università di Bologna, Bologna, Italy

1 Introduction

This paper deviates from the mainstream literature on asymptotic methods in finance; in fact, our main result does not add another formula to the plethora of approximation formulas for the implied volatility (IV) already available in the literature. Rather, we prove an *exact result:* a rigorous derivation of the exact Taylor formula of IV, as a *function of both strike and maturity*, in a parabolic region close to expiry and at-the-money (ATM).

This is done under general assumptions that allow including popular models, such as the CEV and the Heston models, as particular cases; indeed, we consider a multivariate model driven by a stochastic process that is a *local diffusion* in a sense that suitably generalizes the classical notion of diffusion as given by [18, 19, 45].

The literature on IV asymptotics is extensive and exploits a diverse range of mathematical techniques. Focusing on short-time asymptotics, well-known results were obtained by [6, 7, 15]. Deferring precise definitions until the body of this paper, we denote by $\sigma(t,x;T,k)$ the IV related to a call option with log-strike k and maturity T, where x is the spot log-price of the underlying asset at time t. [7] uses PDEs techniques to prove the existence of the limits $\lim_{T\to t+}\sigma(t,x;T,k)$ in a generic stochastic volatility model and to characterize such limits in terms of Varadhan's geodesic distance (see also [22] for related results). More recently, [15] gives conditions under which it is possible to recover the ATM limits $\lim_{T\to t+}\partial_T^q\partial_k^m\sigma(t,k;T,k)$ using a semi-martingale decomposition of implied volatilities; although this approach performs also in non-Markovian settings, the validity of the conditions for the existence of the limits is verified only under Markovian assumptions and employing the results in [7].

While it is common practice to consider the IV as a function of maturity and strike (T,k), the aforementioned papers examine only the *vertical limits* (see Fig. 1), as $T\to t+$, of $\sigma(t,x;T,k)$. The aim of this paper is to give conditions for the existence and an explicit representation of the limits of $\partial_T^q\partial_k^m\sigma(t,x;T,k)$, at any order m, q, as $(T-t,x-k)$ approaches the origin within the parabolic region $\mathcal{P}_\lambda:=\{|x-k|\le\lambda\sqrt{T-t}\}$; here λ is an arbitrarily large positive parameter. From a practical perspective, $\mathcal{P}_\lambda$ is the region of interest where implied volatility data are typically observed in the market. As a by-product, we also provide a rigorous and explicit derivation of the exact Taylor formula (see formula (1.3) below) for the implied volatility $\sigma(t,x;\cdot,\cdot)$ in $\mathcal{P}_\lambda$, around $(T,k)=(t,x)$.

The starting point is the analysis of the transition density first developed in a scalar setting in [37] and later extended to asymptotic IV expansions in multiple dimensions in [33], where the authors derived a fully explicit approximation, hereafter denoted by $\bar\sigma_N$, for the IV at any given order $N\in\mathbb{N}$. Our main result, Theorem 5.1 below, gives a sharp error bound on $\partial_T^q\partial_k^m(\sigma-\bar\sigma_N)$ and leads to the existence of the limits

$$\lim_{\substack{(T,k)\to(t,x)\\ |x-k|\le\lambda\sqrt{T-t}}}\partial_T^q\partial_k^m(\sigma-\bar\sigma_N)(t,x;T,k)=0,\qquad 2q+m\le N. \tag{1.1}$$

In the one-dimensional case and for derivatives of order less than or equal to two, similar results were proved in [8] by using Malliavin calculus techniques. Our results

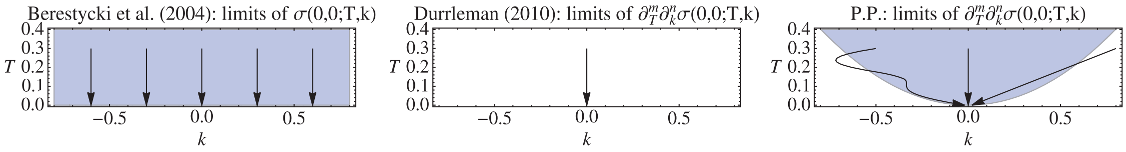

Fig. 1 Directions along which the limits are computed in [7], in [15] and in this paper, respectively

are proved under mild conditions on the driving stochastic process, which is assumed to be a Feller process and an inhomogeneous local diffusion. Loosely speaking, we assume that the infinitesimal generator of the diffusion is only *locally elliptic* (i.e., elliptic on a certain domain $D \subseteq \mathbb{R}^d$) and its coefficients satisfy suitable regularity conditions; note that no ellipticity condition is imposed on the complementary set $\mathbb{R}^d \setminus D$. Results under such general hypotheses appear to be novel compared to the existing literature. In particular, our analysis includes processes with killing and/or degenerate processes: *our assumptions do not even imply that the law of the underlying process has a density* and therefore our results apply to many degenerate cases of interest, such as the well-known CEV, Heston and SABR models, among others.

Formula (1.1) implies that the limits of the derivatives $\partial_T^q \partial_k^m \sigma$ exist if and only if the limits of $\partial_T^q \partial_k^m \bar{\sigma}_N$ do exist, and in that case we have

$$\lim_{\substack{(T,k)\to(t,x)\\ |x-k|\le\lambda\sqrt{T-t}}} \partial_T^q \partial_k^m \sigma(t,x;T,k) = \lim_{\substack{(T,k)\to(t,x)\\ |x-k|\le\lambda\sqrt{T-t}}} \partial_T^q \partial_k^m \bar{\sigma}_N(t,x;T,k). \tag{1.2}$$

Note that in general, the limits in (1.2) do not exist; a simple example is given in [43, Sect. 6], which exhibits a lognormal model with oscillating time-dependent volatility. In that case, the results by [6, 7, 15] do not apply, while the approximation $\bar{\sigma}_N$ in [32] turns out to be exact at order $N = 0$. More generally, we provide simple and explicit conditions ensuring the existence of the limits of $\partial_T^q \partial_k^m \bar{\sigma}_N$, and consequently the existence of those of $\partial_T^q \partial_k^m \sigma$ in (1.2). A particular case is when the underlying diffusion is time-homogeneous; in that case, $\bar{\sigma}_N$ is polynomial in time and thus smooth up to $T = t$.

Denoting by $\partial_T^q \partial_k^m \bar{\sigma}_N(t,x)$ the limits in (1.2), whose explicit expression is known at any order, we get for σ the exact *parabolic* Taylor formula

$$\sigma(t,x;T,k) = \sum_{2q+m\le N} \frac{\partial_T^q \partial_k^m \bar{\sigma}_N(t,x)}{q!m!}(T-t)^q(k-x)^m$$
$$+ \mathrm{o}\Big((T-t)^{\frac{N}{2}} + |k-x|^N\Big) \tag{1.3}$$

as $(T,k) \to (t,x)$ in $\mathcal{P}_\lambda$. Here, the meaning of the adjective *parabolic* is twofold. On the one hand, it refers to the parabolic domain $\mathcal{P}_\lambda$ on which the Taylor formula is proved; on the other hand, it refers to the nature of the remainder, which is expressed in terms of the homogeneous norm typically used to describe the geometry induced by a parabolic differential operator. Note that this formula describes the behavior of σ in a joint regime of small log-moneyness and/or small maturity. This result appears to be novel compared to the existing literature and complementary to [9, 20, 35]. In [20], the asymptotic behavior of σ in a joint regime of extreme strikes and short/long time-to-maturity is studied; [35] studied, in an exponential Lévy model, the small-time asymptotic behavior of σ along relevant curves lying outside the parabolic region $\mathcal{P}_\lambda$ for any $\lambda > 0$; eventually, in a very general setting, [9] studied the asymptotics of σ for different regimes of log-strikes and maturities, including the region $\mathcal{P}_\lambda$ where their result coincides with ours at order zero.

Apart from the mere interest of having at hand a Taylor formula like (1.3), additional advantages of having two-dimensional limits, as opposed to vertical ones, might come from applications such as the asymptotic study of the IV generated by VIX options (see [2]). In this case, the underlying value, given by the price of the future VIX, is not fixed but varies in time, meaning that the log-moneyness of an ATM VIX call is not constantly zero, but approaches zero for small times to maturity along a curve which is not a straight line.

The proof of our result proceeds in several steps. We first introduce a notion of local diffusion (Assumption 2.1); we study its basic properties and the existence of a local transition density. We provide a double characterization of the local density in terms of the forward and backward Kolmogorov equations (Theorem 2.6); the forward representation follows from Hörmander's theorem and is coherent with the classical results by [29]. On the other hand, the backward representation appears to be novel at this level of generality. Indeed, its proof is more delicate and requires the use of the Feller property combined with classical pointwise estimates by [36] for weak solutions of parabolic PDEs. Then we derive sharp asymptotic estimates for the derivatives $\partial_T^q \partial_k^m u(t, x; T, k)$, with u representing the pricing function of a call option with maturity T and log-strike k. This is done first in a uniformly parabolic framework and is then extended to a *locally* parabolic setting to include the majority of the models used in mathematical finance. The second step is particularly interesting due to the very weak assumptions imposed on the generator $\mathcal{A}_t$ of the underlying diffusion. The main idea is to extend $\mathcal{A}_t$ to an operator $\widetilde{\mathcal{A}}_t$ which is globally parabolic and then to prove that locally in space, the difference between the fundamental solution of $\widetilde{\mathcal{A}}_t$ and the local density of the underlying process decays exponentially as the time-to-maturity approaches zero. This last step requires a non-trivial use of some techniques first introduced by [44]. Finally, the estimates on the derivatives $\partial_T^q \partial_k^m u$ are combined with some sharp estimates on the inverse of the BS pricing function and its sensitivities to obtain the main results, Theorem 5.1 and the Taylor formula (1.3).

The paper is organized as follows. In Sect. 2, we describe the general setting and show some illustrative examples of popular models satisfying our standing assumptions. In Sect. 3, we briefly recall the asymptotic expansion procedure proposed by [33]. In Sect. 4, we derive error estimates for prices and sensitivities, first under the strong assumption of uniform parabolicity (Sect. 4.1) and then in the general case (Sect. 4.2). In Sect. 5, we prove our main result (Theorem 5.1) on the error estimates of the IV and its derivatives, and the consequent parabolic Taylor formula. Finally, the Appendix contains the proof of Theorem 4.4 and other auxiliary results, namely some short-time/small-volatility asymptotic estimates for the Black–Scholes sensitivities (Appendix C), an explicit representation formula for the terms appearing in the proxy $\bar{\sigma}_N$ (Appendix D), and a multivariate version of Faà di Bruno's formula (Appendix E).

2 Local diffusions and local transition densities

In this section, we describe the general setting and state the standing assumptions under which the main results of the paper are carried out. We also show some examples and prove some conditions under which such assumptions are satisfied. Generally we adopt definitions and notations from [18, 19].

We fix a time horizon $T_0 > 0$ and consider a continuous $\mathbb{R}^d$-valued Markov process $Z = (Z_t)_{t\in[0,T_0]}$ with transition probability function $\bar{p} = \bar{p}(t, z; T, d\zeta)$, defined on the space $(\Omega, \mathcal{F}, (\mathcal{F}_T^t)_{0\le t\le T\le T_0}, (P_{t,z})_{0\le t\le T_0})$. For any bounded Borel-measurable function φ, we denote by

$$
\begin{aligned}
E_{t,z}[\varphi(Z_T)] &:= (\mathbf{T}_{t,T}\varphi)(z) \\
&:= \int_{\mathbb{R}^d} \bar{p}(t, z; T, d\zeta)\varphi(\zeta), \quad 0 \le t < T \le T_0, z \in \mathbb{R}^d,
\end{aligned}
\tag{2.1}
$$

the $P_{t,z}$-expectation and the semigroup associated with the transition probability function $\bar{p}$, respectively (cf. [18, Chap. 2.1]).

We assume that $Z = (S, Y)$, where S is a nonnegative martingale[1] and Y takes values in $\mathbb{R}^{d-1}$; here S represents the risk-neutral price of a financial asset and Y models a number of stochastic factors in the market. For simplicity, we assume zero interest rates and no dividends.[2]

Throughout the paper, we assume the existence of a domain[3] $D \subseteq \mathbb{R}_{++} \times \mathbb{R}^{d-1}$ on which the following three standing assumptions hold. We emphasize that in the following assumptions, we impose only *local conditions,* satisfied by all the most popular financial models.

Assumption 2.1 The process Z is a *local diffusion on* D, meaning that for any $t \in [0, T_0)$, $\delta > 0$, $1 \le i, j \le d$ and H, compact subset of D, there exist the limits

$$
\lim_{h\to 0+} \int_{\{|z-\zeta|>\delta\}\cap H} \frac{\bar{p}(t, z; t+h, d\zeta)}{h} = \lim_{h\to 0+} \int_{\{|z-\zeta|>\delta\}\cap H} \frac{\bar{p}(t-h, z; t, d\zeta)}{h} = 0, \tag{2.2}
$$

uniformly with respect to $z \in \mathbb{R}_+ \times \mathbb{R}^{d-1}$, and the limits

$$
\lim_{h\to 0+} \int_{\{|z-\zeta|>\delta\}} \frac{\bar{p}(t, z; t+h, d\zeta)}{h} = \lim_{h\to 0+} \int_{\{|z-\zeta|>\delta\}} \frac{\bar{p}(t-h, z; t, d\zeta)}{h} = 0, \tag{2.3}
$$

[1] We assume that S is a martingale in order to ensure that the financial model is well posed; however, this assumption will not be used in the proof of our main results.

[2] The case of deterministic interest rates and/or dividends can be easily included by performing the analysis on the forward prices.

[3] Connected and open set.

$$\lim_{h\to 0+} \int\limits_{\{|z-\zeta|<\delta\}} (\zeta_i - z_i) \frac{\bar{p}(t,z;t+h,d\zeta)}{h}$$

$$= \lim_{h\to 0+} \int\limits_{\{|z-\zeta|<\delta\}} (\zeta_i - z_i) \frac{\bar{p}(t-h,z;t,d\zeta)}{h} =: \bar{a}_i(t,z), \tag{2.4}$$

$$\lim_{h\to 0+} \int\limits_{\{|z-\zeta|<\delta\}} (\zeta_i - z_i)(\zeta_j - z_j) \frac{\bar{p}(t,z;t+h,d\zeta)}{h}$$

$$= \lim_{h\to 0+} \int\limits_{\{|z-\zeta|<\delta\}} (\zeta_i - z_i)(\zeta_j - z_j) \frac{\bar{p}(t-h,z;t,d\zeta)}{h} =: \bar{a}_{ij}(t,z), \tag{2.5}$$

uniformly with respect to $z \in H$.

The following lemma, whose proof is deferred to Sect. 2.3, collects some useful consequences of Assumption 2.1.

Lemma 2.2 *Under Assumption 2.1, for any $\varphi \in C_0([0,T_0] \times D)$ and for any $f \in C_0^2([0,T_0] \times D)$, we have*

$$\lim_{T-t\to 0+} \|\mathbf{T}_{t,T}\varphi(T,\cdot) - \varphi(t,\cdot)\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^{d-1})} = 0, \tag{2.6}$$

$$\lim_{T-t\to 0+} \left\| \frac{\mathbf{T}_{t,T} f(T,\cdot) - f(t,\cdot)}{T-t} - (\partial_t + \bar{\mathcal{A}}_t) f(t,\cdot) \right\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^{d-1})} = 0, \tag{2.7}$$

where

$$\bar{\mathcal{A}}_t := \frac{1}{2} \sum_{i,j=1}^{d} \bar{a}_{ij}(t,z) \partial_{z_i z_j} + \sum_{i=1}^{d} \bar{a}_i(t,z) \partial_{z_i}, \qquad t \in [0,T_0), z \in D. \tag{2.8}$$

Moreover, for any $0 \le t < T < T_0$ and $z \in \mathbb{R}_+ \times \mathbb{R}^{d-1}$, we have

$$\frac{d}{dT} \left(\mathbf{T}_{t,T} f(T,\cdot) \right)(z) = \mathbf{T}_{t,T} \left((\partial_T + \bar{\mathcal{A}}_T) f(T,\cdot) \right)(z). \tag{2.9}$$

Many financial models are defined in terms of (stopped) solutions of stochastic differential equations. We refer to Sect. 2.2 in [18] for the definition and basic results about $(\mathcal{F}^t)$-stopping times with respect to a given Markov process. The following result shows that stopped solutions of SDEs satisfy Assumption 2.1.

Lemma 2.3 *Let $(Z_t)_{t\in[0,T_0]}$ be a continuous Markov process defined as $Z_t = \hat{Z}_{t\wedge\tau}$, where*

(i) *$\hat{Z}$ is a solution of the SDE*

$$d\hat{Z}_t = \mu(t,\hat{Z}_t)dt + \sigma(t,\hat{Z}_t)dW_t,$$

where W is a multidimensional Brownian motion and the coefficients of the SDE are continuous and bounded on $[0, T_0] \times D$, with D a domain of $\mathbb{R}^d$;

(ii) *τ is the first exit time of $\hat{Z}$ from a domain $D' \subseteq \mathbb{R}_+ \times \mathbb{R}^{d-1}$ containing D.*

Then Z is a local diffusion on D in the sense of Assumption 2.1, with

$$\bar{a}_i = \mu_i, \qquad \bar{a}_{ij} = (\sigma\sigma^*)_{ij}, \qquad 1 \leq i, j \leq d. \tag{2.10}$$

The proof of Lemma 2.3 is deferred to Sect. 2.3.

We refer to the operator $\bar{\mathcal{A}}_t$ in (2.8) as the *infinitesimal generator of Z on D*. In the second standing assumption, we require that $\bar{\mathcal{A}}_t$ be a nondegenerate operator. Notice that $\bar{\mathcal{A}}_t$ is defined only locally, on the domain D. In the following assumption and **throughout the paper**, $N \geq 2$ is a fixed integer.[4]

Assumption 2.4 The operator $\bar{\mathcal{A}}_t$ satisfies the following conditions:

(i) the coefficients $\bar{a}_{ij}$, $\bar{a}_i$ are in $C_P^{N,1}([0, T_0) \times D)$, where $C_P^{N,\alpha}$ denotes the usual parabolic Hölder space (see for instance [19, Chap. 10.1]);

(ii) $\bar{\mathcal{A}}_t$ is elliptic on D, i.e., there exist $M > 0$ and $\varepsilon \in (0, 1)$ such that

$$\varepsilon M|\zeta|^2 \leq \sum_{i,j=1}^{d} \bar{a}_{ij}(t, z)\zeta_i\zeta_j \leq M|\zeta|^2, \qquad t \in [0, T_0), z \in D, \zeta \in \mathbb{R}^d.$$

Finally, we state the third standing assumption.

Assumption 2.5 Z is a *Feller process on D*, i.e., for any $T \in (0, T_0)$ and $\varphi \in C_0(\mathbb{R}^d)$, the function $(t, z) \mapsto (\mathbf{T}_{t,T}\varphi)(z)$ is continuous on $[0, T) \times D$.

The following result summarizes some properties of the law of Z. In particular, it states the existence of a *local transition density* for Z on D, which is a nonnegative measurable function $\bar{\Gamma} = \bar{\Gamma}(t, z; T, \zeta)$, defined for $0 \leq t < T < T_0$ and $z, \zeta \in D$, such that for any $H \in \mathcal{B}(D)$ (Borel subset of D),

$$\bar{p}(t, z; T, H) = \int_H \bar{\Gamma}(t, z; T, \zeta)d\zeta.$$

Moreover, it provides a double characterization of such a local density, first as a solution to a *forward* Kolmogorov equation (with respect to the *end point* (T, ζ)) and then as a solution to a *backward* Kolmogorov equation (with respect to the *initial point* (t, z)). The existence and the *forward* representation follow from Hörmander's theorem [25], after proving that the law is a local solution, in the distributional sense, of the adjoint of the infinitesimal generator of Z. This result is rather classical and is coherent with the well-known results in [29] (see also the more recent paper [11]). In order to prove the *backward* formulation, we still employ Hörmander's theorem,

[4]To simplify the presentation, we assume $N \geq 2$. However, the proofs of neither the results in dimension one (i.e., $d = 1$), nor the results for the derivatives of order one or two in a generic dimension, do require this condition.

but in this case the proof is more delicate and technically involved. In fact, to prove that the law is a distributional solution of the generator of Z, it will be crucial to use the Feller property combined with the classical pointwise estimates [36] for weak solutions of parabolic PDEs. At this level of generality, the resulting *backward* representation for the transition local density appears to be novel and of independent interest.

Theorem 2.6 *Let Assumptions 2.1 and 2.4 be in force. Then Z has a local transition density $\bar{\Gamma}$ on D such that for any $(t,z) \in [0,T_0) \times D$, $\bar{\Gamma}(t,z;\cdot,\cdot)$ is in $C_P^{N,1}((t,T_0)\times D)$ and solves the forward Kolmogorov equation*

$$(\partial_T - \bar{\mathcal{A}}_T^*) f = 0 \qquad \text{on } (t,T_0)\times D. \tag{2.11}$$

Here $\bar{\mathcal{A}}_T^*$ denotes the formal adjoint of $\bar{\mathcal{A}}_T$, acting as

$$\bar{\mathcal{A}}_T^* f = \frac{1}{2}\sum_{i,j=1}^{d} \partial_{z_i z_j}\big(\bar{a}_{ij}(T,\cdot) f\big) - \sum_{i=1}^{d} \partial_{z_i}\big(\bar{a}_i(T,\cdot) f\big).$$

If in addition also Assumption 2.5 is satisfied, then

$$\bar{\Gamma}(\cdot,\cdot;T,\zeta) \text{ is in } C_P^{N+2,1}\big([0,T)\times D\big)$$

for any $(T,\zeta) \in (0,T_0)\times D$, and solves the backward Kolmogorov equation

$$(\partial_t + \bar{\mathcal{A}}_t) f = 0 \qquad \text{on } [0,T)\times D. \tag{2.12}$$

We give a detailed proof of Theorem 2.6 in Sect. 2.3. Before, in Sects. 2.1 and 2.2, we provide illustrative examples of popular models that satisfy Assumptions 2.1, 2.4 and 2.5, and to which our analysis applies. Only in order to deal with the derivatives of a call option price with respect to the strike, in Sect. 4.2, we introduce additional assumptions to ensure existence and local boundedness of such derivatives.

2.1 The CEV model

Consider the SDE

$$d\hat{S}_t = \sigma \hat{S}_t^{\beta} dW_t, \tag{2.13}$$

where $\sigma > 0$ and $0 < \beta < 1$. A solution to (2.13) can be represented, through the transformation $X_t = \frac{\hat{S}_t^{2(1-\beta)}}{\sigma^2(1-\beta)^2}$, in terms of the squared Bessel process

$$dX_t = \delta dt + 2\sqrt{X_t} dW_t,$$

with $\delta = \frac{1-2\beta}{1-\beta}$. The process $\hat{S}$ has different properties according to the parameter regimes $\beta < \frac{1}{2}$ and $\beta \geq \frac{1}{2}$. To describe these, we first introduce the functions

$$\bar{\Gamma}_{\pm}(t,s;T,S) = \frac{s^{\frac{1}{2}-2\beta}\sqrt{S}e^{-\frac{s^{2(1-\beta)}+S^{2(1-\beta)}}{2(1-\beta)^2\sigma^2(T-t)}}}{(1-\beta)\sigma^2(T-t)} I_{\pm\frac{1}{2(1-\beta)}}\left(\frac{(sS)^{1-\beta}}{(1-\beta)^2\sigma^2(T-t)}\right), \tag{2.14}$$

where $I_\nu(x)$ is the modified Bessel function of the first kind defined by

$$I_\nu(x) = \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty} \frac{x^{2k}}{2^{2k} k! \Gamma_E(\nu + k + 1)},$$

and Γ_E represents the Euler gamma function. Both $\bar{\Gamma}_+$ and $\bar{\Gamma}_-$ are fundamental solutions of $\partial_t + \bar{\mathcal{A}}$, where $\bar{\mathcal{A}}$ is the infinitesimal generator of $\hat{S}$, i.e.,

$$\bar{\mathcal{A}} = \frac{\sigma^2 s^{2\beta}}{2} \partial_{ss}. \tag{2.15}$$

Precisely, we have

$$(\partial_t + \bar{\mathcal{A}})\bar{\Gamma}_\pm(\cdot, \cdot; T, S) = 0 \qquad \text{on } [0, T) \times \mathbb{R}_{++}$$

and

$$\lim_{\substack{(t,s)\to(T,\bar{s}) \\ t<T}} \int_{\mathbb{R}_{++}} \bar{\Gamma}_\pm(t, s; T, S)\varphi(S)dS = \varphi(\bar{s}), \qquad \bar{s} \in \mathbb{R}_{++},$$

for any continuous and bounded function φ.

