ASYMPTOTICS OF FORWARD IMPLIED VOLATILITY

ANTOINE JACQUIER AND PATRICK ROOME

Abstract. We prove here a general closed-form expansion formula for forward-start options and the forward implied volatility smile in a large class of models, including Heston and time-changed exponential Lévy models. This expansion applies to both small and large maturities and is based solely on the knowledge of the forward characteristic function of the underlying process. The method is based on sharp large deviations techniques, and allows us to recover (in particular) many results for the spot implied volatility smile. In passing we show (i) that the small-maturity exploding behaviour of forward smiles depends on whether the quadratic variation of the underlying is bounded or not, and (ii) that the forward-start date also has to be rescaled in order to obtain non-trivial small-maturity asymptotics.

1. Introduction

Consider an asset price process $\left(e^{X_t}\right)_{t\geq 0}$ with $X_0 = 0$, paying no dividend, defined on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq0}, \mathbb{P})$ with a given risk-neutral measure $\mathbb{P}$, and assume that interest rates are zero. In the Black-Scholes-Merton (BSM) model, the dynamics of the logarithm of the asset price are given by

$$\mathrm{d}X_t = -\frac{1}{2}\sigma^2 \mathrm{d}t + \sigma \mathrm{d}W_t, \tag{1.1}$$

where $\sigma > 0$ is the instantaneous volatility and W is a standard Brownian motion. The no-arbitrage price of the call option at time zero is then given by the famous BSM formula [12, 44]: $C_{\mathrm{BS}}(\tau, k, \sigma) := \mathbb{E}\left(e^{X_\tau} - e^k\right)_+ = \mathcal{N}(d_+) - e^k \mathcal{N}(d_-)$, with $d_\pm := -\frac{k}{\sigma\sqrt{\tau}} \pm \frac{1}{2}\sigma\sqrt{\tau}$, where $\mathcal{N}$ is the standard normal distribution function. For a given market price $C^{\mathrm{obs}}(\tau, k)$ of the option at strike e^k and maturity τ we define the spot implied volatility $\sigma_\tau(k)$ as the unique solution to the equation $C^{\mathrm{obs}}(\tau, k) = C_{\mathrm{BS}}(\tau, k, \sigma_\tau(k))$.

For any $t, \tau > 0$ and $k \in \mathbb{R}$, we define [10, 43] a Type-I forward-start option with forward-start date t, maturity τ and strike e^k as a European option with payoff $\left(e^{X_{t+\tau}}/e^{X_t} - e^k\right)^+$. In the BSM model (1.1) its value is simply worth $C_{\mathrm{BS}}(\tau, k, \sigma)$. For a given market price $C^{\mathrm{obs}}(t, \tau, k)$ of the option at strike e^k, forward-start date t and maturity τ we define the forward implied volatility smile $\sigma_{t,\tau}(k)$ as the unique solution to $C^{\mathrm{obs}}(t, \tau, k) = C_{\mathrm{BS}}(\tau, k, \sigma_{t,\tau}(k))$ since $\partial_\sigma C_{\mathrm{BS}}(\tau, k, \sigma) > 0$ [10, 34]. A second type of forward-start option exists [43] and corresponds to a European option with payoff $\left(e^{X_{t+\tau}} - e^{k+X_t}\right)^+$. In the BSM model (1.1) the value of the Type-II forward-start option is worth $C_{\mathrm{BS}}(\tau, k, \sigma)$ [47]. Again, for a given market price $C^{\mathrm{obs,II}}(\tau, t, k)$ of such an option, we define the Type-II forward implied volatility smile $\widetilde{\sigma}_{t,\tau}(k)$ as the unique solution to $C^{\mathrm{obs,II}}(\tau, t, k) = C_{\mathrm{BS}}(\tau, k, \widetilde{\sigma}_{t,\tau}(k))$. Both definitions of the forward smile are generalisations of the spot implied volatility smile since they reduce to the spot smile when $t = 0$.

The literature on implied volatility asymptotics is extensive and has been studied using a diverse range of mathematical techniques. In particular, small-maturity asymptotics have historically received wide attention

Date: December 18, 2012.

The authors would like to thank Peter Tankov for useful discussions.

due to earlier results from the eighties on expansions of the heat kernel [7]. PDE methods for continuous-time diffusions [9], large deviations [19, 16], saddlepoint methods [21] and differential geometry [32] are among the main methods used to tackle the small-maturity case. Extreme strike asymptotics arose with the seminal paper by Roger Lee [42] and have been further extended by Benaim and Friz [5, 6] and in [30, 31, 24, 16]. Comparatively, large-maturity asymptotics have only been studied in [51, 20, 36, 35, 22] using large deviations and saddlepoint methods. Fouque et al. [23] have also successfully introduced perturbation techniques in order to study slow and fast mean-reverting stochastic volatility models. Models with jumps (including Lévy processes), studied in the above references for large maturities and extreme strikes, 'explode' in small time, in a precise sense investigated in [1, 2, 50, 46, 45, 18].

A collection of implied volatility smiles over a time horizon $(0, T]$ is also known to be equivalent to the marginal distributions of the asset price process over $(0, T]$. Implied volatility asymptotics has therefore provided a set of tools to analytically understand the marginal distributions of a model and their relationships to market observable quantities such as implied volatility smiles. However many models can calibrate to implied volatility smiles (static information) with the same degree of precision and produce radically different prices and risk sensitivities for exotic securities. This can usually be traced back to a complex and often non-transparent dependence on model transitional probabilities or equivalently on the model generated dynamics of implied volatility smiles. The model dynamics of implied volatility smiles is therefore a key model risk associated with these products and any model used for pricing and risk management should produce realistic dynamics that are in line with trader expectations and historical dynamics. One metric that can be used to understand the dynamics of implied volatility smiles ([10] calls it a 'global measure' of the dynamics of implied volatilities) is to use the forward smile defined above. The forward smile is also a market defined quantity and naturally extends the notion of the spot implied volatility smile. Forward-start options also serve as natural hedging instruments for several exotic securities (such as Cliquets, Ratchets and Napoleons; see [26, Chapter 10]) and it is therefore important for a model to be able to calibrate to liquid forward smiles. Despite the significant research on implied volatility asymptotics, there are virtually no results on the asymptotics of the forward smile: Glasserman and Wu [28] introduced different notions of forward volatilities to assess their predictive values in determining future option prices and future implied volatility, Keller-Ressel [40] studies a very specific type of asymptotic (when the forward-start date becomes large), and empirical results have been carried out by practitioners in [10, 26, 13].

We consider below a continuous-time stochastic process (Z_ε) and prove an expansion of option prices on (Z_ε) as ε tends to zero. Setting $Z_\varepsilon \equiv X_\varepsilon$ or $Z_\varepsilon \equiv \varepsilon X_{1/\varepsilon}$ then yields small or large-maturity expansions of option prices. This main result is presented in Section 2 as well as corollaries applying it to forward-start option asymptotics. We also translate these results into closed-form asymptotic expansions for the forward implied volatility smile (Type I and Type II). In Section 3, we provide explicit examples for the Heston, multi-Heston, Schöbel-Zhu and time-changed exponential Lévy processes. Section 4 provides numerical evidence supporting the practical relevance of these results and we leave the proofs of the main results to Section 5.

Notations: $\mathcal{N}(\mu, \sigma^2)$ shall represent the Gaussian distribution with mean μ and variance σ^2. Furthermore $\mathbb{E}$ shall always denote expectation under a risk-neutral measure $\mathbb{P}$ given a priori. We shall refer to the standard (as opposed to the forward) implied volatility as the spot smile and denote it σ_τ. The (Type-I) forward implied volatility will be denoted $\sigma_{t,\tau}$ as above. In the remaining of this paper ε shall always denote a strictly positive small quantity.

2. General Results

This section gathers the main notations of the paper as well as the general results. The main result is Theorem 2.6, which provides an asymptotic expansion—up to virtually any arbitrary order—of option prices on a given process (X_ε), as ε tends to zero. This general formulation allows us, by a suitable scaling, to obtain both small-time and large-time expansions. Indeed, setting $X_\varepsilon \equiv X_\varepsilon$ or $X_\varepsilon \equiv \varepsilon X_{1/\varepsilon}$ yields two expansions for different regimes, small and large maturities. The first rescaling is detailed in Section 2.1.1 and the second one in Section 2.1.2. In each case, we shall make the computations explicit for the BSM (1.1) case, which will also be needed to translate these expansions into expansions for the forward implied volatility in Section 2.2. Such expansions for European option prices and their corresponding spot implied volatilities are known for many models, and we shall consider here forward-start options (which clearly reduce to standard vanilla options when the forward-start date is null).

2.1. **Forward-start option asymptotics.** Let (X_ε) be a stochastic process with re-normalised moment generating function (mgf)

$$\Lambda_\varepsilon(u) := \varepsilon \log \mathbb{E}\left[\exp\left(\frac{uX_\varepsilon}{\varepsilon}\right)\right], \qquad \text{for all } u \in \mathcal{D}_\varepsilon, \tag{2.1}$$

where $\mathcal{D}_\varepsilon := \{u \in \mathbb{R} : |\Lambda_\varepsilon(u)| < \infty\}$. We now introduce the following critical assumptions.

Assumption 2.1. For each $u \in \mathcal{D}_{0,0}$ the re-normalised mgf can be represented as

$$\Lambda_\varepsilon(u) = \sum_{i=0}^{2} \Lambda_{i,0}(u)\,\varepsilon^i + \mathcal{O}\left(\varepsilon^3\right), \quad \text{as } \varepsilon \text{ tends to } 0. \tag{2.2}$$

Further we suppose that for all $\varepsilon > 0$ the map $\Lambda_\varepsilon : \mathcal{D}_\varepsilon \mapsto \mathbb{R}$ is infinitely differentiable on $\mathcal{D}^o_{0,0} \subseteq \mathcal{D}_\varepsilon$ and $0 \in \mathcal{D}^o_{0,0}$ where we define $\mathcal{D}_{0,0} := \{u \in \mathbb{R} : |\Lambda_{0,0}(u)| < \infty\}$ and $\mathcal{D}^o_{0,0}$ is the interior of $\mathcal{D}_{0,0}$ in $\mathbb{R}$.

The infinite differentiability assumption of the map Λ_ε could also be relaxed by a $\mathcal{C}^4(\mathcal{D}^o_{0,0})$ condition but this hardly makes any difference in practice and does, however, render some formulations awkward. If the expansion (2.2) holds up to some higher order $n \geq 3$, one can in principle show that both forward-start option prices and the forward implied volatility expansions below hold to order n as well. However expressions for the coefficients of higher order are extremely cumbersome and scarcely useful in practice.

Definition 2.2. [15, Definition 2.3.5] A convex function $h : \mathbb{R} \supset \mathcal{D}_h \to (-\infty, \infty]$ is essentially smooth if

(i) $\mathcal{D}^o_h$ is non-empty;
(ii) h is differentiable in $\mathcal{D}^o_h$;
(iii) h is steep, in other words $\lim_{n\to\infty} |h'(u_n)| = \infty$ for every sequence $(u_n)_{n\in\mathbb{N}}$ in $\mathcal{D}^o_h$ that converges to a boundary point of $\mathcal{D}^o_h$.

Assumption 2.3. $\Lambda_{0,0}$ is strictly convex and essentially smooth on $\mathcal{D}^o_{0,0}$.

Define the function $\Lambda^* : \mathbb{R} \to \mathbb{R}_+$ as the Fenchel-Legendre transform of $\Lambda_{0,0}$:

$$\Lambda^*(k) := \sup_{u\in\mathcal{D}_{0,0}} \{uk - \Lambda_{0,0}(u)\}, \qquad \text{for all } k \in \mathbb{R}. \tag{2.3}$$

For ease of exposition in the paper we will use the notation

$$\Lambda_{i,l}(u) := \partial^l_u \Lambda_{i,0}(u) \quad \text{for } l \geq 1, \tag{2.4}$$

where $\Lambda_{i,0}$ is defined in (2.2) for $i=0,1,2$. The following lemma gathers some immediate properties of the functions Λ^* and Λ_{ij} which will be needed later.

Lemma 2.4. *Under Assumptions 2.1 and 2.3, the following properties hold:*

(i) For any $k\in\mathbb{R}$, there exists a unique $u^*(k)\in\mathcal{D}^o_{0,0}$ such that

$$\Lambda_{0,1}(u^*(k))=k, \tag{2.5}$$

$$\Lambda^*(k)=u^*(k)k-\Lambda_{0,0}\left(u^*(k)\right); \tag{2.6}$$

(ii) Λ^* is strictly convex and differentiable on $\mathbb{R}$;

(iii) if $a\in\mathcal{D}^o_{0,0}$ such that $\Lambda_{0,0}(a)=0$, then $\Lambda^*(k)>ak$ for all $k\in\mathbb{R}\setminus\{\Lambda_{0,1}(a)\}$ and $\Lambda^*(\Lambda_{0,1}(a))=a\Lambda_{0,1}(a)$.

Proof.

(i) By Assumption 2.3 and 2.1 $\Lambda_{0,1}$ is a strictly increasing differentiable function from $-\infty$ to ∞ on $\mathcal{D}_{0,0}$.

(ii) By (i), $\partial_k\Lambda^*(k)=\Lambda_{0,1}^{-1}(k)$ for all $k\in\mathbb{R}$. In particular $\partial_k\Lambda^*$ is strictly increasing on $\mathbb{R}$.

(iii) Since $\Lambda_{0,1}$ is strictly increasing, $\Lambda_{0,1}(a)=k$ if and only if $u^*(k)=a$ and then $\Lambda^*(\Lambda_{0,1}(a))=a\Lambda_{0,1}(a)$ using (2.6). Using the definition (2.3) with $a\in\mathcal{D}^o_{0,0}$ and $\Lambda_{0,0}(a)=0$ gives $\Lambda^*(k)\geq ak$. Since Λ^* is strictly convex from (ii) it follows that $\Lambda^*(k)>ak$ for all $k\in\mathbb{R}\setminus\{\Lambda_{0,1}(a)\}$.

□

For ease of notation we shall write $\Lambda_{j,l}$ in place of $\Lambda_{j,l}\left(u^*(k)\right)$. Let $f:\mathbb{R}_+\to\mathbb{R}_+$ be a function such that

$$f(\varepsilon)\varepsilon=c+\mathcal{O}(\varepsilon),\quad \text{for some } c\geq 0,\ \text{as } \varepsilon \text{ tends to zero.} \tag{2.7}$$

For any $b\geq 0$ we now define the functions $A_b,\bar{A}_b:\mathbb{R}\setminus\{\Lambda_{0,1}(0),\Lambda_{0,1}(b)\}\times(0,\infty)\to\mathbb{R}$ by

$$\bar{A}_b(k,\varepsilon):=\frac{b\sqrt{\varepsilon}\mathbb{1}_{\{b>0\}}+\varepsilon^{3/2}f(\varepsilon)\mathbb{1}_{\{b=0\}}}{u^*(k)\left(u^*(k)-b\right)\sqrt{2\pi\Lambda_{0,2}}}\exp(\Lambda_{1,0}), \tag{2.8}$$

$$A_b(k,\varepsilon):=\mathrm{e}^{-\Lambda^*(k)/\varepsilon+kf(\varepsilon)}\bar{A}_b(k,\varepsilon)\left(1+\Upsilon(b,k)\varepsilon+\frac{u^*(k)(\varepsilon f(\varepsilon)-b)}{\left(u^*(k)-b\right)b}\mathbb{1}_{\{b>0\}}+\frac{\varepsilon f(\varepsilon)}{u^*(k)}\mathbb{1}_{\{b=0\}}+\mathcal{O}\left(\varepsilon^2\right)\right), \tag{2.9}$$

where $\Upsilon:[0,\infty)\times\mathbb{R}\backslash\{\Lambda_{0,1}(0),\Lambda_{0,1}(b)\}\to\mathbb{R}$ is given by

$$\begin{aligned}\Upsilon(b,k):=\Lambda_{2,0}-\frac{5\Lambda_{0,3}^2}{24\Lambda_{0,2}^3}+\frac{4\Lambda_{1,1}\Lambda_{0,3}+\Lambda_{0,4}}{8\Lambda_{0,2}^2}-\frac{\Lambda_{1,1}^2+\Lambda_{1,2}}{2\Lambda_{0,2}}-\frac{\Lambda_{0,3}}{2u^*(k)\Lambda_{0,2}^2}-\frac{\Lambda_{0,3}}{2\left(u^*(k)-b\right)\Lambda_{0,2}^2}\\ -\frac{\Lambda_{1,1}\left(b-2u^*(k)\right)+3}{u^*(k)\left(u^*(k)-b\right)\Lambda_{0,2}}-\frac{b^2}{u^*(k)^2\left(u^*(k)-b\right)^2\Lambda_{0,2}}.\end{aligned} \tag{2.10}$$

Remark 2.5. The domain of definition of A_b excludes the set $\{\Lambda_{0,1}(0),\Lambda_{0,1}(b)\}=\{k\in\mathbb{R}:u^*(k)\in\{0,b\}\}$. For all k in this domain, $\Lambda_{0,2}(u^*(k))>0$ by Assumption 2.3, so that A_b is a well-defined real-valued function.

The main result of the section is the following theorem on asymptotics of option prices. A quick glimpse at the proof of Theorem 2.6 in Section 5.1 shows that this result can be extended to any arbitrary order.

Theorem 2.6. *Let (X_ε) satisfy Assumptions 2.1 and 2.3, and $f:\mathbb{R}_+\to\mathbb{R}_+$ be a function satisfying* (2.7) *with constant $c\in\mathcal{D}^o_{0,0}\cap\mathbb{R}_+$. Then the following expansion holds for all $k\in\mathbb{R}\backslash\{\Lambda_{0,1}(0),\Lambda_{0,1}(c)\}$ as $\varepsilon\searrow 0$:*

$$\begin{aligned}A_c(k,\varepsilon)&=\mathbb{E}\left[\left(\mathrm{e}^{X_\varepsilon f(\varepsilon)}-\mathrm{e}^{kf(\varepsilon)}\right)^+\right]\mathbb{1}_{\{k>\Lambda_{0,1}(c)\}}+\mathbb{E}\left[\left(\mathrm{e}^{kf(\varepsilon)}-\mathrm{e}^{X_\varepsilon f(\varepsilon)}\right)^+\right]\mathbb{1}_{\{k<\Lambda_{0,1}(0)\}}\\&\quad-\mathbb{E}\left[\mathrm{e}^{X_\varepsilon f(\varepsilon)}\wedge\mathrm{e}^{kf(\varepsilon)}\right]\mathbb{1}_{\{\Lambda_{0,1}(0)<k<\Lambda_{0,1}(c)\}},\end{aligned}$$

where $\Lambda_{0,1}$ is defined in (2.4) *and $A_c(k,\varepsilon)$ in* (2.9).

Remark 2.7. In the case $c = 0$, the expansion holds for all $k \in \mathbb{R}\backslash\{\Lambda_{0,1}(0)\}$ and the last term on the right-hand side disappears, since the indicator function is taken over an empty set.

Let $(X_t)_{t\geq 0}$ be a stochastic process. For any $t \geq 0$, we define (pathwise) the process $(X_\tau^{(t)})_{\tau\geq 0}$ by

$$X_\tau^{(t)} := X_{t+\tau} - X_t \tag{2.11}$$

We now specialise Theorem 2.6 to forward-start option asymptotics.

2.1.1. *Diagonal small-maturity asymptotics.* We first consider asymptotics when both t and τ are small, which we term *diagonal small-maturity asymptotics.*

Corollary 2.8. *If* $\left(X_{\varepsilon\tau}^{(\varepsilon t)}\right)_{\varepsilon>0}$ *satisfies Assumptions 2.1, 2.3, then the following expansion holds:*

$$\frac{\mathrm{e}^{-\Lambda^*(k)/\varepsilon+k+\Lambda_{1,0}}\varepsilon^{3/2}}{u^*(k)^2\sqrt{2\pi\Lambda_{0,2}}}\left(1+\left(\Upsilon(0,k)+\frac{1}{u^*(k)}\right)\varepsilon+\mathcal{O}\left(\varepsilon^2\right)\right) = \mathbb{E}\left[\left(\mathrm{e}^{X_{\varepsilon\tau}^{(\varepsilon t)}}-\mathrm{e}^k\right)^+\right]\mathbb{1}_{\{k>\Lambda_{0,1}(0)\}} + \mathbb{E}\left[\left(\mathrm{e}^k-\mathrm{e}^{X_{\varepsilon\tau}^{(\varepsilon t)}}\right)^+\right]\mathbb{1}_{\{k<\Lambda_{0,1}(0)\}},$$

as ε *tends to zero, where* $\Lambda_{0,1}$, Λ^*, $u^*(k)$, Υ *and* $\Lambda_{i,l}$ *are defined in* (2.4)*,* (2.3)*,* (2.5)*,* (2.10) *and* (2.4)*.*

Proof. Set $(X_\varepsilon) := \left(X_{\varepsilon\tau}^{(\varepsilon t)}\right)$ and $f \equiv 1$. Then $c = 0$ and the corollary follows from Theorem 2.6. □

Corollary 2.9. *In the BSM model* (1.1) *the following expansion holds as* ε *tends to zero:*

$$\mathbb{E}\left[\left(\mathrm{e}^{X_{\varepsilon\tau}^{(\varepsilon t)}}-\mathrm{e}^k\right)^+\right]\mathbb{1}_{\{k>0\}}+\mathbb{E}\left[\left(\mathrm{e}^k-\mathrm{e}^{X_{\varepsilon\tau}^{(\varepsilon t)}}\right)^+\right]\mathbb{1}_{\{k<0\}} = \frac{\mathrm{e}^{k/2-k^2/(2\sigma^2\tau\varepsilon)}\left(\sigma^2\tau\varepsilon\right)^{3/2}}{k^2\sqrt{2\pi}}\left[1-\left(\frac{3}{k^2}+\frac{1}{8}\right)\sigma^2\tau\varepsilon+\mathcal{O}(\varepsilon^2)\right].$$

Proof. For the rescaled (forward) process $\left(X_{\varepsilon\tau}^{(\varepsilon t)}\right)_{\varepsilon>0}$ in the BSM model (1.1) we have $\Lambda_\varepsilon(u) = \Lambda_{0,0}(u)+\varepsilon\Lambda_{1,0}(u)$ for $u \in \mathbb{R}$, where $\Lambda_{0,0}(u) = u^2\sigma^2\tau/2$ and $\Lambda_{1,0}(u) = -u\sigma^2\tau/2$. It follows that $\Lambda_{0,1}(u) = u\sigma^2\tau$, $\Lambda_{0,2}(u) = \sigma^2\tau$ and $\Lambda_{1,1}(u) = -\sigma^2\tau/2$. For any $k \in \mathbb{R}$, $u^*(k) := k/(\sigma^2\tau)$ is the unique solution to the equation $\Lambda_{0,1}(u^*(k)) = k$ and $\Lambda^*(k) = k^2/(2\sigma^2\tau)$. $\Lambda_{0,0}$ is essentially smooth and convex on $\mathbb{R}$ and the BSM model satisfies Assumptions 2.1 and 2.3. Since $0 \in \mathcal{D}_{0,0}^o$ and $\Lambda_{0,1}(0) = 0$, the result follows from Corollary 2.8. □

It is natural to wonder why we considered diagonal small-maturity asymptotics and not the small-maturity asymptotic of $\sigma_{t,\tau}$ for fixed $t > 0$. In this case it turns out that in many cases of interest (stochastic volatility models, time-changed exponential Lévy models), the forward smile blows up to infinity (except at-the-money) as τ tends to zero. However under the assumptions given above, this degenerate behaviour does not occur in the diagonal small-maturity regime (Corollary 2.8). We leave the precise study of this degeneracy for future research, but provide a preliminary conjecture explaining the origin of this exploding behaviour. Consider a two-state Markov-chain $\mathrm{d}X_t = -\frac{1}{2}V\mathrm{d}t + \sqrt{V}\mathrm{d}W_t$, starting at $X_0 = 0$, where W is a standard Brownian motion and where V is independent of W and takes value V_1 with probability $p \in (0,1)$ and value $V_2 \in (0, V_1)$ with probability $1-p$. Conditioning on V and by the independence assumption, we have

$$\mathbb{E}\left(\mathrm{e}^{u(X_{t+\tau}-X_t)}\right) = p\mathrm{e}^{V_1u\tau(u-1)/2} + (1-p)\mathrm{e}^{V_2u\tau(u-1)/2}, \qquad \text{for all } u \in \mathbb{R}.$$

Consider now the small-maturity regime where $\varepsilon = \tau$, $f(\varepsilon) = 1$ and $X_\varepsilon := X_\varepsilon^{(t)}$ for a fixed $t > 0$. In this case an expansion for the re-scaled mgf in (2.2) as τ tends to zero is given by

$$\Lambda_\varepsilon(u) = \tau\log\mathbb{E}\left(\mathrm{e}^{u(X_{t+\tau}-X_t)/\tau}\right) = \frac{V_1}{2}u^2 + \tau\log\left(p\mathrm{e}^{-V_1u/2}\right) + \tau\mathcal{O}\left(\mathrm{e}^{-u^2(V_1-V_2)/(2\tau)}\right), \qquad \text{for all } u \in \mathbb{R}.$$

Since $V_1 > V_2$ the remainder tends to zero exponentially fast as τ tends to zero. The assumptions of Theorem 2.6 are clearly satisfied and a simple calculation shows that $\lim_{\tau\searrow 0}\sigma_{t,\tau}(k) = \sqrt{V_1}$. This example naturally extends to an n-state Markov chain, and a natural conjecture is hence that the small-maturity forward smile does not blow up if and only if the quadratic variation of the process is bounded. In practice, most models are of unbounded quadratic variation (see the examples in Section 3 below), and hence the diagonal small-maturity asymptotic is a natural scaling.

2.1.2. *Large-maturity asymptotics.* We now consider large-maturity asymptotics, when τ is large and t is fixed. Define the function $B^\infty : \mathbb{R}\backslash\{\Lambda_{0,1}(0), \Lambda_{0,1}(1)\} \times (0,\infty) \to \mathbb{R}$ by

$$B^\infty(k,\tau) := \frac{\mathrm{e}^{-\tau(\Lambda^*(k)-k)+\Lambda_{1,0}}\tau^{-1/2}}{u^*(k)\left(u^*(k)-1\right)\sqrt{2\pi\Lambda_{0,2}}}\left(1+\frac{\Upsilon(1,k)}{\tau}+\mathcal{O}\left(\frac{1}{\tau^2}\right)\right), \tag{2.12}$$

where Λ^*, u^*, Υ, $\Lambda_{i,l}$ are defined in (2.3), (2.5), (2.10), (2.4). From Remark 2.5, the function B^∞ is well-defined. We now have the following large-maturity asymptotic for forward-start options.

