ARCHIVIO ISTITUZIONALE DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

LEVERAGED ETF IMPLIED VOLATILITIES FROM ETF DYNAMICS

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Leung, T., Lorig, M., Pascucci, A. (2017). LEVERAGED ETF IMPLIED VOLATILITIES FROM ETF DYNAMICS. MATHEMATICAL FINANCE, 27(4), 1035-1068 [10.1111/mafi.12128].

Availability:
This version is available at: https://hdl.handle.net/11585/613992 since: 2017-12-19

Published:
DOI: http://doi.org/10.1111/mafi.12128

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

16 September 2026

This is the final peer-reviewed accepted manuscript of:

Leung, T., Lorig, M. and Pascucci, A. (2017), LEVERAGED ETF IMPLIED VOLATILITIES FROM ETF DYNAMICS. Mathematical Finance, 27: 1035-1068.

The final published version is available online at : **https://doi.org/10.1111/mafi.12128**

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/)

When citing, please refer to the published version.

Author Manuscript

Leveraged ETF implied volatilities from ETF dynamics

Tim Leung * Matthew Lorig † Andrea Pascucci ‡

This version: December 11, 2015

Abstract

The growth of the exchange-traded fund (ETF) industry has given rise to the trading of options written on ETFs and their leveraged counterparts (LETFs). We study the relationship between the ETF and LETF implied volatility surfaces when the underlying ETF is modeled by a general class of local-stochastic volatility models. A closed-form approximation for prices is derived for European-style options whose payoffs depend on the terminal value of the ETF and/or LETF. Rigorous error bounds for this pricing approximation are established. A closed-form approximation for implied volatilities is also derived. We also discuss a scaling procedure for comparing implied volatilities across leverage ratios. The implied volatility expansions and scalings are tested in three settings: Heston, limited CEV and limited SABR; the last two are regularized versions of the well-known CEV and SABR models.

Keywords: implied volatility, local-stochastic volatility, leveraged exchange-traded fund, implied volatility scaling

1 Introduction

The market of exchange-traded funds (ETFs) has been growing at a robust pace since their introduction in 1993[1]. As of the end of 2012, the global ETF industry has over $1.8 trillion in assets under management (AUM) comprised of 4,272 products, and has seen close to $200 billion of positive capital inflows[2]. In recent years, a sub-class of ETFs, called leveraged ETFs (LETFs), has gained popularity among investors for their accessibility and liquidity for leveraged positions. These funds are designed to replicate multiples of the daily returns of some reference index or asset. For instance, the ProShares S&P 500 Ultra (SSO) and UltraPro (UPRO) are advertised to generate, respectively, 2 and 3 times of the daily returns of the S&P 500 index, minus a small expense fee. On the other hand, a LETF with a negative leverage ratio allows investors to

*Industrial Engineering & Operations Research Department, Columbia University, New York, NY 10027. **E-mail:** leung@ieor.columbia.edu.

†Department of Applied Mathematics, University of Washington, Seattle, WA 98195. **E-mail:** mattlorig@gmail.com. Work partially supported by NSF grant DMS-0739195.

‡Dipartimento di Matematica, Università di Bologna, Bologna, Italy. **E-mail:** andrea.pascucci@unibo.it.

[1] The first US-listed ETF, the SPDR S&P 500 ETF (SPY), was launched on January 29th, 1993.

[2] "2013 ETF & Investment Outlook" by David Mazza, SPDR ETF Strategy & Consulting, State Street Global Advisors. Available at http://www.spdr-etfs.com.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mafi.12128

take a bearish position on the underlying index by longing the fund. An example is the ProShares S&P 500 UltraShort (SDS) with leverage ratio of -2. The most typical leverage ratios are $\{-3, -2, -1, 2, 3\}$. With the same reference, such as the S&P 500, these LETFs share very similar sources of randomness, but they also exhibit different path behaviors (see Cheng and Madhavan (2009) and Avellaneda and Zhang (2010)).

The use of ETFs has also led to increased trading of options written on ETFs. During 2012, the total options contract volume traded at Chicago Board Options Exchange (CBOE) is 1.06 billion contracts, of which 282 million contracts are ETF options while 473 million are equity options. This leads to an important question of consistent pricing of options on ETFs and LETFs with the same reference. As options are commonly quoted and compared in terms of implied volatility, it is natural to consider the implied volatility relationships among LETF options, not only across strikes and maturities, but also for various leverage ratios.

In this paper, we analyze the implied volatility surfaces associated with European-style LETF options in a general class of local-stochastic volatility (LSV) models. Our approach is to (i) find an expansion for approximate LETF option prices (ii) establish rigorous error bounds for this approximation and (iii) translate the price approximation into approximate implied volatilities. Exact pricing and implied volatility formulas in a general LSV setting are obviously impossible to obtain. There are a number of approaches one could feasibly take in order to approximate European-style option prices and their associated implied volatilities. We review some recent approaches for unleveraged products here. Gatheral et al. (2012) use heat kernel methods in a local volatility setting. Benhamou et al. (2010) use a small volatility of volatility expansion for the time-dependent Heston model. More recently, Bompis and Gobet (2013) use Malliavin calculus to obtain approximations in a quite general LSV setting. And, Forde and Jacquier (2011) use the Freidlin-Wentzell theory of large deviations to analyze an uncorrelated LSV model.

In this paper, we use a polynomial operator expansion technique to obtain approximate prices and implied volatilities. The polynomial operator expansion technique was first introduced in Pagliarani and Pascucci (2012) and Pagliarani et al. (2013) to compute option prices in a scalar jump-diffusion setting. It was further developed in Lorig et al. (2015b) to obtain approximate prices and implied volatilities in a multidimensional local-stochastic volatility setting (see also Lorig et al. (2015c) for pricing approximations for models with jumps). The reason for basing our expansions on the methods developed in Lorig et al. (2015b) is that these methods allow us to consider a large class of LSV models for the ETF; many of the above mentioned methods work only for specific ETF dynamics. However, without further development, the methods described in Lorig et al. (2015b) are not sufficient for the rigorous error bounds we establish in this paper. Indeed, in Lorig et al. (2015b), error bounds are established under a uniform ellipticity assumption. As we shall see, the generator of the joint ETF/LETF process is not elliptic. As such, to establish rigorous error bounds for LETF option prices, we must work in this challenging non-elliptic setting.

Perhaps the most useful result of our analysis is the general expression we obtain for the implied volatility expansion. This expansion allows us to pinpoint the non-trivial role played by the leverage ratio β, and thus, relate the implied volatility surfaces between (unleveraged) ETF and LETF options. This also motivates us to apply the idea of *log-moneyness scaling*, with the objective to view the implied volatilities across leverage ratios on the same scale and orientation. In particular, for a negative leverage ratio and up to the

Author Manuscript

first-order in log-moneyness, the LETF implied volatility is known to be upward sloping while the ETF and long-LETF implied volatilities are downward sloping (see Leung and Sircar (2015)). The scaling is capable of appropriately adjusting the level and shape of the implied volatility so that the ETF and LETF implied volatilities match closely under a given model. For illustration, we test our implied volatility expansions and the log-moneyness scaling in three settings: Heston, limited CEV, and limited SABR; the limited CEV and limited SABR models are, respectively, regularized versions of the well-known CEV and SABR models. Our numerical results demonstrate a high degree of accuracy of our analytical implied volatility approximation in the models tested.

In a recent paper, Leung and Sircar (2015) apply asymptotic techniques to understand the link between implied volatilities of the ETF and LETFs of different leverage ratios within a multiscale stochastic volatility framework (see Fouque et al. (2011) for a review of multiscale methods). They also introduce implied volatility scaling procedure, different from our own, in order to identify possible price discrepancies in the ETF and LETF options markets. In contrast to their work, the current paper studies the problem in a general LSV framework, which naturally includes models such as limited CEV, Heston and limited SABR, among others. Moreover, while Leung and Sircar (2015) obtain an implied volatility approximation that is linear in log-moneyness, we provide a general expression for LETF implied volatilities that is quadratic in log-moneyness. We also provide formulas for three specific models (limited CEV, Heston and limited SABR that are cubic in log-moneyness.

Ahn et al. (2012) propose a method to compute exact LETF option prices when the underlying ETF has Heston stochastic volatility dynamics. They also derive approximate LETF option prices when the underlying ETF has Heston dynamics and jumps. While they do not investigate the implied volatilities, they point out that if the underlying ETF admits the Heston dynamics (no jumps), then the LETF also has Heston dynamics with different parameters. As a particular example of LSV models, we also obtain the same result revealed through our implied volatility expansions (see Section 6.2). Lee and Wang (2015) rigorously relate the implied volatility surfaces between LETF and ETF under various asymptotic regimes in a family of stochastic volatility frameworks, including those driven by a fractional Brownian motion, as well as Lévy models.

The rest of this paper proceeds as follows. In Section 2 we review how LETF dynamics are related to ETF dynamics in a general diffusion setting. We then introduce Markov dynamics for a general class of LSV models for the ETF. Next, in Section 3, we formally construct an asymptotic expansion for European-style options whose payoffs depend on the terminal value of the ETF and/or LETF. Rigorous error bounds for our pricing approximation are established in Section 4. In Section 5 we translate our asymptotic expansion for prices into an asymptotic expansion for implied volatilities. We also discuss some natural scalings of the implied volatility surface of the LETF. Finally, in Section 6 we implement our implied volatility expansion in three settings: limited CEV, Heston and limited SABR. Some concluding remarks are given in Section 7.

Author Manuscript

Author Manuscript

2 Leveraged ETF dynamics

We take as given an equivalent martingale measure $\mathbb{Q}$, chosen by the market on a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, \mathbb{Q})$. The filtration $\{\mathcal{F}_t, t \geq 0\}$ represents the history of the market. All stochastic processes defined below live on this probability space and all expectations are taken with respect to $\mathbb{Q}$. For simplicity, we assume a frictionless market, no arbitrage, zero interest rates and no dividends. We will discuss how to relax these assumptions in Remark 3.2.

Let S be the price process of an Exchange-Traded Fund (ETF). We assume S can be modeled under $\mathbb{Q}$ as a strictly positive Itô diffusion. Specifically, we have

$$\text{ETF}: \qquad S_t = \mathrm{e}^{X_t}, \qquad \mathrm{d}X_t = -\frac{1}{2}\sigma_t^2 \mathrm{d}t + \sigma_t \, \mathrm{d}W_t^x, \tag{2.1}$$

where σ is a non-negative stochastic process. Note that the drift is fixed by the volatility so that S is a martingale. Let L be the price process of a Leveraged Exchange-Traded Fund (LETF) with underlying S and with leverage ratio β. Typical values of β are $\{-3, -2, -1, 2, 3\}$. The LETF is managed as follows: for every unit of currency a trader invests in L, the LETF manager borrows $(\beta - 1)$ units of currency and invests β units of currency in S. The fund manager also typically charges the trader a small expense rate, which, for simplicity, we assume is zero. Then the dynamics of L are related to S as follows

$$\frac{\mathrm{d}L_t}{L_t} = \beta \, \frac{\mathrm{d}S_t}{S_t} = \beta \sigma_t \, \mathrm{d}W_t^x,$$

and thus we have

$$\text{LETF}: \qquad L_t = \mathrm{e}^{Z_t}, \qquad \mathrm{d}Z_t = -\frac{1}{2}\beta^2 \sigma_t^2 \, \mathrm{d}t + \beta \sigma_t \, \mathrm{d}W_t^x. \tag{2.2}$$

Comparing (2.1) with (2.2), we observe that the volatility of L is scaled by a factor of β. Moreover, as shown by Avellaneda and Zhang (2010), one can solve explicitly the SDE for Z in order to obtain an expression for Z_t in terms of X_t and the quadratic variation (integrated variance) of X up to time t. Specifically, we have

$$Z_t - Z_0 = \beta \, (X_t - X_0) - \frac{\beta(\beta - 1)}{2} \int_0^t \sigma_s^2 \, \mathrm{d}s. \tag{2.3}$$

Equation (2.3) shows that the log returns of a LETF is the sum of two terms. The first term is proportional to the log returns of the underlying ETF. The second term is proportional to the integrated variance of X, and highlights the fact that options on LETFs are *path dependent* options. Note that, for leverage ratio $\beta \in \{-3, -2, -1, 2, 3\}$, the coefficient $-\frac{\beta(\beta-1)}{2}$ of the realized variance is strictly negative. Note also that $-\frac{\beta(\beta-1)}{2}$ is an asymmetric function of β.

2.1 Local-stochastic volatility framework

We now specialize to the Markov setting. We introduce an auxiliary process Y, which is intended to capture effects such as stochastic volatility. We assume that the triple (X, Y, Z) can be modeled by the following

Stochastic Differential Equation (SDE):

$$\begin{aligned} \mathrm{d}X_t &= -\frac{1}{2}\sigma^2(t, X_t, Y_t)\mathrm{d}t + \sigma(t, X_t, Y_t)\mathrm{d}W_t^x, \\ \mathrm{d}Y_t &= c(t, X_t, Y_t)\mathrm{d}t + g(t, X_t, Y_t)\mathrm{d}W_t^y, \\ \mathrm{d}Z_t &= -\frac{1}{2}\beta^2\sigma^2(t, X_t, Y_t)\mathrm{d}t + \beta\sigma(t, X_t, Y_t)\mathrm{d}W_t^x, \\ \mathrm{d}\langle W^x, W^y\rangle_t &= \rho(t, X_t, Y_t)\mathrm{d}t. \end{aligned} \tag{2.4}$$

We assume that SDE (2.4) has a unique strong solution. Sufficient conditions for a unique strong solution are given in Ikeda and Watanabe (1989). The class of models described by (2.4) enjoys the following features:

1. **Stochastic Volatility**: When σ and ρ are functions of (t, y) only (as they would be in a stochastic volatility model such as Heston), then the pairs (X, Y) and (Y, Z) are Markov processes. From a mathematical point of view, the lack of x-dependence in the correlation ρ and volatility σ greatly simplifies the pricing and implied volatility analysis, as calls written on Z can be analyzed independently from calls on X.

2. **Local Volatility**: If both σ and ρ are dependent on (t, x) only (as they would be in a local volatility model such as limited CEV), then X alone and the pair (X, Z) are Markov processes. In this case, calls on X can be analyzed separately from Z. However, calls on Z must be analyzed in conjunction with X.

3. **Local-Stochastic Volatility**: If σ and ρ depend on (x, y) (as would be the case in a local-stochastic volatility setting such as limited SABR), then the pair (X, Y) is a Markov process, as is the triple (X, Y, Z). In this case, options on X can be analyzed independently from Z. In contrast, to analyze options on Z, one must consider the triple (X, Y, Z).

4. If $\beta = 1$, then from (2.4) we see that $\mathrm{d}X_t = \mathrm{d}Z_t$. Thus, we need only obtain prices and implied volatilities for options written on Z. Options written on X can always be obtained by considering the special case $\beta = 1$.

3 Option pricing

Using risk-neutral pricing and the Markov property of the process (X, Y, Z), we can write the time t price of an option $u(t, x, y, z)$ with expiration date $T > t$ and payoff $\varphi(Z_T)$ as the risk-neutral expectation of the payoff

$$u(t, x, y, z) = \mathbb{E}[\varphi(Z_T)|X_t = x, Y_t = y, Z_t = z]. \tag{3.1}$$

Under mild assumptions, the function u satisfies the Kolmogorov backward equation

$$(\partial_t + \mathcal{A}(t))\, u(t) = 0, \qquad\qquad u(T) = \varphi, \tag{3.2}$$

Author Manuscript

Author Manuscript

where the operator $\mathcal{A}(t)$ is given by

$$\begin{aligned}\mathcal{A}(t) = a(t,x,y)\left(\left(\partial_x^2 - \partial_x\right) + \beta^2\left(\partial_z^2 - \partial_z\right) + 2\beta\,\partial_x\partial_z\right)\\ + b(t,x,y)\partial_y^2 + c(t,x,y)\partial_y + f(t,x,y)\left(\partial_x\partial_y + \beta\,\partial_y\partial_z\right),\end{aligned} \tag{3.3}$$

with the functions (a, b, f) defined as

$$a(t,x,y) = \tfrac{1}{2}\sigma^2(t,x,y), \qquad b(t,x,y) = \tfrac{1}{2}g^2(t,x,y), \qquad f(t,x,y) = g(t,x,y)\sigma(t,x,y)\rho(t,x,y).$$

For general (a, b, c, f), an explicit solution to (3.2) is not available. Thus, our goal is to find a closed form approximation for the option price u and derive rigorous error bounds for our approximation.

Remark 3.1. We note that the matrix of second-order derivatives of $\mathcal{A}(t)$

$$\frac{1}{2}\begin{pmatrix} 2a & f & 2\beta a \\ f & 2b & \beta f \\ 2\beta a & \beta f & 2\beta^2 a \end{pmatrix} \tag{3.4}$$

is singular; the eigenvector $(\beta, 0, -1)$ corresponds to eigenvalue zero. Therefore, the operator $\mathcal{A}(t)$ is *not* elliptic. This gives rise to an additional mathematical challenge in establishing error bounds for the pricing approximation, which we will carry out in Section 4.

