Projection on a Quadratic Model by Asymptotic Expansion with an Application to LMM Swaption.

A. Antonov* and T. Misirpashaev†

June 16, 2008

Abstract

We develop a technique of parameter averaging and Markovian projection on a quadratic volatility model based on a term-by-term matching of the asymptotic expansions of option prices in volatilities. In doing so, we revisit the procedure of asymptotic expansion and show that the use of the product formula for iterated Itô integrals leads to a considerable simplification in comparison with the approach currently prevalent in the literature. Results are applied to the classic problem of LIBOR Market Model (LMM) swaption pricing. We confirm numerically that the retention of the quadratic term gives a marked improvement over the standard approximation based on the projection on a displaced diffusion.

Keywords: asymptotic expansion, Markovian projection, skew averaging, quadratic volatility model, LIBOR Market Model, swaption, Wiener chaos

1 Introduction

The purpose of this work is to enrich the set of tools available for calibration of multidimensional models by merging two powerful techniques of applied quantitative finance: asymptotic expansion for stochastic differential equations (SDEs) and model mapping. The idea of model mapping is to replace a more complicated model by a simpler one. Two common incarnations of model mapping identified by Piterbarg include parameter averaging (replacing time-dependent coefficients by time-independent ones [17]) and Markovian projection (replacing a non-Markovian model by a Markovian one [18]).

As in [3], [4] we start with a generic statement of the problem of finding an efficient way to compute European options $E[(S(T)-K)^+]$ for a complicated, generally non-Markovian process, $S(t)$, following an SDE in the martingale measure, $dS(t) = \theta(t) \cdot dW(t)$, with a volatility process $\theta(t)$ obeying certain restrictions to make an expansion in small volatilities possible. We then look for a simpler, Markovian process, $S^*(t)$, with marginal distributions at fixed t sufficiently close to those of the original process. Such a process exists under mild technical restrictions. Indeed, Gyöngy [9] showed that it is sufficient to require $S^*(0) = S(0)$ and $dS^*(t) = \Theta(S^*(t), t)dW(t)$, where $|\Theta(x,t)|^2 = E[|\theta(t)|^2 | S(t) = x]$, for the distributions of $S(t)$ and $S^*(t)$ to be identical for any t.

*Numerix Software Ltd., 41 Eastcheap, 2nd Floor London EC3M 1DT UK

†Bank of America Securities, One Bryant Park, New York NY 10036 USA

The calculation of the conditional expectation of a stochastic instantaneous variance presents a challenge for the practical use of Gyöngy's result. Several techniques blending rigorous theory, analytical approximations, and numerics have been proposed so far. We mention geometrical methods [5] (see also [10] for a more extended development), reduction to unconditional expectations enabled by Malliavin calculus [8], and orthogonal projection on a low-dimensional subspace of local volatility functions [3], [4]. In this work we revisit the idea of projecting on a low-dimensional subspace but do it differently, namely by looking for a local volatility of a given functional form which would reproduce exactly several leading terms in the asymptotic expansion of option prices in small volatilities.

It is well understood that the local volatility function should be at least quadratic in order to have a chance to adequately describe the volatility smile (unless a separate linear volatility model is used for each value of option strike). We therefore focus our attention on the quadratic model with time-dependent coefficients,

$$\Theta(x,t) = \left(1 + \beta(t)(x - x_0) + \frac{1}{2}\gamma(t)(x - x_0)^2\right)\lambda(t),$$

or time-independent (apart from a common factor) coefficients,

$$\Theta(x,t) = \left(1 + \bar{\beta}(x - x_0) + \frac{1}{2}\bar{\gamma}(x - x_0)^2\right)\lambda(t),$$

treating the averaging of coefficients over time as a particular case of model projection, developed for the linear case by Piterbarg [17] and recently extended to the quadratic case by Andersen and Hutchings [2]. Quadratic volatility model leads to a closed form option pricing formula but has not been widely used because of a known difficulty in dealing with the cases where solutions fail to be martingales. However, after the exhaustive analysis of [1], solving the model no longer presents a problem.

We apply projection on a quadratic volatility model to swaption pricing in the LIBOR Market Model (LMM) [6], [13]. Asymptotic expansion for the swaption price was derived by Kawai [14] (see also [15]). We confirmed the results of [14] and improved their practical usability not only by clarifying the structure of resulting analytical expressions and simplifying multidimensional integrals and summations but also by using the asymptotic expansion as an intermediate result for mapping on a quadratic volatility model. We found that using the exact solution of the mapped quadratic volatility model significantly improves the accuracy of option pricing in comparison with a direct use of the asymptotic expansion or with a strike-dependent log-normal approximation introduced in [14].

The paper is organized as follows. We begin with a review of the asymptotic expansion method in Sect. 2. We achieve a considerable simplification of the calculations in comparison with most of recently published works by a direct use of the product formula for iterated Itô integrals for the calculation of unconditional averages. In Sect. 3.1 we work out the first three terms of the asymptotic expansion for the quadratic volatility model, which serves both as a preparation for more complicated calculations in multifactor models and the basis for the mapping of more complicated models on a quadratic volatility model. A simple version of model mapping is parameter averaging. We show in Sect. 3.2 that a straightforward approach to the averaging results in a strike-dependent curvature parameter $\bar{\gamma}$. An alternative approach resulting in strike-independent parameters is presented in Sect. 3.3.

In Sect. 4.1 we develop the asymptotic expansion for a so-called *separable* model template, which covers many models encountered in practice, including the LMM. Projection on the quadratic volatility model is discussed in Sect. 4.2, followed by a specialization of the results to the case of LMM in Sect. 4.3. Numerical results are given in Sect. 5, consisting of the tests for parameter averaging in the quadratic volatility model in Sect. 5.1 and LMM swaption pricing in Sect. 5.2. We conclude in Sect. 6. The general technique of calculating the averages using the product formula for iterated Itô integrals is presented in Appendix A. Additional details for the analytics of LMM swaption pricing, including a comparison with [14], are found in Appendix B.

2 Asymptotic expansion for option prices

The asymptotic expansion method applies to an SDE with a small parameter ε and constitutes a technique for finding successive perturbative corrections in powers of ε for the functionals of the solution (see [20], [21] for a rigorous mathematical introduction). The method found numerous applications in quantitative finance, e.g., [14]–[16] and [19]. In this work we do not address mathematical foundations of the method and focus on obtaining manageable expressions for the coefficients of the asymptotic expansion of option prices in an efficient way. The validity of the approximations is ultimately tested numerically.

