Small-time Smile for the Multi-factor Volatility Heston Model and its Applications in the Pension Fund Management*

Kyoung-Kuk Kim†

Korea Advanced Institute of Science and Technology

and

Younghoon Kim‡

University of North Carolina at Chapel Hill

Oct 2018

Abstract

Black-Scholes implied volatilities are one important ingredient in the derivatives markets. When it comes to manufacturing and managing of financial products, a production model needs to be calibrated to market implied volatilities in order to find model parameters that could reproduce option prices data. In this procedure, it has been long known that an asymptotic expansion of model implied volatilities can be very useful rather than brute force matching of model option prices with market data. In this paper, we extend the existing small-time asymptotics for implied volatilities under the Heston stochastic volatility model to the multi-factor volatility Heston model, which is also known as the Wishart multi-dimensional stochastic volatility model. More explicitly, we show that the approaches taken in Forde and Jacquier [11] and Forde et al. [13] are applicable to the WMSV model under mild conditions, and obtain explicit small-time expansions of implied volatilities. Our results open new possibilities for the use of multi-factor stochastic volatilities, and they are potentially useful for understanding the behaviors of option portfolios.

Keywords: implied volatility; Wishart process; asymptotic expansion; large deviation; Laplace expansion;

---

*This work was supported by the 2018 Academic Research Support Program of the Korea Derivative Association funded by the National Pension Research Institute.

†Corresponding author, Industrial and Systems Engineering, E-mail: kkim3128@kaist.ac.kr

‡Statistics and Operations Research, E-mail: yhkim0225@unc.edu

1 Introduction

Multi-factor volatility Heston model or Wishart multidimensional stochastic volatility (WMSV) model has received much attention in the quantitative finance community. Since Duffie et al. [10], the general theory of affine processes has been developed rapidly, which provides powerful analytical tools in asset pricing via the transform analysis. In the relevant literature, multiple volatility factors have played key roles. Stochastic volatility based on Wishart process was initially suggested by Gourieroux and Sufana [19], while maintaining computational convenience. Such a modeling approach appears to be useful and effective as its enlarged flexibility helps explain stylized facts observed in the markets. Many studies such as Chrisoffersen et al. [4] or Gruber et al. [21] support the strength of multiple stochastic volatility modeling empirically. The model proposed by Da Fonseca et al. [8] is a successful modeling specification in this direction in that their model captures the essential ingredients of multivariate volatility factors and provides an explicit and tractable analytical framework.

In this paper, we study the small-time asymptotics of implied volatilities for the model of Da Fonseca et al. [8]. Asymptotic expansions of implied volatilities have been studied for more than a decade due to their practical usage for model calibration. This procedure of matching model parameters to observed market data is performed on a regular basis, and reliable and fast operations are needed for successful model implementation in practice. For this reason, if there is a closed form of implied volatilities, then it could facilitate the calibration procedure and could provide beneficial insights about model behaviors. There is indeed a large literature on implied volatility expansions for a wide range of models, including Lee [25] one of early works in this area. Particularly for the Heston model and its variants, small and large maturity smiles are extensively studied by Forde and Jacquier [11, 12], Forde et al. [13], etc. For a more complete historical account of volatility expansion, we refer the reader to the previously mentioned references.

On the other hand, when it comes to the multi-dimensional version of the Heston model, resources are somewhat limited. Benabid et al. [1] embark on the study of volatility asymptotics in small-time inspired by Fouque et al. [15, 16]. Da Fonseca and Grasselli [7] report the form of limiting implied volatility in terms of the so called vol-of-vol scale factor α as well as asymptotics for the biHeston model and the Wishart affine stochastic correlation model. Compared to these approaches based on perturbation method, we take a different route which helps us avoid the use of extra parameter α. Additionally, we derive a correction term which accounts for the small but nonzero maturity. Essentially, ours are based on the developments of asymptotic expansions taken by Forde and Jacquier [11] and Forde et al. [13].

Our main contributions can be summarized as follows. First, the large deviations approach which has been quite popular for asymptotic expansions is proved applicable to the WMSV model. For this, the explicit form of the limiting cumulant generating function of the log stock price is derived for which a similar idea can be found in Gauthier and Possamai [17]. One could utilize the results of Gnoatto and Grasselli [18] on the explicit moment generating function under the WMSV model which works for any time t, but our approach imposes a weaker condition. Second, Laplace type expansion is shown to be applicable to the WMSV model as done for the Heston model [13]. With a slightly more stringent condition on model parameters, we obtain an explicit and higher order expansion of the cumulant generating function. And this helps a better fit to the true implied volatility curve.

The structure of this paper is as follows. In Section 2, we introduce the WMSV model first, and derive the limiting cumulant generating function based on which the large deviation principle is stated. In the last subsection, the expansion of the limiting implied volatility is computed. Section 3 then reports a more elaborate form of the cumulant generating function under a small time scale so that we can derive the explicit form of small-time smiles. The performance of the resulting formulae is examined in Section 4. Section 6 concludes.

2 Limiting Implied Volatilities

2.1 Wishart multifactor stochastic volatility process

Da Fonseca et al. [8] suggest that in an arbitrage-free frictionless financial market, the Wishart multidimensional stochastic volatility model (WMSV) can represent a dynamics of a return of a risky asset price S_t,

$$\begin{aligned} \mathrm{d}Y_t &= -\frac{1}{2}\mathrm{Tr}\left[\Sigma_t\right]\mathrm{d}t + \mathrm{Tr}\left[\sqrt{\Sigma_t}\left(\mathrm{d}W_t R^\top + \mathrm{d}B_t\sqrt{\mathbb{I} - RR^\top}\right)\right], \\ \mathrm{d}\Sigma_t &= \left(\Omega\Omega^\top + M\Sigma_t + \Sigma_t M^\top\right)\mathrm{d}t + \sqrt{\Sigma_t}\mathrm{d}W_t Q + Q^\top\left(\mathrm{d}W_t\right)^\top\sqrt{\Sigma_t}, \end{aligned} \tag{1}$$

where $Y_t = \ln S_t$, $\mathbb{I}$ is the n-dimensional identity matrix, $\Omega, R, M, Q \in \mathcal{M}_n$ (the set of square matrices), and $W_t,\ B_t \in \mathcal{M}_n$ are composed of n^2 independent Brownian motions under the risk-neutral measure. Here, the risk-free rate r is assumed to be zero without loss of any generality. As noted in Da Fonseca et al [8], (1) represents the matrix analogue of the square-root diffusion of the Heston model [22]. The random shocks to the stock return and the volatility process are correlated by the matrix R.

There are additional assumptions on model parameters. The matrix M is negative semi-definite, which is related to the mean-reverting feature in the typical volatility modeling practice. It is also assumed that Q is nonsingular and $\Omega\Omega^\top = \beta Q^\top Q$ with real parameter $\beta > n-1$. These standard conditions ensure that the resulting symmetric matrix Σ_t is positive semi-definite. For the rest of this paper, such conditions are treated as part of the WMSV model.

The usefulness of the WMSV model lies in that the transform techniques widely studied in the literature, including Duffie et al. [10], are still applicable. The transform formula is well documented in Da Fonseca et al. [6]. We here record a version that fits our purpose: for a scalar p,

$$\ln \mathsf{E}\left[e^{pY_t}\right] = pY_0 + \mathrm{Tr}\left[A(p,t)\Sigma_0\right] + b(p,t). \tag{2}$$

Here $A(p,t)$ is a solution to the matrix Riccati differential equation given as

$$\frac{\mathrm{d}A}{\mathrm{d}t} = A\left(M + pQ^\top R^\top\right) + \left(M^\top + pRQ\right)A + 2AQ^\top QA + \frac{p(p-1)}{2}\mathbb{I} \tag{3}$$

with the initial condition $A(p,0) = \mathbf{0}$. Lastly, $b(p,t) = \beta\int_0^t \mathrm{Tr}\left[Q^\top QA(p,s)\right]\mathrm{d}s$. There are many results in the literature regarding solutions to the above differential equation. For completeness, we provide a statement that is used in the next section.

Proposition 1 *Consider a matrix differential equation for a fixed p:*

$$\frac{\mathrm{d}}{\mathrm{d}s}\begin{pmatrix} G & F \end{pmatrix} = \begin{pmatrix} G & F \end{pmatrix}\begin{pmatrix} M + pQ^\top R^\top & -2Q^\top Q \\ \frac{p(p-1)}{2}\mathbb{I} & -\left(M^\top + pRQ\right) \end{pmatrix}$$

with $G(p,0) = 0$ and $F(p,0) = \mathbb{I}$. As long as $F(p,t)$ is invertible, (3) *has a solution $A(p,s) = F(p,s)^{-1}G(p,s)$ on $[0,t]$. Conversely, if* (3) *has a solution $A(p,s)$ on $[0,t]$, then $F(p,s)$ is invertible on $[0,t]$ so that $A(p,s) = F(p,s)^{-1}G(p,s)$. In such a case, $b(p,t)$ is simplified to*

$$b(p,t) = -\frac{\beta}{2}\mathrm{Tr}\left[\ln F(p,t) + t\left(M + pQ^\top R^\top\right)\right].$$

The first statement is already noted in Da Fonseca et al. [8] and Da Fonseca [5] where full details are referred to Grasselli and Tebaldi [20]. The second statement is a slight modification of a result reported in Choi [3] and Reid [27]. Hence, we have the equivalence between the existence of A and the non-singularity of $F(p,t)$. In addition, the validity of the affine transform formula (2) for general affine processes is proved in Keller-Ressel and Mayerhofer [24], that is, the region for p in which exponential moments of Y_t exist coincides with the region for p such that the solution to (3) exists at a given time t. Consequently, we see that the blow-up condition of exponential moments at time t is specified as $\det F(p,t) = 0$.