The point 0 is an attainable state for $\hat{S}$. In particular, if $\beta \geq \frac{1}{2}$, then 0 is absorbing: if we denote by $\tau_s := \inf\{\tau : \hat{S}_\tau = 0\}$ the first time $\hat{S}$ hits 0 starting from $\hat{S}_0 = s \geq 0$, then we have $\hat{S}_t = 0$ for $t \geq \tau_s$. The law of $\hat{S}$ has a Dirac delta component at the origin, and the function $\bar{\Gamma}_+$ in (2.14) is the transition *semi-density* of $\hat{S}$ on $\mathbb{R}_{++}$; more precisely, denoting by $\hat{p}$ the transition probability function of $\hat{S}$, we have

$$\hat{p}(t, s; T, H) = \int_H \bar{\Gamma}_+(t, s; T, S)dS$$

for any Borel subset H of $\mathbb{R}_{++}$ and

$$\int_0^{+\infty} \bar{\Gamma}_+(t, s; T, S)dS < 1.$$

On the other hand, if $\beta < \frac{1}{2}$, then $\hat{S}$ reaches 0 but it is reflected; in this case $\bar{\Gamma}_-$, which integrates to one on $\mathbb{R}_{++}$, is the transition density of $\hat{S}$. Moreover, $\hat{S}$ is a strict local martingale (cf. [13] or [24]) that "cannot" represent the risk-neutral price of an asset; the intuitive idea is that arbitrage opportunities would arise investing in an asset whose price is zero at the stopping time τ_s, but later becomes positive.

For this reason, in the CEV model introduced by [10], the asset price is defined as the process obtained by stopping the solution $\hat{S}$, starting from $\hat{S}_0 = s$, of the SDE (2.13) at τ_s, that is,

$$S_t := \hat{S}_{t \wedge \tau_s}, \qquad t \geq 0.$$

For any $0 < \beta < 1$, the transition semi-density of S is $\bar{\Gamma}_+$ in (2.14). For this model, the authors of [13] show that for any $0 < \beta < 1$, the process is a nonnegative martingale.

Table 1 ATM IV time-derivative

β	Numerical approx.	Taylor expansion	Durrleman
0.1	0.0337524	0.03375	−1.0125
0.2	0.0266639	0.0266667	−0.8
0.3	0.0204115	0.0204167	−0.6125
0.4	0.0149955	0.015	−0.45
0.5	0.0104115	0.0104167	−0.3125
0.6	0.00666029	0.00666667	−0.2
0.7	0.00374753	0.00375	−0.1125
0.8	0.00136839	0.00166667	−0.05
0.9	0.000415421	0.000416667	−0.0125

Now let D be any domain compactly contained in $\mathbb{R}_{++}$. By Lemma 2.3, the stopped process S is a local diffusion on D and satisfies Assumption 2.1. The infinitesimal generator $\bar{\mathcal{A}}$ is the operator in (2.15), has smooth coefficients, and is uniformly elliptic on D; thus Assumption 2.4 is satisfied for any $N \in \mathbb{N}$. Moreover, the Feller property on D (Assumption 2.5) follows from the explicit expression of the transition semi-density or from the general results in [16, Chap. 8] (see Problem 3 and Theorem 2.1).

The CEV model (and also its stochastic volatility counterpart, the popular SABR model used in interest rate modeling) is an interesting example of a degenerate model because the infinitesimal generator is *not globally uniformly elliptic and the law of the price process is not absolutely continuous with respect to Lebesgue measure*.

Remark 2.7 Durrleman [15, Sect. 5], provided formulas for the implied volatility in a local volatility (LV) model with LV function $\sigma = \sigma(s)$. His expression for the time-derivative of the ATM implied volatility, denoted by Σ, is equal to

$$\partial_t \Sigma(t,s)|_{t=0} = \frac{1}{12} s^2 \sigma(s)^2 \sigma''(s) - \frac{4}{3} s^2 \sigma(s) \sigma'(s)^2 + \frac{1}{12} s \sigma(s)^2 \sigma'(s).$$

The latter is slightly different from the expression we get from our Taylor expansion that, in this particular case, can be computed as in Sect. 3.2 and reads as

$$\partial_t \Sigma(t,s)|_{t=0} = \frac{1}{12} s^2 \sigma(s)^2 \sigma''(s) - \frac{1}{24} s^2 \sigma(s) \sigma'(s)^2 + \frac{1}{12} s \sigma(s)^2 \sigma'(s). \tag{2.16}$$

Actually, simple numerical tests performed in the CEV model confirm that formula (2.16) is correct. As a matter of example, in Table 1 we show the values of $\partial_t \Sigma(t,1)|_{t=0}$ in the CEV model with $\sigma = S_0 = 1$ (cf. (2.13)) and $\beta = 0.1, \dots, 0.9$.

2.2 Multifactor local-stochastic volatility models

We consider a pricing model defined as the solution of a system of SDEs of the form

$$\begin{cases} dS_t = \eta_1(t, S_t, Y_t) S_t \, dW_t^{(1)}, \\ S_0 = s \in \mathbb{R}_{++}, \end{cases} \tag{2.17}$$

$$\begin{cases} dY_t^{(i)} = \mu_i(t, S_t, Y_t)dt + \eta_i(t, S_t, Y_t)dW_t^{(i)}, & i = 2, \dots, d, \\ Y_0 = y \in \mathbb{R}^{d-1}, \end{cases} \tag{2.18}$$

where W is a d-dimensional correlated Brownian motion with

$$d\langle W^{(i)}, W^{(j)} \rangle_t = \rho_{ij}(t, S_t, Y_t)dt, \qquad i, j = 1, \dots, d.$$

In the most classical setting, one assumes that the coefficients of the SDEs are measurable functions, locally Lipschitz-continuous in the spatial variables (s, y) uniformly with respect to $t \in [0, T_0]$, and have sublinear growth in (s, y); for more details, we refer, for instance, to condition (A′) of Chap. 5.3 in [18]. In this case, a unique global-in-time solution (S, Y) exists, which is a Feller process[5] and a diffusion (see [18, Theorems 5.3.4 and 5.4.2]).

Usually, however, the above conditions are considered too restrictive and of limited practical use. Actually, we shall see that Assumptions 2.1, 2.4 and 2.5 are satisfied under much weaker conditions. To see this, we first note that the infinitesimal generator $\bar{\mathcal{A}}$ of (S, Y) is the operator of the form (2.8) with coefficients given by

$$\bar{a}_1 = 0, \quad \bar{a}_i = \mu_i, \quad \bar{a}_{11} = \rho_{11}\eta_1^2 s^2, \quad \bar{a}_{1i} = \bar{a}_{i1} = \rho_{1i}\eta_i\eta_1 s, \quad \bar{a}_{ij} = \bar{a}_{ji} = \rho_{ij}\eta_i\eta_j$$

for any $i, j = 2, \dots, d$. Now, Assumption 2.4 is straightforward to verify and applies to the great majority of the models used in finance, and thus, by Lemma 2.3, Assumption 2.1 is also satisfied provided that a solution to the system (2.17), (2.18) exists. The Feller property in Assumption 2.5 has to be verified case by case. Results ensuring the Feller property for the solution of an SDE under weak regularity conditions on the coefficients (Hölder- or local Lipschitz-continuity) have been recently proved in [48] (see Proposition 2.1) and by [47]. Moreover, the results of [16, Chap. 8] cover several SDEs related to financial models.

As a matter of example, we analyze the classical model proposed by [23]. Set $d = 2$ and

$$\begin{aligned} dS_t &= S_t\sqrt{Y_t}dW_t^{(1)}, & S_0 &\in \mathbb{R}_{++}, \\ dY_t &= \kappa(\theta - Y_t)dt + \delta\sqrt{Y_t}dW_t^{(2)}, & Y_0 &\in \mathbb{R}_{++}, \end{aligned}$$

where δ is a positive constant (the so-called vol-of-vol parameter), $\kappa, \theta > 0$ are the drift-mean and the mean-reverting term of the variance process, respectively, and W is a 2-dimensional Brownian motion with correlation $\rho \in (-1, 1)$. It is well known that the joint transition probability function $\bar{p}$ in (2.1) admits an explicit characterization in terms of its Fourier–Laplace transform. Precisely, setting $X_t = \log S_t$ and assuming for simplicity $\delta = 1$, we have

$$\hat{p}(t, x, y; T, \xi, \eta) := E_{t,x,y}[e^{\mathrm{i}\xi X_T - \eta Y_T}] = e^{\mathrm{i}x\xi - yA(T-t,\xi,\eta)}B(T-t, \xi, \eta), \tag{2.19}$$

[5]The definition of Feller process given in [18, Chap. 2.2] is slightly different from ours. However, the Feller property for solutions of SDEs is proved in [18] as a consequence of Lemma 5.3.3; this lemma also implies the Feller property as given in Assumption 2.5.

where

$$A(u,\xi,\eta)=\frac{b(\xi)g(\xi,\eta)e^{-D(\xi)(u-s)}-a(\xi)}{g(\xi,\eta)e^{-D(\xi)(u-s)}-1},$$

$$B(u,\xi,\eta)=e^{-\kappa\theta a(\xi)u}\left(\frac{g(\xi,\eta)-1}{g(\xi,\eta)e^{-D(\xi)u}-1}\right)^{2\kappa\theta}$$

with

$$g(\xi,\eta)=\frac{a(\xi)-\eta}{b(\xi)-\eta},\qquad a(\xi)=\mathrm{i}\xi\rho-\kappa+D(\xi),\qquad b(\xi)=\mathrm{i}\xi\rho-\kappa-D(\xi),$$

$$D(\xi)=\sqrt{(\mathrm{i}\xi\rho-\kappa)^2+\xi\,(\xi+\mathrm{i})}.$$

Using the explicit knowledge of the characteristic function of S, [1, Proposition 2.5] proves that S is a martingale and can reach neither ∞ nor 0 in finite time (see also [30] for related results in a more general setting). The variance process Y can reach the boundary with positive probability if the Feller condition $2\kappa\theta\geq\delta^2$ is violated, and in this case, the origin is a reflecting boundary. In any case, the distribution of Y_t has no mass at 0 for any positive t.

By Lemma 2.3, Assumption 2.1 is verified on any domain D compactly contained in $\mathbb{R}_{++}\times\mathbb{R}_{++}$, and the generator $\bar{\mathcal{A}}$ of (S,Y) reads as

$$\bar{\mathcal{A}}=\frac{ys^2}{2}\partial_{ss}+\frac{\delta^2 y}{2}\partial_{yy}+\rho\delta ys\,\partial_{sy}+\kappa(\theta-y)\partial_y,\qquad (s,y)\in\mathbb{R}_{++}\times\mathbb{R}_{+}.$$

It is also clear that Assumption 2.4 is satisfied on D for any $N\in\mathbb{N}$. Finally, the Feller property follows by the explicit expression of the characteristic function in (2.19), and thus Assumption 2.5 is also satisfied.

Remark 2.8 By Theorem 2.6, the couple (S,Y) in the Heston model has a smooth local transition density on any domain D compactly contained in $\mathbb{R}_{++}\times\mathbb{R}_{++}$. Therefore, since $p(t,z;T,\mathbb{R}^2\setminus(\mathbb{R}_{++}\times\mathbb{R}_{++}))=0$, the process (S,Y) has a transition density on $\mathbb{R}^2$, which is smooth on $\mathbb{R}_{++}\times\mathbb{R}_{++}$. In particular, the marginal distribution of S_t has a smooth density on $\mathbb{R}_{++}$, which is consistent with [12].

2.3 Proofs of Lemmas 2.2, 2.3 and Theorem 2.6

Proof of Lemma 2.2 We first remark that in the statement of the lemma, the short notation (see (2.6))

$$\lim_{T-t\to 0+}\|\mathbf{T}_{t,T}\varphi(T,\cdot)-\varphi(t,\cdot)\|_\infty=0$$

must be interpreted as

$$\lim_{h\to 0+}\|\mathbf{T}_{t,t+h}\varphi(t+h,\cdot)-\varphi(t,\cdot)\|_\infty=\lim_{h\to 0+}\|\mathbf{T}_{t-h,t}\varphi(t,\cdot)-\varphi(t-h,\cdot)\|_\infty=0,$$

and analogously for (2.7). Hereafter, for greater convenience, we use this abbreviation systematically. Now let us prove (2.6). For a given $\varphi \in C_0([0, T_0] \times D)$, we denote by H_φ the support of φ and consider a compact subset H of D such that $H_\varphi \subseteq [0, T_0] \times H$ and $\bar{\delta} := \mathrm{dist}(H_\varphi, [0, T_0] \times (\mathbb{R}^d \setminus H)) > 0$. Then we have

$$\mathbf{T}_{t,T}\varphi(T, z) - \varphi(t, z) = I_{t,T,1}(z) + I_{t,T,2}(z) + I_{t,T,3}(z),$$

where

$$I_{t,T,1}(z) = \int_H \bar{p}(t, z; T, d\zeta)\big(\varphi(T, \zeta) - \varphi(T, z)\big),$$

$$I_{t,T,2}(z) = \big(\varphi(T, z) - \varphi(t, z)\big) \int_H \bar{p}(t, z; T, d\zeta),$$

$$I_{t,T,3}(z) = -\varphi(t, z) \int_{(\mathbb{R}_+ \times \mathbb{R}^{d-1}) \setminus H} \bar{p}(t, z; T, d\zeta).$$

Since φ is uniformly continuous, for any $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that

$$|I_{t,T,1}(z)| \le \varepsilon \int_{\{|z-\zeta| \le \delta_\varepsilon\}} \bar{p}(t, z; T, d\zeta) + 2\|\varphi\|_\infty \int_{H \cap \{|z-\zeta| > \delta_\varepsilon\}} \bar{p}(t, z; T, d\zeta)$$

and therefore, by (2.2),

$$\limsup_{T-t \to 0+} |I_{t,T,1}(z)| \le \varepsilon$$

uniformly with respect to $z \in \mathbb{R}_+ \times \mathbb{R}^{d-1}$. Moreover, we have

$$|I_{t,T,2}(z)| \le |\varphi(T, z) - \varphi(t, z)| \longrightarrow 0$$

as $T - t \to 0+$, uniformly with respect to z. On the other hand, by (2.3), we have

$$|I_{t,T,3}(z)| \le \|\varphi\|_\infty \int_{\{|z-\zeta| > \bar{\delta}\}} \bar{p}(t, z; T, d\zeta) \longrightarrow 0$$

as $T - t \to 0+$, uniformly with respect to $z \in H_\varphi$, and $I_{t,T,3}(z) \equiv 0$ if $z \notin H_\varphi$. This concludes the proof of (2.6). Notice that for any $z \in D$ and $r > 0$ such that $B(z, r) := \{\zeta : |z - \zeta| < r\} \subseteq D$, we have

$$\lim_{T-t \to 0+} \int_{B(z,r)} \bar{p}(t, z; T, d\zeta) = 1; \tag{2.20}$$

indeed, for any $\varphi \in C_0(B(z, r))$ such that $|\varphi| \le 1$ and $\varphi(z) = 1$, by (2.6) we have

$$1 \ge \int_{B(z,r)} \bar{p}(t, z; T, d\zeta) \ge \mathbf{T}_{t,T}\varphi(z) \longrightarrow \varphi(z) = 1$$

as $T - t \to 0+$.

The proof of (2.7) is similar: for any $f \in C_0^2([0,T_0]\times D)$, we have

$$\frac{\mathbf{T}_{t,T} f(T,z) - f(t,z)}{T-t} = I_{t,T,1}(z) + I_{t,T,2}(z),$$

where

$$\begin{aligned} I_{t,T,1}(z) &= \int_H \bar{p}(t,z;T,d\zeta)\frac{f(T,\zeta)-f(t,z)}{T-t}, \\ I_{t,T,2}(z) &= \frac{f(t,z)}{T-t}\int_{(\mathbb{R}_+\times\mathbb{R}^{d-1})\setminus H} \bar{p}(t,z;T,d\zeta), \end{aligned} \tag{2.21}$$

with H defined analogously to how it was defined in the proof of (2.6). Again, by (2.3) the term $I_{t,T,2}(z)$ is negligible in the limit. As for $I_{t,T,1}(z)$, it suffices to plug the Taylor formula

$$\begin{aligned} f(T,\zeta) - f(t,z) = (T-t)\partial_t f(t,z) &+ \sum_{i=1}^d (\zeta_i - z_i)\partial_{z_i} f(t,z) \\ &+ \frac{1}{2}\sum_{i,j=1}^d (\zeta_i - z_i)(\zeta_j - z_j)\partial_{z_i z_j} f(t,z) \\ &+ \mathrm{o}(|T-t|) + \mathrm{o}(|z-\zeta|^2) \end{aligned}$$

into (2.21) and pass to the limit using (2.20), (2.4) and (2.5). This proves (2.7).

Finally, we have

$$\begin{aligned} &\left\| \frac{\mathbf{T}_{t,T+h} f(T+h,\cdot) - \mathbf{T}_{t,T} f(T,\cdot)}{h} - \mathbf{T}_{t,T}\big((\partial_T + \bar{\mathcal{A}}_T) f(T,\cdot)\big) \right\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^{d-1})} \\ &= \left\| \mathbf{T}_{t,T}\left(\frac{\mathbf{T}_{T,T+h} f(T+h,\cdot) - f(T,\cdot)}{h} - (\partial_T + \bar{\mathcal{A}}_T) f(T,\cdot)\right) \right\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^{d-1})} \\ &\le \left\| \frac{\mathbf{T}_{T,T+h} f(T+h,\cdot) - f(T,\cdot)}{h} - (\partial_T + \bar{\mathcal{A}}_T) f(T,\cdot) \right\|_{L^\infty(\mathbb{R}_+\times\mathbb{R}^{d-1})} \longrightarrow 0 \end{aligned}$$

as $h \to 0+$, where the last limit follows from (2.7). This proves the existence of the right derivative. For the left derivative, it suffices to use the identity

$$\begin{aligned} &\frac{\mathbf{T}_{t,T-h} f(T-h,\cdot) - \mathbf{T}_{t,T} f(T,\cdot)}{-h} - \mathbf{T}_{t,T}\left((\partial_T + \bar{\mathcal{A}}_T) f(T,\cdot)\right) \\ &= \mathbf{T}_{t,T-h}\left(\frac{\mathbf{T}_{T-h,T} - I}{h} - (\partial_T + \bar{\mathcal{A}}_T)\right) f(T,\cdot) \\ &\quad + (\mathbf{T}_{t,T-h} - \mathbf{T}_{t,T})\big((\partial_T + \bar{\mathcal{A}}_T) f(T,\cdot)\big), \end{aligned}$$

where I is the identity operator. This concludes the proof. □

Proof of Lemma 2.3 *Step 1.* We prove (2.2). Fix $\delta > 0$ and H, a compact subset of D. Consider a family of functions $(\varphi_z)_{z\in\mathbb{R}^d}$ such that $\varphi_z(z) = 0$, $\varphi_z(\zeta) \equiv 1$ for $\zeta \in H \cap \{|\zeta - z| > \delta\}$ and $\varphi_z \in C_0^\infty(D)$ with all derivatives bounded by a constant C_1 which depends on D, H and δ, but not on z. By the Itô formula, we have

$$\varphi_z(\hat{Z}_T) = \varphi_z(\hat{Z}_t) + \int_t^T \bar{\mathcal{A}}_s \varphi_z(\hat{Z}_s)ds + \int_t^T \nabla\varphi_z(\hat{Z}_s)\sigma(s, \hat{Z}_s)dW_s \tag{2.22}$$

with $\bar{\mathcal{A}}_s$ as defined in (2.8) and $\bar{a}_i$, $\bar{a}_{ij}$ as in (2.10). Notice that

$$|\bar{\mathcal{A}}_s\varphi_z(\hat{Z}_s)| + |\nabla\varphi_z(\hat{Z}_s)\sigma(s, \hat{Z}_s)| \le C_2, \qquad s \in [0, T_0], z \in \mathbb{R}^d,$$

with C_2 dependent only on C_1 and the $L^\infty([0, T_0] \times D)$-norm of the coefficients of the SDE. Let $\bar{p}(t, z; T, d\zeta)$ denote the transition probability of the stopped process $Z_T = \hat{Z}_{T\wedge\tau}$. Then, by recalling the definition of τ and since $D \subseteq D'$ and φ_z has compact support in D, we have

$$\int_{\{|z-\zeta|>\delta\}\cap H} \bar{p}(t, z; T, d\zeta) \le E_{t,z}[\varphi_z^4(\hat{Z}_{T\wedge\tau})] \le E_{t,z}[\varphi_z^4(\hat{Z}_T)],$$

and (2.2) follows from (2.22), the Hölder inequality and Doob's maximal inequality (in the form of [18, Corollary 6.4] with $m = 2$). The proof of (2.3) is analogous and omitted.

Step 2. We prove (2.4). Fix $1 \le i \le d$ and H, a compact subset of D. We first remark that it is sufficient to prove the claim for $\delta < \bar{\delta} := \text{dist}(H, \partial D)$. Indeed, we have

$$\frac{1}{T-t}\int_{\{|z-\zeta|<\delta\}}(\zeta_i - z_i)\bar{p}(t, z; T, d\zeta) = \frac{1}{T-t}\int_{\{|z-\zeta|<\bar{\delta}\}}(\zeta_i - z_i)\bar{p}(t, z; T, d\zeta) + I_{t,T},$$

where by (2.3),

$$I_{t,T} = \frac{1}{T-t}\int_{\{\bar{\delta}\le|z-\zeta|<\delta\}}(\zeta_i - z_i)\bar{p}(t, z; T, d\zeta) \longrightarrow 0$$

as $T - t \to 0+$, uniformly with respect to $z \in H$.

Next, we consider a family of functions $(\varphi_z)_{z\in H}$ such that $\varphi_z(\zeta) = \zeta_i - z_i$ for $|\zeta - z| < \delta$ and $\varphi_z \in C_0^\infty(D)$ with all the derivatives bounded by a constant C_1 which depends on D, H and δ, but not on z. Note that

$$|\nabla\varphi_z(Z_s)\sigma(s, Z_s)| \le C_2, \qquad s \in [0, T_0], z \in H, \tag{2.23}$$

with C_2 dependent only on C_1 and the $L^\infty([0, T_0] \times D)$-norm of the coefficients of the SDE. Now, we set $\Psi_z(t, \cdot) = \bar{\mathcal{A}}_t\varphi_z$ and note that $\Psi_z(t, \zeta) = a_i(t, \zeta)$ for $|\zeta - z| < \delta$. Denoting again by $\bar{p}(t, z; T, d\zeta)$ the transition probability of the stopped process $(\hat{Z}_{T\wedge\tau})$, we have

$$\frac{1}{T-t}\int_{\{|z-\zeta|<\delta\}}(\zeta_i - z_i)\bar{p}(t, z; T, d\zeta) - \bar{a}_i(t, z) = I_{1,t,T,z} + I_{2,t,T,z},$$

where by (2.3),

$$I_{1,t,T,z} := -\frac{1}{T-t}\int_{\{|z-\zeta|\geq\delta\}} \bar{p}(t,z;T,d\zeta)\varphi_z(\zeta) \longrightarrow 0$$

as $T - t \to 0+$, uniformly in H, and

$$\begin{aligned}
I_{2,t,T,z} &:= E_{t,z}\left[\frac{\varphi_z(\hat{Z}_{T\wedge\tau})}{T-t} - \Psi_z(t,z)\right] \\
&= E_{t,z}\left[\frac{1}{T-t}\int_t^T \bar{\mathcal{A}}_s\varphi_z(\hat{Z}_{s\wedge\tau})ds - \Psi_z(t,z)\right] \\
&= E_{t,z}\left[\int_0^1 \Psi_z\big(t+\rho(T-t), \hat{Z}_{(t+\rho(T-t))\wedge\tau}\big)d\rho - \Psi_z(t,z)\right] \\
&= \int_0^1 \Big(\big(\mathbf{T}_{t,t+\rho(T-t)}\Psi_z\big(t+\rho(T-t),\cdot\big)\big)(z) - \Psi_z(t,z)\Big)d\rho.
\end{aligned}$$

Here, the second equality holds since by assumption $D \subseteq D'$ and φ_z has compact support in D, and we use (2.22) and the fact that by (2.23), the stochastic integral is a true martingale, while the fourth equality uses Fubini's theorem. Thus, by (2.6) and the fact that $\Psi_z(t,\cdot) \in C_0([0,T_0]\times D)$ by definition, we infer that $I_{2,t,T,z}$ converges to zero as $T-t \to 0+$, uniformly with respect to $z \in H$. We remark here explicitly that (2.6) in Lemma 2.2 is proved using (2.2) and (2.3) only, which in turn have already been proved for the stopped process in the previous step; therefore, no circular argument has been used. The proof of (2.5) is based on analogous arguments; thus we leave the details to the reader. □

Proof of Theorem 2.6 We fix $(t,z) \in [0,T_0)\times D$ and $f \in C_0^2([0,T_0)\times D)$ and show that the process

$$M_T^t := f(T,Z_T) - f(t,Z_t) - \int_t^T (\partial_u + \bar{\mathcal{A}}_u) f(u,Z_u)du, \qquad t \leq T < T_0, \tag{2.24}$$

is an $(\mathcal{F}^t)$-martingale. First observe that integrating (2.9), we get the identity

$$\big(\mathbf{T}_{t,T} f(T,\cdot)\big)(z) - f(t,z) = \int_t^T \mathbf{T}_{t,\tau}\big((\partial_\tau + \bar{\mathcal{A}}_\tau)f(\tau,\cdot)\big)(z)d\tau, \qquad T \in (t,T_0). \tag{2.25}$$

Note that the integrand in (2.25) is bounded as a function of τ because of Assumption 2.4 and since $f \in C_0^2([0,T_0)\times D)$ and $\mathbf{T}_{t,\tau}$ is a contraction. Now for $\tau \in [t,T]$, we have

$$\begin{aligned}
E_{t,z}[M_T^t|\mathcal{F}_\tau^t] - M_\tau^t &= E_{t,z}\left[f(T,Z_T) - f(\tau,Z_\tau) - \int_\tau^T (\partial_u + \bar{\mathcal{A}}_u)f(u,Z_u)du\,\middle|\,\mathcal{F}_\tau^t\right] \\
&= \Phi(\tau,Z_\tau),
\end{aligned}$$

where by the Markov property and Fubini's theorem,

$$\begin{aligned}\Phi(\tau,z) &= E_{\tau,z}\left[f(T,Z_T)-f(\tau,z)-\int_\tau^T(\partial_u+\bar{\mathcal{A}}_u)f(u,Z_u)du\right]\\ &= \big(\mathbf{T}_{\tau,T}f(T,\cdot)\big)(z)-f(\tau,z)-\int_\tau^T\mathbf{T}_{\tau,u}\big((\partial_u+\bar{\mathcal{A}}_u)f(u,\cdot)\big)(z)du\end{aligned}$$

which is 0 by (2.25).