Corollary 2.10. *If the process* $\left(\tau^{-1}X_\tau^{(t)}\right)_{\tau>0}$ *satisfies Assumptions 2.1, 2.3 with* $\varepsilon = \tau^{-1}$ *and* $1 \in \mathcal{D}_{0,0}^o$, *then the following expansion holds as* τ *tends to infinity:*

$$\mathbb{E}\left[\left(\mathrm{e}^{X_\tau^{(t)}}-\mathrm{e}^{k\tau}\right)^+\right]\mathbb{1}_{\{k>\Lambda_{0,1}(1)\}}+\mathbb{E}\left[\left(\mathrm{e}^{k\tau}-\mathrm{e}^{X_\tau^{(t)}}\right)^+\right]\mathbb{1}_{\{k<\Lambda_{0,1}(0)\}}-\mathbb{E}\left[\mathrm{e}^{X_\tau^{(t)}}\wedge\mathrm{e}^{k\tau}\right]\mathbb{1}_{\{\Lambda_{0,1}(0)<k<\Lambda_{0,1}(1)\}} = B^\infty(k,\tau),$$

where $\Lambda_{0,1}$ *is defined in* (2.4) *and* B^∞ *in* (2.12).

Proof. Let $(X_\varepsilon) := \left(\varepsilon X_{1/\varepsilon}^{(t)}\right)$, $\varepsilon := 1/\tau$ and $f(\varepsilon) \equiv 1/\varepsilon$ $(c=1)$, then the result follows from Proposition 2.6. □

In the BSM case (1.1), define the function $B_{BS}^\infty : \mathbb{R}\backslash\{-\sigma^2/2, \sigma^2/2\} \times (0,\infty) \to \mathbb{R}$ by

$$B_{BS}^\infty(k,\tau) := \exp\left(-\tau\left(\frac{\left(k+\sigma^2/2\right)^2}{2\sigma^2}-k\right)\right)\frac{4\sigma^3\tau^{-1/2}}{\left(4k^2-\sigma^4\right)\sqrt{2\pi}}\left(1-\frac{4\sigma^2\left(\sigma^4+12k^2\right)}{\left(4k^2-\sigma^4\right)^2\tau}+\mathcal{O}\left(\frac{1}{\tau^2}\right)\right), \tag{2.13}$$

and we have the following corollary.

Corollary 2.11. *In the BSM model* (1.1) *the following expansion holds as* τ *tends to infinity:*

$$\mathbb{E}\left[\left(\mathrm{e}^{X_\tau^{(t)}}-\mathrm{e}^{k\tau}\right)^+\right]\mathbb{1}_{\{k>\sigma^2/2\}}-\mathbb{E}\left[\mathrm{e}^{X_\tau^{(t)}}\wedge\mathrm{e}^{k\tau}\right]\mathbb{1}_{\{-\sigma^2/2<k<\sigma^2/2\}}+\mathbb{E}\left[\left(\mathrm{e}^{k\tau}-\mathrm{e}^{X_\tau^{(t)}}\right)^+\right]\mathbb{1}_{\{k<-\sigma^2/2\}} = B_{BS}^\infty(k,\tau).$$

Proof. Consider the process $\left(X_\tau^{(t)}/\tau\right)_{\tau>0}$ and set $\varepsilon = \tau^{-1}$. In the BSM model (1.1) for any $u \in \mathbb{R}$, we have $\Lambda_\varepsilon(u) = \tau^{-1}\log\mathbb{E}(\exp(uX_\tau^{(t)})) = \Lambda_{0,0}(u) = \frac{1}{2}\sigma^2u(u-1)$. Thus $\Lambda_{0,1}(u) = \sigma^2(u-1/2)$ and $\Lambda_{0,2}(u) = \sigma^2$. For any $k \in \mathbb{R}$, the equation $\Lambda_{0,1}(u^*(k)) = k$ has a unique solution $u^*(k) = 1/2 + k/\sigma^2$ and therefore $\Lambda^*(k) = \left(k+\sigma^2/2\right)^2/(2\sigma^2)$. $\Lambda_{0,0}$ is essentially smooth and strictly convex on $\mathbb{R}$ and Assumptions 2.1 and that 2.3 are satisfied. Since $\{0,1\} \subset \mathcal{D}_{0,0}^o$ the result follows from Corollary 2.10. □

2.2. **Forward smile asymptotics.** In this section we translate our results on forward-start options into asymptotics of the forward implied volatility smile $k \mapsto \sigma_{t,\tau}(k)$. We first focus on the diagonal small-maturity case.

For $i=0,1,2$ we define the functions $v_i:\mathbb{R}^*\times[0,\infty)\times(0,\infty)\to\mathbb{R}$ by

$$
\begin{aligned}
v_0(k,t,\tau) &:= \frac{k^2}{2\tau\Lambda^*(k)},\\
v_1(k,t,\tau) &:= \frac{2\tau v_0(k,t,\tau)^2}{k^2}\log\left(\frac{k^2\mathrm{e}^{\Lambda_{1,0}(u^*(k))}}{u^*(k)^2\sqrt{\Lambda_{0,2}(u^*(k))}\,(\tau v_0(k,t,\tau))^{3/2}}\right)+\frac{\tau v_0(k,t,\tau)^2}{k},\\
v_2(k,t,\tau) &:= \frac{2\tau^2 v_0^3(k,t,\tau)}{k^2}\left(\frac{3}{k^2}+\frac{1}{8}\right)+\frac{2\tau v_0^2(k,t,\tau)}{k^2}\left(\Upsilon(0,k)+\frac{1}{u^*(k)}\right)\\
&\quad+\frac{v_1^2(k,t,\tau)}{v_0(k,t,\tau)}-\frac{3\tau}{k^2}v_0(k,t,\tau)v_1(k,t,\tau),
\end{aligned}
\tag{2.14}
$$

where Λ^*, u^*, $\Lambda_{i,l}$, Υ are defined in (2.3) (2.5), (2.4), (2.10). Also if $\Lambda_{0,1}(0)=0$ then $\Lambda^*(k)>0$ for $k\in\mathbb{R}^*$ and $\Lambda^*(0)=0$ from Assumption 2.1 and Lemma 2.4(iii) so that v_0 is always strictly positive. A direct application of L'Hôpitals rule, together with Lemma 2.4(i)(ii), Assumption 2.1 and $\Lambda_{0,1}(0)=0$, shows that for any $t\geq 0$, $\tau>0$, the map $v_0(\cdot,t,\tau)$ can be extended by continuity at the origin with $v_0(0,t,\tau)=1/(\tau u^{*'}(0))$. All the v_i ($i=0,1,2$) are hence well-defined real-valued functions (see also Remark 2.5). The diagonal small-maturity forward smile asymptotic is now given in the following proposition, proved in Section 5.1.

Proposition 2.12. *Suppose that* $\left(X_{\varepsilon\tau}^{(\varepsilon t)}\right)_{\varepsilon>0}$ *satisfies Assumptions 2.1 and 2.3, and that* $\Lambda_{0,1}(0)=0$ *(defined in (2.4)). The following expansion holds for the corresponding forward smile for all* $k\in\mathbb{R}^*$ *as* ε *tends to zero:*

$$
\sigma^2_{\varepsilon t,\varepsilon\tau}(k)=v_0(k,t,\tau)+v_1(k,t,\tau)\varepsilon+v_2(k,t,\tau)\varepsilon^2+\mathcal{O}\left(\varepsilon^3\right),
\tag{2.15}
$$

where v_0, v_1 *and* v_2 *are given in (2.14).*

In the large-maturity case, define for $i=0,1,2$, the functions $v_i^\infty:\mathbb{R}\backslash\{\Lambda_{0,1}(0),\Lambda_{0,1}(1)\}\times[0,\infty)\to\mathbb{R}$ by

$$
\begin{aligned}
v_0^\infty(k,t) &:= \begin{cases} 2\left(2\Lambda^*(k)-k-2\sqrt{\Lambda^*(k)(\Lambda^*(k)-k)}\right), & \text{if } k\in\mathbb{R}\backslash\left[\Lambda_{0,1}(0),\Lambda_{0,1}(1)\right],\\ 2\left(2\Lambda^*(k)-k+2\sqrt{\Lambda^*(k)(\Lambda^*(k)-k)}\right), & \text{if } k\in\left(\Lambda_{0,1}(0),\Lambda_{0,1}(1)\right),\end{cases}\\
v_1^\infty(k,t) &:= \frac{8v_0^\infty(k,t)^2}{4k^2-v_0^\infty(k,t)^2}\left(\Lambda_{1,0}(u^*(k))+\log\left(\frac{4k^2-v_0^\infty(k,t)^2}{4(u^*(k)-1)u^*(k)v_0^\infty(k,t)^{3/2}\sqrt{\Lambda_{0,2}(u^*(k))}}\right)\right),\\
v_2^\infty(k,t) &:= \frac{4}{v_0^\infty(k,t)\left(v_0^\infty(k,t)^2-4k^2\right)^3}\Big[8k^4v_1^\infty(k,t)v_0^\infty(k,t)^2\left(v_1^\infty(k,t)+6\right)-16k^6v_1^\infty(k,t)^2\\
&\quad-2\Upsilon(1,k)v_0^\infty(k,t)^3\left(v_0^\infty(k,t)^2-4k^2\right)^2-k^2v_0^\infty(k,t)^4\left(96+v_1^\infty(k,t)^2+8v_1^\infty(k,t)\right)\\
&\quad-v_0^\infty(k,t)^6\left(v_1^\infty(k,t)+8\right)\Big].
\end{aligned}
\tag{2.16}
$$

Λ^* is defined in 2.3, u^* in (2.5), $\Lambda_{i,l}$ in (2.4) and Υ in (2.10). Since $\{0,1\}\subset\mathcal{D}^o_{0,0}$ and $\Lambda_{0,0}(1)=\Lambda_{0,0}(0)=0$, we always have $\Lambda^*(k)\geq\max(0,k)$ from Lemma 2.4(iii). One can also check that $0<v_0^\infty(k,t)<2|k|$ for $k\in\mathbb{R}\backslash\left[\Lambda_{0,1}(0),\Lambda_{0,1}(1)\right]$ and $v_0^\infty(k,t)>2|k|$ for $k\in(\Lambda_{0,1}(0),\Lambda_{0,1}(1))$. Together with Remark 2.5, this implies that the functions v_i^∞ ($i=0,1,2$) are always well-defined and real-valued.

Remark 2.13. By Assumption 2.1 and Lemma 2.4(iii) we have $\Lambda^*(\Lambda_{0,1}(0))=0$. Further by the assumptions in Proposition 2.14 below we have $1\in\mathcal{D}^o_{0,0}$ and $\Lambda_{0,0}(1)=0$. Again from Lemma 2.4(iii) this implies that $\Lambda^*(\Lambda_{0,1}(1))=\Lambda_{0,1}(1)$. Hence for all $t\geq 0$, $v_0^\infty(\cdot,t)$ can be extended by continuity on $\mathbb{R}$.

The large-maturity forward smile asymptotic is given in the following proposition, proved in Section 5.1.

Proposition 2.14. *Suppose that* $\left(\tau^{-1}X_\tau^{(t)}\right)_{\tau>0}$ *satisfies Assumptions 2.1, 2.3 with* $\varepsilon=\tau^{-1}$ *and that* $\Lambda_{0,0}(1)=0$ *with* $1\in\mathcal{D}_{0,0}^o$ *(all defined in Assumption 2.1). The following expansion then holds for the forward smile as* τ *tends to infinity for* $k\in\mathbb{R}\backslash\{\Lambda_{0,1}(0),\Lambda_{0,1}(1)\}$*:*

$$\sigma_{t,\tau}^2(k\tau)=v_0^\infty(k,t)+\frac{v_1^\infty(k,t)}{\tau}+\frac{v_2^\infty(k,t)}{\tau^2}+\mathcal{O}\left(\frac{1}{\tau^3}\right), \tag{2.17}$$

where v_0^∞*,* v_1^∞ *and* v_2^∞ *are defined in* (2.16)*.*

Remark 2.15.

(i) If we set $t=0$ in (2.15) and (2.17) then we recover—and actually improve—the implied volatility asymptotics obtained in [17], [19], [20], [21], [22].

(ii) The forward smile results can be extended to a deterministic interest rate setting by considering the forward price instead of the stock price and re-scaling the strike appropriately.

For the $(\mathcal{F}_u)$-martingale price $(\mathrm{e}^{X_u})_{u\geq 0}$ (under $\mathbb{P}$) define the stopped process $\widetilde{X}_u^t:=X_{t\wedge u}$ for any $t>0$. Following [43] define a new measure $\widetilde{\mathbb{P}}$ by

$$\widetilde{\mathbb{P}}(A):=\mathbb{E}\left(\mathrm{e}^{\widetilde{X}_{t+\tau}^t}\mathbb{1}_A\right),\qquad \text{for every } A\in\mathcal{F}_{t+\tau}. \tag{2.18}$$

The stopped process $(\mathrm{e}^{\widetilde{X}_u^t})_{u\geq 0}$ is a $(\mathcal{F}_{t\wedge u})_u$-martingale and (2.18) defines the stopped-share-price measure $\widetilde{\mathbb{P}}$. The following proposition shows how the Type-II forward smile $\widetilde{\sigma}_{t,\tau}$ can be incorporated into our framework.

Proposition 2.16. *If* $\left(\mathrm{e}^{X_t}\right)_{t\geq 0}$ *is a* $(\mathcal{F}_t)$*-martingale under* $\mathbb{P}$*, then Propositions 2.12 and 2.14 hold for the Type-II forward smile* $\widetilde{\sigma}_{t,\tau}$ *with the mgf* (2.1) *calculated under* $\widetilde{\mathbb{P}}$*.*

Proof. We can write the value of our Type-II forward-start call option as

$$\mathbb{E}\left[\left(\mathrm{e}^{X_{t+\tau}}-\mathrm{e}^{k+X_t}\right)^+\right]=\mathbb{E}\left[\mathrm{e}^{X_t}\left(\mathrm{e}^{X_{t+\tau}-X_t}-\mathrm{e}^k\right)^+\right]=\mathbb{E}\left[\mathrm{e}^{\widetilde{X}_{t+\tau}^t}\left(\mathrm{e}^{X_{t+\tau}-X_t}-\mathrm{e}^k\right)^+\right]=\widetilde{\mathbb{E}}\left[\left(\mathrm{e}^{X_{t+\tau}-X_t}-\mathrm{e}^k\right)^+\right].$$

Proposition 2.6 and Corollaries 2.8, 2.10 hold in this case with all expectations (and the mgf in (2.1)) calculated under the stopped measure $\widetilde{\mathbb{P}}$. An easy calculation shows that under $\widetilde{\mathbb{P}}$, the forward BSM mgf remains the same as under $\mathbb{P}$. Thus all the previous results carry over and the proposition follows. □

3. Applications

3.1. **Heston.** In this section, we apply our general results to the Heston model, in which the (log) stock price process is the unique strong solution to the following SDEs:

$$\begin{aligned} \mathrm{d}X_t &= -\frac{1}{2}V_t\mathrm{d}t+\sqrt{V_t}\mathrm{d}W_t, && X_0=0,\\ \mathrm{d}V_t &= \kappa\left(\theta-V_t\right)\mathrm{d}t+\xi\sqrt{V_t}\mathrm{d}Z_t, && V_0=v>0,\\ \mathrm{d}\left\langle W,Z\right\rangle_t &= \rho\mathrm{d}t, \end{aligned} \tag{3.1}$$

with $\kappa>0$, $\xi>0$, $\theta>0$ and $|\rho|<1$. The Feller SDE for the variance process has a unique strong solution by the Yamada-Watanabe conditions [38, Proposition 2.13, page 291]). The X process is a stochastic integral of the V process and is therefore well-defined. The Feller condition, $2\kappa\theta\geq\xi^2$, ensures that the origin is unattainable. Otherwise the origin is regular (hence attainable) and strongly reflecting (see [39, Chapter 15]). We do not require the Feller condition in our analysis since we work with the forward mgf of X which is always well-defined.

3.1.1. *Diagonal Small-Maturity Heston Forward Smile.* The objective of this section is to apply Proposition 2.12 to the Heston forward smile. We define the function $\Lambda : \mathcal{D}_{t,\tau} \times [0,\infty) \times (0,\infty) \to \mathbb{R}$ by

$$\Lambda(u,t,\tau) := \frac{uv}{\xi\left(\bar{\rho}\cot\left(\frac{1}{2}\xi\bar{\rho}\tau u\right) - \rho\right) - \frac{1}{2}\xi^2 tu}, \qquad \text{for all } u \in \mathcal{D}_{t,\tau}, \tag{3.2}$$

where

$$\mathcal{D}_{t,\tau} := \left\{u \in \mathbb{R} : \Lambda(u,0,\tau) < \frac{2v}{\xi^2 t}\right\} \qquad \text{and} \qquad \bar{\rho} := \sqrt{1-\rho^2}. \tag{3.3}$$

Further we let the function $L : \mathcal{D}_{t,\tau} \times [0,\infty) \times (0,\infty) \to \mathbb{R}$ be defined as

$$L(u,t,\tau) := L_0(u,\tau) + \Lambda(u,t,\tau)^2\left(\frac{vL_1(u,\tau)}{\Lambda(u,0,\tau)^2} - \frac{\kappa\xi^2 t^2}{4v}\right) - \kappa t\Lambda(u,t,\tau) - \frac{2\theta\kappa}{\xi^2}\log\left(1 - \frac{\xi^2 t\Lambda(u,0,\tau)}{2v}\right), \tag{3.4}$$

where the functions $L_i : \mathcal{D}_{t,\tau} \times (0,\infty) \to \mathbb{R}$ for $i = 0,1$, are defined by

(3.5)

$$L_0(u,\tau) := \frac{\kappa\theta}{\xi^2}\left((\mathtt{i}\xi\rho - d_0)\mathtt{i}\tau u - 2\log\left(\frac{1 - g_0\mathrm{e}^{-\mathtt{i}d_0\tau u}}{1-g_0}\right)\right),$$

$$L_1(u,\tau) := \frac{\mathrm{e}^{-\mathtt{i}d_0\tau u}}{\xi^2\left(1 - g_0\mathrm{e}^{-\mathtt{i}d_0\tau u}\right)}\left((\mathtt{i}\xi\rho - d_0)\mathtt{i}d_1\tau u + (d_1-\kappa)\left(1 - \mathrm{e}^{\mathtt{i}d_0\tau u}\right) + \frac{(\mathtt{i}\xi\rho - d_0)\left(1 - \mathrm{e}^{-\mathtt{i}d_0\tau u}\right)(g_1 - \mathtt{i}d_1 g_0\tau u)}{1 - g_0\mathrm{e}^{-\mathtt{i}d_0\tau u}}\right),$$

with

$$d_0 := \xi\bar{\rho}, \quad d_1 := \frac{\mathtt{i}\,(2\kappa\rho - \xi)}{2\bar{\rho}}, \qquad g_0 := \frac{\mathtt{i}\rho - \bar{\rho}}{\mathtt{i}\rho + \bar{\rho}} \qquad \text{and} \qquad g_1 := \frac{2\kappa - \xi\rho}{\xi\bar{\rho}\,(\bar{\rho} + \mathtt{i}\rho)^2}.$$

Remark 3.1. For any $t \geq 0, \tau > 0$ the functions L_0 and L_1 are well-defined real-valued functions for all $u \in \mathcal{D}_{t,\tau}$ (see Remark 5.6 for technical details). Also since $\Lambda(0,t,\tau)/\Lambda(0,0,\tau) = 1$, L is well-defined at $u = 0$.

Proposition 3.2. *In Heston, Corollary 2.8 and Proposition 2.12 hold with $\mathcal{D}_{0,0} = \mathcal{D}_{t,\tau}$, $\Lambda_{0,0} = \Lambda$ and $\Lambda_{1,0} = L$.*

Proof. We simply outline the proof of the proposition, and we refer the reader to Section 5.2.1 for the details.

- (i) In Lemma 5.3 we show that $\mathcal{D}_{0,0} = \mathcal{D}_{t,\tau}$ and $0 \in \mathcal{D}_{0,0}^o$;
- (ii) In Lemma 5.5 we show that the Heston diagonal small-maturity process has an expansion of the form given in Assumption 2.1 with $\Lambda_{0,0} = \Lambda$ and $\Lambda_{1,0} = L$;
- (iii) In Lemma 5.7 we show that Λ is strictly convex and essentially smooth on $\mathcal{D}_{t,\tau}^o$, i.e. Assumption 2.3;
- (iv) Λ_ε is infinitely differentiable and $\Lambda_{0,1}(0) = 0$.

We now apply Proposition 2.12 and this completes the proof. □

In order to gain some intuition on the role of the Heston parameters on the forward smile we expand (2.15) around the ATM point in terms of the log strike k. We now define the following functions:

$$\begin{aligned}
\nu_0(t,\tau) &:= \frac{\tau}{48}\left(24\kappa\theta + \xi^2\left(\rho^2 - 4\right) + 12v(\xi\rho - 2\kappa)\right) - \frac{t}{4}\left(\xi^2 + 4\kappa(v - \theta)\right), \\
\nu_1(t,\tau) &:= \frac{\rho\xi\tau}{24v}\left(\xi^2\left(1-\rho^2\right) - 2\kappa(v+\theta) + \xi\rho v\right) + \frac{\rho\xi^3 t}{8v}, \\
\nu_2(t,\tau) &:= \left(80\kappa\theta\left(13\rho^2 - 6\right) + \xi^2\left(521\rho^4 - 712\rho^2 + 176\right) + 40\rho^2 v(\xi\rho - 2\kappa)\right)\frac{\xi^2\tau}{7680v^2} \\
&\quad - \frac{\xi^2 t}{192v^2}\left(4\kappa\theta\left(16 - 7\rho^2\right) + \left(7\rho^2 - 4\right)\left(9\xi^2 + 4\kappa v\right)\right) + \frac{\xi^2 t^2}{32\tau v^2}\left(4\kappa(v - 3\theta) + 9\xi^2\right).
\end{aligned} \tag{3.6}$$

The proof of the following corollary is given in Section 5.2.1.

Corollary 3.3. *The following expansion holds for the Heston forward smile as ε and k tend to zero:*

$$\sigma^2_{\varepsilon t,\varepsilon\tau}(k) = v + \varepsilon\nu_0(t,\tau) + \left(\frac{\rho\xi}{2} + \varepsilon\nu_1(t,\tau)\right)k + \left(\frac{(4-7\rho^2)\xi^2}{48v} + \frac{\xi^2 t}{4\tau v} + \varepsilon\nu_2(t,\tau)\right)k^2 + \mathcal{O}(k^3) + \mathcal{O}(\varepsilon k^3) + \mathcal{O}(\varepsilon^2). \tag{3.7}$$

Remark 3.4. The following remarks should convey some practical intuition about the results above:

(i) For $t = 0$ this expansion perfectly lines up with Corollary 4.3 in [21] for the implied volatility smile.

(ii) Corollary 3.3 implies $\sigma_{\varepsilon t,\varepsilon\tau}(0) - \sigma_{0,\varepsilon\tau}(0) = -\frac{\varepsilon t}{8\sqrt{v}}\left(\xi^2 + 4\kappa(v-\theta)\right) + \mathcal{O}(\varepsilon^2)$, as ε tends to zero. For small enough maturity, the spot ATM volatility is higher than the forward one if and only if $\xi^2 + 4\kappa(v-\theta) > 0$. In particular, when $v \geq \theta$ the forward ATM volatility is lower than the corresponding spot ATM volatility and this difference is increasing in the forward-start dates and volatility of variance. In Figure 2 we plot this effect using $\theta = v$ and $\theta > v + \xi^2/(4\kappa)$. The relative values of v and θ impact the level of the forward smile vs spot smile.

(iii) Similarly, we can deduce some information on the forward skew from Corollary 3.3:

$$\partial_k\sigma_{\varepsilon t,\varepsilon\tau}(0) = \frac{\xi\rho}{4\sqrt{v}} + \frac{(4\nu_1(t,\tau)v - \xi\rho\nu_0(t,\tau))}{8v^{3/2}}\varepsilon + \mathcal{O}(\varepsilon^2),$$

and hence

$$\partial_k\sigma_{\varepsilon t,\varepsilon\tau}(0) - \partial_k\sigma_{0,\varepsilon\tau}(0) = \frac{\xi\rho t\left(3\xi^2 + 4\kappa(v-\theta)\right)}{32v^{3/2}}\varepsilon + \mathcal{O}(\varepsilon^2).$$

(iv) Likewise an expansion for the Heston forward convexity as ε tends to zero is given by

$$\partial_k^2\sigma_{\varepsilon t,\varepsilon\tau}(0) = \frac{\xi^2((2-5\rho^2)\tau + 6t)}{24\tau v^{3/2}} - \frac{\nu_0(t,\tau)\xi^2(3t + (1-4\rho^2)\tau) + 6\tau v(\rho\xi\nu_1(t,\tau) - 4\nu_2(t,\tau)v)}{24\tau v^{5/2}}\varepsilon + \mathcal{O}(\varepsilon^2),$$

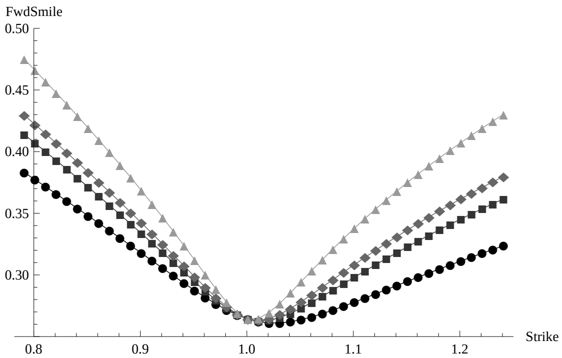
and in particular $\partial_k^2\sigma_{\varepsilon t,\varepsilon\tau}(0) - \partial_k^2\sigma_{0,\varepsilon\tau}(0) = \xi^2 t/(4\tau v^{3/2}) + \mathcal{O}(\varepsilon)$. For fixed maturity the forward convexity is always greater than the spot implied volatility convexity (see Figure 2) and this difference is increasing in the forward-start dates and volatility of variance. At zeroth order in ε the wings of the forward smile increase to arbitrarily high levels with decreasing maturity. (see Figure 1(a)) This effect has been mentioned qualitatively by practitioners [13]. As it turns out for fixed $t > 0$ the Heston forward smile blows up to infinity (except ATM) as the maturity tends to zero. This is clearly outside the scope of our main theorem, and we leave this degenerate case for future research.

In the Heston model $(\mathrm{e}^{X_t})_{t\geq 0}$ is a true martingale [3, Proposition 2.5]. Applying Proposition 2.16 with Lemma 5.8, giving the Heston forward mgf under the stopped-share-price measure, we derive the following asymptotic for the Type-II Heston forward smile $\widetilde{\sigma}_{t,\tau}$. The proof of Corollary 3.5 is omitted as it is analogous to the proofs of Proposition 3.2 and Corollary 3.3. Set

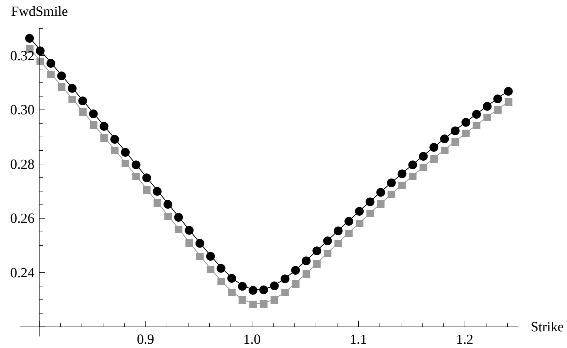
$$\widetilde{\nu}_0(t,\tau) := \nu_0(t,\tau) + \xi\rho v t, \qquad \widetilde{\nu}_1(t,\tau) := \nu_1(t,\tau), \qquad \widetilde{\nu}_2(t,\tau) := \nu_2(t,\tau) + \frac{\rho\xi^3 t}{48v}\left(7\rho^2 - 4\right) - \frac{\rho\xi^3 t^2}{8v\tau},$$

with ν_0, ν_1 and ν_2 defined in 3.6. In particular when $\rho = 0$ or $t = 0$, $\nu_i = \widetilde{\nu}_i$ $(i = 1,\dots,3)$, and hence the Heston forward smiles Type-I and Type-II are the same as shown in the following corollary.