2.2 Large deviations principle

A large deviations principle for the Heston stochastic volatility model is nicely derived in Forde and Janquier [11]. The authors compute the limiting cumulant generating function (CGF)

$$\Lambda(p) = \lim_{t\downarrow 0} t \ln \mathsf{E}\left[e^{p(Y_t - Y_0)/t}\right]$$

on some open interval (p_-, p_+) including 0. Subsequently, its Fenchel-Legendre transform $\Lambda^*(x) = \sup_{p\in(p_-,p_+)}\{px - \Lambda(p)\}$ is shown to determine the small-time limit of implied volatilities under the Heston model.

Lemma 1 *Suppose that RQ is symmetric in the WMSV model. We further assume that $\mathbb{I} - R^\top R$ is invertible. Then, the limiting cumulant generating function $\Lambda(p)$ is given by*

$$\Lambda(p) = \mathrm{Tr}\left[\left(\cos(p\Gamma) - \Gamma^{-1}\sin(p\Gamma)RQ\right)^{-1}\left(\frac{p}{2}\Gamma^{-1}\sin(p\Gamma)\right)\Sigma_0\right]$$

on some open interval (p_-, p_+) which includes 0. Here, Γ is a square root of $Q^\top(\mathbb{I} - R^\top R)Q$. The scalars $p_\pm$ are the largest negative and the smallest positive numbers such that

$$\det\left(\cos(p\Gamma) - \Gamma^{-1}\sin(p\Gamma)RQ\right) = 0.$$

Proof: By definition, the target function is given as follows:

$$\begin{aligned}
\Lambda(p) &= \lim_{t\to 0} t\ln \mathsf{E}\left[e^{p(Y_t-Y_0)/t}\right] \\
&= \lim_{t\to 0}\left\{\mathrm{Tr}\left[tA\left(\frac{p}{t},t\right)\Sigma_0\right] + tb\left(\frac{p}{t},t\right)\right\} \\
&= \mathrm{Tr}\left[\lim_{t\to 0} tA\left(\frac{p}{t},t\right)\Sigma_0\right] + \lim_{t\to 0} tb\left(\frac{p}{t},t\right)
\end{aligned}$$

provided the limits exist. We compute those limits by looking at the solution to the matrix differential equation in Proposition 1. It is easy to see that the solution is given as

$$\begin{pmatrix} G\left(\frac{p}{t},t\right) & F\left(\frac{p}{t},t\right)\end{pmatrix} = \begin{pmatrix}\mathbf{0} & \mathbb{I}\end{pmatrix} e^{t\mathsf{M}+\mathsf{N}_t}, \quad \mathsf{M} = \begin{pmatrix} M & \mathbf{0} \\ \mathbf{0} & -M^\top\end{pmatrix}, \mathsf{N}_t = \begin{pmatrix} p\mathsf{A} & t\mathsf{B} \\ \mathsf{c}(t)\mathbb{I} & -p\mathsf{A}\end{pmatrix}$$

where $\mathsf{A} = Q^\top R^\top = RQ$, $\mathsf{B} = -2Q^\top Q$, and $\mathsf{c}(t) = \frac{p(p-t)}{2t}$. For notational simplicity, we simply write c if there is no risk of confusion.

We will later show that M does not affect the limit as t decreases to zero. Provided this is valid, it is enough to compute $\lim_{t\to 0} e^{\mathsf{N}_t}$. Straightforward computations yield that

$$\mathsf{N}_t^{2k} = \begin{pmatrix}\mathsf{Z}_t^k & \Pi_k \\ \mathbf{0} & \mathsf{Z}_t^k\end{pmatrix}, \quad \mathsf{N}_t^{2k+1} = \begin{pmatrix} * & * \\ \mathsf{c}\mathsf{Z}_t^k & -p\mathsf{Z}_t^k\mathsf{A}\end{pmatrix}$$

where $\mathsf{Z}_t = p^2\mathsf{A}^2 + \mathsf{c}t\mathsf{B}$ and Π_k is given by $\mathsf{c}\Pi_k = p\left(\mathsf{A}\mathsf{Z}_t^k - \mathsf{Z}_t^k\mathsf{A}\right)$ with $\Pi_1 = pt(\mathsf{AB} - \mathsf{BA})$. Since $\lim_t \mathsf{Z}_t = p^2\mathsf{A}^2 + \frac{p^2}{2}\mathsf{B} =: \mathsf{Z}_0$, the matrix norm $|\mathsf{Z}_t|$ is bounded by some constant for all small t. Consequently, by looking at the lower block matrices of e^{N_t}, we obtain

$$
\begin{aligned}
G\left(\frac{p}{t}, t\right) &= \mathsf{c}\sum_{k=0}^{\infty} \frac{1}{(2k+1)!}\mathsf{Z}_t^k, \\
F\left(\frac{p}{t}, t\right) &= \sum_{k=0}^{\infty} \frac{1}{(2k)!}\mathsf{Z}_t^k - p\sum_{k=0}^{\infty} \frac{1}{(2k+1)!}\mathsf{Z}_t^k\mathsf{A}.
\end{aligned}
$$

Since the convergence of infinite sums is uniform on a finite interval, we get the following limits:

$$
\begin{aligned}
G^*(p) := \lim_{t\to 0} tG\left(\frac{p}{t}, t\right) &= \frac{p^2}{2}\sum_{k=0}^{\infty} \frac{1}{(2k+1)!}\mathsf{Z}_0^k, \\
F^*(p) := \lim_{t\to 0} F\left(\frac{p}{t}, t\right) &= \sum_{k=0}^{\infty} \frac{1}{(2k)!}\mathsf{Z}_0^k - p\sum_{k=0}^{\infty} \frac{1}{(2k+1)!}\mathsf{Z}_0^k\mathsf{A}.
\end{aligned}
$$

On the other hand, we notice that $\mathsf{Z}_0 = -p^2 Q^\top \left(\mathbb{I} - R^\top R\right) Q$. Hence, it is represented by $-p^2\Gamma^2$. The final expression is easily obtained by the definitions of matrix sine and cosine.

Let us turn our attention to $e^{t\mathsf{M}+\mathsf{N}_t}$. We note that

$$
\begin{aligned}
e^{t\mathsf{M}+\mathsf{N}_t} - e^{\mathsf{N}_t} &= \sum_{l=0}^{\infty} \frac{1}{(2l)!}\left\{(t\mathsf{M}+\mathsf{N}_t)^{2l} - \mathsf{N}_t^{2l}\right\} + \sum_{l=0}^{\infty} \frac{1}{(2l+1)!}\left\{(t\mathsf{M}+\mathsf{N}_t)^{2l+1} - \mathsf{N}_t^{2l+1}\right\} \\
&= \sum_{l=0}^{\infty} \frac{1}{(2l)!}\left\{(t\mathsf{M}+\mathsf{N}_t)^{2l} - \mathsf{N}_t^{2l}\right\} + \sum_{l=0}^{\infty} \frac{1}{(2l+1)!}\left(t\mathsf{M}+\mathsf{N}_t\right)^{2l} t\mathsf{M} \\
&\quad + \sum_{l=0}^{\infty} \frac{1}{(2l+1)!}\left\{(t\mathsf{M}+\mathsf{N}_t)^{2l} - \mathsf{N}_t^{2l}\right\}\mathsf{N}_t.
\end{aligned}
$$

It is easy to verify that the matrix norm $|(t\mathsf{M}+\mathsf{N}_t)^2|$ is bounded by a constant for all small t values. This implies that the second term converges to zero as t decreases to zero. Indeed, $(t\mathsf{M}+\mathsf{N}_t)^2 = t^2\mathsf{M}^2 + \mathsf{M}(t\mathsf{N}_t) + t\mathsf{N}_t\mathsf{M} + \mathsf{N}_t^2$ and each term is at most $\mathcal{O}(1)$ in t.

Regarding the first and third terms on the right hand side, we are concerned with their lower left, say (i), and lower right (ii) block matrices of size n-by-n as they are relevant to matrices G and F. Simple calculations show that

$$
(t\mathsf{M}+\mathsf{N}_t)^2 = \left(\begin{array}{c|c} \mathcal{O}(t) + \mathsf{Z}_t & \mathcal{O}(t) \\ \hline \mathcal{O}(1) & \mathcal{O}(t) + \mathsf{Z}_t \end{array}\right), \quad \mathsf{N}_t^2 = \left(\begin{array}{c|c} \mathsf{Z}_t & \mathcal{O}(t) \\ \hline \mathbf{0} & \mathsf{Z}_t \end{array}\right).
$$

As a result, it can be verified that (i) and (ii) are of $\mathcal{O}(1)$ and $\mathcal{O}(t)$, respectively. Therefore, as t decreases to zero, (i) multiplied by t does not affect the limit $G^*(p)$. Likewise, $F^*(p)$ is independent of (ii). Similar arguments can be made for the third summation. Hence, the limit in the statement is valid for any M.