Notice that $M_t^t=0$; thus for any $f\in C_0^2((t,T_0)\times D)$, we have

$$0=E_{t,z}[M_{T_0}^t]=\int_t^{T_0}\int_D\bar{p}(t,z;T,d\zeta)(\partial_T+\bar{\mathcal{A}}_T)f(T,\zeta)dT. \tag{2.26}$$

Since f is arbitrary, (2.26) means that $\bar{p}(t,z;\cdot,\cdot)$ satisfies (2.11) on $(t,T_0)\times D$ in the sense of distributions. If the coefficients of the generator are smooth functions, then from Hörmander's theorem (see for instance [42, Sect. V.38]), we infer that $\bar{p}(t,z;\cdot,\cdot)$ admits a local density $\bar{\Gamma}(t,z;\cdot,\cdot)$ which is a smooth function and solves the forward Kolmogorov PDE on $(t,T_0)\times D$. In the general case, it suffices to use a standard regularization argument by smoothing the coefficients and then applying Schauder's interior estimates (cf. [19, Chap. 10.1]); for this, we refer, for instance, to [28]. The first part of the statement then follows since z and r are arbitrary.

Next, we use the classical Moser pointwise estimates (see [36] and the more recent and general formulation in [40, Corollary 1.4]) to prove an L^∞_{loc}-estimate of $\bar{\Gamma}$ that is used in the second part of the proof. More precisely, let us fix $(t,z)\in[0,T_0)\times D$, $T\in(t,T_0)$ and a compact subset H of D, and set

$$r=\frac{1}{2}\min\left\{\sqrt{T_0-T},\sqrt{T-t},\mathrm{dist}(H,\partial D)\right\}.$$

Since $\bar{\Gamma}(t,z;\cdot,\cdot)$ solves the PDE $(\partial_T-\bar{\mathcal{A}}_T^*)\bar{\Gamma}(t,z;\cdot,\cdot)=0$ on $(t,T_0)\times D$, Moser's estimate gives that

$$\bar{\Gamma}(t,z;T,\zeta)\le\frac{c_0}{r^{d+2}}\int_{T-r^2}^{T+r^2}\int_{B(\zeta,r)}\bar{\Gamma}(t,z;\bar{T},\bar{\zeta})d\bar{\zeta}d\bar{T}\le 2c_0r^{-d},\quad \zeta\in H, \tag{2.27}$$

where *the constant c_0 depends only on the dimension d and the local ellipticity constant M of Assumption 2.4(ii)*. We notice explicitly that the constant c_0 in (2.27) is independent of $z\in D$ and $\zeta\in H$.

To prove the second part of Theorem 2.6, we adapt the argument of Theorem 2.7 in [26]. We fix $\varphi\in C_0(D)$, $T\in(0,T_0)$, $z_0\in D$ and $r>0$ such that the closure of the ball $B(z_0,r)$ is contained in D. Then we denote by f the smooth solution of

$$\begin{cases}(\partial_t+\bar{\mathcal{A}}_t)f=0 & \text{on } [0,T)\times B(z_0,r),\\ f(t,z)=(\mathbf{T}_{t,T}\varphi)(z) & \text{for } (t,z)\in\partial_P([0,T]\times B(z_0,r)),\end{cases} \tag{2.28}$$

where

$$\partial_P\big([0,T]\times B(z_0,r)\big):=\Big([0,T]\times\partial B(z_0,r)\cup\big(\{T\}\times B(z_0,r)\big)\Big)$$

is the parabolic boundary of the cylinder $[0, T] \times B(z_0, r)$. Such a solution exists because $\bar{\mathcal{A}}_t$ is uniformly elliptic on $[0, T_0) \times D$ and $(t, z) \mapsto (\mathbf{T}_{t,T}\varphi)(z)$ is continuous on $[0, T] \times D$ by the Feller property (cf. Assumption 2.5) and (2.6).

Now we fix $t \in [0, T)$ and denote by τ_0 the $(\mathcal{F}^t)$-stopping time defined as $\tau_0 = T \wedge \tau_1$, where τ_1 is the first exit time after t of Z from $B(z_0, r)$. By the $(\mathcal{F}^t)$-martingale property of the process M^t in (2.24), with f as in (2.28), and the optional sampling theorem, we have the stochastic representation

$$f(t, z) = E_{t,z}[(\mathbf{T}_{\tau_0,T}\varphi)(Z_{\tau_0})].$$

On the other hand, for $(t, z) \in [0, T) \times B(z_0, r)$, we have by the strong Markov property that

$$\begin{aligned}(\mathbf{T}_{t,T}\varphi)(z) &= E_{t,z}[\varphi(Z_T)] = E_{t,z}\left[E_{t,z}[\varphi(Z_T)|\mathcal{F}^t_{\tau_0}]\right] \\ &= E_{t,z}[(\mathbf{T}_{\tau_0,T}\varphi)(Z_{\tau_0})] = f(t, z), \qquad (2.29)\end{aligned}$$

and in particular $(t, z) \mapsto (\mathbf{T}_{t,T}\varphi)(z)$ solves the backward equation (2.12).

Finally, we consider a sequence $(\varphi_n)_{n\in\mathbb{N}}$ of functions in $C_0(D)$, approximating a Dirac delta $\delta_{\bar{z}}$ for a fixed $\bar{z} \in D$. We also fix a test function $\psi \in C_0^\infty((0, T) \times D)$ and integrate by parts to obtain

$$\begin{aligned}0 &= \int_0^T \int_D (\partial_t + \bar{\mathcal{A}}_t)(\mathbf{T}_{t,T}\varphi_n)(z)\psi(t, z)dtdz \\ &= \int_0^T \int_D (\mathbf{T}_{t,T}\varphi_n)(z)(-\partial_t + \bar{\mathcal{A}}_t^*)\psi(t, z)dtdz \\ &= \int_0^T \int_D \int_D \bar{\Gamma}(t, z; T, \zeta)\varphi_n(\zeta)d\zeta(-\partial_t + \bar{\mathcal{A}}_t^*)\psi(t, z)dtdz. \qquad (2.30)\end{aligned}$$

Note that $\zeta \mapsto \bar{\Gamma}(t, z; T, \zeta)$ is a continuous function for $t < T$, and therefore

$$\int_D \bar{\Gamma}(t, z; T, \zeta)\varphi_n(\zeta)d\zeta \longrightarrow \bar{\Gamma}(t, z; T, \bar{\zeta})$$

pointwise. On the other hand, the L^∞_{loc}-estimate (2.27) of $\bar{\Gamma}$ allows passing to the limit as $n \to \infty$ in (2.30), using the dominated convergence theorem, to get

$$\int_0^T \int_D \bar{\Gamma}(t, z; T, \bar{z})(-\partial_t + \bar{\mathcal{A}}_t^*)\psi(t, z)dtdz = 0.$$

This shows that $\bar{\Gamma}(\cdot, \cdot; T, \zeta)$ is a distributional solution of (2.12) on $[0, T) \times D$, and we conclude using again Hörmander's theorem. □

Remark 2.9 The same argument used to prove (2.29) applies also to

$$\varphi(s, y) = (s - K)^+,$$

and allows us to prove that the expectation $E_{t,s,v}[(S_T - K)^+]$ solves the backward equation (2.12) as a function of (t, s, v). Indeed, it suffices to use a standard localization technique and the fact that the call payoff $(S_T - K)^+$ is integrable because S is a martingale by assumption.

3 Analytical approximations of prices and implied volatilities

Here we briefly recall the construction proposed in [33] of an explicit approximating series for option prices, along with a consequent polynomial expansion for the related implied volatility. Such a construction relies on a singular perturbation technique that allows, in its most general form, carrying out closed-form expansions for the local transition density; this leads to an approximation of the solution to the related backward Cauchy problem with generic final datum φ. Such a technique has been recently fully described in [32] in the uniformly parabolic setting, and subsequently extended in [38] to the case of locally parabolic operators and in [34] to models with jumps. Moreover, a recent extension of this technique to *utility indifference pricing* was proposed by [31].

We consider a model $Z = (S, Y)$ that satisfies Assumptions 2.1, 2.4 and 2.5 in Sect. 2. We denote by $C_{t,T,K}$ the time-t no-arbitrage value of a European call option with positive strike K and maturity $T \le T_0$, defined as $C_{t,T,K} = v(t, S_t, Y_t; T, K)$, where

$$v(t, s, y; T, K) := E_{t,s,y}[(S_T - K)^+], \qquad (t, s, y) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}^{d-1}. \quad (3.1)$$

Clearly,[6] we have $v(t, 0, y; T, K) \equiv 0$ and therefore, to avoid trivial situations, we may assume a positive initial price, i.e., $s > 0$. As a consequence of Theorem 2.6 (see also Remark 2.9), for any positive K, the function v in (3.1) is such that

$$v(\cdot, \cdot; T, K) \in C_P^{N+2,1}\big((0, T) \times D\big) \cap C([0, T] \times D)$$

and solves the backward Kolmogorov equation (2.12), i.e.,

$$(\partial_t + \bar{\mathcal{A}}_t) v(\cdot, \cdot; T, K) = 0 \qquad \text{on } (0, T) \times D.$$

As will be shown in Sect. 3.2, in order to obtain an explicit expansion of the implied volatility, it is crucial to expand the call price around a Black–Scholes price. Since the perturbation technique that we employ naturally yields Gaussian approximations at the leading term, we work in logarithmic variables. Therefore, for any $T \in (0, T_0]$ and $k \in \mathbb{R}$, we set

$$u(t, x, y; T, k) = v(t, e^x, y; T, e^k), \qquad 0 \le t \le T, \ (x, y) \in \mathbb{R} \times \mathbb{R}^{d-1}, \quad (3.2)$$

where v is the pricing function in (3.1). Here, x and k are meant to represent the spot log-price of the underlying asset and the log-strike of the option, respectively. Note that the function u is well defined regardless of the process S hitting zero or not.

After switching to log-variables, the generator $\bar{\mathcal{A}}_t$ in (2.8) is transformed into the second order operator

$$\mathcal{A} := \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(t, z) \partial_{z_i z_j} + \sum_{i=1}^{d} a_i(t, z) \partial_{z_i} \quad (3.3)$$

[6] Simply note that $(S_T - K)^+ \le S_T$ and S is a martingale by assumption.

with $t \in [0, T_0]$, $z = (x, y) \in \mathbb{R} \times \mathbb{R}^{d-1}$,

$$a_{11}(t, x, y) = e^{-2x} \bar{a}_{11}(t, e^x, y), \qquad a_1(t, x, y) = -\frac{e^{-2x}}{2} \bar{a}_{11}(t, e^x, y)$$

and for $i, j = 2, \dots, d$,

$$a_{1i}(t, x, y) = e^{-x} \bar{a}_{1i}(t, e^x, y), \qquad a_{ij}(t, x, y) = \bar{a}_{ij}(t, e^x, y),$$
$$a_i(t, x, y) = \bar{a}_i(t, e^x, y).$$

For the reader's convenience, we also recall the classical definitions of the Black–Scholes price and the implied volatility in terms of the spot log-price and the log-strike.

Definition 3.1 We denote by u^{BS} the *Black–Scholes price* function defined as

$$u^{\mathrm{BS}}(\sigma; \tau, x, k) := e^x \mathcal{N}(d_+) - e^k \mathcal{N}(d_-), \qquad d_\pm := \frac{1}{\sigma \sqrt{\tau}} \left(x - k \pm \frac{\sigma^2 \tau}{2} \right)$$

for any $x, k \in \mathbb{R}$ and $s, \tau > 0$, where $\mathcal{N}$ is the CDF of a standard normal random variable.

Definition 3.2 The *implied volatility* $\sigma = \sigma(t, x, y; T, k)$ of the price $u(t, x, y; T, k)$ as in (3.2) is the unique positive solution of the equation

$$u^{\mathrm{BS}}(\sigma; T - t, x, k) = u(t, x, y; T, k).$$

Note that Definition 3.2 is well posed because $C_{t,T,K}$ is a no-arbitrage price and thus $u(t, x, y; T, k)$ belongs to the no-arbitrage interval $((e^x - e^k)^+, e^x)$.

The computations in the following two subsections are meant to be formal and not rigorous. They only serve the purpose to lead us through the definition of an approximating expansion for prices and implied volatilities. The well-posedness of such definitions will be clarified under rigorous assumptions in Sect. 4.

3.1 Price expansion

We fix $\bar{z} = (\bar{x}, \bar{y}) \in \mathbb{R} \times \mathbb{R}^{d-1}$ such that $(e^{\bar{x}}, \bar{y}) \in D$ with D as in Assumption 2.4, and expand the operator $\mathcal{A}_t$ by replacing the functions $a_{ij}(t, \cdot)$, $a_i(t, \cdot)$ with their Taylor series around $\bar{z}$. We formally obtain

$$\mathcal{A}_t = \sum_{n=0}^{\infty} \mathcal{A}_{t,n}^{(\bar{z})},$$

where

$$\mathcal{A}_{t,n}^{(\bar{z})} = \sum_{|\beta|=n} \left(\sum_{i,j=1}^{d} \frac{D^\beta a_{ij}(t, \bar{z})}{\beta!} (z - \bar{z})^\beta \partial_{z_i z_j} + \sum_{i=1}^{d} \frac{D^\beta a_i(t, \bar{z})}{\beta!} (z - \bar{z})^\beta \partial_{z_i} \right). \quad (3.4)$$

The intuitive idea underlying the following procedure is inspired by the fact that typically, the pricing function $u(\cdot,\cdot;T,k)$ solves the backward Cauchy problem

$$\begin{cases}(\partial_t+\mathcal{A}_t)u(\cdot,\cdot;T,k)=0 & \text{on } [0,T)\times\mathbb{R}\times\mathbb{R}^{d-1},\\ u(T,x,y;T,k)=(e^x-e^k)^+ & \text{for } (x,y)\in\mathbb{R}\times\mathbb{R}^{d-1}.\end{cases} \tag{3.5}$$

Actually, (3.5) holds automatically true if the operator $(\partial_t+\mathcal{A}_t)$ is uniformly parabolic and can be also proved to be satisfied, case by case, in many degenerate cases of interest in mathematical finance, such as the CEV model. Nevertheless, the validity of (3.5) is not necessary for our analysis and *it is not required as an assumption*.

Next we assume that the pricing function u can be expanded as

$$u=\sum_{n=0}^{\infty}u_n^{(\bar z)}. \tag{3.6}$$

Inserting (3.4) and (3.6) into (3.5), we find that the functions $(u_n(\cdot,\cdot;T,k))_{n\geq 0}$ satisfy the sequence of nested Cauchy problems

$$\begin{cases}(\partial_t+\mathcal{A}_{t,0})u_0^{(\bar z)}(\cdot,\cdot;T,k)=0 & \text{on } [0,T)\times\mathbb{R}^d,\\ u_0^{(\bar z)}(T,x,y;T,k)=(e^x-e^k)^+ & \text{for } (x,y)\in\mathbb{R}\times\mathbb{R}^{d-1}\end{cases} \tag{3.7}$$

and

$$\begin{cases}(\partial_t+\mathcal{A}_{t,0})u_n^{(\bar z)}(\cdot,\cdot;T,k)=-\sum_{h=1}^{n}\mathcal{A}_{t,h}^{(\bar z)}u_{n-h}^{(\bar z)}(\cdot,\cdot;T,k) & \text{on } [0,T)\times\mathbb{R}^d,\\ u_n^{(\bar z)}(T,z;T,k)=0 & \text{for } z\in\mathbb{R}^d.\end{cases} \tag{3.8}$$

Note that by Assumption 2.4, $\mathcal{A}_{t,0}$ is an elliptic operator with time-dependent coefficients and therefore problem (3.7) can be solved to obtain

$$u_0^{(\bar z)}(t,x,y;T,k)=u^{\mathrm{BS}}(\sigma_0^{(\bar z)};T-t,x,k), \tag{3.9}$$

$$\sigma_0^{(\bar z)}\equiv\sigma_0^{(\bar z)}(t,T)=\sqrt{\frac{1}{T-t}\int_t^T a_{11}(\tau,\bar z)d\tau} \tag{3.10}$$

for any $t\in[0,T]$ and $(x,y)\in\mathbb{R}\times\mathbb{R}^{d-1}$. As for the nth order correcting term $u_n^{(\bar z)}$, an explicit representation in terms of differential operators acting on $u_0^{(\bar z)}$ is available (see Theorem D.1).

Definition 3.3 For fixed maturity date T and log-strike k, we define the *Nth order approximations* of $u(\cdot,\cdot;T,k)$ as

$$\bar u_N(t,z;T,k)=\sum_{n=0}^{N}u_n^{(z)}(t,z;T,k),\qquad t\in[0,T],z\in\mathbb{R}\times\mathbb{R}^{d-1}, \tag{3.11}$$

where the functions $u_n^{(z)}$ are explicitly defined as in (3.9) and (D.1).

We recall that similar price expansions have been developed by [5, 46] using Malliavin calculus techniques and by [3] using heat kernel methods.

3.2 Implied volatility expansion

We briefly recall how to derive a formal polynomial IV expansion from the price expansion (3.6)–(3.8). To ease notation, we sometimes suppress the dependence on $(t, x, y; T, k)$. Consider the family of approximate call prices indexed by δ

$$u^{(\bar{z})}(\delta) = u^{\mathrm{BS}}(\sigma_0^{(\bar{z})}) + \sum_{n=1}^{N} \delta^n u_n^{(\bar{z})} + \delta^{N+1}\Big(u - \sum_{n=0}^{N} u_n^{(\bar{z})}\Big), \qquad \delta \in [0,1], \tag{3.12}$$

with $\sigma_0^{(\bar{z})}$ as in (3.10) and the functions $u_n^{(\bar{z})}$ as in Sect. 3.1. Note that setting $\delta = 1$ yields the true pricing function u. Defining

$$g(\delta) := (u^{\mathrm{BS}})^{-1}\big(u(\delta)\big), \qquad \delta \in [0,1], \tag{3.13}$$

we seek the implied volatility $\sigma = g(1)$. We show in Lemma 5.8 below that under suitable assumptions, $u(\delta) \in ((e^x - e^k)^+, e^x)$ for any $\delta \in [0,1]$. This guarantees that $g(\delta)$ in (3.13) is well defined. By expanding both sides of (3.13) as a Taylor series in δ, we see that σ admits an expansion of the form

$$\sigma = g(1) = \sigma_0 + \sum_{n=1}^{\infty} \sigma_n, \qquad \sigma_n = \frac{1}{n!}\partial_\delta^n g(\delta)|_{\delta=0}. \tag{3.14}$$

Note that by (3.12), we also have

$$u_n = \frac{1}{n!}\partial_\delta^n u^{\mathrm{BS}}\big(g(\delta)\big)\big|_{\delta=0}, \qquad 1 \le n \le N,$$

and by applying Faà di Bruno's formula (Proposition E.1), one can find the recursive representation

$$\sigma_n^{(\bar{z})} = \frac{u_n^{(\bar{z})}}{\partial_\sigma u^{\mathrm{BS}}(\sigma_0^{(\bar{z})})} - \frac{1}{n!}\sum_{h=2}^{n} \mathbf{B}_{n,h}\big(1!\,\sigma_1^{(\bar{z})}, 2!\,\sigma_2^{(\bar{z})}, \ldots, (n-h+1)!\,\sigma_{n-h+1}^{(\bar{z})}\big)\frac{\partial_\sigma^h u^{\mathrm{BS}}(\sigma_0^{(\bar{z})})}{\partial_\sigma u^{\mathrm{BS}}(\sigma_0^{(\bar{z})})} \tag{3.15}$$

for any $1 \le n \le N$, where $\mathbf{B}_{n,h}$ denote the so-called Bell polynomials. It was shown in [33] (see also Proposition D.3) that each term $\sigma_n^{(\bar{z})}$ is a polynomial in the log-moneyness $k - x$. Moreover, if the coefficients of the model are time-independent, the expansion turns out to be also polynomial in time.

Definition 3.4 For a call option with log-strike k and maturity T, we define the *Nth order approximation of the implied volatility* $\sigma(t, x, y; T, k)$ as

$$\bar{\sigma}_N(t, x, y; T, k) := \sum_{n=0}^{N} \sigma_n^{(x,y)}(t, x, y; T, k), \tag{3.16}$$

where $\sigma_n^{(x,y)}$ are as defined in (3.15).

We recall that similar implied volatility expansions have been developed by [4, 14, 17, 21], among others.

4 Error estimates for prices and sensitivities

In this section, we derive error estimates for prices and sensitivities. Let us introduce the following

Notation 4.1 For $z_0 = (x_0, y_0) \in \mathbb{R} \times \mathbb{R}^{d-1}$ and $0 < r \le +\infty$, we set

$$D(z_0, r) = B(x_0, r) \times B(y_0, r)$$

with $B(x_0, r) = \{x \in \mathbb{R} : |x - x_0| < r\}$ and $B(y_0, r) = \{y \in \mathbb{R}^{d-1} : |y - y_0| < r\}$. Moreover, for $T \in (0, T_0)$, we consider the cylinders $H(T, z_0, r)$, $\bar{H}(T, z_0, r)$ and the lateral boundary $\Sigma(T, z_0, r)$ defined by

$$H(T, z_0, r) := (0, T) \times D(z_0, r), \qquad \bar{H}(T, z_0, r) := [0, T) \times D(z_0, r),$$
$$\Sigma(T, z_0, r) := [0, T) \times \partial D(z_0, r),$$

respectively.

Since we work with logarithmic variables, we restate Assumption 2.4 in terms of conditions on the operator $\mathcal{A}_t$ as defined in (3.3). We recall that $N \ge 2$ is an integer constant that is fixed **throughout the paper**.

Assumption 4.2 There exist $M_0 > 0$, $0 < r \le +\infty$ and $z_0 = (x_0, y_0) \in \mathbb{R} \times \mathbb{R}^{d-1}$ such that the operator $\mathcal{A}_t$ as in (3.3) coincides with $\widetilde{\mathcal{A}}_t$ on $\bar{H}(T_0, z_0, r)$, where $\widetilde{\mathcal{A}}_t$ is a differential operator of the form

$$\widetilde{\mathcal{A}}_t = \frac{1}{2} \sum_{i,j=1}^{d} \widetilde{a}_{ij}(t, z)\partial_{z_i z_j} + \sum_{i=1}^{d} \widetilde{a}_i(t, z)\partial_{z_i}, \qquad t \in [0, T_0), \ z \in \mathbb{R}^d,$$

such that for some $M \in (0, M_0]$ and $\varepsilon \in (0, 1)$, we have:

(i) *Regularity and boundedness:* the coefficients $\widetilde{a}_{ij}, \widetilde{a}_i$ are in $C_P^{N+1}([0, T_0) \times \mathbb{R}^d)$, with partial derivatives up to order $N + 1$ bounded by M.