Remark 3.2 (Deterministic interest rates, dividends and expense ratios)**.** Suppose interest rates are a deterministic function $r(t)$ of time. Suppose also that the ETF holder receives a dividend $q(t)S_t$ per unit time, and the LETF provider charges an expense rate $c(t)L_t$ per unit time where $q(t)$ and $c(t)$ are deterministic functions. In this case options prices are computed as discounted expectations of the form

$$\begin{aligned}\widetilde{u}(t,\widetilde{x},y,\widetilde{z}) &:= \mathbb{E}[\mathrm{e}^{-\int_t^T \mathrm{d}s\, r(s)}\varphi(\widetilde{Z}_T)|\widetilde{X}_t = \widetilde{x}, Y_t = y, \widetilde{Z}_t = z],\\ \mathrm{d}\widetilde{X}_t &= \mathrm{d}X_t + (r(t) - q(t))\,\mathrm{d}t,\\ \mathrm{d}\widetilde{Z}_t &= \mathrm{d}Z_t + (r(t) - c(t) - \beta q(t))\,\mathrm{d}t,\end{aligned}$$

with (X, Y, Z) as given in (2.4). Upon making the following change of variables

$$\begin{aligned}u(t, x(t,\widetilde{x}), y, z(t,\widetilde{z})) &:= \mathrm{e}^{\int_t^T \mathrm{d}s\, r(s)}\widetilde{u}(t,\widetilde{x},y,\widetilde{z}),\\ x(t,\widetilde{x}) &:= \widetilde{x} + \int_t^T \mathrm{d}s\, r(s),\\ z(t,\widetilde{z}) &:= \widetilde{z} + \int_t^T \mathrm{d}s\, \left(r(s) - c(s) - \beta q(s)\right),\end{aligned} \tag{3.5}$$

a simple application of the chain rule reveals that u as defined in (3.5) satisfies Cauchy problem (3.2). Thus, the current framework allows us to readily accommodate these additional features.

3.1 Asymptotic prices via Taylor and Dyson series

In this section, we show how Taylor and Dyson series can be combined in order to formally construct an asymptotic approximation of the solution u of Cauchy problem (3.2). Throughout the derivation that follows

Author Manuscript

we assume that for every t the coefficients (a, b, c, f) of the operator $\mathcal{A}(t)$ are analytic in (x, y) so we can expand each of these functions as a Taylor series. As we will see, this assumption is not necessary for the Nth-order approximation of u, which we will give in Definition 3.3. However, making this assumption will simplify the derivation that follows.

Let $(\bar{x}(\cdot), \bar{y}(\cdot)) : [0, T] \to \mathbb{R}^2$ be a piecewise continuous map. For any (t, x, y) we have:

$$\chi(t, x, y) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \chi_{n-k,k}(t)(x - \bar{x}(t))^{n-k}(y - \bar{y}(t))^k,$$
$$\chi_{n-k,k}(t) = \frac{\partial_x^{n-k} \partial_y^k \chi(t, \bar{x}(t), \bar{y}(t))}{(n-k)!k!}, \qquad \chi \in \{a, b, c, f\}.$$

Formally, the operator $\mathcal{A}(t)$ can now be written as

$$\mathcal{A}(t) = \mathcal{A}_0(t) + \mathcal{B}_1(t), \qquad \mathcal{B}_1(t) = \sum_{n=1}^{\infty} \mathcal{A}_n(t), \qquad \mathcal{A}_n(t) = \sum_{k=0}^{n} (x - \bar{x}(t))^{n-k}(y - \bar{y}(t))^k \mathcal{A}_{n-k,k}(t), \tag{3.6}$$

where

$$\begin{aligned}\mathcal{A}_{n-k,k}(t) = a_{n-k,k}(t) \left(\left(\partial_x^2 - \partial_x \right) + \beta^2 \left(\partial_z^2 - \partial_z \right) + 2\beta \, \partial_x \partial_z \right) \\ + b_{n-k,k}(t) \partial_y^2 + c_{n-k,k}(t) \partial_y + f_{n-k,k}(t) \left(\partial_x \partial_y + \beta \, \partial_y \partial_z \right),\end{aligned}$$

Inserting expansion (3.6) for $\mathcal{A}(t)$ into Cauchy problem (3.2) we find

$$(\partial_t + \mathcal{A}_0(t))u(t) = -\mathcal{B}_1(t)u(t), \qquad u(T) = \varphi.$$

By construction, the operator $\mathcal{A}_0(t)$ is the generator of a diffusion with coefficients that are deterministic functions of time only. By Duhamel's principle, we therefore have

$$u(t) = \mathcal{P}_0(t, T)\varphi + \int_t^T \mathrm{d}t_1 \, \mathcal{P}_0(t, t_1)\mathcal{B}_1(t_1)u(t_1), \tag{3.7}$$

where $\mathcal{P}_0(t, T) = \exp \int_t^T \mathrm{d}s \, \mathcal{A}_0(s)$, is the *semigroup* of operators generated by $\mathcal{A}_0(t)$; we will provide an explicit form for $\mathcal{P}_0(t, T)$ in Section 3.2. Inserting expression (3.7) for u back in to the right-hand side of (3.7) and iterating we obtain

$$\begin{aligned}u(t) &= \mathcal{P}_0(t, T)\varphi + \sum_{k=1}^{\infty} \int_t^T \mathrm{d}t_1 \int_{t_1}^T \mathrm{d}t_2 \cdots \int_{t_{k-1}}^T \mathrm{d}t_k \\ &\qquad \mathcal{P}_0(t, t_1)\mathcal{B}_1(t_1)\mathcal{P}_0(t_1, t_2)\mathcal{B}_1(t_2) \cdots \mathcal{P}_0(t_{k-1}, t_k)\mathcal{B}_1(t_k)\mathcal{P}_0(t_k, T)\varphi \qquad (3.8) \\ &= \mathcal{P}_0(t, T)\varphi + \sum_{n=1}^{\infty} \sum_{k=1}^{n} \int_t^T \mathrm{d}t_1 \int_{t_1}^T \mathrm{d}t_2 \cdots \int_{t_{k-1}}^T \mathrm{d}t_k \\ &\qquad \sum_{i \in I_{n,k}} \mathcal{P}_0(t, t_1)\mathcal{A}_{i_1}(t_1)\mathcal{P}_0(t_1, t_2)\mathcal{A}_{i_2}(t_2) \cdots \mathcal{P}_0(t_{k-1}, t_k)\mathcal{A}_{i_k}(t_k)\mathcal{P}_0(t_k, T)\varphi, \qquad (3.9)\end{aligned}$$

$$I_{n,k} = \{i = (i_1, i_2, \cdots, i_k) \in \mathbb{N}^k : i_1 + i_2 + \cdots + i_k = n\}. \tag{3.10}$$

Note that the second to last equality (3.8) is the classical Dyson series expansion of u corresponding to order zero generator $\mathcal{A}_0(t)$ and perturbation $\mathcal{B}_1(t)$. To obtain (3.9) from (3.8) we have used the fact that, by

(3.6), the operator $\mathcal{B}_1(t)$ is an infinite sum. Rigorous justification for exchanging infinite sums and integrals, which would require additional assumptions, is not intended at this point. It will be clear in Definition 3.3 that the Nth-order approximation for u contains only finite sums. Expression (3.9) motivates the following definition:

Definition 3.3. Let u be given by (3.1). Assume that for every $t \in [0,T]$ the coefficients (a,b,c,f) of the operator $\mathcal{A}(t)$ are N-times differentiable in the spatial variables (x,y) in some domain $D \subseteq \mathbb{R}^2$. For a fixed piecewise continuous map $(\bar{x}(\cdot),\bar{y}(\cdot)) : [0,T] \to D \subseteq \mathbb{R}^2$, the *$N$th-order approximation of u*, denoted $\bar{u}_N$, is defined as

$$\bar{u}_N = \sum_{n=0}^{N} u_n, \qquad \text{where} \qquad u_0(t) := \mathcal{P}_0(t,T)\varphi, \tag{3.11}$$

and

$$\begin{aligned} u_n(t) := \sum_{k=1}^{n} \int_t^T \mathrm{d}t_1 \int_{t_1}^T \mathrm{d}t_2 \cdots \int_{t_{k-1}}^T \mathrm{d}t_k \\ \sum_{i \in I_{n,k}} \mathcal{P}_0(t,t_1)\mathcal{A}_{i_1}(t_1)\mathcal{P}_0(t_1,t_2)\mathcal{A}_{i_2}(t_2)\cdots\mathcal{P}_0(t_{k-1},t_k)\mathcal{A}_{i_k}(t_k)\mathcal{P}_0(t_k,T)\varphi. \end{aligned} \tag{3.12}$$

Here, $\mathcal{A}_i(t)$ and $I_{n,k}$ are as given in (3.6) and (3.10), respectively, and $\mathcal{P}_0(t,T)$ is the semigroup generated by $\mathcal{A}_0(t)$.

3.2 Expression for u_0

The action of the semigroup $\mathcal{P}_0(t,T)$ generated by $\mathcal{A}_0(t)$ when acting on a function $\theta : \mathbb{R}^3 \to \mathbb{R}$ is

$$\mathcal{P}_0(t,T)\theta(x,y,z) = \int_{\mathbb{R}^3} \mathrm{d}\xi\mathrm{d}\eta\mathrm{d}\zeta\, \delta_{\bar{z}}(\zeta)\, \Gamma_0(t,x,y;T,\xi,\eta)\theta(\xi,\eta,\zeta), \tag{3.13}$$

where $\delta_{\bar{z}}$ is a Dirac mass centered at

$$\bar{z} = z + \beta(\xi - x) - \beta(\beta-1)\int_0^t a_{0,0}(s)\mathrm{d}s, \tag{3.14}$$

and

$$\Gamma_0(t,x,y;T,\xi,\eta) = \frac{1}{2\pi\sqrt{|\mathbf{C}|}} \exp\left(-\frac{1}{2}\mathbf{m}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{m}\right), \tag{3.15}$$

with the covariance matrix $\mathbf{C}$ and vector $\mathbf{m}$ given by:

$$\mathbf{C} = \begin{pmatrix} 2\int_t^T a_{0,0}(s)\mathrm{d}s & \int_t^T f_{0,0}(s)\mathrm{d}s \\ \int_t^T f_{0,0}(s)\mathrm{d}s & 2\int_t^T b_{0,0}(s)\mathrm{d}s \end{pmatrix}, \qquad \mathbf{m} = \begin{pmatrix} \xi - x + \int_t^T a_{0,0}(s)\mathrm{d}s \\ \eta - y - \int_t^T c_{0,0}(s)\mathrm{d}s \end{pmatrix}.$$

Using (3.11), we have $u_0(t) = \mathcal{P}_0(t,T)\varphi$. Hence, from (3.13) a direct computation gives the zeroth-order approximation

$$u_0(t,z) = \int_{\mathbb{R}} \mathrm{d}\zeta\, \frac{1}{\sqrt{2\pi s^2(t,T)}} \exp\left(\frac{-(\zeta - m(t,T))^2}{2s^2(t,T)}\right)\varphi(\zeta), \tag{3.16}$$

where the mean $m(t,T)$ and variance $s^2(t,T)$ are given by

$$m(t,T) = z - \beta^2 \int_t^T \mathrm{d}t_1\, a_{0,0}(t_1), \qquad s^2(t,T) = 2\beta^2 \int_t^T \mathrm{d}t_1\, a_{0,0}(t_1).$$

Author Manuscript

Author Manuscript

3.3 Expression for u_n

The following theorem, and the ensuing proof, show that $u_n(t)$ can be written as a differential operator $\mathcal{L}_n(t,T)$ acting on $u_0(t)$. The theorem is written specifically for Put options, which play an important role in derivative markets. Call prices, which are also important in derivative markets, can be obtained from Put prices via Put-Call parity.

Theorem 3.4. *Assume that for every $t \in [0,T]$ the coefficients (a,b,c,f) of the operator $\mathcal{A}(t)$ are n-times differentiable in the spatial variables (x,y) in some domain $D \subseteq \mathbb{R}^2$. Assume also that φ is the payoff of a Put option on Z. That is, $\varphi(z) = \left(\mathrm{e}^k - \mathrm{e}^z\right)^+$. Then, for a fixed piecewise continuous map $(\bar{x}(\cdot), \bar{y}(\cdot)) : [0,T] \to D \subseteq \mathbb{R}^2$, the function u_n defined in* (3.12) *is given explicitly by*

$$u_n(t) = \mathcal{L}_n(t,T)u_0(t), \tag{3.17}$$

where u_0 is given by (3.16) *and*

$$\mathcal{L}_n(t,T) = \sum_{k=1}^{n} \int_t^T \mathrm{d}t_1 \int_{t_1}^T \mathrm{d}t_2 \cdots \int_{t_{k-1}}^T \mathrm{d}t_k \sum_{i \in I_{n,k}} \mathcal{G}_{i_1}(t,t_1)\mathcal{G}_{i_2}(t,t_2)\cdots\mathcal{G}_{i_k}(t,t_k), \tag{3.18}$$

with $I_{n,k}$ as defined in (3.10) *and*

$$\begin{aligned}
\mathcal{G}_n(t,t_i) &:= \sum_{k=0}^{n} \left(\mathcal{M}_x(t,t_i) - \bar{x}(t_i)\right)^{n-k} \left(\mathcal{M}_y(t,t_i) - \bar{y}(t_i)\right)^k \mathcal{A}_{n-k,k}(t_i) \qquad (3.19)\\
\mathcal{M}_x(t,t_i) &:= x + \int_t^{t_i} \mathrm{d}s \Big(a_{0,0}(s)\left(2\partial_x + 2\beta\partial_z - 1\right) + f_{0,0}(s)\partial_y \Big),\\
\mathcal{M}_y(t,t_i) &:= y + \int_t^{t_i} \mathrm{d}s \Big(f_{0,0}(s)\left(\partial_x + \beta\partial_z\right) + 2b_{0,0}(s)\partial_y + c_{0,0}(s) \Big).
\end{aligned}$$

Proof. The proof consists of showing that the operator $\mathcal{G}_i(t,t_k)$ in (3.19) satisfies

$$\mathcal{P}_0(t,t_k)\mathcal{A}_i(t_k) = \mathcal{G}_i(t,t_k)\mathcal{P}_0(t,t_k). \tag{3.20}$$

Assuming (3.20) holds, we can use the fact that $\mathcal{P}_0(t,T)$ satisfies the semigroup property

$$\mathcal{P}_0(t,T) = \mathcal{P}_0(t,t_1)\mathcal{P}_0(t_1,t_2)\cdots\mathcal{P}_0(t_{k-1},t_k)\mathcal{P}_0(t_k,T), \tag{3.21}$$

and we can re-write (3.12) as

$$\begin{aligned}
u_n(t) = \sum_{k=1}^{n} \int_t^T \mathrm{d}t_1 \int_{t_1}^T \mathrm{d}t_2 \cdots \int_{t_{k-1}}^T \mathrm{d}t_k \\
\sum_{i \in I_{n,k}} \mathcal{G}_{i_1}(t,t_1)\mathcal{G}_{i_2}(t,t_2)\cdots\mathcal{G}_{i_k}(t,t_k)\mathcal{P}_0(t,T)\varphi.
\end{aligned} \tag{3.22}$$

Note, in deriving (3.22), we have repeatedly used (3.20) to move the semigroup operators $\mathcal{P}_0(t_i,t_{i+1})$ in (3.12) past the $\mathcal{A}_i(t_{i+1})$ operators. Then, we used (3.21). Finally, using that $\mathcal{P}_0(t,T)\varphi = u_0(t)$, equations (3.17)-(3.18) follow directly from (3.22) Thus, we only need to show that $\mathcal{G}_i(t,t_k)$ satisfies (3.20).