Corollary 3.5. *The diagonal small-maturity expansion of the Heston Type-II forward smile as ε and k tend to zero is the same as the one in Corollary 3.3 with ν_0, ν_1 and ν_2 replaced by $\widetilde{\nu}_0$, $\widetilde{\nu}_1$ and $\widetilde{\nu}_2$.*



(a) Small-maturity forward smile explosion.



(b) Type I vs Type II forward smile.

FIGURE 1. In (a) we plot forward smiles with forward-start date $t = 1/2$ and maturities $\tau = 1/6, 1/12, 1/16, 1/32$ given by circles, squares, diamonds and triangles respectively using the Heston parameters $v = 0.07, \theta = 0.07, \kappa = 1, \rho = -0.6, \xi = 0.5$ and the asymptotic in Proposition 3.2. In (b) we plot the Type I (circles) vs Type 2 (squares) forward smile with $t = 1/2$, $\tau = 1/12$ and the Heston parameters $v = 0.07, \theta = 0.07, \kappa = 1, \rho = -0.2, \xi = 0.34$ using Corollaries 3.3 and 3.5.

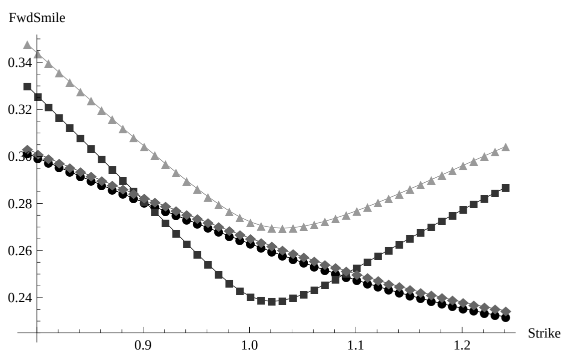


FIGURE 2. Forward smile vs spot smile with $v = \theta$ and $\theta > v + \xi^2/(4\kappa)$. Circles ($t = 0, \tau = 1/12$) and squares ($t = 1/2, \tau = 1/12$) use the Heston parameters $v = \theta = 0.07$,$\kappa = 1$, $\rho = -0.6$, $\xi = 0.3$. Diamonds ($t = 0, \tau = 1/12$) and triangles ($t = 1/2$, $\tau = 1/12$) use the same parameters but with $\theta = 0.1$. Plots use the asymptotic in Proposition 3.2.

3.1.2. *Large-maturity Heston forward smile.* We apply here Proposition 2.14 to the Heston forward smile. We shall use the standing assumption $\kappa > \rho\xi$, needed in the proof of Proposition 3.8. If this condition fails then we have a finite explosion time for moments greater than one for the price process (3.1) and consequently the limiting mgf is not essentially smooth on its effective domain and Assumption 2.3 is violated. This a standard assumption in the large-maturity implied volatility asymptotics literature [20, 22, 36]. It has been relaxed in [35] for the Heston model to study the limiting spot smile, but not the higher-order terms. This restriction bears no consequences in markets where the implied volatility skew is downward sloping, such as equity markets, where the correlation is negative. Define the quantities

$$u_{\pm} := \frac{\xi - 2\kappa\rho \pm \eta}{2\xi(1-\rho^2)} \qquad \text{and} \qquad u^*_{\pm} := \frac{\psi \pm \nu}{2\xi(\mathrm{e}^{\kappa t} - 1)}, \tag{3.8}$$

with

$$\eta := \sqrt{\xi^2(1-\rho^2)+(2\kappa-\rho\xi)^2}, \qquad \nu := \sqrt{\psi^2 - 16\kappa^2 \mathrm{e}^{\kappa t}} \qquad \text{and} \qquad \psi := \xi(\mathrm{e}^{\kappa t}-1) - 4\kappa\rho\mathrm{e}^{\kappa t}, \tag{3.9}$$

as well as the interval $\mathcal{D}_\infty \subset \mathbb{R}$ and the real numbers ρ_- and ρ_+ by

$$\mathcal{D}_\infty := \begin{cases} [u_-, u_+^*), & \text{if } -1 < \rho < \rho_- \text{ and } t > 0, \\ (u_-^*, u_+], & \text{if } \rho_+ < \rho < \min(1, \kappa/\xi), t > 0 \text{ and } \kappa > \rho_+\xi, \\ [u_-, u_+], & \text{if } \rho_- \le \rho \le \min(\rho_+, \kappa/\xi), \end{cases} \tag{3.10}$$

$$\rho_\pm := \frac{\mathrm{e}^{-2\kappa t}\left(\xi(\mathrm{e}^{2\kappa t}-1) \pm (\mathrm{e}^{\kappa t}+1)\sqrt{16\kappa^2\mathrm{e}^{2\kappa t} + \xi^2(1-\mathrm{e}^{\kappa t})^2}\right)}{8\kappa}. \tag{3.11}$$

Remark 3.6. The following remarks are proved in Lemmas 5.9, 5.10 and 5.12 and we summarise them here.

(i) When $\kappa/\xi < \rho_+$, the second case in (3.10) never occurs. From the proof of Lemma 5.10(i) if $\kappa \ge \xi$ then $\rho_+ \le 1$ so that $\min(\rho_+, \kappa/\xi) \le 1$.

(ii) If $t > 0$ and $\rho \le \rho_-$, then $u_+ > u_+^* > 1$ and if $t > 0$ and $\rho \ge \rho_+$, then $u_- < u_-^* < 0$.

(iii) We always have $-1 \le \rho_- < 0$ and if $\kappa > \rho_+\xi$ then $1/2 < \rho_+ \le 1$. Also $\rho_- = -1$ if and only if $t = 0$. When $\kappa > \rho_+\xi$ then $\rho_+ = 1$ if and only if $t = 0$. Finally ν defined in (3.9) is a well-defined real number for all $\rho \in [-1, \rho_-] \cup [\rho_+, 1]$.

We define the functions V and H from $\mathcal{D}_\infty$ to $\mathbb{R}$ by

$$V(u) := \frac{\kappa\theta}{\xi^2}\left(\kappa - \rho\xi u - d(u)\right) \qquad \text{and} \qquad H(u) := \frac{V(u) v \mathrm{e}^{-\kappa t}}{\kappa\theta - 2\beta_t V(u)} - \frac{2\kappa\theta}{\xi^2}\log\left(\frac{\kappa\theta - 2\beta_t V(u)}{\kappa\theta\left(1-\gamma(u)\right)}\right), \tag{3.12}$$

with

$$d(u) := \left((\kappa-\rho\xi u)^2 + u\xi^2(1-u)\right)^{1/2}, \quad \gamma(u) := \frac{\kappa - \rho\xi u - d(u)}{\kappa - \rho\xi u + d(u)}, \quad \text{and } \beta_t := \frac{\xi^2}{4\kappa}\left(1 - \mathrm{e}^{-\kappa t}\right). \tag{3.13}$$

Remark 3.7. We have $(\kappa - \rho\xi u)^2 + u(1-u)\xi^2 \ge 0$ and $\kappa\theta - 2\beta_t V(u) > 0$ for all $u \in \mathcal{D}_\infty$ from the proof of Proposition 5.13. Further by the definition of γ in (3.13) we have $\gamma(u) \in (-1, 1)$ for all $u \in \mathcal{D}_\infty$ using (5.38) and (5.39) in the proof of Proposition 5.13. So V and H are always well-defined real-valued functions.

Finally we define the functions $q^* : \mathbb{R} \to [u_-, u_+]$ and $V^* : \mathbb{R} \to \mathbb{R}_+$ by

$$q^*(x) := \frac{\xi - 2\kappa\rho + (\kappa\theta\rho + x\xi)\,\eta\left(x^2\xi^2 + 2x\kappa\theta\rho\xi + \kappa^2\theta^2\right)^{-1/2}}{2\xi\left(1-\rho^2\right)} \qquad \text{and} \qquad V^*(x) := q^*(x)x - V\left(q^*(x)\right), \tag{3.14}$$

where V and η are defined in (3.12) and in (3.9). The following proposition gives the large-maturity forward Heston smile in Case (iii) in (3.10).

Proposition 3.8. *Suppose that $\rho_- \le \rho \le \min(\rho_+, \kappa/\xi)$ in the Heston model with $\rho_\pm$ defined in* (3.11). *Then Corollary 2.10 and Proposition 2.14 hold with $\Lambda_{0,0} = V$, $\Lambda^* = V^*$, $u^* = q^*$, $\Lambda_{1,0} = H$, $\Lambda_{2,0} = 0$ and $\mathcal{D}_{0,0} = [u_-, u_+]$, where V, H, V^*, q^* and $u_\pm$ are defined in* (3.12), (3.14) *and* (3.8).

Proof. We simply outline the proof of the proposition, and we refer the reader to Section 5.2.2 for the details.

(i) In Proposition 5.13 we show that $\mathcal{D}_{0,0} = \mathcal{D}_\infty$ and that $\{0, 1\} \subset \mathcal{D}_\infty^o$;

(ii) In Lemma 5.14 and Remark 5.16 we show that the process has an expansion of the form given in Assumption 2.1 with $\Lambda_{0,0} = V$, $\Lambda_{1,0} = H$ and $\Lambda_{2,0} = 0$;

(iii) By Proposition 5.13 and Lemma 5.9, V is strictly convex and essentially smooth on $\mathcal{D}_\infty^o$ if $\rho_- \leq \rho \leq \min(\rho_+, \kappa/\xi)$; see also Remark 3.9(ii);

(iv) Λ_ε is infinitely differentiable and $V(1) = 0$ from Lemma 5.9;

(v) u^* can be computed in closed-form and is given by q^* in (3.14).

A direct application of Proposition 2.14 completes the proof. □

Remark 3.9.

(i) In the Heston model there is no t-dependence for v_0^∞ in (2.17) since V^* does not depend on t. Therefore under the conditions of the proposition, the limiting (zeroth order) smile is exactly of SVI form (see [27]).

(ii) For Cases (i) and (ii) in (3.10) the essential smoothness property in Assumption 2.3 is not satisfied and a different strategy needs to be employed to derive a sharp large deviation result for large-maturity forward-start options. We leave this analysis for future research.

(iii) For $t = 0$ we have $\rho_\pm = \pm 1$ and Proposition 3.8 agrees with and extends the Heston large-maturity implied volatility asymptotics in [20] and [22].

(iv) The condition $\rho \in [0, \min(1/2, \kappa/\xi)]$ is stronger than the condition in Proposition 3.8.

(v) Even though the rate function V^* does not depend on t, there is t-dependence through $\rho_\pm$ and the function H (see the ATM example below). That said, to zeroth order and correlation close to zero, the large-maturity forward smile is the same as the large-maturity spot smile. This is a very different result compared to the Heston small-maturity forward smile, as mentioned in Remark 3.4(iv), where large differences emerge between the forward smile and the spot smile at zeroth order.

We now give an example illustrating some of the differences between the Heston large-maturity forward smile and the large-maturity spot smile due to first order differences in the asymptotic (2.17). This ties in with Remark 3.9(v). Specifically we look at the forward ATM volatility which, when using Proposition 3.8 with $\rho_- \leq \rho \leq \min(\rho_+, \kappa/\xi)$, has the asymptotic

$$\sigma_{t,\tau}^2(0) = v_0^\infty(0) + v_1^\infty(0,t)/\tau + \mathcal{O}\left(1/\tau^2\right), \quad \text{as } \tau \text{ tends to infinity,}$$

with

$$\begin{aligned} v_0^\infty(0) &= \frac{4\theta\kappa(\eta - 2\kappa + \xi\rho)}{\xi^2(1-\rho^2)}, \\ v_1^\infty(0,t) &= \frac{16\kappa v(\rho\xi - 2\kappa + \eta)}{\Delta\xi^2} + \frac{16\kappa\theta}{\xi^2}\log\left(\frac{\Delta \mathrm{e}^{-\kappa t}\left(2\kappa - \xi\rho + \left(1-2\rho^2\right)\eta\right)}{8\kappa\left(1-\rho^2\right)^2\eta}\right) \\ &\quad - 8\log\left(\frac{\xi\left(1-\rho^2\right)^{3/2}\sqrt{\eta\left(2\xi\rho - 4\kappa + 2\eta\right)}}{\left(\xi\left(1-2\rho^2\right) - \rho(\eta - 2\kappa)\right)\left(\rho(\eta-2\kappa)+\xi\right)}\right), \end{aligned}$$

η is defined in (3.9) and $\Delta := 2\kappa\left(1 + \mathrm{e}^{\kappa t}\left(1 - 2\rho^2\right)\right) - \left(1 - \mathrm{e}^{\kappa t}\right)(\rho\xi + \eta)$. To get an idea of the t-dependence of the ATM forward volatility we set $\rho = 0$ (since Proposition 3.8 is valid for correlations near zero) and perform a Taylor expansion of $v_1^\infty(0,t)$ around $t = 0$:

$$v_1^\infty(0,t) = v_1^\infty(0,0) + \left(\frac{2\theta}{1 + \sqrt{1 + \xi^2/4\kappa^2}} - v\right)t + \mathcal{O}\left(t^2\right).$$

When $v \geq \theta$ then at this order the large τ-maturity forward ATM volatility is lower than the corresponding large τ-maturity ATM implied volatility and this difference is increasing in t and in the ratio ξ/κ. This is similar in spirit to Remark 3.4(ii) for the small-maturity Heston forward smile.

3.2. **Multivariate Heston.** The n-Heston model ($n \in \mathbb{N}$) is defined as the unique strong solution to the following SDE:

$$
\begin{aligned}
\mathrm{d}X_t &= -\frac{1}{2}\sum_{i=1}^n V_t^{(i)}\mathrm{d}t + \sum_{i=1}^n \sqrt{V_t^{(i)}}\mathrm{d}W_t^{(i)}, && X_0 = 0 \in \mathbb{R},\\
\mathrm{d}V_t^{(i)} &= \kappa_i\left(\theta_i - V_t^{(i)}\right)\mathrm{d}t - \xi_i\sqrt{V_t^{(i)}}\mathrm{d}Z_t^{(i)}, i = 1,\dots,n && V_0^{(i)} = v_0^{(i)} \in \mathbb{R}_+^*,\\
\mathrm{d}\left\langle W^{(i)}, Z^{(j)}\right\rangle_t &= \rho_i \mathbb{1}_{\{i=j\}}\mathrm{d}t, && i,j = 1,\dots,n,\\
\mathrm{d}\left\langle W^{(i)}, W^{(j)}\right\rangle_t &= \mathrm{d}\left\langle Z^{(i)}, Z^{(j)}\right\rangle_t = \mathbb{1}_{\{i=j\}}\mathrm{d}t, && i,j = 1,\dots,n,
\end{aligned}
\tag{3.15}
$$

where $\boldsymbol{\kappa} = (\kappa_1,\dots,\kappa_n) \in (0,\infty)^n$, $\boldsymbol{\rho} \in (-1,1)^n$, $\boldsymbol{\theta} \in (0,\infty)^n$, $\boldsymbol{\xi} \in (0,\infty)^n$, $\boldsymbol{v_0} \in (0,\infty)^n$. The independence assumption of the variances allows the forward mgf of the n-Heston to be written as the sum of the forward mgf's of the individual Heston models. The asymptotics of the forward smile are then given directly in terms of the results in Section 3.1. For ease of notation in this section, whenever we reference a function or variable used in the Heston analysis in Section 3.1 and use the index i, it means that function or variable defined using the i-th Heston model parameters in (3.15). So for example when we reference the function L_i in (3.4) we mean that function L evaluated using the i-th Heston model parameters in (3.15).

Proposition 3.10. *In the n-Heston model Corollary 2.8 and Proposition 2.12 hold with $\mathcal{D}_{0,0} = \bigcap_{i=1}^n \mathcal{D}_{t,\tau}^i$, $\Lambda_{0,0} = \sum_{i=1}^n \Lambda_i$ and $\Lambda_{1,0} = \sum_{i=1}^n L_i$, where $\mathcal{D}_{t,\tau}^i$, Λ_i and L_i are defined in 3.3, 3.2 and 3.4.*

Consider for instance $n = 2$, and define the functions:

$$
\begin{aligned}
\nu_0(t,\tau) :=& \frac{\tau}{48\left(v_1+v_2\right)^2}\Big[\Big(\sum_{i=1}^2 24\theta_i\kappa_i\left(v_1+v_2\right)^2 + 12v_i\left(v_1+v_2\right)^2\left(\xi_i\rho_i - 2\kappa_i\right)\\
&+ \xi_i^2\left(\rho_i^2-4\right)\left(v_1+v_2\right)^2\Big) - v_1^2\xi_2^2\left(\rho_2^2-4\right) - v_2^2\xi_1^2\left(\rho_1^2-4\right) - 2v_1v_2\left(\xi_1^2\left(5\rho_1^2-2\right)\right.\\
&\left. - 9\xi_2\xi_1\rho_1\rho_2 + \xi_2^2\left(5\rho_2^2-2\right)\right)\Big] - \frac{t}{4\left(v_1+v_2\right)}\left(\sum_{i=1}^2 4\kappa_i\left(v_1+v_2\right)\left(v_i-\theta_i\right) + \xi_i^2 v_i\right),
\end{aligned}
\tag{3.16}
$$

$$
\begin{aligned}
\nu_1(t,\tau) :=& \frac{1}{48\left(v_1+v_2\right)^3}\left(\sum_{i=1}^2 \xi_i^2\left(4-7\rho_i^2\right)v_i^2 + 2v_1v_2\left(\xi_1^2\left(4\rho_1^2+2\right) - 15\xi_2\xi_1\rho_1\rho_2 + \xi_2^2\left(4\rho_2^2+2\right)\right)\right)\\
&+ \frac{t}{4\tau\left(v_1+v_2\right)^2}\left(\sum_{i=1}^2 \xi_i^2 v_i\right).
\end{aligned}
\tag{3.17}
$$

In order to gain some intuition on the role of the Heston parameters on the forward smile we expand our solution around the ATM point in terms of the log strike k.

Corollary 3.11. *The following expansion holds for the* 2*-Heston forward smile as ε and k tend to zero:*

$$
\sigma_{\varepsilon t,\varepsilon\tau}^2(k) = v_1 + v_2 + \varepsilon\nu_0(t,\tau) + \frac{\xi_1\rho_1 v_1 + \xi_2\rho_2 v_2}{2v_1 + 2v_2}k + \nu_1(t,\tau)k^2 + \mathcal{O}(k^3) + \mathcal{O}(\varepsilon k) + \mathcal{O}(\varepsilon^2).
$$

Remark 3.12. Remarks similar to Remark 3.4(ii)-(iv) for the Heston forward smile also apply to the 2-Heston.

Proposition 3.13. *If $\rho_-^i \le \rho_i \le \min(\rho_+^i, \kappa/\xi)$ for $i = 1,2,...,n$ ($\rho_\pm^i$ defined in* (3.11)*), then Corollary 2.10 and Proposition 2.14 hold with $\Lambda_{0,0} = \sum_{i=1}^n V_i$, $\Lambda_{1,0} = \sum_{i=1}^n H_i$, $\Lambda_{2,0} = 0$ and $\mathcal{D}_{0,0} = \bigcap_{i=1}^n \left[u_-^i, u_+^i\right]$, with V_i, H_i, $u_\pm^i$ defined in* (3.12) *and* (3.8)*.*

3.3. **Schöbel-Zhu.** The Schöbel-Zhu (SZ) stochastic volatility model [48] is an extension to non-zero correlation of the Stein & Stein [49] model in which the logarithmic spot price process $(X_t)_{t\geq 0}$ satisfies the following system of SDEs:

$$\begin{aligned} \mathrm{d}X_t &= -\frac{1}{2}\sigma_t^2\mathrm{d}t + \sigma_t\mathrm{d}W_t, && X_0 = x_0 \in \mathbb{R}, \\ \mathrm{d}\sigma_t &= \kappa(\theta - \sigma_t)\,\mathrm{d}t + \frac{1}{2}\xi\mathrm{d}Z_t, && \sigma_0 = \sqrt{v} > 0, \\ \mathrm{d}\langle W, Z\rangle_t &= \rho\mathrm{d}t, \end{aligned} \tag{3.18}$$

where κ, θ and ξ are strictly positive real numbers, $\rho \in (-1,1)$ and $(W_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$ are two standard Brownian motions. The volatility process $(\sigma_t)_{t\geq 0}$ is Gaussian and hence the SDE is well-defined. The process $(X_t)_{t\geq 0}$ is simply the integrated volatility process and hence is well-defined as well. The analysis in this section is very similar to the diagonal small-maturity Heston analysis and therefore the proofs are omitted. Note that although in some cases we use the same variables as in the Heston analysis they may have a different definition in this section. We present limited results to highlight the similarities and differences between the Heston and Schöbel-Zhu forward smiles.

Proposition 3.14. *In the Schöbel-Zhu model Corollary 2.8 and Proposition 2.12 hold with $\mathcal{D}_{0,0} = \mathcal{D}_{t,\tau}$ and $\Lambda_{0,0} = \Lambda$, where $\mathcal{D}_{t,\tau}$ and Λ are defined in 3.3 and 3.2.*

Remark 3.15. At zeroth order in ε the SZ diagonal small-maturity forward smile is the same as in Heston modulo a re-scaling of the volatility of volatility. The first-order asymptotic is used in Corollary 3.16 below to highlight differences with the Heston model.

Let us now define the following functions:

$$\begin{aligned} \nu_0(t,\tau) &:= \tau\left(\frac{1}{48}\xi^2\left(\rho^2+2\right) + \kappa\theta\sqrt{v} + \frac{1}{4}v(\xi\rho - 4\kappa)\right) + 2\kappa t\sqrt{v}\left(\theta - \sqrt{v}\right), \\ \nu_1(t,\tau) &:= \frac{\rho\xi\tau\left(\xi^2\left(1-2\rho^2\right) - 8\kappa v + 2\xi\rho v\right)}{48v} + \frac{\xi^3\rho t}{8v}, \\ \nu_2(t,\tau) &:= \Big(\left(521\rho^4 - 452\rho^2 + 56\right)\xi^2 + 480\kappa\theta\sqrt{v}\left(2\rho^2-1\right) + 40\rho^2 v(\rho\xi - 4\kappa)\Big)\frac{\xi^2\tau}{7680v^2} \\ &\quad - \frac{\xi^2 t}{48v^2}\Big(\left(14\rho^2-5\right)\xi^2 + 2\kappa\theta\sqrt{v}\left(10-7\rho^2\right) + 2\kappa v\left(7\rho^2-4\right)\Big) + \frac{\xi^2t^2}{16\tau v^2}\Big(3\xi^2 + 4\kappa\sqrt{v}\left(\sqrt{v}-2\theta\right)\Big). \end{aligned} \tag{3.19}$$

In order to gain some intuition on the role of the Schöbel-Zhu parameters on the forward smile we expand our solution (to first order in ε) around the ATM point in terms of the log strike k.

Corollary 3.16. *The following expansion holds for the Schöbel-Zhu forward smile as ε and k tend to zero:*

$$\begin{aligned} \sigma^2_{\varepsilon t,\varepsilon\tau}(k) &= v + \varepsilon\nu_0(t,\tau) + \left(\frac{\xi\rho}{2} + \varepsilon\nu_1(t,\tau)\right)k + \left(\frac{(4-7\rho^2)\xi^2}{48v} + \frac{\xi^2 t}{4\tau v} + \varepsilon\nu_2(t,\tau)\right)k^2 \\ &\quad + \mathcal{O}(k^3) + \mathcal{O}(\varepsilon k^3) + \mathcal{O}(\varepsilon^2), \end{aligned} \tag{3.20}$$

where ν_0, ν_1 and ν_2 are defined in (3.19).

Remark 3.17. At this order we can make the following remarks concerning the SZ forward smile:

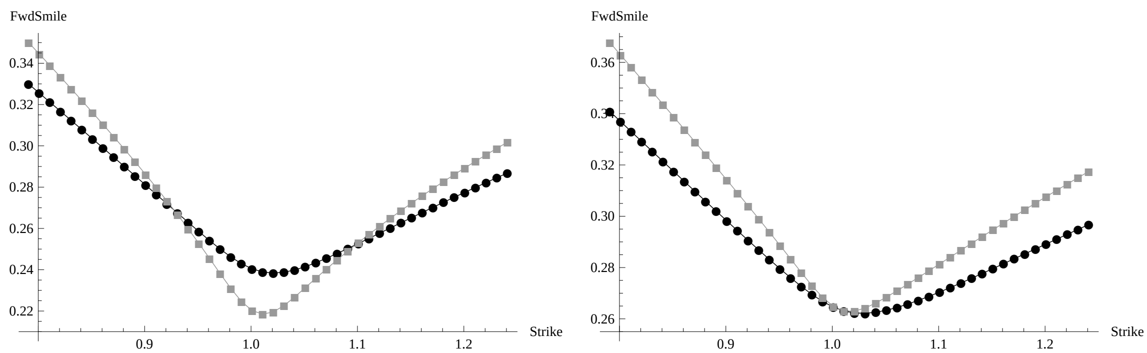
(i) Remark 3.4(iv) for the Heston forward smile also applies to the SZ forward smile.

(ii) The forward ATM volatility has a different dependence on the volatility of volatility ξ in Heston and SZ. In Heston (Remark 3.4(ii)), $\sigma_{\varepsilon t,\varepsilon\tau}(0) - \sigma_{0,\varepsilon\tau}(0)$ is decreasing in ξ. In the SZ model, Corollary 3.16 implies

$$\sigma_{\varepsilon t,\varepsilon\tau}(0) - \sigma_{0,\varepsilon\tau}(0) = (\theta - \sqrt{v})\kappa t\varepsilon + \mathcal{O}(\varepsilon^2), \qquad \text{as } \varepsilon \text{ tends to zero,}$$

which does not depend on ξ (up to an error of order $\mathcal{O}(\varepsilon^2)$). Also for realistic parameter choices (eg. $\rho \le 0$) the Heston ATM forward volatility is decreasing in ξ while in the SZ model (for example when $\xi > 2v$) it is increasing in ξ and the impact is small. This effect is illustrated in Figure 3.

Remark 3.18. An analysis analogous to that of the Heston model can be conducted for the large-maturity SZ forward smile. We shall omit it here though for brevity.