(ii) *Uniform ellipticity:*

$$\varepsilon M|\zeta|^2 \leq \sum_{i,j=1}^{d} \widetilde{a}_{ij}(t,z)\zeta_i\zeta_j \leq M|\zeta|^2, \qquad t \in [0,T_0), z, \zeta \in \mathbb{R}^d.$$

Note that if Assumption 4.2 is satisfied with $r = +\infty$, then the operator $\mathcal{A}_t$ is uniformly elliptic with bounded coefficients. The forthcoming error bounds will be asymptotic in the limit of small $M(T-t)$; in particular, the constant C appearing in the error estimates will be dependent on M_0, but not on M.

Assumption 4.2 is (locally) equivalent to Assumption 2.4. Precisely, the former implies the latter on the domain $D = (e^{x_0-r}, e^{x_0+r}) \times B(y_0, r)$. Therefore, when Assumptions 2.1, 2.5 and 4.2 are in force, in light of Theorem 2.6, there exists a local transition density $\bar{\Gamma}$ on D for the process (S, Y). We then define the *logarithmic local density* Γ as

$$\Gamma(t,x,y;T,\xi,\eta) = e^{\xi}\, \bar{\Gamma}(t,e^x,y;T,e^{\xi},\eta)$$

for any $(T,\xi,\eta) \in H(T_0,z_0,r)$ and $(t,x,y) \in \bar{H}(T,z_0,r)$.

Remark 4.3 Clearly, Lemma 2.2 and Theorem 2.6 can be extended to Γ through the logarithmic change of variables. In particular, in this section, we use that

(i) $\Gamma(t,z;\cdot,\cdot)$ is in $C_P^{N,1}((t,T_0) \times D(z_0,r))$ for any $(t,z) \in \bar{H}(T_0,z_0,r)$;
(ii) $\Gamma(\cdot,\cdot;T,\zeta)$ is in $C_P^{N+2,1}(\bar{H}(T,z_0,r))$ for any $(T,\zeta) \in H(T_0,z_0,r)$ and solves the backward Kolmogorov equation

$$(\partial_t + \mathcal{A}_t) f = 0 \qquad \text{on } \bar{H}(T,z_0,r). \tag{4.1}$$

Moreover, for any $(T,\bar{z}) \in H(T_0,z_0,r)$ and $\varphi \in C_b(D(z_0,r))$, we have

$$\lim_{\substack{(t,z)\to(T,\bar{z}) \\ t<T}} \int_{D(z_0,r)} \Gamma(t,z;T,\zeta)\varphi(\zeta)d\zeta = \varphi(\bar{z});$$

(iii) if u is the function defined in (3.2), then for any $T \in (0,T_0)$ and $k \in \mathbb{R}$, we have that $u(\cdot,\cdot;T,k)$ is in $C_P^{N+2,1}(\bar{H}(T,z_0,r)) \cap C([0,T] \times D(z_0,r))$ and solves (4.1).

Next we prove sharp error estimates for the derivatives $\partial_k^m(u - \bar{u}_N)$. In Sect. 4.1, we prove some global bounds in the case $r = +\infty$, and then in Sect. 4.2, we prove analogous local bounds in the general case $r < +\infty$.

4.1 Error estimates for uniformly parabolic equations

Throughout this section, we assume that Assumption 4.2 is satisfied with $r = +\infty$. Under this assumption, u is the unique[7] classical solution of the Cauchy problem

[7] The solution is unique within the class of non-rapidly increasing functions.

(3.5) and can be represented as

$$u(t,z)=\int_{\mathbb{R}^d}\Gamma(t,z;T,\xi,\eta)(e^{\xi}-e^{k})^{+}d\xi d\eta,\qquad t\in[0,T),\,z\in\mathbb{R}^d,$$

where Γ is the fundamental solution of the uniformly parabolic operator $\partial_t+\mathcal{A}_t$. In the following statement, $\bar{u}_N$ is the Nth order approximation of u as defined in (3.11).

Theorem 4.4 *Let Assumptions 2.1, 2.5 and 4.2 hold with $r=+\infty$. Then for any $m,q\in\mathbb{N}_0$ with $m+2q\le N$, we have*

$$|\partial_T^q\partial_k^m(u-\bar{u}_N)(t,x,y;T,k)|\le Ce^xM^q\big(M(T-t)\big)^{\frac{N-m-2q+2}{2}}\tag{4.2}$$

for $0\le t<T<T_0$, $x,k\in\mathbb{R}$ and $y\in\mathbb{R}^{d-1}$. The constant C in (4.2) depends only on T_0, M_0, ε, N and the dimension d. In particular, C is independent of M.

The proof of Theorem 4.4, which is postponed to Appendix A, is based on the following classical Gaussian estimates (see for instance [39, Theorem 8.10]).

Lemma 4.5 *Let $\Gamma=\Gamma(t,z;T,\zeta)$ be the fundamental solution of $\mathcal{A}_t+\partial_t$. Then for any $c>1$, $q\in\mathbb{N}_0$ and $\beta,\gamma\in\mathbb{N}_0^d$ with $|\beta|+2q\le N$, we have*

$$|(z-\zeta)^{\gamma}\partial_T^qD_\zeta^\beta\Gamma(t,z;T,\zeta)|\le CM^q\big(M(T-t)\big)^{\frac{|\gamma|-|\beta|-2q}{2}}\Gamma_0\big(cM(T-t),z-\zeta\big)$$

for $0\le t<T\le T_0$ and $z,\zeta\in\mathbb{R}^d$, where Γ_0 is the d-dimensional standard Gaussian function

$$\Gamma_0(t,z)=(2\pi t)^{-\frac{d}{2}}\exp\left(-\frac{|z|^2}{2t}\right),\qquad t\in\mathbb{R}_{++},\ z\in\mathbb{R}^d,\tag{4.3}$$

and C is a positive constant that depends only on c, T_0, M_0, ε, N and the dimension d.

4.2 Error estimates for locally parabolic equations

We now relax the global parabolicity assumption of Sect. 4.1 by assuming that the pricing operator $\mathcal{A}_t$ is only *locally elliptic*; precisely, **throughout this section**, we impose that Assumptions 2.1, 2.5 and 4.2 hold for some $r>0$. We first state the result in the one-dimensional case.

Theorem 4.6 *Let $d=1$. Under Assumptions 2.1, 2.5 and 4.2, for any $\delta\in(0,1)$, $T\in(0,T_0)$ and $m\le N$, we have*

$$|\partial_k^mu(t,z;T,k)-\partial_k^m\bar{u}_N(t,z;T,k)|\le C\big(M(T-t)\big)^{\frac{N-m+2}{2}}$$

for $(t,z)\in\bar{H}(T,z_0,\delta r)$ and $|k-x_0|<\delta r$, where C is a positive constant that depends only on r, z_0, δ, d, M_0, ε, N and T_0. In particular, C is independent of M.

The proof of Theorem 4.6 is a simple modification of that of Theorem 4.9 below and therefore omitted. Theorem 4.9 is the main result of this section; it gives estimates for the derivatives of the price function with respect to the log-strike k in dimension $d \geq 2$.

For the rest of the section, we fix $\hat{N} \in \mathbb{N}_0$ with $\hat{N} \leq N$ and consider $d \geq 2$. By our general assumptions (see in particular Remark 4.3), we have that for any $T \in (0, T_0)$, $(t, z) \in \bar{H}(T, z_0, r)$, $|k - x_0| < r$ and $\delta \in [0, 1]$, the pricing function u can be represented as

$$u(t, z; T, k) = I_{1,\delta}(t, z; T, k) + I_{2,\delta}(t, z; T, k), \tag{4.4}$$

where

$$I_{1,\delta}(t, z; T, k) = \int_{D(z_0,\delta r)} (e^{\xi} - e^{k})^{+} \Gamma(t, z; T, \xi, \eta) d\xi d\eta,$$

$$I_{2,\delta}(t, z; T, k) = \int_{\mathbb{R}^d \setminus D(z_0,\delta r)} (e^{\xi} - e^{k})^{+} p(t, z; T, d\xi, d\eta),$$

and p denotes the transition distribution of the process $(\log S, Y)$. We note explicitly that even if $\log S$ takes values in $[-\infty, +\infty)$ (due to the possibility for S to reach 0), we can exclude $\{-\infty\} \times \mathbb{R}^{d-1}$ from the domain of integration of $I_{2,\delta}$ because the call payoff function is zero for $\xi \leq k$.

Formula (4.4) is useful to study the regularity properties of u with respect to k and T. In fact, by (i) of Remark 4.3, $I_{1,\delta}$ is twice differentiable in k, with $\partial_k^2 I_{1,\delta}(t, z; \cdot, \cdot)$ being in $C_P^N((t, T_0) \times D(z_0, r))$, and we have

$$\partial_T^q \partial_k^m I_{1,\delta}(t, z; T, k) = U_{1,q,m,\delta}(t, z; T, k) + U_{2,q,m,\delta}(t, z; T, k), \tag{4.5}$$

where

$$U_{1,q,m,\delta}(t, z; T, k) = e^k \int_k^{x_0+\delta r} \int_{\{|\eta - y_0| < \delta r\}} \partial_T^q \Gamma(t, z; T, \xi, \eta) d\xi d\eta,$$

$$U_{2,q,m,\delta}(t, z; T, k) = e^k \sum_{j=1}^{m-1} \binom{m-1}{j} \int_{\{|\eta - y_0| < \delta r\}} \partial_T^q \partial_k^{j-1} \Gamma(t, z; T, k, \eta) d\eta$$

for $(t, z) \in \bar{H}(T, z_0, r)$ and $k \in B(x_0, \delta r)$. However, the assumptions imposed in Sect. 2 are not sufficient to ensure the existence of the derivatives $\partial_T^q \partial_k^m I_{2,\delta}$ (and consequently of $\partial_T^q \partial_k^m u$). Indeed, a formal computation gives

$$\partial_T^q \partial_k^m I_{2,\delta}(t, z; T, k) = U_{3,q,m,\delta}(t, z; T, k) + U_{4,q,m,\delta}(t, z; T, k), \tag{4.6}$$

where

$$U_{3,q,m,\delta}(t, z; T, k) = \partial_T^q e^k \int_{[x_0+\delta r, +\infty) \times \mathbb{R}^{d-1}} p(t, z; T, d\xi, d\eta),$$

$$U_{4,q,m,\delta}(t, z; T, k) = \partial_T^q \partial_k^m \int_{(k, x_0+\delta r) \times (\mathbb{R}^{d-1} \setminus B(y_0, \delta r))} p(t, z; T, d\xi, d\eta)(e^{\xi} - e^{k}).$$

Now, it is clear that $U_{3,q,m,\delta}$ depends smoothly on k. In contrast, the existence and boundedness properties of the derivatives $U_{4,q,m,\delta}$ depend on the tails of the distribution and cannot be deduced from the general assumptions of Sect. 2 because of the *local nature* of those assumptions. Notice that this problem only arises when $d \geq 2$, and therefore, in order to prove results in the most general setting, we need to impose the following additional

Assumption 4.7 For any $(t, z) \in \bar{H}(T_0, z_0, r)$, we have

$$u(t, z; \cdot, \cdot) \in C_P^{\hat{N}}\big((t, T_0) \times D(z_0, r)\big).$$

Moreover, in the case $\hat{N} \geq 2$, there exist $\delta \in (0, 1)$ and some positive constants $\widetilde{C}$ and $\bar{C}$ such that

$$|\partial_T^q \partial_k^m \Gamma(t, z; T, k, \eta)| \leq \widetilde{C}, \qquad 2q + m \leq \hat{N}, \tag{4.7}$$

for any $(T, k, \eta) \in H(T_0, z_0, \delta^2 r)$, $(t, z) \in \bar{H}(T, z_0, r) \setminus \bar{H}(T, z_0, \delta r)$, and

$$|U_{3,q,m,\delta^2}(t, z; T, k)| + |U_{4,q,m,\delta^2}(t, z; T, k)| \leq \bar{C}, \qquad 2q + m \leq \hat{N}, \tag{4.8}$$

for any $(T, k) \in (0, T_0) \times B(x_0, \delta^2 r)$ and $(t, z) \in \bar{H}(T, z_0, \delta^3 r)$.

Remark 4.8 If $\log S_T$ (or equivalently, S_T) has a *marginal* local density $\Gamma_S(t, z; T, k)$ such that

$$\partial_T^q \partial_k^m \Gamma_S(t, z; \cdot, \cdot) \in C\big((t, T_0) \times B(x_0, r)\big), \qquad 2q + m \leq \hat{N},$$

then the first part of Assumption 4.7 is satisfied; in fact, we have

$$u(t, z; \cdot, \cdot) \in C_P^{\hat{N}}\big((t, T_0) \times B(x_0, r)\big)$$

because it can be represented as

$$u(t, z; T, k) = \int_k^{\bar{k}} \Gamma_S(t, z; T, \xi)(e^{\xi} - e^k) d\xi + \int_{[\bar{k}, +\infty)} p_S(t, z; T, d\xi)(e^{\xi} - e^k)$$

for some $\bar{k} > k$, where p_S denotes the marginal transition probability of $\log S$. This is the case, for instance, for the Heston model, where S_T has a smooth marginal density (see Remark 2.8).

The need for conditions (4.7) and (4.8) will be clarified in the proofs of Lemma 4.11 and Theorem 4.9, respectively. Condition (4.7) is intuitively easy to understand: roughly speaking, it states that the derivatives of the local density $\Gamma(t, z; T, \zeta)$ are locally bounded, away from the pole, all the way up to $t = T$. This looks like a sensible condition, given the boundedness hypothesis for the diffusion coefficients on the whole cylinder. By contrast, condition (4.8) might seem a little bit cryptic at a first glance; however, in most cases of interest, such a hypothesis turns out to be substantially simplified. For instance, in many financial models such as the Heston model, the local density Γ is defined on the whole strip $B(x_0, r) \times \mathbb{R}^{d-1}$ (see Remark 2.8), i.e., we have

$$p(t, z; T, H) = \int_H \Gamma(t, z; T, \zeta) d\zeta, \qquad H \in \mathcal{B}\big(B(x_0, r) \times \mathbb{R}^{d-1}\big).$$

In this case, condition (4.8) is automatically satisfied for $q = 0$ and $m = 0, 1$, whereas for $2 \le m + 2q \le \hat{N}$, it reduces to

$$\left| \int_{[x_0+\delta r, +\infty)\times\mathbb{R}^{d-1}} \partial_T^q \Gamma(t, z; T, \zeta) d\zeta \right| + \left| \int_{\{|\eta - y_0| > \delta^2 r\}} \partial_T^q \partial_k^{(m-2)\vee 0} \Gamma(t, z; T, k, \eta) d\eta \right| \le \bar{C}$$

for any $(T, k) \in (0, T_0) \times B(x_0, \delta^2 r)$, $(t, z) \in \bar{H}(T, z_0, \delta^3 r)$.

We are now ready to state the main result of this section.

Theorem 4.9 *Let $d \ge 2$, and let Assumptions 2.1, 2.5, 4.2 and 4.7 be in force. Then for any $m, q \in \mathbb{N}_0$ with $m + 2q \le \hat{N}$ and $T \in (0, T_0)$, we have*

$$|\partial_T^q \partial_k^m (u - \bar{u}_N)(t, z; T, k)| \le C M^q \big(M(T - t)\big)^{\frac{N-m-2q+2}{2}}$$

for $(t, z) \in \bar{H}(T, z_0, \delta^4 r)$ and $|k - x_0| < \delta^4 r$, where $\delta \in (0, 1)$ is as in Assumption 4.7, and the positive constant C depends only on $r, z_0, d, M_0, \varepsilon, N, T_0$ and, only if $\hat{N} \ge 2$, also on δ and the constants $\widetilde{C}$ and $\bar{C}$ in (4.7) and (4.8). In particular, C is independent of M.

Lemma 4.10 *Let D_0 be a domain of $\mathbb{R}^n$ and*

$$h(\cdot, \cdot; T, \theta) : \overline{H(T, z_0, r)} \to \mathbb{R}, \qquad (T, \theta) \in (0, T_0) \times D_0,$$

such that

(i) *for any $(t, z) \in [0, T_0) \times \overline{D(z_0, r)}$, the function $h(t, z; \cdot, \cdot)$ is in $C^p((t, T_0) \times D_0)$ with derivatives $\partial_T^q D_\theta^\beta h(t, z; T, \theta)$ locally bounded in (T, θ), uniformly with respect to $(t, z) \in [0, T) \times (\overline{D(z_0, r)} \setminus D(z_0, \varrho_0 r))$ for a certain $\varrho_0 \in (0, 1)$;*
(ii) *for any $(T, \theta) \in (0, T_0) \times D_0$, the function $h(\cdot, \cdot; T, \theta)$ belongs to the space $C^{1,2}(\bar{H}(T, z_0, r)) \cap C(\overline{H(T, z_0, r)})$ and verifies*

$$\begin{cases} (\partial_t + \widetilde{\mathcal{A}}_t) h(t, z; T, \theta) = 0 & \text{for } (t, z) \in \bar{H}(T, z_0, r), \\ h(T, z; T, \theta) = 0 & \text{for } (t, z) z \in D(z_0, r). \end{cases} \tag{4.9}$$

Then for any multi-index $\beta \in \mathbb{N}_0^n$ and any $q \in \mathbb{N}_0$ with $q + |\beta| \le p$, we have

$$\lim_{\substack{(t,z)\to(T,\bar{z}) \\ t<T}} \partial_T^q D_\theta^\beta h(t, z; T, \theta) = 0, \qquad \bar{z} \in D(z_0, r), (T, \theta) \in (0, T_0) \times D_0. \tag{4.10}$$

Proof By induction on q, we prove (4.10) and that for any $\varrho \in [\varrho_0, 1)$, we have

$$\partial_T^q D_\theta^\beta h(t, z; T, \theta) = \int_t^T \int_{\partial D(z_0, \varrho r)} P_{\varrho r}(t, z; s, \zeta) \partial_T^q D_\theta^\beta h(s, \zeta; T, \theta) d\zeta ds \tag{4.11}$$

for $(t, z) \in H(T, z_0, \varrho r)$, where $P_{\varrho r}$ denotes the Poisson kernel of the uniformly parabolic operator $\partial_t + \widetilde{\mathcal{A}}_t$ on $H(T, z_0, \varrho r)$.

For $q = 0$, differentiating the representation formula

$$h(t, z; T, \theta) = \int_t^T \int_{\partial D(z_0, \varrho r)} P_{\varrho r}(t, z; s, \zeta) h(s, \zeta; T, \theta) d\zeta ds$$

for $(t, z) \in H(T, z_0, \varrho r)$ and using the terminal condition in (4.9), we obtain

$$\begin{aligned} |D_\theta^\beta h(t, z; T, \theta)| \le & \|D_\theta^\beta h(\cdot, \cdot; T, \theta)\|_{L^\infty(\Sigma(T, z_0, \varrho r))} \\ & \times \int_t^T \int_{\partial D(z_0, \varrho r)} P_{\varrho r}(t, z; s, \zeta) d\xi ds \end{aligned}$$

for $(t, z) \in H(T, z_0, \varrho r)$, which in turn implies (4.10) with $q = 0$.

Next, we assume (4.10) and (4.11) true for q; by differentiating (4.11), we get by (4.10) that

$$\begin{aligned} \partial_T^{q+1} D_\theta^\beta h(t, z; T, \theta) = & \int_{\partial D(z_0, \varrho r)} P_{\varrho r}(t, z; T, \zeta) \partial_T^q D_\theta^\beta h(T, \zeta; T, \theta) d\zeta \\ & + \int_t^T \int_{\partial D(z_0, \varrho r)} P_{\varrho r}(t, z; s, \zeta) \partial_T^{q+1} D_\theta^\beta h(s, \zeta; T, \theta) d\zeta d \\ = & \int_t^T \int_{\partial D(z_0, \varrho r)} P_{\varrho r}(t, z; s, \zeta) \partial_T^{q+1} D_\theta^\beta h(s, \zeta; T, \theta) d\zeta ds \end{aligned}$$

for $(t, z) \in H(T, z_0, \varrho r)$. Then for $(t, z) \in H(T, z_0, \varrho r)$, we have

$$\begin{aligned} |\partial_T^{q+1} D_\theta^\beta h(t, z; T, \theta)| \le & \|\partial_T^{q+1} D_\theta^\beta h(\cdot, \cdot; T, \theta)\|_{L^\infty(\Sigma(T, z_0, \varrho r))} \\ & \times \int_t^T \int_{\partial D(z_0, \varrho r)} P_{\varrho r}(t, z; s, \zeta) d\xi ds, \end{aligned}$$

which concludes the proof. □

The following lemma is preparatory for the proof of Theorem 4.9, but it may also have an independent interest. It shows that the difference between Γ and $\widetilde{\Gamma}$, and between their derivatives, decays exponentially on $H(T, z_0, r)$ as t approaches T.

Lemma 4.11 *Let $\hat{N} \ge 2$ and let $\widetilde{\Gamma}$ be the fundamental solution of the uniformly parabolic operator $\partial_t + \widetilde{\mathcal{A}}_t$. Then under the assumptions of Theorem 4.9, for any $m, q \in \mathbb{N}_0$ with $m + 2q \le \hat{N}$, we have*

$$|\partial_T^q \partial_k^m (\Gamma - \widetilde{\Gamma})(t, z; T, k, \eta)| \le C e^{-\frac{1}{C\sqrt{M(T-t)}}} \tag{4.12}$$

for $(T,k,\eta) \in H(T_0, z_0, \delta^2 r)$ *and* $(t,z) \in \bar{H}(T, z_0, \delta^2 r)$, *where* C *is a positive constant that depends only on* $z_0, \delta, N, d, M_0, \varepsilon, T_0$, *and on* $\widetilde{C}, \bar{C}$ *in* (4.7) *and* (4.8).

Proof Step 1. Fix $(T,k,\eta) \in H(T_0, z_0, \delta^2 r)$ and consider the function

$$w_{q,m}(t,z) := \partial_T^q \partial_k^m (\Gamma - \widetilde{\Gamma})(t,z;T,k,\eta), \qquad (t,z) \in \bar{H}(T, z_0, r).$$

We prove that

$$\begin{cases} (\partial_t + \widetilde{\mathcal{A}}_t) w_{q,m} = 0 & \text{on } \bar{H}(T, z_0, r), \\ \lim\limits_{\substack{(t,z)\to(T,\bar{z}) \\ t<T}} w_{q,m}(t,z) = 0 & \text{for } \bar{z} \in D(z_0, r). \end{cases} \tag{4.13}$$

The first equality in (4.13) follows because $\mathcal{A}_t$ and $\widetilde{\mathcal{A}}_t$ coincide on $\bar{H}(T_0, z_0, r)$. To prove the second one, we set

$$h(t,z;k) := \int_{D(z_0,r)} \big(\Gamma(t,z;T,\zeta) - \widetilde{\Gamma}(t,z;T,\zeta)\big)\, \psi\big(\zeta - (k,\eta)\big) d\zeta$$

for $(t,z) \in \bar{H}(T, z_0, r)$, where

$$\psi(z) := \prod_{i=1}^d \zeta_i^+, \qquad \zeta = (\zeta_1, \dots, \zeta_d) \in \mathbb{R}^d.$$

Notice that $h(\cdot,\cdot;T,k,\eta)$ satisfies

$$\begin{cases} (\partial_t + \widetilde{\mathcal{A}}_t) h(t,z;k) = 0 & \text{for } (t,z) \in \bar{H}(T, z_0, r), \\ h(t,z;k) = 0 & \text{for } z \in D(z_0, r). \end{cases}$$

Moreover, we have

$$\partial_k^2 \partial_{\eta_2}^2 \cdots \partial_{\eta_d}^2 h(t,z;k) = \Gamma(t,z;T,k,\eta) - \widetilde{\Gamma}(t,z;T,k,\eta)$$

and therefore also

$$\partial_T^q \partial_k^{2+m} \partial_{\eta_2}^2 \cdots \partial_{\eta_d}^2 h(t,z;k) = w_{q,m}(t,z).$$

Hence by applying Lemma 4.10 to h, we obtain the limit in (4.13).