We begin our examination with a Markovian local martingale,

$$dX_\varepsilon(t) = \varepsilon\sigma(t, X_\varepsilon(t))\, dW(t), \tag{1}$$

with a formal parameter ε which makes it possible to treat volatility as a small correction in the limit $\varepsilon \ll 1$ (although nothing will prevent us from setting $\varepsilon = 1$ in the final result and checking numerically if the approximation is still viable). This case covers the quadratic volatility model. A much more general case, sufficient for the application to LMM, will be considered in Sect. 4.

We proceed by writing Eq. (1) in an integral form,

$$X_\varepsilon(T) = X_\varepsilon(0) + \varepsilon \int_0^T dW(t)\, \sigma(t, X_\varepsilon(t)), \tag{2}$$

and taking the formal derivative of order n with respect to ε,

$$\frac{d^n X_\varepsilon(T)}{d\varepsilon^n} = n \int_0^T dW(t)\, \frac{d^{n-1}}{d\varepsilon^{n-1}}\sigma(t, X_\varepsilon(t)) + \varepsilon \int_0^T dW(t)\, \frac{d^n}{d\varepsilon^n}\sigma(t, X_\varepsilon(t)). \tag{3}$$

Setting $\varepsilon = 0$, we get a recursive formula for the derivatives $Y_n(T) = \frac{d^{(n+1)}X}{d\varepsilon^{(n+1)}}|_{\varepsilon=0}$ entering the formal Taylor series at $\varepsilon = 0$,

$$Y_n(T) = (n+1) \int_0^T dW(t)\ \frac{d^n}{d\varepsilon^n}\sigma(t, X_\varepsilon(t))\bigg|_{\varepsilon=0}. \tag{4}$$

Explicit expressions for the first three terms are

$$Y_0(T) = \int_0^T dW(t)\, \sigma(t, X_0), \tag{5}$$

$$Y_1(T) = 2 \int_0^T dW(t)\, \sigma'(t, X_0)\, Y_0(t), \tag{6}$$

$$Y_2(T) = 3 \int_0^T dW(t)\, \left(\sigma''(t, X_0)\, {Y_0}^2(t) + \sigma'(t, X_0)\, Y_1(t)\right), \tag{7}$$

where we introduced the notations $\sigma'(t, X) = \partial\sigma(t, X)/\partial X$ and $X_0 = X_\varepsilon(t)|_{\varepsilon=0} \equiv X_\varepsilon(0)$.

In this way we have obtained explicit expressions for first several terms of the Taylor series for $X_\varepsilon(T)$ (understood at least in the formal sense) in terms of iterated Itô integrals of the model volatility function $\sigma(t, X)$ and its derivatives,

$$X_\varepsilon(T) = X_0 + \varepsilon\, Y_0(T) + \frac{1}{2}\,\varepsilon^2\, Y_1(T) + \frac{1}{6}\,\varepsilon^3\, Y_2(T) + O(\varepsilon^4). \tag{8}$$

This series is the starting point for the remainder of this section, which will rely only on certain properties of $Y_0(T)$, $Y_1(T)$, and $Y_2(T)$ but not on their specific form, and will also be applied to the case of a non-Markovian model of Sect. 4. The key assumption that we make (verified in all cases to which the result will be applied) is for $Y_0(T)$ to be a Gaussian random variable with mean value 0 and some variance which we denote $v(T)$. This is obviously so for $Y_0(T)$ defined by Eq. (5) with $v(T) = \int_0^T dt|\sigma(t, X_0)|^2$ due to Itô isometry. We will also require existence of certain expected values, as will be seen from the derivation below.

We use Eq. (8) to calculate the corresponding terms in the expansion for the call option, defined as $\mathcal{C}(T,k) = \varepsilon^{-1}\, E[(X_\varepsilon(T) - K)^+]$ with a centered and ε-resized strike $K = X_0 + \varepsilon\, k$,

$$\mathcal{C}(T,k) = E\left[\left(Y_0(T) + \frac{1}{2}\,\varepsilon\, Y_1(T) + \frac{1}{6}\,\varepsilon^2\, Y_2(T) + O(\epsilon^3) - \, k\right)^+\right]. \tag{9}$$

We introduce the moment-generating function of $\varepsilon^{-1}(X_\varepsilon(T) - X_0)$ in the plane of complex-valued ξ,

$$\Phi_\varepsilon(\xi) = E\left[e^{\xi(X_\varepsilon(T)-X_0)/\varepsilon}\right] = E\left[e^{\xi\left(Y_0(T)+\frac{1}{2}\,\varepsilon\, Y_1(T)+\frac{1}{6}\,\varepsilon^2\, Y_2(T)+O(\varepsilon^3)\right)}\right]. \tag{10}$$

The option price can be written using inverse Laplace transform as a contour integral,

$$\mathcal{C}(T,k) = \frac{1}{2\pi i}\int_{C^+} \frac{d\xi}{\xi^2}\,\Phi_\varepsilon(\xi)\, e^{-k\xi}, \tag{11}$$

where the contour of integration C^+ is chosen to go along the imaginary axis from $-i\infty$ to $+i\infty$ far from the origin, bypassing the origin counterclockwise along a semicircle of a sufficiently small radius to stay to the left from all possible singularities of $\Phi_\varepsilon(\xi)$ with positive real part.

Expanding the exponent in (10), we obtain

$$\begin{aligned} \Phi_\varepsilon(\xi) &= E\left[e^{\xi\, Y_0(T)}\left(1 + \frac{1}{2}\,\varepsilon\,\xi\, Y_1(T) + \frac{1}{6}\,\varepsilon^2\,\xi\, Y_2(T) + \frac{1}{8}\varepsilon^2\,\xi^2\, Y_1^2(T) + O(\varepsilon^3)\right)\right] \\ &= \Phi_0(\xi) + \varepsilon\,\Phi_1(\xi) + \varepsilon^2\,\Phi_2(\xi) + O(\varepsilon^3), \end{aligned} \tag{12}$$

where

$$\Phi_0(\xi) = E\left[e^{\xi\, Y_0(T)}\right], \tag{13}$$

$$\Phi_1(\xi) = \frac{1}{2}\,\xi\, E\left[e^{\xi\, Y_0(T)} Y_1(T)\right], \tag{14}$$

$$\Phi_2(\xi) = \frac{1}{6}\,\xi\, E\left[e^{\xi\, Y_0(T)} Y_2(T)\right] + \frac{1}{8}\,\xi^2\, E\left[e^{\xi\, Y_0(T)} Y_1^2(T)\right]. \tag{15}$$