Now we notice that $F^*(0) = \mathbb{I}$ and thus there is a maximal open interval (p_-, p_+) including zero such that $\det F^*(p) \neq 0$. The equivalence of this and the finiteness of $\Lambda(p)$ comes from the limit

$$\Lambda(p) = \mathrm{Tr}\left[F^*(p)^{-1}G^*(p)\Sigma_0\right] = \frac{1}{\det F^*(p)} \times \text{some non-blow-up function of } p,$$

and $\det F^*(p) = \lim_t \det F\left(\frac{p}{t}, t\right)$, in addition to that the matrix A exists if and only if $\det F$ is nonzero. Hence, for $\Lambda(p)$ to be finite, each $\det F\left(\frac{p}{t}, t\right)$ must be nonzero for all small t values with a nonzero limit.[1] The last statement then follows.

So far we have not discussed the limit $\lim_{t\to 0} tb\left(\frac{p}{t}, t\right)$. However, from Proposition 1, we notice that

$$\begin{aligned}\lim_{t\to 0} b\left(\frac{p}{t}, t\right) &= -\frac{\beta}{2}\lim_t \mathrm{Tr}\left[\ln F\left(\frac{p}{t}, t\right) + tM + pQ^\top R^\top\right] \\ &= -\frac{\beta}{2}\mathrm{Tr}\left[\ln F^*(p) + pQ^\top R^\top\right]\end{aligned}$$

as long as $p \in (p_-, p_+)$. As a result, this term does not appear in $\Lambda(p)$. ■

Remark 1 The above lemma imposes the assumption that RQ is symmetric. This condition is arguably mild. Da Fonseca et al. [8] show that the stochastic correlation ρ_t between the stock noise and the volatility noise depends on the product RQ by $\rho_t = \frac{\mathrm{Tr}[RQ\Sigma_t]}{\sqrt{\mathrm{Tr}[\Sigma_t]\mathrm{Tr}[Q^\top Q\Sigma_t]}}$. Symmetric RQ does not affect the flexibility of this model to reflect the stochastic skew effect. The same condition appears in other related works. For instance, in Gnoatto and Grasselli [18] or Da Fonseca [5], the authors derive the explicit Laplace transform for the WMSV model under technical conditions including the symmetric RQ. ▮

The CGF $\ln \mathsf{E}\left[e^{p(Y_t - Y_0)/t}\right]$ is convex in p for each t by Hölder's inequality. Therefore, its limit $\Lambda(p)$ is convex as well. On the other hand, its derivative $\Lambda'(p)$ can be computed as

$$\Lambda'(p) = \mathrm{Tr}\left[F^*(p)^{-1}\left(-F^{*\prime}(p)F^*(p)^{-1}G^*(p) + G^{*\prime}(p)\right)\Sigma_0\right] = \frac{f(p)}{\det F^*(p)^2} \tag{4}$$

for some non-blow-up function f. Here the derivative of a matrix with respect to a scalar p is computed by componentwise differentiation. Recall G^* and F^* are the limits of $tG\left(\frac{p}{t}, t\right)$ and $F\left(\frac{p}{t}, t\right)$ defined in Lemma 1. Consequently, $|\Lambda'(p)| \to \infty$ at the boundary of (p_-, p_+) as long as $f(p)$ does not vanish at the boundary points. Based on this essential smoothness of Λ and its convexity, we obtain the next result.

[1] Alternatively, one might apply arguments similar to Lemma B.1 of Forde and Jacquier [11].

Proposition 2 *Assume that all of the conditions in Lemma 1 hold and that $f(p)$ of (4) does not vanish at $p_\pm$. Then, $(Y_t - Y_0)$ satisfies the large deviation principle as t approaches zero, with the rate function $\Lambda^*(x) = \sup_{p\in(p_-,p_+)}\{px - \Lambda(p)\}$ for all $x \in \mathbb{R}$. As a result, if $\mathrm{Tr}[\Sigma_0] > 0$, then we get*

$$\Lambda^*(k) = \begin{cases} -\lim\limits_{t\to 0} t \ln \mathsf{P}\Big(Y_t - Y_0 > k\Big), & \text{if } k \geq 0, \\ -\lim\limits_{t\to 0} t \ln \mathsf{P}\Big(Y_t - Y_0 < k\Big), & \text{if } k \leq 0. \end{cases}$$

Proof: The large deviations principle follows from the well known Gärtner-Ellis Theorem in Dembo and Zeitouni [9]. The Fenchel-Legendre transform Λ^* is a good rate function. A version in our context for $\{Y_t - Y_0\}$ is

$$-\inf_{x>k}\Lambda^*(x) \leq \liminf_{t\to 0} t\ln \mathsf{P}\Big(Y_t - Y_0 > k\Big) \leq \limsup_{t\to 0} t\ln \mathsf{P}\Big(Y_t - Y_0 > k\Big) \leq -\inf_{x\geq k}\Lambda^*(x)$$

for $k \geq 0$. This version is similar to the one summarized in Forde et al. [14][2]. Since Λ is convex, Λ^* is also convex. Furthermore, the essential smoothness of Λ implies it is a closed proper convex function whose convex dual Λ^* is closed and convex with $\Lambda^{**} = \Lambda$. Theorem 26.3 of Rockafellar [28] then tells us that Λ^* is strictly convex.

Straightforward computations yield $F^*(0) = \mathbb{I}$, $G^*(0) = G^{*\prime}(0) = \mathbf{0}$, and $G^{*\prime\prime}(0) = \mathbb{I}$. And this results in $\Lambda(0) = \Lambda'(0) = 0$ and $\Lambda''(0) = \mathrm{Tr}[\Sigma_0]$. Therefore if $\mathrm{Tr}[\Sigma_0] > 0$, then there is a small interval around 0, say $\mathcal{I} = [-\varepsilon, \varepsilon]$ on which $\Lambda'(p)$ is monotonically increasing and $\Lambda''(p)$ is strictly positive. Outside of $\mathcal{I}$, the convexity of Λ implies Λ' is non-decreasing. Now let us consider $x \in [\Lambda'(-\varepsilon), \Lambda'(\varepsilon)]$. Note this last interval includes 0 in its interior. For such x, $\Lambda^*(x) = p^*x - \Lambda(p^*)$ where $p^* = p^*(x)$ is a unique solution to the equation $\Lambda'(p) = x$. It is clear that $p^*(0) = 0$.

Non-degeneracy of Λ'' around 0 then implies the differentiability of $p^*(x)$ thanks to the Implicit Function Theorem. Finally, we observe

$$\Lambda^{*\prime}(x) = p^{*\prime}(x)x + p^*(x) - \Lambda'(p^*)p^{*\prime}(x) = p^*(x), \quad \Lambda^*(0) = \Lambda^{*\prime}(0) = 0.$$

From the strict convexity of Λ^*, we can conclude that $\Lambda^{*\prime}$ is positive on the positive real and negative on the negative real. In turn, we see that Λ^* is non-decreasing on the positive real and non-increasing on the negative real. As a consequence, $-\inf_{x\geq k}\Lambda^*(x) = -\inf_{x>k}\Lambda^*(x) = -\Lambda^*(k)$ for $k \geq 0$. We can draw a similar conclusion for $k \leq 0$. The desired statement is immediate. ■

In the above proposition, the assumption that $f(p)$ does not vanish at the boundary points is a mild one as it is expected that the two nonlinear functions f and $\det F^*$ do not share zeros in common settings. We note that the large deviations results above are the extension of Theorem

[2] We refer the reader to an extended version available at the authors' websites

2.1 in Forde and Jacquier [11] to the WMSV model. Since their subsequent analyses are general in nature, we adopt those results in order to connect Proposition 2 to implied volatilities in the next subsection.

2.3 First order expansion

For the reader's convenience, we record relevant results.

Proposition 3 (Forde and Jacquier [11]) *Assume that all of the assumptions in Proposition 2 hold. Then, we have the following small-time behaviors of vanilla option prices and implied volatilities:*

1. *for out-of-the-money call option with the log-moneyness $x = \ln \frac{K}{S_0} \geq 0$,*
$$\Lambda^*(x) = -\lim_{t \to 0} t \ln C(S_0, K, t);$$
2. *for out-of-the-money put option with the log-moneyness $x \leq 0$,*
$$\Lambda^*(x) = -\lim_{t \to 0} t \ln P(S_0, K, t);$$
3. *for the option implied volatility $\sigma_t(x)$ with the log-moneyness $x \neq 0$,*
$$I(x) = \lim_{t \to 0} \sigma_t(x) = \frac{|x|}{\sqrt{2\Lambda^*(x)}}.$$

Here, $C(S_0, K, t)$ and $P(S_0, K, t)$ are the respective prices of call and put options with initial price S_0, strike K, and maturity t.