Step 2. It suffices to prove the thesis for $T-t$ suitably small and positive. In [38, Theorem 3.1], we proved that there exist $\tau > 0$ and a nonnegative function v such that

$$\begin{cases} (\partial_t + \widetilde{\mathcal{A}}_t) v(t,z) = 0 & \text{for } (t,z) \in [T-\tau, T) \times D(z_0, r), \\ v(t,z) \geq 1 & \text{for } (t,z) \in [T-\tau, T) \times \partial D(z_0, r) \end{cases} \tag{4.14}$$

and

$$0 < v(t,z) \leq C e^{-\frac{r^2}{C\sqrt{M(T-t)}}} \qquad \text{for } (t,z) \in [T-\tau, T) \times D(z_0, \delta^2 r), \tag{4.15}$$

where the positive constant C depends only on δ, M_0, ε, T_0, z_0 and d. Now, by (4.14), (4.15), and by the limit in (4.13) together with the bound (4.7), one has

$$\liminf_{\substack{(t,z)\to(\bar t,\bar z)\\(t,z)\in[T-\tau,T)\times D(z_0,r)}} (\widetilde{C}v - w_{q,m})(t,z) \geq 0$$

for $(\bar t,\bar z) \in (\{T\}\times D(z_0,r)) \cup ([T-\tau,T)\times \partial D(z_0,r))$. Therefore, the maximum principle yields

$$|w_{q,m}(t,z)| \leq \widetilde{C}v(t,z), \qquad (t,z)\in[T-\tau,T)\times D(z_0,r),$$

and eventually (4.12) follows from (4.15). □

Proof of Theorem 4.9 We only prove the statement for $2 \leq m \leq \hat N$, the other cases being simpler. **Throughout the proof**, we denote by C every positive constant that depends at most on r, z_0, δ, d, M_0, ε, N, T_0 and on $\widetilde{C}$, $\bar C$ in (4.7) and (4.8).

Step 1. We fix $T \in (0,T_0)$ and prove that

$$|w_{q,m}(t,z;T,k)| \leq C, \qquad (t,z)\in \bar H(T,z_0,\delta^3 r),\ k\in B(x_0,\delta^3 r), \tag{4.16}$$

where $w_{q,m} := \partial_T^q \partial_k^m (u-\widetilde{u})$ and for $(t,z)\in[0,T)\times\mathbb{R}^d$,

$$\widetilde{u}(t,z;T,k) := \int_k^\infty \int_{\mathbb{R}^{d-1}} \widetilde{\Gamma}(t,z;T,\xi,\eta)(e^\xi - e^k)d\xi d\eta. \tag{4.17}$$

Differentiating (4.4) and recalling (4.5) and (4.6), we get

$$\partial_T^q \partial_k^m u(t,z;T,k) = \sum_{i=1}^4 (-1)^i U_{i,q,m,\delta}(t,z;T,k).$$

Analogously, differentiating (4.17), we obtain

$$\begin{aligned}\partial_T^q \partial_k^m \widetilde{u}(t,z;T,k) = &- e^k \int_k^\infty \int_{\mathbb{R}^{d-1}} \partial_T^q \widetilde{\Gamma}(t,z;T,\xi,\eta)d\xi d\eta \\ &+ \sum_{j=1}^{m-1} \binom{m-1}{j} e^k \int_{\mathbb{R}^{d-1}} \partial_T^q \partial_k^{j-1} \widetilde{\Gamma}(t,z;T,k,\eta)d\eta.\end{aligned}$$

Thus we have by (4.8)

$$\begin{aligned}|w_{q,m}(t,z;T,k)| &\leq C\left(1 + U_{4,q,m,\delta^2}(t,z;T,k)\right) \\ &\quad + C\sum_{j=1}^{m-1}(J_{1,q,j,\delta^2} + J_{2,q,j,\delta^2})(t,z;T,k) \\ &\leq C\left(1 + \sum_{j=1}^{m-1}(J_{1,q,j,\delta^2} + J_{2,q,j,\delta^2})(t,z;T,k)\right)\end{aligned}$$

for any $k \in B(x_0, \delta^2 r)$ and $(t, z) \in \bar{H}(T, z_0, \delta^3 r)$, where

$$J_{1,q,j,\delta^2}(t, z; T, k) = \int_{\{|\eta - y_0| < \delta^2 r\}} |\partial_T^q \partial_k^{j-1} (\Gamma - \widetilde{\Gamma})(t, z; T, k, \eta)| d\eta,$$

$$J_{2,q,j,\delta^2}(t, z; T, k) = \int_{\{|\eta - y_0| \geq \delta^2 r\}} |\partial_T^q \partial_k^{j-1} \widetilde{\Gamma}(t, z; T, k, \eta)| d\eta.$$

Now, by applying Lemma 4.11 and standard Gaussian estimates on the functions J_{1,q,j,δ^2} and J_{2,q,j,δ^2}, respectively, we obtain that the latter are bounded by a constant C for any $k \in B(x_0, \delta^2 r)$ and $(t, z) \in \bar{H}(T, z_0, \delta^3 r)$. This proves (4.16).

Step 2. Fix now $(T, k) \in (0, T_0] \times B(x_0, \delta^3)$. Clearly, $\widetilde{u}(\cdot, \cdot; T, k)$ in (4.17) is a classical solution to the Cauchy problem

$$\begin{cases} (\partial_t + \widetilde{\mathcal{A}}_t)\widetilde{u}(\cdot, \cdot; T, k) = 0 & \text{on } [0, T) \times \mathbb{R}^d, \\ \widetilde{u}(T, x, y; T, k) = (e^x - e^k)^+ & \text{for } (x, y) \in \mathbb{R}^d. \end{cases}$$

We set $h(t, z; k) := (u - \widetilde{u})(t, z; T, k)$ and notice that by Remark 4.3(iii), we have

$$(\partial_t + \widetilde{\mathcal{A}}_t)h(\cdot, \cdot; k) = 0 \qquad \text{on } \bar{H}(T, z_0, r), \tag{4.18}$$

because $\mathcal{A}_t$ and $\widetilde{\mathcal{A}}_t$ coincide on $\bar{H}(T_0, z_0, r)$; moreover, we have

$$h(T, z; k) = 0 \qquad \text{for } z \in D(z_0, r).$$

Now, by (4.16), the derivatives $\partial_T^q \partial_k^m h = w_{q,m}$ are bounded on $\Sigma(T, z_0, \delta^3 r)$ for $k \in B(x_0, \delta^3)$. Then from Lemma 4.10 applied to h on $\bar{H}(T, z_0, \delta^3 r)$, we infer

$$\lim_{\substack{(t,z)\to(T,\bar{z}) \\ t<T}} w_{q,m}(t, z; T, k) = 0 \qquad \text{for } \bar{z} \in D(z_0, \delta^3 r). \tag{4.19}$$

By differentiating (4.18), we also obtain that $(\partial_t + \widetilde{\mathcal{A}}_t)w_{q,m}(\cdot, \cdot; T, k) = 0$ on the set $\bar{H}(T, z_0, \delta^2 r)$. Thus we can use the same argument as in Step 2 of the proof of Lemma 4.11; precisely, we consider the function v satisfying (4.14) and (4.15) and by the maximum principle, (4.19) and (4.16), we infer that

$$|w_{q,m}(t, z; T, k)| \leq \|w_{q,m}(\cdot, \cdot; T, k)\|_{L^\infty(\Sigma(T, z_0, \delta^3 r))} e^{-\frac{r^2}{C\sqrt{M(T-t)}}}$$

for $(t, z) \in \bar{H}(T, z_0, \delta^4 r)$. Eventually, by the triangular inequality, we get

$$|\partial_k^m (u - \bar{u}_N)| \leq |w_{q,m}| + |\partial_k^m (\widetilde{u} - \bar{u}_N)| \leq C e^{-\frac{r^2}{C\sqrt{M(T-t)}}} + |\partial_k^m (\widetilde{u} - \bar{u}_N)|$$

on $\bar{H}(T, z_0, \delta^4 r)$, and the statement follows from the asymptotic estimate of Theorem 4.4 applied to the uniformly parabolic operator $\partial_t + \widetilde{\mathcal{A}}_t$. ☐

5 Error estimates and Taylor formula of the implied volatility

In this section, we establish error estimates for the Nth order implied volatility approximation $\bar{\sigma}_N(t,x,y;T,k)$ in Definition 3.4 and for its derivatives with respect to k and T. Such bounds are proved under the assumptions of Sect. 4.2, and are valid in the *parabolic* domain $|x-k| \le \lambda\sqrt{M(T-t)}$, for any $\lambda > 0$ and suitably small time to maturity $T-t$, with M being the local ellipticity constant in Assumption 4.2. We recall that $N, \hat{N} \in \mathbb{N}_0$ are fixed **throughout the paper** and such that $N \ge 2$ and $\hat{N} \le N$. Moreover, $z_0 = (x_0, y_0) \in \mathbb{R} \times \mathbb{R}^{d-1}$ is the center of the cylinder in Assumptions 4.2 and 4.7.

Theorem 5.1 *Let $d=1$ ($d \ge 2$) and let the assumptions of Theorem 4.6 (Theorem 4.9) be in force. Then for any $\lambda > 0$ and $m, q \in \mathbb{N}_0$ with $2q+m \le \hat{N}$, there exist two positive constants C and τ_0 such that*

$$|\partial_T^q \partial_k^m \sigma(t,x_0,y_0;T,k) - \partial_T^q \partial_k^m \bar{\sigma}_N(t,x_0,y_0;T,k)| \le C M^{q+\frac{1}{2}} \big(M(T-t)\big)^{\frac{N-m-2q+1}{2}}$$

for any $0 \le t < T < T_0$ and k such that $T-t \le \tau_0$ and $|x_0-k| \le \lambda\sqrt{M(T-t)}$. The constants C and τ_0 depend only on $r, z_0, d, M_0, \varepsilon, N, T_0, \lambda$ and, if both $d, \hat{N} \ge 2$, also on δ and the constants $\widetilde{C}$ and $\bar{C}$ in (4.7) and (4.8). In particular, C and τ_0 are independent of M.

Before proving Theorem 5.1, we show the following remarkable corollary which is the main result of the paper.

Corollary 5.2 *Let the assumptions of Theorem 5.1 hold, and for simplicity assume $N = \hat{N}$. Then for any $q, m \in \mathbb{N}_0$ with $2q+m \le N$, the two limits*

$$\partial_T^q \partial_k^m \bar{\sigma}_N(t,x_0,y_0;t,x_0) := \lim_{\substack{(T,k)\to(t,x_0)\\ |x_0-k|\le\lambda\sqrt{T-t}}} \partial_T^q \partial_k^m \bar{\sigma}_N(t,x_0,y_0;T,k), \tag{5.1}$$

$$\partial_T^q \partial_k^m \sigma(t,x_0,y_0;t,x_0) := \lim_{\substack{(T,k)\to(t,x_0)\\ |x_0-k|\le\lambda\sqrt{T-t}}} \partial_T^q \partial_k^m \sigma(t,x_0,y_0;T,k) \tag{5.2}$$

exist, are finite, and coincide for any $\lambda > 0$ and $t \in [0, T_0)$. Consequently, we have the parabolic Nth order Taylor expansion

$$\sigma(t,x_0,y_0;T,k) = \sum_{2q+m\le N} \frac{(T-t)^q (k-x_0)^m}{q!m!} \partial_T^q \partial_k^m \bar{\sigma}_N(t,x_0,y_0;t,x_0) + R_N(t,x_0,y_0,T,k) \tag{5.3}$$

with

$$R_N(t,x_0,y_0,T,k) = \mathrm{o}(|T-t|^{\frac{N}{2}} + |k-x_0|^N)$$

as $(T,k) \to (t,x_0)$, with $|x_0-k| \le \lambda\sqrt{T-t}$.

Proof By Theorem 5.1, we have

$$\lim_{\substack{(T,k)\to(t,x_0)\\ |x_0-k|\le\lambda\sqrt{T-t}}} \partial_T^q \partial_k^m (\sigma-\bar{\sigma}_N)(t,x_0,y_0;T,k)=0, \qquad t\in[0,T_0),\lambda>0,$$

for any $q, m \in \mathbb{N}_0$ with $2q+m\le N$. Therefore, the limit in (5.1) exists if and only if the limit (5.2) exists, and in that case, they coincide. Now, by the representation formulas in Theorem D.1 and Proposition D.3, $\bar{\sigma}_N(t,x_0,y_0;\cdot,\cdot)$ is in $C_P^N([0,T_0)\times\mathbb{R})$ and thus the limit in (5.2) exists. □

Remark 5.3 The derivatives appearing in the Taylor formula (5.3) can be computed explicitly (possibly with the aid of a symbolic computation software) by means of the representation formulas of Theorem D.1 and Proposition D.3.

Remark 5.4 A direct computation shows that at order $N=0$, formula (5.3) is consistent with the well-known results in [6, 7]. Furthermore, again by direct computation, one can check that in the special case $d=1$, formula (5.3) with $q=0$ and $m=1$ is consistent with the well-known practitioners' $1/2$ *slope rule*, according to which the at-the-money slope of the implied volatility is one-half the slope of the local volatility function.

The rest of the section is devoted to the proof of Theorem 5.1. Hereafter $\lambda>0$ is fixed and we assume the hypotheses of Theorem 5.1 to be in force. In particular, the center $z_0=(x_0,y_0)$ of the cylinder $H(T_0,z_0,r)$ in Assumptions 4.2 and 4.7 is fixed from now on.

Notation 5.5 If not explicitly stated, C and τ_0 will always denote two positive constants dependent at most on λ, on $r, z_0, d, M_0, \varepsilon, N, T_0, \delta$ appearing in Assumptions 2.1, 2.5, and, only if both $\hat{N}, d\ge 2$, also on $\widetilde{C}, \bar{C}$ in (4.7) and (4.8). Note that in particular, *neither C nor τ_0 depend on M*.

The proof of Theorem 5.1 is based on some preliminary results.

Lemma 5.6 *For any positive constants $c, \bar{\sigma}, \lambda, \mu$ with $\mu<1$, there exists a positive $\bar{\tau}$ only dependent on $c, \bar{\sigma}, \lambda, \mu$ such that*

$$u^{\mathrm{BS}}(\mu\sigma;\tau,x,k)+ce^x\sigma^2\tau\le u^{\mathrm{BS}}(\sigma;\tau,x,k) \tag{5.4}$$

for any $\tau\in[0,\bar{\tau}]$, $\sigma\le\bar{\sigma}$ and $|x-k|\le\lambda\sigma\sqrt{\tau}$.

Proof We recall for the Black–Scholes price the expression (see for instance [43])

$$u^{\mathrm{BS}}(\sigma;\tau,x,k)=(e^x-e^k)^+ + e^x\sqrt{\frac{\tau}{2\pi}}\int_0^\sigma e^{-\frac{1}{2}\left(\frac{x-k}{w\sqrt{\tau}}+\frac{w\sqrt{\tau}}{2}\right)^2}dw.$$

Then we have, by using $|x-k| \le \lambda\sigma\sqrt{\tau}$ and $\sigma \le \bar{\sigma}$,

$$u^{\mathrm{BS}}(\sigma;\tau,x,k) - u^{\mathrm{BS}}(\mu\sigma;\tau,x,k) = e^x\sqrt{\frac{\tau}{2\pi}}\int_{\mu\sigma}^{\sigma} e^{-\frac{1}{2}\left(\frac{x-k}{w\sqrt{\tau}}+\frac{w\sqrt{\tau}}{2}\right)^2} dw$$
$$\ge e^x\sqrt{\frac{\tau}{2\pi}} e^{-\frac{1}{2}\left(\frac{\lambda}{\mu}+\frac{\bar{\sigma}\sqrt{\tau}}{2}\right)^2}\sigma(1-\mu) \ge c e^x \sigma^2 \tau$$

for any $\tau \in [0,\bar{\tau}]$, where $\bar{\tau}$ is a positive and suitably small constant, depending only on $c, \lambda, \bar{\sigma}$ and μ. □

Notation 5.7 Sometimes, in order to simplify the notation, for $k \in \mathbb{R}$ and $T \ge t$, we use the shortcuts

$$\begin{aligned} u^{\mathrm{BS}}(\sigma,k,T) &:= u^{\mathrm{BS}}(\sigma;T-t,x_0,k) && \text{for } \sigma > 0,\\ \sigma^{\mathrm{BS}}(u,k,T) &:= \left(u^{\mathrm{BS}}(\cdot;T-t,x_0,k)\right)^{-1}(u) && \text{for } u \in \left((e^{x_0}-e^k)^+, e^{x_0}\right) \end{aligned}$$

for the Black–Scholes price and its inverse function with respect to the volatility variable. To ease notation, for any function F of three variables z_1, z_2, z_3, we also set $\partial_i F = \frac{\partial F}{\partial z_i}$, $i = 1,2,3$. Derivatives of compositions of u^{BS} and σ^{BS} are expressed according to this notation; for example, first order derivatives are given by

$$\frac{d}{dk}u^{\mathrm{BS}}\left(\sigma^{\mathrm{BS}}(u,k,T),k,T\right) = (\partial_1 u^{\mathrm{BS}})\left(\sigma^{\mathrm{BS}}(u,k,T),k,T\right)\partial_2\sigma^{\mathrm{BS}}(u,k,T) + (\partial_2 u^{\mathrm{BS}})\left(\sigma^{\mathrm{BS}}(u,k,T),k,T\right),$$
$$\frac{d}{dT}u^{\mathrm{BS}}\left(\sigma^{\mathrm{BS}}(u,k,T),k,T\right) = (\partial_1 u^{\mathrm{BS}})\left(\sigma^{\mathrm{BS}}(u,k,T),k,T\right)\partial_3\sigma^{\mathrm{BS}}(u,k,T) + (\partial_3 u^{\mathrm{BS}})\left(\sigma^{\mathrm{BS}}(u,k,T),k,T\right).$$

For any $\delta \in [0,1]$, we introduce the functions

$$u(\delta,k,T) \equiv u(\delta;t,x_0,y_0,T,k) := u^{\mathrm{BS}}\left(\sigma_0^{(x_0,y_0)}(t,T);T-t,x_0,k\right) + R(\delta;t,x_0,y_0,T,k), \tag{5.5}$$
$$R(\delta,k,T) \equiv R(\delta;t,x_0,y_0,T,k) := \sum_{n=1}^{N}\delta^n u_n^{(x_0,y_0)}(t,x_0,y_0;T,k) + \delta^{N+1}(u-\bar{u}_N)(t,x_0,y_0;T,k).$$

Recall that $\sigma_0^{(x_0,y_0)}(t,T)$ and $u_n^{(x_0,y_0)}(t,x_0,y_0;T,k)$ are defined for $0 \le t < T \le T_0$ and $k \in \mathbb{R}$, as indicated by (3.9), (3.10) and (3.8), respectively. Consequently, by Theorem 4.9 and (D.6) in Corollary D.2 there exist C and τ_0 as in Notation 5.5 such that

$$|R(\delta,k,T)| \le C e^{x_0} M(T-t), \tag{5.6}$$

and for any $q,m,h \in \mathbb{N}_0$ and $j \in \mathbb{N}$ with $q+m+h > 0$, $h, j \leq N+1$ and $m+2q \leq \hat{N}$,

$$\left|\partial_T^q \partial_k^m \left(\left(\partial_\delta^h u(\delta, k, T)\right)^j\right)\right| \leq C e^{x_0} M^q \left(M(T-t)\right)^{\frac{j(h+1)-m-2q}{2}} \tag{5.7}$$

for any $0 \leq t < T < T_0$ and k such that $T - t \leq \tau_0$ and $|x_0 - k| \leq \lambda\sqrt{M(T-t)}$.

Lemma 5.8 *There exists a positive τ_0 as in Notation 5.5 such that*

$$u^{\mathrm{BS}}(\sqrt{\varepsilon M}; T-t, x_0, k) \leq u(\delta, k, T) \leq u^{\mathrm{BS}}(\sqrt{4M}; T-t, x_0, k),$$

or equivalently,

$$\sqrt{\varepsilon M} \leq (u^{\mathrm{BS}})^{-1}\left(u(\delta, k, T); T-t, x_0, k\right) \leq \sqrt{4M}, \tag{5.8}$$

for any $\delta \in [0, 1]$, $0 \leq t < T < T_0$ and $k \in \mathbb{R}$ such that $T - t \leq \tau_0$ and $|x_0 - k| \leq \lambda\sqrt{M(T-t)}$.

Proof Since $u(\delta, k, T) - u^{\mathrm{BS}}(\sigma_0^{(x_0,y_0)}(t, T); T-t, x_0, k) = R(\delta, k, T)$, we infer from the estimate (5.6) that

$$\begin{aligned} &u^{\mathrm{BS}}\left(\sigma_0^{(x_0,y_0)}(t, T); T-t, x_0, k\right) - C e^{x_0} M(T-t) \\ &\leq u(\delta, k, T) \\ &\leq u^{\mathrm{BS}}\left(\sigma_0^{(x_0,y_0)}(t, T); T-t, x_0, k\right) + C e^{x_0} M(T-t) \end{aligned} \tag{5.9}$$

with C as in Notation 5.5. Now recall that by Assumption 4.2 along with the definition (3.10), we have

$$\sqrt{2\varepsilon M} \leq \sigma_0^{(x_0,y_0)}(t, T) \leq \sqrt{2M} \leq \sqrt{2M_0}$$

and therefore, for any fixed $\lambda > 0$, the thesis follows by combining (5.9) with the estimate (5.4) with $\mu = \frac{1}{2}$. □

Remark 5.9 In light of Lemma 5.8, the function $\sigma^{\mathrm{BS}}(u(\delta, k, T), k, T)$ is well defined for any $\delta \in [0, 1]$, $0 \leq t < T < T_0$ and $k \in \mathbb{R}$ such that $T - t \leq \tau_0$ and $|x_0 - k| \leq \lambda\sqrt{M(T-t)}$.

Lemma 5.10 *For any $q,m,n \in \mathbb{N}_0$, there exist $C, \tau_0 > 0$ as in Notation 5.5 such that*

$$\left|(\partial_1^n \partial_2^m \partial_3^q \sigma^{\mathrm{BS}})\left(u(\delta, k, T), k, T\right)\right| \leq C M^{q+\frac{1}{2}} \left(M(T-t)\right)^{-\frac{m+2q+n}{2}} e^{-nk} \tag{5.10}$$

for any $\delta \in [0, 1]$, $0 \leq t < T < T_0$ and $k \in \mathbb{R}$ such that $T - t \leq \tau_0$ and $|x_0 - k| \leq \lambda\sqrt{M(T-t)}$. Here C also depends on m, q and n.

Proof See Appendix B. □

Lemma 5.11 *For any $q, m, n \in \mathbb{N}_0$ with $2q+m \le \hat{N}$, there exist $C, \tau_0 > 0$ as in Notation 5.5 such that*

$$\left| \frac{d^{q+m}}{dT^q \, dk^m} (\partial_1^n \sigma^{\mathrm{BS}})\big(u(\delta, k, T), k, T\big) \right| \le C M^{q+\frac{1}{2}} \big(M(T-t)\big)^{-\frac{m+2q+n}{2}} e^{-nk} \quad (5.11)$$

for any $\delta \in [0, 1]$, $0 \le t < T < T_0$ and $k \in \mathbb{R}$ such that $T - t \le \tau_0$ and $|x_0 - k| \le \lambda\sqrt{M(T-t)}$. Here the constant C also depends on n.

Proof See Appendix B. □

We are now ready to prove Theorem 5.1.