With the assumption of a Gaussian $Y_0(T)$, the leading term delivers the Gaussian approximation with the moment generating function

$$\Phi_0(\xi) = E\left[e^{\xi\, Y_0(T)}\right] = e^{\frac{1}{2}\,\xi^2\, v(T)}, \tag{16}$$

and density

$$P_G(k) = \frac{1}{2\pi i}\int_{C^+} d\xi\,\Phi_0(\xi)\, e^{-k\xi} = \frac{e^{-\frac{1}{2}\frac{k^2}{v(T)}}}{\sqrt{2\pi\, v(T)}}, \tag{17}$$

while $\Phi_1(\xi)$ and $\Phi_2(\xi)$ lead to corrections to the Gaussian approximation. The expectations can be evaluated analytically using the Wiener chaos as described in Appendix A and turn out to be equal to a product of $\Phi_0(\xi)$ and a polynomial in ξ. More specifically, we will see that for the cases of our interest, the terms $\Phi_1(\xi)$ and $\Phi_2(\xi)$ can be written as

$$\Phi_1(\xi) = \phi_{13}\,\xi^3\,\Phi_0(\xi), \tag{18}$$
$$\Phi_2(\xi) = \left(\phi_{22}\,\xi^2 + \phi_{24}\,\xi^4 + \phi_{26}\,\xi^6\right)\Phi_0(\xi), \tag{19}$$

with the coefficients $\phi_{13}, \phi_{22}, \phi_{24}, \phi_{26}$ expressible via the model parameters.

This brings us into a position to evaluate the terms in the expansion of the option price,

$$\mathcal{C}(T,k) = C_0(T,k) + \varepsilon\, C_1(T,k) + \varepsilon^2 C_2(T,k) + O(\varepsilon^3), \tag{20}$$

including the leading Gaussian term (option price in the Bachelier model),

$$C_0(T,k) = \frac{1}{2\pi i}\int_{C+}\frac{d\xi}{\xi^2}\,\Phi_0(\xi)\,e^{-k\xi} = \sqrt{\frac{v(T)}{2\pi}}e^{-k^2/2v(T)} - k\mathcal{N}(-k/\sqrt{v(T)}), \tag{21}$$

and two corrections,

$$C_1(k) = \frac{1}{2\pi i}\int_{C+}\frac{d\xi}{\xi^2}\,\phi_{13}\,\xi^3\,\Phi_0(\xi)\,e^{-k\xi} = \phi_{13}\,h_1(k)\,P_G(k), \tag{22}$$
$$\begin{aligned} C_2(k) &= \frac{1}{2\pi i}\int_{C+}\frac{d\xi}{\xi^2}\,\left(\phi_{22}\,\xi^2 + \phi_{24}\,\xi^4 + \phi_{26}\,\xi^6\right)\Phi_0(\xi)\,e^{-k\xi} \\ &= \left(\phi_{22} + \phi_{24}\,h_2(k) + \phi_{26}\,h_4(k)\right)\,P_G(k). \end{aligned} \tag{23}$$

Here $\mathcal{N}(x) = (1/\sqrt{2\pi})\int_{-\infty}^{x} dt\, e^{-t^2/2}$ is the standard normal CDF, and $h_n(x)$ are rescaled Hermite polynomials,

$$h_n(x) = v(T)^{-\frac{n}{2}}\mathrm{H}_n\left(\frac{x}{\sqrt{v(T)}}\right). \tag{24}$$

We recall that Hermite polynomials are defined by

$$\mathrm{H}_n(x) = (-1)^n e^{\frac{1}{2}x^2}\frac{d^n}{dx^n}\left(e^{-\frac{1}{2}x^2}\right), \tag{25}$$

with the first few being $\mathrm{H}_0(x) = 1$, $\mathrm{H}_1(x) = x$, $\mathrm{H}_2(x) = x^2 - 1$, $\mathrm{H}_3(x) = x^3 - 3\,x$, $\mathrm{H}_4(x) = x^4 - 6\,x^2 + 3$, so that

$$h_1(k) = k/v(T), \tag{26}$$
$$h_2(k) = (k^2 - v(T))/v^2(T), \tag{27}$$
$$h_4(k) = (k^4 - 6k^2 v(T) + 3v^2(T))/v^4(T). \tag{28}$$

Thus we succeeded in finding closed form expressions for the two highest order corrections to the Gaussian approximation for the option price in a very generic form in terms of the Gaussian density (17), Hermite polynomials (24), and coefficient functions ϕ_{ij} resulting from the expectations (14–15). In comparison with the approach of [14], we avoided the step of computing corrections to the density, which makes the analytical structure of the corrections to the option price more transparent and easier to handle. We now proceed to applications, where the coefficients ϕ_{ij} will be computed from model parameters.

3 Quadratic volatility model

We consider a quadratic volatility model with time-dependent parameters, which we write in the form

$$dX_\varepsilon(t) = \varepsilon\left(1 + \beta(t)\,\Delta X_\varepsilon(t) + \frac{1}{2}\gamma(t)\,\Delta X_\varepsilon^2(t)\right)\lambda(t)\,dW(t), \tag{29}$$

where $\Delta X_\varepsilon(t) = X_\varepsilon(t) - X_\varepsilon(0)$. This form is convenient because it separates the time-dependent volatility $\lambda(t)$, which can be easily eliminated by a time change in the Brownian motion, and leaves no other parameters in the volatility function evaluated at $X_\varepsilon(t) \equiv X_\varepsilon(0)$.

3.1 Asymptotic expansion

The model is of the type (1), therefore we can immediately identify $Y_0(T)$, $Y_1(T)$, and $Y_2(T)$ in the expansion (8) from Eqs. (5–7) with $\sigma(t,X) = (1 + \beta(t)(X - X_0) + \frac{1}{2}\gamma(t)(X - X_0)^2)\lambda(t)$. We get

$$Y_0(T) = \int_0^T dW(t)\,\lambda(t), \tag{30}$$

$$Y_1(T) = 2\int_0^T dW(t)\,\lambda(t)\,\beta(t)\,Y_0(t), \tag{31}$$

$$Y_2(T) = 3\int_0^T dW(t)\,\lambda(t)\gamma(t)\,{Y_0}^2(t) + 3\int_0^T dW(t)\,\lambda(t)\beta(t)\,Y_1(t). \tag{32}$$