Theorem 1 *Assume that all of the conditions in Proposition 2 hold. Then, in some neighborhood of zero for the log-moneyness $x = \ln \frac{K}{S_0}$, the following expansion is valid: with $y = \frac{x}{\mathrm{Tr}[\Sigma_0]}$,*

$$\begin{aligned} I(x) &= \sqrt{\mathrm{Tr}[\Sigma_0]} \Bigg[1 + \frac{1}{2} \frac{\mathrm{Tr}[RQ\Sigma_0]}{\mathrm{Tr}[\Sigma_0]} y + \Bigg(\frac{1}{6} \frac{\mathrm{Tr}\left[Q^\top Q \Sigma_0\right]}{\mathrm{Tr}[\Sigma_0]} \\ &\quad + \frac{1}{3} \frac{\mathrm{Tr}\left[(RQ)^2 \Sigma_0\right]}{\mathrm{Tr}[\Sigma_0]} - \frac{3}{4} \frac{\mathrm{Tr}[RQ\Sigma_0]^2}{\mathrm{Tr}[\Sigma_0]^2} \Bigg) y^2 + \mathcal{O}(y^3) \Bigg]. \end{aligned}$$

Proof: The formula above is derived by tedious but straightforward computations of Taylor expansions. In the proof of Proposition 2, we already argued that Λ' is smooth and strictly monotone in a neighborhood of $p = 0$. Its convex dual Λ^* is then given by $\Lambda^*(x) = p^*(x)x - \Lambda(p^*(x))$ where

p^* is a smooth solution to $\Lambda'(p) = x$ in a small neighborhood of $x = 0$. Then, Taylor expansions yield

$$\begin{aligned}\Lambda(p) &= \frac{1}{2}\mathrm{Tr}[\Sigma_0]p^2 + \frac{1}{2}\mathrm{Tr}[RQ\Sigma_0]p^3 + \frac{1}{6}\mathrm{Tr}\left[\left(Q^\top Q + 2(RQ)^2\right)\Sigma_0\right]p^4 + \mathcal{O}(p^5),\\ p^*(x) &= \frac{1}{\mathrm{Tr}[\Sigma_0]}x - \frac{3}{2}\frac{\mathrm{Tr}[RQ\Sigma_0]}{\mathrm{Tr}[\Sigma_0]^3}x^2 + \mathcal{O}(x^3).\end{aligned}$$

For notational convenience, we denote $\Lambda(p) = p^2\sum_{i=0}^{2} a_i p^i + \mathcal{O}(p^5)$ and $p^*(x) = x\sum_{i=0}^{2} b_i x^i + \mathcal{O}(x^4)$. It can be shown that

$$b_2 = \frac{9}{2}\frac{\mathrm{Tr}\left[RQ\Sigma_0\right]^2}{\mathrm{Tr}\left[\Sigma_0\right]^5} - \frac{2}{3}\frac{\mathrm{Tr}\left[\left(Q^\top Q + 2(RQ)^2\right)\Sigma_0\right]}{\mathrm{Tr}\left[\Sigma_0\right]^4}.$$

Recall that in Proposition 3, for all sufficiently small nonzero log-moneyness x, we have $I(x)^2 = \frac{x^2}{2\Lambda^*(x)}$. On the other hand, the above expansions give us

$$\frac{2\Lambda^*(x)}{x^2} = \frac{2}{x^2}\Big(p^*(x)x - \Lambda\left(p^*(x)\right)\Big) = c_0 + c_1 x + c_2 x^2 + \mathcal{O}(x^3)$$

for some constants c_i's. Those values are readily obtained by collecting relevant terms carefully as functions of a_i's and b_i's. To be explicit, they are given by $c_0 = \frac{1}{\mathrm{Tr}[\Sigma_0]}$, $c_1 = -\frac{\mathrm{Tr}[RQ\Sigma_0]}{\mathrm{Tr}[\Sigma_0]^3}$, and lastly

$$c_2 = \frac{9}{4}\frac{\mathrm{Tr}[RQ\Sigma_0]^2}{\mathrm{Tr}[\Sigma_0]^5} - \frac{1}{3}\frac{\mathrm{Tr}\left[\left(Q^\top Q + 2(RQ)^2\right)\Sigma_0\right]}{\mathrm{Tr}[\Sigma_0]^4}.$$

The reciprocal of the above expansion is not difficult to get:

$$\begin{aligned}I(x)^2 &= \frac{1}{c_0} - \frac{c_1}{c_0^2}x + \left(\frac{c_1^2}{c_0^3} - \frac{c_2}{c_0^2}\right)x^2 + \mathcal{O}(x^3),\\ I(x) &= \frac{1}{\sqrt{c_0}} - \frac{c_1}{2c_0^{1.5}}x + \frac{3c_1^2 - 4c_0c_2}{8c_0^{2.5}}x^2 + \mathcal{O}(x^3).\end{aligned}$$

The formula in the statement then easily follows. ■

The expansion in Theorem 1 reduces to that of the Heston model in Theorem 3.2 of Forde and Jacquier [11]. We also note that the above result, in particular $I(x)^2$, is analogous to one of the main results in Da Fonseca and Grasselli [7], Proposition 3.6 for the WMSV model. Their method is based on the direct expansion of the call option price in terms of the vol-of-vol scale factor α. In fact, two formulae are the same if $\alpha = 1$ and RQ is symmetric. Nevertheless, we find the large deviations approach is useful because, first, the validity of the expansion is obtained without introducing an extra control parameter α and, second, the information about tail probability behaviors of Y_t is provided via the convex dual of the limiting CGF $\Lambda(p)$.

3 Small-time Smiles

The program in Forde et al. [13] applies to the WMSV model as well. This extension is the main focus of the current section.

Lemma 2 *Assume that all of the conditions in Proposition 2 hold. We further assume that the matrix M is symmetric. Then, for each $p \in (p_-, p_+)\backslash\{0\}$, we have*

$$\mathsf{E}\left[e^{\frac{p}{t}(Y_t - Y_0)}\right] = U(p)e^{\frac{\Lambda(p)}{t}}(1 + \mathcal{O}(t))$$

as t decreases to zero where $\Lambda(p)$ is the limiting cumulant generating function and $U(p)$ is given by

$$\begin{aligned}
\ln U(p) &= \operatorname{Tr}\left[F^{*-1}\left[-\frac{1}{2}\sin p\Gamma \cdot \Gamma^{-1} + \frac{1}{2}(p\mathsf{D}_1 - \mathsf{D}_2)\right]\Sigma_0\right] \\
&\quad + \operatorname{Tr}\left[F^{*-1}\left[\sin p\Gamma \cdot \Upsilon_0' + \frac{1}{p}\sin p\Gamma \cdot \Gamma^{-1}M + \left(\mathsf{D}_1 - \frac{1}{p}\mathsf{D}_2\right)RQ\right]F^{*-1}G^*\Sigma_0\right] \\
&\quad - \frac{\beta}{2}\operatorname{Tr}\left[\ln F^* + pRQ\right].
\end{aligned}$$

Here, $F^*(p)$, $G^*(p)$ and Γ are as in Lemma 1. New symbols $\mathsf{D}_1, \mathsf{D}_2, \Upsilon_0'$ satisfy

$$\begin{aligned}
\mathsf{D}_1 &= \cos p\Gamma \cdot \Upsilon_0'\Gamma^{-1}, \\
\mathsf{D}_2 &= \sin p\Gamma \cdot \Gamma^{-1}\Upsilon_0'\Gamma^{-1}, \\
\operatorname{vec}\Upsilon_0' &= -(\Gamma \oplus \Gamma)^{-1}\operatorname{vec}\left(MRQ + RQM + Q^\top Q\right)
\end{aligned}$$

where vec *is the vectorization operator and $\oplus$ means the Kronecker sum.*

Proof: In the proof of Lemma 1, we showed that the matrix M does not affect the limiting CGF. However, for small-time smiles, we need to calculate the behaviors of the CGF for small t values, which turn out to depend on M. Let us repeat the arguments of the lemma as follows. Recall the solution $A\left(\frac{p}{t}, t\right)$ is $F^{-1}G$ where

$$\begin{pmatrix} G\left(\frac{p}{t}, t\right) & F\left(\frac{p}{t}, t\right) \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbb{I} \end{pmatrix} e^{\mathsf{L}_t}, \quad \mathsf{L}_t = \begin{pmatrix} tM + p\mathsf{A} & t\mathsf{B} \\ \mathsf{c}\mathbb{I} & -(tM + p\mathsf{A}) \end{pmatrix}$$

where $\mathsf{A} = RQ$, $\mathsf{B} = -2Q^\top Q$, and $\mathsf{c} = \frac{p^2}{2t} - \frac{p}{2}$. For notational convenience, we denote $(tM + p\mathsf{A})^2 + \mathsf{c}t\mathsf{B}$ by Z_t. Then, thanks to the symmetry of M, it can be readily shown that

$$\mathsf{L}_t^{2k} = \begin{pmatrix} \mathsf{Z}_t^k & \Pi_k \\ \mathbf{0} & \mathsf{Z}_t^k \end{pmatrix}$$

for some matrix Π_k. More explicitly, it is given by $\mathsf{c}\Pi_k = (tM + p\mathsf{A})\mathsf{Z}_t^k - \mathsf{Z}_t^k(tM + p\mathsf{A})$. We can rewrite

$$\mathsf{Z}_t = \mathsf{Z}_0 + t\mathsf{C} + t^2M^2, \quad \mathsf{Z}_0 = p^2\mathsf{A}^2 + \frac{p^2}{2}\mathsf{B}, \quad \mathsf{C} = p\left(M\mathsf{A} + \mathsf{A}M - \frac{1}{2}\mathsf{B}\right).$$

From these calculations, we can find the matrices G and F as

$$\begin{aligned} G\left(\frac{p}{t}, t\right) &= \mathsf{c}\sum_{k=0}^{\infty} \frac{1}{(2k+1)!}\mathsf{Z}_t^k, \\ F\left(\frac{p}{t}, t\right) &= \sum_{k=0}^{\infty} \frac{1}{(2k)!}\mathsf{Z}_t^k - \sum_{k=0}^{\infty} \frac{1}{(2k+1)!}\mathsf{Z}_t^k(tM + p\mathsf{A}). \end{aligned}$$