Proof of Theorem 5.1 We set

$$G(\delta, k, T) = \sigma^{\mathrm{BS}}\big(u(\delta, k, T), k, T\big)$$

with $\sigma^{\mathrm{BS}} = \sigma^{\mathrm{BS}}(u, k, T)$ and $u = u(\delta, k, T)$ defined in Notation 5.7 and (5.5), respectively. By definition, we have

$$\sigma(k, T) = g(1, k, T), \quad (5.12)$$

where $\sigma(k, T) := \sigma(t, x_0, y_0, k, T)$ is the exact implied volatility. Moreover, for $\bar{\sigma}_N(k, T) := \bar{\sigma}_N(t, x_0, y_0; k, T)$ as defined in (3.16), we have

$$\bar{\sigma}_N(k, T) = \sum_{n=0}^{N} \sigma_n^{(x_0, y_0)}(t, x_0, y_0; k, T) = \sum_{n=0}^{N} \frac{1}{n!} \partial_\delta^n g(\delta, k, T)|_{\delta=0} \quad (5.13)$$

as by (5.5) and (3.14), $g(\delta, k, T)|_{\delta=0} = \sigma_0^{(x_0, y_0)}(t, T)$ and

$$\partial_\delta^n g(\delta, k, T)|_{\delta=0} = \sigma_n^{(x_0, y_0)}(t, x_0, y_0; k, T)$$

for $1 \le n \le N$. Now, by (5.12) and (5.13), there exists $\bar{\delta} \in [0, 1]$ such that

$$\begin{aligned}
&\sigma(k, T) - \bar{\sigma}_N(k, T) \\
&= \frac{1}{(N+1)!} \partial_\delta^{N+1} g(\bar{\delta}, k, T) \\
&= \frac{1}{(N+1)!} \sum_{h=1}^{N+1} (\partial_1^h \sigma^{\mathrm{BS}})\big(u(\bar{\delta}, k, T), k, T\big) \\
&\quad \times \mathbf{B}_{N+1,h}\big(\partial_\delta u(\bar{\delta}, k, T), \partial_\delta^2 u(\bar{\delta}, k, T), \dots, \partial_\delta^{N-h+2} u(\bar{\delta}, k, T)\big),
\end{aligned}$$

where the last equality stems from Faà di Bruno's formula (E.4). Now, differentiating both the left- and the right-hand side m and q times with respect to k and T, respectively, we get

$$
\begin{aligned}
&|\partial_T^q \partial_k^m \sigma(k,T) - \partial_T^q \partial_k^m \bar{\sigma}_N(k,T)| \\
&\leq C \sum_{h=1}^{N+1} \sum_{\ell=0}^{q} \sum_{j=0}^{m} \left| \frac{d^{q-\ell+m-j}}{dT^{q-\ell} dk^{m-j}} (\partial_1^h \sigma^{\mathrm{BS}})\big(u(\bar{\delta},k,T),k,T\big) \right| \\
&\qquad \times \left| \frac{d^{\ell+j}}{dT^\ell dk^j} \mathbf{B}_{N+1,h}\big(\partial_\delta u(\bar{\delta},k,T),\ldots,\partial_\delta^{N-h+2} u(\bar{\delta},k,T)\big) \right|. \qquad (5.14)
\end{aligned}
$$

Again using Faà di Bruno's formula, we have by (5.7) and the identities in (E.6) that

$$
\begin{aligned}
&\left| \frac{d^{\ell+j}}{dT^\ell dk^j} \mathbf{B}_{N+1,h}\big(\partial_\delta u(\bar{\delta},k,T),\ldots,\partial_\delta^{N-h+2} u(\bar{\delta},k,T)\big) \right| \\
&\leq C \sum_{\substack{j_1,\ldots,j_{N-h+2} \\ i_1+\cdots+i_{N-h+2}=j \\ \ell_1+\cdots+\ell_{N-h+2}=\ell}} \prod_{r=1}^{N-h+2} \left| \partial_T^{\ell_r} \partial_k^{i_r} \big(\partial_\delta^r u(\bar{\delta},k,T)\big)^{j_r} \right| \\
&\leq C \sum_{j_1,\ldots,j_{N-h+2}} e^{(j_1+\cdots+j_{N-h+2})x_0} \\
&\quad \times M^\ell \big(M(T-t)\big)^{-\frac{j+2\ell}{2}+\frac{j_1+\cdots+j_{N-h+2}}{2}+\frac{j_1+2j_2+\cdots+(N-h+2)j_{N-h+2}}{2}} \\
&= C \sum_{j_1,\ldots,j_{N-h+2}} e^{hx_0} \big(M(T-t)\big)^{\frac{-j+h+N+1}{2}} \\
&= C e^{hx_0} M^\ell \big(M(T-t)\big)^{\frac{-j-2\ell+h+N+1}{2}}. \qquad (5.15)
\end{aligned}
$$

Combining Lemma 5.11 and (5.15) with (5.14), we obtain

$$
|\partial_k^m \sigma(k,T) - \partial_k^m \bar{\sigma}_N(k,T)| \leq C M^{q+\frac{1}{2}} \big(M(T-t)\big)^{\frac{N+1-m-2q}{2}} \sum_{h=1}^{N+1} e^{h(x_0-k)}.
$$

The statement then follows from the assumption $|x_0-k| \leq \lambda\sqrt{M(T-t)} \leq \lambda T_0$. $\square$

Acknowledgements The authors are grateful to Enrico Priola, Jian Wang and an anonymous referee for their valuable comments and suggestions to improve the quality of the paper.

Appendix A: Proof of Theorem 4.4

First observe that for any $z, \bar{z} \in \mathbb{R}^d$, $t < T$ and $m \leq N$, we have

$$
\begin{aligned}
&\partial_k^m u(t,z;T,k) - \partial_k^m \bar{u}_N^{(\bar{z})}(t,z;T,k) \\
&= \int_t^T \int_{\mathbb{R}^d} \Gamma(t,z;s,\zeta) \sum_{n=0}^{N} (\mathcal{A}_s - \bar{\mathcal{A}}_{s,n}^{(\bar{z})}) \partial_k^m u_{N-n}^{(\bar{z})}(s,\zeta;T,k) d\zeta ds, \qquad (A.1)
\end{aligned}
$$

where

$$\bar{\mathcal{A}}^{(\bar{z})}_{t,n} = \sum_{i=0}^{n} \mathcal{A}^{(\bar{z})}_{t,i}.$$

In fact, when $m = 0$, the identity (A.1) reduces to Lemma 6.23 in [32]. The general case easily follows by applying the operator ∂_k^m to (A.1) with $m = 0$ and then shifting ∂_k^m onto $u^{(\bar{z})}_{N-n}$. For clarity, we split the proof in two separate steps.

Step 1: Case $q = 0$ and $0 \le m \le N$. Let

$$\mathbb{T}^{a_\alpha(s,\cdot)}_{z,n}(\zeta) := \sum_{|\beta|\le n} \frac{D^\beta a_\alpha(s,z)}{\beta!}(\zeta - z)^\beta$$

be the nth order Taylor polynomial of the function $\zeta \mapsto a_\alpha(s,\zeta)$, centered at z. Setting $\bar{z} = z$ and by the definition of $(\mathcal{A}_{t,i})_{0\le i\le N}$, we obtain from (A.1) that

$$\partial_k^m u(t,z;T,k) - \partial_k^m \bar{u}_N(t,z;T,k) = \sum_{\substack{0\le n\le N\\ |\alpha|\le 2}} I_{n,\alpha},$$

where by Corollary D.2 and integrating by parts m times,

$$\begin{aligned}
I_{n,\alpha} &= \int_t^T \int_{\mathbb{R}^d} \Gamma(t,z;s,\zeta)\big(a_\alpha(s,\zeta) - \mathbb{T}^{a_\alpha(s,\cdot)}_{z,n}(\zeta)\big) D^\alpha_\zeta \partial_k^m u^{(z)}_{N-n}(s,\zeta;T,k)\,d\zeta\,ds\\
&= \sum_{\substack{|\gamma|\le N-n\\ 1\le j\le 3(N-n)}} \int_t^T \int_{\mathbb{R}^d} \Gamma(t,z;s,\zeta)\big(a_\alpha(s,\zeta) - \mathbb{T}^{a_\alpha(s,\cdot)}_{z,n}(\zeta)\big)(\zeta - z)^\gamma\\
&\qquad\qquad \times f^{(N-n,0,m,\alpha)}_{\gamma,j}(z;s,T)\partial_{\zeta_1}^{j+m+\alpha_1} u^{(z)}_0(s,\zeta;T,k)\,d\zeta\,ds\\
&= \sum_{\substack{|\gamma|\le N-n\\ 1\le j\le 3(N-n)}} \int_t^T \int_{\mathbb{R}^d} (-1)^m R^{\alpha,\gamma,m}_{n,1} R^{\alpha,\gamma,m,j}_{n,2}\,d\zeta\,ds
\end{aligned} \tag{A.2}$$

with

$$\begin{aligned}
R^{\alpha,\gamma,m}_{n,1} &= \partial^m_{\zeta_1}\Big(\Gamma(t,z;s,\zeta)\big(a_\alpha(s,\zeta) - \mathbb{T}^{a_\alpha(s,\cdot)}_{z,n}(\zeta)\big)(\zeta - z)^\gamma\Big),\\
R^{\alpha,\gamma,m,j}_{n,2} &= f^{(N-n,0,m,\alpha)}_{\gamma,j}(z;s,T)\partial^{j+\alpha_1}_{\zeta_1} u^{(z)}_0(s,\zeta;T,k).
\end{aligned}$$

Note that $R_{n,1}$ is well defined because $a_\alpha(s,\cdot) \in C^{N+1}(\mathbb{R}^d)$ by hypothesis and $m \le N$. Now on the one hand, by repeatedly applying the Leibniz rule, the mean value theorem and Lemma 4.5 with $c = 2$, we obtain

$$|R_{n,1}^{\alpha,\gamma,m}| \le CM\big(M(s-t)\big)^{\frac{n-m+|\gamma|+1}{2}}\Gamma_0\big(2M(s-t),\zeta - z\big). \tag{A.3}$$

On the other hand, by (D.4) and by Lemma C.3, we have

$$|R_{n,2}^{\alpha,\gamma,m,j}| \le Ce^{\zeta_1}\big(M(T-s)\big)^{\frac{N-n-|\gamma|-\alpha_1+1}{2}} \le Ce^{\zeta_1}\big(M(T-s)\big)^{\frac{N-n-|\gamma|-1}{2}} \tag{A.4}$$

since $\alpha_1 \le 2$. To conclude, it is enough to combine (A.4) and (A.3) with (A.2). In particular, by using

$$\int_{\mathbb{R}^d}\Gamma_0\big(2M(s-t),\zeta - z\big)e^{\zeta_1}d\zeta = e^{z_1+M(s-t)/2},$$

we get

$$\begin{aligned}|I_{n,\alpha}| &\le Ce^{z_1}M^{\frac{N-m+2}{2}}\int_t^T (s-t)^{\frac{n-m+|\gamma|+1}{2}}(T-s)^{\frac{N-n-|\gamma|-1}{2}}ds\\ &\le Ce^{z_1}\big(M(T-t)\big)^{\frac{N-m+2}{2}},\end{aligned}$$

where we used the identity

$$\int_t^T (T-s)^n(s-t)^j\,ds = \frac{\Gamma_E(j+1)\Gamma_E(n+1)}{\Gamma_E(j+n+2)}(T-t)^{j+n+1},$$

with Γ_E representing the Euler gamma function.

Step 2: Case $0 < m + 2q \le N$. We first prove that for any $\bar{m}, \bar{q} \in \mathbb{N}_0$ with $\bar{m} + 2\bar{q} \le N - 2$, one has

$$\begin{aligned}&\lim_{s\to T-}\int_{\mathbb{R}^d}\Gamma(t,z;s,\zeta)\sum_{n=0}^{N}(\mathcal{A}_s - \bar{\mathcal{A}}_{s,n}^{(z)})\partial_T^{\bar{q}}\partial_k^{\bar{m}}u_{N-n}^{(z)}(s,\zeta;T,k)d\zeta\\ &= \left(\frac{a_{11}(T,z)}{2}\right)^{\bar{q}}e^k\\ &\quad\times\int_{\mathbb{R}^{d-1}}(\partial_k^2+\partial_k)^{\bar{q}}(1+\partial_k)^{\bar{m}}\left(\Gamma(t,z;T,k,\eta)\big(a_{11}(T,k,\eta) - \mathbb{T}_{z,N}^{a_{11}(T,\cdot)}(k,\eta)\big)\right)d\eta.\end{aligned} \tag{A.5}$$

For $0 \le n \le N$, set

$$\begin{aligned}I_n(t,z) := \sum_{|\alpha|\le 2}\int_{\mathbb{R}^d}&\Gamma(t,z;s,\zeta)\big(a_\alpha(s,\zeta) - \mathbb{T}_{z,n}^{a_\alpha(s,\cdot)}(\zeta)\big)\\ &\times D_\zeta^\alpha\partial_T^{\bar{q}}\partial_k^{\bar{m}}u_{N-n}^{(z)}(s,\zeta;T,k)d\zeta.\end{aligned}$$

Now by applying (D.3) and integrating by parts $\bar{m}+2\bar{q}+2$ times with respect to ζ_1 (this is possible because $a_\alpha(s,\cdot)\in C^{N+1}(\mathbb{R}^d)$), we get for $n\le N-1$ that

$$I_n(t,z)=(-1)^{\bar{m}+2\bar{q}+2}\sum_{|\alpha|\le 2}\sum_{\substack{|\gamma|\le N-n\\ 1\le j\le 3(N-n)}}\int_{\mathbb{R}^d}\widetilde{R}_n^{\alpha,\gamma,\bar{q},\bar{m},j}R_n^{\alpha,\gamma,\bar{q},\bar{m},j}\,d\zeta$$

with

$$\begin{aligned}\widetilde{R}_n^{\alpha,\gamma,\bar{q},\bar{m},j}&=\partial_{\zeta_1}^{\bar{m}+2\bar{q}+2}\Big(\big(a_\alpha(s,\zeta)-\mathbb{T}_{z,n}^{a_\alpha(s,\cdot)}(\zeta)\big)\Gamma(t,z;s,\zeta)(\zeta-z)^\gamma\Big),\\ R_n^{\alpha,\gamma,\bar{q},\bar{m},j}&=f_{\gamma,j}^{(N-n,\bar{q},\bar{m},\alpha)}(z;s,T)\partial_{\zeta_1}^{j+\alpha_1-2}u_0^{(z)}(s,\zeta;T,k),\end{aligned}$$

and $f_{\gamma,j}^{(N-n,\bar{q},\bar{m},\alpha)}$ as in Corollary D.2. Moreover, by (D.4) and Lemma C.3, we get

$$|R_n^{\alpha,\gamma,\bar{q},\bar{m},j}|\le CM^{\bar{q}}e^{\zeta_1}\sqrt{M(T-s)}$$

and thus

$$\lim_{s\to T-}I_n(t,z)=0,\qquad 0\le n\le N-1,\ t<T,\ z\in\mathbb{R}^d. \tag{A.6}$$

On the other hand, by (C.6) and (D.9), we have by integrating by parts that

$$\begin{aligned}I_N(t,z)&:=\int_{\mathbb{R}^d}\Gamma(t,z;s,\zeta)(\mathcal{A}_s-\bar{\mathcal{A}}_{s,N}^{(z)})\partial_T^{\bar{q}}\partial_k^{\bar{m}}u_0^{(z)}(s,\zeta;T,k)d\zeta\\ &=\left(\frac{a_{11}(T,z)}{2}\right)^{\bar{q}}\int_{\mathbb{R}^d}\Gamma(t,z;s,\zeta)\big(a_{11}(s,\zeta)-\mathbb{T}_{z,N}^{a_{11}(s,\cdot)}(\zeta)\big)\\ &\qquad\times(\partial_{\zeta_1}^2-\partial_{\zeta_1})^{\bar{q}+1}(1-\partial_{\zeta_1})^{\bar{m}}u_0^{(z)}(s,\zeta;T,k)d\zeta\\ &=\left(\frac{a_{11}(T,z)}{2}\right)^{\bar{q}}\\ &\quad\times\int_{\mathbb{R}^d}(\partial_{\zeta_1}^2-\partial_{\zeta_1})u_0^{(z)}(s,\zeta;T,k)\\ &\qquad\times(\partial_{\zeta_1}^2+\partial_{\zeta_1})^{\bar{q}}(1+\partial_{\zeta_1})^{\bar{m}}\Big(\Gamma(t,z;s,\zeta)\big(a_{11}(s,\zeta)-\mathbb{T}_{z,N}^{a_{11}(s,\cdot)}(\zeta)\big)\Big)d\zeta.\end{aligned}$$

From (3.9), (3.10) and (C.5), we have

$$(\partial_{\zeta_1}^2-\partial_{\zeta_1})u_0^{(z)}(s,\zeta;T,k)=e^k\Gamma_0\bigg(\int_s^T a_{11}(r,z)dr,\zeta_1-\frac{\int_s^T a_{11}(r,z)dr}{2}-k\bigg),$$

where Γ_0 denotes the Gaussian density in (4.3) with $d=1$. Noting that

$$\Gamma_0\bigg(\int_s^T a_{11}(r,z)dr, \zeta_1 - \frac{\int_s^T a_{11}(r,z)dr}{2} - k\bigg) \longrightarrow \delta_k \qquad \text{as } s\to T-,$$

we obtain

$$\begin{aligned}
&\lim_{s\to T-} I_N(t,z)\\
&= \left(\frac{a_{11}(T,z)}{2}\right)^{\bar q} e^k\\
&\quad\times \int\limits_{\mathbb{R}^{d-1}} (\partial_k^2+\partial_k)^{\bar q}(1+\partial_k)^{\bar m}\Big(\Gamma(t,z;T,k,\eta)\big(a_{11}(T,k,\eta)-\mathbb{T}_{z,N}^{a_{11}(T,\cdot)}(k,\eta)\big)\Big)\,d\eta.
\end{aligned} \tag{A.7}$$

Finally, (A.6) and (A.7) yield (A.5).

We now prove (4.2). By repeatedly applying the Leibniz rule on (A.1) and (A.5), we get

$$\begin{aligned}
&\partial_T^q\partial_k^m(u-\bar u_N)(t,x,y;T,k)\\
&= \int_t^T \int\limits_{\mathbb{R}^d} \Gamma(t,z;s,\zeta)\sum_{n=0}^{N}(\mathcal{A}_s-\bar{\mathcal{A}}_{s,n}^{(\bar z)})\partial_T^q\partial_k^m u_{N-n}^{(\bar z)}(s,\zeta;T,k)d\zeta ds + \sum_{i=0}^{q-1} J_i
\end{aligned}$$

with

$$J_i = \partial_T^{q-1-i}\left(\left(\frac{a_{11}(T,z)}{2}\right)^i e^k \int\limits_{\mathbb{R}^{d-1}} (\partial_k^2+\partial_k)^i(1+\partial_k)^m\hat\Gamma(k,\eta)d\eta\right)$$

and

$$\hat\Gamma(k,\eta) = \Gamma(t,z;T,k,\eta)\big(a_{11}(T,k,\eta)-\mathbb{T}_{z,N}^{a_{11}(T,\cdot)}(k,\eta)\big).$$

Now, by proceeding as in Step 1, it is easy to show that

$$\begin{aligned}
&\left|\int_t^T \int\limits_{\mathbb{R}^d} \Gamma(t,z;s,\zeta)\sum_{n=0}^{N}(\mathcal{A}_s-\bar{\mathcal{A}}_{s,n}^{(\bar z)})\partial_T^q\partial_k^m u_{N-n}^{(\bar z)}(s,\zeta;T,k)d\zeta ds\right|\\
&\le Ce^x M^q\big(M(T-t)\big)^{\frac{N-m-2q+2}{2}}.
\end{aligned}$$

Analogously, by repeatedly applying the Leibniz rule along with Faà di Bruno's formula (Proposition E.1) and Lemma 4.5, and by using that

$$e^k \int_{\mathbb{R}^{d-1}} \Gamma_0\big(2M(T-t), x-k, y-\eta\big) d\eta = \frac{e^k}{\sqrt{4\pi M(T-t)}} e^{-\frac{(k-x)^2}{4M(T-t)}}$$
$$\leq \frac{Ce^x}{\sqrt{M(T-t)}}$$

with Γ_0 as in (4.3), one can also show

$$|J_i| \leq Ce^x M^q \big(M(T-t)\big)^{\frac{N-m-2q+2}{2}}, \qquad 0 \leq i \leq q-1,$$

which concludes the proof. □

Appendix B: Proof of Lemmas 5.10 and 5.11

Proof of Lemma 5.10 The case $n = m = 0$ has been already proved in (5.8). To prove the general case, we proceed by induction on m and n.

Step 1: Case $m = q = 0$**.** By (C.9) and using $|x_0 - k| \leq \lambda\sqrt{M(T-t)}$, we have

$$\partial_\sigma u^{\mathrm{BS}}(\sigma, k, T) \geq \frac{e^k\sqrt{T-t}}{\sqrt{2\pi}} \exp\left(-\frac{\lambda^2 M}{2\sigma^2} - \frac{\sigma^2(T-t)}{8} - \frac{\lambda\sqrt{M(T-t)}}{2}\right)$$
$$\geq \frac{e^k\sqrt{T-t}}{\sqrt{2\pi}} \exp\left(-\frac{\lambda^2 M}{2\sigma^2} - \frac{\sigma^2 T_0}{8} - \frac{\lambda\sqrt{M_0 T_0}}{2}\right),$$

which by (5.8) implies

$$(\partial_1 u^{\mathrm{BS}})\Big(\sigma^{\mathrm{BS}}\big(u(\delta, k, T), k, T\big)\Big) \geq \frac{e^k\sqrt{T-t}}{\sqrt{2\pi}} \exp\left(-\frac{\lambda^2}{2\varepsilon} - \frac{M_0 T_0}{2} - \frac{\lambda\sqrt{M_0 T_0}}{2}\right). \tag{B.1}$$

Therefore, we obtain

$$0 < (\partial_1 \sigma^{\mathrm{BS}})\big(u(\delta, k, T), k, T\big) = \frac{1}{(\partial_1 u^{\mathrm{BS}})(\sigma^{\mathrm{BS}}(u(\delta, k, T), k, T))} \leq \frac{C}{e^k\sqrt{T-t}},$$

which is (5.10) for $m = 0$ and $n = 1$.

We now fix $\bar{n} \in \mathbb{N}$, assume (5.10) to hold true for any $n \in \mathbb{N}_0$ with $n \leq \bar{n}$ and prove it for $\bar{n}+1$. Differentiating the identity $u = u^{\mathrm{BS}}(\sigma^{\mathrm{BS}}(u, k, T), k, T)$ and applying the univariate version of Faà di Bruno's formula (E.4), we obtain

$$\partial_1^{\bar{n}+1} \sigma^{\mathrm{BS}}(u, k, T) = -\sum_{h=2}^{\bar{n}+1} \frac{(\partial_1^h u^{\mathrm{BS}})(\sigma^{\mathrm{BS}}(u, k, T), k, T)}{(\partial_1 u^{\mathrm{BS}})(\sigma^{\mathrm{BS}}(u, k, T), k, T)}$$
$$\times \mathbf{B}_{\bar{n}+1,h}\big(\partial_1 \sigma^{\mathrm{BS}}(u, k, T), \ldots, \partial_1^{\bar{n}-h+2} \sigma^{\mathrm{BS}}(u, k, T)\big).$$

Now by (B.1), Lemma C.5 and recalling the estimate of Lemma 5.8 for $u = u(\delta, k, T)$, we get

$$\left| \frac{(\partial_1^h u^{\mathrm{BS}})(\sigma^{\mathrm{BS}}(u(\delta,k,T),k,T),k,T)}{(\partial_1 u^{\mathrm{BS}})(\sigma^{\mathrm{BS}}(u(\delta,k,T),k,T),k,T)} \right| \leq C M^{-\frac{h-1}{2}}.$$

Moreover, for any $h = 2, \dots, \bar{n}+1$, we have by (E.5) and the inductive hypothesis that

$$\begin{aligned}
&\left| \mathbf{B}_{\bar{n}+1,h}\big(\partial_1 \sigma^{\mathrm{BS}}(u,k,T), \dots, \partial_1^{\bar{n}-h+2}\sigma^{\mathrm{BS}}(u,k,T)\big)\big|_{u=u(\delta,k)} \right| \\
&\leq C \sum_{j_1,\dots,j_{\bar{n}-h+2}} \left|(\partial_1\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big)\right|^{j_1} \cdots \left|(\partial_1^{\bar{n}-h+2}\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big)\right|^{j_{\bar{n}-h+2}} \\
&\leq C \sum_{j_1,\dots,j_{\bar{n}-h+2}} \sqrt{M}\big(e^k\sqrt{M(T-t)}\big)^{-j_1} \cdots \sqrt{M}\big(e^k\sqrt{M(T-t)}\big)^{-(\bar{n}-h+2)j_{\bar{n}-h+2}} \\
&\leq C M^{\frac{h}{2}}\big(e^k\sqrt{M(T-t)}\big)^{-\bar{n}-1},
\end{aligned}$$

where the last inequality follows from the identities (E.6) in Appendix E. This concludes the proof of (5.10) with $m = 0$.