To establish (3.20), we note that

$$\mathcal{M}_x(t,T)\Big(\delta_{\bar{z}}(\zeta)\Gamma_0(t,x,y;T,\xi,\eta)\Big) = \xi\Big(\delta_{\bar{z}}(\zeta)\Gamma_0(t,x,y;T,\xi,\eta)\Big), \tag{3.23}$$

$$\mathcal{M}_y(t,T)\Big(\delta_{\bar{z}}(\zeta)\Gamma_0(t,x,y;T,\xi,\eta)\Big) = \eta\Big(\delta_{\bar{z}}(\zeta)\Gamma_0(t,x,y;T,\xi,\eta)\Big), \tag{3.24}$$

where $\bar{z}$ is defined in (3.14) and Γ_0 is defined in (3.15). This is a direct computation, which can be checked by hand. It follows from repeated application of (3.23) and (3.24) that if $p : \mathbb{R}^2 \to \mathbb{R}$ is a polynomial function, we have

$$p\left(\mathcal{M}_x(t,T),\mathcal{M}_y(t,T)\right)\Big(\delta_{\bar{z}}(\zeta)\Gamma_0(t,x,y;T,\xi,\eta)\Big) = p(\xi,\eta)\Big(\delta_{\bar{z}}(\zeta)\Gamma_0(t,x,y;T,\xi,\eta)\Big). \tag{3.25}$$

In what follows we write $\mathcal{A}^{\xi,\eta,\zeta}_{n-k,k}(s)$ and $\mathcal{A}^{x,y,z}_{n-k,k}(s)$ in order to indicate explicitly which variables these operators act on. We also denote by $(\mathcal{A}^{\xi,\eta,\zeta}_{n-k,k}(s))^*$ the formal adjoint of $\mathcal{A}^{\xi,\eta,\zeta}_{n-k,k}(s)$. Suppose $\theta : \mathbb{R}^3 \to \mathbb{R}$ is $C^2(\mathbb{R}^3)$ and at most exponentially growing. Then we have

$$\begin{aligned}
&\mathcal{P}_0(t,s)\mathcal{A}_i(s)\theta(x,y,z)\\
&= \int_{\mathbb{R}^3} \mathrm{d}\xi\mathrm{d}\eta\mathrm{d}\zeta\, \delta_{\bar{z}}(\zeta)\Gamma_0(t,x,y;s,\xi,\eta)\sum_{k=0}^{n}\left(\xi-\bar{x}(s)\right)^{n-k}\left(\eta-\bar{y}(s)\right)^{k}\mathcal{A}^{\xi,\eta,\zeta}_{n-k,k}(s)\theta(\xi,\eta,\zeta)\\
&= \sum_{k=0}^{n}\left(\mathcal{M}_x(t,s)-\bar{x}(s)\right)^{n-k}\left(\mathcal{M}_y(t,s)-\bar{y}(s)\right)^{k}\int_{\mathbb{R}^3}\mathrm{d}\xi\mathrm{d}\eta\mathrm{d}\zeta\,\delta_{\bar{z}}(\zeta)\Gamma_0(t,x,y;s,\xi,\eta)\mathcal{A}^{\xi,\eta,\zeta}_{n-k,k}(s)\theta(\xi,\eta,\zeta)\\
&= \sum_{k=0}^{n}\left(\mathcal{M}_x(t,s)-\bar{x}(s)\right)^{n-k}\left(\mathcal{M}_y(t,s)-\bar{y}(s)\right)^{k}\int_{\mathbb{R}^3}\mathrm{d}\xi\mathrm{d}\eta\mathrm{d}\zeta\,\delta_{\bar{z}}(\zeta)\theta(\xi,\eta,\zeta)\left(\mathcal{A}^{\xi,\eta,\zeta}_{n-k,k}(s)\right)^*\Gamma_0(t,x,y;s,\xi,\eta)\\
&= \sum_{k=0}^{n}\left(\mathcal{M}_x(t,s)-\bar{x}(s)\right)^{n-k}\left(\mathcal{M}_y(t,s)-\bar{y}(s)\right)^{k}\mathcal{A}^{x,y,z}_{n-k,k}(s)\int_{\mathbb{R}^3}\mathrm{d}\xi\mathrm{d}\eta\mathrm{d}\zeta\,\delta_{\bar{z}}(\zeta)\theta(\xi,\eta,\zeta)\Gamma_0(t,x,y;s,\xi,\eta)\\
&= \mathcal{G}_i(t,s)\mathcal{P}_0(t,s)\theta(x,y,z).
\end{aligned}$$

The first equality follows from the definitions of $\mathcal{P}_0(t,s)$ and $\mathcal{A}_i(s)$. In the second equality we have used (3.25) and pulled the operators $\mathcal{M}_x$ and $\mathcal{M}_y$ out of the integral as they act on the backward variables (x,y,z). In the third equality we have integrated by parts. In the fourth equality we have used the symmetry property of the kernel $\delta_{\bar{z}}(z)\Gamma_0(t,x,y;s,\xi,\eta)$ to replace $\left(\mathcal{A}^{\xi,\eta,\zeta}_{n-k,k}(s)\right)^*$ with $\mathcal{A}^{x,y,z}_{n-k,k}(s)$. We then pulled $\mathcal{A}^{x,y,z}_{n-k,k}(s)$ out of the integral as it acts on the backward variables (x,y,z). The last equality follows from the definitions of $\mathcal{G}_i(t,s)$ and $\mathcal{P}_0(t,s)$. Thus, we have established $\mathcal{P}_0(t,s)\mathcal{A}_i(s) = \mathcal{G}_i(t,s)\mathcal{P}_0(t,s)$, when acting on a function θ that is $C^2(\mathbb{R}^3)$ that is at most exponentially growing.

To complete the proof we must show that terms of the form

$$\mathcal{P}_0(t,t_1)\mathcal{A}_{i_1}(t_1)\mathcal{P}_0(t_1,t_2)\mathcal{A}_{i_2}(t_2)\cdots\mathcal{P}_0(t_{k-1},t_k)\mathcal{A}_{i_k}(t_k)\mathcal{P}_0(t_k,T)\varphi, \tag{3.26}$$

are at least $C^2(\mathbb{R}^3)$ and at most exponentially growing. In fact, we will show that such terms are $C^\infty_e(\mathbb{R}^3)$, where $C^\infty_e(\mathbb{R}^3)$ denotes the space of functions that are $C^\infty(\mathbb{R}^3)$ with derivatives of all orders that are at most exponentially growing. To see this, we note that $\mathcal{P}_0(t_k,T)\varphi = u^{\mathrm{BS}}(t_k)$, where u^{BS} is the Black-Scholes price of a put option. As derivatives of the Black-Scholes put price with respect to z are $C^\infty_e(\mathbb{R})$ it follows

Author Manuscript

Author Manuscript

that $\mathcal{P}_0(t_k, T)\varphi \in C_e^\infty(\mathbb{R}^3)$. Now, note that $C_e^\infty(\mathbb{R}^3)$ is invariant under differentiation, multiplication by a polynomial, and transformation by the semigroup operator $\mathcal{P}_0(t, s)$. It follows that any term of the form (3.26) is a member of $C_e^\infty(\mathbb{R}^3)$. □

Remark 3.5. In fact, Theorem 3.4 holds directly for Call options as well, as derivatives of the Black-Scholes call price with respect to z are $C_e^\infty(\mathbb{R})$.

In the following proposition, we provide an alternative characterization of the approximating sequence (u_n) as the solution of a nested sequence of PDEs. This alternative characterization, which was derived using alternative methods in Lorig et al. (2015a), will be used in Section 4 for the analysis of the accuracy of the approximation.

Proposition 3.6. *Let φ be the payoff of a put option: $\varphi(z) = (\mathrm{e}^k - \mathrm{e}^z)^+$. The sequence of functions (u_n) in (3.17) solves the following nested sequence of Cauchy problems*

$$(\partial_t + \mathcal{A}_0(t))u_0 = 0, \qquad u_0(T) = \varphi, \tag{3.27}$$

$$(\partial_t + \mathcal{A}_0(t))u_n = -\sum_{k=1}^{n} \mathcal{A}_k(t)u_{n-k}, \qquad u_n(T) = 0, \qquad n \geq 1. \tag{3.28}$$

Proof. The proof is by induction. By Duhamel's principle, the solution to (3.27) and the solution to (3.28) with $n = 1$ are

$$u_0(t) = \mathcal{P}_0(t, T)\varphi, \qquad u_1(t) = \int_{t_1}^{T} \mathrm{d}t_1 \mathcal{P}_0(t, t_1)\mathcal{A}_1(t_1)\mathcal{P}_0(t_1, T)\varphi,$$

in agreement with (3.11) and (3.12). We now assume expression (3.12) holds for the first $(n-1)$ terms and show that it holds for the nth term. Once again, using Duhamel's principle, the solution to (3.28) is

$$\begin{aligned}
u_n(t) &= \sum_{k=1}^{n} \int_t^T \mathrm{d}t_0 \mathcal{P}_0(t, t_0)\mathcal{A}_k(t_0)u_{n-k}(t_0) \\
&= \int_t^T \mathrm{d}t_0 \mathcal{P}_0(t, t_0)\mathcal{A}_n(t_0)\mathcal{P}_0(t_0, T)\varphi \\
&\quad + \sum_{k=1}^{n-1} \int_t^T \mathrm{d}t_0 \mathcal{P}_0(t, t_0)\mathcal{A}_k(t_0) \sum_{m=1}^{n-k} \int_{t_0}^T \mathrm{d}t_1 \int_{t_1}^T \mathrm{d}t_2 \cdots \int_{t_{m-1}}^T \mathrm{d}t_m \\
&\quad \sum_{i \in I_{n-k,m}} \mathcal{P}_0(t_0, t_1)\mathcal{A}_{i_1}(t_1)\mathcal{P}_0(t_1, t_2)\mathcal{A}_{i_2}(t_2) \cdots \mathcal{P}_0(t_{m-1}, t_m)\mathcal{A}_{i_m}(t_m)\mathcal{P}_0(t_m, T)\varphi \\
&= \int_t^T \mathrm{d}t_0 \mathcal{P}_0(t, t_0)\mathcal{A}_n(t_0)\mathcal{P}_0(t_0, T)\varphi \\
&\quad + \sum_{k=1}^{n-1} \int_t^T \mathrm{d}t_0 \int_{t_0}^T \mathrm{d}t_1 \int_{t_1}^T \mathrm{d}t_2 \cdots \int_{t_{m-1}}^T \mathrm{d}t_m \\
&\quad \sum_{m=1}^{n-k} \sum_{i \in I_{n-k,m}} \mathcal{P}_0(t, t_0)\mathcal{A}_k(t_0)\mathcal{P}_0(t_0, t_1)\mathcal{A}_{i_1}(t_1)\mathcal{P}_0(t_1, t_2)\mathcal{A}_{i_2}(t_2) \cdots \mathcal{P}_0(t_{m-1}, t_m)\mathcal{A}_{i_m}(t_m)\mathcal{P}_0(t_m, T)\varphi \\
&= \sum_{k=1}^{n} \int_t^T \mathrm{d}t_1 \int_{t_1}^T \mathrm{d}t_2 \cdots \int_{t_{k-1}}^T \mathrm{d}t_k
\end{aligned}$$

$$\sum_{i\in I_{n,k}} \mathcal{P}_0(t,t_1)\mathcal{A}_{i_1}(t_1)\mathcal{P}_0(t_1,t_2)\mathcal{A}_{i_2}(t_2)\cdots\mathcal{P}_0(t_{k-1},t_k)\mathcal{A}_{i_k}(t_k)\mathcal{P}_0(t_k,T)\varphi,$$

which agrees with expression (3.12). □

Remark 3.7. Note that, by (3.16), the order zero price u_0 is simply an integral of the option payoff φ versus a Gaussian kernel Γ_0, just as in the Black-Scholes model. From Theorem 3.4 we see that higher order terms u_n can be obtained by applying the differential operator $\mathcal{L}_n$ to u_0. The operator $\mathcal{L}_n$ acts on the backward variable z, which is present only in the Gaussian kernel Γ_0, producing (Hermite) polynomials in the forward variable ζ multiplied by Γ_0. Thus, every term in the price expansion is of the form

$$u_n(t,z) = \int_{\mathbb{R}} \mathrm{d}\zeta \, \frac{p_n(\zeta)}{\sqrt{2\pi s^2(t,T)}} \exp\left(\frac{-(\zeta - m(t,T))^2}{2s^2(t,T)}\right)\varphi(\zeta).$$

where the function p_n is a polynomial. As such, computation times for approximate prices are comparable to the Black-Scholes model.

4 Accuracy of the option-pricing approximation

The goal of this section is to establish a rigorous error bound for the Nth-order pricing approximation described in the previous sections. We will adapt the methods from Pagliarani and Pascucci (2014), who treat operators $\mathcal{A}(t)$ that are *locally elliptic*, to our current case, where the operator $\mathcal{A}(t)$ is singular (see Remark 3.1). Our main error bound is given in Theorem 4.8 at the end of this section. In order to prove this theorem we introduce $A(t,x,y)$, the symmetric and positive semi-definite diffusion matrix of the (X,Y) process:

$$A(t,x,y) := \frac{1}{2}\begin{pmatrix} 2a(t,x,y) & f(t,x,y) \\ f(t,x,y) & 2b(t,x,y) \end{pmatrix}.$$

We also introduce $D_r(x_0,y_0)$, the Euclidean ball

$$D_r(x_0,y_0) = \{(x,y)\in\mathbb{R}^2 : |(x,y)-(x_0,y_0)| < r\},$$

which is defined for any $(x_0,y_0)\in\mathbb{R}^2$ and $r>0$. Throughout Section 4 we assume the following:

Assumption 4.1. The function u in (3.1) solves the backward Cauchy problem

$$\begin{aligned} (\partial_t + \mathcal{A}(t))\, u(t,x,y,z) &= 0, & (t,x,y,z)&\in[0,T)\times D, \\ u(T,x,y,z) &= \varphi(z), & (x,y,z)&\in D, \end{aligned}$$

where D is a domain in $\mathbb{R}^3$. It is possible, but not required, that $D=\mathbb{R}^3$.

Assumption 4.2.

i) *Local boundedness and global regularity*: the coefficients a,b,c,f belong to $L^\infty_{\mathrm{loc}}([0,T]\times D)$ and satisfy $a(t,\cdot,\cdot), b(t,\cdot,\cdot), c(t,\cdot,\cdot), f(t,\cdot,\cdot)\in C^{N+1}(D)$ for any $t\in[0,T]$.

Author Manuscript

ii) *Local non-degeneracy*: the diffusion matrix of $A = A(t,x,y)$ is positive definite on some cylinder $[0,T]\times D_r(x_0,y_0)$. More precisely, $A = \widetilde{A}$ in $[0,T]\times D_r(x_0,y_0)$ where $\widetilde{A} \in L^\infty([0,T]\times\mathbb{R}^2)$ is a matrix of the form

$$\widetilde{A}(t,x,y) = \frac{1}{2}\begin{pmatrix} 2\widetilde{a}(t,x,y) & \widetilde{f}(t,x,y) \\ \widetilde{f}(t,x,y) & 2\widetilde{b}(t,x,y)\end{pmatrix},$$

such that $\widetilde{A}(t,\cdot,\cdot) \in C_{\mathrm{b}}^{N+1}(\mathbb{R}^2)$ for any $t\in[0,T]$, where C_{b}^N denotes the space of continuously differentiable functions with bounded derivatives up to order N, and

$$M^{-1}|\xi|^2 \le \sum_{i,j=1}^{2} \widetilde{A}_{ij}(t,x,y)\xi_i\xi_j \le M|\xi|^2, \qquad t\in[0,T],\ (x,y),\xi\in\mathbb{R}^2,$$

for some positive constant M. We also require the existence of a function $\widetilde{c}\in L^\infty([0,T]\times\mathbb{R}^2)$ such that $\widetilde{c}(t,\cdot,\cdot)\in C_{\mathrm{b}}^{N+1}(\mathbb{R}^2)$ for any $t\in[0,T]$ and $c=\widetilde{c}$ in $[0,T]\times D_r(x_0,y_0)$.

Assumption 4.3. We assume the payoff function φ is that of a Put option on Z. That is, $\varphi(z) = \left(\mathrm{e}^k - \mathrm{e}^z\right)^+$.

Remark 4.4. Assumptions 4.1 and 4.2 are satisfied by a number of well-known models including the Heston (Heston (1993)), exponential Ornstein-Uhlenbeck (Perelló et al. (2008)) and three-halves stochastic volatility (Carr and Sun (2007)) models.

Remark 4.5. Note that, although the diffusion matrix A of the process (X,Y) is locally positive definite, the diffusion matrix (3.4) of the process (X,Y,Z) remains singular. Thus, Cauchy problem (3.2) is *not parabolic* at any point. This issue, which is not handled in Pagliarani and Pascucci (2014), presents a technical challenge that must be overcome in order to establish error estimates for our pricing approximation $\bar{u}_n$.

In order to cope with the double degeneracy of the pricing operator (recall, we have a partial degeneracy in the (x,y) variables and a global degeneracy in the z variable), we now use an elliptic regularization technique, which is a classical method in the theory of degenerate PDEs; see, e.g., Theorem 5.2 of Bony (1969). Specifically, we introduce a process Z^ε, which is a modification of the dynamics of Z in (2.4). We define

$$\mathrm{d}Z_t^\varepsilon := \mathrm{d}Z_t - \frac{1}{2}\varepsilon^2\mathrm{d}t + \varepsilon\mathrm{d}W_t^z, \qquad \mathrm{d}\langle W^x, W^z\rangle = 0, \qquad \mathrm{d}\langle W^y, W^z\rangle = 0, \qquad \varepsilon\ge 0.$$

We denote by $\mathcal{A}^\varepsilon(t)$ the infinitesimal generator of the Markov process (X,Y,Z^ε) and by u^ε the solution of Cauchy problem related to $\mathcal{A}^\varepsilon(t)$, with final datum φ. Specifically

$$\begin{aligned} (\partial_t + \mathcal{A}^\varepsilon(t))\, u^\varepsilon(t,x,y,z) &= 0, & (t,x,y,z)&\in[0,T)\times D_\varepsilon, \\ u^\varepsilon(T,x,y,z) &= \varphi(z), & (x,y,z)&\in D_\varepsilon, \end{aligned}$$

where D_ε is some domain of $\mathbb{R}^3$. Thus, u^ε represents the price of a European Put option written on (X,Y,Z^ε). Assumption 4.2-i) guarantees we can construct $\bar{u}_N^\varepsilon$, the N-th order approximation of u^ε, by replacing $\mathcal{A}(t)$ with $\mathcal{A}^\varepsilon(t)$ in Definition 3.3. Moreover, for any $\varepsilon>0$, $\mathcal{A}^\varepsilon(t)$ and φ satisfy the assumptions of Theorem 3.1 in Pagliarani and Pascucci (2014), in which local error estimates for $|u^\varepsilon - \bar{u}_N^\varepsilon|$ are established. Below we

Author Manuscript

Author Manuscript

prove that such error estimates are *uniform in* ε and therefore error bounds for the price approximation $\bar{u}_N$ of options written on the Markov process (X, Y, Z) will follow in the limit as $\varepsilon \to 0$.