Under the assumed conditions, $-\mathsf{Z}_t$ converges to a nonsingular $-\mathsf{Z}_0 = (p\Gamma)^2$ as t approaches zero. Here Γ is as given in Lemma 1. Since Z_t is symmetric, we can find a nonsingular Υ_t such that $\mathsf{Z}_t = -\Upsilon_t^2$ for some symmetric and nonsingular Υ_t and for all sufficiently small t. Clearly, $\lim_{t\to 0} \Upsilon_t = p\Gamma$. Consequently, more compact representations are possible:

$$\begin{aligned} G\left(\frac{p}{t}, t\right) &= \mathsf{c}\sin\Upsilon_t \cdot \Upsilon_t^{-1}, \\ F\left(\frac{p}{t}, t\right) &= \cos\Upsilon_t - \sin\Upsilon_t \cdot \Upsilon_t^{-1}(tM + p\mathsf{A}). \end{aligned}$$

Let us now make the next observations: direct differentiations yield

$$\begin{aligned} G\left(\frac{p}{t}, t\right) &= \mathsf{c}\sin\Upsilon_0 \cdot \Upsilon_0^{-1} + \mathsf{c}\mathsf{D}t + \mathsf{c}\mathcal{O}(t^2) \\ &= \frac{1}{t}G^*(p) - \frac{p}{2}\sin\Upsilon_0 \cdot \Upsilon_0^{-1} + \frac{p^2}{2}\mathsf{D} + \mathcal{O}(t), \\ F\left(\frac{p}{t}, t\right) &= F^*(p) - \sin\Upsilon_0 \cdot \Upsilon_0' t - \mathsf{D}t(tM + p\mathsf{A}) - \sin\Upsilon_0 \cdot \Upsilon_0^{-1}Mt + \mathcal{O}(t^2) \\ &= F^*(p) - \left[\sin\Upsilon_0 \cdot \Upsilon_0' + \sin\Upsilon_0 \cdot \Upsilon_0^{-1}M + p\mathsf{D}\mathsf{A}\right]t + \mathcal{O}(t^2). \end{aligned}$$

Here, D is defined as $\frac{\mathrm{d}}{\mathrm{d}t}\sin\Upsilon_t \cdot \Upsilon_t^{-1}\Big|_{t=0}$ and computed as

$$\mathsf{D} = \cos\Upsilon_0 \cdot \Upsilon_0'\Upsilon_0^{-1} - \sin\Upsilon_0 \cdot \Upsilon_0^{-1}\Upsilon_0'\Upsilon_0^{-1}.$$

We note that Υ_0' satisfies the special case of Sylvester's equation: $\Upsilon_0'\Upsilon_0 + \Upsilon_0\Upsilon_0' = -\mathsf{C}$ by differentiating the defining equation of Υ_t^2. Thanks to the symmetry and the positive definiteness of Υ_t for small t, it is known that there exists a unique solution Υ_0' and the solution is obtained by $\mathrm{vec}\Upsilon_0' = -\left(\Upsilon_0 \otimes \mathbb{I} + \mathbb{I} \otimes \Upsilon_0\right)^{-1}\mathrm{vec}\mathsf{C}$ where $\otimes$ is the Kronecker product operator and vec is the vectorization operator.

Using the definition of Kronecker sum, we get $\Upsilon_0 \oplus \Upsilon_0 = \Upsilon_0 \otimes \mathbb{I} + \mathbb{I} \otimes \Upsilon_0$. It is then immediate to see $\mathrm{vec}\Upsilon_0' = -(\Gamma \oplus \Gamma)^{-1}\mathrm{vec}\left(M\mathsf{A} + \mathsf{A}M - \frac{1}{2}\mathsf{B}\right)$. We also write for notational simplicity,

$$\mathsf{D} = \frac{1}{p}\mathsf{D}_1 - \frac{1}{p^2}\mathsf{D}_2, \quad \begin{cases} \mathsf{D}_1 = \cos(p\Gamma)\Upsilon_0'\Gamma^{-1} \\ \mathsf{D}_2 = \sin(p\Gamma)\Gamma^{-1}\Upsilon_0'\Gamma^{-1}. \end{cases}$$

Finally, it only takes several simple algebraic manipulations until we arrive at

$$A\left(\frac{p}{t},t\right) = \frac{1}{t}F^{*-1}G^* + F^{*-1}\left[-\frac{1}{2}\sin p\Gamma\cdot\Gamma^{-1} + \frac{1}{2}(p\mathsf{D}_1 - \mathsf{D}_2)\right]$$
$$+F^{*-1}\left[\sin p\Gamma\cdot\Upsilon_0' + \frac{1}{p}\sin p\Gamma\cdot\Gamma^{-1}M + \left(\mathsf{D}_1 - \frac{1}{p}\mathsf{D}_2\right)\mathsf{A}\right]F^{*-1}G^* + \mathcal{O}(t).$$

This first order expansion of A combined with the known formula $b\left(\frac{p}{t},t\right)$ in the proof of Lemma 1 helps us get

$$\ln\mathsf{E}\left[e^{\frac{p}{t}(Y_t-Y_0)}\right] = \frac{1}{t}\Lambda(p) + \mathrm{Tr}\left[F^{*-1}\left[-\frac{1}{2}\sin p\Gamma\cdot\Gamma^{-1} + \frac{1}{2}(p\mathsf{D}_1 - \mathsf{D}_2)\right]\Sigma_0\right]$$
$$+\mathrm{Tr}\left[F^{*-1}\left[\sin p\Gamma\cdot\Upsilon_0' + \frac{1}{p}\sin p\Gamma\cdot\Gamma^{-1}M + \left(\mathsf{D}_1 - \frac{1}{p}\mathsf{D}_2\right)\mathsf{A}\right]F^{*-1}G^*\Sigma_0\right]$$
$$-\frac{\beta}{2}\mathrm{Tr}\left[\ln F^* + p\mathsf{A}\right] + \mathcal{O}(t).$$

■

It is possible to obtain an asymptotic expansion of $\ln U(p)$ which becomes handy for approximating small-time smiles. Since the derivation is tedious and long, we briefly sketch some relevant computations.

Corollary 1 *Assume that all of the conditions in Lemma 2 hold. Then, in a small neighborhood of zero for p, the following asymptotic expansion is valid:*

$$\ln U(p) = -\frac{1}{2}\mathrm{Tr}[\Sigma_0]p + \left\{-\frac{1}{2}\mathrm{Tr}[RQ\Sigma_0] + \frac{1}{2}\mathrm{Tr}\left[M\Sigma_0\right] + \frac{\beta}{4}\mathrm{Tr}\left[Q^\top Q\right]\right\}p^2$$
$$+\left\{\mathrm{Tr}\left[\left(-\frac{1}{6}\Gamma^2 - \frac{1}{6}\Gamma^2\Upsilon_0'\Gamma^{-1} - \frac{1}{2}(RQ)^2\right)\Sigma_0\right]\right.$$
$$\left.+\frac{1}{2}\mathrm{Tr}\left[\left(\Gamma\Upsilon_0' + MRQ + RQM\right)\Sigma_0\right] + \frac{\beta}{6}\mathrm{Tr}\left[Q^\top QRQ\right]\right\}p^3 + \mathcal{O}(p^4)$$

where Γ and Υ_0' are as in Lemma 2.

Proof: Useful expansions are given as follows:

$$F^{*-1} = \mathbb{I} + \mathsf{A}p + \left(\mathsf{A}^2 + \frac{1}{2}\Gamma^2\right)p^2 + \mathcal{O}(p^3),$$
$$G^* = \frac{p^2}{2}\mathbb{I} - \frac{1}{12}\Gamma^2p^4 + \mathcal{O}(p^6),$$
$$\sin p\Gamma = \Gamma p - \frac{1}{6}\Gamma^3p^3 + \mathcal{O}(p^5),$$
$$\cos p\Gamma = \mathbb{I} - \frac{1}{2}\Gamma^2p^2 + \frac{1}{24}\Gamma^4p^4 + \mathcal{O}(p^6).$$

Then, expansions for D_1, D_2 can be calculated accordingly. Regarding the last term in $\ln U(p)$ of Lemma 2, we do the formal expansion of $\ln F^* = \alpha_0 + \alpha_1 p + \cdots$ and find matching coefficients via

$$e^{\ln F^*} = F^* = \mathbb{I} - \mathsf{A}p - \frac{1}{2}\Gamma^2 p^2 + \frac{1}{6}\Gamma^2 \mathsf{A} p^3 + \mathcal{O}(p^4).$$

This leads to $\alpha_0 = \mathbf{0}$, $\alpha_1 = -\mathsf{A}$, $\alpha_2 = -\frac{1}{2}\Gamma^2 - \frac{1}{2}\mathsf{A}^2$, and $\alpha_3 = -\frac{1}{12}\Gamma^2\mathsf{A} - \frac{1}{4}\mathsf{A}\Gamma^2 - \frac{1}{3}\mathsf{A}^3$. Then, we eventually find that

$$\begin{aligned} -\frac{\beta}{2}\mathrm{Tr}\left[\ln F^* + \mathsf{A}p\right] &= -\frac{\beta}{2}\mathrm{Tr}\left[\alpha_2 p^2 + \alpha_3 p^3\right] + \mathcal{O}(p^4) \\ &= \frac{\beta}{4}\mathrm{Tr}\left[Q^\top Q\right] p^2 + \frac{\beta}{6}\mathrm{Tr}\left[Q^\top Q\mathsf{A}\right] p^3 + \mathcal{O}(p^4). \end{aligned}$$

Carefully collecting relevant terms, we obtain the desired result. ■

Lemma 3 *Assume that all of the conditions in Lemma 1 hold. Let us define* $h(q) = \mathrm{Re}\big(\Lambda(p^*(x) + \mathrm{i}q)\big)$, $q \in \mathbb{R}$, *where* $p^*(x)$ *is a solution to the equation* $\Lambda'(p) = x$ *for* x *in a small neighborhood of zero. If* $\mathrm{Tr}[\Sigma_0] > 0$, *then* $h(q)$ *attains a unique maximum at* $q = 0$.