The expectations entering the expressions (14) and (15) can be evaluated in an ad-hoc manner with repeated applications of Itô's lemma, however, the Wiener chaos technique leads to final answers almost instantly. In Appendix A we work out this technique specifically for the calculation of the expectations of the type

$$E\left[\exp\left(\xi\int_0^T dW(t)\lambda(t)\right)P\right],$$

where P is a polynomial in one or several iterated Itô integrals with respect to $dW(t)$. The result is

$$E[e^{\xi Y_0(T)}Y_1(T)] = 2\xi^2 c(T)\Phi_0(\xi), \tag{33}$$

$$E[e^{\xi Y_0(T)}Y_2(T)] = 6\left(\xi g_1(T) + \xi^3\left(g_2(T) + b(T)\right)\right)\Phi_0(\xi), \tag{34}$$

$$E[e^{\xi Y_0(T)}Y_1^2(T)] = 8\left(a_1(T) + \xi^2(a_2(T) - b(T)) + \frac{1}{2}\xi^4 c^2(T)\right)\Phi_0(\xi), \tag{35}$$

with $\Phi_0(\xi)$ given by Eq. (16) and the following notations for the time-dependent coefficients (note that the variance $v(T)$ is consistent with Eq. (16)),

$$\begin{aligned}
v(T) &= \int_0^T dt\, \lambda^2(t), && (36)\\
c(T) &= \int_0^T dt\, \lambda^2(t)\, \beta(t)\, v(t), && (37)\\
a_1(T) &= \frac{1}{2}\int_0^T dt\, \lambda^2(t)\beta^2(t)\, v(t), && (38)\\
a_2(T) &= 3\int_0^T dt\, \lambda^2(t)\beta(t)\, c(t) + \frac{1}{2}\int_0^T dt\, \lambda^2(t)\beta^2(t)\, v^2(t), && (39)\\
b(T) &= \int_0^T dt\, \lambda^2(t)\beta(t)\, c(t), && (40)\\
g_1(T) &= \frac{1}{2}\int_0^T dt\, \lambda^2(t)\gamma(t)\, v(t), && (41)\\
g_2(T) &= \frac{1}{2}\int_0^T dt\, \lambda^2(t)\gamma(t)\, v^2(t). && (42)
\end{aligned}$$

Now we can complete the calculation of (14) and (15) and convince ourselves that the result has the form of Eqs. (18–19) with the coefficients

$$\begin{aligned}
\phi_{13} &= c(T), && (43)\\
\phi_{22} &= a_1(T) + g_1(T), && (44)\\
\phi_{24} &= a_2(T) + g_2(T), && (45)\\
\phi_{26} &= \frac{1}{2}c^2(T). && (46)
\end{aligned}$$

The result for the correction to options prices follows from Eqs. (22–23),

$$\begin{aligned}
C_1(k) &= c(T)\, h_1(k) P_G(k), && (47)\\
C_2(k) &= \left(\frac{1}{2}c^2(T)\, h_4(k) + (g_2(T) + a_2(T))\; h_2(k) \; + g_1(T) + a_1(T)\right) P_G(k). && (48)
\end{aligned}$$

Note that in contrast with [14], [15], [19] we skipped the superfluous step of computing averages conditional on $Y_0(T)$, going straight for the unconditional averages and thereby achieving significant operational and notational simplifications.

3.2 Parameter averaging

While it is possible to use option price corrections (47–48) directly, we prefer to use them only as a matching target to calibrate a simpler model with the same $C_0(k)$, $C_1(k)$ and $C_2(k)$. The simpler model can then be solved exactly, without asymptotic expansion. In this work, the target model is a version of the quadratic volatility model

$$dX(t) = \left(1 + \bar{\beta}\, \Delta X(t) + \frac{1}{2}\bar{\gamma}\, \Delta X^2(t)\right) \lambda(t)\, dW(t), \tag{49}$$

with time-independent parameters $\bar{\beta}$ and $\bar{\gamma}$. It is convenient to keep a time-dependent $\lambda(t)$ because it can be easily eliminated by a time change and therefore does not introduce any difficulties in the solution of the quadratic volatility model.

The asymptotic expansion of the time-independent (apart from $\lambda(t)$) quadratic volatility model is a special case of the expansion derived in the previous section. Using bars to denote the corresponding coefficients, we find[1]

$$
\begin{aligned}
\bar{c}(T) &= \frac{1}{2}\bar{\beta}\, v^2(T), && (50)\\
\bar{a}_1(T) &= \frac{1}{4}\bar{\beta}^2\, v^2(T), && (51)\\
\bar{a}_2(T) &= \frac{2}{3}\bar{\beta}^2\, v^3(T), && (52)\\
\bar{g}_1(T) &= \frac{1}{4}\bar{\gamma}\, v^2(T), && (53)\\
\bar{g}_2(T) &= \frac{1}{6}\bar{\gamma}\, v^3(T), && (54)
\end{aligned}
$$

where $v(T) = \int_0^T dt\lambda^2(t)$ does not need a bar as it remains unchanged together with the Gaussian approximation (21) to option prices.

Finding the values of $\bar{\beta}$ and $\bar{\gamma}$ reproducing the same corrections $C_1(k)$ and $C_2(k)$ to option prices as time-dependent $\beta(t)$ and $\gamma(t)$ can be regarded as parameter averaging in the sense discussed in [17] and [2]. Equality of the first corrections gives $c(T) = \bar{c}(T)$, which immediately reproduces the averaging formula for the skew in the displaced-diffusion model derived in [17] and [3] from different considerations,

$$
\bar{\beta} = \frac{\int_0^T dt\, \lambda^2(t)\, v(t)\, \beta(t)}{\int_0^T dt\, \lambda^2(t)\, v(t)}. \tag{55}
$$

We note that this averaging is *strike-independent*.

The averaging formula of the curvature parameter, $\bar{\gamma}$, follows from the equality of the second order corrections for the option price, which gives

$$
\left(g_2(T) + a_2(T)\right)\, h_2(k) \; + g_1(T) + a_1(T) = \left(\bar{g}_2(T) + \bar{a}_2(T)\right)\, h_2(k) + \bar{g}_1(T) + \bar{a}_1(T). \tag{56}
$$

This leads to an expression for the averaged $\bar{\gamma}$,

$$
\bar{\gamma} = 2\frac{\left(g_2(T) + a_2(T) - \bar{a}_2(T)\right) h_2(k) + g_1(T) + a_1(T) - \bar{a}_1(T)}{v^2(T)\left(\frac{1}{3}\, v(T)\, h_2(k) + \frac{1}{2}\right)}, \tag{57}
$$

which turns out to be a ratio of two quadratic functions of strike k. While it is certainly a less convenient result, meaning that the simplified model will have to be different for different values of option strike, practical applications are not necessarily severely hampered if the simplified model can be solved efficiently. However we will show in Sect. 3.3 that it is also possible to find an efficient strike-independent quadratic volatility model.