It is useful at this point to introduce the process V^ε, which satisfies the following SDE:

$$\mathrm{d}V_t^\varepsilon = \frac{1}{2}\left(\beta(1-\beta)\sigma^2(t, X_t, Y_t) - \varepsilon^2\right)\mathrm{d}t + \varepsilon \mathrm{d}W_t^z.$$

We note that the dynamics of Z^ε can be written as follows

$$\mathrm{d}Z_t^\varepsilon = \beta\,\mathrm{d}X_t + \mathrm{d}V_t^\varepsilon.$$

Therefore, rather than considering the generator of (X, Y, Z^ε), we can consider the generator of (X, Y, V^ε), which (with a slight abuse of notation) we denote again by $\mathcal{A}^\varepsilon(t)$. This operator separates into an operator $\mathcal{X}(t)$, which takes derivatives with respect to (x, y), and an operator $\mathcal{V}^\varepsilon(t)$, which takes derivatives with respect to v. That is,

$$\begin{aligned}
\mathcal{A}^\varepsilon(t) &= \mathcal{X}(t) + \mathcal{V}^\varepsilon(t),\\
\mathcal{X}(t) &= a(t,x,y)\left(\partial_x^2 - \partial_x\right) + b(t,x,y)\partial_y^2 + f(t,x,y)\partial_x\partial_y + c(t,x,y)\partial_y,\\
\mathcal{V}^\varepsilon(t) &= \frac{\varepsilon^2}{2}\left(\partial_v^2 - \partial_v\right) + a(t,x,y)\beta(1-\beta)\partial_v.
\end{aligned}$$

The first step in the proof of Theorem 3.1 in Pagliarani and Pascucci (2014) consists of extending operator $\mathcal{A}^\varepsilon(t)$, which is defined on $[0,T]\times D$, to a uniformly elliptic operator $\widetilde{\mathcal{A}}^\varepsilon(t)$ on $[0,T]\times\mathbb{R}^3$. This can be done by virtue of Assumption 4.2-ii). Indeed, for any $\varepsilon \geq 0$, it suffices to define

$$\begin{aligned}
\widetilde{\mathcal{A}}^\varepsilon(t) &= \widetilde{\mathcal{X}}(t) + \widetilde{\mathcal{V}}^\varepsilon(t),\\
\widetilde{\mathcal{X}}(t) &= \widetilde{a}(t,x,y)\left(\partial_x^2 - \partial_x\right) + \widetilde{b}(t,x,y)\partial_y^2 + \widetilde{f}(t,x,y)\partial_x\partial_y + \widetilde{c}(t,x,y)\partial_y,\\
\widetilde{\mathcal{V}}^\varepsilon(t) &= \frac{\varepsilon^2}{2}\left(\partial_v^2 - \partial_v\right) + \widetilde{a}(t,x,y)\beta(1-\beta)\partial_v.
\end{aligned}$$

By Assumption 4.2, $\mathcal{A}^\varepsilon(t) = \widetilde{\mathcal{A}}^\varepsilon(t)$ and $\mathcal{X}(t) = \widetilde{\mathcal{X}}(t)$ in $[0,T]\times D_r(x_0,y_0)\times\mathbb{R}$. Notice that $\widetilde{\mathcal{A}}^\varepsilon(t)$ and $\widetilde{\mathcal{X}}(t)$ are uniformly elliptic operators on $[0,T]\times\mathbb{R}^3$ and $[0,T]\times\mathbb{R}^2$ respectively. Moreover, $(\partial_t + \widetilde{\mathcal{A}}^\varepsilon(t))$ is uniformly parabolic and has a fundamental solution, denoted by

$$\widetilde{\Gamma}^\varepsilon = \widetilde{\Gamma}^\varepsilon(t,x,y,v;T,x',y',v'), \qquad t < T, \tag{4.1}$$

which (by definition) is the solution to

$$\begin{aligned}
(\partial_t + \widetilde{\mathcal{A}}^\varepsilon(t))\widetilde{\Gamma}^\varepsilon(t,x,y,v;T,x',y',v') &= 0, \qquad && (t,x,y,v)\in[0,T)\times\mathbb{R}^3,\\
\widetilde{\Gamma}^\varepsilon(T,\cdot,\cdot,\cdot;T,x',y',v') &= \delta_{x',y',v'}.
\end{aligned}$$

In the following lemma, we show that $\widetilde{\Gamma}^\varepsilon$ satisfies some Gaussian estimates.

Lemma 4.6. *Let $i,j,h,k\in\mathbb{N}_0$ with $h+k\leq N+2$, and $\bar{T}>0$. Then, under Assumption 4.2, we have*

$$\left|(x-x')^i(y-y')^j\partial_x^h\partial_y^k\widetilde{\Gamma}^\varepsilon(t,x,y,v;T,x',y',v')\right| \leq \mathbf{c}_0(T-t)^{\frac{i+j-h-k}{2}}\Gamma_{\text{heat}}^{(M,\varepsilon)}(t,x,y,v;T,x',y',v') \tag{4.2}$$

for any $x, y, v, x', y', v' \in \mathbb{R}$, $0 \le t < T \le \bar{T}$ *and* $\varepsilon \in (0,1]$. *Here,* $\Gamma_{\text{heat}}^{(M,\varepsilon)}$ *denotes the fundamental solution of the heat operator*

$$\partial_t + M(\partial_{xx} + \partial_{yy}) + \frac{\varepsilon^2}{2}\partial_{vv},$$

and $\mathbf{c}_0$ *is a positive constant that depends only on* M, N, i, j *and* $\bar{T}$. *In particular, the constant* $\mathbf{c}_0$ *is independent of* ε.

Proof. Estimate (4.2) differs slightly from the classical Gaussian estimates for parabolic equations (cf. Friedman (1964); see also Di Francesco and Pascucci (2005), Pascucci (2011) for a more recent and general presentation) because the operator $(\partial_t + \widetilde{\mathcal{A}}^{\varepsilon}(t))$, while parabolic, is *not uniformly parabolic* with respect to $\varepsilon \in (0,1]$. Nevertheless, the thesis can be proved by mimicking the classical argument which is based on the parametrix method and carefully checking that the constant $\mathbf{c}_0$ is independent of ε. In particular, the main ingredients in the parametrix construction are some uniform-in-ε, Gaussian estimates (see, for instance, Proposition 3.1 in Di Francesco and Pascucci (2005)), which we now describe. For any fixed $(\bar{x}, \bar{y}) \in \mathbb{R}^2$, we denote by $\widetilde{\mathfrak{X}}_{\bar{x},\bar{y}}(t)$ the operator obtained by freezing at $(\bar{x}, \bar{y})$ the coefficients of $\widetilde{\mathfrak{X}}(t)$ and we set

$$\widetilde{\mathcal{A}}^{\varepsilon}_{\bar{x},\bar{y}}(t) := \widetilde{\mathfrak{X}}_{\bar{x},\bar{y}}(t) + \frac{\varepsilon^2}{2}\partial_v^2.$$

Let $\widetilde{\Gamma}^{\varepsilon}_{\bar{x},\bar{y}}$ and $\widetilde{\Gamma}_{\bar{x},\bar{y}}$ be the fundamental solutions corresponding to $(\partial_t + \widetilde{\mathcal{A}}^{\varepsilon}_{\bar{x},\bar{y}})$ and $(\partial_t + \widetilde{\mathfrak{X}}_{\bar{x},\bar{y}})$ respectively. Then for every $\bar{x}, \bar{y}, x, y, v, x', y', v' \in \mathbb{R}$, $0 \le t < T \le \bar{T}$ and $\varepsilon \in (0,1]$, we have

$$M^{-2}\Gamma_{\text{heat}}^{(M^{-1},\varepsilon)}(t,x,y,v;T,x',y',v') \le \widetilde{\Gamma}^{\varepsilon}_{\bar{x},\bar{y}}(t,x,y,v;T,x',y',v') \le M^2\Gamma_{\text{heat}}^{(M,\varepsilon)}(t,x,y,v;T,x',y',v'). \quad (4.3)$$

Estimate (4.3) can be readily proved as in Proposition 3.1 in Di Francesco and Pascucci (2005), by noting that $\widetilde{\Gamma}^{\varepsilon}_{\bar{x},\bar{y}} = \widetilde{\Gamma}_{\bar{x},\bar{y}}\Gamma_{\varepsilon}$ where Γ_{ε} is the fundamental solution of the one-dimensional heat (parabolic) operator $(\partial_t + \frac{\varepsilon^2}{2}\partial_{vv})$. Notice that (4.3) is uniform in ε (i.e. the constants in the estimates are independent of ε). Based on this fact, the estimate (4.2), with $\mathbf{c}_0$ independent of ε, follows by the parametrix method. □

Lemma 4.7. *Let Assumption 4.2 hold. Denote by* $\widetilde{\Gamma}^{\varepsilon}$ *the fundamental solution in* (4.1) *corresponding to* $(\partial_t + \widetilde{\mathcal{A}}^{\varepsilon}(t))$. *Denote by* $\bar{\Gamma}^{\varepsilon}_N$ *the* N*th-order approximation of* $\widetilde{\Gamma}^{\varepsilon}$, *constructed using* $(\bar{x}(\cdot), \bar{y}(\cdot)) = (x, y)$. *Then we have*

$$\left|\widetilde{\Gamma}^{\varepsilon}(t,x,y,v;T,x',y',v') - \bar{\Gamma}^{\varepsilon}_N(t,x,y,v;T,x',y',v')\right| \le \mathbf{c}_1(T-t)^{\frac{N+1}{2}}\Gamma_{\text{heat}}^{(M,\varepsilon)}(t,x,y,v;T,x',y',v'), \quad (4.4)$$

for any $x, y, v, x', y', v' \in \mathbb{R}$, $0 \le t < T$ *and* $\varepsilon \in (0,1]$, *where* $\mathbf{c}_1$ *is a positive constant that depends on* M, N, T *but is independent of* ε.

Proof. Using the uniform in ε estimate (4.2) and the ellipticity of $\mathcal{A}^{\varepsilon}(t)$, we can repeat step by step the proof of (Lorig et al., 2015a, Theorem 3.10). The key ingredient in the modified proof is to verify that, as $\mathbf{c}_0$ in (4.2) does not depend on ε, neither does $\mathbf{c}_1$. □

We are now in a position to state our main error estimate. For any $\varepsilon \ge 0$, let $\widetilde{u}^{\varepsilon}$ be the classical bounded solution of Cauchy problem

$$\left(\partial_t + \widetilde{\mathcal{A}}^{\varepsilon}(t)\right)\widetilde{u}^{\varepsilon}(t,x,y,z) = 0, \qquad (t,x,y,z) \in [0,T) \times \mathbb{R}^3,$$

Author Manuscript

$$\widetilde{u}^{\varepsilon}(T,x,y,z) = \left(\mathrm{e}^{k} - \mathrm{e}^{z}\right)^{+}, \qquad (x,y,z) \in \mathbb{R}^3.$$

For $\varepsilon = 0$ we will generally omit the superscript and simply write $\widetilde{u}$ instead of $\widetilde{u}^0$.

Theorem 4.8. *Let Assumptions 4.1, 4.2 and 4.3 hold. Let $\bar{u}^{\varepsilon}_N$, $\varepsilon \geq 0$, denote the N-th order approximation of $\widetilde{u}^{\varepsilon}$, which is constructed as in (3.11) with $(\bar{x}(\cdot), \bar{y}(\cdot)) = (x,y)$ and with $\mathcal{A}(t)$ replaced by $\widetilde{\mathcal{A}}^{\varepsilon}(t)$. Let $\bar{u}_N := \bar{u}^{\varepsilon}_N|_{\varepsilon=0}$. Note that $\bar{u}_N$ coincides with the N-th order approximation of u when $(x,y,z) \in (D_{\delta r}(x_0,y_0) \times \mathbb{R}) \cap D$. Then for any $\delta \in (0,1)$ we have*

$$|u(t,x,y,z) - \bar{u}_N(t,x,y,z)| \leq \mathbf{c}_2 (T-t)^{\frac{N+2}{2}}, \qquad 0 \leq t < T,\ (x,y,z) \in (D_{\delta r}(x_0,y_0) \times \mathbb{R}) \cap D.$$

The constant $\mathbf{c}_2$ depends only on δ, k, M, N and T.

Proof. Firstly, we remark explicitly that if $(x,y) \in D_r(x_0,y_0)$ then, for any $\varepsilon \geq 0$, $\bar{u}^{\varepsilon}_N$ coincides with the N-th order approximation of u^{ε} because $\mathcal{A}(t) \equiv \widetilde{\mathcal{A}}(t)$ in $[0,T] \times D_r(x_0,y_0) \times \mathbb{R}$. Then, integrating estimate (4.4) against the payoff function we obtain

$$|\widetilde{u}^{\varepsilon}(t,x,y,z) - \bar{u}^{\varepsilon}_N(t,x,y,z)| \leq \mathbf{c}_1 (T-t)^{\frac{N+1}{2}}, \qquad 0 \leq t < T,\ (x,y,z) \in \mathbb{R}^3. \tag{4.5}$$

By exploiting the Lipschitz regularity and boundedness of the Put payoff, we have a more refined estimate with the power $\frac{N+2}{2}$ replacing $\frac{N+1}{2}$ in the exponent of $(T-t)$ in (4.5). As the operator $\mathcal{A}^{\varepsilon}(t)$, $\varepsilon > 0$, and the payoff function φ satisfy the assumptions of Theorem 3.1 in Pagliarani and Pascucci (2014), we can pass from the global error estimate for $\widetilde{u}^{\varepsilon}$ to the local estimate for u^{ε}

$$|u^{\varepsilon}(t,x,y,z) - \bar{u}^{\varepsilon}_N(t,x,y,z)| \leq \mathbf{c}_2 (T-t)^{\frac{N+2}{2}}, \qquad 0 \leq t < T,\ (x,y,z) \in (D_{\delta r}(x_0,y_0) \times \mathbb{R}) \cap D.$$

By Lemma 4.6, the above estimate is uniform in ε. Therefore, we can use the maximum principle and pass to $\varepsilon = 0$ as in the proof of Theorem 5.2 of Bony (1969) (see also Theorem 2.5 of Lanconelli and Pascucci (1999)) to conclude the proof. □

5 Implied volatility

In this section, we translate our price expansion for a call option with payoff function $\varphi(z) = (\mathrm{e}^z - \mathrm{e}^k)^+$ into an expansion in implied volatility. We emphasize that our implied volatility expansion is *heurisitic*; we shall not attempt to provide a rigorous accuracy result for this expansion. For the interested reader, accuracy results for implied volatility of the ETF (not the LETF) have been established in Lorig et al. (2015b).

To ease notation, in what follows, we shall suppress much of the dependence on (t,T,x,y,z,k). However, one should keep in mind that prices and implied volatilities do depend on these quantities, even if this is not explicitly indicated. We begin our analysis by recalling the definitions of the Black-Scholes call price and implied volatility.

Definition 5.1. The *Black-Scholes Call price* $u^{\mathrm{BS}} : \mathbb{R}^+ \to \mathbb{R}^+$ is given by

$$u^{\mathrm{BS}}(\sigma) := \mathrm{e}^z \mathcal{N}(d_+(\sigma)) - \mathrm{e}^k \mathcal{N}(d_-(\sigma)), \qquad d_{\pm}(\sigma) := \frac{1}{\sigma\sqrt{\tau}} \left(z - k \pm \frac{\sigma^2 \tau}{2} \right), \qquad \tau := T - t, \tag{5.1}$$

where $\mathcal{N}$ is the CDF of a standard normal random variable.

Author Manuscript

Author Manuscript

Definition 5.2. For fixed (t, T, z, k), the *implied volatility* corresponding to a call price $u \in ((\mathrm{e}^z - \mathrm{e}^k)^+, \mathrm{e}^z)$ is defined as the unique strictly positive real solution σ of the equation

$$u^{\mathrm{BS}}(\sigma) = u, \tag{5.2}$$

where u^{BS} is given by (5.1).

Theorem 5.3. *For a European call option with payoff function $\varphi(z) = (\mathrm{e}^z - \mathrm{e}^k)^+$ we have*

$$u_0 = u^{\mathrm{BS}}(\sigma_0), \qquad\qquad \sigma_0^2 = \frac{2\beta^2}{T-t}\int_t^T \mathrm{d}s\, a_{0,0}(s). \tag{5.3}$$

Proof. The proof follows directly from (3.16) with $\varphi(z) = (\mathrm{e}^z - \mathrm{e}^k)^+$. □

From Theorem 5.3 we note that the price expansion (3.11) is of the form

$$u = \bar{u}_N + \delta u_N = u^{\mathrm{BS}}(\sigma_0) + \sum_{n=1}^{N} u_n + \delta u_N, \tag{5.4}$$

where $\delta u_N = u - \bar{u}_N$. As shown in Lorig et al. (2015b) and Jacquier and Lorig (2013), the special form (5.4) lends itself to an expansion

$$\sigma = \sigma_0 + \eta, \qquad\qquad \eta = \sum_{n=1}^{N} \sigma_n + \delta\sigma_N,$$

of implied volatility. To see this, one expands $u^{\mathrm{BS}}(\sigma)$ as a Taylor series about the point σ_0. For η small enough (i.e., within the radius of convergence of the Taylor series expansion of u^{BS} about the point σ_0) we have

$$\begin{aligned} u^{\mathrm{BS}}(\sigma) &= u^{\mathrm{BS}}(\sigma_0 + \eta) \\ &= u^{\mathrm{BS}}(\sigma_0) + \eta\,\partial_\sigma u^{\mathrm{BS}}(\sigma_0) + \frac{1}{2!}\eta^2\partial_\sigma^2 u^{\mathrm{BS}}(\sigma_0) + \frac{1}{3!}\eta^3\partial_\sigma^3 u^{\mathrm{BS}}(\sigma_0) + \dots. \end{aligned} \tag{5.5}$$

Inserting expansions (5.4) and (5.5) into equation (5.2), one can solve iteratively for every term in the sequence $(\sigma_n)_{1\le n\le N}$. We define the *Nth-order approximation of implied volatility* as

$$\bar{\sigma}_N = \sum_{n=0}^{N} \sigma_n.$$

The first four terms in the sum, which are enough to provide an accurate approximation of implied volatility, are σ_0, given by (5.3), and

$$\sigma_1 = \frac{u_1}{\partial_\sigma u^{\mathrm{BS}}(\sigma_0)}, \qquad \sigma_2 = \frac{u_2 - \frac{1}{2}\sigma_1^2\partial_\sigma^2 u^{\mathrm{BS}}(\sigma_0)}{\partial_\sigma u^{\mathrm{BS}}(\sigma_0)}, \qquad \sigma_3 = \frac{u_3 - \left(\sigma_2\sigma_1\partial_\sigma^2 + \frac{1}{3!}\sigma_1^3\partial_\sigma^3\right)u^{\mathrm{BS}}(\sigma_0)}{\partial_\sigma u^{\mathrm{BS}}(\sigma_0)}. \tag{5.6}$$

A general expression for the nth-order term can be found in Lorig et al. (2015b); Jacquier and Lorig (2013).