(a) Heston forward smile ξ dependence

(b) SZ forward smile ξ dependence

FIGURE 3. Here $t = 1/2$ and $\tau = 1/12$ and we apply Corollaries 3.16 and 3.3. Circles use the Heston parameters $v = 0.07, \theta = 0.07, \kappa = 1, \rho = -0.6, \xi = 0.3$ and SZ parameters $v = 0.07, \theta = \sqrt{0.07}, \kappa = 1, \rho = -0.6, \xi = 0.3$. Squares use the same parameters but with $\xi = 0.4$.

3.4. **Time-changed exponential Lévy.** It is well-known (see for example in [14, Proposition 11.2]) that the forward smile in exponential Lévy models is time-homogeneous in the sense that $\sigma_{t,\tau}$ does not depend on t for any fixed $\tau > 0$, since the process has stationary increments. This is not necessarily true in time-changed exponential Lévy models as we shall now see. Let $(Y_t)_{t\ge 0}$ be a Lévy process with mgf given by $\log \mathbb{E}\left(\mathrm{e}^{uY_t}\right) = t\phi(u)$ for all $t \ge 0$ and $u \in \mathcal{D}_\phi := \{u \in \mathbb{R} : |\phi(u)| < \infty\}$. We consider models where $(X_t)_{t\ge 0} := (Y_{V_t})_{t\ge 0}$ pathwise and the time-change is given by $V_t := \int_0^t v_s \mathrm{d}s$ with v being a strictly positive process independent of Y. We shall consider the two following examples:

$$\mathrm{d}v_t = \kappa\left(\theta - v_t\right)\mathrm{d}t + \xi\sqrt{v_t}\mathrm{d}W_t, \tag{3.21}$$

$$\mathrm{d}v_t = -\lambda v_t \mathrm{d}t + \mathrm{d}Z_t, \tag{3.22}$$

with $v_0 = v > 0$ and $\kappa, \xi, \theta, \lambda > 0$. Here W is a standard Brownian motion and Z is a compound Poisson subordinator with exponential jump size distribution and Lévy exponent $l(u) := \lambda du/(\alpha - u)$ for all $u < \alpha$ with $d > 0$ and $\alpha > 0$. In (3.21), v is a Feller diffusion and in (3.22), it is a Γ-OU process. We now define the functions $\overline{V}$ and $\overline{H}$ from $\widehat{\mathcal{D}}_\infty$ to $\mathbb{R}$ by

$$\overline{V}(u) := \frac{\kappa\theta}{\xi^2}\left(\kappa - \sqrt{\kappa^2 - 2\phi(u)\xi^2}\right) \quad \text{and} \quad \overline{H}(u) := \frac{\overline{V}(u)v\mathrm{e}^{-\kappa t}}{\kappa\theta - 2\beta_t\overline{V}(u)} - \frac{2\kappa\theta}{\xi^2}\log\left(\frac{\kappa\theta - 2\beta_t\overline{V}(u)}{\kappa\theta\left(1 - \gamma(\phi(u))\right)}\right), \tag{3.23}$$

and the functions $\widetilde{V}$ and $\widetilde{H}$ from $\widetilde{\mathcal{D}}_\infty$ to $\mathbb{R}$ by

$$\begin{aligned} \widetilde{V}(u) &:= \frac{\phi(u)\lambda d}{\alpha\lambda - \phi(u)}, \\ \widetilde{H}(u) &:= \frac{\lambda\alpha d}{\alpha\lambda - \phi(u)}\log\left(1 - \frac{\phi(u)}{\alpha\lambda}\right) + \frac{\phi(u)v\mathrm{e}^{-\lambda t}}{\lambda} + d\log\left(\frac{\phi(u) - \alpha\lambda\mathrm{e}^{\lambda t}}{\mathrm{e}^{t\lambda}(\phi(u) - \alpha\lambda)}\right), \end{aligned} \tag{3.24}$$

where we set

$$\widehat{\mathcal{D}}_\infty := \left\{u : \phi(u) \leq \kappa^2/(2\xi^2)\right\}, \qquad \widetilde{\mathcal{D}}_\infty := \left\{u : \phi(u) < \alpha\lambda\right\}, \tag{3.25}$$

ϕ is the Lévy exponent of Y and β_t and γ are defined in (5.46). The following proposition—proved in Section 5.5—is the main result of the section.

Proposition 3.19. *Suppose that ϕ is essentially smooth (Definition 2.2), strictly convex and infinitely differentiable on $\mathcal{D}_\phi^o$ with $\{0,1\} \subset \mathcal{D}_\phi^o$ and $\phi(1) = 0$. Then Corollary 2.10 and Proposition 2.14 hold:*

(i) if v follows (3.21)*, with $\Lambda_{0,0} = \overline{V}$, $\Lambda_{1,0} = \overline{H}$, $\Lambda_{2,0} = 0$ and $\mathcal{D}_{0,0} = \widehat{\mathcal{D}}_\infty$;*
(ii) if v follows (3.22)*, with $\Lambda_{0,0} = \widetilde{V}$, $\Lambda_{1,0} = \widetilde{H}$, $\Lambda_{2,0} = 0$ and $\mathcal{D}_{0,0} = \widetilde{\mathcal{D}}_\infty$;*
(iii) if $v_t \equiv 1$, with $\Lambda_{0,0} = \phi$, $\Lambda_{1,0} = 0$, $\Lambda_{2,0} = 0$ and $\mathcal{D}_{0,0} = \mathcal{D}_\phi$.

Remark 3.20.

(i) The uncorrelated Heston model (3.1) can be represented as $Y_t := -t/2+W_t$ time-changed by an integrated Feller diffusion (3.21). With $\phi(u) \equiv u(u-1)/2$ and $\mathcal{D}_\phi = \mathbb{R}$, Proposition 3.19(i) agrees with Proposition 3.8.
(ii) The zeroth order large-maturity forward smile is the same as its corresponding zeroth order large-maturity spot smile and differences only emerge at first order. It seems plausible that this will always hold if there exists a stationary distribution for v and if v is independent of the Lévy process Y;
(iii) Case (iii) in the proposition corresponds to the standard exponential Lévy case (without time-change).

We now use Proposition 3.19 to highlight the first-order differences in the large-maturity forward smile (2.17) and the corresponding spot smile. If v follows (3.21) then a Taylor expansion of v_1^∞ in (2.16) around $t = 0$ gives

$$v_1^\infty(t,k) - v_1^\infty(0,k) = \frac{8v_0^\infty(k)^2}{4k^2 - v_0^\infty(k)^2}\overline{V}(u^*(k))\left(\frac{\xi^2 v\overline{V}(u^*(k))}{2\theta^2\kappa^2} + 1 - \frac{v}{\theta}\right)t + \mathcal{O}(t^2), \quad \text{for all } k \in \mathbb{R}\backslash\{\overline{V}'(0), \overline{V}'(1)\}.$$

Using simple properties of v_0^∞ discussed below (2.16) and $\overline{V}$ we see that the large-maturity forward smile is lower than the corresponding spot smile for $k \in (\overline{V}'(0), \overline{V}'(1))$ (which always include the at-the-money) if $v \geq \theta$. The forward smile is higher than the corresponding spot smile for $k \in \mathbb{R}\backslash(\overline{V}'(0), \overline{V}'(1))$ (OTM options) if $v \leq \theta$, and these differences are increasing in ξ/κ and t. This effect is illustrated in Figure 4 and $k \in (\overline{V}'(0), \overline{V}'(1))$ corresponds to strikes in the region $(0.98, 1.02)$ in the figure.

If v follows (3.22) then a simple Taylor expansion of $v_1^\infty(\cdot,k)$ in (2.16) around $t = 0$ gives

$$v_1^\infty(t,k) - v_1^\infty(0,k) = \frac{8v_0^\infty(k)^2}{4k^2 - v_0^\infty(k)^2}\frac{\phi(u^*(k))\left[\lambda(d-\alpha v) + v\phi(u^*(k))\right]}{\alpha\lambda - \phi(u^*(k))}t + \mathcal{O}(t^2), \quad \text{for all } k \in \mathbb{R}\setminus\{\widetilde{V}'(0), \widetilde{V}'(1)\}.$$

Similarly we deduce that the large-maturity forward smile is lower than the corresponding spot smile for $k \in (\widetilde{V}'(0), \widetilde{V}'(1))$ if $v \geq d/\alpha$. The forward smile is higher than the corresponding spot smile for $k \in \mathbb{R}\backslash(\widetilde{V}'(0), \widetilde{V}'(1))$ (OTM options) if $v \leq d/\alpha$, and these differences are increasing in t.

If v follows (3.21)((3.22)) then the stationary distribution is a gamma distribution with mean θ (d/α), see [14, page 475 and page 487]. The above results seem to indicate that the differences in level between the large-maturity forward smile and the corresponding spot smile depend on the relative values of v_0 and the mean of the stationary distribution of the process v. This is also similar to Remark 3.4(ii) and the analysis below Remark 3.9 for the Heston forward smile. These observations are also independent of the choice of ϕ indicating that the fundamental quantity driving the non-stationarity of the large-maturity forward smile over the corresponding spot implied volatility smile is the choice of time-change.

As an example of a Lévy process satisfying the assumptions of Proposition 3.19, consider the Variance-Gamma model, defined by

$$\phi(u) = \mu u + C \log\left(\frac{GM}{(M-u)(G+u)}\right), \qquad \text{for all } u \in (-G, M),$$

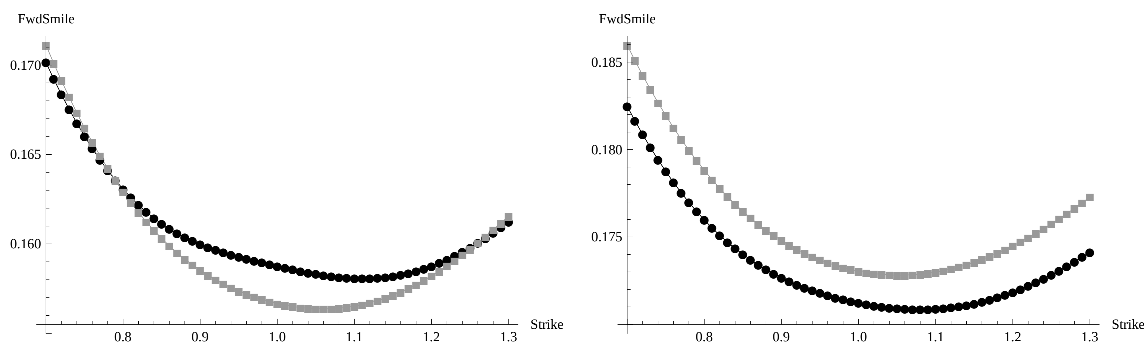
with $C > 0$, $G > 0$, $M > 1$ and $\mu := -C \log\left(\frac{GM}{(M-1)(G+1)}\right)$ ensures that $(\mathrm{e}^{X_t})_{t\geq 0}$ is a true martingale ($\phi(1) = 0$). We immediately obtain

$$\phi'(u) = \frac{C(G - M + 2u) + \mu(G+u)(M-u)}{(G+u)(M-u)} \qquad \text{and} \qquad \phi''(u) = \frac{C\left((G+M)^2 + (2u + G - M)^2\right)}{2(G+u)^2(M-u)^2},$$

for $u \in (-G, M)$, so that ϕ is essentially smooth and strictly convex on $(-G, M)$. It is also infinitely differentiable on $(-G, M)$ with $\{0, 1\} \subset (-G, M)$ and Proposition 3.19 applies. For Proposition 3.19(iii) we can compute $u^* : \mathbb{R} \to (-G, M)$ through (2.5) in closed-form. The solutions to $\phi'(u^*(k)) = k$ are $u^*(\mu) = (M - G)/2$ and

$$u^*_{\pm}(k) = \frac{-2C - (G - M)(k - \mu) \pm \sqrt{4C^2 + (G + M)^2(k - \mu)^2}}{2(k - \mu)} \quad \text{for all } k \neq \mu.$$

The sign condition $(M - u)(G + u) > 0$ imposes $-2C \pm \sqrt{4C^2 + (G + M)^2(k - \mu)^2} > 0$ for all $k \neq \mu$. Hence u^*_+ (continuous on the whole real line) is the only valid solution and the rate function is then given in closed-form as $\Lambda^*(k) = ku^*_+(k) - \phi(u^*_+(k))$ for all real k.

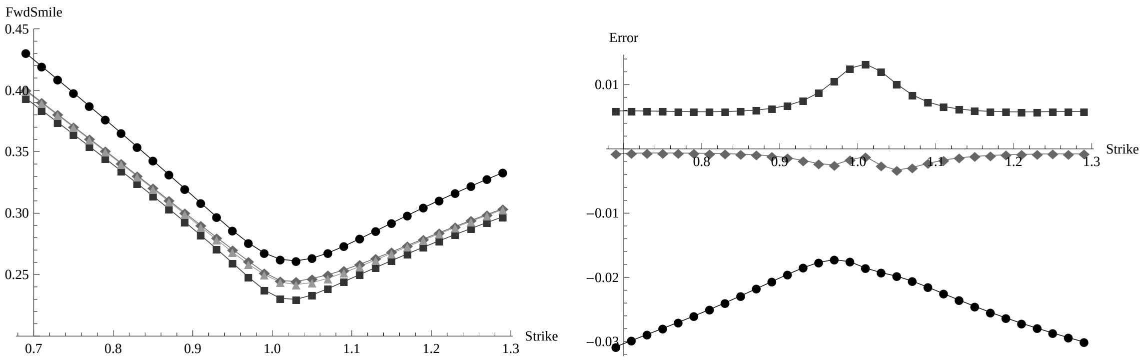


(a) Feller time-change: forward smile vs spot smile $v > \theta$. (b) Feller time-change: forward smile vs spot smile $v < \theta$.

FIGURE 4. Circles represent $t = 0$ and $\tau = 2$ and squares represent $t = 1/2$ and $\tau = 2$ using a Variance-Gamma model time-changed by a Feller diffusion and the asymptotic in Proposition 3.19. In (a) the parameters are $C = 58.12$, $G = 50.5$, $M = 69.37$, $\kappa = 1.23$, $\theta = 0.9$, $\xi = 1.6$, $v = 1$ and (b) uses the same parameters but with $\theta = 1.1$.

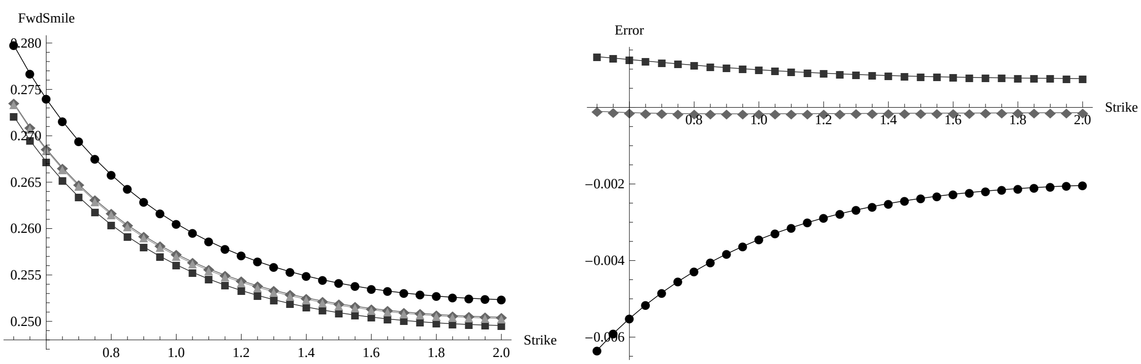
4. NUMERICS

We compare here the true forward smile in various models and the asymptotics developed in Propositions 2.12 and 2.14. We calculate forward-start option prices using the inverse Fourier transform representation in [41, Theorem 5.1] and a global adaptive Gauss-Kronrod quadrature scheme. We then compute the forward smile $\sigma_{t,\tau}$ and compare it to the zeroth, first and second order asymptotics given in Propositions 2.12 and 2.14 for various models. In Figure 5 we compare the Heston diagonal small-maturity asymptotic in Proposition 3.2 with the true forward smile. Figure 6 tests the accuracy of the Heston large-maturity asymptotic from Proposition 3.8. In order to use this proposition we require $\rho_- \leq \rho \leq \min(\rho_+, \kappa/\xi)$ with $\rho_\pm$ defined in (3.11). For the parameter choice in the figure we have $\rho_- = -0.65$ and the condition is satisfied. Finally in Figure 7 we consider the Variance Gamma model time-changed by a Γ-OU process using Proposition 3.19. Results are in line with expectations and the higher the order of the asymptotic the closer we match the true forward smile.



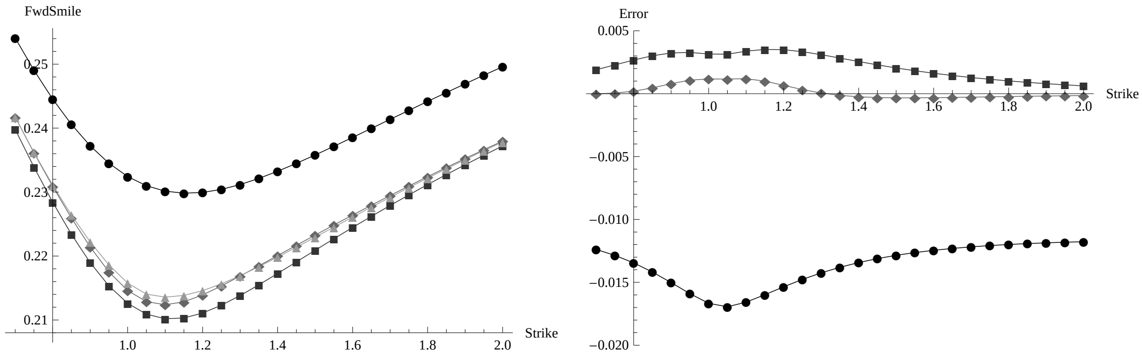
(a) Heston diagonal small-maturity vs Fourier inversion. (b) Errors

Figure 5. In (a) circles, squares and diamonds represent the zeroth, first and second order asymptotics respectively in Proposition 3.2 and triangles represent the true forward smile using Fourier inversion. In (b) we plot the differences between the true forward smile and the asymptotic. We use $t = 1/2$ and $\tau = 1/12$ and the Heston parameters $v = 0.07$, $\theta = 0.07$, $\kappa = 1$, $\xi = 0.34$, $\rho = -0.8$.



(a) Heston Large-Maturity vs Fourier Inversion. (b) Errors

Figure 6. In (a) circles, squares and diamonds represent the zeroth, first and second order asymptotics respectively in Proposition 3.8 and triangles represent the true forward smile using Fourier inversion. In (b) we plot the differences between the true forward smile and the asymptotic. We use $t = 1$ and $\tau = 5$ and the Heston parameters $v = 0.07$, $\theta = 0.07$, $\kappa = 1.5$, $\xi = 0.34$, $\rho = -0.25$.



(a) Γ-OU time-change large-maturity vs Fourier inversion.

(b) Errors

FIGURE 7. In (a) circles, squares and diamonds represent the zeroth, first and second order asymptotics respectively in proposition 3.19 and triangles represent the true forward smile using Fourier inversion for a variance gamma model time-changed by a Γ-OU process. In (b) we plot the differences between the true forward smile and the asymptotic. We use $t = 1$ and $\tau = 3$ with the parameters $C = 6.5$, $G = 11.1$, $M = 33.4$, $v = 1$, $\alpha = 0.6$, $d = 0.6$, $\lambda = 1.8$.

5. PROOFS

5.1. **Proofs of Section 2.** We define a change of measure by

$$\frac{\mathrm{d}\mathbb{Q}_{k,\varepsilon}}{\mathrm{d}\mathbb{P}} = \exp\left(\frac{u^*(k)X_\varepsilon}{\varepsilon} - \frac{\Lambda_\varepsilon\left(u^*(k)\right)}{\varepsilon}\right), \tag{5.1}$$

with $u^*(k)$ defined in (2.5). By Lemma 2.4(i), $u^*(k) \in \mathcal{D}_{0,0}^o$ for all $k \in \mathbb{R}$. Since $\mathcal{D}_{0,0}^o \subseteq \mathcal{D}_\varepsilon^0$ this means that $\Lambda_\varepsilon\left(u^*(k)\right) < \infty$. Also $\mathrm{d}\mathbb{Q}_{k,\varepsilon}/\mathrm{d}\mathbb{P}$ is almost surely strictly positive and by the very definition $\mathbb{E}\left(\mathrm{d}\mathbb{Q}_k/\mathrm{d}\mathbb{P}\right) = 1$. Therefore (5.1) is a valid measure change for all $k \in \mathbb{R}$. We define the random variable

$$Z_{k,\varepsilon} := \left(X_\varepsilon - k\right)/\sqrt{\varepsilon} \tag{5.2}$$

and set the characteristic function $\Phi_{Z_{k,\varepsilon}} : \mathbb{R} \to \mathbb{C}$ of $Z_{k,\varepsilon}$ in the $\mathbb{Q}_{k,\varepsilon}$-measure as follows

$$\Phi_{Z_{k,\varepsilon}}(u) = \mathbb{E}^{\mathbb{Q}_{k,\varepsilon}}\left(\mathrm{e}^{\mathtt{i}uZ_{k,\varepsilon}}\right). \tag{5.3}$$

Recall from Section 2 that for ease of exposition $\Lambda_{i,l} := \Lambda_{i,l}\left(u^*(k)\right)$ with $\Lambda_{i,l}$ defined in (2.4). This notation will be used throughout the section. We now have the following important technical lemma.

Lemma 5.1. *The following expansion holds for the characteristic function $\Phi_{Z_{k,\varepsilon}}$ defined in (5.3) as $\varepsilon \to 0$:*

$$\log \Phi_{Z_{k,\varepsilon}}(u) = -\frac{1}{2}\Lambda_{0,2}u^2 + \eta_1(u)\sqrt{\varepsilon} + \eta_2(u)\varepsilon + \eta_3(u)\varepsilon^{3/2} + \mathcal{O}\left(\varepsilon^2\right),$$

where the functions η_i, $i = 1, 2, 3$ are defined in (5.4).

Remark 5.2. By Lévy's Convergence Theorem [52, Page 185, Theorem 18.1], $Z_{k,\varepsilon}$ defined in (5.2) converges weakly to a normal random variable with mean 0 and variance $\Lambda_{0,2}$ in the $\mathbb{Q}_{k,\varepsilon}$-measure as ε tends to zero.

Proof. By Lemma 2.4, $u^*(k) \in \mathcal{D}^o_{0,0}$, therefore using the definition of the $\mathbb{Q}_{k,\varepsilon}$-measure in (5.1) we have

$$\begin{aligned}
\log \Phi_{Z_{k,\varepsilon}}(u) &= \log \mathbb{E}^{\mathbb{P}}\left(\frac{\mathrm{d}\mathbb{Q}_{k,\varepsilon}}{\mathrm{d}\mathbb{P}} \mathrm{e}^{\mathtt{i}uZ_{k,\varepsilon}}\right) = \log \mathbb{E}^{\mathbb{P}}\left[\exp\left(\frac{u^*(k)X_\varepsilon}{\varepsilon} - \frac{\Lambda_\varepsilon(u^*(k))}{\varepsilon}\right)\exp\left(\mathtt{i}u\sqrt{\varepsilon}\left(\frac{X_\varepsilon}{\varepsilon}\right) - \frac{\mathtt{i}ku}{\sqrt{\varepsilon}}\right)\right] \\
&= -\frac{1}{\varepsilon}\Lambda_\varepsilon\left(u^*(k)\right) - \frac{\mathtt{i}uk}{\sqrt{\varepsilon}} + \log \mathbb{E}^{\mathbb{P}}\left[\exp\left(\left(\frac{X_\varepsilon}{\varepsilon}\right)\left(\mathtt{i}u\sqrt{\varepsilon} + u^*(k)\right)\right)\right] \\
&= -\frac{\mathtt{i}uk}{\sqrt{\varepsilon}} + \frac{1}{\varepsilon}\left(\Lambda_\varepsilon\left(\mathtt{i}u\sqrt{\varepsilon} + u^*(k)\right) - \Lambda_\varepsilon\left(u^*(k)\right)\right) \\
&= -\frac{\mathtt{i}uk}{\sqrt{\varepsilon}} + \frac{1}{\varepsilon}\sum_{n=1}^{\infty}\frac{1}{n!}\left(\Lambda_{0,n} + \varepsilon\Lambda_{1,n} + \varepsilon^2\Lambda_{2,n} + \mathcal{O}\left(\varepsilon^3\right)\right)\left(\mathtt{i}u\sqrt{\varepsilon}\right)^n,
\end{aligned}$$

where the last equality holds by the differentiation and expansion properties in Assumption 2.1. We now write

$$\begin{aligned}
\log \Phi_{Z_{k,\varepsilon}}(u) &= -\frac{\mathtt{i}uk}{\sqrt{\varepsilon}} + \frac{\mathtt{i}\Lambda_{0,1}u}{\sqrt{\varepsilon}} - \frac{1}{2}\Lambda_{0,2}u^2 + \frac{1}{\varepsilon}\left(\sum_{n=3}^{\infty}\frac{1}{n!}\Lambda_{0,n}\left(\mathtt{i}u\sqrt{\varepsilon}\right)^n + \sum_{n=1}^{\infty}\frac{1}{n!}\left(\varepsilon\Lambda_{1,n} + \varepsilon^2\Lambda_{2,n} + \mathcal{O}\left(\varepsilon^3\right)\right)\left(\mathtt{i}u\sqrt{\varepsilon}\right)^n\right) \\
&= -\frac{1}{2}\Lambda_{0,2}u^2 + \sum_{n=1}^{\infty}\frac{1}{n!}\left(\left(\frac{n!}{(n+2)!}\Lambda_{0,n+2}\left(\mathtt{i}u\right)^2 + \Lambda_{1,n}\right) + \varepsilon\Lambda_{2,n} + \mathcal{O}\left(\varepsilon^2\right)\right)\left(\mathtt{i}u\sqrt{\varepsilon}\right)^n \\
&= -\frac{1}{2}\Lambda_{0,2}u^2 + \eta_1(u)\sqrt{\varepsilon} + \eta_2(u)\varepsilon + \eta_3(u)\varepsilon^{3/2} + \mathcal{O}\left(\varepsilon^2\right),
\end{aligned}$$

where we used (2.5) from Lemma 2.4 and where

$$\begin{aligned}
\eta_1(u) &:= \mathtt{i}u\Lambda_{1,1} - \frac{1}{6}\mathtt{i}u^3\Lambda_{0,3}, \\
\eta_2(u) &:= -\frac{1}{2}u^2\Lambda_{1,2} + \frac{1}{24}u^4\Lambda_{0,4}, \\
\eta_3(u) &:= \mathtt{i}u\Lambda_{2,1} - \frac{1}{6}\mathtt{i}u^3\Lambda_{1,3} + \frac{1}{120}\mathtt{i}u^5\Lambda_{0,5}.
\end{aligned} \tag{5.4}$$

□

Proof of Theorem 2.6. For $j = 1, 2, 3$, let us define the functions $g_j : \mathbb{R}^2_+ \to \mathbb{R}_+$ by

$$g_j(x, y) := \begin{cases} (x-y)^+, & \text{if } j = 1, \\ (y-x)^+, & \text{if } j = 2, \\ \min(x, y), & \text{if } j = 3. \end{cases}$$