Step 2: Case $q = 0$. We proceed by induction on m. The subcase $m = 0$ has already been proved in Step 1. Now fix $\bar{m} \in \mathbb{N}$, assume (5.10) to hold for any $n, m \in \mathbb{N}_0$, with $m \leq \bar{m}$ and prove it for $m = \bar{m}+1$ and $n \in \mathbb{N}_0$. First note that differentiating with respect to k the identity

$$\sigma = \sigma^{\mathrm{BS}}\big(u^{\mathrm{BS}}(\sigma,k,T),k,T\big), \qquad \sigma > 0, \tag{B.2}$$

we get

$$(\partial_2\sigma^{\mathrm{BS}})\big(u^{\mathrm{BS}}(\sigma,k,T),k,T\big) = -(\partial_1\sigma^{\mathrm{BS}})\big(u^{\mathrm{BS}}(\sigma,k,T),k,T\big)\partial_2 u^{\mathrm{BS}}(\sigma,k,T),$$

or equivalently, setting $u = u^{\mathrm{BS}}(\sigma,k,T)$, that is, $\sigma = \sigma^{\mathrm{BS}}(u,k,T)$,

$$\partial_2\sigma^{\mathrm{BS}}(u,k,T) = -\partial_1\sigma^{\mathrm{BS}}(u,k,T)(\partial_2 u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big) \tag{B.3}$$

for $u \in ((e^{x_0}-e^k)^+, e_0^x)$. Fix $n \in \mathbb{N}_0$; differentiating (B.3) n times with respect to u and $\bar{m}$ times with respect to k, we get

$$\begin{aligned}
\partial_1^n\partial_2^{\bar{m}+1}\sigma^{\mathrm{BS}}(u,k,T) &= -\frac{d^{n+\bar{m}}}{du^n dk^{\bar{m}}}\Big(\partial_1\sigma^{\mathrm{BS}}(u,k,T)(\partial_2 u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big)\Big) \\
&= -\sum_{i=0}^{n}\sum_{j=0}^{\bar{m}}\binom{n}{i}\binom{\bar{m}}{j}\big(\partial_1^{n+1-i}\partial_2^{\bar{m}-j}\sigma^{\mathrm{BS}}(u,k,T)\big) \\
&\qquad \times \frac{d^{i+j}}{du^i dk^j}(\partial_2 u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big).
\end{aligned} \tag{B.4}$$

Now by the inductive hypothesis, for any $i, j, n \in \mathbb{N}_0$ with $i \leq n$ and $j \leq \bar{m}$, we have

$$\left|(\partial_1^{n+1-i}\partial_2^{\bar{m}-j}\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big)\right| \leq \frac{C\sqrt{M}e^{-(n+1-i)k}}{(M(T-t))^{\frac{n+1-i+\bar{m}-j}{2}}}. \tag{B.5}$$

The proof will be concluded once we show that

$$\left|\frac{d^{i+j}}{du^i dk^j}(\partial_2 u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big)\big|_{u=u(\delta,k,T)}\right| \leq C\big(M(T-t)\big)^{-\frac{i+j}{2}}e^{-(i-1)k}. \tag{B.6}$$

Indeed (B.6), combined with (B.5) and (B.4), yields (5.10) for $\bar{m}+1$.

More generally, we prove that for any $i, j, \gamma_1, \gamma_2, \gamma_3 \in \mathbb{N}_0$ with $\gamma_1+\gamma_2+\gamma_3 > 0$ and $j \leq \bar{m}$ (here $\bar{m}$ is fixed in the inductive hypothesis at the beginning of Step 2), we have

$$\begin{aligned}&\left|\frac{d^{i+j}}{du^i dk^j}(\partial_1^{\gamma_1}\partial_2^{\gamma_2}\partial_3^{\gamma_3}u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big)\big|_{u=u(\delta,k,T)}\right|\\ &\leq CM^{\gamma_3-\frac{\gamma_1}{2}}\big(M(T-t)\big)^{\frac{1-i-j-\gamma_2-2\gamma_3}{2}}e^{(1-i)k}.\end{aligned} \tag{B.7}$$

We prove (B.7) by using another inductive argument on j.

Step 2-a: Case $j=0$**.** By the univariate version of Faà di Bruno's formula (E.4), we have for any $i, \gamma_1, \gamma_2 \in \mathbb{N}_0$ that

$$\begin{aligned}&\frac{d^i}{du^i}(\partial_1^{\gamma_1}\partial_2^{\gamma_2}\partial_3^{\gamma_3}u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big)\\ &=\sum_{h=1}^{i}(\partial_1^{h+\gamma_1}\partial_2^{\gamma_2}\partial_3^{\gamma_3}u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big)\\ &\qquad\times\mathbf{B}_{i,h}\big(\partial_1\sigma^{\mathrm{BS}}(u,k,T),\partial_1^2\sigma^{\mathrm{BS}}(u,k,T),\ldots,\partial_1^{i-h+1}\sigma^{\mathrm{BS}}(u,k,T)\big).\end{aligned} \tag{B.8}$$

By Lemmas C.5 and 5.8, using that $\gamma_1+\gamma_2+\gamma_3 > 0$, we have

$$\left|(\partial_1^{h+\gamma_1}\partial_2^{\gamma_2}\partial_3^{\gamma_3}u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big)\big|_{u=u(\delta,k,T)}\right| \leq \frac{Ce^k M^{\gamma_3-\frac{h+\gamma_1}{2}}}{(M(T-t))^{\frac{\gamma_2+2\gamma_3-1}{2}}}. \tag{B.9}$$

Moreover, by (5.10) with $m=0$ (already proved in Step 1) and by the relations (E.6), we have

$$\begin{aligned}&\left|\mathbf{B}_{i,h}\big(\partial_1\sigma^{\mathrm{BS}}(u,k,T),\partial_1^2\sigma^{\mathrm{BS}}(u,k,T),\ldots,\partial_1^{i-h+1}\sigma^{\mathrm{BS}}(u,k,T)\big)\big|_{u=u(\delta,k,T)}\right|\\ &\leq CM^{\frac{h}{2}}\big(M(T-t)\big)^{-\frac{i}{2}}e^{-ik},\end{aligned}$$

which combined with (B.9) and (B.8) proves (B.7) for $j=0$ and any $i, \gamma_1, \gamma_2 \in \mathbb{N}_0$ with $\gamma_1+\gamma_2+\gamma_3 > 0$.

Step 2-b: Case $1 \le j \le \bar{m}$**.** Fix $j_0 \in \mathbb{N}$ with $j_0 \le \bar{m} - 1$; we assume (B.7) to hold for any $i, \gamma_1, \gamma_2, \gamma_3 \in \mathbb{N}_0$ with $\gamma_1 + \gamma_2 + \gamma_3 > 0$ and $0 \le j \le j_0$ and prove it for $i, \gamma_1, \gamma_2, \gamma_3 \in \mathbb{N}_0$ with $\gamma_1 + \gamma_2 + \gamma_3 > 0$ and $j = j_0 + 1$. We have

$$
\begin{aligned}
&\frac{d^{i+j_0+1}}{du^i dk^{j_0+1}} (\partial_1^{\gamma_1} \partial_2^{\gamma_2} \partial_3^{\gamma_3} u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big) \\
&= \frac{d^{i+j_0}}{du^i dk^{j_0}} \Big((\partial_1^{1+\gamma_1} \partial_2^{\gamma_2} \partial_3^{\gamma_3} u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big) \partial_2 \sigma^{\mathrm{BS}}(u,k,T) \\
&\qquad\qquad + (\partial_1^{\gamma_1} \partial_2^{1+\gamma_2} \partial_3^{\gamma_3} u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big) \Big) \\
&= \sum_{h=0}^{i} \sum_{q=0}^{j_0} \binom{i}{h} \binom{j_0}{q} \Big(\frac{d^{h+q}}{du^h dk^q} (\partial_1^{1+\gamma_1} \partial_2^{\gamma_2} \partial_3^{\gamma_3} u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big) \Big) \\
&\qquad \times \partial_1^{i-h} \partial_2^{j_0-q+1} \sigma^{\mathrm{BS}}(u,k,T) \\
&\quad + \frac{d^{i+j_0}}{du^i dk^{j_0}} (\partial_1^{\gamma_1} \partial_2^{1+\gamma_2} \partial_3^{\gamma_3} u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big).
\end{aligned} \tag{B.10}
$$

By the inductive hypothesis, we have

$$
\begin{aligned}
&\left| \frac{d^{h+q}}{du^h dk^q} (\partial_1^{1+\gamma_1} \partial_2^{\gamma_2} \partial_3^{\gamma_3} u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big)\big|_{u=u(\delta,k,T)} \right| \\
&\le C M^{\gamma_3 - \frac{\gamma_1+1}{2}} \big(M(T-t)\big)^{-\frac{h+q+\gamma_2+2\gamma_3-1}{2}} e^{-(h-1)k}
\end{aligned}
$$

and

$$
\begin{aligned}
&\left| \frac{d^{i+j_0}}{du^i dk^{j_0}} (\partial_1^{\gamma_1} \partial_2^{1+\gamma_2} \partial_3^{\gamma_3} u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big)\big|_{u=u(\delta,k,T)} \right| \\
&\le C M^{\gamma_3 - \frac{\gamma_1}{2}} \big(M(T-t)\big)^{-\frac{i+j_0+\gamma_2+2\gamma_3}{2}} e^{-(i-1)k}.
\end{aligned}
$$

Now recall that we are assuming, by the inductive hypothesis, that (5.10) holds for any $n \in \mathbb{N}_0$ and $m \le \bar{m}$; thus, since $j_0 - q + 1 \le \bar{m}$ by assumption, we get

$$
\left| \partial_1^{i-h} \partial_2^{j_0-q+1} \sigma^{\mathrm{BS}}(u,k,T)|_{u=u(\delta,k,T)} \right| \le C M^{\frac{1}{2}} \big(M(T-t)\big)^{-\frac{i-h+j_0-q+1}{2}} e^{-(i-h)k}.
$$

The last three estimates combined with (B.10) yield (B.7) for $j = j_0 + 1$.

Step 3: Case $q \in \mathbb{N}$**.** This is analogous to Step 2. For simplicity, we only prove the case $q = 1$. By the identity (B.2), we get

$$
(\partial_3 \sigma^{\mathrm{BS}})\big(u^{\mathrm{BS}}(\sigma,k,T),k,T\big) = -(\partial_1 \sigma^{\mathrm{BS}})\big(u^{\mathrm{BS}}(\sigma,k,T),k,T\big) \partial_3 u^{\mathrm{BS}}(\sigma,k,T),
$$

or equivalently, setting $u = u^{\mathrm{BS}}(\sigma,k,T)$, that is, $\sigma = \sigma^{\mathrm{BS}}(u,k,T)$,

$$
\partial_3 \sigma^{\mathrm{BS}}(u,k,T) = -\partial_1 \sigma^{\mathrm{BS}}(u,k,T) (\partial_3 u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big) \tag{B.11}
$$

for $u \in ((e^{x_0} - e^k)^+, e_0^x)$. Fix $n, m \in \mathbb{N}_0$; differentiating (B.11) n and m times with respect to u and k, respectively, and once with respect to T, we get

$$\begin{aligned}
\partial_1^n \partial_2^m \partial_3 \sigma^{\mathrm{BS}}(u,k,T) &= -\frac{d^{n+m}}{du^n dk^m}\Big(\partial_1 \sigma^{\mathrm{BS}}(u,k,T)(\partial_3 u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big)\Big) \\
&= -\sum_{i=0}^{n}\sum_{j=0}^{m}\binom{n}{i}\binom{m}{j}\big(\partial_1^{n+1-i}\partial_2^{m-j}\sigma^{\mathrm{BS}}(u,k,T)\big) \\
&\qquad \times \frac{d^{i+j}}{du^i dk^j}(\partial_3 u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big). \qquad \text{(B.12)}
\end{aligned}$$

Now by (5.10) with $q = 0$, for any $i, j, n \in \mathbb{N}_0$ with $i \le n$ and $j \le m$, we have

$$\big|(\partial_1^{n+1-i}\partial_2^{m-j}\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big)\big| \le \frac{C M^{\frac{1}{2}} e^{-(n+1-i)k}}{(M(T-t))^{\frac{n+1-i+m-j}{2}}}, \qquad \text{(B.13)}$$

whereas by (B.7), we obtain

$$\left|\frac{d^{i+j}}{du^i dk^j}(\partial_3 u^{\mathrm{BS}})\big(\sigma^{\mathrm{BS}}(u,k,T),k,T\big)\big|_{u=u(\delta,k,T)}\right| \le \frac{C M e^{-(i-1)k}}{(M(T-t))^{\frac{i+j+1}{2}}}. \qquad \text{(B.14)}$$

Finally, (B.13) and (B.14) combined with (B.12) prove (5.10) for $q = 1$. □

Remark B.1 The inductive argument of the previous proof shows that the estimate (B.7) is valid for any $i, j, \gamma_1, \gamma_2, \gamma_3 \in \mathbb{N}_0$ with $\gamma_1 + \gamma_2 + \gamma_3 > 0$ and any $\delta \in [0, 1]$, $0 \le t < T < T_0$ and $k \in \mathbb{R}$ such that $T - t \le \tau_0$ and $|x_0 - k| \le \lambda\sqrt{M(T-t)}$. In this case, the constant C in (B.7) also depends on i, j, γ_1, γ_2 and γ_3.

Proof of Lemma (5.11) For simplicity, we split the proof in two separate steps.

Step 1: Case $q = 0$. By the bivariate version of Faà di Bruno's formula (see Proposition E.1), we obtain by exploiting the first relation in (E.6) that

$$\begin{aligned}
&\frac{d^m}{dk^m}(\partial_1^n \sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big) \\
&= \sum_{h=1}^{m}(\nabla^h \partial_1^n \sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big) \\
&\qquad * \mathbf{B}_{m,h}\left(\begin{pmatrix}\partial_k u(\delta,k,T)\\ 1\end{pmatrix}, \begin{pmatrix}\partial_k^2 u(\delta,k,T)\\ 0\end{pmatrix}, \dots, \begin{pmatrix}\partial_k^{m-h+1} u(\delta,k,T) 0\end{pmatrix}\right), \\
&= \sum_{h=1}^{m}\sum_{j_1=0}^{h} g_{h,j_1}(\delta,k,T)(\nabla^{j_1}\partial_1^{n+h-j_1}\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big) \\
&\qquad * \begin{pmatrix}\partial_k u(\delta,k,T)\\ 1\end{pmatrix}^{j_1}, \qquad \text{(B.15)}
\end{aligned}$$

where $*$ denotes the tensorial scalar product (see (E.2)) and

$$g_{h,j_1}(\delta,k,T)=\sum_{j_2,\ldots,j_{m-h+1}} c^{m,h}_{j_1,\ldots,j_{m-h+1}} \prod_{i=2}^{m-h+1} \left(\partial_k^i u(\delta,k,T)\right)^{j_i} \tag{B.16}$$

for some constants $c^{m,h}_{j_1,\ldots,j_{m-h+1}}$ and the sum in (B.16) is taken over all sequences $j_2,\ldots,j_{m-h+1}$ of nonnegative integers verifying the identities in (E.6). Now, by the estimate (5.7) and the relations (E.6), we obtain

$$|g_{h,j_1}(\delta,k,T)| \le Ce^{(h-j_1)x_0}\big(M(T-t)\big)^{-\frac{m-h}{2}}. \tag{B.17}$$

Moreover, we have

$$\left|(\nabla^{j_1}\partial_1^{n+h-j_1}\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big) * \begin{pmatrix}\partial_k u(\delta,k,T)\\ 1\end{pmatrix}^{j_1}\right|$$
$$\le C\sum_{q=0}^{j_1}\left|(\partial_1^{n+h-q}\partial_2^{q}\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big)\right|\left|\big(\partial_k u(\delta,k,T)\big)^{j_1-q}\right|,$$

and therefore by Lemma 5.10 and the estimate (5.7), we get

$$\left|(\nabla^{j_1}\partial_1^{n+h-j_1}\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big) * \begin{pmatrix}\partial_k u(\delta,k,T)\\ 1\end{pmatrix}^{j_1}\right|$$
$$\le Ce^{-(n+h-q)k+(j_1-q)x_0}\sqrt{M}\big(M(T-t)\big)^{-\frac{n+h}{2}}. \tag{B.18}$$

Finally, (5.11) follows by combining (B.17) and (B.18) with (B.15) and observing that

$$e^{(h-q)(x_0-k)} \le e^{m|x_0-k|} \le e^{m\lambda\sqrt{M(T-t)}},$$

since $|x_0-k|\le\lambda\sqrt{M(T-t)}$.

Step 2: Case $q\in\mathbb{N}$. This is analogous to Step 1. For simplicity, we only prove the case $q=1$. The Leibniz rule yields

$$\frac{d^m}{dk^m}\frac{d}{dT}(\partial_1^n\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big)$$
$$=\sum_{i=0}^{m}\binom{m}{i}\big(\partial_k^{m-i}\partial_T u(\delta,k,T)\big)\frac{d^i}{dk^i}(\partial_1^{n+1}\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big)$$
$$+\frac{d^m}{dk^m}(\partial_1^n\partial_3\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big). \tag{B.19}$$

By (5.11) with $q=0$, by (5.7) and by using that $|x_0-k|\le\lambda(T-t)$, we get

$$\left|\big(\partial_k^{m-i}\partial_T u(\delta,k,T)\big)\frac{d^i}{dk^i}(\partial_1^{n+1}\sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big)\right| \le \frac{CM^{1+\frac12}e^{-nk}}{(M(T-t))^{\frac{m+2+n}{2}}}. \tag{B.20}$$

On the other hand, by proceeding exactly as in Step 1, one can show that

$$\left|\frac{d^m}{dk^m}(\partial_1^n \partial_3 \sigma^{\mathrm{BS}})\big(u(\delta,k,T),k,T\big)\right| \leq \frac{C M^{1+\frac{1}{2}} e^{-nk}}{(M(T-t))^{\frac{m+2+n}{2}}},$$

which combined with (B.20) and (B.19) proves (5.11) for $q=1$. □

Appendix C: Short-time/small-noise estimates in the Black–Scholes model

We collect here the short-time estimates for the sensitivities with respect to σ, x and k of the Black–Scholes function $u^{\mathrm{BS}}(\sigma) = u^{\mathrm{BS}}(\sigma;\tau,x,k)$ needed to prove the results of Sect. 5. In this appendix, Γ_0 denotes the Gaussian density in (4.3) with $d=1$.

Lemma C.1 *For any $n \in \mathbb{N}_0$ and $c > 1$, we have*

$$\left(\frac{|x|}{\sqrt{t}}\right)^n \Gamma_0(t,x) \leq \sqrt{c}\left(\frac{cn}{(c-1)\sqrt{e}}\right)^{\frac{n}{2}} \Gamma_0(ct,x), \qquad t \in \mathbb{R}_{++}, x \in \mathbb{R}.$$

Proof Set $z = \frac{|x|}{\sqrt{t}}$. For any $c > 1$, we have

$$\left(\frac{|x|}{\sqrt{t}}\right)^n \Gamma_0(t,x) = \frac{z^n}{\sqrt{2\pi t}} \exp\left(-\frac{z^2}{2}\right) = \sqrt{c}\, g(z)\Gamma_0(ct,x)$$

with

$$g(z) = z^n \exp\left(-\frac{z^2}{2}\left(1-\frac{1}{c}\right)\right), \qquad z \geq 0.$$

The claim now follows by observing that g attains a global maximum at $z_n = \sqrt{\frac{cn}{c-1}}$ and that

$$g(z_n) = e^{-\frac{n}{2}}\left(\frac{cn}{c-1}\right)^{n/2}.$$

□

In what follows, we make use of the representation of the Black–Scholes price in terms of the Gaussian density Γ_0 in (4.3), i.e.,

$$u^{\mathrm{BS}}(\sigma) = u^{\mathrm{BS}}(\sigma;\tau,x,k) = \int_k^{+\infty} \Gamma_0\left(\sigma^2\tau, x - \frac{\sigma^2\tau}{2} - y\right)(e^y - e^k)dy, \qquad \text{(C.1)}$$

and of the family of Hermite polynomials defined as

$$\mathbf{H}_n(x) := e^{x^2}\partial_x^n e^{-x^2}, \qquad n \in \mathbb{N}_0. \qquad \text{(C.2)}$$

Lemma C.2 *For any $n \in \mathbb{N}_0$ and $c > 1$, we have*

$$|\partial_x^n \Gamma_0(t,x)| \leq C\, t^{-\frac{n}{2}} \Gamma_0(ct,x), \qquad t \in \mathbb{R}_{++}, x \in \mathbb{R}, \tag{C.3}$$

where C is a positive constant only dependent on n and c.

Proof By the definition (4.3), we have

$$\partial_x^n \Gamma_0(t,x) = t^{-\frac{n}{2}} \mathbf{H}_n\left(\frac{x}{\sqrt{2t}}\right) \Gamma_0(t,x),$$

and thus the statement easily follows from Lemma C.1. □

Lemma C.3 *For any $m, n \in \mathbb{N}_0$ and $M > 0$, we have*

$$|\partial_x^n \partial_k^m u^{\mathrm{BS}}(\sigma;\tau,x,k)| \leq C e^x (\sigma\sqrt{\tau})^{(1-m-n)\wedge 0} \tag{C.4}$$

for $x, k \in \mathbb{R}$ and $0 < \sigma\sqrt{\tau} \leq M$, where C is a positive constant only dependent on m, n and M.

Proof Throughout this proof, we denote by C any generic constant that depends at most on m, n and M. We first prove the statement for $m = 0$. If also $n = 0$, the thesis easily follows by writing u^{BS} as an expectation. If $n \geq 1$, then by (C.1), we have from $\partial_x \Gamma_0 = -\partial_y \Gamma_0$ and integrating by parts that

$$\begin{aligned}
\partial_x^n u^{\mathrm{BS}}(\sigma;\tau,x,k) &= \int_k^{+\infty} \partial_x^n \Gamma_0\left(\sigma^2\tau, x - \frac{\sigma^2\tau}{2} - y\right)(e^y - e^k)dy \\
&= \int_k^{\infty} \partial_x^{n-1} \Gamma_0\left(\sigma^2\tau, x - \frac{\sigma^2\tau}{2} - y\right) e^y dy.
\end{aligned} \tag{C.5}$$

Thus, by the Gaussian estimate (C.3) with $c = 2$, we obtain

$$\begin{aligned}
|\partial_x^n u^{\mathrm{BS}}(\sigma;\tau,x,k)| &\leq C(\sigma\sqrt{\tau})^{-n+1} \int_{\mathbb{R}} \Gamma_0\left(2\sigma^2\tau, x - \frac{\sigma^2\tau}{2} - y\right) e^y dy \\
&= C\, e^{x + \frac{\sigma^2\tau}{2}} (\sigma\sqrt{\tau})^{-n+1}
\end{aligned}$$

which proves the statement for $m = 0$. The case $m \geq 1$ now trivially follows from the identity

$$\partial_k u^{\mathrm{BS}}(\sigma;\tau,x,k) = u^{\mathrm{BS}}(\sigma;\tau,x,k) - \partial_x u^{\mathrm{BS}}(\sigma;\tau,x,k), \tag{C.6}$$

along with (C.4) with $m = 0$. □

Proposition C.4 *Fix* (t, T, k, σ) *and let* $\zeta = \frac{x-k-\frac{\sigma^2\tau}{2}}{\sigma\sqrt{2\tau}}$ *and* $\tau = T - t$. *Then for any* $n \geq 2$, *we have*

$$\frac{\partial_\sigma^n u^{\mathrm{BS}}(\sigma)}{\partial_\sigma u^{\mathrm{BS}}(\sigma)} = \sum_{q=0}^{\lfloor n/2 \rfloor} \sum_{p=0}^{n-q-1} c_{n,n-2q}\sigma^{n-2q-1}\tau^{n-q-1}$$

$$\times \binom{n-q-1}{p} \left(\frac{1}{\sigma\sqrt{2\tau}}\right)^{p+n-q-1} \mathbf{H}_{p+n-q-1}(\zeta),$$

where the coefficients $c_{n,n-2k}$ *are defined recursively by*

$$c_{n,n} = 1 \quad \textit{and} \quad c_{n,n-2q} = (n-2q+1)c_{n-1,n-2q+1} + c_{n-1,n-2q-1}$$

for $q \in \{1, 2, \ldots, \lfloor n/2 \rfloor\}$.

Proof See [33, Proposition 3.5]. □

Lemma C.5 *For any* $m, q, n \in \mathbb{N}_0$ *with* $m + q + n > 0$, *we have*

$$|\partial_\sigma^n \partial_\tau^q \partial_k^m u^{\mathrm{BS}}(\sigma; \tau, x, k)| \leq C e^k \sigma^{-n+2q} (\sigma\sqrt{\tau})^{1-m-2q} \tag{C.7}$$

for $x, k \in \mathbb{R}$ *and* $0 < \sigma\sqrt{\tau} \leq M$, *where* C *is a positive constant only dependent on* m, q, n *and* M. *If* $q = 0$, *then* C *is independent of* M.

Proof We split the proof in three steps.