3.3 Strike-independent parameter averaging with an adjusted volatility function

We saw in Sect 3.2 that a term-by-term matching of asymptotic expansions for option prices for time-dependent and time-independent quadratic volatility models leads to a strike-independent averaged linear term but strike-dependent averaged quadratic term. It turns out that a fully

[1] Iterated integrals were simplified using Eq. (159) with $n = 2, 3$.

strike-independent approximation of the same or better quality is possible if we write the target model in the form,

$$dX_\varepsilon(t) = \varepsilon \left(1 + \bar\beta\, \Delta X_\varepsilon(t) + \frac{1}{2}\bar\gamma \Delta X_\varepsilon(t)\right) \left(1 + \varepsilon^2 \alpha\right) \lambda(t)\, dW(t), \tag{58}$$

introducing an additional matching parameter α and effectively allowing for a second order correction to the local volatility function $\lambda(t)$. We call this model a quadratic model with adjusted volatility to distinguish it from the quadratic volatility model considered in Sect. 3.2. Of course, after substituting $\varepsilon = 1$ in the final results, any functional difference between the models of this and previous sections disappears.

The coefficients in the expansion (8) are

$$\begin{aligned}
Y_0(T) &= \int_0^T dW(t)\, \lambda(t), && (59)\\
Y_1(T) &= 2 \int_0^T dW(t)\, \lambda(t)\, \bar\beta\, Y_0(t), && (60)\\
Y_2(T) &= 3 \int_0^T dW(t)\, \lambda(t) \bar\gamma\, {Y_0}^2(t) + 3 \int_0^T dW(t)\, \lambda(t) \bar\beta\, Y_1(t) + 6\alpha\, Y_0(T). && (61)
\end{aligned}$$

In comparison with the expressions (30–32) for the standard quadratic volatility model, there is an additional term $6\alpha\, Y_0(T)$ in $Y_2(T)$, leading to

$$E[e^{\xi Y_0(T)} Y_2(T)] = 6 \left(\xi(g_1(T) + \alpha\, v(T)) + \xi^3 \left(\bar g_2(T) + \bar b(T)\right)\right) \Phi_0(\xi). \tag{62}$$

The corrections to the option price take the standard form,

$$\begin{aligned}
C_1(k) &= \bar\phi_{13}\, h_1(k) P_G(k), && (63)\\
C_2(k) &= \left[\bar\phi_{22} + \bar\phi_{24}\, h_2(k) + \bar\phi_{26}\, h_4(k)\right] P_G(k), && (64)
\end{aligned}$$

where

$$\begin{aligned}
\bar\phi_{13} &= \frac{1}{2} \bar\beta\, v^2(T), && (65)\\
\bar\phi_{22} &= \frac{1}{4} \bar\gamma\, v^2(T) + \frac{1}{4} \bar\beta^2\, v^2(T) + \alpha v(T), && (66)\\
\bar\phi_{24} &= \frac{1}{6} \bar\gamma\, v^3(T) + \frac{2}{3} \bar\beta^2\, v^3(T). && (67)
\end{aligned}$$

Model fit is expressed by the equations

$$\begin{aligned}
\phi_{13} &= \bar\phi_{13}, && (68)\\
\phi_{22} + \phi_{24}\, h_2(k) &= \bar\phi_{22} + \bar\phi_{24}\, h_2(k), && (69)
\end{aligned}$$

which can be satisfied with *strike-independent* parameters α, $\bar\beta$, $\bar\gamma$ such that

$$\begin{aligned}
\phi_{13} &= \bar\phi_{13}, && (70)\\
\phi_{22} &= \bar\phi_{22}, && (71)\\
\phi_{24} &= \bar\phi_{24}. && (72)
\end{aligned}$$

These equations are solved by

$$\begin{aligned}
\bar\beta &= \frac{2\phi_{13}}{v^2(T)}, && (73)\\
\bar\gamma &= \frac{6\phi_{24}}{v^3(T)} - 4\bar\beta^2, && (74)\\
\alpha &= \frac{\phi_{22}}{v(T)} - \frac{3\phi_{24}}{2v^2(T)} + \frac{3}{4} \bar\beta^2\, v(T). && (75)
\end{aligned}$$

A similar approach to strike-independent averaging of the time-dependent quadratic volatility model was put forward in [2]. Our result (75) has a different form but agrees with Lemmas 3 and 4 of [2] up to the order in which the asymptotic expansion was derived.

4 Separable model and LMM

We now proceed to the calculation of the first two corrections in the asymptotic expansion of the option price in non-Markovian models. We focus on a subclass of separable models introduced in [3] as a convenient generic template for more specialized multifactor models such as LMM. After deriving the generic result we apply it to the case of LMM swaption, confirming the result of [14] and presenting it in a more compact form possible due to the structure introduced by the formulation of LMM as a separable model.

We narrow the class of separable models for our current purposes to models for a finite set of processes $X_\varepsilon(t)$, $\{S_n(t)\}$, $\{L_m(t)\}$ obeying the following dynamics,

$$
\begin{aligned}
dX_\varepsilon(t) &= \varepsilon \sum_n S_n(t)\, a_n(t) \cdot dW(t), &(76)\\
dS_n(t) &= \varepsilon^2 \mu_n(t, \{L_k(t)\})\, dt + \varepsilon \sigma_n(t, \{L_k(t)\}) \cdot dW(t), &(77)\\
dL_m(t) &= \varepsilon^2 \nu_m(t, \{L_k(t)\})\, dt + \varepsilon \lambda_m(t, \{L_k(t)\}) \cdot dW(t), &(78)
\end{aligned}
$$

where a_n are deterministic functions of time, μ_n, σ_n, ν_m, λ_m are deterministic functions of time t and basis processes $\{L_k(t)\}$ which taken together form a multi-dimensional Markov process. In case of a multi-dimensional driving Brownian motion, $W(t) = (W_1(t), \ldots, W_F(t))$, each of the functions a_n, σ_n, and λ_m has F components, and dot denotes the scalar product in the space of factors. Notations are chosen with the view to an application to LMM where $\{L_k(t)\}$ will be LIBOR rates. Note that the scaling of the drift term as ε^2 is also consistent with LMM. We are interested in options on the process $X_\varepsilon(t)$, which will represent the swap rate in the case of LMM.