Proof: Under the given assumptions, it is argued in the proof of Proposition 2 that there is a small interval $x \in [\Lambda'(-\varepsilon), \Lambda'(\varepsilon)]$ for a small $\varepsilon > 0$ with the following properties. First, this interval includes 0 in its interior. Second, Λ'' is strictly positive on $[-\varepsilon, \varepsilon]$. Third, for such x, there is a unique solution $p^*(x)$ to $\Lambda'(p) = x$ thanks to the strict monotonicity of Λ'.

Let us denote $z = p^* + \mathrm{i}q$. The CGF of $(Y_t - Y_0)$ is defined as $\Lambda_t\left(\frac{z}{t}\right) = \ln \mathsf{E}\left[e^{\frac{z}{t}(Y_t - Y_0)}\right]$. Then, observe that

$$\left|\mathsf{E}\left[e^{\frac{z}{t}(Y_t-Y_0)}\right]\right| = \left|\exp\left(\mathrm{Re}\Lambda_t\left(\frac{z}{t}\right) + \mathrm{i}\,\mathrm{Im}\Lambda_t\left(\frac{z}{t}\right)\right)\right| = \exp\left(\mathrm{Re}\Lambda_t\left(\frac{z}{t}\right)\right).$$

On the other hand, the left hand side of the above equation is clearly less than or equal to

$$\mathsf{E}\left[\left|e^{\frac{z}{t}(Y_t-Y_0)}\right|\right] = \mathsf{E}\left[e^{\frac{p^*}{t}(Y_t-Y_0)}\right] = \exp\left(\Lambda_t\left(\frac{p^*}{t}\right)\right).$$

Since this holds for all t, we obtain the limiting result after multiplying t, $\mathrm{Re}\Lambda(p^* + \mathrm{i}q) \le \Lambda(p^*)$, making $q = 0$ a maximum.

In order to see that $q = 0$ is a unique maximum, we note that

$$h''(0) = \mathrm{Re}\Big(\Lambda''(p^* + \mathrm{i}q)\mathrm{i}^2\Big)\Big|_{q=0} = -\Lambda''(p^*).$$

The strict convexity of Λ on $[-\varepsilon, \varepsilon]$ implies that $-\Lambda''(p^*) < 0$ so that the function h becomes strictly concave at $q = 0$. This proves the uniqueness of the maximum. ■

Based on the two lemmas above, the procedure of Forde et al. [13] can be applied. For the reader's convenience, we summarize its outline. Assume that the log-moneyness $x = \ln \frac{K}{S_0}$ is in the small neighborhood of Lemma 3. Then, for sufficiently small t, it holds that

$$\frac{1}{S_0}\mathsf{E}\left[\left(e^{Y_t} - S_0 e^x\right)^+\right] = (1-e^x)\mathbf{1}_{\{x<0\}} - \frac{e^x t}{2\pi}\mathrm{Re}\left(\int_{\mathscr{P}} e^{\mathrm{i}xu/t}\phi_t\left(-\frac{u}{t}\right)\left(\frac{1}{u^2}+\mathcal{O}(t)\right)\mathrm{d}u\right)$$

where $\phi_t(z) = \mathsf{E}\left[e^{\mathrm{i}z(Y_t - Y_0)}\right]$ for complex number z with $-\mathrm{Im}(z) \in (p_-, p_+)$. The integration path $\mathscr{P}$ is from $-\infty + \mathrm{i}p^*(x)$ to $\infty + \mathrm{i}p^*(x)$. Then, Lemma 2 implies

$$\begin{aligned}
&\int_{\mathscr{P}} e^{\mathrm{i}xu/t}\phi_t\left(-\frac{u}{t}\right)\left(\frac{1}{u^2}+\mathcal{O}(t)\right)\mathrm{d}u \\
&= \int_{\mathscr{P}} e^{\mathrm{i}xu/t} U(-\mathrm{i}u) e^{\Lambda(-\mathrm{i}u)/t}(1+\mathcal{O}(t))\left(\frac{1}{u^2}+\mathcal{O}(t)\right)\mathrm{d}u \\
&= \int_{\mathscr{P}} e^{-H(u)/t}\frac{U(-\mathrm{i}u)}{u^2}\mathrm{d}u(1+\mathcal{O}(t))
\end{aligned}$$

where $H(u) = -\mathrm{i}xu - \Lambda(-\mathrm{i}u)$. With nonzero x, Lemma 3 is then utilized to apply the Laplace expansion of Theorem 7.1, Chapter 4, of Olver [29]. More specifically, $\mathrm{Re}(H(u) - H(u_0))$ is positive on $\mathscr{P}$ for $u_0 = \mathrm{i}p^*(x)$. The first order expansion of the resulting formula is

$$2\sqrt{\pi t}e^{-H(u_0)/t}\frac{U(-\mathrm{i}u_0)}{u_0^2\sqrt{2F''(u_0)}}(1+\mathcal{O}(t)) = -2\sqrt{\pi t}e^{-(xp^* - \Lambda(p^*))/t}\frac{U(p^*)}{p^{*2}\sqrt{2\Lambda''(p^*)}}(1+\mathcal{O}(t)).$$

By definition of the convex dual of Λ, we see that $xp^* - \Lambda(p^*) = \Lambda^*(x)$. The next result is simply the WMSV version of Theorem 3.1 of Forde et al. [13].

Proposition 4 *Assume that all of the conditions in Lemma 2 hold. Then, for a nonzero log-moneyness $x = \ln \frac{K}{S_0}$ in a small neighborhood of zero, the asymptotic behavior for European call options is given by*

$$\frac{1}{S_0}\mathsf{E}\left[\left(e^{Y_t} - S_0 e^x\right)^+\right] = (1-e^x)\mathbf{1}_{\{x<0\}} + e^{-\Lambda^*(x)/t}\left(\frac{A(x)}{\sqrt{2\pi}}t^{1.5} + \mathcal{O}(t^{2.5})\right)$$

where $A(x) = \frac{e^x U(p^*(x))}{p^*(x)^2\sqrt{\Lambda''(p^*(x))}}$ as t decreases to zero.

Comparison of the above expansion with the call price expansion under the Black-Scholes model leads us to the asymptotic expansion of the small-time smile as in Theorem 4.2 of Forde et al. [13]: for nonzero and small x,

$$\sigma_t^2(x) \approx I(x)^2 + \frac{2I(x)^4}{x^2}\left[\ln\frac{A(x)x^2}{I(x)^3} - \frac{x}{2}\right]t.$$

Here, $I(x)$ is as in Proposition 3. Since this step is straightforward, we refer the reader to the main reference. This expression still depends on the implicitly defined function $p^*(x)$. Nevertheless, we can utilize asymptotic results obtained so far. Specifically, recall that we derived expansions for $\Lambda(p)$, $p^*(x)$, and $I(x)$ previously. By plugging in estimates for Λ'' and U using Lemma 1 and Corollary 1, we get an approximation to $A(x)$. And Theorem 1 provides us with an approximation to $I(x)$. Alternatively, an asymptotic expression can be produced by expanding necessary quantities to suitable degrees.

Theorem 2 *Assume that all of the conditions in Lemma 2 hold. Then, for a nonzero log-moneyness $x = \ln \frac{K}{S_0}$ in a small neighborhood of zero, the small-time smile is approximated by*

$$\sigma_t^2 \approx I(x)^2 + 2I(x)^4(d_0 + d_1 x)t$$

for some constants d_0 and d_1. In particular, d_0 is equal to

$$\frac{1}{4}\frac{\mathrm{Tr}\left[RQ\Sigma_0\right]}{\mathrm{Tr}\left[\Sigma_0\right]^2} + \frac{1}{2}\frac{\mathrm{Tr}\left[M\Sigma_0\right]}{\mathrm{Tr}\left[\Sigma_0\right]^2} + \frac{\beta}{4}\frac{\mathrm{Tr}\left[Q^\top Q\right]}{\mathrm{Tr}\left[\Sigma_0\right]^2} + \frac{3}{8}\frac{\mathrm{Tr}\left[RQ\Sigma_0\right]^2}{\mathrm{Tr}\left[\Sigma_0\right]^4} - \frac{1}{6}\frac{\mathrm{Tr}\left[\left(Q^\top Q + 2(RQ)^2\right)\Sigma_0\right]}{\mathrm{Tr}\left[\Sigma_0\right]^3}.$$