Step 1: Case $q = n = 0$. Here we denote by C any generic constant that depends at most on m. For any $m \in \mathbb{N}$, we have by (C.1) that

$$\begin{aligned} \partial_k^m u^{\mathrm{BS}}(\sigma; \tau, x, k) &= \partial_k^{m-1}\left(e^k \int_k^\infty \Gamma_0\left(\sigma^2\tau, y - x + \frac{\sigma^2\tau}{2}\right)dy\right) \\ &= \sum_{i=0}^{m-1} \binom{m-1}{i} e^k \partial_k^i \int_k^\infty \Gamma_0\left(\sigma^2\tau, y - x + \frac{\sigma^2\tau}{2}\right)dy. \end{aligned} \tag{C.8}$$

Now $\int_k^\infty \Gamma_0(\sigma^2\tau, y - x + \frac{\sigma^2\tau}{2})dy \in (0, 1)$ and for $i \geq 1$, we have

$$\partial_k^i \int_k^\infty \Gamma_0\left(\sigma^2\tau, y - x + \frac{\sigma^2\tau}{2}\right)dy = -\partial_k^{i-1}\Gamma_0\left(\sigma^2\tau, k - x + \frac{\sigma^2\tau}{2}\right).$$

Thus by applying the Gaussian estimate (C.3) with $c = 2$, we obtain

$$\begin{aligned} \left|\partial_k^i \int_k^\infty \Gamma_0\left(\sigma^2\tau, y - x + \frac{\sigma^2\tau}{2}\right)dy\right| &\leq C(\sigma\sqrt{2\tau})^{-i+1}\Gamma_0\left(2\sigma^2\tau, k - x + \frac{\sigma^2\tau}{2}\right) \\ &\leq C(\sigma\sqrt{\tau})^{-i}, \end{aligned}$$

which combined with (C.8) proves (C.7).

Step 2: Case $q=0, n\geq 1$. Here we denote by C any generic constant that depends at most on m and n. A direct computation shows that

$$\partial_\sigma u^{\rm BS}(\sigma;\tau,x,k)=e^k\sigma\tau\Gamma_0\left(\sigma^2\tau,x-k-\frac{\sigma^2\tau}{2}\right)=e^k\sqrt{\tau}\,\Gamma_0\,(1,\zeta) \qquad \text{(C.9)}$$

with $\zeta=\frac{x-k-\frac{\sigma^2\tau}{2}}{\sigma\sqrt{2\tau}}$. Therefore we have

$$0<\partial_\sigma u^{\rm BS}(\sigma;\tau,x,k)\leq\frac{e^k\sqrt{\tau}}{\sqrt{2\pi}},\qquad x,k\in\mathbb{R},\sigma,\tau\in\mathbb{R}_{++},$$

which proves (C.7) for $n=1$ and $m=0$. Notice that

$$|\partial_k^m\Gamma_0(1,\zeta)|=\frac{1}{(\sigma\sqrt{2\tau})^m}|\partial_\zeta^m\Gamma_0(1,\zeta)|\leq C(\sigma\sqrt{\tau})^{-m},\qquad m\in\mathbb{N}_0,$$

where the last inequality follows from (C.3). Then by differentiating (C.9), it is straightforward to show that

$$|\partial_\sigma\partial_k^m u^{\rm BS}(\sigma;\tau,x,k)|\leq Ce^k\sqrt{\tau}(\sigma\sqrt{\tau})^{-m},\qquad m\in\mathbb{N}_0.$$

For $n\geq 2$, by combining Proposition C.4 with (C.9), we have

$$\begin{aligned}&\partial_\sigma^n u^{\rm BS}(\sigma;\tau,x,k)\\&=e^k\sqrt{\tau}\sum_{q=0}^{\lfloor n/2\rfloor}\sum_{p=0}^{n-q-1}c_{n,n-2q}\sigma^{n-2q-1}\tau^{n-q-1}\binom{n-q-1}{p}\\&\qquad\times\left(\frac{1}{\sigma\sqrt{2\tau}}\right)^{p+n-q-1}\Gamma_0\,(1,\zeta)\,\mathbf{H}_{p+n-q-1}(\zeta).\end{aligned}\qquad \text{(C.10)}$$

Now notice that

$$\begin{aligned}\left|\partial_k^m\left(\Gamma_0\,(1,\zeta)\,\mathbf{H}_p(\zeta)\right)\right|&=|\partial_k^m\partial_\zeta^p\Gamma_0(1,\zeta)|\\&=\frac{1}{(\sigma\sqrt{\tau})^m}|\partial_\zeta^{m+p}\Gamma_0(1,\zeta)|\leq C(\sigma\sqrt{\tau})^{-m}.\end{aligned}\qquad \text{(C.11)}$$

Then the thesis follows by differentiating (C.10) and using (C.11).

Step 3: Case $q\geq 1$. Here we denote by C any generic constant that depends at most on m, q, n and M. By applying the identity

$$\partial_\tau u^{\rm BS}(\sigma;\tau,x,k)=\frac{\sigma^2}{2}(\partial_x^2-\partial_x^2)u^{\rm BS}(\sigma;\tau,x,k)=\frac{\sigma^2}{2}(\partial_k^2-\partial_k^2)u^{\rm BS}(\sigma;\tau,x,k),$$

we get

$$\partial_\sigma^n\partial_\tau^q\partial_k^m u^{\rm BS}(\sigma;\tau,x,k)=\partial_k^m(\partial_k^2-\partial_k)^q\,\partial_\sigma^n\left(\left(\frac{\sigma^2}{2}\right)^q u^{\rm BS}(\sigma;\tau,x,k)\right).$$

The statement now follows by applying Faà di Bruno's formula (Proposition E.1) along with (C.7) for $q = 0$. □

Appendix D: Explicit representation for the volatility expansion

Here we recall an explicit representation formula for the nth order correcting terms u_n and σ_n appearing in the price expansion (3.6) and the implied volatility expansion (3.14), respectively. The following result is a particular case of [32, Theorem 3.2].

Theorem D.1 *Let $N \in \mathbb{N}$, $\bar{z} \in \mathbb{R}^d$ and assume that $D_z^{\beta} a_{\alpha}(\cdot, \bar{z}) \in L^{\infty}([0, T])$ for any $1 \le |\alpha| \le 2$ and $|\beta| \le N$. Then for any $1 \le n \le N$, the function u_n in (3.8) is given by*

$$u_n^{(\bar{z})}(t, z) = \mathcal{L}_n^{(\bar{z})}(t, T, z) u_0^{(\bar{z})}(t, z), \qquad t \in [0, T), z \in \mathbb{R}^d. \tag{D.1}$$

In (D.1), $\mathcal{L}_n^{(\bar{z})}(t, T, z)$ denotes the differential operator acting on the z-variable and defined as

$$\sum_{h=1}^{n} \int_t^T ds_1 \int_{s_1}^T ds_2 \cdots \int_{s_{h-1}}^T ds_h \sum_{i \in I_{n,h}} \mathcal{G}_{i_1}^{(\bar{z})}(t, s_1, z) \cdots \mathcal{G}_{i_h}^{(\bar{z})}(t, s_h, z), \tag{D.2}$$

where[8]

$$I_{n,h} = \{i = (i_1, \ldots, i_h) \in \mathbb{N}^h : i_1 + \cdots + i_h = n\}, \qquad 1 \le h \le n,$$

and the operator $\mathcal{G}_n^{(\bar{z})}(t, s, z)$ is defined as

$$\mathcal{G}_n^{(\bar{z})}(t, s, z) := \mathcal{A}_n^{(\bar{z})}\big(s, z - \bar{z} + \mathbf{m}^{(\bar{z})}(t, s) + \mathbf{C}^{(\bar{z})}(t, s) \nabla_z\big)$$

with $\mathbf{m}^{(\bar{z})}(t, s)$ and $\mathbf{C}^{(\bar{z})}(t, s)$ being respectively the vector and matrix whose components are given by

$$\mathbf{m}_i^{(\bar{z})}(t, s) = \int_t^s a_i(r, \bar{z}) dr, \qquad \mathbf{C}_{ij}^{(\bar{z})}(t, s) = \int_t^s a_{ij}(r, \bar{z}) dr, \qquad i, j = 1, \ldots, d.$$

Corollary D.2 *Let $N \in \mathbb{N}_0$, and let Assumption 4.2 be in force. Then we have for any $n, m, q \in \mathbb{N}_0$ with $n, 2q \le N$ and for any multi-index $\alpha \in \mathbb{N}_0^d$ that*

$$\begin{aligned} &\partial_T^q \partial_k^m D_z^{\alpha} u_n^{(\bar{z})}(t, z; T, k) \\ &= \sum_{\substack{0 \le |\gamma| \le n \\ 1 \le j \le 3n}} f_{\gamma, j}^{(n,q,m,\alpha)}(\bar{z}; t, T)(z - \bar{z})^{\gamma} \partial_{z_1}^{j+m+2q+\alpha_1} u_0^{(\bar{z})}(t, z; T, k) \end{aligned} \tag{D.3}$$

with

$$|f_{\gamma, j}^{(n,q,m,\alpha)}(\bar{z}; t, T)| \le C M^q \big(M(T - t)\big)^{\frac{n - |\gamma| + j}{2}} \tag{D.4}$$

---

[8] For instance, for $n = 3$, we have $I_{3,3} = \{(1, 1, 1)\}$, $I_{3,2} = \{(1, 2), (2, 1)\}$, and $I_{3,1} = \{(3)\}$.

for any $0 \le t < T < T_0$, $z, \bar{z} \in D(z_0, r)$ *and* $k \in \mathbb{R}$. *Consequently, we have*

$$|\partial_T^q \partial_k^m u_0^{(z)}(t, z; T, k)| \le C e^x M^q \big(M(T-t)\big)^{\frac{(1-m-2q)\wedge 0}{2}}, \tag{D.5}$$

and for $n \ge 1$,

$$|\partial_T^q \partial_k^m u_n^{(z)}(t, z; T, k)| \le C e^x M^q \big(M(T-t)\big)^{\frac{n+1-m-2q}{2}}. \tag{D.6}$$

In (D.4)–(D.6), C *is a positive constant only dependent on* ε, M_0, T_0, N, $|\alpha|$, *and* m.

Proof Using the explicit formulas (D.1) and (D.2) and noting that $u_0^{(\bar{z})}(t, z; T, k)$ does not depend on $z_2, \dots, z_d$, it is straightforward to prove that

$$u_n^{(\bar{z})}(t, z; T, k) = \sum_{\substack{|\gamma| \le n \\ 0 \le j \le 3n}} f_{\gamma,j}^{(n)}(\bar{z}; t, T)(z-\bar{z})^{\gamma} \partial_{z_1}^j u_0^{(\bar{z})}(t, z; T, k) \tag{D.7}$$

with

$$|\partial_T^i f_{\gamma,j}^{(n)}(\bar{z}; t, T)| \le C M^i \big(M(T-t)\big)^{\frac{n-|\gamma|+j-2i}{2}}, \qquad 0 \le 2i \le N. \tag{D.8}$$

The general statement now follows from (D.7) and (D.8) along with the identities (C.6) and

$$\partial_T u_0^{(\bar{z})}(t, z; T, k) = \frac{a_{11}(T, \bar{z})}{2}(\partial_{z_1}^2 - \partial_{z_1}) u_0^{(\bar{z})}(t, z; T, k). \tag{D.9}$$

The estimate (D.5) follows from Lemma C.3. By combining (D.3) with (C.4), we finally get the estimate (D.6). □

Furthermore, we recall the following result [33, Proposition 3.6].

Proposition D.3 *For every* $n \in \mathbb{N}$ *and* $\bar{z} \in \mathbb{R}^d$, *the ratio* $u_n^{(\bar{z})}/\partial_\sigma u^{\mathrm{BS}}(\sigma_0^{(\bar{z})})$ *in* (3.15) *is a finite sum of the form*

$$\sum_m \Big(\sigma_0^{(\bar{z})}\sqrt{2(T-t)}\Big)^{-m} \chi_{m,n}^{(\bar{z})}\, \mathbf{H}_m(\zeta), \qquad \zeta = \frac{x - k - \frac{1}{2}\sigma_0^2(T-t)}{\sigma_0\sqrt{2(T-t)}}$$

for any $t < T$, *any* $z = (x, y) \in \mathbb{R}^d$ *and any* $k \in \mathbb{R}$, *where the coefficients* $\chi_{m,n}^{(\bar{z})} = \chi_{m,n}^{(\bar{z})}(t, z; T, k)$ *are explicit functions which are polynomial in the log-moneyness* $k - x$. *Here,* $\mathbf{H}_m$ *represents the* m*th order Hermite polynomial defined in* (C.2).

Appendix E: Multivariate Faà di Bruno formula and Bell polynomials

In this section, we recall a multivariate version of the well-known Faà di Bruno formula (see Riordan [41] and Johnson [27]) and more precisely, its Bell polynomial version.

For greater convenience, we recall some elements of tensorial calculus. For any $n, h \in \mathbb{N}$, we denote by Λ a *rank-h tensor on* $\mathbb{R}^n$, i.e., an array $\Lambda = (\Lambda_i)_{i \in \{1,\dots,n\}^h}$, with $\Lambda_i \in \mathbb{R}$. Moreover, by definition, a *rank-0 tensor* is a real number, independently of the dimension n.

Let us now fix the dimension $n \in \mathbb{N}$. For any couple of tensors Λ, Θ of rank h_1 and h_2 respectively, we define the *tensorial product* $\Lambda \otimes \Theta$ as the rank-$(h_1 + h_2)$ tensor given by

$$\Lambda \otimes \Theta_{i_1,\dots,i_{h_1}, i_{h_1+1},\dots,i_{h_1+h_2}} = \Theta_{i_1,\dots,i_{h_1}} \Lambda_{i_1,\dots,i_{h_2}}, \qquad i \in \{1,\dots,n\}^{h_1+h_2}. \tag{E.1}$$

We also set $\Lambda^0 = 1$, $\Lambda^1 = \Lambda$ and

$$\Lambda^i := \Lambda \underbrace{\otimes \Lambda \otimes \cdots \otimes}_{(i-1)\text{ times}} \Lambda, \qquad i \geq 2.$$

Furthermore, if Λ and Θ have the same rank h, we define the *tensorial scalar product* $\Lambda * \Theta$ as the rank-0 tensor given by

$$\Lambda * \Theta = \sum_{i \in \{1,\dots,n\}^h} \Lambda_i \Theta_i. \tag{E.2}$$

We say that a rank-h tensor Λ is *symmetric* if $\Lambda_i = \Lambda_{\nu(i)}$ for any $i \in \{1,\dots,n\}^h$ and any permutation ν of the indexes $(i_1,\dots,i_h)$.

Consider now a polynomial p in the variables $x = (x_1,\dots,x_j)$, homogeneous of degree h, of the form

$$p(x) = \sum_{\substack{\beta \in \mathbb{N}_0^j \\ |\beta| = h}} b_\beta x_1^{\beta_1} \cdots x_j^{\beta_j}. \tag{E.3}$$

For any rank-h symmetric tensor Λ and any family of rank-1 tensors $\{\Theta_1,\dots,\Theta_j\}$, the scalar

$$\Lambda * p(\Theta_1,\dots,\Theta_j) = \Lambda * \sum_{\substack{\beta \in \mathbb{N}_0^j \\ |\beta| = h}} b_\beta \Theta_1^{\beta_1} \otimes \cdots \otimes \Theta_j^{\beta_j}$$

is well defined. Note that the tensor $p(\Theta_1,\dots,\Theta_j)$ is not well defined on its own because *the tensorial product* (E.1) *is not commutative*. Nevertheless, by assuming Λ to be symmetric, the scalar product (E.3) is well defined as it does not depend on the specific order of the tensorial products inside the sum.

We are ready to state the following

Proposition E.1 Multivariate Faà di Bruno formula *Let* $G : \mathbb{R} \to \mathbb{R}^n$ *and* $F : \mathbb{R}^n \to \mathbb{R}$ *be two smooth functions. Then for any* $m \in \mathbb{N}$, *we have*

$$\frac{d^m F(G(x))}{dx^m} = \sum_{h=1}^{m} (\nabla^h F)\big(G(x)\big) * \mathbf{B}_{m,h}\left(\frac{dG(x)}{dx}, \frac{d^2G(x)}{dx^2}, \dots, \frac{d^{m-h+1}G(x)}{dx^{m-h+1}}\right), \tag{E.4}$$

where $\nabla^h F$ *is the rank-*h *tensor with dimension* n *of the* h*th order partial derivatives of* F, *i.e.,*

$$\nabla^h F_i = \partial_{i_1} \cdots \partial_{i_h} F, \qquad i \in \{1, \dots, n\}^h,$$

and $\mathbf{B}_{m,h}$ *is the family of the Bell polynomials defined as*

$$\mathbf{B}_{m,h}(z) = m! \sum_{j_1, j_2, \dots, j_{m-h+1}} \prod_{i=1}^{m-h+1} \frac{1}{j_i!} \left(\frac{z_i}{i!} \right)^{j_i} \tag{E.5}$$

for $1 \le h \le m$, *where the sum is taken over all sequences* $j_1, j_2, \dots, j_{m-h+1}$ *of non-negative integers such that*

$$j_1 + j_2 + \cdots + j_{m-h+1} = h \quad \text{and} \quad j_1 + 2j_2 + \cdots + (m-h+1)\, j_{m-h+1} = m. \tag{E.6}$$

References

1. Andersen, L.B.G., Piterbarg, V.V.: Moment explosions in stochastic volatility models. Finance Stoch. **11**, 29–50 (2007)
2. Barletta, A., Nicolato, E., Pagliarani, S.: The short-time behavior of VIX implied volatilities in a multifactor stochastic volatility framework (2017). Preprint. Available online at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2942262
3. Bayer, C., Laurence, P.: Asymptotics beats Monte Carlo: the case of correlated local vol baskets. Commun. Pure Appl. Math. **67**, 1618–1657 (2014)
4. Ben Arous, G., Laurence, P.: Second order expansion for implied volatility in two factor local-stochastic volatility models and applications to the dynamic λ-SABR model. In: Friz, P.K., et al. (eds.) Large Deviations and Asymptotic Methods in Finance. Springer Proceedings in Mathematics & Statistics, vol. 110, pp. 89–136 (2015)
5. Benhamou, E., Gobet, E., Miri, M.: Expansion formulas for European options in a local volatility model. Int. J. Theor. Appl. Finance **13**, 603–634 (2010)
6. Berestycki, H., Busca, J., Florent, I.: Asymptotics and calibration of local volatility models. Quant. Finance **2**, 61–69 (2002)
7. Berestycki, H., Busca, J., Florent, I.: Computing the implied volatility in stochastic volatility models. Commun. Pure Appl. Math. **57**, 1352–1373 (2004)
8. Bompis, R., Gobet, E.: Asymptotic and non-asymptotic approximations for option valuation. In: Gerstner, T., Kloeden, P. (eds.) Recent Developments in Computational Finance: Foundations, Algorithms and Applications, pp. 1–80. World Scientific, Singapore (2012)
9. Caravenna, F., Corbetta, J.: General smile asymptotics with bounded maturity. SIAM J. Financ. Math. **7**, 720–759 (2016)
10. Cox, J.C.: The constant elasticity of variance option pricing model. J. Portf. Manag. **23**(5), 15–17 (1996), Special issue
11. De Marco, S.: Smoothness and asymptotic estimates of densities for SDEs with locally smooth coefficients and applications to square root-type diffusions. Ann. Appl. Probab. **21**, 1282–1321 (2011)
12. Del Baño Rollin, S., Ferreiro-Castilla, A., Utzet, F.: On the density of log-spot in the Heston volatility model. Stoch. Process. Appl. **120**, 2037–2063 (2010)
13. Delbaen, F., Shirakawa, H.: A note on option pricing for the constant elasticity of variance model. Asia-Pac. Financ. Mark. **9**, 85–99 (2002)
14. Deuschel, J.D., Friz, P.K., Jacquier, A., Violante, S.: Marginal density expansions for diffusions and stochastic volatility I: theoretical foundations. Commun. Pure Appl. Math. **67**, 40–82 (2014)
15. Durrleman, V.: From implied to spot volatilities. Finance Stoch. **14**, 157–177 (2010)

16. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley, New York (1986)
17. Forde, M., Jacquier, A., Lee, R.: The small-time smile and term structure of implied volatility under the Heston model. SIAM J. Financ. Math. **3**, 690–708 (2012)
18. Friedman, A.: Stochastic Differential Equations and Applications. Vol. 1. Academic Press/Harcourt Brace Jovanovich, New York/London (1975)
19. Friedman, A.: Stochastic Differential Equations and Applications. Vol. 2. Academic Press/Harcourt Brace Jovanovich, New York/London (1976)
20. Gao, K., Lee, R.: Asymptotics of implied volatility to arbitrary order. Finance Stoch. **18**, 349–392 (2014)
21. Gatheral, J., Hsu, E.P., Laurence, P., Ouyang, C., Wang, T.-H.: Asymptotics of implied volatility in local volatility models. Math. Finance **22**, 591–620 (2012)
22. Gavalas, G.R., Yortsos, Y.C.: Short-time asymptotic solutions of the heat conduction equation with spatially varying coefficients. J. Inst. Math. Appl. **26**, 209–219 (1980)
23. Heston, S.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Financ. Stud. **6**, 327–343 (1993)
24. Heston, S.L., Loewenstein, M., Willard, G.A.: Options and bubbles. Rev. Financ. Stud. **20**, 359–390 (2007)
25. Hörmander, L.: Hypoelliptic second order differential equations. Acta Math. **119**, 147–171 (1967)
26. Janson, S., Tysk, J.: Feynman–Kac formulas for Black–Scholes-type operators. Bull. Lond. Math. Soc. **38**, 269–282 (2006)
27. Johnson, W.P.: The curious history of Faà di Bruno's formula. Am. Math. Mon. **109**, 217–234 (2002)
28. Kusuoka, S.: Hölder continuity and bounds for fundamental solutions to nondivergence form parabolic equations. Anal. PDE **8**, 1–32 (2015)
29. Kusuoka, S., Stroock, D.: Applications of the Malliavin calculus. II. J. Fac. Sci., Univ. Tokyo, Sect. 1A, Math. **32**, 1–76 (1985)
30. Lions, P.-L., Musiela, M.: Correlations and bounds for stochastic volatility models. Ann. Inst. Henri Poincaré, Anal. Non Linéaire **24**, 1–16 (2007)
31. Lorig, M.: Indifference prices and implied volatilities. Math. Finance (2015, forthcoming). Available online at http://onlinelibrary.wiley.com/doi/10.1111/mafi.12129/epdf
32. Lorig, M., Pagliarani, S., Pascucci, A.: Analytical expansions for parabolic equations. SIAM J. Appl. Math. **75**, 468–491 (2015)
33. Lorig, M., Pagliarani, S., Pascucci, A.: Explicit implied volatilities for multifactor local-stochastic volatility models. Math. Finance (2015, forthcoming). Available online at http://onlinelibrary.wiley.com/doi/10.1111/mafi.12105/epdf
34. Lorig, M., Pagliarani, S., Pascucci, A.: A family of density expansions for Lévy-type processes. Ann. Appl. Probab. **25**, 235–267 (2015)
35. Mijatović, A., Tankov, P.: A new look at short-term implied volatility in asset price models with jumps. Math. Finance **26**, 149–183 (2016)
36. Moser, J.: On a pointwise estimate for parabolic differential equations. Commun. Pure Appl. Math. **24**, 727–740 (1971)
37. Pagliarani, S., Pascucci, A.: Analytical approximation of the transition density in a local volatility model. Cent. Eur. J. Math. **10**, 250–270 (2012)
38. Pagliarani, S., Pascucci, A.: Asymptotic expansions for degenerate parabolic equations. C. R. Math. **352**, 1011–1016 (2014)
39. Pascucci, A.: PDE and Martingale Methods in Option Pricing. Springer/Bocconi University Press, Milan (2011)
40. Pascucci, A., Polidoro, S.: The Moser's iterative method for a class of ultraparabolic equations. Commun. Contemp. Math. **6**, 395–417 (2004)
41. Riordan, J.: Derivatives of composite functions. Bull. Am. Math. Soc. **52**, 664–667 (1946)
42. Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes, and Martingales. Vol. 2, Itô Calculus. Wiley, New York (1987)
43. Roper, M., Rutkowski, M.: On the relationship between the call price surface and the implied volatility surface close to expiry. Int. J. Theor. Appl. Finance **12**, 427–441 (2009)
44. Safonov, M.: Estimates near the boundary for solutions of second order parabolic equations. In: Proceedings of the International Congress of Mathematicians, Vol. I, Berlin, 1998. Doc. Math. 1998, Extra Vol. I, pp. 637–647 (1998),
45. Stroock, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes. Grundlehren der mathematischen Wissenschaften., vol. 233. Springer, Berlin (1979)

46. Takahashi, A., Yamada, T.: On error estimates for asymptotic expansions with Malliavin weights: application to stochastic volatility model. Math. Oper. Res. **40**, 513–541 (2015)
47. Wang, F.-Y., Zhang, X.: Degenerate SDE with Hölder–Dini drift and non-Lipschitz noise coefficient. SIAM J. Math. Anal. **48**, 2189–2226 (2016)
48. Wang, J.: Regularity of semigroups generated by Lévy type operators via coupling. Stoch. Process. Appl. **120**, 1680–1700 (2010)