4.1 Asymptotic expansion

We begin by identifying the expansion terms in Eq. (8) (all sums here and below run over a finite set of indexes),

$$
\begin{aligned}
Y_0(T) &= \sum_n \int_0^T dW(t) \cdot a_n(t)\, S_n(0) \equiv \int_0^T dW(t) \cdot \lambda(t), &(79)\\
Y_1(T) &= 2 \sum_n \int_0^T dW(t_1) \cdot a_n(t_1) \int_0^{t_1} dW(t_2) \cdot \hat{\sigma}_n(t_2), &(80)\\
Y_2(T) &= 6 \sum_{n,m} \int_0^T dW(t_1) \cdot a_n(t_1) \int_0^{t_1} dW(t_2) \cdot \frac{\partial \hat{\sigma}_n}{\partial l_m}(t_2) \int_0^{t_2} dW(t_3) \cdot \hat{\lambda}_m(t_3) &(81)\\
&+ 6 \sum_n \int_0^T dW(t_1) \cdot a_n(t_1) \int_0^{t_1} dt_2 \hat{\mu}_n(t_2), &(82)
\end{aligned}
$$

where $l_m = L_m(0)$, $\lambda(t) = \sum_n a_n(t) S_n(0)$ is the volatility of $X(t)$ "frozen" at $S_n(0)$, and hat is used to distinguish parameters evaluated at the initial values of the processes,

$$
\begin{aligned}
\hat{\sigma}_n(t) &= \sigma_n(t, \{l_k\}), &(83)\\
\hat{\mu}_n(t) &= \mu_n(t, \{l_k\}), &(84)\\
\hat{\lambda}_n(t) &= \lambda_n(t, \{l_k\}), &(85)\\
\frac{\partial \hat{\sigma}_n}{\partial l_m}(t) &= \left.\frac{\partial \sigma_n(t, \{L_k\})}{\partial L_m}\right|_{\{L_k\}=\{l_k\}}. &(86)
\end{aligned}
$$

As required, $Y_0(T)$ is a Gaussian random variable with the moment generating function $\Phi_0(\xi)$ given by Eq. (16) with

$$
v(T) = \int_0^T dt |\lambda(t)|^2 \tag{87}
$$

(absolute value sign is essential for the multifactor case). Using the technique described in Appendix A, we calculate the expectations

$$
\begin{aligned}
E\left[e^{\xi Y_0(T)} Y_1(T)\right] &= 2\xi^2\, C(T) \Phi_0(\xi), &(88)\\
E\left[e^{\xi Y_0(T)} Y_2(T)\right] &= 6\left(\xi\, M(T) + \xi^3\, D(T)\right) \Phi_0(\xi), &(89)\\
E\left[e^{\xi Y_0(T)} Y_1^2(T)\right] &= 4\left(R(T) + \xi^2 \int_0^T dt |Q(t)|^2 + \xi^4\, C^2(T)\right) \Phi_0(\xi), &(90)
\end{aligned}
$$

with the following notations for the coefficients,

$$
\begin{aligned}
C(T) &= \sum_n \int_0^T dt_1 \lambda(t_1) \cdot a_n(t_1) \int_0^{t_1} dt_2 \lambda(t_2) \cdot \hat{\sigma}_n(t_2), &(91)\\
D(T) &= \sum_{n,m} \int_0^T dt_1 \lambda(t_1) \cdot a_n(t_1) \int_0^{t_1} dt_2 \lambda(t_2) \cdot \frac{\partial \hat{\sigma}_n}{\partial l_m}(t_2) \int_0^{t_2} dt_3 \lambda(t_3) \cdot \hat{\lambda}_m(t_3), &(92)\\
M(T) &= \sum_n \int_0^T dt_1 \lambda(t_1) \cdot a_n(t_1) \int_0^{t_1} dt_2 \hat{\mu}_n(t_2), &(93)\\
Q(t) &= \sum_n a_n(t) \int_0^t dt_1 \lambda(t_1) \cdot \hat{\sigma}_n(t_1) + \sum_n \hat{\sigma}_n(t) \int_t^T dt_2 \lambda(t_2) \cdot a_n(t_2), &(94)\\
R(T) &= \sum_{n,k} \int_0^T dt_1 a_k(t_1) \cdot a_n(t_1) \int_0^{t_1} dt_2 \hat{\sigma}_k(t_2) \cdot \hat{\sigma}_n(t_2). &(95)
\end{aligned}
$$

This is sufficient to determine the coefficients ϕ_{ij} in Eqs. (18-19),

$$
\begin{aligned}
\phi_{13} &= C(T), &(96)\\
\phi_{22} &= M(T) + \frac{1}{2} R(T), &(97)\\
\phi_{24} &= D(T) + \frac{1}{2} \int_0^T dt |Q(t)|^2, &(98)\\
\phi_{26} &= \frac{1}{2} C^2(T), &(99)
\end{aligned}
$$

which completes the derivation of the corrections to the option prices in the separable model, given by Eqs. (22–23).

4.2 Mapping to the quadratic volatility model

We are looking for an approximation of the separable model with time-dependent coefficients by a quadratic volatility model given by Eq. (49),

$$dX(t) = \left(1 + \bar{\beta}\, \Delta X(t) + \frac{1}{2}\bar{\gamma}\, \Delta X^2(t)\right)\, |\lambda(t)|\, dW(t),$$

based on the matching of the asymptotic expansions for option prices. The Gaussian term is matched by virtue of the choice of $|\lambda(t)|$ for the volatility function (note that the Wiener process $dW(t)$ in the quadratic volatility model has only one component). Coefficients $\bar{\beta}$ and $\bar{\gamma}$ are determined from the match of the first and second corrections to the option, Eqs. (22–23), which is equivalent to the pair of conditions,

$$\begin{aligned} \phi_{13} &= \bar{\phi}_{13}, & (100)\\ \phi_{22} + \phi_{24}\, h_2(k) &= \bar{\phi}_{22} + \bar{\phi}_{24}\, h_2(k), & (101) \end{aligned}$$

where ϕ_{ij} are given by Eqs. (96-98), and $\bar{\phi}_{ij}$ are computed in the quadratic model using Eqs. (43–45) and Eqs. (50–54) with $v(T) = \int_0^T dt |\lambda(t)|^2$,

$$\begin{aligned} \bar{\phi}_{13} &= \frac{1}{2}\bar{\beta}\, v^2(T), & (102)\\ \bar{\phi}_{22} &= \bar{g}_1(T) + \bar{a}_1(T) = \frac{1}{4}\bar{\gamma}\, v^2(T) + \frac{1}{4}\bar{\beta}^2\, v^2(T), & (103)\\ \bar{\phi}_{24} &= \bar{g}_2(T) + \bar{a}_2(T) = \frac{1}{6}\bar{\gamma}\, v^3(T) + \frac{2}{3}\bar{\beta}^2\, v^3(T). & (104) \end{aligned}$$

We omitted the terms with ϕ_{26}, $\bar{\phi}_{26}$ from the condition (101) because these satisfy $\phi_{26} = \frac{1}{2}\phi_{13}^2$, $\bar{\phi}_{26} = \frac{1}{2}\bar{\phi}_{13}^2$, and are matched automatically if the condition (100) is satisfied.