Proof: For notational convenience, we write $\ln U(p) = u_0 p + u_1 p^2 + u_2 p^3 + \mathcal{O}(p^4)$. We also retrieve the expressions $\Lambda(p) = \sum_{i=0}^3 a_i p^{i+2} + \mathcal{O}(p^6)$ and $p^*(x) = \sum_{i=0}^3 b_i x^{i+1} + \mathcal{O}(x^5) = x\tilde{p}(x)$. Lastly, $I(x)^{-2} = \frac{2\Lambda^*(x)}{x^2} = \sum_{i=0}^3 c_i x^i + \mathcal{O}(x^4)$. Then, we have

$$\begin{aligned}
\left[\ln \frac{A(x)x^2}{I(x)^3} - \frac{x}{2}\right] &= \frac{x}{2} + \ln U(p^*) + 2\ln x - 2\ln p^* - \frac{1}{2}\ln \Lambda''(p^*) - 3\ln I(x) \\
&= \frac{x}{2} + \ln U(p^*) - 2\ln \tilde{p}(x) + \frac{1}{2}\ln p^{*\prime}(x) + \frac{3}{2}\ln I(x)^{-2} \\
&= \frac{x}{2} + u_0(b_0 x + b_1 x^2 + b_2 x^3) + u_1(b_0^2 x^2 + 2b_0 b_1 x^3) + u_2 b_0^3 x^3 + \mathcal{O}(x^4) \\
&\quad -2\ln\left[b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \mathcal{O}(x^4)\right] \\
&\quad +\frac{1}{2}\ln\left[b_0 + 2b_1 x + 3b_2 x^2 + 4b_3 x^3 + \mathcal{O}(x^4)\right] \\
&\quad +\frac{3}{2}\ln\left[c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \mathcal{O}(x^4)\right].
\end{aligned}$$

Here we utilized the relation $\Lambda'(p^*) = x$ so that $\Lambda''(p^*)p^{*\prime}(x) = 1$ by differentiating with respect to x. Our previous computations show us that $a_0 = \frac{1}{2}\mathrm{Tr}[\Sigma_0]$ and $b_0 = c_0 = \frac{1}{\mathrm{Tr}[\Sigma_0]}$. We also have $u_0 = -\frac{1}{2}\mathrm{Tr}[\Sigma_0]$. This simplifies the formula a bit and leads us to

$$\begin{aligned}
\left[\ln \frac{A(x)x^2}{I(x)^3} - \frac{x}{2}\right] &= u_0(b_1 x^2 + b_2 x^3) + u_1(b_0^2 x^2 + 2b_0 b_1 x^3) + u_2 b_0^3 x^3 \\
&\quad -2\ln\left[1 + \frac{b_1}{b_0}x + \frac{b_2}{b_0}x^2 + \frac{b_3}{b_0}x^3 + \mathcal{O}(x^4)\right] \\
&\quad +\frac{1}{2}\ln\left[1 + 2\frac{b_1}{b_0}x + 3\frac{b_2}{b_0}x^2 + 4\frac{b_3}{b_0}x^3 + \mathcal{O}(x^4)\right] \\
&\quad +\frac{3}{2}\ln\left[1 + \frac{c_1}{b_0}x + \frac{c_2}{b_0}x^2 + \frac{c_3}{b_0}x^3 + \mathcal{O}(x^4)\right].
\end{aligned}$$

It is easy to check that $c_1 = \frac{2}{3}b_1$. Hence, the Taylor expansions of the log functions cancel out x terms because $-2b_1 + b_1 + \frac{3}{2}c_1 = 0$. Finally, $\frac{1}{x^2}\left[\ln \frac{A(x)x^2}{I(x)^3} - \frac{x}{2}\right]$ boils down to $d_0 + d_1 x + \mathcal{O}(x^2)$ for some suitable constants d_0 and d_1.

Carefully collecting relevant terms and utilizing the relationships $b_1 = \frac{3}{2}c_1$, $b_2 = 2c_2$, we can see that $d_0 = u_0 b_1 + u_1 b_0^2 + \frac{b_2}{4b_0} - \frac{b_1^2}{3b_0^2}$ from which we obtain the expression in the statement. ■

In the previous theorem, it is possible to compute d_1 explicitly, in which case b_3 and c_3 are required. Depending on the time scale for which the user applies implied volatility expansions, however, the expansion $I(x)^2 + 2I(x)^4 d_0 t$ might suffice. In our numerical tests in the next section, we use d_0 only with the quadratic expansion in Theorem 1 for $I(x)$.

4 Numerical Results

In this section, we conduct several numerical comparisons to test the effectiveness of the derived formulae in Theorems 1 and 2. The limiting implied volatility $I(x)$ is also denoted by σ_0 and the small-time smile by σ_t. Our experiments use the following parameters: $\beta = 3$, $S_0 = 100$, $r = 0$, $T = 0.1, 0.05, 0.01$, and

$$M = \begin{bmatrix} -3 & 0 \\ 0 & -4 \end{bmatrix}, \quad Q = \begin{bmatrix} .5 & 0 \\ 0 & .25 \end{bmatrix}, \quad R = \begin{bmatrix} -.5 & 0 \\ 0 & -.5 \end{bmatrix}, \quad \Sigma_0 = \begin{bmatrix} .09 & 0 \\ 0 & .09 \end{bmatrix}.$$

Since the original paper by Da Fonseca et al. [8] presents detailed sensitivity analyses of parameters on option prices, we focus on the numerical performance of our implied volatility expansions. In order to assess the accuracy of σ_0 and σ_t, we numerically calculate call option prices based on the WMSV model via the widely used Fourier transform method, which we refer the reader to any of many available references such as Da Fonseca et al. [8], Carr and Madan [2], or Duffie et al. [10] to name a few. Nevertheless, for the reader's convenience, we document the details of our Matlab implementation. The documentation can be found on the corresponding author's personal web page or can be requested via email.

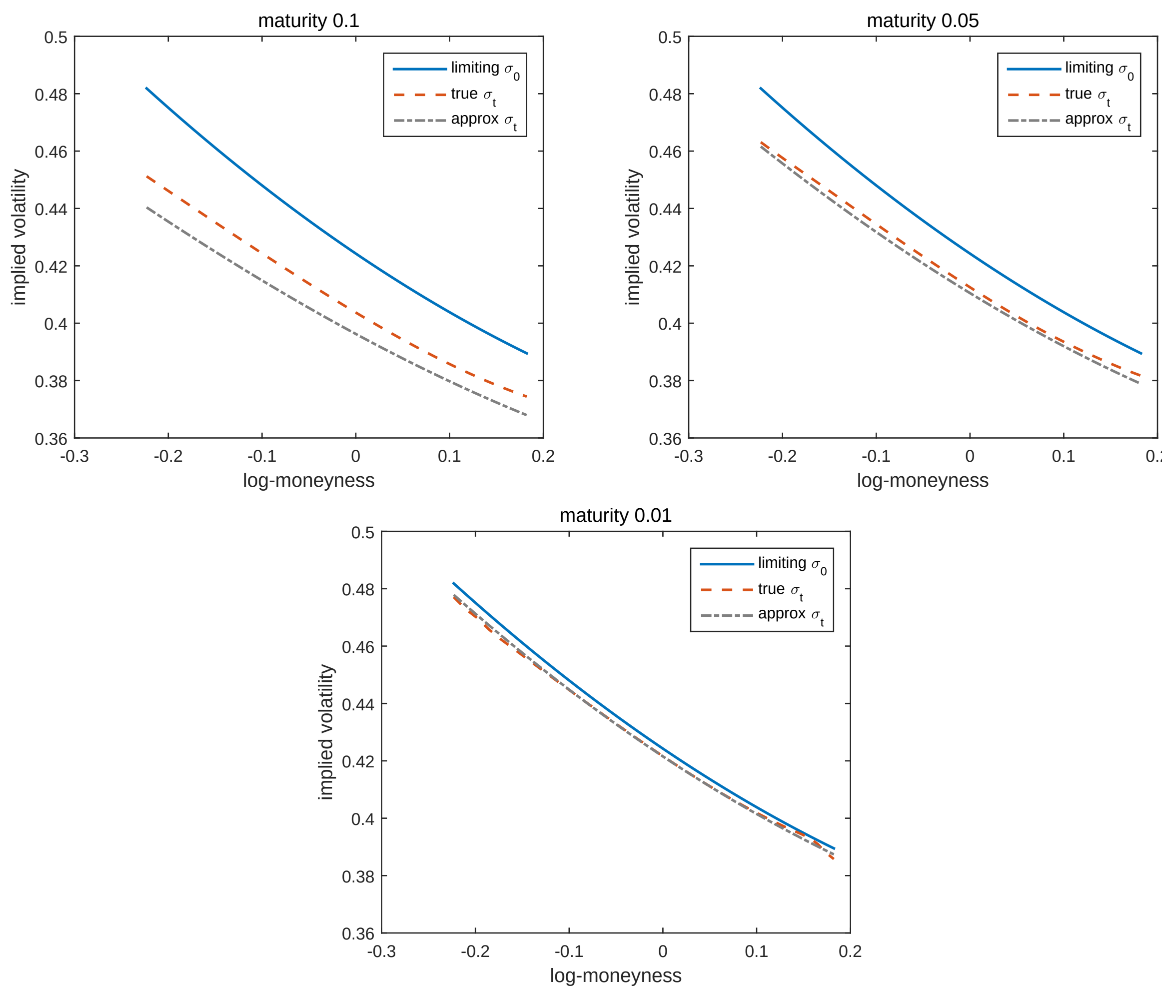

Figure 1: Comparison of true implied volatility, limiting implied volatility σ_0, and small-time smile σ_t for (1) option maturity 0.1 (2) option maturity 0.05 (3) option maturity 0.01.