Using the definition of the $\mathbb{Q}_{k,\varepsilon}$-measure in (5.1) we can re-write the option price as

$$\begin{aligned}
\mathbb{E}\left[g_j\left(\mathrm{e}^{X_\varepsilon f(\varepsilon)}, \mathrm{e}^{kf(\varepsilon)}\right)\right] &= \mathrm{e}^{\frac{1}{\varepsilon}\Lambda_\varepsilon(u^*(k))}\mathbb{E}^{\mathbb{Q}_{k,\varepsilon}}\left[\mathrm{e}^{-\frac{u^*(k)}{\varepsilon}X_\varepsilon}g_j\left(\mathrm{e}^{X_\varepsilon f(\varepsilon)}, \mathrm{e}^{kf(\varepsilon)}\right)\right] \\
&= \mathrm{e}^{-\frac{1}{\varepsilon}[ku^*(k) - \Lambda_\varepsilon(u^*(k))]}\mathbb{E}^{\mathbb{Q}_{k,\varepsilon}}\left[\mathrm{e}^{-\frac{u^*(k)}{\varepsilon}(X_\varepsilon - k)}g_j\left(\mathrm{e}^{X_\varepsilon f(\varepsilon)}, \mathrm{e}^{kf(\varepsilon)}\right)\right].
\end{aligned}$$

By the rescaled mgf expansion in Assumption 2.1 and Equality (2.6) we immediately have

$$\exp\left(-\frac{1}{\varepsilon}\left(ku^*(k) - \Lambda_\varepsilon\left(u^*(k)\right)\right)\right) = \exp\left(-\frac{1}{\varepsilon}\Lambda^*\left(k\right) + \Lambda_{1,0} + \Lambda_{2,0}\,\varepsilon + \mathcal{O}\left(\varepsilon^2\right)\right). \tag{5.5}$$

From the definition of the random variable $Z_{k,\varepsilon}$ in (5.2) we have

$$\mathbb{E}^{\mathbb{Q}_{k,\varepsilon}}\left[\mathrm{e}^{-\frac{u^*(k)}{\varepsilon}(X_\varepsilon - k)}g_j\left(\mathrm{e}^{X_\varepsilon f(\varepsilon)}, \mathrm{e}^{kf(\varepsilon)}\right)\right] = \mathrm{e}^{kf(\varepsilon)}\mathbb{E}^{\mathbb{Q}_{k,\varepsilon}}\left[\tilde{g}_j(Z_{k,\varepsilon})\right],$$

where for $j = 1, 2, 3$, we define the modified payoff functions $\tilde{g}_j : \mathbb{R} \to \mathbb{R}_+$ by

$$\tilde{g}_j(z) := \mathrm{e}^{-u^*(k)z/\sqrt{\varepsilon}}g_j\left(\mathrm{e}^{z\sqrt{\varepsilon}f(\varepsilon)}, 1\right).$$

For a function $f \in L^2(\mathbb{R})$ denote its Fourier transform by $(\mathcal{F}f)(u) := \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}uz} f(z)\mathrm{d}z$, for any $u \in \mathbb{R}$. Assuming (for now) that $\tilde{g}_j \in L^2(\mathbb{R})$, we have for any $u \in \mathbb{R}$,

$$(\mathcal{F}\tilde{g}_j)(u) = \int_{-\infty}^{\infty} \exp\left(-\frac{u^*(k)z}{\sqrt{\varepsilon}}\right) g_j\left(\mathrm{e}^{z\sqrt{\varepsilon}f(\varepsilon)}, 1\right) \mathrm{e}^{\mathrm{i}uz}\mathrm{d}z, \quad \text{for } j = 1, 2, 3.$$

For $j = 1$ we can write

$$\begin{aligned}
\int_{-\infty}^{\infty} \tilde{g}_1(z)\mathrm{e}^{\mathrm{i}uz}\mathrm{d}z &= \left[\frac{\exp\left(z\left(\sqrt{\varepsilon}f(\varepsilon) - u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u\right)\right)}{\sqrt{\varepsilon}f(\varepsilon) - u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u}\right]_0^{\infty} - \left[\frac{\exp\left(z\left(-u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u\right)\right)}{-u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u}\right]_0^{\infty} \\
&= \frac{\varepsilon^{3/2}f(\varepsilon)}{(u^*(k) - \mathrm{i}u\sqrt{\varepsilon})(u^*(k) - \varepsilon f(\varepsilon) - \mathrm{i}u\sqrt{\varepsilon})},
\end{aligned}$$

which is valid for $u^*(k) > \varepsilon f(\varepsilon)$. For ε sufficiently small and by the definition of f in (2.7) this holds for $u^*(k) > c$. For $j = 2$ we can write

$$\begin{aligned}
\int_{-\infty}^{\infty} \tilde{g}_2(z)\mathrm{e}^{\mathrm{i}uz}\mathrm{d}z &= \left[\frac{\exp\left(z\left(-u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u\right)\right)}{-u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u}\right]_{-\infty}^{0} - \left[\frac{\exp\left(z\left(\sqrt{\varepsilon}f(\varepsilon) - u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u\right)\right)}{\sqrt{\varepsilon}f(\varepsilon) - u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u}\right]_{-\infty}^{0} \\
&= \frac{\varepsilon^{3/2}f(\varepsilon)}{(u^*(k) - \mathrm{i}u\sqrt{\varepsilon})(u^*(k) - \varepsilon f(\varepsilon) - \mathrm{i}u\sqrt{\varepsilon})},
\end{aligned}$$

which is valid for $u^*(k) < 0$ as ε tends to zero. Finally, for $j = 3$ we have

$$\begin{aligned}
\int_{-\infty}^{\infty} \tilde{g}_3(z)\mathrm{e}^{\mathrm{i}uz}\mathrm{d}z &= \int_{-\infty}^{0} \mathrm{e}^{-\frac{u^*(k)}{\sqrt{\varepsilon}}z} g_3\left(\mathrm{e}^{z\sqrt{\varepsilon}f(\varepsilon)}, 1\right)\mathrm{e}^{\mathrm{i}uz}\mathrm{d}z + \int_{0}^{\infty} \mathrm{e}^{-\frac{u^*(k)}{\sqrt{\varepsilon}}z} g_3\left(\mathrm{e}^{z\sqrt{\varepsilon}f(\varepsilon)}, 1\right)\mathrm{e}^{\mathrm{i}uz}\mathrm{d}z \\
&= \left[\frac{\exp\left(z\left(\sqrt{\varepsilon}f(\varepsilon) - u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u\right)\right)}{\sqrt{\varepsilon}f(\varepsilon) - u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u}\right]_{-\infty}^{0} + \left[\frac{\exp\left(z\left(-u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u\right)\right)}{-u^*(k)/\sqrt{\varepsilon} + \mathrm{i}u}\right]_0^{\infty} \\
&= -\frac{\varepsilon^{3/2}f(\varepsilon)}{(u^*(k) - \mathrm{i}u\sqrt{\varepsilon})(u^*(k) - \varepsilon f(\varepsilon) - \mathrm{i}u\sqrt{\varepsilon})},
\end{aligned}$$

which is valid for $0 < u^*(k) < \varepsilon f(\varepsilon)$. For ε sufficiently small and by the assumption on f in (2.7) this is true for $0 < u^*(k) < c$. In this context $u^*(k)$ comes out naturally in the analysis as a classical dampening factor. Note that in order for these strips of regularity to exist we require that $\{0, c\} \subset \mathcal{D}_{0,0}^o$, as assumed in the theorem. By the differentiability property in Assumption 2.1 and the strict convexity and essential smoothness property in Assumption 2.3 we have

$$\begin{aligned}
0 < u^*(k) < c \quad &\text{if and only if} \quad \Lambda_{0,1}(0) < k < \Lambda_{0,1}(c), \\
u^*(k) < 0 \quad &\text{if and only if} \quad k < \Lambda_{0,1}(0), \qquad (5.6)\\
u^*(k) > c \quad &\text{if and only if} \quad k > \Lambda_{0,1}(c).
\end{aligned}$$

For ε sufficiently small and the strips of regularity defined above, the modified payoffs $\tilde{g}_j$ are in $L^2(\mathbb{R})$. By Remark 5.2, $Z_{k,\varepsilon}$ converges weakly to a zero-mean Gaussian random variable as ε tends to zero and the Gaussian density is in $L^2(\mathbb{R})$. For ε sufficiently small we can apply Parseval's Theorem [29, Theorem 13E] to write

$$
\mathbb{E}^{\mathbb{Q}_{k,\varepsilon}}\left[\tilde{g}_j(Z_{k,\varepsilon})\right] = \begin{cases} \dfrac{1}{2\pi}\displaystyle\int_{\mathbb{R}} \dfrac{\varepsilon^{3/2} f(\varepsilon)\Phi_{Z_{k,\varepsilon}}(u)\mathrm{d}u}{(u^*(k)+\mathtt{i}u\sqrt{\varepsilon})\,(u^*(k)-\varepsilon f(\varepsilon)+\mathtt{i}u\sqrt{\varepsilon})}, & \text{if } j=1, u^*(k)>c, \\ \dfrac{1}{2\pi}\displaystyle\int_{\mathbb{R}} \dfrac{\varepsilon^{3/2} f(\varepsilon)\Phi_{Z_{k,\varepsilon}}(u)\mathrm{d}u}{(u^*(k)+\mathtt{i}u\sqrt{\varepsilon})\,(u^*(k)-\varepsilon f(\varepsilon)+\mathtt{i}u\sqrt{\varepsilon})}, & \text{if } j=2, u^*(k)<0, \\ -\dfrac{1}{2\pi}\displaystyle\int_{\mathbb{R}} \dfrac{\varepsilon^{3/2} f(\varepsilon)\Phi_{Z_{k,\varepsilon}}(u)\mathrm{d}u}{(u^*(k)+\mathtt{i}u\sqrt{\varepsilon})\,(u^*(k)-\varepsilon f(\varepsilon)+\mathtt{i}u\sqrt{\varepsilon})}, & \text{if } j=3, 0<u^*(k)<c, \end{cases} \tag{5.7}
$$

where we have used that

$$
\overline{\frac{\varepsilon^{3/2} f(\varepsilon)}{(u^*(k)-\mathtt{i}u\sqrt{\varepsilon})\,(u^*(k)-\varepsilon f(\varepsilon)-\mathtt{i}u\sqrt{\varepsilon})}} = \frac{\varepsilon^{3/2} f(\varepsilon)}{(u^*(k)+\mathtt{i}u\sqrt{\varepsilon})\,(u^*(k)-\varepsilon f(\varepsilon)+\mathtt{i}u\sqrt{\varepsilon})},
$$

with $\overline{a}$ denoting the complex conjugate for $a \in \mathbb{C}$. Now using Lemma 5.1 we write

$$
\int_{\mathbb{R}} \frac{\varepsilon^{3/2} f(\varepsilon)\Phi(u)\mathrm{d}u}{(u^*(k)+\mathtt{i}u\sqrt{\varepsilon})\,(u^*(k)-\varepsilon f(\varepsilon)+\mathtt{i}u\sqrt{\varepsilon})} = \int_{\mathbb{R}} \exp\left(-\frac{\Lambda_{0,2}u^2}{2}\right) H(\varepsilon,u)\,\mathrm{d}u, \tag{5.8}
$$

where the function $H : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{C}$ is defined as

$$
H(\varepsilon,u) := \frac{\exp\left(\eta_1(u)\sqrt{\varepsilon}+\eta_2(u)\varepsilon+\eta_3(u)\varepsilon^{3/2}+\mathcal{O}\left(\varepsilon^2\right)\right)}{(u^*(k)+\mathtt{i}u\sqrt{\varepsilon})\,(u^*(k)-\varepsilon f(\varepsilon)+\mathtt{i}u\sqrt{\varepsilon})} f(\varepsilon)\varepsilon^{3/2},
$$

with η_i $(i=1,2,3)$ defined in (5.4). A Taylor expansion of H around $\varepsilon=0$ for $c=0$ yields

$$
\begin{aligned}
H(\varepsilon,u) &= \frac{f(\varepsilon)\varepsilon^{3/2}}{u^*(k)^2}\Big(1+h_1(u,0)\sqrt{\varepsilon}+h_2(u,0)\varepsilon+h_3(u,0)\varepsilon^{3/2}+\frac{f(\varepsilon)\varepsilon}{u^*(k)}-\frac{3\mathtt{i}u\varepsilon^{3/2}f(\varepsilon)}{u^*(k)}+\mathcal{O}\left(\varepsilon^2\right)\Big) \\
&\quad \left(1+\eta_1(u)\sqrt{\varepsilon}+\left(\frac{1}{2}\eta_1(u)^2+\eta_2(u)\right)\varepsilon+\left(\frac{1}{6}\eta_1(u)^3+\eta_2(u)\eta_1(u)+\eta_3(u)\right)\varepsilon^{3/2}+\mathcal{O}\left(\varepsilon^2\right)\right) \\
&= \frac{f(\varepsilon)\varepsilon^{3/2}}{u^*(k)^2}\left(1+\bar{h}_1(u,0)\sqrt{\varepsilon}+\bar{h}_2(u,0)\varepsilon+\bar{h}_3(u,0)\varepsilon^{3/2}+\frac{\varepsilon f(\varepsilon)}{u^*(k)}+\left(\frac{\eta_1(u)}{u^*(k)}-\frac{3\mathtt{i}u}{u^*(k)^2}\right)\varepsilon^{3/2}f(\varepsilon)+\mathcal{O}\left(\varepsilon^2\right)\right),
\end{aligned}
$$

where we define the following functions:

(5.9)

$$
\begin{aligned}
h_1(u,c) &:= \frac{\mathtt{i}u}{u^*(k)-c}\left(\frac{c}{u^*(k)}-2\right), \qquad h_2(u,c) := -\frac{u^2\left(c^2-3cu^*(k)+3u^*(k)^2\right)}{u^*(k)^2\left(u^*(k)-c\right)^2}, \\
h_3(u,c) &:= \frac{\mathtt{i}u^3\left(4u^*(k)^3-c^3+4c^2u^*(k)-6cu^*(k)^2\right)}{u^*(k)^3\left(u^*(k)-c\right)^3}, \qquad \bar{h}_1(u,c) := \eta_1(u)+\frac{\mathtt{i}u\left(c-2u^*(k)\right)}{u^*(k)\left(u^*(k)-c\right)}, \\
\bar{h}_2(u,c) &:= \frac{\eta_1^2(u)}{2}+\eta_2(u)-\frac{u^2\left(c^2-3cu^*(k)+3\left(u^*(k)\right)^2\right)}{u^*(k)^2\left(u^*(k)-c\right)^2}+\frac{\mathtt{i}u\eta_1(u)}{u^*(k)-c}\left(\frac{c}{u^*(k)}-2\right) \\
\bar{h}_3(u,c) &:= \frac{\eta_1^3(u)}{6}+\eta_1(u)\eta_2(u)+\eta_3(u)+\frac{\mathtt{i}u^3}{u^*(k)^3}-\frac{u^2\eta_1(u)}{u^*(k)^2}+\frac{\mathtt{i}u^3}{\left(u^*(k)-c\right)^3} \\
&\quad -\frac{\mathtt{i}u}{u^*(k)-c}\left(\frac{\eta_1^2(u)}{2}+\eta_2(u)-\frac{\mathtt{i}u\eta_1(u)}{u^*(k)}-\frac{u^2}{u^*(k)^2}\right)+\frac{\mathtt{i}u^3-u^2u^*(k)\eta_1(u)}{u^*(k)\left(u^*(k)-c\right)^2}-\frac{\mathtt{i}u}{u^*(k)}\left(\frac{\eta_1^2(u)}{2}+\eta_2(u)\right),
\end{aligned}
$$

with the η_i for $i = 1, 2, 3$, defined in (5.4). A Taylor expansion of H around $\varepsilon = 0$ for $c > 0$ yields

$$H(\varepsilon, u) = \frac{c\sqrt{\varepsilon}}{u^*(k)\left(u^*(k) - c\right)}\left\{1 + \eta_1(u)\sqrt{\varepsilon} + \left(\frac{\eta_1^2(u)}{2} + \eta_2(u)\right)\varepsilon + \left(\frac{\eta_1^3(u)}{6} + \eta_2(u)\eta_1(u) + \eta_3(u)\right)\varepsilon^{3/2} + \mathcal{O}\left(\varepsilon^2\right)\right\}$$
$$\left\{1 + h_1(u,c)\sqrt{\varepsilon} + h_2(u,c)\varepsilon + h_3(u,c)\varepsilon^{3/2} + \frac{u^*(k)(\varepsilon f(\varepsilon) - c)}{c\left(u^*(k) - c\right)} - \frac{2\mathtt{i}u\sqrt{\varepsilon}(\varepsilon f(\varepsilon) - c)u^*(k)}{c\left(u^*(k) - c\right)^2} + \mathcal{O}\left(\varepsilon^2\right)\right\}$$
$$= \frac{c\sqrt{\varepsilon}}{u^*(k)\left(u^*(k) - c\right)}\left\{1 + \bar{h}_1(u,c)\sqrt{\varepsilon} + \bar{h}_2(u,c)\varepsilon + \bar{h}_3(u,c)\varepsilon^{3/2} + \frac{u^*(k)(\varepsilon f(\varepsilon) - c)}{c\left(u^*(k) - c\right)}\right.$$
$$\left. + \frac{u^*(k)\sqrt{\varepsilon}(\varepsilon f(\varepsilon) - c)}{c\left(u^*(k) - c\right)}\left(\eta_1(u) - \frac{2\mathtt{i}u}{u^*(k) - c}\right) + \mathcal{O}\left(\varepsilon^2\right)\right\},$$

with the h_i, $\bar{h}_i$ defined in (5.9) and the η_i defined in (5.4). We will shortly be integrating H against a zero-mean Gaussian characteristic function over $\mathbb{R}$ and as such all odd powers of u will have a null contribution. In particular we note that the polynomials

$$\bar{h}_1, \quad \bar{h}_3, \quad \left(\frac{\eta_1(u)}{u^*(k)} - \frac{3\mathtt{i}u}{\left(u^*(k)\right)^2}\right)\varepsilon^{3/2} f(\varepsilon) \quad \text{and} \quad \frac{u^*(k)\sqrt{\varepsilon}(\varepsilon f(\varepsilon) - c)}{c\left(u^*(k) - c\right)}\left(\eta_1(u) - \frac{2\mathtt{i}u}{u^*(k) - c}\right)$$

are odd functions of u and hence have zero contribution. The major quantity is $\bar{h}_2$, which we can rewrite as

$$\bar{h}_2(u,c) = \bar{h}_{2,1}(c)u^2 + \bar{h}_{2,2}(c)u^4 - \frac{\Lambda_{0,3}^2 u^6}{72}, \tag{5.10}$$

where

$$\bar{h}_{2,1}(c) := \frac{2\Lambda_{1,1}}{u^*(k) - c} - \frac{c\Lambda_{1,1}}{u^*(k)\left(u^*(k) - c\right)} - \frac{\Lambda_{1,1}^2 + \Lambda_{1,2}}{2} - \frac{c^2 - 3cu^*(k) + 3u^*(k)^2}{u^*(k)^2\left(u^*(k) - c\right)^2},$$
$$\bar{h}_{2,2}(c) := \frac{c\Lambda_{0,3}}{6u^*(k)\left(u^*(k) - c\right)} - \frac{\Lambda_{0,3}}{3\left(u^*(k) - c\right)} + \frac{\Lambda_{1,1}\Lambda_{0,3}}{6} + \frac{\Lambda_{0,4}}{24}.$$

Let

$$\phi_\varepsilon(c) \equiv \frac{c\sqrt{\varepsilon}\mathbb{1}_{\{c>0\}} + \varepsilon^{3/2} f(\varepsilon)\mathbb{1}_{\{c=0\}}}{u^*(k)\left(u^*(k) - c\right)}.$$

Using simple properties of moments of a Gaussian random variable we finally compute the following

$$\int_{\mathbb{R}} \exp\left(-\frac{\Lambda_{0,2}u^2}{2}\right) H\left(\varepsilon, u\right) \mathrm{d}u$$
$$= \phi_\varepsilon(c)\int_{\mathbb{R}} \mathrm{e}^{-\frac{1}{2}\Lambda_{0,2}u^2}\left(1 + \bar{h}_{2,1}(c)u^2 + \bar{h}_{2,2}(c)u^4 - \frac{\Lambda_{0,3}^2 u^6}{72} + \frac{u^*(k)(\varepsilon f(\varepsilon) - c)}{c\left(u^*(k) - c\right)}\mathbb{1}_{\{c>0\}} + \frac{\varepsilon f(\varepsilon)}{u^*(k)}\mathbb{1}_{\{c=0\}} + \mathcal{O}\left(\varepsilon^2\right)\right)\mathrm{d}u$$
$$= \phi_\varepsilon(c)\sqrt{\frac{2\pi}{\Lambda_{0,2}}}\left(1 + \frac{\bar{h}_{2,1}(c)}{\Lambda_{0,2}} + \frac{3\bar{h}_{2,2}(c)}{\Lambda_{0,2}^2} - \frac{5\Lambda_{0,3}^2}{24\Lambda_{0,2}^3} + \frac{u^*(k)(\varepsilon f(\varepsilon) - c)}{c\left(u^*(k) - c\right)}\mathbb{1}_{\{c>0\}} + \frac{\varepsilon f(\varepsilon)}{u^*(k)}\mathbb{1}_{\{c=0\}} + \mathcal{O}\left(\varepsilon^2\right)\right).$$

In the second line we have dropped all odd powers of u and have used (5.10). Combining this with (5.8), (5.7) and (5.5) with the property (5.6), the proposition follows. □

In [25] Gao and Lee have obtained representations for asymptotic implied volatility for small and large-maturity regimes in terms of the assumed asymptotic behaviour of certain unspecified option prices, outlining the general procedure for transforming option price asymptotics into implied volatility asymptotics.

The same methodology can be followed to transform our forward-start option asymptotics (Corollary 2.8 and Corollary 2.10) into forward smile asymptotics. In the proofs of Proposition 2.12 and Proposition 2.14

we therefore assume for brevity the existence of an ansatz for the forward smile asymptotic and solve for the coefficients. We refer the reader to [25] for the complete methodology.

Proof of Proposition 2.12. Using $\Lambda_{0,1}(0)=0$ and substituting the ansatz $\sigma^2_{\varepsilon t,\varepsilon\tau}(k)=v_0(k,t,\tau)+v_1(k,t,\tau)\varepsilon+v_2(k,t,\tau)\varepsilon^2+\mathcal{O}\left(\varepsilon^3\right)$ into Corollary 2.9, we get that forward-start option prices have the asymptotics

$$\mathbb{E}\left[\left(\mathrm{e}^{X^{(\varepsilon t)}_{\varepsilon\tau}}-\mathrm{e}^k\right)^+\right]\mathbb{1}_{\{k>0\}}+\mathbb{E}\left[\left(\mathrm{e}^k-\mathrm{e}^{X^{(\varepsilon t)}_{\varepsilon\tau}}\right)^+\right]\mathbb{1}_{\{k<0\}}$$
$$=\exp\left(-\frac{k^2}{2\tau v_0(k,t,\tau)\varepsilon}+\frac{k^2 v_1(k,t,\tau)}{2\tau v_0(k,t,\tau)^2}+\frac{k}{2}\right)\frac{\left(v_0(k,t,\tau)\varepsilon\tau\right)^{3/2}}{k^2\sqrt{2\pi}}\left(1+\gamma(k,t,\tau)\varepsilon+\mathcal{O}\left(\varepsilon^2\right)\right),$$

where we set

$$\gamma(k,t,\tau):=-\tau v_0(k,t,\tau)\left(\frac{3}{k^2}+\frac{1}{8}\right)+\frac{k^2 v_2(k,t,\tau)}{2\tau v_0(k,t,\tau)^2}-\frac{k^2 v_1(k,t,\tau)^2}{2\tau v_0(k,t,\tau)^3}+\frac{3v_1(k,t,\tau)}{2v_0(k,t,\tau)}.$$

The result follows after equating orders with the general formula in Corollary 2.8. □

Proof of Proposition 2.14. Substituting the ansatz

$$\sigma^2_{t,\tau}(k)=v_0^\infty(k,t)+v_1^\infty(k,t)/\tau+v_2^\infty(k,t)/\tau^2+\mathcal{O}\left(1/\tau^3\right), \tag{5.11}$$

into Corollary 2.11 we obtain the following asymptotic expansions for forward-start options:

$$\mathbb{E}\left[\left(\mathrm{e}^{X^{(t)}_\tau}-\mathrm{e}^{k\tau}\right)^+\right]\mathbb{1}_{\mathrm{A}}-\mathbb{E}\left[\mathrm{e}^{X^{(t)}_\tau}\wedge\mathrm{e}^{k\tau}\right]\mathbb{1}_{\mathrm{B}}+\mathbb{E}\left[\left(\mathrm{e}^{k\tau}-\mathrm{e}^{X^{(t)}_\tau}\right)^+\right]\mathbb{1}_{\mathrm{C}}$$
$$=\exp\left(-\tau\left(\frac{k^2}{2v_0(k,t)}-\frac{k}{2}+\frac{v_0(k,t)}{8}\right)+\frac{v_1(k,t)k^2}{2v_0(k,t)^2}-\frac{v_1(k,t)}{8}\right)$$
$$\frac{4\tau^{-1/2}v_0(k,t)^{3/2}}{\left(4k^2-v_0(k,t)^2\right)\sqrt{2\pi}}\left(1+\frac{\gamma^\infty(k,t)}{\tau}+\mathcal{O}\left(\frac{1}{\tau^2}\right)\right),$$

where

$$\mathrm{A}:=\left\{k>\frac{1}{2}\sigma^2_{t,\tau}(k)\right\},\qquad \mathrm{B}:=\left\{-\frac{1}{2}\sigma^2_{t,\tau}(k)<k<\frac{1}{2}\sigma^2_{t,\tau}(k)\right\},\qquad \mathrm{C}:=\left\{k<-\frac{1}{2}\sigma^2_{t,\tau}(k)\right\}, \tag{5.12}$$

$$\gamma^\infty(k,t):=\frac{\left(12k^2+v_0^2(k,t)\right)\left(4k^2 v_1(k,t)-v_0^2(k,t)\left(v_1(k,t)+8\right)\right)}{2v_0(k,t)\left(v_0^2(k,t)-4k^2\right)^2}-\frac{v_1^2(k,t)k^2}{2v_0^3(k,t)}+\frac{v_2(k,t)k^2}{2v_0^2(k,t)}-\frac{v_2(k,t)}{8}.$$

We obtain the expressions for v_1^∞ and v_2^∞ by equating orders with the formula in Corollary 2.10. However it is not clear which is the correct root for the zeroth order term v_0^∞. In order to do so, we have to match the domains in (5.12) and in Corollary 2.10. Indeed, suppose that we choose the roots according to v_0^∞ in (2.16). For τ sufficiently large the condition $k>\sigma^2/2$ is equivalent to $k>v_0^\infty(k,t)/2$. Now for $k>\Lambda_{0,1}(1)$ or $k<\Lambda_{0,1}(0)$, the definition of v_0^∞ in (2.16) implies

$$k>\sigma^2/2\qquad\text{if and only if}\qquad\sqrt{\left(\Lambda^*(k)-k\right)^2+k\left(\Lambda^*(k)-k\right)}>\Lambda^*(k)-k, \tag{5.13}$$

which is always true since $\Lambda^*(k)>k$ by Lemma 2.4(iii). Now, for $k\in(\Lambda_{0,1}(0),\Lambda_{0,1}(1))$, the definition of v_0^∞ in (2.16) implies

$$k>\sigma^2/2\qquad\text{if and only if}\qquad-\sqrt{\left(\Lambda^*(k)-k\right)^2+k\left(\Lambda^*(k)-k\right)}>\Lambda^*(k)-k, \tag{5.14}$$

which never holds. By the assumption in the proposition 2.14 and Assumption 2.1 we have $\{0,1\}\subset\mathcal{D}^o_{0,0}$ and $\Lambda_{0,0}(0)=\Lambda_{0,0}(1)=0$. The differentiability and strict convexity of $\Lambda_{0,0}$ (Assumptions 2.3 and 2.1) then imply

$\Lambda_{0,1}(0) < 0$ and $\Lambda_{0,1}(1) > 0$. Since $v_0^\infty > 0$ we can ignore the case $k < \Lambda_{0,1}(0) < 0$ and hence $k > \sigma^2/2$ if and only if $k > \Lambda_{0,1}(1)$. Similarly the definition of v_0^∞ in (2.16) implies that for τ large enough τ,

$$-\sigma^2/2 < k < \sigma^2/2 \quad \text{if and only if} \quad \Lambda_{0,1}(0) < k < \Lambda_{0,1}(1) \text{ and } k < -\sigma^2/2 \quad \text{if and only if} \quad k < \Lambda_{0,1}(0).$$

This lines up the domains in (5.12) with the domains in Corollary 2.10. Had we specified the roots in any other way, it is easy to check that a contradiction would have occurred. □

5.2. **Proofs of Section 3.1.** For a stochastic process $(X_t)_{t\geq 0}$ we define the forward mgf of the process by

$$\Lambda(u) := \log \mathbb{E}\left(\mathrm{e}^{uX_\tau^{(t)}}\right), \qquad \text{for all } u \in \mathcal{D}_\Lambda, \tag{5.15}$$

where $\mathcal{D}_\Lambda := \{u \in \mathbb{R} : |\Lambda(u)| < \infty\}$ and $X_\tau^{(t)}$ is defined in (2.11). We now let $(X_t)_{t\geq 0}$ be the Heston process satisfying the SDE (3.1). By a straightforward application of the tower property for expectations (see also [34]), the forward mgf defined in (5.15) is given by

$$\Lambda(u) = A(u,\tau) + \frac{B(u,\tau)}{1-2\beta_t B(u,\tau)} v\mathrm{e}^{-\kappa t} - \frac{2\kappa\theta}{\xi^2}\log\left(1-2\beta_t B(u,\tau)\right), \quad \text{for all } u \in \mathcal{D}_\Lambda, \tag{5.16}$$

where

$$\begin{aligned} A(u,\tau) &:= \frac{\kappa\theta}{\xi^2}\left(\left(\kappa-\rho\xi u-d(u)\right)\tau - 2\log\left(\frac{1-\gamma(u)\exp\left(-d(u)\tau\right)}{1-\gamma(u)}\right)\right), \\ B(u,\tau) &:= \frac{\kappa-\rho\xi u-d(u)}{\xi^2}\frac{1-\exp\left(-d(u)\tau\right)}{1-\gamma(u)\exp\left(-d(u)\tau\right)}, \end{aligned} \tag{5.17}$$

and d, γ and β were introduced in (3.13). In the next two sections we develop the tools needed to apply Propositions 2.12 and 2.14 to the Heston model.