Eq. (100) reproduces the skew formula of Markovian projection to displaced diffusion derived in [3] in a very different way,

$$\bar{\beta} = \frac{2\phi_{13}}{v^2(T)} = \frac{2\sum_n \int_0^T dt_1 \lambda(t_1)\cdot a_n(t_1) \int_0^{t_1} dt_2 \lambda(t_2)\cdot \hat{\sigma}_n(t_2)}{v^2(T)}. \tag{105}$$

Note that the resulting skew parameter does not depend on option strike k.

The effective curvature parameter $\bar{\gamma}$ follows from Eq. (101),

$$\bar{\gamma} = 2\frac{(\phi_{24} - \bar{a}_2(T))\, h_2(k) + \phi_{22} - \bar{a}_1(T)}{v^2(T)\left(\frac{1}{3}\, v(T)\, h_2(k) + \frac{1}{2}\right)}, \tag{106}$$

or, substituting the expressions for $h_2(k)$, $\bar{a}_1(T)$, $\bar{a}_2(T)$,

$$\bar{\gamma} = 6v^{-3}(T)\frac{\left(\phi_{24} - \frac{2}{3}\bar{\beta}^2\, v^3(T)\right)\, \frac{k^2}{v(T)} + v(T)\, \phi_{22} - \phi_{24} + \frac{5}{12}\bar{\beta}^2\, v^3(T)}{\frac{k^2}{v(T)} + \frac{1}{2}}, \tag{107}$$

where a dependence on the option strike k is present.

Alternatively, we can use a strike-independent projection on a model

$$dX(t) = \left(1 + \bar{\beta}\, \Delta X(t) + \frac{1}{2}\bar{\gamma}\, \Delta X^2(t)\right)\, (1+\alpha)|\lambda(t)|\, dW(t),$$

as described in Sect. 3.3 The coefficients α, $\bar{\beta}$, and $\bar{\gamma}$ are found by putting the values of ϕ_{13}, ϕ_{22}, ϕ_{24} into the formulas (73–75).

4.3 Application to the LMM model

In this section we cast the LMM swap rate process as the component $X(t) = X_\varepsilon(t)|_{\varepsilon=1}$ of a separable model and apply the generic results to swaption pricing. We fix maturity dates T_1, T_2, $\ldots$, $T_{N_{\max}}$ and denote by $P(t, T_n)$ the price of the zero-coupon bonds with maturity T_n observed at time t. Forward LIBOR rates are defined by

$$L_n(t) = \frac{1}{\delta_n}\left(\frac{P(t,T_n)}{P(t,T_{n+1})} - 1\right), \tag{108}$$

where δ_n is the daycount fraction for interest accrued from T_n to T_{n+1}. We postulate a displaced diffusion for the evolution of each LIBOR rate $L_n(t)$ in its natural measure (called T_{n+1}-forward measure) with the numeraire $P(t, T_{n+1})$ and the Brownian motion $W^{(n+1)}(t)$,

$$dL_n(t) = (1 + \beta_n(t)\,\Delta L_n(t))\,\gamma_n(t)\cdot dW^{(n+1)}(t). \tag{109}$$

This form is equivalent to $dL_n(t) = (L_n(t) + \tilde{\beta}_n(t)\,\Delta L_n(t))\,\tilde{\gamma}_n(t)\cdot dW^{(n+1)}(t)$ used in [17] for positive initial rates and has the advantage of remaining well defined for non-positive ones. Here $\beta_n(t)$ is a time-dependent shift parameter, $\gamma_n(t)$ is a vector of time-dependent volatility functions, and $\Delta L_n(t) = L_n(t) - L_n(0)$. The log-normal case corresponds to $\beta_n(t) \equiv 1/L_n(0)$ while the normal case is $\beta_n(t) \equiv 0$. Note that we use a superscript to distinguish Brownian motions in different measures. Each of these Brownian motions can be multidimensional; we usually omit the corresponding index.

We introduce a process for the forward swap rate spanning the period from T_{n_0} to T_N,

$$X(t) = \frac{P(t,T_{n_0}) - P(t,T_N)}{A(t)}, \tag{110}$$

where $A(t)$ is the annuity process given by a linear combination of zero-coupon bonds,

$$A(t) = \sum_{j=n_0+1}^{N} \delta_j P(t,T_j). \tag{111}$$

Swaption price with strike K is given by an expectation

$$A(0)\,E_A[(X(T_{n_0}) - K)^+], \tag{112}$$

in the measure with the numeraire $A(t)$ and a certain Brownian motion $W^A(t)$, called swap measure [13]. We will have to convert the evolution of each LIBOR from its natural measure to the common swap measure.

The process involved in the change from T_n-forward measure to swap measure (Radon-Nikodym derivative) is the ratio of the numeraires,

$$M_{n,A}(t) = \frac{P(t,T_n)}{A(t)}, \tag{113}$$

which can be expressed as a function of LIBORs,

$$M_{n,A}(t) = \frac{R_{nN}(t)}{\sum_{j=n_0+1}^{N} \delta_j R_{jN}(t)}, \tag{114}$$

where

$$R_{s_1 s_2}(t) = \begin{cases} \prod_{j=s_1}^{s_2-1}(1+\delta_j L_j(t)) & \text{for} \quad s_1 < s_2, \\ 1 & \text{for} \quad s_1 = s_2, \\ \prod_{j=s_2}^{s_1-1}(1+\delta_j L_j(t))^{-1} & \text{for} \quad s_1 > s_2. \end{cases} \tag{115}$$

The process $M_{n,A}(t)$ is a martingale in the swap measure, obeying a driftless SDE,

$$dM_{n,A}(t) = M_{n,A}(t)\sigma_{n,A}\cdot dW^A(t), \tag{116}$$

with the log-normal volatility

$$\sigma_{n,A}(t) = \sum_k \frac{\partial \ln M_{n,A}(t)}{\partial L_k} \left(1 + \beta_k(t)\, \Delta L_k(t)\right) \gamma_k(t). \tag{117}$$

On the other hand, $1/M_{n,A}(t)$ is a martingale in the T_n-forward measure, which leads to the following relationship for the Brownian motions,

$$dW^A(t) = dW^{(n)}(t) + \sigma_{n,A}\, dt, \tag{118}$$

and the evolution of the LIBORs in the swap measure,

$$dL_n(t) = \nu_n(t)\, dt + \lambda_n(t)\cdot dW^A(t), \tag{119}$$

with

$$\nu_n(t) \;=\; -\left(1 + \beta_n(t)\, \Delta L_n(t)\right) \gamma_n(t) \cdot \sigma_{n+1,A}, \tag{120}$$

$$\lambda_n(t) \;=\; \left(1 + \beta_n(t)\, \Delta L_n(t)\right) \gamma_n(t). \tag{121}$$