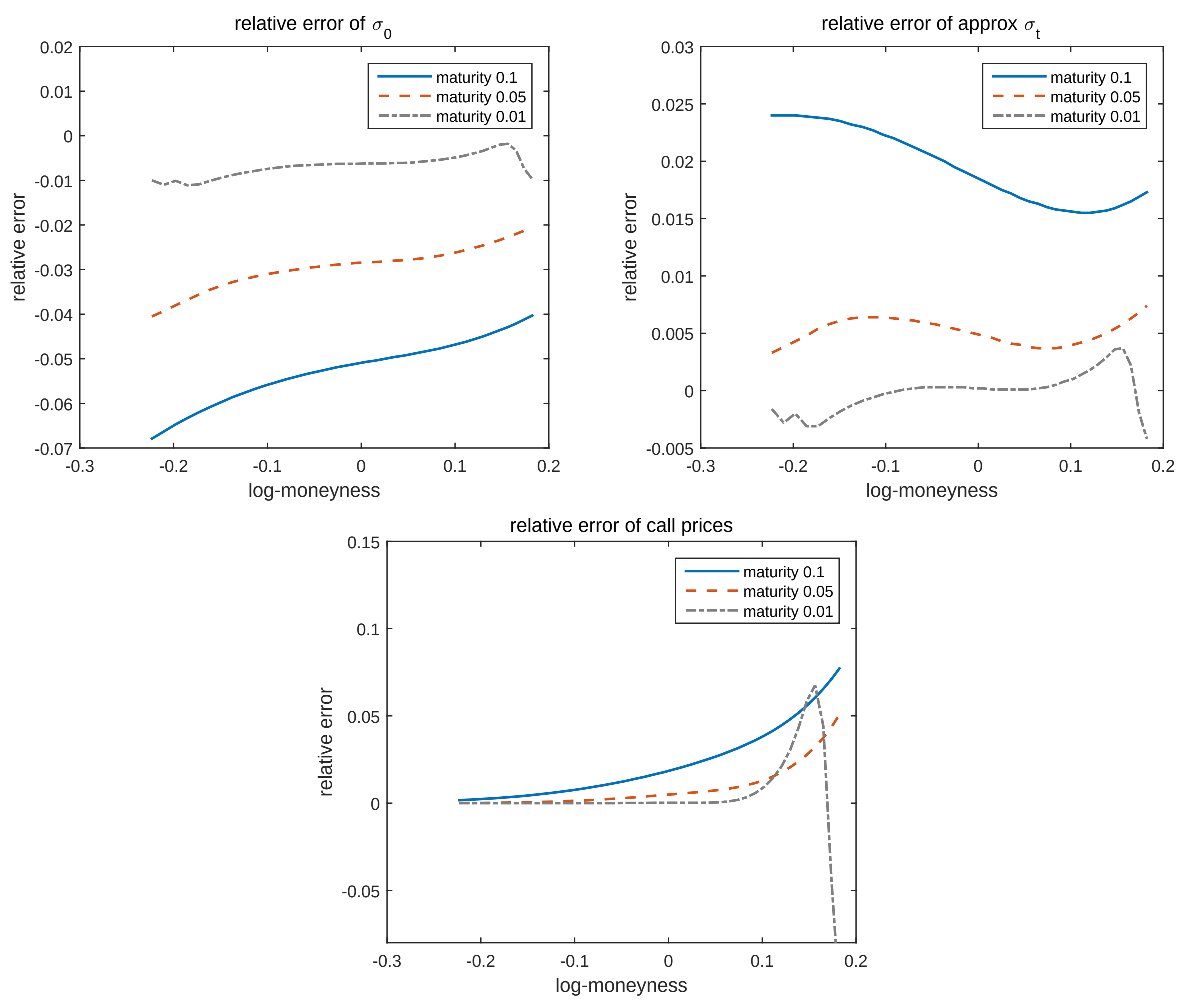

Figure 2: Relative differences `(true_val - approx_val)/true_val` for (1) limiting implied volatility σ_0 (2) small-time smile σ_t (3) true call prices vs. approximate call prices from σ_t.

5 Implications for the Pension Fund Management

From the manager's viewpoint, there are several potentially useful directions that implied volatility asymptotics can be utilized. Their primary use lies in improving calibration efficiency. By matching asymptotic implied volatilities with model parameters to market implied volatilities, we obtain numerically more stable solutions when compared to matching model option prices with market option prices as we circumvent computational steps in option pricing. The computational gain is large. A similar example where approximate schemes prove to be powerful is American options. In managing portfolios of financial products, it is vital to maintain suitably calibrated models and the results in this paper could be of use for products whose underlyings are assumed to follow the multi-factor volatility Heston model.

Another, possibly more important, application of implied volatility asymptotics is that they provide information about derivatives prices in certain regions. For given model parameters, the small-time smile in this paper tells us the behavior of vanilla option prices near expiry. More generally, if the price of a derivatives product can be written using implied volatility, the price movement can be better understood by decomposing it in terms of option greeks (e.g., vega) and implied volatility asymptotics. A related application is optimal construction of stock portfolios. It has been well known that option implied information can improve the portfolio performance compared to when historical volatilities are used in the model. As a future research topic, it would be interesting to see the explanatory power of the near-expiry implied volatility asymptotics on the behavior of optimal portfolios where assets follow the geometric Brownian motion.

6 Conclusion

This paper showed that it is possible to extend the large deviations approach and the Laplace expansion approach for the small-time asymptotics for implied volatilities under the Heston stochastic volatility model to its multi-dimensional version. For this, all the arguments underlying the existing approaches were re-examined and some proofs were simplified. It should also be acknowledged that all the analyses were possible due to the recent developments of the theory of affine processes in general. The tractability of the resulting formulae depends on the specific assumptions on model parameters, that is, RQ and M are symmetric. Such assumptions do not restrict the flexibility of the Wishart process based model. However, the full consequences of those assumptions are beyond the scope of this paper.

References

[1] Benabid, A., Bensusan, H., and El Karoui, N. (2010) Wishart stochastic volatility: Asymptotic smile and numerical framework. Working Paper.

[2] Carr, P. and Madan, D. (1999) Option valuation using the fast Fourier transform. *Journal of Computational Finance*, **2**:61–73.

[3] Choi, C. H. (1990) A survey of numerical methods for solving matrix Riccati differential equations. *IEEE Proceedings on Southeastcon*, 696–700, vol.2.

[4] Christoffersen, P., Heston, S., and Jacobs, K. (2009) The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, **55**:1914–1932.

[5] Da Fonseca, J. (2016) On moment non-explosions for Wishart-based stochastic volatility models. *European Journal of Operational Research*, **254**:889–894.

[6] Da Fonseca, J., Gnoatto, A., and Grasselli, M. (2015) Analytic pricing of volatility-equity options within Wishart-based stochastic volatility models. *Operations Research Letters*, **43**:601–607.

[7] Da Fonseca, J. and Grasselli, M. (2011) Riding on the smiles. *Quantitative Finance*, **11**:1609–1632.

[8] Da Fonseca, J., Grasselli, M., and Tebaldi, C. (2008) A multifactor volatility Heston model. *Quantitative Finance*, **8**:591–8604.

[9] Dembo, A. and Zeitouni, O. (1998) *Large Deviations Techniques and Applications*. 2nd edition, Springer, New York.

[10] Duffie, D., Pan, J., and Singleton, K. (2000) Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, **68**:1343–1376.

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

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

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

[14] Forde, M., Jacquier, A., and Mijatović, A. (2011) A note on essential smoothness in the Heston model. *Finance and Stochastics*, **15**:781–784.

[15] Fouque, J. P., Papanicolaou, G., Sircar, R., and Solna, K. (2003) Singular perturbations in option pricing. *SIAM Journal on Applied Mathematics*, **63**:1648–1665.

[16] Fouque, J. P., Papanicolaou, G., Sircar, R., and Solna, K. (2003) Multiscale stochastic volatility asymptotics. *Multiscale Modeling & Simulation*, **2**:22–42.

[17] Gauthier, P. and Possamai, D. (2009) Prices expansions in the Wishart model. *The IUP Journal of Computational Mathematics*, **4**:44–71.

[18] Gnoatto, M. and Grasselli, M. (2014) The explicit Laplace transform for the Wishart process. *Journal of Applied Probability*, **51**:640–656.

[19] Gourieroux, C. and Sufana, R. (2010) Derivative pricing with Wishart multivariate stochastic volatility. *Journal of Business & Economic Statistics*, **28**:438–451.

[20] Grasselli, M. and Tebaldi, C. (2008) Solvable affine term structure models. *Mathematical Finance*, **18**:135-153.

[21] Gruber, P., Tebaldi, C., and Trojani, F. (2010) Three make a smiledynamic volatility, skewness and term structure components in option valutation. Working Paper.

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

[23] Kang, C., Kang, W., and Lee, J. M. (2017) Exact simulation of the Wishart multidimensional stochastic volatility model. *Operations Research*, **65**:1190-1206.

[24] Keller-Ressel, M. and Mayerhofer, E. (2015) Exponential moments of affine processes. *Annals of Applied Probability*, **25**: 714–752.

[25] Lee, R. (2004) The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, **14**:469–480.

[26] Lee, R. and Wang, D. (2012) Displaced lognormal volatility skews: Analysis and applications to stochastic volatility simulations. *Annals of Finance*, **8**:159–181.

[27] Reid, W. T. (1972) *Riccati Differential Equations*. Academic Press, New York.

[28] Rockafellar, R. T. (1970) *Convex Analysis*. Princeton University Press, Princeton.

[29] Olver, F. W. (1974) *Asymptotics and Special Functions*. Academic Press, New York.