5.2.1. *Proofs of Section 3.1.1.* We consider here the Heston diagonal small-maturity process $\left(X_{\varepsilon\tau}^{(\varepsilon t)}\right)_{\varepsilon>0}$ with X defined in (3.1) and $\left(X_\tau^{(t)}\right)_{\tau>0}$ in (2.11). The forward rescaled mgf Λ_ε in (2.1) is easily determined from (5.16).

Lemma 5.3. *For the Heston diagonal small-maturity process we have $\mathcal{D}_{0,0} = \mathcal{D}_{t,\tau}$ and $0 \in \mathcal{D}_{0,0}^o$ with $\mathcal{D}_{t,\tau}$ defined in* (3.3) *and $\mathcal{D}_{0,0}$ and $\mathcal{D}_{0,0}^o$ defined in Assumption 2.1.*

Proof. For any $t > 0$, the random variable V_t in (3.1) is distributed as β_t times a non-central chi-square random variable with $q = 4\kappa\theta/\xi^2 > 0$ degrees of freedom and non-centrality parameter $\lambda = v\mathrm{e}^{-\kappa t}/\beta_t > 0$. It follows that the corresponding mgf is given by

$$\Lambda_t^V(u) := \mathbb{E}\left(\mathrm{e}^{uV_t}\right) = \exp\left(\frac{\lambda\beta_t u}{1-2\beta_t u}\right)\left(1-2\beta_t u\right)^{-q/2}, \qquad \text{for all } u < \frac{1}{2\beta_t}. \tag{5.18}$$

The re-normalised Heston forward mgf is then computed as

$$\frac{\Lambda_\varepsilon(u)}{\varepsilon} = \mathbb{E}\left[\mathrm{e}^{\frac{u}{\varepsilon}(X_{\varepsilon t+\varepsilon\tau}-X_{\varepsilon t})}\right] = \mathbb{E}\left[\mathbb{E}\left(\mathrm{e}^{\frac{u}{\varepsilon}(X_{\varepsilon t+\varepsilon\tau}-X_{\varepsilon t})}|\mathcal{F}_{\varepsilon t}\right)\right] = \mathbb{E}\left(\mathrm{e}^{A\left(\frac{u}{\varepsilon},\varepsilon\tau\right)+B\left(\frac{u}{\varepsilon},\varepsilon\tau\right)V_{\varepsilon t}}\right) = \mathrm{e}^{A\left(\frac{u}{\varepsilon},\varepsilon\tau\right)}\Lambda_{\varepsilon t}^V\left(B\left(u/\varepsilon,\varepsilon\tau\right)\right),$$

which agrees with (5.16). This only makes sense in some effective domain $\mathcal{D}_{\varepsilon t,\varepsilon\tau} \subset \mathbb{R}$. The mgf for $V_{\varepsilon t}$ is well-defined in $\mathcal{D}_{\varepsilon t}^V := \{u \in \mathbb{R} : B(u/\varepsilon,\varepsilon\tau) < \frac{1}{2\beta_{\varepsilon t}}\}$, and hence $\mathcal{D}_{\varepsilon t,\varepsilon\tau} = \mathcal{D}_{\varepsilon t}^V \cap \mathcal{D}_{\varepsilon\tau}^H$, where $\mathcal{D}_{\varepsilon\tau}^H$ is the effective domain of the (spot) Heston mgf. Let us first consider $\mathcal{D}_{\varepsilon\tau}^H$ for small ε. From [3, Proposition 3.1] if $\xi^2(u/\varepsilon-1)u > (\kappa-\xi\rho u/\varepsilon)^2$ then the explosion time $\tau^*(u) := \sup\{t \geq 0 : \mathbb{E}(\mathrm{e}^{uX_t}) < \infty\}$ of the Heston mgf is

$$\tau_H^*\left(\frac{u}{\varepsilon}\right) = \frac{2}{\sqrt{\xi^2(u/\varepsilon-1)u/\varepsilon-(\kappa-\rho\xi u/\varepsilon)^2}}\left(\pi\mathbb{1}_{\{\rho\xi u/\varepsilon-\kappa<0\}} + \arctan\left(\frac{\sqrt{\xi^2(u/\varepsilon-1)u/\varepsilon-(\kappa-\rho\xi u/\varepsilon)^2}}{\rho\xi u/\varepsilon-\kappa}\right)\right).$$

Recall the following Taylor series expansions, for x close to zero:

$$\begin{aligned}\arctan\left(\frac{1}{\rho\xi u/x-\kappa}\sqrt{\xi^2\left(\frac{u}{x}-1\right)\frac{u}{x}-\left(\kappa-\xi\rho\frac{u}{x}\right)^2}\right) &= \operatorname{sgn}(u)\arctan\left(\frac{\bar\rho}{\rho}\right)+\mathcal{O}\left(x\right), && \text{if } \rho\neq 0,\\ \arctan\left(-\frac{1}{\kappa}\sqrt{\xi^2\left(\frac{u}{x}-1\right)\frac{u}{x}-\kappa^2}\right) &= -\frac{\pi}{2}+\mathcal{O}(x), && \text{if } \rho=0.\end{aligned}$$

As ε tends to zero $\xi^2(u/\varepsilon-1)u/\varepsilon>(\kappa-\rho\xi u/\varepsilon)^2$ is satisfied since $\xi^2>\xi^2\rho^2$ and hence

$$\tau_H^*\left(u/\varepsilon\right)=\begin{cases}\dfrac{\varepsilon}{\xi|u|}\left(\pi\mathbb{1}_{\{\rho=0\}}+\dfrac{2}{\bar\rho}\left(\pi\mathbb{1}_{\{\rho u\leq 0\}}+\operatorname{sgn}(u)\arctan\left(\dfrac{\bar\rho}{\rho}\right)\right)\mathbb{1}_{\{\rho\neq 0\}}+\mathcal{O}(\varepsilon)\right), & \text{if } u\neq 0,\\ \infty, & \text{if } u=0.\end{cases}$$

Therefore, for ε small enough, we have $\tau_H^*\left(\frac{u}{\varepsilon}\right)>\varepsilon\tau$ for all $u\in(u_-,u_+)$, where

$$\begin{aligned}u_- &:= \frac{2}{\bar\rho\xi\tau}\arctan\left(\frac{\bar\rho}{\rho}\right)\mathbb{1}_{\{\rho<0\}}-\frac{\pi}{\xi\tau}\mathbb{1}_{\{\rho=0\}}+\frac{2}{\bar\rho\xi\tau}\left(\arctan\left(\frac{\bar\rho}{\rho}\right)-\pi\right)\mathbb{1}_{\{\rho>0\}},\\ u_+ &:= \frac{2}{\bar\rho\xi\tau}\left(\arctan\left(\frac{\bar\rho}{\rho}\right)+\pi\right)\mathbb{1}_{\{\rho<0\}}+\frac{\pi}{\xi\tau}\mathbb{1}_{\{\rho=0\}}+\frac{2}{\bar\rho\xi\tau}\arctan\left(\frac{\bar\rho}{\rho}\right)\mathbb{1}_{\{\rho>0\}}.\end{aligned}$$

So as ε tends to zero, $\mathcal{D}^H_{\varepsilon\tau}$ shrinks to (u_-,u_+). Regarding $\mathcal{D}^V_{\varepsilon t}$, we have (see (5.22) for details on the expansion computation) $\beta_{\varepsilon t}B\left(u/\varepsilon,\varepsilon\tau\right)=\frac{\xi^2 t}{4v}\Lambda\left(u,0,\tau\right)+\mathcal{O}(\varepsilon)$ for any $u\in(u_-,u_+)$, with Λ defined in (3.2). Therefore $\lim_{\varepsilon\searrow 0}\mathcal{D}^V_{\varepsilon t}=\{u\in\mathbb{R}:\Lambda\left(u,0,\tau\right)<\frac{2v}{\xi^2 t}\}$ and hence $\lim_{\varepsilon\searrow 0}\mathcal{D}_{\varepsilon t,\varepsilon\tau}=\{u\in\mathbb{R}:\Lambda\left(u,0,\tau\right)<\frac{2v}{\xi^2 t}\}\cap(u_-,u_+)$. It is easily checked that $\Lambda\left(u,0,\tau\right)$ is strictly positive except at $u=0$ where it is zero, $\Lambda'\left(u,0,\tau\right)>0$ for $u>0$, $\Lambda'\left(u,0,\tau\right)<0$ for $u<0$ and that $\Lambda\left(u,0,\tau\right)$ tends to infinity as u approaches $u_\pm$. Since v and ξ are strictly positive and $t\geq 0$ it follows that $\{u\in\mathbb{R}:\Lambda\left(u,0,\tau\right)<\frac{2v}{\xi^2 t}\}\subseteq(u_-,u_+)$ with equality only if $t=0$. So $\mathcal{D}_{0,0}$ is an open interval around zero and the lemma follows with $\mathcal{D}_{0,0}=\mathcal{D}_{t,\tau}$. □

Remark 5.4. For $u\in\mathbb{R}^*$ the inequality $0<\Lambda(u,0,\tau)<\frac{2v}{\xi^2 t}$ is equivalent to $\Lambda(u,t,\tau)\in(0,\infty)$, where Λ is defined in (3.2). In Lemma 5.5 below we show that Λ is the limiting mgf of the rescaled Heston forward mgf and so the condition for the limiting forward domain is equivalent to ensuring that the limiting forward mgf does not blow up and is strictly positive except at $u=0$ where it is zero.

Lemma 5.5. *For any $t\geq 0$, $\tau>0$, $u\in\mathcal{D}_{t,\tau}$, the following expansion holds as ε tends to zero:*

$$\Lambda_\varepsilon(u)=\Lambda(u,t,\tau)+L(u,t,\tau)\varepsilon+\mathcal{O}\left(\varepsilon^2\right),$$

where $\mathcal{D}_{t,\tau}$, Λ and L are defined in (3.3), (3.2) and (3.4) and Λ_ε is the rescaled mgf in Assumption 2.1 for the Heston diagonal small-maturity process $\left(X^{(\varepsilon t)}_{\varepsilon\tau}\right)_{\varepsilon>0}$.

Remark 5.6. For any $u\in\mathcal{D}_{t,\tau}$, Lemma 5.3 implies that $\Lambda_\varepsilon(u)$ is a real number for any $\varepsilon>0$. Therefore L defined in (3.4) and used in Lemma 5.5 is a real-valued function on $\mathcal{D}_{t,\tau}$.

Proof. All expansions below for d, γ and β_t defined in (3.13) hold for any $u\in\mathcal{D}_{t,\tau}$:

$$\begin{aligned}d\left(u/\varepsilon\right) &= \frac{1}{\varepsilon}\left(\kappa^2\varepsilon^2+u\varepsilon\left(\xi-2\kappa\rho\right)-u^2\xi^2\bar\rho\right)^{1/2}=\frac{\mathtt{i}u}{\varepsilon}d_0+d_1+\mathcal{O}(\varepsilon),\\ \gamma\left(u/\varepsilon\right) &= \frac{\kappa\varepsilon-\rho\xi u-\mathtt{i}ud_0-d_1\varepsilon+\mathcal{O}\left(\varepsilon^2\right)}{\kappa\varepsilon-\rho\xi u+\mathtt{i}ud_0+d_1\varepsilon+\mathcal{O}\left(\varepsilon^2\right)}=g_0-\frac{\mathtt{i}\varepsilon}{u}g_1+\mathcal{O}\left(\varepsilon^2\right),\end{aligned}$$

$$\beta_{\varepsilon t}=\frac{1}{4}\xi^2 t\varepsilon-\frac{1}{8}\kappa\xi^2 t^2\varepsilon^2+\mathcal{O}\left(\varepsilon^3\right), \tag{5.19}$$

where we have set

$$d_0 := \bar{\rho}\xi\,\mathrm{sgn}(u), \qquad d_1 := \frac{\mathtt{i}\,(2\kappa\rho-\xi)\,\mathrm{sgn}(u)}{2\bar{\rho}}, \qquad g_0 := \frac{\mathtt{i}\rho-\bar{\rho}\,\mathrm{sgn}(u)}{\mathtt{i}\rho+\bar{\rho}\,\mathrm{sgn}(u)} \qquad g_1 := \frac{(2\kappa-\xi\rho)\,\mathrm{sgn}(u)}{\xi\bar{\rho}\,(\bar{\rho}+\mathtt{i}\rho\,\mathrm{sgn}(u))^2}, \tag{5.20}$$

where $\mathrm{sgn}(u)=1$ if $u\geq 0$ and -1 otherwise. From the definition of A in (5.17) we obtain

$$\begin{aligned} A\left(u/\varepsilon,\varepsilon\tau\right) &= \frac{\kappa\theta}{\xi^2}\left(\left(\kappa-\rho\xi u/\varepsilon-d(u/\varepsilon)\right)\varepsilon\tau-2\log\left(\frac{1-\gamma(u/\varepsilon)\exp\left(-d(u/\varepsilon)\varepsilon\tau\right)}{1-\gamma(u/\varepsilon)}\right)\right) \\ &= \frac{\kappa\theta}{\xi^2}\left(\left(\kappa\varepsilon-\rho\xi u-\mathtt{i}ud_0-\varepsilon d_1+\mathcal{O}(\varepsilon^2)\right)\tau\right. \\ &\left.-2\log\left(\frac{1-\left(g_0-\mathtt{i}\varepsilon g_1/u+\mathcal{O}(\varepsilon^2)\right)\exp\left(-\mathtt{i}ud_0\tau-\varepsilon d_1\tau+\mathcal{O}(\varepsilon^2)\right)}{1-\left(g_0-\mathtt{i}\varepsilon g_1/u+\mathcal{O}(\varepsilon^2)\right)}\right)\right) \\ &= L_0(u,\tau)+\mathcal{O}(\varepsilon), \end{aligned} \tag{5.21}$$

where L_0 is defined in (3.5). Substituting the asymptotics for d and γ above we further obtain

$$\frac{1-\exp\left(-d(u/\varepsilon)\varepsilon\tau\right)}{1-\gamma(u/\varepsilon)\exp\left(-d(u/\varepsilon)\varepsilon\tau\right)} = \frac{1-\exp\left(-\mathtt{i}ud_0\tau-\varepsilon d_1\tau+\mathcal{O}(\varepsilon^2)\right)}{1-\left(g_0-\mathtt{i}\varepsilon g_1/u+\mathcal{O}(\varepsilon^2)\right)\exp\left(-\mathtt{i}ud_0\tau-\varepsilon d_1\tau+\mathcal{O}(\varepsilon^2)\right)},$$

and therefore using the definition of B in (5.17) we obtain

$$\begin{aligned} B(u/\varepsilon,\varepsilon\tau) &= \frac{\kappa-\rho\xi u/\varepsilon-d(u/\varepsilon)}{\xi^2}\,\frac{1-\exp\left(-d\left(u/\varepsilon\right)\varepsilon\tau\right)}{1-\gamma\left(u/\varepsilon\right)\exp\left(-d\left(u/\varepsilon\right)\varepsilon\tau\right)} \\ &= -\frac{\rho\xi u+\mathtt{i}ud_0}{\varepsilon\xi^2}\,\frac{1-\exp\left(-\mathtt{i}ud_0\tau\right)}{1-g_0\exp\left(-\mathtt{i}ud_0\tau\right)}+L_1(u,\tau)+\mathcal{O}(\varepsilon) \\ &= \frac{\Lambda(u,0,\tau)}{v\varepsilon}+L_1(u,\tau)+\mathcal{O}(\varepsilon), \end{aligned} \tag{5.22}$$

with L_1 defined in (3.5) and Λ in (3.2). Combining (5.19) and (5.22) we deduce

$$\beta_{\varepsilon t}B\left(u/\varepsilon,\varepsilon\tau\right) = \frac{\xi^2 t\Lambda(u,0,\tau)}{4v}+\left(\frac{L_1(u,\tau)\xi^2 t}{4}-\frac{\Lambda(u,0,\tau)\kappa\xi^2t^2}{8v}\right)\varepsilon+\mathcal{O}(\varepsilon^2), \tag{5.23}$$

and therefore as ε tends to zero,

$$\begin{aligned} \frac{\varepsilon B(u/\varepsilon,\varepsilon\tau)v\mathrm{e}^{-\kappa\varepsilon t}}{1-2\beta_{\varepsilon t}B(u/\varepsilon,\varepsilon\tau)} &= \frac{\left[\Lambda(u,0,\tau)+vL_1(u,\tau)\varepsilon+\mathcal{O}\left(\varepsilon^2\right)\right]\left(1-t\kappa\xi+\mathcal{O}(\varepsilon^2)\right)}{1-\xi^2t\Lambda(u,0,\tau)/2v+\left(\Lambda(u,0,\tau)\kappa\xi^2t^2/4v-L_1(u,\tau)\xi^2t/2\right)\varepsilon+\mathcal{O}\left(\varepsilon^2\right)} \\ &= \Lambda(u,t,\tau)+\left(\Lambda(u,t,\tau)^2\left(\frac{vL_1(u,\tau)}{\Lambda(u,0,\tau)^2}-\frac{\kappa\xi^2t^2}{4v}\right)-\kappa t\Lambda(u,t,\tau)\right)\varepsilon+\mathcal{O}(\varepsilon^2). \end{aligned} \tag{5.24}$$

Again using (5.23) we have

$$-\frac{2\kappa\theta\varepsilon}{\xi^2}\log\left(1-2\beta_{\varepsilon t}B\left(u/\varepsilon,\varepsilon\tau\right)\right) = -\frac{2\kappa\theta}{\xi^2}\log\left(1-\frac{\Lambda(u,0,\tau)\xi^2t}{2v}\right)\varepsilon+\mathcal{O}(\varepsilon^2). \tag{5.25}$$

Recalling that

$$\Lambda_\varepsilon(u) = \varepsilon A\left(u/\varepsilon,\varepsilon\tau\right)+\frac{\varepsilon B\left(u/\varepsilon,\varepsilon\tau\right)}{1-2\beta_{\varepsilon t}B\left(u/\varepsilon,\varepsilon\tau\right)}v\mathrm{e}^{-\kappa\varepsilon t}-\frac{2\kappa\theta\varepsilon}{\xi^2}\log\left(1-2\beta_{\varepsilon t}B\left(u/\varepsilon,\varepsilon\tau\right)\right),$$

the lemma follows by combining (5.21), (5.24) and (5.25). □

Lemma 5.7. *For all $t\geq 0$, $\tau>0$, Λ (given in* (3.2)*) is convex and essentially smooth on $\mathcal{D}_{t,\tau}$, defined in* (3.3).

Proof. The first derivative of Λ is given by

$$\frac{\partial\Lambda(u,t,\tau)}{\partial u} = \frac{-v}{\rho\xi+\frac{1}{2}\xi^2 tu-\bar{\rho}\xi\cot\left(\frac{1}{2}\xi\bar{\rho}\tau u\right)} + \frac{uv\left(\frac{\xi^2 t}{2}+\frac{1}{2}\xi^2\bar{\rho}^2\tau\csc^2\left(\frac{1}{2}\xi\bar{\rho}\tau u\right)\right)}{\left(-\xi\rho-\frac{1}{2}\xi^2 tu+\xi\bar{\rho}\cot\left(\frac{1}{2}\xi\bar{\rho}\tau u\right)\right)^2}$$
$$= \frac{\Lambda(u,t,\tau)}{u}\left(1+\frac{\Lambda(u,t,\tau)}{v}\left(\frac{\xi^2 t}{2}+\frac{1}{2}\xi^2\bar{\rho}^2\tau\csc^2\left(\frac{1}{2}\bar{\rho}\xi\tau u\right)\right)\right).$$

Any sequence tending to the boundary satisfies $\Lambda(u,0,\tau)\to 2v/\xi^2 t$ which implies $\Lambda(u,t,\tau)\nearrow\infty$ from Remark 5.4 and hence $|\partial\Lambda(u,t,\tau)/\partial u|\nearrow\infty$. Therefore $\Lambda(\cdot,t,\tau)$ is essential smooth. Now,

$$\frac{\partial^2\Lambda(u,t,\tau)}{\partial u^2} = \frac{2uv\left(-\frac{\xi^2 t}{2}-\frac{1}{2}\xi^2\bar{\rho}^2\tau\csc^2(\psi_u)\right)^2}{\left(-\xi\rho-\frac{1}{2}\xi^2 tu+\bar{\rho}\xi\cot(\psi_u)\right)^3} + \frac{\xi^2 tv+\bar{\rho}^2\xi^2\tau v\left(1-\psi_u\cot(\psi_u)\right)\csc^2(\psi_u)}{\left(-\rho\xi-\frac{1}{2}\xi^2 tu+\bar{\rho}\xi\cot(\psi_u)\right)^2}$$
$$= \frac{\xi^2}{2}\Lambda(u,t,\tau)\frac{\left(t+\bar{\rho}^2\tau\csc^2(\psi_u)\right)^2}{\left(\rho+\frac{1}{2}\xi tu-\bar{\rho}\cot(\psi_u)\right)^2} + \frac{v+\bar{\rho}^2\tau v\left(1-\psi_u\cot(\psi_u)\right)\csc^2(\psi_u)}{\left(\rho+\frac{1}{2}\xi tu-\bar{\rho}\cot(\psi_u)\right)^2}.$$

where $\psi_u := \frac{1}{2}\bar{\rho}\xi\tau u$. For $u\in\mathcal{D}_{t,\tau}\setminus\{0\}$, we have $\Lambda(u,t,\tau)>0$ and $\Lambda(0,t,\tau)=0$ from Remark 5.4. Also we have the identity that $1-\theta/2\cot(\theta/2)\geq 0$ for $\theta\in(-2\pi,2\pi)$, so that Λ is strictly convex on $\mathcal{D}_{t,\tau}$. □

Proof of Corollary 3.3. We first look for a Taylor expansion of $u^*(k)$ around $k=0$ using $\Lambda'(u^*(k),t,\tau)=k$. Differentiating this equation iteratively and setting $k=0$ (and using $u^*(0)=0$) gives an expansion for u^* in terms of the derivatives of Λ. In particular, $\Lambda''(0,t,\tau)u^{*'}(0)=1$ and $\Lambda'''(0,t,\tau)(u^{*'}(0))^2+\Lambda''(0,t,\tau)u^{*''}(0)=0$, which implies that $u^{*'}(0)=1/\Lambda''(0,t,\tau)$ and $u^{*''}(0)=-\Lambda'''(0,t,\tau)/\Lambda''(0,t,\tau)^3$. From the explicit expression of Λ in (3.2), we then obtain

$$u^*(k) = \frac{k}{\tau v}-\frac{3\xi\rho}{4\tau v^2}k^2+\frac{\xi^2\left(\left(19\rho^2-4\right)\tau-12t\right)}{24\tau^2 v^3}k^3+\frac{5\xi^3\rho\left(48t+\left(16-37\rho^2\right)\tau\right)}{192\tau^2 v^4}k^4$$
$$+\frac{\xi^4\left(1080t^2+\left(2437\rho^4-1604\rho^2+112\right)\tau^2-180\left(27\rho^2-4\right)\tau t\right)}{1920\tau^3 v^5}k^5+\mathcal{O}(k^6).$$

Using this series expansion and the fact that $\Lambda^*(k)=u^*(k)k-\Lambda(u^*(k),t,\tau)$, the corollary follows from tedious but straightforward Taylor expansions of v_0 and v_1 defined in (2.14). □

Lemma 5.8. *Under the stopped-share-price measure* (2.18) *the forward Heston mgf defined in* (5.15) *reads*

$$\Lambda(u) = A(u,\tau)+\frac{B(u,\tau)}{1-2\widetilde{\beta}_t B(u,\tau)}v\mathrm{e}^{-\widetilde{\kappa}t}-\frac{2\kappa\theta}{\xi^2}\log\left(1-2\widetilde{\beta}_t B(u,\tau)\right), \qquad \textit{for all } u\in\mathcal{D}_\Lambda,$$

where A and B are defined in (5.17)*,*

$$\widetilde{\beta}_t := \frac{\xi^2}{4\widetilde{\kappa}}(1-\mathrm{e}^{-\widetilde{\kappa}t}) \qquad \textit{and} \qquad \widetilde{\kappa} := \kappa-\xi\rho. \tag{5.26}$$

Proof. Under the stopped-share-price measure (2.18) the Heston dynamics are given by

$$\begin{aligned} \mathrm{d}X_u &= \left(-\tfrac{1}{2}V_u+V_u\mathbb{1}_{u\leq t}\right)\mathrm{d}u+\sqrt{V_u}\mathrm{d}W_u, && X_0\in\mathbb{R}, \\ \mathrm{d}V_u &= \left(\kappa\theta-\kappa V_u+\rho\xi V_u\mathbb{1}_{u\leq t}\right)\mathrm{d}u+\xi\sqrt{V_u}\mathrm{d}Z_u, && V_0=v>0, \\ \mathrm{d}\left\langle W,Z\right\rangle_u &= \rho\mathrm{d}u. \end{aligned} \tag{5.27}$$

We now compute

$$\widetilde{\mathbb{E}}\left(\mathrm{e}^{u(X_{t+\tau}-X_t)}\right) = \widetilde{\mathbb{E}}\left(\widetilde{\mathbb{E}}\left(\mathrm{e}^{u(X_{t+\tau}-X_t)}|\mathcal{F}_t\right)\right) = \widetilde{\mathbb{E}}\left(\mathrm{e}^{A(u,\tau)+B(u,\tau)V_t}\right) \tag{5.28}$$

$$= \mathrm{e}^{A(u,\tau)}\widetilde{\mathbb{E}}\left(\mathrm{e}^{B(u,\tau)V_t}\right) = \mathrm{e}^{A(u,\tau)}\widetilde{\Lambda}_t^V(B(u,\tau)), \tag{5.29}$$

where

$$\widetilde{\Lambda}_t^V(u) = \exp\left(\frac{\widetilde{\beta}_t u \widetilde{\lambda}}{1-2\widetilde{\beta}_t u}\right)\left(1-2\widetilde{\beta}_t u\right)^{-q/2} \quad \text{for all } u < \frac{1}{2\widetilde{\beta}_t}, \tag{5.30}$$

with $q := 4\kappa\theta/\xi^2$, $\widetilde{\lambda} := v\mathrm{e}^{-\widetilde{\kappa}t}/\widetilde{\beta}_t$. Over $(t, t+\tau]$ the Heston dynamics in (5.27) remain the same as under the risk-neutral measure and so we can apply the standard spot Heston mgf in (5.28). On $[0,t]$ we use the modified chi-squared mgf in (5.30) to obtain (5.29), corresponding to the modified Heston process in (5.27) on $[0,t]$. □

5.2.2. *Proofs of Section 3.1.2.* In this section we shall use the standing assumption that $\kappa > \rho\xi$. Let $\varepsilon = \tau^{-1}$ and consider the Heston process $\left(\tau^{-1}X_\tau^{(t)}\right)_{\tau>0}$ with $(X_t)_t$ defined in (3.1) and $\left(X_\tau^{(t)}\right)_\tau$ defined in (2.11). Specifically Λ_ε defined in (2.1) is then given by $\Lambda_\varepsilon(u) = \tau^{-1}\mathbb{E}\left(\mathrm{e}^{uX_\tau^{(t)}}\right)$, and for ease of notation we set

$$\Lambda_\tau^{(t)}(u) = \Lambda_\varepsilon(u) \quad \text{for all } u \in \mathcal{D}_\varepsilon. \tag{5.31}$$

The following lemma [20, page 13] recalls some elementary facts about the function V in (3.12), which will be used throughout the section. We then proceed with two technical results needed in the proof of Proposition 5.13.