As written, the expressions in (5.6) are not particularly useful. Indeed $u^{\mathrm{BS}}(\sigma_0)$ and u_n are Gaussian integrals, which are not numerically intensive to compute, but do not give much explicit information about how implied volatility depends on $(t, T, x, y, z, k, \beta)$. However, using (5.1) a direct computation shows

$$\frac{\partial_\sigma^2 u^{\mathrm{BS}}(\sigma)}{\partial_\sigma u^{\mathrm{BS}}(\sigma)} = \frac{(k-z)^2}{\tau\sigma^3} - \frac{\tau\sigma}{4}, \qquad \frac{\partial_\sigma^3 u^{\mathrm{BS}}(\sigma)}{\partial_\sigma u^{\mathrm{BS}}(\sigma)} = \frac{(k-z)^4}{\tau^2\sigma^6} - \left(\frac{3}{\tau\sigma^4} - \frac{1}{2\sigma^2}\right)(k-z)^2 + \frac{\tau^2\sigma^2}{16} - \frac{\tau}{4}. \tag{5.7}$$

Author Manuscript

In general, every term of the form $\partial_\sigma^n u^{\mathrm{BS}}(\sigma_0)/\partial_\sigma u^{\mathrm{BS}}(\sigma_0)$ can be computed explicitly. Moreover, terms of the form $u_n/\partial_\sigma u^{\mathrm{BS}}(\sigma_0)$ can also be computed explicitly. To see this, we note from Theorems 3.4 and 5.3 that

$$u_n = \mathcal{L}_n(t,T)u_0 = \widetilde{\mathcal{L}}_n(t,T)u^{\mathrm{BS}}(\sigma_0),$$

where

$$\widetilde{\mathcal{L}}_n(t,T) = \sum_{k=1}^{n}\int_t^T \mathrm{d}t_1 \int_{t_1}^T \mathrm{d}t_2 \cdots \int_{t_{k-1}}^T \mathrm{d}t_k \sum_{i\in I_{n,k}} \mathcal{G}_{i_1}(t,t_1)\cdots\mathcal{G}_{i_{k-1}}(t,t_{k-1})\widetilde{\mathcal{G}}_{i_k}(t,t_k),$$
$$\widetilde{\mathcal{G}}_n(t,t_i) := \sum_{k=0}^{n}\left(\mathcal{M}_x(t,t_i)-\bar{x}(t_i)\right)^{n-k}\left(\mathcal{M}_y(t,t_i)-\bar{y}(t_i)\right)^{k} a_{n-k,k}(t_i)\beta^2(\partial_z^2-\partial_z)u^{\mathrm{BS}}(\sigma_0).$$

Thus, u_n is a finite sum of the form

$$u_n = \sum_m \mathcal{X}_{n,m}\partial_z^m(\partial_z^2-\partial_z)u^{\mathrm{BS}}(\sigma_0), \tag{5.8}$$

where the coefficients $(\mathcal{X}_{n,m})$ are (t,T,x,y)-dependent constants, which can be computed from Theorem 3.4. Now, using (5.1), a direct computation shows

$$\frac{\partial_z^m(\partial_z^2-\partial_z)u^{\mathrm{BS}}(\sigma_0)}{\partial_\sigma u^{\mathrm{BS}}(\sigma_0)} = \left(\frac{-1}{\sqrt{2\sigma_0^2\tau}}\right)^m \frac{H_n(w)}{\tau\sigma_0}, \qquad w := \frac{z-k-\frac{1}{2}\sigma_0^2\tau}{\sigma\sqrt{2\sigma_0^2\tau}}, \tag{5.9}$$

where $H_n(z) := (-1)^n \mathrm{e}^{z^2}\partial_z^n \mathrm{e}^{-z^2}$ is the n-th Hermite polynomial. Combining (5.8) with (5.9) we have

$$\frac{u_n}{\partial_\sigma u^{\mathrm{BS}}(\sigma_0)} = \sum_m \mathcal{X}_{n,m}\left(\frac{-1}{\sqrt{2\sigma_0^2\tau}}\right)^m \frac{H_n(w)}{\tau\sigma_0}. \tag{5.10}$$

Finally, from (5.7) and (5.10), we see that all terms in the implied volatility expansion (5.6) are polynomials in log-*moneyness* $\lambda := (k-z)$. Explicit expressions for $(\sigma_n)_{n\le 3}$ under different models will be given in Section 6. A general expression for $(\sigma_n)_{n\le 2}$ in the time-homogeneous LSV setting is given below. We denote by

$$\lambda = k-z, \qquad \tau = T-t, \qquad (X_t,Y_t) = (x,y),$$

and we choose the expansion point of our Taylor series approximation as $(\bar{x}(\cdot),\bar{y}(\cdot)) = (x,y)$. We have

$$\sigma_0 = |\beta|\sqrt{2a_{0,0}}, \qquad \sigma_1 = \sigma_{1,0}+\sigma_{0,1}, \qquad \sigma_2 = \sigma_{2,0}+\sigma_{1,1}+\sigma_{0,2},$$

where

$$\sigma_{1,0} = \frac{\tau}{4}\left((\beta-1)\sigma_0 a_{1,0}\right) + \frac{1}{2\sigma_0}\left(\beta a_{1,0}\right)\lambda,$$
$$\sigma_{0,1} = \frac{\tau}{4\sigma_0}\left(\beta^2 a_{0,1}\left(2c_{0,0}+\beta f_{0,0}\right)\right) + \frac{1}{2\sigma_0^3}\left(\beta^3 a_{0,1} f_{0,0}\right)\lambda,$$
$$\sigma_{2,0} = \frac{\tau}{24\sigma_0}\left(2\sigma_0^2 a_{2,0} - 3\beta^2 a_{1,0}^2\right) + \frac{\tau^2}{96\beta^2}\left(\beta^2(2\beta(2\beta-5)+5)\sigma_0 a_{1,0}^2 + 4(\beta-1)^2\sigma_0^3 a_{2,0}\right)$$
$$+ \frac{\tau}{24\beta\sigma_0}\left(-(\beta-1)\left(\beta^2 a_{1,0}^2 - 4\sigma_0^2 a_{2,0}\right)\right)\lambda + \frac{1}{12\sigma_0^3}\left(2\sigma_0^2 a_{2,0} - 3\beta^2 a_{1,0}^2\right)\lambda^2,$$

$$
\begin{aligned}
\sigma_{1,1} = & \frac{\tau}{12\sigma_0^3}\left(\beta^2\left(a_{0,1}\left(\beta^2 a_{1,0}f_{0,0} - 2\sigma_0^2 f_{1,0}\right) + \sigma_0^2 a_{1,1} f_{0,0}\right)\right) \\
& + \frac{\tau^2}{48\sigma_0}\left(a_{0,1}\left(\beta^2 a_{1,0}\left(2(\beta-1)c_{0,0} - \beta f_{0,0}\right) + 2(\beta-1)\sigma_0^2\left(2c_{1,0} + \beta f_{1,0}\right)\right) + 2(\beta-1)\sigma_0^2 a_{1,1}\left(2c_{0,0} + \beta f_{0,0}\right)\right) \\
& + \frac{\tau}{24\sigma_0^3}\left(\beta\left(a_{0,1}\left(5\beta^2 a_{1,0}\left((1-2\beta)f_{0,0} - 2c_{0,0}\right) + 2\sigma_0^2\left(2c_{1,0} + (2\beta-1)f_{1,0}\right)\right) + 2\sigma_0^2 a_{1,1}\left(2c_{0,0} + (2\beta-1)f_{0,0}\right)\right)\right)\lambda \\
& + \frac{1}{6\sigma_0^5}\left(\beta^2\left(a_{0,1}\left(\sigma_0^2 f_{1,0} - 5\beta^2 a_{1,0}f_{0,0}\right) + \sigma_0^2 a_{1,1} f_{0,0}\right)\right)\lambda^2, \\
\sigma_{0,2} = & \frac{\tau}{24\sigma_0^5}\left(12\beta^2\sigma_0^4 a_{0,2}b_{0,0} - 4\beta^4\sigma_0^2\left(2a_{0,1}^2 b_{0,0} + a_{0,1}f_{0,0}f_{0,1} + a_{0,2}f_{0,0}^2\right) + 9\beta^6 a_{0,1}^2 f_{0,0}^2\right) \\
& + \frac{\tau^2}{24\sigma_0^3}\left(\beta^2\left(\sigma_0^2\left(-2\beta^2 a_{0,1}^2 b_{0,0} + a_{0,1}\left(2c_{0,0} + \beta f_{0,0}\right)\left(2c_{0,1} + \beta f_{0,1}\right) + a_{0,2}\left(2c_{0,0} + \beta f_{0,0}\right)^2\right) - 3\beta^2 a_{0,1}^2 c_{0,0}\left(c_{0,0} + \beta f_{0,0}\right)\right)\right) \\
& + \frac{\tau}{24\sigma_0^5}\left(\beta^3\left(-9\beta^2 a_{0,1}^2 f_{0,0}\left(2c_{0,0} + \beta f_{0,0}\right) + 4\sigma_0^2 a_{0,2}f_{0,0}\left(2c_{0,0} + \beta f_{0,0}\right) + 4\sigma_0^2 a_{0,1}\left(f_{0,1}\left(c_{0,0} + \beta f_{0,0}\right) + c_{0,1}f_{0,0}\right)\right)\right)\lambda \\
& + \frac{1}{12\sigma_0^7}\left(\beta^4\left(2\sigma_0^2\left(2a_{0,1}^2 b_{0,0} + a_{0,1}f_{0,0}f_{0,1} + a_{0,2}f_{0,0}^2\right) - 9\beta^2 a_{0,1}^2 f_{0,0}^2\right)\right)\lambda^2.
\end{aligned}
$$

5.1 Comparison to other implied volatility expansions

As previously mentioned, when $\beta = 1$, options written on the log LETF Z are equivalent to options written the log ETF X. In this special case, the implied volatility expansion discussed in this manuscript reduces to the implied volatility expansion developed in Lorig et al. (2015b). If one additionally chooses $(\bar{x}, \bar{y}) = (x, y)$, then the implied volatility approximation given in Lorig et al. (2015b) is equivalent to the implied volatility expansion given in Bompis and Gobet (2013). However, the expansion presented here and in Lorig et al. (2015b) is derived using PDE methods, whereas the expansion presented in Bompis and Gobet (2013) is developed using tools from Malliavin calculus. As of yet, the implied volatility approximation of Bompis and Gobet (2013) have not been extended to options on LETFs.

We note that, for options written on the ETF X, extensive comparisons to other implied volatility expansions have been carried out in Lorig et al. (2015b). In particular, for the Heston model, the approximation method presented here is compared to the approximation method in Forde et al. (2012), for CEV, it is compared to the approximation method of Hagan and Woodward (1999), and for SABR it is compared to the approximation of Hagan et al. (2002). However, neither Forde et al. (2012), Hagan and Woodward (1999) nor Hagan et al. (2002) develop approximations for implied volatilities written on the LETF Z, as we do here.

Two other methods one might conceivably use to compute approximate options prices and implied volatilities on LETFs are the *heat kernel* and *large deviations* methods, which are discussed, for example, in Armstrong et al. (2014); Henry-Labordère (2009); Gatheral et al. (2012). Generally speaking, these methods all rely on computing geodesic distances on a Riemannian manifold whose metric is the inverse of the covariance matrix of the underlying diffusion. It is not clear how one would compute large deviation estimates and geodesic distances for the metric associated with the process (X, Y, Z) as the diffusion matrix is singular (see Remark 3.1).

Author Manuscript

Author Manuscript

5.2 Implied volatility and log-moneyness scaling

Let us continue to work in the time-homogeneous setting. Let $\sigma_Z(\tau,\lambda)$ be the implied volatility of a call written on the LETF Z with time to maturity τ and log-moneyness $\lambda=(k-z)$ and let $\sigma_X(\tau,\lambda)$ be the implied volatility of a call written on the ETF X time to maturity τ and log-moneyness $\lambda=(k-x)$. The expressions above provide an explicit approximation for $\sigma_Z(\tau,\lambda)$ and $\sigma_X(\tau,\lambda)$ in a general time-homogeneous LSV setting (for $\sigma_X(\tau,\lambda)$, simply set $\beta=1$). These expressions show the highly non-trivial dependence of the implied volatility on the leverage ratio β, and are useful for the purposes of calibration. The implied volatility surfaces $(\tau,\lambda)\mapsto\sigma_Z(\tau,\lambda)$ and $(\tau,\lambda)\mapsto\sigma_X(\tau,\lambda)$ can potentially behave very differently. Nevertheless, for price comparison across leverage ratios, it would practical to relate them, albeit heuristically or approximately. To this end, we now introduce some intuitive scalings. Examining the lowest-order terms σ_0 and σ_1 we observe

$$\text{LETF}: \qquad \sigma_Z \approx |\beta|\sqrt{2a_{0,0}} + |\beta|\left(\frac{a_{1,0}}{2\sqrt{2a_{0,0}}} + \frac{a_{0,1}f_{0,0}}{2(2a_{0,0})^{3/2}}\right)\frac{\lambda}{\beta} + \mathcal{O}(\tau), \tag{5.11}$$

$$\text{ETF}: \qquad \sigma_X \approx \sqrt{2a_{0,0}} + \left(\frac{a_{1,0}}{2\sqrt{2a_{0,0}}} + \frac{a_{0,1}f_{0,0}}{2(2a_{0,0})^{3/2}}\right)\lambda + \mathcal{O}(\tau). \tag{5.12}$$

Comparing σ_Z with σ_X, we see two effects from the leverage ratio β. First, the vertical axis of σ_Z is scaled by a factor of $|\beta|$. Second, the horizontal axis is scaled by a factor of $1/\beta$. In particular, this means that if $\beta<0$ the slopes of σ_X and σ_Z will have opposite signs. For small τ the contribution of the $\mathcal{O}(\tau)$ terms in the expansion will be insignificant. In light of the above observations, it is natural to introduce $\sigma_X^{(\beta)}$ and $\sigma_Z^{(1/\beta)}$, the *scaled implied volatilities*, which we define as

$$\sigma_X^{(\beta)}(\tau,\lambda) := |\beta|\sigma_X(\tau,\lambda/\beta), \qquad\qquad \sigma_Z^{(1/\beta)}(\tau,\lambda) := \frac{1}{|\beta|}\sigma_Z(\tau,\beta\,\lambda). \tag{5.13}$$

These definitions offer two ways to link the implied volatilities surfaces σ_X and σ_Z. Viewed one way, the ETF implied volatility $\sigma_X(\tau,\lambda)$ should roughly coincide with the LETF implied volatility $\frac{1}{|\beta|}\sigma_Z(\tau,\beta\lambda)$. Conversely, the LETF implied volatility $\sigma_Z(\tau,\lambda)$ should be close to the ETF implied volatility $|\beta|\sigma_X(\tau,\lambda/\beta)$. In other words, from (5.11), (5.12) and (5.13), we see that for small τ

$$\sigma_Z(\tau,\lambda)\approx\sigma_X^{(\beta)}(\tau,\lambda), \qquad\qquad \sigma_X(\tau,\lambda)\approx\sigma_Z^{(1/\beta)}(\tau,\lambda). \tag{5.14}$$

In Figure 1, using empirical options data from the S&P500-based ETF and LETFs, we plot σ_Z and $\sigma_Z^{(1/\beta)}$, the unscaled and scaled implied volatilities, respectively. The figure demonstrates the pronounced effect of the scaling argument. Prior to scaling (left panel), the implied volatilities of the LETFs, SSO ($\beta=+2$) and SDS ($\beta=-2$), have much higher values than those of the unleveraged ETF SPY ($\beta=+1$). Moreover, the SDS implied volatility is increasing in log-moneyness. After scaling the LETF implied volatilities according to (5.13) (right panel), they are brought very close to the ETF implied volatility and they are now all downward sloping. In Section 6, we will compute explicit approximations for $\sigma_X(\tau,\lambda)$ and $\sigma_Z^{(1/\beta)}(\tau,\lambda)$ for three models: limited CEV, Heston and limited SABR. As we shall see, although these three models induce distinct implied volatility surfaces, for small τ the role of β in relating σ_X to σ_Z will be captured by (5.14).

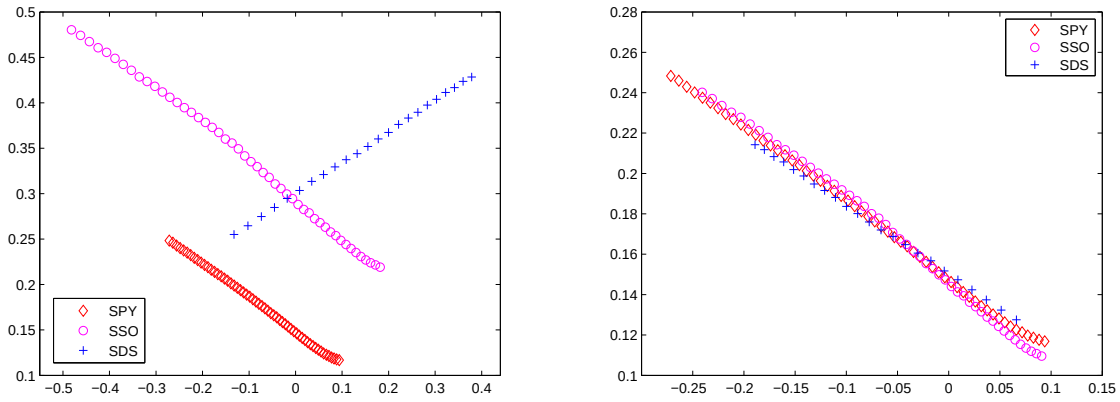

Figure 1: Left: Empirical implied volatilities $\sigma_Z(\tau,\lambda)$ plotted as a function of log-moneyness λ for SPY (red diamonds, $\beta=+1$), SSO (purple circles, $\beta=+2$), and SDS (blue crosses, $\beta=-2$) on August 15, 2013 with $\tau=155$ days to maturity. Note that the implied volatility of SDS is increasing in the LETF log-moneyness. Right: Using the same data, the scaled LETF implied volatilities $\sigma_Z^{(1/\beta)}(\tau,\lambda)$ nearly coincide.