Lemma 5.9. *The function V defined in* (3.12) *is infinitely differentiable, strictly convex, essentially smooth on the open interval (u_-, u_+) with $u_\pm$ defined in* (3.8) *and*

$$V(u_-) := \lim_{u\searrow u_-} V(u) = \frac{\kappa\theta\left(2\kappa-\rho\xi+\rho\eta\right)}{2\xi^2\left(1-\rho^2\right)} < \infty,$$
$$V(u_+) := \lim_{u\nearrow u_+} V(u) = \frac{\kappa\theta\left(2\kappa-\rho\xi-\rho\eta\right)}{2\xi^2\left(1-\rho^2\right)} < \infty,$$

with $u_- < 0$, $u_+ > 1$ and $V(0) = V(1) = 0$. Furthermore, it has a unique minimum at $\frac{\xi-2\rho\kappa-\rho\eta}{2\xi(1-\rho^2)} \in (0,1)$.

Lemma 5.10. *Let $\rho_\pm$ be defined as in* (3.11), *β_t in* (3.13), *and recall the standing assumption $\rho < \kappa/\xi$. Assume further that $t > 0$ and define the functions g_+ and g_- by*

$$g_\pm(\rho) := (2\kappa-\rho\xi) \pm \rho\sqrt{\xi^2\left(1-\rho^2\right)+\left(2\kappa-\rho\xi\right)^2} - \frac{\xi^2(1-\rho^2)}{\beta_t}.$$

(i) *The inequality $-1 < \rho_- < 0$ always holds and if $\kappa/\xi > \rho_+$ then $1/2 < \rho_+ < 1$;*
(ii) *the inequality $g_+(\rho) > 0$ holds if and only if $\rho \in (\rho_+, 1)$ and $\kappa/\xi > \rho_+$;*
(iii) *the inequality $g_-(\rho) > 0$ holds if and only if $\rho \in (-1, \rho_-)$.*

Remark 5.11. From the proof, we have the equality $\rho_- = -1$ if and only if $t = 0$. Also if $\kappa > \rho_+\xi$, then $\rho_+ = 1$ if and only if $t = 0$.

Proof. We first prove Lemma 5.10(i) and consider the inequality $-1 < \rho_- < 0$. Using the definition of ρ_- in (3.11) this is equivalent to proving that

$$\frac{\xi-(8\kappa+\xi)\mathrm{e}^{2\kappa t}}{\mathrm{e}^{\kappa t}+1} < -\sqrt{16\kappa^2\mathrm{e}^{2\kappa t}+\xi^2\left(1-\mathrm{e}^{\kappa t}\right)^2} < \xi\left(1-\mathrm{e}^{\kappa t}\right) = \frac{\xi(1-\mathrm{e}^{2\kappa t})}{1+\mathrm{e}^{\kappa t}}.$$

The upper bound then follows trivially. Also we can write

$$\sqrt{16\kappa^2\mathrm{e}^{2\kappa t}+\xi^2\left(1-\mathrm{e}^{\kappa t}\right)^2} = \sqrt{\frac{\left(\xi-(8\kappa+\xi)\mathrm{e}^{2\kappa t}\right)^2}{\left(\mathrm{e}^{\kappa t}+1\right)^2} - \frac{16\kappa\mathrm{e}^{2\kappa t}\left(\mathrm{e}^{\kappa t}-1\right)\left(\kappa+\xi+\xi\mathrm{e}^{\kappa t}+3\kappa\mathrm{e}^{\kappa t}\right)}{\left(\mathrm{e}^{\kappa t}+1\right)^2}},$$

and the lower bound follows. We now prove that $\rho_+ > 1/2$. From (3.11) this is equivalent to

$$\sqrt{16\kappa^2 \mathrm{e}^{2\kappa t} + \xi^2 \left(1 - \mathrm{e}^{\kappa t}\right)^2} > \frac{4\xi + (\kappa - 4\xi)\mathrm{e}^{2\kappa t}}{4\left(\mathrm{e}^{\kappa t} + 1\right)}.$$

Tedious rearrangements show that the left-hand side can be written as

$$\sqrt{16\kappa^2 \mathrm{e}^{2\kappa t} + \xi^2 \left(1 - \mathrm{e}^{\kappa t}\right)^2} = \sqrt{\frac{\left(4\xi + (\kappa - 4\xi)\mathrm{e}^{2\kappa t}\right)^2}{16\left(\mathrm{e}^{\kappa t} + 1\right)^2} + \frac{\kappa \mathrm{e}^{2\kappa t}\left(8\xi\left(\mathrm{e}^{2\kappa t} - 1\right) + \kappa\left(512\mathrm{e}^{\kappa t} + 255\mathrm{e}^{2\kappa t} + 256\right)\right)}{16\left(\mathrm{e}^{\kappa t} + 1\right)^2}},$$

and the result follows. We now prove that the upper bound $\rho_+ < 1$ holds if $\kappa/\xi > \rho_+$. Assume that $\kappa/\xi > \rho_+$. Since $\frac{\mathrm{e}^{\kappa t}+1}{3\mathrm{e}^{\kappa t}+1}$ is always strictly smaller than $1/2$, we immediately obtain the inequality

$$\frac{\mathrm{e}^{\kappa t} + 1}{3\mathrm{e}^{\kappa t} + 1} < \frac{\kappa}{\xi}. \tag{5.32}$$

Using the definition of ρ_+ in (3.11) the statement $\rho_+ < 1$ is equivalent to

$$\sqrt{16\kappa^2 \mathrm{e}^{2\kappa t} + \xi^2 \left(1 - \mathrm{e}^{\kappa t}\right)^2} < \frac{\xi + (8\kappa - \xi)\mathrm{e}^{2\kappa t}}{\mathrm{e}^{\kappa t} + 1},$$

which can be written as

$$\sqrt{\frac{\left(\xi + (8\kappa - \xi)\mathrm{e}^{2\kappa t}\right)^2}{\left(\mathrm{e}^{\kappa t} + 1\right)^2} - \frac{16\kappa \mathrm{e}^{2\kappa t}\left(\mathrm{e}^{\kappa t} - 1\right)\left(\kappa - \xi\left(\mathrm{e}^{\kappa t} + 1\right) + 3\kappa \mathrm{e}^{\kappa t}\right)}{\left(\mathrm{e}^{\kappa t} + 1\right)^2}} < \frac{\xi + (8\kappa - \xi)\mathrm{e}^{2\kappa t}}{\mathrm{e}^{\kappa t} + 1}. \tag{5.33}$$

This statement is true if $\kappa - \xi\left(\mathrm{e}^{\kappa t} + 1\right) + 3\kappa \mathrm{e}^{\kappa t} > 0$ and if the rhs is positive. The former inequality is precisely (5.32), and immediately implies the positivity condition. Therefore $\rho_+ < 1$.

We now prove Lemma 5.10(ii). The equation $g_+(\rho) = 0$ implies (by squaring and rearranging the terms):

$$4\kappa(\rho^2 - 1)\left(4\kappa \mathrm{e}^{2\kappa t}\rho^2 + \xi(1 - \mathrm{e}^{2\kappa t})\rho - \kappa(1 + 2\mathrm{e}^{\kappa t} + \mathrm{e}^{2\kappa t})\right) = 0. \tag{5.34}$$

The roots of this equation are ± 1 and $\rho_\pm$ defined in (3.11). Clearly some of these solutions are extraneous. The two possible positive roots are $\{\rho_+, 1\}$ and the two possible negative ones are $\{\rho_-, -1\}$. Clearly $g_+(-1) = 0$. Straightforward computations show that $g_+'(-1) < 0$ and $g_+'(0) > 0$. Since g_+ is continuous on $(-1, 0)$ with $g_+(0) < 0$, it cannot have a single root in this interval, and $\rho_- \in (-1, 0)$ (by Lemma 5.10(i)) is hence not a valid root. Consider now $\rho \in (0, \kappa/\xi)$. From Lemma 5.10(i) the only possible roots are $\rho \in \{1, \rho_+\}$. If $\rho = 1$ then clearly $\kappa/\xi > 1$ by the standing assumption. But $g_+(1) = 2\kappa - \xi + \sqrt{(2\kappa - \xi)^2}$ is null if and only if $\kappa/\xi < 1/2$, which is a contradiction. Therefore the only possible positive solution is ρ_+. Now, on $(0, \kappa/\xi)$ we have

$$\begin{aligned}
g_+'(\rho) &= -\xi - \frac{2\kappa\xi\rho}{\sqrt{4\kappa^2 - 4\kappa\xi\rho + \xi^2}} + \sqrt{4\kappa^2 - 4\kappa\xi\rho + \xi^2} + \frac{8\kappa\rho}{1 - \mathrm{e}^{-\kappa t}} \\
&> -\xi - \frac{2\kappa\xi\rho}{\sqrt{4\kappa^2 - 4\kappa\xi\left(\kappa/\xi\right) + \xi^2}} + \sqrt{4\kappa^2 - 4\kappa\xi\left(\kappa/\xi\right) + \xi^2} + \frac{8\kappa\rho}{1 - \mathrm{e}^{-\kappa t}} \\
&= \frac{8\kappa\rho}{1 - \mathrm{e}^{-\kappa t}} - 2\kappa\rho > 0.
\end{aligned}$$

In summary, g_+ is strictly increasing on $(0, \kappa/\xi)$ with a unique zero at ρ_+ satisfying $\kappa > \rho_+\xi$. On the interval $(-1, \kappa/\xi)$, $g_+(\rho) > 0$ if and only if $\rho \in (\rho_+, 1)$ and $\rho_+\xi < \kappa$. The proof of (iii) is analogous to the proof of (ii) and we omit it for brevity.

□

Lemma 5.12. *Let $\rho_\pm$ and $u_\pm^*$ be as in* (3.11) *and* (3.8) *and $t > 0$. Then $u_+^* > 1$ if $\rho \leq \rho_-$, and $u_-^* < 0$ if $\rho \geq \rho_+$.*

Proof. From (3.9) write $\nu = \sqrt{z(\rho)}$, where

$$z(\rho) := \xi^2 - 2\mathrm{e}^{\kappa t}\left(8\kappa^2 - 4\kappa\xi\rho + \xi^2\right) + \mathrm{e}^{2\kappa t}(\xi - 4\kappa\rho)^2.$$

The two numbers u_-^* and u_+^* in (3.8) are well-defined in $\mathbb{R}$ if and only if $z(\rho) \geq 0$ and $t > 0$. The two roots of this quadratic polynomial are given by $\chi_\pm := \frac{1}{4\kappa}\left[\mathrm{e}^{-\kappa t}\left(\xi(\mathrm{e}^{\kappa t} - 1) \pm 4\kappa\mathrm{e}^{\kappa t/2}\right)\right]$. We now claim that $\rho_- \leq \chi_-$ and $\rho_+ \geq \chi_+$. From the expression of ρ_- given in (3.11), the inequality $\rho_- \leq \chi_-$ can be rearranged as

$$-\sqrt{\xi^2 + 16\kappa^2\mathrm{e}^{2\kappa t} - 2\xi^2\mathrm{e}^{\kappa t} + \xi^2\mathrm{e}^{2\kappa t}} \leq \frac{\xi - 2\xi\mathrm{e}^{\kappa t} + \xi\mathrm{e}^{2\kappa t} - 8\kappa\mathrm{e}^{\frac{3\kappa t}{2}}}{\mathrm{e}^{\kappa t} + 1}.$$

Noting the the square root term is equal to

$$\sqrt{\xi^2 + 16\kappa^2\mathrm{e}^{2\kappa t} - 2\xi^2\mathrm{e}^{\kappa t} + \xi^2\mathrm{e}^{2\kappa t}} = \sqrt{\frac{4\mathrm{e}^{\kappa t}\left(\mathrm{e}^{\kappa t} - 1\right)^2\left(\xi + 2\kappa\mathrm{e}^{\frac{\kappa t}{2}}\right)^2}{\left(\mathrm{e}^{\kappa t} + 1\right)^2} + \frac{\left(\xi - 2\xi\mathrm{e}^{\kappa t} + \xi\mathrm{e}^{2\kappa t} - 8\kappa\mathrm{e}^{\frac{3\kappa t}{2}}\right)^2}{\left(\mathrm{e}^{\kappa t} + 1\right)^2}},$$

proves the claim. Analogous manipulations imply the other claim $\rho_+ \geq \chi_+$, and hence $z(\rho)$ is a well-defined real number for $\rho \in [-1, \rho_-] \cup [\rho_+, 1]$.

The claim $u_-^* < 0$ is equivalent to $-\sqrt{\xi^2 - 2\mathrm{e}^{\kappa t}\left(8\kappa^2 - 4\kappa\xi\rho + \xi^2\right) + \mathrm{e}^{2\kappa t}(\xi - 4\kappa\rho)^2} < \xi\left(1 - \mathrm{e}^{\kappa t}\right) + 4\kappa\rho\mathrm{e}^{\kappa t}$, which holds as soon as $\xi\left(1 - \mathrm{e}^{\kappa t}\right) + 4\kappa\rho\mathrm{e}^{\kappa t} > 0$, or $\rho > \frac{\xi}{4\kappa}\left(1 - \mathrm{e}^{-\kappa t}\right)$. Therefore the claim follows for any $\rho \geq \rho_+$ if and only if $\rho_+ > \frac{\xi}{4\kappa}\left(1 - \mathrm{e}^{-\kappa t}\right)$. This inequality simplifies to

$$\sqrt{\xi^2 + 16\kappa^2\mathrm{e}^{2\kappa t} - 2\xi^2\mathrm{e}^{\kappa t} + \xi^2\mathrm{e}^{2\kappa t}} > \frac{\xi\left(\mathrm{e}^{\kappa t} - 1\right)^2}{\mathrm{e}^{\kappa t} + 1},$$

which can be written as

$$\sqrt{\frac{4\mathrm{e}^{\kappa t}\left(4\kappa^2\mathrm{e}^{\kappa t}\left(\mathrm{e}^{\kappa t} + 1\right)^2 + \xi^2\left(\mathrm{e}^{\kappa t} - 1\right)^2\right)}{\left(\mathrm{e}^{\kappa t} + 1\right)^2} + \frac{\xi^2\left(\mathrm{e}^{\kappa t} - 1\right)^4}{\left(\mathrm{e}^{\kappa t} + 1\right)^2}} > \frac{\xi\left(\mathrm{e}^{\kappa t} - 1\right)^2}{\mathrm{e}^{\kappa t} + 1},$$

which is always true.

Now straightforward manipulations show that the inequality $u_+^* > 1$ is equivalent to

$$\sqrt{\left(\xi\left(\mathrm{e}^{\kappa t} - 1\right) + 4\kappa\rho\mathrm{e}^{\kappa t}\right)^2 - 16\kappa\mathrm{e}^{\kappa t}\left(\kappa + \xi\rho\left(\mathrm{e}^{\kappa t} - 1\right)\right)} > \xi\left(\mathrm{e}^{\kappa t} - 1\right) + 4\kappa\rho\mathrm{e}^{\kappa t},$$

which is true if $\rho < -\dfrac{\kappa}{\xi\left(\mathrm{e}^{\kappa t} - 1\right)}$ or $\rho < -\dfrac{\xi\left(1 - \mathrm{e}^{-\kappa t}\right)}{4\kappa}$. And of course the claim $\left(u_+^* > 1 \text{ if } \rho \leq \rho_-\right)$ shall hold if

$$\rho_- < -\frac{\kappa}{\xi\left(\mathrm{e}^{\kappa t} - 1\right)} \tag{5.35}$$

$$\text{or} \quad \rho_- < -\frac{\xi\left(1 - \mathrm{e}^{-\kappa t}\right)}{4\kappa}. \tag{5.36}$$

Consider (5.35). This inequality, which can be re-written as

$$-\sqrt{\frac{16\kappa^2\mathrm{e}^{3\kappa t}\left(\xi^2\left(\mathrm{e}^{\kappa t} - 1\right)^2\left(\mathrm{e}^{\kappa t} + 1\right) - 4\kappa^2\mathrm{e}^{\kappa t}\right)}{\xi^2\left(\mathrm{e}^{2\kappa t} - 1\right)^2} + \left(\frac{\xi^2(1 - \mathrm{e}^{\kappa t})(1 - \mathrm{e}^{2\kappa t}) + 8\kappa^2\mathrm{e}^{2\kappa t}}{\xi(\mathrm{e}^{\kappa t} + 1)(1 - \mathrm{e}^{\kappa t})}\right)^2} < \frac{\xi^2(1 - \mathrm{e}^{\kappa t})(1 - \mathrm{e}^{2\kappa t}) + 8\kappa^2\mathrm{e}^{2\kappa t}}{\xi(\mathrm{e}^{\kappa t} + 1)(1 - \mathrm{e}^{\kappa t})},$$

holds if $\xi^2\left(\mathrm{e}^{\kappa t} - 1\right)^2\left(\mathrm{e}^{\kappa t} + 1\right) - 4\kappa^2\mathrm{e}^{\kappa t} > 0$, or $\frac{\left(\mathrm{e}^{\kappa t} - 1\right)^2\left(1 + \mathrm{e}^{-\kappa t}\right)}{4} > \frac{\kappa^2}{\xi^2}$. Quick manipulations turn (5.36) into

$$-\sqrt{\frac{4\mathrm{e}^{\kappa t}\left(4\kappa^2\mathrm{e}^{\kappa t}\left(\mathrm{e}^{\kappa t} + 1\right)^2 - \xi^2\left(\mathrm{e}^{\kappa t} - 1\right)^2\left(2\mathrm{e}^{\kappa t} + 1\right)\right)}{\left(\mathrm{e}^{\kappa t} + 1\right)^2} + \frac{\xi^2\left(2\mathrm{e}^{\kappa t} - 3\mathrm{e}^{2\kappa t} + 1\right)^2}{\left(\mathrm{e}^{\kappa t} + 1\right)^2}} < \frac{\xi\left(2\mathrm{e}^{\kappa t} - 3\mathrm{e}^{2\kappa t} + 1\right)}{\mathrm{e}^{\kappa t} + 1}.$$

Again this trivially holds if $4\kappa^2\mathrm{e}^{\kappa t}\left(\mathrm{e}^{\kappa t}+1\right)^2-\xi^2\left(\mathrm{e}^{\kappa t}-1\right)^2\left(2\mathrm{e}^{\kappa t}+1\right)>0$, which is in turn equivalent to $\frac{\kappa^2}{\xi^2}>\frac{\left(\mathrm{e}^{\kappa t}-1\right)^2\left(2+\mathrm{e}^{-\kappa t}\right)}{4\left(\mathrm{e}^{\kappa t}+1\right)^2}$. Since the inequality $\frac{\left(\mathrm{e}^{\kappa t}-1\right)^2\left(2+\mathrm{e}^{-\kappa t}\right)}{4\left(\mathrm{e}^{\kappa t}+1\right)^2}<\frac{\left(\mathrm{e}^{\kappa t}-1\right)^2\left(1+\mathrm{e}^{-\kappa t}\right)}{4}$, is clearly true, it follows that for any valid choice of parameters either (5.35) or (5.36) (or both) hold, and the claim follows. □

Proposition 5.13. *Let $\varepsilon=\tau^{-1}$ and consider the large-maturity Heston forward process $\left(\tau^{-1}X_\tau^{(t)}\right)_{\tau>0}$ with X_t defined in* (3.1) *and $X_\tau^{(t)}$ defined in* (2.11). *Then $\mathcal{D}_{0,0}=\mathcal{D}_\infty$ and $\{0,1\}\subset\mathcal{D}_{0,0}^o$ with $\mathcal{D}_\infty$, $\mathcal{D}_{0,0}$ and $\mathcal{D}_{0,0}^o$ defined in* (3.10) *and in Assumption 2.1.*

Proof. We write

$$\mathbb{E}\left(\mathrm{e}^{u(X_{t+\tau}-X_t)}\right)=\mathbb{E}\left[\mathbb{E}\left(\mathrm{e}^{u(X_{t+\tau}-X_t)}|\mathcal{F}_t\right)\right]=\mathbb{E}\left(\mathrm{e}^{A(u,\tau)+B(u,\tau)V_t}\right)=\mathrm{e}^{A(u,\tau)}\mathbb{E}\left(\mathrm{e}^{B(u,\tau)V_t}\right).$$

For any fixed $t\geq 0$ we require that

$$\mathbb{E}\left(\mathrm{e}^{u(X_{t+\tau}-X_t)}|\mathcal{F}_t\right)<\infty\quad\text{for all}\quad\tau>0.\tag{5.37}$$

Andersen and Piterbarg [3, Proposition 3.1] proved that if the following conditions are satisfied

$$\kappa>\rho\xi u,\tag{5.38}$$

$$(\kappa-\rho\xi u)^2+u\left(1-u\right)\xi^2\geq 0,\tag{5.39}$$

then the explosion time is infinite and (5.37) is satisfied. In [20] the authors proved that these conditions are equivalent to $\kappa>\rho\xi$ and $u\in[u_-,u_+]$, with $u_-<0$ and $u_+>1$ ($u_\pm$ defined in (3.8)). Further we require that

$$\mathbb{E}\left(\mathrm{e}^{B(u,\tau)V_t}\right)<\infty,\quad\text{for all }\tau>0.\tag{5.40}$$

Now denote

$$\mathcal{D}_V:=\{u\in\mathbb{R}:\mathbb{E}\left(\mathrm{e}^{B(u,\tau)V_t}\right)<\infty,\ \text{for all }\tau>0\}.$$

Then the limiting forward mgf domain is given by $\mathcal{D}_\infty=[u_-,u_+]\cap\mathcal{D}_V$ and $\kappa>\rho\xi$. The condition (5.40) is equivalent to $B\left(u,\tau\right)<1/(2\beta_t)$ for all $\tau>0$, where

$$B\left(u,\tau\right):=\xi^{-2}\left(\kappa-\rho\xi u-d\left(u\right)\right)\frac{1-\exp\left(-d\left(u\right)\tau\right)}{1-\gamma\left(u\right)\exp\left(-d\left(u\right)\tau\right)},$$

and where d and γ are given in (3.13). Now a simple calculation gives $B\left(0,\tau\right)=0$ and $B\left(1,\tau\right)=0$ for all $\tau>0$. Furthermore for $u\in(0,1)$, and given Conditions (5.38) and (5.39), we have $d\left(u\right)>\kappa-\rho\xi u$ and $\gamma\left(u\right)<0$. This implies that $B\left(u,\tau\right)$ is strictly negative for $u\in(0,1)$ and $\tau>0$. In particular $[0,1]\subset\mathcal{D}_\infty$ (recall that the process is a martingale). For a fixed $u\in\mathbb{R}$ we calculate

$$\frac{\partial B\left(u,\tau\right)}{\partial\tau}=\frac{2u(u-1)d(u)^2\mathrm{e}^{d(u)\tau}}{\left(\kappa-\kappa\mathrm{e}^{d(u)\tau}+\xi\rho u\left(\mathrm{e}^{d(u)\tau}-1\right)-d(u)\left(\mathrm{e}^{d(u)\tau}+1\right)\right)^2},$$

so that for any $u\notin[0,1]$, $B\left(u,\cdot\right)$ is strictly increasing. Therefore the limiting domain is given by

$$\mathcal{D}_\infty=\left\{u\in\mathbb{R}:\lim_{\tau\to\infty}B\left(u,\tau\right)<\frac{1}{2\beta_t}\right\}\cap\left[u_-,u_+\right],$$

with $[0,1]\subset\mathcal{D}_\infty$. We first concentrate on the first condition and check at the end that our solution always contains $[0,1]$. We have $\lim_{\tau\to\infty}B\left(u,\tau\right)=\xi^{-2}\left(\kappa-\rho\xi u-d\left(u\right)\right)=V(u)/(\kappa\theta)$, with V defined in (3.12). So the condition is equivalent to $\kappa-\rho\xi u-d(u)<2\kappa/(1-\mathrm{e}^{-\kappa t})$. If $\rho\leq 0$ ($\rho\geq 0$) and $u\leq 0$ ($u\geq 0$) then

$$\kappa-\rho\xi u-d(u)\leq\kappa-\rho\xi u\leq\kappa<\frac{2\kappa}{1-\mathrm{e}^{-\kappa t}},$$

and the condition is always satisfied. So if $\rho = 0$ the domain is given by $[u_-, u_+]$. If $\rho < 0$ ($\rho > 0$), the domains contain $[u_-, 0]$ ($[0, u_+]$). Now suppose that $\rho < 0$ and $u > 0$. The condition above is equivalent to $V(u) < \kappa\theta/(2\beta_t)$. From Lemma 5.9, on $(0, u_+]$, the function V attains its maximum at u_+ with

$$V(u_+) = \frac{\kappa\theta\,(2\kappa - \rho\xi - \rho\eta)}{2\xi^2\,(1-\rho^2)}.$$

Using the properties in Lemma 5.9, there exists $u_+^* \in (1, u_+)$ solving the equation

$$\frac{V(u_+^*)}{\kappa\theta} = \frac{1}{2\beta_t}, \tag{5.41}$$

if and only if $(2\kappa - \rho\xi) - \rho\sqrt{\xi^2\,(1-\rho^2) + (2\kappa-\rho\xi)^2} > \xi^2\left(1-\rho^2\right)/\beta_t$. This condition has been shown in Lemma 5.10 to be equivalent to $-1 < \rho < \rho_-$ and $t > 0$. The solution to (5.41) has two roots u_-^* and u_+^* defined in (3.8), and the correct solution in this case is u_+^* since from Lemma 5.12 we know that $u_+^* > 1$ and $u_-^* < 0$. So if $\rho_- \le \rho < 0$ then the effective limiting domain is $[u_-, u_+]$. If $-1 < \rho < \rho_-$ and $t > 0$ then the effective limiting domain is given by $\left[u_-, u_+^*\right)$. Following a similar procedure we can show for $0 < \rho \le \min(\kappa/\xi, \rho_+)$ the effective limiting domain is given by $[u_-, u_+]$. If $\rho_+ < \rho < \min(\kappa/\xi, 1)$, $t > 0$ and $\kappa > \rho_+\xi$ then the effective limiting domain is given by $\left(u_-^*, u_+\right]$, with $u_- < u_-^* < 0$. □

Lemma 5.14. *The following expansion holds for the forward mgf $\Lambda_\tau^{(t)}$ defined in* (5.31)*:*

$$\Lambda_\tau^{(t)}(u) = V(u) + \frac{H(u)}{\tau}\left(1 + \mathcal{O}\left(\mathrm{e}^{-d(u)\tau}\right)\right), \quad \textit{for all } u \in \mathcal{D}_\infty, \textit{ as } \tau \textit{ tends to infinity,}$$

where the functions V, H, d and the interval $\mathcal{D}_\infty$ are defined in (3.12)*,* (3.13) *and* (3.10)*.*

Remark 5.15. For any $u \in \mathcal{D}_\infty$, $d(u) > 0$. Indeed $\mathcal{D}_\infty \subseteq [u_-, u_+]$ by Proposition 5.13. Furthermore since $\kappa > \rho\xi$, $u \in [u_-, u_+]$ implies (5.39) which in turn implies $d(u) > 0$.