We emphasize, however, that the scaling alone is not sufficient to capture the complexity of the LETF implied volatility surface. Indeed, as τ increases, we expect $\sigma_Z^{(1/\beta)}$ to diverge from σ_X. This discrepancy is due to the integrated variance contribution to the terminal value of Z, as can be seen from (2.3). Thus, for longer maturities, an accurate approximation of the LETF implied volatility surface must include higher terms in τ. From the general implied volatility expression, we can see that the role of β in the $\mathcal{O}(\tau)$ terms is complicated and does not lend itself to a simple scaling argument. For this reason the full implied volatility expansion – not just the scaling argument – is important.

Remark 5.4. Zhang (2010) and Leung and Sircar (2015) postulate an alternative implied volatility scaling based on stochastic arguments. Given the terminal ETF value $X_\tau=k$, they compute the expected future log-moneyness $Z_\tau - z$

$$\mathbb{E}_{x,y,z}[Z_T - z | X_T = k] = \beta(k-x) - \frac{1}{2}\beta(\beta-1)\int_0^\tau \mathbb{E}_{x,y,z}[\sigma^2(s,X_s,Y_s)|X_\tau=k]\,\mathrm{d}s, \tag{5.15}$$

where $\mathbb{E}_{x,y,z}[\cdot]=\mathbb{E}[\cdot|X_0=x,Y_0=y,Z_0=z]$. Zhang (2010) calls the LETF option strike corresponding to (5.15) the "most likely" strike. They also note, from the ETF and LETF SDEs that the volatility of Z is $|\beta|$ times the volatility of X. Using the above as heuristic, the authors propose to scale implied volatilities as follows

$$\sigma_Z(\tau,\lambda) = |\beta|\sigma_X(\tau,\beta\lambda - \frac{1}{2}\beta(\beta-1)I(\tau)),$$
$$I(\tau) = \int_0^\tau \mathbb{E}_{x,y,z}[\sigma^2(s,X_s,Y_s)|X_\tau=k]\,\mathrm{d}s.$$

In Zhang (2010) and Leung and Sircar (2015), the value of $I(\tau)$ is estimated using an average from observed implied volatility. In contrast, the scaling proposed in (5.13) does not attempt to account for the integral in

Author Manuscript

(5.15). Nevertheless, the effect of the integrated variance is captured by $\mathcal{O}(\tau)$ terms in the general implied volatility expansion.

6 Examples

In this Section, we provide explicit expressions for implied volatilities under three different model dynamics: limited CEV, Heston and limited SABR. Special attention will be paid to the role of β, the leverage ratio. In the examples below, we fix $(\bar{x}(\cdot), \bar{y}(\cdot)) = (X_0, Y_0)$ and we evaluate implied volatilities at time $t = 0$ and maturing at time $T = \tau$.

We note that, although Theorem 4.8 establishes the order of accuracy of our pricing approximation as $\tau \to 0$, our numerical tests indicate that the implied volatility expansion gives an accurate approximation of $\sigma_Z^{(1/\beta)}$ for maturities of multiple years. Nevertheless, options on LETFs only currently trade with maturities of less than 1.25 years[3]. Leung and Sircar (2015) have plotted the empirical implied volatilities for four S&P500 based LETF options ($\beta = \pm 2, \pm 3$), all with maturities of less than a year. Hence, in the numerical examples below, we focus on these maturities.

6.1 Limited CEV

In the Constant Elasticity of Variance (CEV) local volatility model of Cox (1975), the dynamics of the underlying S are given by

$$\mathrm{d}S_t = \delta S_t^{\gamma-1} S_t \mathrm{d}W_t^x, \qquad S_0 > 0,$$

where, to ensure the martingale property of the process S, the parameter γ is assumed to satisfy $\gamma \leq 1$ (Heston et al. (2007)).

Note that when $\gamma < 1$, the process S can be absorbed at zero and yet we have assumed that S remains strictly positive. Thus, rather the consider the CEV model, we consider instead the *limited* CEV model (LCEV), which was introduced in Andersen and Andreasen (2000). In the LCEV model, the ETF, denoted here by S^ε, has regularized dynamics

$$\mathrm{d}S_t^\varepsilon = \delta((S_t^\varepsilon)^{\gamma-1} \wedge \varepsilon^{\gamma-1}) S_t^\varepsilon \mathrm{d}W_t^x, \qquad S_0^\varepsilon \gg \varepsilon > 0,$$

where $\varepsilon > 0$ is a fixed, arbitrarily small parameter. Because the volatility process in the LCEV model is bounded by $\delta\varepsilon^{\gamma-1}$ the process S^ε remains strictly positive.

The dynamics of $(X, Z) = (\log S^\varepsilon, \log L)$ in the LCEV setting are

$$\begin{aligned}
\mathrm{d}X_t &= -\frac{1}{2}\delta^2 \mathrm{e}^{2(\gamma-1)(X_t \vee \log\varepsilon)}\mathrm{d}t + \delta\, \mathrm{e}^{(\gamma-1)(X_t\vee\log\varepsilon)}\mathrm{d}W_t^x, & X_0 = x := \log S_0^\varepsilon \gg \log\varepsilon.\\
\mathrm{d}Z_t &= -\frac{1}{2}\beta^2\delta^2 \mathrm{e}^{2(\gamma-1)(X_t \vee \log\varepsilon)}\mathrm{d}t + \beta\delta\, \mathrm{e}^{(\gamma-1)(X_t\vee\log\varepsilon)}\mathrm{d}W_t^x, & Z_0 = z := \log L_0.
\end{aligned}$$

The generator of (X, Z) is given by

$$\mathcal{A} = \frac{1}{2}\delta^2 \mathrm{e}^{2(\gamma-1)(x\vee\log\varepsilon)}\left((\partial_x^2 - \partial_x) + \beta^2(\partial_z^2 - \partial_z) + 2\beta\partial_x\partial_z\right).$$

[3]Delayed quotes for ETF and LETF options of all traded maturities are available on the CBOE and Yahoo Finance websites.

Author Manuscript

Thus, from (3.3), we identify

$$a(x,y) = \frac{1}{2}\delta^2 \mathrm{e}^{2(\gamma-1)(x \vee \log \varepsilon)}, \qquad b(x,y) = 0, \qquad c(x,y) = 0, \qquad f(x,y) = 0. \tag{6.1}$$

As we assume without loss of generality that $x > \log \varepsilon$, the implied volatility approximation is independent of ε. Using equations (5.6), (5.7) and (5.10), we compute

$$\begin{aligned}
\sigma_0 &= |\beta| \sqrt{\mathrm{e}^{2x(\gamma-1)}\delta^2}, \\
\sigma_1 &= \tau \left(\frac{(\beta-1)(\gamma-1)\sigma_0^3}{4\beta^2} \right) + \left(\frac{(\gamma-1)\sigma_0}{2\beta} \right)(k-z), \\
\sigma_2 &= \tau \left(\frac{(\gamma-1)^2\sigma_0^3}{24\beta^2} \right) + \tau^2 \left(\frac{(2\beta(6\beta-13)+13)(\gamma-1)^2\sigma_0^5}{96\beta^4} \right) \\
&\quad + \tau \left(\frac{7(\beta-1)(\gamma-1)^2\sigma_0^3}{24\beta^3} \right)(k-z) + \left(\frac{(\gamma-1)^2\sigma_0}{12\beta^2} \right)(k-z)^2, \\
\sigma_3 &= \tau^2 \left(\frac{5(\beta-1)(\gamma-1)^3\sigma_0^5}{32\beta^4} \right) + \tau^3 \left(\frac{(\beta-1)\left(26\beta^2-70\beta+35\right)(\gamma-1)^3\sigma_0^7}{384\beta^6} \right) \\
&\quad + \tau \left(\frac{(\gamma-1)^3\sigma_0^3}{16\beta^3} \right)(k-z) + \tau^2 \left(\frac{5(2\beta(4\beta-9)+9)(\gamma-1)^3\sigma_0^5}{192\beta^5} \right)(k-z) \\
&\quad + \tau \left(\frac{7(\beta-1)(\gamma-1)^3\sigma_0^3}{48\beta^4} \right)(k-z)^2.
\end{aligned}$$

We observe that the factor $(\gamma - 1)$ appears in every term of these expressions. In particular, when $\gamma = 1$, $\sigma_0 = |\beta|\delta$ and $\sigma_1 = \sigma_2 = \sigma_3 = 0$. The higher order terms also vanish as $a(x,y) = \frac{1}{2}\delta^2$ in this case (see (6.1)). Hence, just as in the Black-Scholes case, the implied volatility expansion becomes flat, as expected.

In Figure 2 we plot our third-order approximation of the scaled implied volatility $\sigma_Z^{(1/\beta)}(\tau,\lambda)$ in the LCEV model with leverages $\beta = \{+2,-2\}$ and with maturities $\tau = \{0.25, 0.5, 1\}$ years. For comparison, we also plot the exact scaled implied volatility $\sigma_Z^{(1/\beta)}(\tau,\lambda)$ and the exact implied volatility of the ETF $\sigma_X(\tau,\lambda)$, which are computed by Monte Carlo simulation.

Remark 6.1. Note that the coefficient a in (6.1) has a kink at $x = \log \varepsilon$. Recall that the Nth order price approximation result in Theorem 4.8 requires (from Assumption 4.2) that the coefficients of $\mathcal{A}$ be C^{N+1}. To be consistent with Assumption 4.2, we can consider an alternative modification of the CEV model that bounds the volatility coefficient in such a way that it is C^{N+1}. For example, such a modification can be obtained by connecting two disjoint C^{N+1} functions with a spline of order $N+2$. However, we do not pursue this here as the aim of this section is simply to provide numerical illustrations of the implied volatility scaling.

6.2 Heston

In the Heston model, due to Heston (1993), the dynamics of the underlying S are given by

$$\begin{aligned}
\mathrm{d}S_t &= \sqrt{V_t} S_t \mathrm{d}W_t^x, & S_0 &> 0, \\
\mathrm{d}V_t &= \kappa(\theta - V_t)\mathrm{d}t + \delta\sqrt{V_t}\mathrm{d}W_t^y, & V_0 &> 0, \\
\mathrm{d}\langle W^x, W^y \rangle_t &= \rho \, \mathrm{d}t.
\end{aligned}$$

Author Manuscript

Author Manuscript

Assumption 6.2. In the numerical example below, we impose the Feller condition $2\kappa\theta > \delta^2$ so that the volatility process V remains strictly positive and the log process $Y = \log V$ is well-defined.

Remark 6.3. We emphasize that imposing the Feller condition is *not* required for the accuracy results of Section 4 nor for the implied volatility expansion of Section 5. When the Feller condition is not satisfied, the V process may hit zero. In this case, one should carry out the approximate option price and implied volatility computations using $(X, Y, Z) := (\log S, V, \log L)$ rather than $(X, Y, Z) := (\log S, \log V, \log L)$.

In log notation $(X, Y, Z) := (\log S, \log V, \log L)$ we have the following dynamics

$$\begin{aligned}
\mathrm{d}X_t &= -\frac{1}{2}\mathrm{e}^{Y_t}\mathrm{d}t + \mathrm{e}^{\frac{1}{2}Y_t}\mathrm{d}W_t^x, && X_0 = x := \log S_0,\\
\mathrm{d}Y_t &= \left((\kappa\theta - \tfrac{1}{2}\delta^2)\mathrm{e}^{-Y_t} - \kappa\right)\mathrm{d}t + \delta\,\mathrm{e}^{-\frac{1}{2}Y_t}\mathrm{d}W_t^y, && Y_0 = y := \log V_0,\\
\mathrm{d}Z_t &= -\beta^2\frac{1}{2}\mathrm{e}^{Y_t}\mathrm{d}t + \beta\mathrm{e}^{\frac{1}{2}Y_t}\mathrm{d}W_t^x, && Z_0 = z := \log L_0,\\
\mathrm{d}\langle W^x, W^y\rangle_t &= \rho\,\mathrm{d}t.
\end{aligned}$$

The generator of (X, Y, Z) is given by

$$\begin{aligned}
\mathcal{A} = \frac{1}{2}\mathrm{e}^y&\left((\partial_x^2 - \partial_x) + \beta^2(\partial_z^2 - \partial_z) + 2\beta\partial_x\partial_z\right)\\
&+ \left((\kappa\theta - \tfrac{1}{2}\delta^2)\mathrm{e}^{-y} - \kappa\right)\partial_y + \frac{1}{2}\delta^2\mathrm{e}^{-y}\partial_y^2 + \rho\,\delta\left(\partial_x\partial_y + \beta\partial_x\partial_z\right).
\end{aligned}$$

Thus, from (3.3), we identify

$$a(x,y) = \frac{1}{2}\mathrm{e}^y, \qquad b(x,y) = \frac{1}{2}\delta^2\mathrm{e}^{-y}, \qquad c(x,y) = \left((\kappa\theta - \tfrac{1}{2}\delta^2)\mathrm{e}^{-y} - \kappa\right), \qquad f(x,y) = \rho\,\delta.$$

Using equations (5.6), (5.7) and (5.10) we obtain

$$\begin{aligned}
\sigma_0 &= |\beta|\sqrt{\mathrm{e}^y}, \qquad (6.2)\\
\sigma_1 &= \frac{\tau}{8\sigma_0}\left(\sigma_0^2(\beta\delta\rho - 2\kappa) - \beta^2\left(\delta^2 - 2\theta\kappa\right)\right) + \frac{1}{4\sigma_0}\left(\beta\delta\rho\right)(k - z),\\
\sigma_2 &= \frac{\tau}{96\sigma_0}\left(\beta^2\delta^2\left(\rho^2 + 8\right)\right)\\
&\quad + \frac{\tau^2}{384\sigma_0^3}\left(-3\beta^4\left(\delta^2 - 2\theta\kappa\right)^2 - 2\beta^2\sigma_0^2\left(\delta^2 - 2\theta\kappa\right)(\beta\delta\rho - 2\kappa) + 4\sigma_0^4\left(\beta\delta\left(\beta\delta\left(2\rho^2 - 1\right) - 5\kappa\rho\right) + 5\kappa^2\right)\right)\\
&\quad + \frac{\tau}{96\sigma_0^3}\left(\beta\delta\rho\left(5\beta^2\left(\delta^2 - 2\theta\kappa\right) + \sigma_0^2(2\kappa - \beta\delta\rho)\right)\right)(k - z) + \frac{1}{48\sigma_0^3}\left(\beta^2\delta^2\left(2 - 5\rho^2\right)\right)(k - z)^2,\\
\sigma_3 &= \frac{\tau^2}{768\sigma_0^3}\left(\beta^2\delta^2\left(\beta^2\left(5\rho^2 + 4\right)\left(\delta^2 - 2\theta\kappa\right) + 3\rho^2\sigma_0^2(\beta\delta\rho - 2\kappa)\right)\right)\\
&\quad + \frac{\tau^3}{3072\sigma_0^5}\left(-3\beta^6\left(\delta^2 - 2\theta\kappa\right)^3 + \beta^4\sigma_0^2\left(\delta^2 - 2\theta\kappa\right)^2(\beta\delta\rho - 2\kappa) + 4\beta^2\kappa\sigma_0^4\left(\delta^2 - 2\theta\kappa\right)(\beta\delta\rho - \kappa)\right)\\
&\quad + \frac{\tau^3}{3072\sigma_0^5}\left(2\sigma_0^6(\beta\delta\rho - 2\kappa)\left(\beta\delta\left(\beta\delta\left(5\rho^2 - 6\right) - 6\kappa\rho\right) + 6\kappa^2\right)\right)\\
&\quad + \frac{\tau}{384\sigma_0^3}\left(-\beta^3\delta^3\rho\left(9\rho^2 + 8\right)\right)(k - z)
\end{aligned}$$

$$+\frac{\tau^2}{1536\sigma_0^5}\left(\beta\delta\rho\left(21\beta^4\left(\delta^2-2\theta\kappa\right)^2-10\beta^2\sigma_0^2\left(\delta^2-2\theta\kappa\right)\left(\beta\delta\rho-2\kappa\right)\right)\right)(k-z)$$
$$+\frac{\tau^2}{1536\sigma_0^5}\left(\beta\delta\rho\left(4\sigma_0^4\left(\beta\delta\left(\beta\left(\delta-2\delta\rho^2\right)+3\kappa\rho\right)-3\kappa^2\right)\right)\right)(k-z)$$
$$+\frac{\tau}{384\sigma_0^5}\left(-\beta^2\delta^2\left(\beta^2\left(23\rho^2-8\right)\left(\delta^2-2\theta\kappa\right)+\left(7\rho^2-2\right)\sigma_0^2(2\kappa-\beta\delta\rho)\right)\right)(k-z)^2$$
$$+\frac{1}{96\sigma_0^5}\left(\beta^3\delta^3\rho\left(8\rho^2-5\right)\right)(k-z)^3. \tag{6.3}$$

For longer maturities, the accuracy of the implied volatility expansion can be improved by setting $Y = V$ rather than $Y = \log V$ and choosing a time-dependent expansion point for the Y process: $\bar{y}(t) = \mathbb{E}[V_t|V_0 = \mathrm{e}^y]$. In this case, the formulas for σ_0, σ_1, σ_2 and σ_3 remain explicit. However, as the expressions are quite long, we omit them.

Ahn et al. (2012) noticed from the SDEs that when X has Heston dynamics with parameters (κ, θ, δ, ρ, y), then Z has Heston dynamics with parameters

$$(\kappa_Z, \theta_Z, \delta_Z, \rho_Z, y_Z) = (\kappa, \beta^2\theta, |\beta|\delta, \mathrm{sign}(\beta)\rho, y + \log\beta^2). \tag{6.4}$$

The characteristic function of X_τ is computed explicitly in Heston (1993) and Bakshi, Cao, and Chen (1997)

$$\begin{aligned}
\eta_X(\tau,x,y,\xi) &:= \log\mathbb{E}[\mathrm{e}^{\mathtt{i}\xi X_\tau}|X_0=x,Y_0=y] = \mathtt{i}\xi x + C(\tau,\xi) + D(\tau,\xi)\mathrm{e}^y,\\
C(\tau,\xi) &= \frac{\kappa\theta}{\delta^2}\left((\kappa-\rho\delta\mathtt{i}\xi+d(\xi))\tau - 2\log\left[\frac{1-f(\xi)\mathrm{e}^{d(\xi)\tau}}{1-f(\xi)}\right]\right),\\
D(\tau,\xi) &= \frac{\kappa-\rho\delta\mathtt{i}\xi+d(\xi)}{\delta^2}\frac{1-\mathrm{e}^{d(\xi)\tau}}{1-f(\xi)\mathrm{e}^{d(\xi)\tau}},\\
f(\xi) &= \frac{\kappa-\rho\delta\mathtt{i}\xi+d(\xi)}{\kappa-\rho\delta\mathtt{i}\xi-d(\xi)},\\
d(\xi) &= \sqrt{\delta^2(\xi^2+\mathtt{i}\xi)+(\kappa-\rho\mathtt{i}\xi\delta)^2}.
\end{aligned}$$

As Z also has Heston dynamics, the characteristic function of Z_τ follows directly

$$\eta_Z(\tau,z,y,\xi) := \log\mathbb{E}[\mathrm{e}^{\mathtt{i}\xi Z_\tau}|Y_0=y,Z_0=z] = \eta_X(\tau,z,y,\xi) \quad \text{with} \quad (\kappa,\theta,\delta,\rho,y)\to(\kappa_Z,\theta_Z,\delta_Z,\rho_Z,y_Z).$$

The price of a European call option with payoff $\varphi(z) = (\mathrm{e}^z - \mathrm{e}^k)^+$ can then be computed using standard Fourier methods

$$u^{\mathrm{Hes}}(\tau,z,y) = \frac{1}{2\pi}\int_{\mathbb{R}}\mathrm{d}\xi_r\,\mathrm{e}^{\eta_Z(\tau,z,y,\xi)}\widehat{\varphi}(\xi), \qquad \widehat{\varphi}(\xi) = \frac{-\mathrm{e}^{k-\mathtt{i}k\xi}}{\mathtt{i}\xi+\xi^2}, \qquad \xi=\xi_r+\mathtt{i}\xi_i, \qquad \xi_i<-1. \tag{6.5}$$

Note, as the call option payoff $\varphi(z) = (\mathrm{e}^z - \mathrm{e}^k)^+$ is not in $L^1(\mathbb{R})$, its Fourier transform $\widehat{h}(\xi)$ must be computed in a generalized sense by fixing an imaginary component of the Fourier variable $\xi_i < -1$. Using (6.5) the exact implied volatility σ can be computed by solving (5.2) numerically.

Moreover, it is worth noting that relationship (6.4) can be inferred from our implied volatility expressions. Indeed, the dependence on β in expansions (6.2)-(6.3) is present *only* in the terms $\beta^2\theta, |\beta|\delta, \mathrm{sign}(\beta)\rho, y+\log\beta^2$ of the coefficients. For instance, we can write the zeroth-order term $\sigma_0 = \sqrt{e^{y+\log\beta^2}} = \sqrt{e^{y_Z}}$, and the

Author Manuscript

coefficient of $(k-z)$ in σ_1 is $\beta\delta\rho/4\sigma_0 = |\beta|\delta\text{sign}(\beta)\rho/4\sigma_0 = \delta_Z\rho_Z/4\sigma_0$, as per the notations in (6.4). Similar verification procedures for other terms confirm relationship (6.4).

In Figure 3 we plot our third-order approximation of the scaled implied volatility $\sigma_Z^{(1/\beta)}(\tau,\lambda)$ in the Heston model with leverages $\beta = \{+2,-2\}$ and with maturities $\tau = \{0.25, 0.5, 1\}$ years. For the longest maturity $\tau = 1$, we use the implied volatility expansion obtained by setting setting $Y = V$ and choosing a time-dependent expansion point for the Y process: $\bar{y}(t) = \mathbb{E}[V_t|V_0 = \mathrm{e}^y]$. For comparison, we also plot the exact scaled implied volatility $\sigma_Z^{(1/\beta)}(\tau,\lambda)$ and the exact implied volatility of the ETF $\sigma_X(\tau,\lambda)$. The exact scaled implied volatility $\sigma_Z^{(1/\beta)}$ of the LETF is computed by obtaining call prices from (6.5) and then by inverting the Black-Scholes formula numerically. The exact implied volatility $\sigma_X(\tau,\lambda)$ of the ETF is computed in the same manner.

6.3 Limited SABR

The SABR model of Hagan, Kumar, Lesniewski, and Woodward (2002) is a local-stochastic volatility model in which the risk-neutral dynamics of S are given by

$$\begin{aligned} \mathrm{d}S_t &= V_t S_t^{\gamma-1} S_t \mathrm{d}W_t^x, & S_0 &> 0, \\ \mathrm{d}V_t &= \delta V_t \mathrm{d}W_t^y, & V_0 &> 0, \\ \mathrm{d}\langle W^x, W^z\rangle_t &= \rho\,\mathrm{d}t, \end{aligned}$$

where, to ensure the martingale property of S, the parameter γ is assumed to satisfy $\gamma < 1$ (Lions and Musiela (2007)).

As in the CEV model, the process S in the SABR model may be absorbed at zero, yet our assumption is that $S > 0$. Thus, we consider instead the *limited* SABR model (LSABR). In the LSABR model, the underlying ETF, denoted S^ε, has the following regularized dynamics

$$\begin{aligned} \mathrm{d}S_t^\varepsilon &= V_t \left((S_t^\varepsilon)^{\gamma-1} \wedge \varepsilon^{\gamma-1}\right) S_t^\varepsilon \mathrm{d}W_t^x, & S_0^\varepsilon &\gg \varepsilon > 0, \\ \mathrm{d}V_t &= \delta V_t \mathrm{d}W_t^y, & V_0 &> 0, \\ \mathrm{d}\langle W^x, W^z\rangle_t &= \rho\,\mathrm{d}t, \end{aligned}$$

where $\varepsilon > 0$ is a fixed, arbitrarily small constant. As, for a given V_t, the volatility of S^ε is bounded by $V_t\varepsilon^{\gamma-1}$, the process S^ε remains strictly positive.

In log notation $(X,Y,Z) := (\log S^\varepsilon, \log V, \log L)$ we have, we have the following dynamics:

$$\begin{aligned} \mathrm{d}X_t &= -\frac{1}{2}\mathrm{e}^{2Y_t+2(\gamma-1)(X_t\vee\log\varepsilon)}\mathrm{d}t + \mathrm{e}^{Y_t+(\gamma-1)(X_t\vee\log\varepsilon)}\mathrm{d}W_t^x, & X_0 &= x := \log S_0 \gg \log\varepsilon, \\ \mathrm{d}Y_t &= -\frac{1}{2}\delta^2\mathrm{d}t + \delta\,\mathrm{d}W_t^y, & Y_0 &= y := \log V_0, \\ \mathrm{d}Z_t &= -\frac{1}{2}\beta^2\mathrm{e}^{2Y_t+2(\gamma-1)(X_t\vee\log\varepsilon)}\mathrm{d}t + \beta\mathrm{e}^{Y_t+(\gamma-1)(X_t\vee\log\varepsilon)}\mathrm{d}W_t^x, & Z_0 &= z := \log L_0, \\ \mathrm{d}\langle W^x, W^y\rangle_t &= \rho\,\mathrm{d}t. \end{aligned}$$

The generator of (X,Y,Z) is given by

$$\mathcal{A} = \frac{1}{2}\mathrm{e}^{2y+2(\gamma-1)(x\vee\log\varepsilon)}\left((\partial_x^2-\partial_x) + \beta^2(\partial_z^2-\partial_z) + 2\beta\partial_x\partial_y\right)$$

Author Manuscript

$$-\frac{1}{2}\delta^2\partial_y+\frac{1}{2}\delta^2\partial_y^2+\rho\,\delta\,\mathrm{e}^{y+(\gamma-1)(x\vee\log\varepsilon)}(\partial_x\partial_y+\beta\partial_y\partial_z).$$

Thus, using (3.3), we identify

$$a(x,y)=\tfrac{1}{2}\mathrm{e}^{2y+2(\gamma-1)(x\vee\log\varepsilon)},\qquad b(x,y)=\tfrac{1}{2}\delta^2,\qquad c(x,y)=-\tfrac{1}{2}\delta^2,\qquad f(x,y)=\rho\delta\mathrm{e}^{y+(\gamma-1)(x\vee\log\varepsilon)}.$$

As we assume without loss of generality that $x>\log\varepsilon$, the implied volatility approximation is independent of ε. Using equations (5.6), (5.7) and (5.10) we compute

$$\sigma_0=|\beta|\sqrt{\mathrm{e}^{2y+2x(-1+\gamma)}},\qquad\qquad \sigma_1=\sigma_{1,0}+\sigma_{0,1},\qquad\qquad \sigma_2=\sigma_{2,0}+\sigma_{1,1}+\sigma_{0,2},$$

where

$$\begin{aligned}
\sigma_{1,0}&=\tau\left(\frac{(\beta-1)(\gamma-1)\sigma_0^3}{4\beta^2}\right)+\left(\frac{(\gamma-1)\sigma_0}{2\beta}\right)(k-z),\\
\sigma_{0,1}&=\tau\left(-\frac{1}{4}\delta\sigma_0\left(\delta-\rho\sigma_0\mathrm{sgn}(\beta)\right)\right)+\left(\frac{1}{2}\delta\rho\mathrm{sgn}(\beta)\right)(k-z),\\
\sigma_{2,0}&=\tau\left(\frac{(\gamma-1)^2\sigma_0^3}{24\beta^2}\right)+\tau^2\left(\frac{(2\beta(6\beta-13)+13)(\gamma-1)^2\sigma_0^5}{96\beta^4}\right)\\
&\quad+\tau\left(\frac{7(\beta-1)(\gamma-1)^2\sigma_0^3}{24\beta^3}\right)(k-z)+\left(\frac{(\gamma-1)^2\sigma_0}{12\beta^2}\right)(k-z)^2,\\
\sigma_{1,1}&=\tau\left(\frac{(\gamma-1)\delta\rho\sigma_0^2}{12\,|\beta|}\right)+\tau^2\left(\frac{(\gamma-1)\delta\sigma_0^3\left(\beta(6\beta-7)\rho\sigma_0-5(\beta-1)\delta\,|\beta|\right)}{48\,|\beta|^3}\right)\\
&\quad+\tau\left(\frac{(\gamma-1)\delta\sigma_0\left(\delta\,|\beta|+(2\beta-1)\rho\sigma_0\right)}{24\beta\,|\beta|}\right)(k-z)+\left(-\frac{(\gamma-1)\delta\rho}{3\,|\beta|}\right)(k-z)^2,\\
\sigma_{0,2}&=\tau\left(\frac{1}{24}\delta^2\left(8-3\rho^2\right)\sigma_0\right)+\tau^2\left(\frac{1}{96}\delta^2\sigma_0\left(5\delta^2+4\left(3\rho^2-1\right)\sigma_0^2-\frac{14\delta\rho\sigma_0}{\mathrm{sgn}(\beta)}\right)\right)\\
&\quad+\tau\left(-\frac{\delta^2\rho\left(\delta-3\rho\sigma_0\mathrm{sgn}(\beta)\right)}{24\mathrm{sgn}(\beta)}\right)(k-z)+\left(\frac{\delta^2\left(2-3\rho^2\right)}{12\sigma_0}\right)(k-z)^2.
\end{aligned}$$

We omit the expression for σ_3 for the sake of brevity. However, an explicit computations shows that σ_3 contains terms of orders τ^2, τ^3, $\tau(k-z)$, $\tau^2(k-z)$, $\tau(k-z)^2$ and $(k-z)^3$. In Figure 4 we plot our third-order approximation of the scaled implied volatility $\sigma_Z^{(1/\beta)}(\tau,\lambda)$ in the LSABR model with leverages $\beta=\{+2,-2\}$ and with maturities $\tau=\{0.25,0.5,1\}$ years. For comparison, we also plot the exact scaled implied volatility $\sigma_Z^{(1/\beta)}(\tau,\lambda)$ and the exact implied volatility of the ETF $\sigma_X(\tau,\lambda)$, which are computed by Monte Carlo simulation. As with the LCEV model, the coefficients in the LSABR model have a kink at $x=\log\varepsilon$. The kink can be smoothed as described in Remark 6.1.

7 Conclusion

In this article, starting from ETF dynamics in a general time-inhomogeneous LSV setting, we derive approximate European-style option prices written on the associated LETFs. The option price approximation requires only a normal CDF to compute. Therefore, computational times for prices are comparable to Black-Scholes. We also establish rigorous error bounds for our pricing approximation. These error bounds

Author Manuscript

are established through a regularization procedure, which allows us to overcome challenges that arise when dealing with a generator $\mathcal{A}(t)$ that is *not* elliptic.

Additionally, we derive an implied volatility expansion that is fully explicit – polynomial in log-moneyness $\lambda = (k - z)$ and (for time-homogeneous models) polynomial in time to maturity. To aid in the analysis of the implied volatility surface, we discuss some natural scalings of implied volatility. Furthermore, we test our implied volatility expansion on three LSV models (limited CEV, Heston and limited SABR) and find that the expansion provides an excellent approximation of the true implied volatility.

The markets for leveraged ETFs and their options continue to grow, not only in equities, but also in other sectors such as commodity, fixed-income, and currency. The question of consistent pricing, as we have investigated for equity LETF options in terms implied volatility, is also relevant to LETF options in other sectors. Naturally, the valuation of LETF options will depend on the dynamics of the LETFs and underlying price process, which may vary significantly across sectors (see e.g. Guo and Leung (2015); Leung and Ward (2015) for commodity LETFs). Nevertheless, it is both practically and mathematically interesting to adapt the techniques in the current paper to investigate the implied volatilities across leverage ratios with different underlyings. From a market stability perspective, it is important for both investors and regulators to understand the risks and dependence structure among ETFs and the price relationships of their traded derivatives.

Acknowledgments The authors are grateful to Peter Carr, Emanuel Derman, Martin Haugh, Sebastian Jaimungal, and Ronnie Sircar for their helpful discussions, and we thank the seminar participants at Morgan Stanley, Columbia University, and Fields Institute for their comments. The authors also wish to express their gratitude to two anonymous referees and one anonymous associate editor, whose comments helped improve the mathematical rigor and clarity of this paper.

References

Ahn, A., M. Haugh, and A. Jain (2012). Consistent pricing of options on leveraged ETFs. *SSRN preprint*.

Andersen, L. and J. Andreasen (2000). Volatility skews and extensions of the LIBOR market model. *Applied Mathematical Finance 7*(1), 1–32.

Armstrong, J., M. Forde, M. Lorig, and H. Zhang (2014). Small-time asymptotics for a general local-stochastic volatility model with a jump-to-default: curvature and the heat kernel expansion. *ArXiv preprint arXiv:1312.2281*.

Avellaneda, M. and S. Zhang (2010). Path-dependence of leveraged ETF returns. *SIAM Journal on Financial Mathematics 1*, 586–603.

Bakshi, G., C. Cao, and Z. Chen (1997, December). Empirical performance of alternative option pricing models. *Journal of Finance 52*(5), 2003–2049.

Benhamou, E., E. Gobet, and M. Miri (2010). Time-dependent Heston model. *SIAM Journal on Financial Mathematics 1*(1), 289–325.

Author Manuscript

Author Manuscript

Bompis, R. and E. Gobet (2013). Asymptotic and non asymptotic approximations for option valuation. In *Recent Developments in Computational Finance. Foundations, Algorithms and Applications*, pp. 159–241. Hackensack, NJ: World Scientific.

Bony, J.-M. (1969). Principe du maximum, inégalite de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés. *Ann. Inst. Fourier (Grenoble) 19*(fasc. 1), 277–304 xii.

Carr, P. and J. Sun (2007). A new approach for option pricing under stochastic volatility. *Review of Derivatives Research 10*(2), 87–150.

Cheng, M. and A. Madhavan (2009). The dynamics of leveraged and inverse exchange traded funds. *Journal Of Investment Management 4*, 43–62.

Cox, J. (1975). Notes on option pricing I: Constant elasticity of diffusions. *Unpublished draft, Stanford University*. A revised version of the paper was published by the Journal of Portfolio Management in 1996.

Di Francesco, M. and A. Pascucci (2005). On a class of degenerate parabolic equations of Kolmogorov type. *AMRX Appl. Math. Res. Express 3*, 77–116.

Forde, M. and A. Jacquier (2011). Small-time asymptotics for an uncorrelated local-stochastic volatility model. *Applied Mathematical Finance 18*(6), 517–535.