Remark 5.16. We note in particular that the exponential decay in the remainder implies that for all $u \in \mathcal{D}_\infty$, $\Lambda_\tau^{(t)}(u) = V(u) + H(u)/\tau + \mathcal{O}\left(1/\tau^3\right)$ for τ large enough, which is used in the proof of Proposition 3.8.

Proof of Lemma 5.14. From the definition of $\Lambda_\tau^{(t)}$ in (5.31) and the Heston forward mgf given in (5.16) we immediately obtain the following asymptotics as τ tends to infinity:

$$A(u,\tau) = \tau V(u) - \frac{2\kappa\theta}{\xi^2}\log\left(\frac{1}{1-\gamma(u)}\right) + \mathcal{O}\left(\mathrm{e}^{-d(u)\tau}\right), \qquad B(u,\tau) = \frac{V(u)}{\kappa\theta} + \mathcal{O}\left(\mathrm{e}^{-d(u)\tau}\right),$$

where A and B are defined in (5.17), V in (3.12) and d and γ in (3.13). In particular this implies that

$$\frac{B(u,\tau)}{1-2\beta_t B(u,\tau)} = \frac{V(u)}{\theta\kappa - 2\beta_t V(u)} + \mathcal{O}\left(\mathrm{e}^{-d(u)\tau}\right),$$
$$\log\left(1 - 2\beta_t B(u,\tau)\right) = \log\left(1 - \frac{2\beta_t V(u)}{\theta\kappa}\right) + \mathcal{O}\left(\mathrm{e}^{-d(u)\tau}\right),$$

which are well-defined for all $u \in \mathcal{D}_\infty$. We therefore obtain

$$H(u) = \frac{V(u)}{\kappa\theta - 2\beta_t V(u)} v\mathrm{e}^{-\kappa t} - \frac{2\kappa\theta}{\xi^2}\log\left(1 - \frac{2\beta_t V(u)}{\kappa\theta}\right) - \frac{2\kappa\theta}{\xi^2}\log\left(\frac{1}{1-\gamma(u)}\right),$$

and the lemma follows from straightforward simplifications. □

5.3. **Proofs of Section 3.2.** For ease of notation here, the subscript i for a function or a variable shall refer to the i-th Heston model, for which we can readily use the results of Section 3.1. For instance the function Λ_i shall refer to the function Λ evaluated using the parameters in the i-th Heston model.

Lemma 5.17. *In the n-dimensional multivariate Heston model* (3.15) *the forward mgf defined in* (5.15) *reads*

$$\Lambda(u) = \log \mathbb{E}\left(\mathrm{e}^{uX_\tau^{(t)}}\right) = \sum_{i=1}^n \Lambda_i(u), \quad \textit{for all } u \in \bigcap_{i=1}^n \mathcal{D}_{\Lambda_i},$$

where Λ_i and $\mathcal{D}_{\Lambda_i}$ are defined in (5.16).

Proof. Conditioning on the filtration $(\mathcal{F}_t)_{t\geq 0}$ and using the tower property, we obtain

$$\Lambda(u) = \log \mathbb{E}\left(\mathbb{E}\left(\mathrm{e}^{uX_\tau^{(t)}}|\mathcal{F}_t\right)\right) = \sum_{i=1}^n A_i(u,\tau) + \log \mathbb{E}\left(\mathrm{e}^{\sum_{i=1}^n B_i(u,\tau)V_t^{(i)}}\right),$$

using the mgf of the standard Heston model. By independence of the variance processes, we obtain

$$\Lambda(u) = \sum_{i=1}^n A_i(u,\tau) + \log \mathbb{E}\left(\mathrm{e}^{\sum_{i=1}^n B_i(u,\tau)V_t^{(i)}}\right) = \sum_{i=1}^n \Lambda_i(u), \quad \text{for all } u \in \bigcap_{i=1}^n \mathcal{D}_{\Lambda_i}.$$

□

5.4. **Proofs of Section 3.3.** Consider the following functions:

(5.42)

$$A(u,\tau) := A_1(u,\tau) + \frac{2\kappa^2\theta^2(\chi(u)-d(u))}{d(u)^3\xi^2}\left(\chi(u)\left(d(u)\tau-2\right) + d(u)\left(d(u)\tau-1\right) + 2\mathrm{e}^{-d(u)\tau}\frac{2\chi(u) + \frac{d(u)^2-2\chi(u)^2}{\chi(u)+d(u)}\mathrm{e}^{-d(u)\tau}}{1-\gamma(u)\mathrm{e}^{-2d(u)\tau}}\right),$$

$$A_1(u,\tau) := \frac{1}{2}\left(\chi(u) - d\left(u\right)\right)\tau - \frac{1}{2}\log\left(\frac{1-\gamma\left(u\right)\exp\left(-2d\left(u\right)\tau\right)}{1-\gamma\left(u\right)}\right),$$

$$B_1(u,\tau) := \frac{4\kappa\theta}{\xi^2}\frac{\chi(u)-d(u)}{d(u)}\frac{\left(1-\exp\left(-d(u)\tau\right)\right)^2}{1-\gamma(u)\exp\left(-2d(u)\tau\right)}, \quad B_2(u,\tau) := \frac{2(\chi(u)-d(u))}{\xi^2}\frac{1-\exp\left(-2d\left(u\right)\tau\right)}{1-\gamma\left(u\right)\exp\left(-2d\left(u\right)\tau\right)},$$

and

$$\text{(5.43)} \qquad M(r,p,q) := \frac{1}{2}\left(\frac{p^2r^2}{1-2rq} - \log\left(1-2rq\right)\right), \quad \beta_t := \frac{\xi^2}{8\kappa}\left(1-\mathrm{e}^{-2\kappa t}\right), \quad \mu_t := \sqrt{v}\mathrm{e}^{-\kappa t} + \theta\left(1-\mathrm{e}^{-\kappa t}\right),$$

$$d(u) := \left(\chi(u)^2 + (1-u)\,\frac{u}{4}\xi^2\right)^{1/2}, \quad \gamma(u) := \frac{\chi(u)-d\left(u\right)}{\chi(u)+d\left(u\right)} \quad \text{and} \quad \chi(u) := \kappa - \frac{\rho\xi u}{2}.$$

Note that although in some cases we use the same variables and functions as in the Heston analysis they may have a different definition in this section. In our analysis we require the following lemma which is a direct consequence of [4, Equation 29.6].

Lemma 5.18. *If $Z \sim \mathcal{N}(0,1)$ and p and q are two constants, then, for M defined in* (5.43),

$$\log \mathbb{E}\left(\mathrm{e}^{u\left(pZ+qZ^2\right)}\right) = M(u,p,q), \qquad \textit{whenever } uq < 1/2.$$

Lemma 5.19. *In the Schöbel-Zhu model* (3.18) *the forward mgf defined in* (5.15) *is given by*

$$\Lambda(u) = A(u,\tau) + B_1(u,\tau)\mu_t + B_2(u,\tau)\mu_t^2 + M\left(1, B_1(u,\tau)\sqrt{\beta_t} + 2B_2(u,\tau)\sqrt{\beta_t}\mu_t, B_2(u,\tau)\beta_t\right),$$

for all $u \in \mathcal{D}_\Lambda$, where all the functions and variables are defined in (5.42) *and* (5.43).

Proof. Conditioning on the filtration $(\mathcal{F}_t)_{t\geq 0}$ and using the tower property we find

$$\Lambda(u) = \log \mathbb{E}\left[\mathbb{E}\left(\mathrm{e}^{uX_\tau^{(t)}}|\mathcal{F}_t\right)\right] = A(u,\tau) + \log \mathbb{E}\left[\exp\left(B_1(u,\tau)\sigma_t + B_2(u,\tau)\sigma_t^2\right)\right],$$

where we have used the standard Schöbel-Zhu mgf from [37] and the functions defined in (5.42). Since σ_t is Gaussian with mean μ_t and variance β_t (given in (5.43)), we obtain

$$\begin{aligned}\Lambda(u) &= A(u,\tau) + \log \mathbb{E}\left(\mathrm{e}^{B_1(u,\tau)\sigma_t + B_2(u,\tau)\sigma_t^2}\right)\\ &= A(u,\tau) + B_1(u,\tau)\mu_t + B_2(u,\tau)\mu_t^2 + \log \mathbb{E}\left(\mathrm{e}^{\left(B_1(u,\tau)\sqrt{\beta_t}+2B_2(u,\tau)\sqrt{\beta_t}\mu_t\right)Z + (B_2(u,\tau)\beta_t)Z^2}\right),\end{aligned}$$

with $Z \sim \mathcal{N}(0,1)$, and the lemma follows directly from Lemma 5.18. □

5.5. **Proofs of Section 3.4.** Let ϕ be the Lévy exponent of the Lévy process Y. If v follows (3.21), by a straightforward application of the tower property for expectations, the forward mgf defined in (5.15) is given by

$$\Lambda(u) = A(\phi(u),\tau) + \frac{B(\phi(u),\tau)}{1-2\beta_t B(\phi(u),\tau)} v\mathrm{e}^{-\kappa t} - \frac{2\kappa\theta}{\xi^2}\log\left(1 - 2\beta_t B(\phi(u),\tau)\right), \quad \text{for all } u \in \mathcal{D}_\Lambda, \tag{5.44}$$

where

$$A(u,\tau) := \frac{\kappa\theta}{\xi^2}\left(\left(\kappa - d(u)\right)\tau - 2\log\left(\frac{1-\gamma(u)\exp\left(-d(u)\tau\right)}{1-\gamma(u)}\right)\right), \tag{5.45}$$

$$B(u,\tau) := \frac{\kappa - d(u)}{\xi^2}\frac{1-\exp\left(-d(u)\tau\right)}{1-\gamma(u)\exp\left(-d(u)\tau\right)},$$

$$d(u) := \left(\kappa^2 - 2u\xi^2\right)^{1/2}, \quad \gamma(u) := \frac{\kappa - d(u)}{\kappa + d(u)} \quad \text{and} \quad \beta_t := \frac{\xi^2}{4\kappa}\left(1 - \mathrm{e}^{-\kappa t}\right). \tag{5.46}$$

Similarly if $(v_t)_{t\geq 0}$ follows (3.22) the forward mgf defined in (5.15) is given by

$$\Lambda(u) = A(\phi(u),\tau) + B(\phi(u),\tau)v\mathrm{e}^{-\lambda t} + d\log\left(\frac{B(\phi(u),\tau) - \mathrm{e}^{t\lambda}\alpha}{\mathrm{e}^{t\lambda}(B(\phi(u),\tau)-\alpha)}\right), \quad \text{for all } u \in \mathcal{D}_\Lambda, \tag{5.47}$$

where

$$A(u,\tau) := \frac{\lambda d}{\alpha\lambda - u}\left[u\tau + \alpha\log\left(1 - \frac{u}{\alpha\lambda}\left(1 - \mathrm{e}^{-\lambda\tau}\right)\right)\right] \quad \text{and} \quad B(u,\tau) := \frac{u}{\lambda}\left(1 - \mathrm{e}^{-\lambda\tau}\right). \tag{5.48}$$

Proof of Proposition 3.19. We show that Proposition 2.14 is applicable given the assumptions of Proposition 3.19. Consider case (i). The expansion for $\Lambda_\tau^{(t)}$ defined in (5.31) is straightforward and analogous to Lemma 5.14. In particular we establish that

$$\Lambda_\tau^{(t)}(u) = \overline{V}(u) + \frac{\overline{H}(u)}{\tau}\left(1 + \mathcal{O}\left(\mathrm{e}^{-d(u)\tau}\right)\right), \quad \text{for all } u \in \widehat{\mathcal{D}}_\infty, \text{ as } \tau \text{ tends to infinity},$$

where the functions $\overline{V}$, $\overline{H}$, d and the domain $\widehat{\mathcal{D}}_\infty$ are defined in (3.23), (5.46) and (3.25). Since ϕ is essentially smooth and strictly convex on $\mathcal{D}_\phi$ and $\widehat{\mathcal{D}}_\infty \subseteq \mathcal{D}_\phi$, then the limiting mgf $\Lambda_{0,0} = \overline{V}$ is essentially smooth and strictly convex on $\widehat{\mathcal{D}}_\infty$. Also $\Lambda_\tau^{(t)}$ is infinitely differentiable on $\widehat{\mathcal{D}}_\infty$ since ϕ is of class $\mathcal{C}^\infty$ on $\widehat{\mathcal{D}}_\infty$. Since $\phi(1) = 0$ we have that $\overline{V}(1) = 0$ and $\{0,1\} \subset \widehat{\mathcal{D}}_\infty^o$. It remains to be checked that the limiting domain is in fact given by $\widehat{\mathcal{D}}_\infty$. We first note that that by conditioning on $(V_u)_{t\leq u\leq t+\tau}$ and using the independence of the time-change and the Lévy process we have $\mathbb{E}\left(\mathrm{e}^{u(X_{t+\tau}-X_t)}\right) = \mathbb{E}\left(\mathrm{e}^{\phi(u)\int_t^{t+\tau} v_s \mathrm{d}s}\right)$ and so any u in the limiting domain must satisfy $\phi(u) < \infty$. Using [14, page 476] and the tower property we compute

$$\mathbb{E}\left(\mathrm{e}^{u(X_{t+\tau}-X_t)}\right) = \mathbb{E}\left[\mathbb{E}\left(\mathrm{e}^{\phi(u)\int_t^{t+\tau} v_s \mathrm{d}s}|\mathcal{F}_t\right)\right] = \mathbb{E}\left(\mathrm{e}^{A(\phi(u),\tau)+B(\phi(u),\tau)v_t}\right) = \mathrm{e}^{A(\phi(u),\tau)}\mathbb{E}\left(\mathrm{e}^{B(\phi(u),\tau)v_t}\right), \tag{5.49}$$

with A and B defined in (5.45). Further from (5.18) we have

$$\log \mathbb{E}\left(\mathrm{e}^{uv_t}\right) = \frac{uv\mathrm{e}^{-\kappa t}}{1-2\beta_t u} - \frac{2\kappa\theta}{\xi^2}\log\left(1-2\beta_t u\right), \qquad \text{for all } u < \frac{1}{2\beta_t}.$$

Following a similar argument to the proof of Proposition 5.13 we can show that for any $t \geq 0$, $B(\phi(u),\tau) < 1/(2\beta_t)$ is always satisfied for each $\tau > 0$. This follows from the independence of the Lévy process Y and the time-change. We also require that for any $t \geq 0$, $\mathbb{E}\left(\mathrm{e}^{\int_t^{t+\tau} v_s \mathrm{d}s \phi(u)} | \mathcal{F}_t\right) < \infty$, for every $\tau > 0$. Here we use [3][Corollary 3.3] with zero correlation to find that we require $\phi(u) \leq \kappa^2/(2\xi^2)$. It follows that $\widehat{\mathcal{D}}_\infty = \left\{u : \phi(u) \leq \kappa^2/(2\xi^2)\right\}$.

Regarding case (ii), arguments analogous to case (i) hold and we focus on showing that the limiting domain is $\widetilde{\mathcal{D}}_\infty$. Using [14, page 488] Equality (5.49) also holds with A and B defined in (5.48). Since we require that for any $t \geq 0$, $\mathbb{E}\left(\mathrm{e}^{\int_t^{t+\tau} v_s \mathrm{d}s \phi(u)} | \mathcal{F}_t\right) < \infty$, for every $\tau > 0$ we have $\phi(u) < \alpha\lambda$. Using [14, page 482] we also have

$$\log \mathbb{E}\left(\mathrm{e}^{uv_t}\right) = uv\mathrm{e}^{-\lambda t} + d\log\left(\frac{u-\alpha\mathrm{e}^{\lambda t}}{(u-\alpha)\mathrm{e}^{\lambda t}}\right), \qquad \text{for all } u < \alpha.$$

But it is straightforward to show that $\phi(u) < \alpha\lambda$ implies $B(\phi(u),\tau) < \alpha$ for any $\tau > 0$ and it follows that $\widetilde{\mathcal{D}}_\infty = \{u : \phi(u) < \alpha\lambda\}$. Case (iii) is straightforward and omitted. □

References

[1] E. Alòs, J. León and J. Vives. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance & Stochastics*, 11: 571-589, 2007.

[2] L.B.G. Andersen and A. Lipton. Asymptotics for exponential Lévy processes and their volatility smile: survey and new results. Preprint available at arXiv:1206.6787, 2012.

[3] L. Andersen and V. Piterbarg. Moment Explosions in Stochastic Volatility Models. *Finance & Stochastics*, 11(1): 29-50, 2007.

[4] N. Balakrishnan, N.L. Johnson and S. Kotz. Continuous Univariate Distributions, Volume 2. Wiley, Second Edition, 1994.

[5] S. Benaim and P. Friz. Smile asymptotics II: models with known moment generating functions. *Journal of Applied Probability*, 45: 16-32, 2008.

[6] S. Benaim and P. Friz. Regular Variation and Smile Asymptotics. *Mathematical Finance*, 19 (1): 1-12, 2009.

[7] G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Annales Scientifiques de l'Ecole Normale Supérieure*, 4 (21): 307-331, 1988.

[8] B. Bercu and A. Rouault. Sharp large deviations for the Ornstein-Uhlenbeck process. *Theory Probab. Appl.*, 46 (1): 1-19, 2002.

[9] H. Berestycki, J. Busca, and I. Florent. Computing the implied volatility in stochastic volatility models. *Communications on Pure and Applied Mathematics*, 57 (10): 1352-1373, 2004.

[10] L. Bergomi. Smile Dynamics I. *Risk*, September, 2004.

[11] P. Beyer and J. Kienitz. Pricing Forward Start Options in models based on (time-changed) Lévy processes. Available at dx.doi.org/10.2139/ssrn.1319703, 2008.

[12] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81 (3): 637-659, 1973.

[13] H. Bühler. Applying Stochastic Volatility Models for Pricing and Hedging Derivatives. Available at quantitative-research.de/dl/021118SV.pdf, 2002.

[14] R. Cont and P. Tankov. Financial Modelling with Jump Processes. Chapman and Hall/CRC, 2007.

[15] A. Dembo and O. Zeitouni. Large deviations techniques and applications. Jones and Bartlet Publishers, Boston, 1993.

[16] J.D. Deuschel, P.K. Friz, A. Jacquier and and S. Violante. Marginal density expansions for diffusions and stochastic volatility, Part II: Applications. Forthcoming in *Communications on Pure and Applied Mathematics*, 2012.

[17] J. Figueroa-López, M. Forde and A. Jacquier. The large-maturity smile and skew for exponential Lévy models. Preprint, 2011.

[18] J. Figueroa-López, R. Gong and C. Houdré. Small-time expansions of the distributions, densities, and option prices of stochastic volatility models with Lévy jumps. *Stochastic Processes and their Applications*, 122: 1808-1839, 2012.

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

[20] M. Forde and A. Jacquier. The large-maturity smile for the Heston model. *Finance & Stochastics*, 15 (4): 755-780, 2011.

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

[22] M. Forde, A. Jacquier and A. Mijatović. Asymptotic formulae for implied volatility in the Heston model. *Proceedings of the Royal Society A*, 466 (2124): 3593-3620, 2010.

[23] J.P. Fouque, G. Papanicolaou, R. Sircar and K. Solna. Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives. CUP, 2011.

[24] P. Friz, S. Gerhold, A. Gulisashvili and S. Sturm. Refined implied volatility expansions in the Heston model. *Quantitative Finance*, 11 (8): 1151-1164, 2011.

[25] K. Gao and R. Lee. Asymptotics of Implied Volatility to Arbitrary Order. SSRN:1768383, 2011.

[26] J. Gatheral. The Volatility Surface: A Practitioner's Guide. John Wiley & Sons, 2006.

[27] J. Gatheral and A. Jacquier. Convergence of Heston to SVI. *Quantitative Finance*, 11 (8): 1129-1132, 2011.

[28] P. Glasserman and Q. Wu. Forward and Future Implied Volatility. *Internat. Journ. of Theor. and App. Fin.*, 14(3), 2011.

[29] R.R. Goldberg. Fourier Transforms. CUP, 1970.

[30] A. Gulisashvili. Asymptotic formulas with error estimates for call pricing functions and the implied volatility at extreme strikes. *SIAM Journal on Financial Mathematics*, 1: 609-641, 2010.

[31] A. Gulisashvili and E. Stein. Asymptotic Behavior of the Stock Price Distribution Density and Implied Volatility in Stochastic Volatility Models. *Applied Mathematics & Optimization*, 61 (3): 287-315, 2010.

[32] P. Henry-Labordère. Analysis, geometry and modeling in finance. Chapman and Hill/CRC, 2008.

[33] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2): 327-342, 1993.

[34] G. Hong. Forward Smile and Derivative Pricing. globalriskguard.com/resources/deriv/fwd3.pdf, 2004.

[35] A. Jacquier and A. Mijatović. Large deviations for the extended Heston model: the large-time case. arXiv:1203.5020, 2012.

[36] A. Jacquier, M. Keller-Ressel and A. Mijatović. Implied volatility asymptotics of affine stochastic volatility models with jumps. Forthcoming in *Stochastics*, 2012.

[37] C. Kahl and R. Lord. Complex Logarithms in Heston-Like Models. *Mathematical Finance*, 20 (4): 671-694, 2010.

[38] I. Karatzas and S.E. Shreve. Brownian Motion and Stochastic Calculus. Springer-Verlag, 1997.

[39] S. Karlin and H. Taylor. A Second Course in Stochastic Processes. Academic Press, 1981.

[40] M. Keller-Ressel. Moment Explosions and Long-Term Behavior of Affine Stochastic Volatility Models. *Mathematical Finance*, 21 (1): 73-98, 2011.

[41] R.W. Lee. Option Pricing by Transform Methods: Extensions, Unification and Error Control. *Journal of Computational Finance*, 7 (3): 51-86, 2004.

[42] R.W. Lee. The Moment Formula for Implied Volatility at Extreme Strikes. *Mathematical Finance*, 14 (3), 469-480, 2004.

[43] V. Lucic. Forward-start options in stochastic volatility models. *Wilmott Magazine*, September, 2003.

[44] R. Merton. The Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science*, 4(1): 141-183, 1973.

[45] A. Mijatović and P. Tankov. A new look at short-term implied volatility in asset price models with jumps. arXiv:1207.0843, 2012.

[46] J. Muhle-Karbe and M. Nutz. Small-time asymptotics of option prices and first absolute moments. *Journal of Applied Probability*, 48: 1003-1020, 2011.

[47] M. Rubinstein. Pay now, choose later. *Risk*, 4(13), 1991.

[48] R. Schöbel and J. Zhu. Stochastic volatility with an Ornstein-Uhlenbeck process: an extension. *European Finance Review*, 3(1): 23-46, 1999.

[49] E. Stein and J. Stein. Stock-price distributions with stochastic volatility - an analytic approach. *Review of Financial studies*, 4(4): 727-752, 1991.

[50] P. Tankov. Pricing and hedging in exponential Lévy models: review of recent results. *Paris-Princeton Lecture Notes in Mathematical Finance*, Springer, 2010.

[51] M. R. Tehranchi. Asymptotics of implied volatility far from maturity. *Journal of Applied Probability*, 46: 629-650, 2009.

[52] D. Williams. Probability With Martingales. CUP, 1991.

Department of Mathematics, Imperial College London

E-mail address: ajacquie@imperial.ac.uk, p.roome11@imperial.ac.uk