Forde, M., A. Jacquier, and R. Lee (2012). The small-time smile and term structure of implied volatility under the Heston model. *SIAM Journal on Financial Mathematics 3*(1), 690–708.

Fouque, J.-P., G. Papanicolaou, R. Sircar, and K. Solna (2011). *Multiscale stochastic volatility for equity, interest rate, and credit derivatives*. Cambridge: Cambridge University Press.

Friedman, A. (1964). *Partial differential equations of parabolic type*. Englewood Cliffs, N.J.: Prentice-Hall Inc.

Gatheral, J., E. P. Hsu, P. Laurence, C. Ouyang, and T.-H. Wang (2012). Asymptotics of implied volatility in local volatility models. *Mathematical Finance 22*(4), 591–620.

Guo, K. and T. Leung (2015). Understanding the tracking errors of commodity leveraged ETFs. In M. Ludkovski, R. Sircar, and R. Aid (Eds.), *Commodities, Energy, and Environmental Finance, Fields Institute Communications*, Volume 74, pp. 39 – 63. Springer.

Hagan, P., D. Kumar, A. Lesniewski, and D. Woodward (2002). Managing smile risk. *Wilmott Magazine 1000*, 84–108.

Hagan, P. and D. Woodward (1999). Equivalent Black volatilities. *Applied Mathematical Finance 6*(3), 147–157.

Henry-Labordère, P. (2009). *Analysis, geometry, and modeling in finance: Advanced methods in option pricing*, Volume 13. Chapman & Hall.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*(2), 327–343.

Heston, S., M. Loewenstein, and G. A. Willard (2007). Options and Bubbles. *Review of Financial Studies 20*(2), 359–390.

Author Manuscript

Ikeda, N. and S. Watanabe (1989). *Stochastic differential equations and diffusion processes* (Second ed.), Volume 24 of *North-Holland Mathematical Library*. Amsterdam: North-Holland Publishing Co.

Jacquier, A. and M. Lorig (2013). The smile of certain Lévy-type models. *SIAM Journal on Financial Mathematics 4*(1), 804–830.

Lanconelli, E. and A. Pascucci (1999). On the fundamental solution for hypoelliptic second order partial differential equations with non-negative characteristic form. *Ricerche Mat. 48*(1), 81–106.

Lee, R. and R. Wang (2015). How leverage shifts and scales a volatility skew: Asymptotics for continuous and jump dynamics.

Leung, T. and R. Sircar (2015). Implied volatility of leveraged ETF options. *Applied Mathematical Finance 22*(2), 162–188.

Leung, T. and B. Ward (2015). The golden target: Analyzing the tracking performance of leveraged gold ETFs. *Studies in Economics and Finance 32*(3), 278 – 297.

Lions, P.-L. and M. Musiela (2007). Correlations and bounds for stochastic volatility models. In *Annales de l'IHP Analyse non linéaire*, Volume 24, pp. 1–16.

Lorig, M., S. Pagliarani, and A. Pascucci (2015a). Analytical expansions for parabolic equations. *SIAM Journal on Applied Mathematics 75*.

Lorig, M., S. Pagliarani, and A. Pascucci (2015b). Explicit implied volatilities for multifactor local-stochastic volatility models. *To appear: Mathematical Finance*.

Lorig, M., S. Pagliarani, and A. Pascucci (2015c). A family of density expansions for Lévy-type processes with default. *Annals of Applied Probability 25*(1), 235–267.

Pagliarani, S. and A. Pascucci (2012). Analytical approximation of the transition density in a local volatility model. *Cent. Eur. J. Math. 10*(1), 250–270.

Pagliarani, S. and A. Pascucci (2014). Asymptotic expansions for degenerate parabolic equations. *C. R. Math. Acad. Sci. Paris 352*(12), 1011–1016.

Pagliarani, S., A. Pascucci, and C. Riga (2013). Adjoint expansions in local Lévy models. *SIAM J. Financial Math. 4*, 265–296.

Pascucci, A. (2011). *PDE and martingale methods in option pricing*, Volume 2 of *Bocconi & Springer Series*. Milan: Springer.

Perelló, J., R. Sircar, and J. Masoliver (2008). Option pricing under stochastic volatility: the exponential ornstein–uhlenbeck model. *Journal of Statistical Mechanics: Theory and Experiment 2008*(06), P06010.

Zhang, J. (2010). *Path dependence properties of leveraged exchange-traded funds: compounding, volatility and option pricing*. Ph. D. thesis, New York University.

Author Manuscript

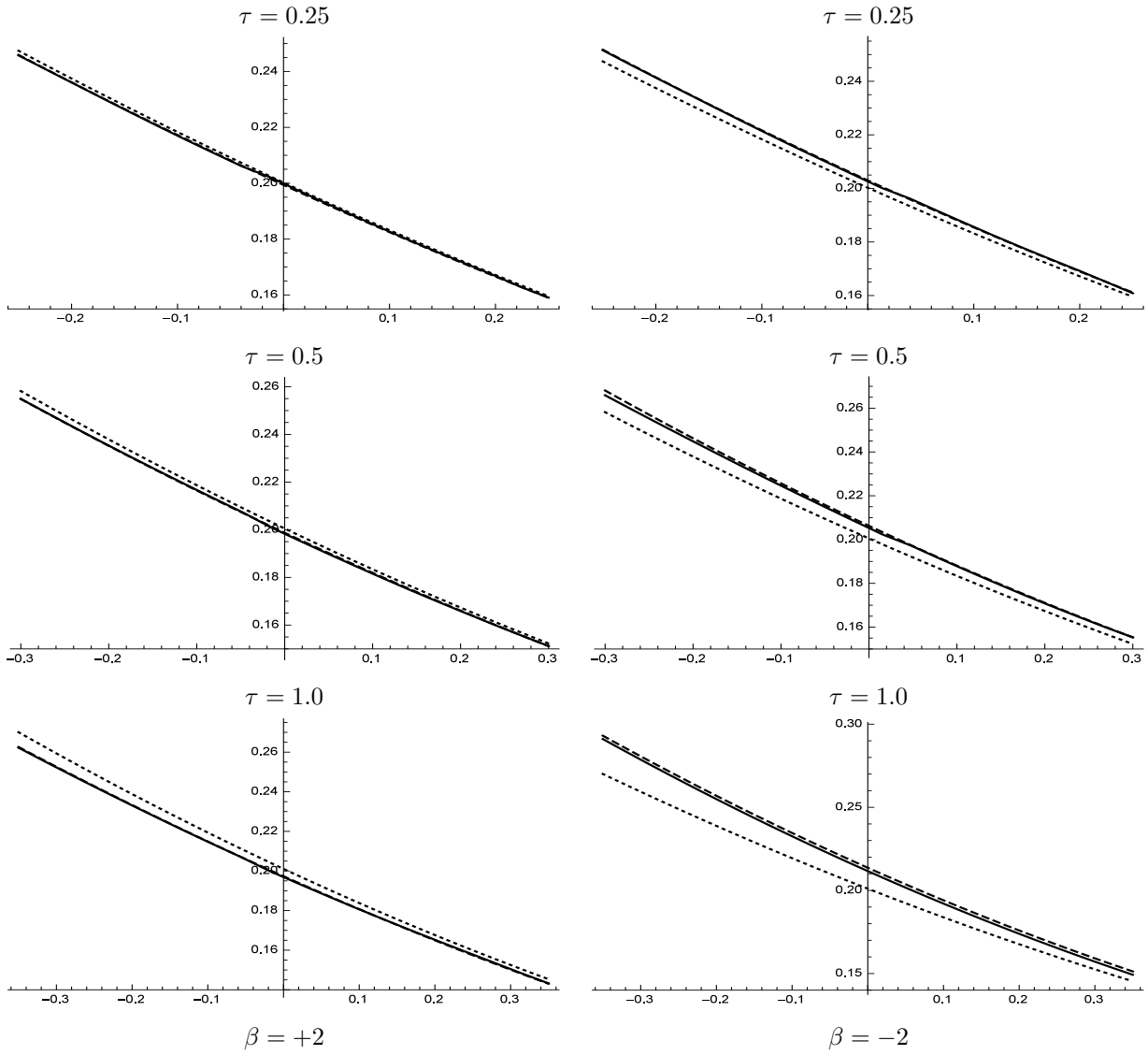

Figure 2: Exact (solid – computed by Monte Carlo) and approximate (dashed) scaled implied volatility $\sigma_Z^{(1/\beta)}(\tau, \lambda)$ under LCEV model dynamics plotted as a function of log-moneyness λ. For comparison, we also plot the exact implied volatility of the LCEV model $\sigma_Z^{(1)}(\tau, \lambda) = \sigma_X(\tau, \lambda)$ (dotted). Parameters: $\delta = 0.2$, $\gamma = -0.75$, $x = 0$. For each leverage ratio ($\beta = \pm 2$), as τ increases, the solid and dotted lines diverge, while the dashed and solid lines remain so close they are nearly indistinguishable.

Author Manuscript

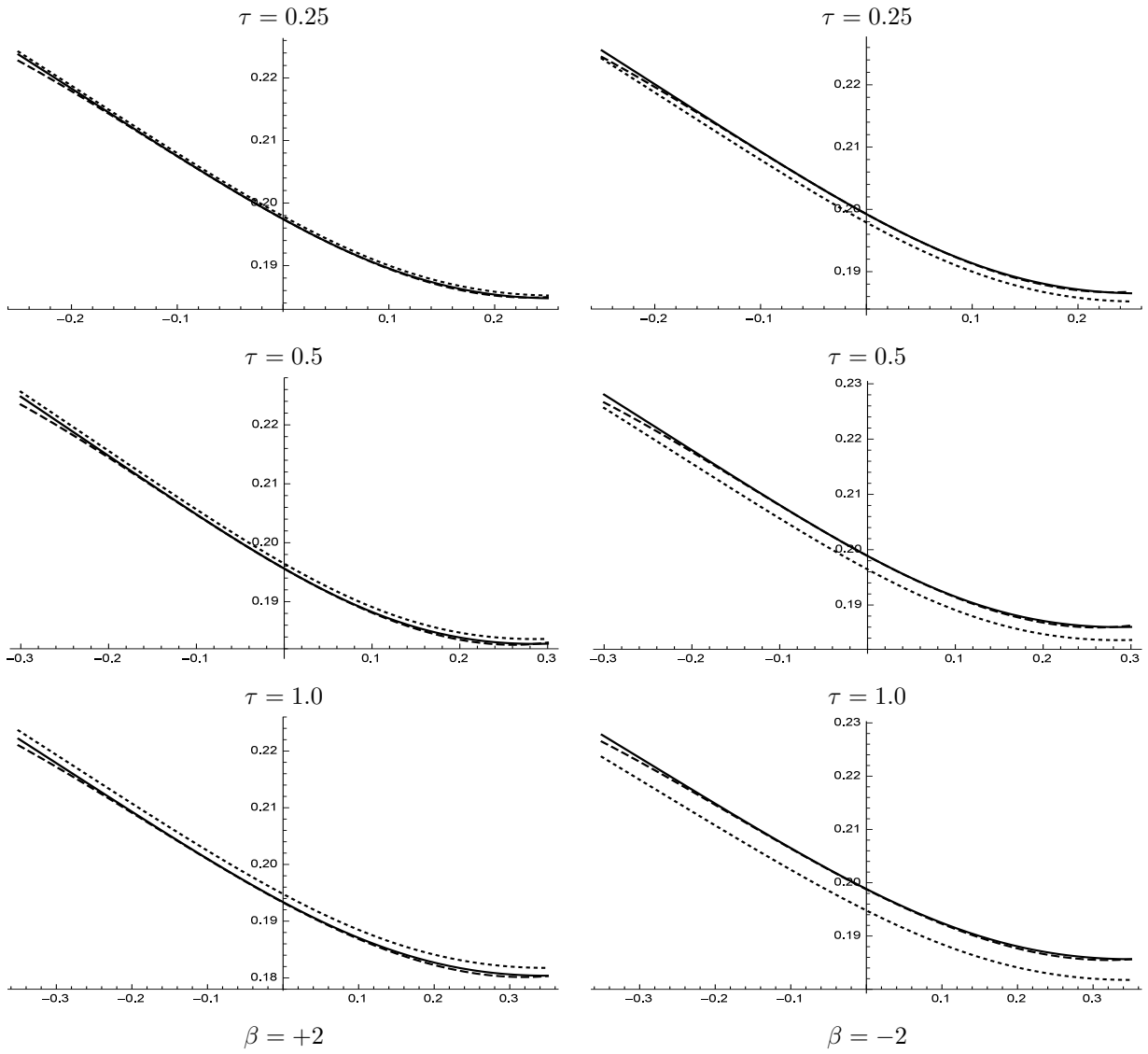

Figure 3: Exact (solid – computed by Fourier inversion) and approximate (dashed) scaled implied volatility $\sigma_Z^{(1/\beta)}(\tau,\lambda)$ under Heston model dynamics plotted as a function of log-moneyness λ. For comparison, we also plot the exact implied volatility of the Heston model $\sigma_Z^{(1)}(\tau,\lambda) = \sigma_X(\tau,\lambda)$ (dotted). Parameters: $\kappa = 1.15$, $\theta = 0.04$, $\delta = 0.2$, $\rho = -0.4$, $y = \log\theta$. For $\beta = \pm 2$, as τ increases, the dotted lines start to deviate from the solid lines, but the dashed and solids lines remain so close they are nearly indistinguishable.

Author Manuscript

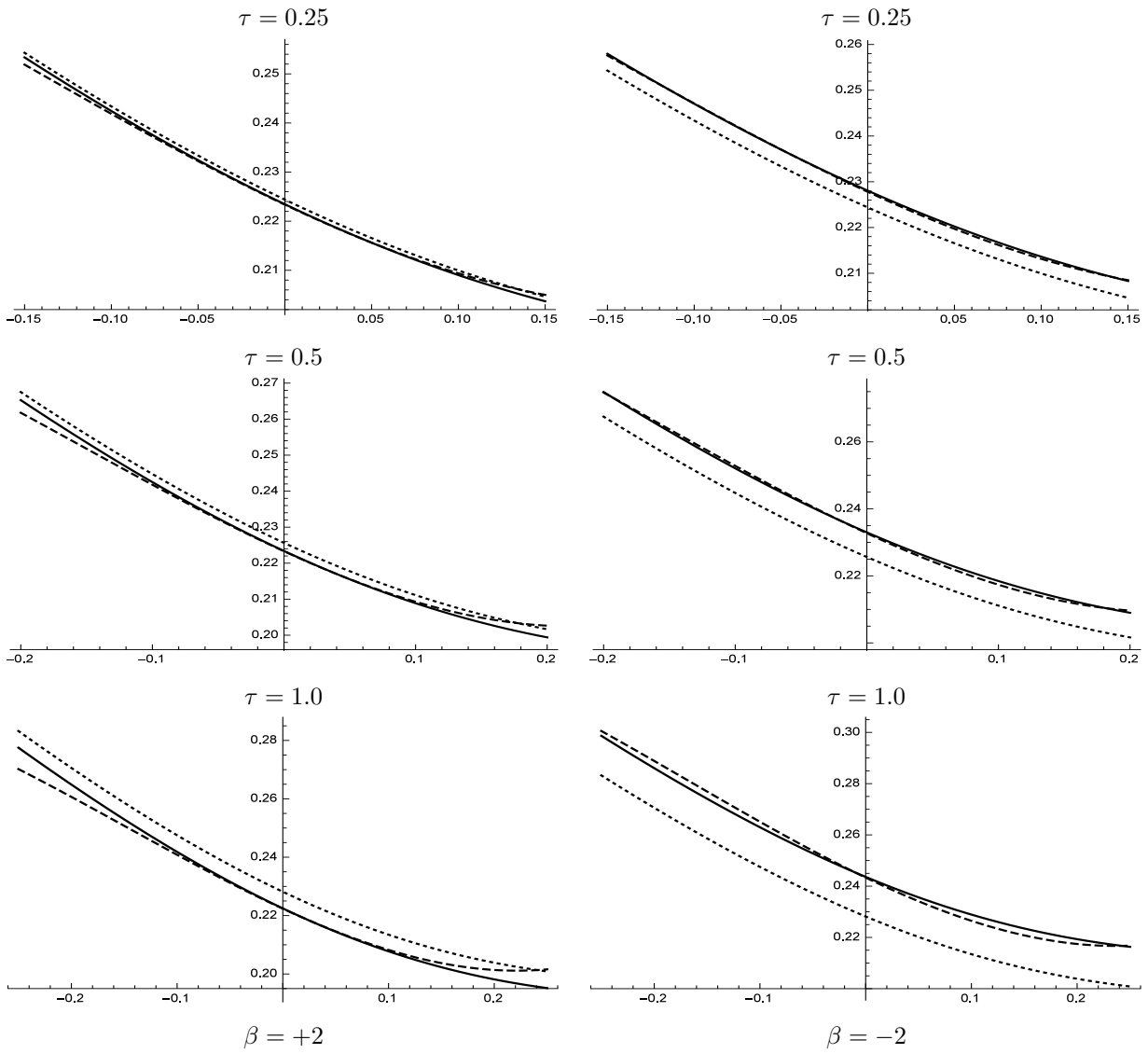

Figure 4: Exact (solid – computed by Monte Carlo) and approximate (dashed) scaled implied volatility $\sigma_Z^{(1/\beta)}(\tau, \lambda)$ under LSABR model dynamics plotted as a function of λ. For comparison, we also plot the exact implied volatility of the LSABR model $\sigma_Z^{(1)}(\tau, \lambda) = \sigma_X(\tau, \lambda)$ (dotted). Parameters: $\delta = 0.5$, $\gamma = -0.5$, $\rho = 0.0$ $x = 0$, $y = -1.5$. As expected, as τ increases, the solid and dotted lines diverge, while the dashed and solid lines remain close.