Consistency of the model, that is cancelation of drift in Eq. (116) after computing the differential of $M_{n,A}(t)$ as a function of LIBORs using Itô's lemma, is ensured by the relationship

$$\frac{\partial M_{k,A}}{\partial L_n} \frac{\partial \ln M_{n+1,A}}{\partial L_m} + \frac{\partial M_{k,A}}{\partial L_m} \frac{\partial \ln M_{m+1,A}}{\partial L_n} = \frac{\partial^2 M_{k,A}}{\partial L_n\, \partial L_m}, \tag{122}$$

which can be verified directly, using the following expression for the logarithmic derivative of $M_{n,A}$,

$$\frac{\partial \ln M_{n,A}}{\partial L_k} = \frac{\delta_k}{1 + \delta_k L_k} \left(\theta_{nN}(k) - \sum_{j=n_0+1}^{N} \delta_j M_{j,A} \theta_{jN}(k) \right), \tag{123}$$

where

$$\theta_{s_1 s_2}(k) = \begin{cases} 1 & \text{for} \quad s_1 \le k < s_2, \\ -1 & \text{for} \quad s_2 \le k < s_1, \\ 0 & \text{otherwise.} \end{cases} \tag{124}$$

The same cancelation holds for the swap rate process $X(t) = M_{n_0,A}(t) - M_{N,A}(t)$, leading to

$$dX(t) = \sum_n \frac{\partial X}{\partial L_n} \left(1 + \beta_n(t)\, \Delta L_n(t)\right) \gamma_n(t)\cdot dW^A(t). \tag{125}$$

From this point we restrict ourselves to the case of time-independent shift functions, $\beta_n(t) \equiv \beta_n$, in which case the results can be presented in a compact matrix form convenient for numerical evaluation. The general case is considered in Appendix B.1. With $X(t)$ and $\{L_n(t)\}$ already present, we only need to identify the set of quantities $\{S_n(t)\}$ to complete the immersion of the swap rate $X(t)$ into a separable model. This is straightforward,

$$S_n(t) = \frac{\partial X}{\partial L_n} \left(1 + \beta_n \Delta L_n(t)\right), \qquad a_n(t) = \gamma_n(t). \tag{126}$$

Using Itô's lemma, we obtain $dS_n(t) = \mu_n(t)\,dt + \sigma_n(t)\cdot dW^A(t)$ with

$$\mu_n(t) = \frac{\partial X}{\partial L_n}\beta_n\nu_n(t) + \sum_k \frac{\partial^2 X}{\partial L_k \partial L_n}\left((1+\beta_n\,\Delta L_n(t))\nu_k(t) + \beta_n\lambda_n(t)\cdot\lambda_k(t)\right)$$
$$+ \frac{1}{2}\sum_{k_1,k_2} \frac{\partial^3 X}{\partial L_{k_1}\partial L_{k_2}\partial L_n}(1+\beta_n\Delta L_n(t))\lambda_{k_1}(t)\cdot\lambda_{k_2}(t), \tag{127}$$

$$\sigma_n(t) = \frac{\partial X}{\partial L_n}\beta_n\lambda_n(t) + \sum_k \frac{\partial^2 X}{\partial L_k \partial L_n}(1+\beta_n\,\Delta L_n(t))\lambda_k(t). \tag{128}$$

We also need the following set of derivatives over LIBORs,

$$\frac{\partial\sigma_n(t)}{\partial L_m} = \frac{\partial X}{\partial L_n}\beta_n^2\gamma_n(t)\delta_{nm} + \frac{\partial^2 X}{\partial L_m\partial L_n}\left(\beta_n\lambda_n(t) + \beta_m\gamma_m(t)(1+\beta_n\Delta L_n(t))\right)$$
$$+ \sum_k \frac{\partial^2 X}{\partial L_k\partial L_n}\beta_n\lambda_k(t)\delta_{nm} + \sum_k \frac{\partial^3 X}{\partial L_k\,\partial L_m\,\partial L_n}(1+\beta_n\Delta L_n(t))\lambda_k(t). \tag{129}$$

Finally we "freeze" stochastic quantities, that is evaluate them for the initial LIBOR values, $L_m(0) = l_m$, keeping the time dependence for deterministic volatility functions of the model. In doing so, we encounter the partial derivatives of the initial swap rate X_0 over initial LIBORs up to the third order. We get

$$S_n(0) = \frac{\partial X_0}{\partial l_n}, \tag{130}$$

$$\lambda(t) = \sum_n \frac{\partial X_0}{\partial l_n}\gamma_n(t), \tag{131}$$

$$\hat{\sigma}_n(t) = \frac{\partial X_0}{\partial l_n}\beta_n\gamma_n(t) + \sum_k \frac{\partial^2 X_0}{\partial l_k\,\partial l_n}\gamma_k(t), \tag{132}$$

$$\hat{\lambda}_k(t) = \gamma_k(t) \tag{133}$$

$$\hat{\nu}_n(t) = -\sum_k \frac{\partial \ln M_{n+1,A}(0)}{\partial l_k}\,\gamma_k(t)\cdot\gamma_n(t), \tag{134}$$

$$\hat{\mu}_n(t) = \frac{\partial X_0}{\partial l_n}\beta_n\hat{\nu}_n + \sum_k \frac{\partial^2 X_0}{\partial l_k\partial l_n}(\hat{\nu}_k(t) + \beta_n\gamma_k(t)\cdot\gamma_n(t))$$
$$+ \frac{1}{2}\sum_{k_1,k_2}\frac{\partial^3 X_0}{\partial l_{k_1}\partial l_{k_2}\partial l_n}\gamma_{k_1}(t)\cdot\gamma_{k_2}(t), \tag{135}$$

$$\frac{\partial\hat{\sigma}_n}{\partial l_m}(t) = \delta_{nm}\left(\frac{\partial X_0}{\partial l_n}\beta_n^2\gamma_n(t) + \sum_k \frac{\partial^2 X_0}{\partial l_k\partial l_n}\beta_n\gamma_k(t)\right)$$
$$+ \frac{\partial^2 X_0}{\partial l_n\partial l_m}(\beta_m\gamma_m(t) + \beta_n\gamma_n(t)) + \sum_k \frac{\partial^3 X_0}{\partial l_k\partial l_n\partial l_m}\gamma_k(t). \tag{136}$$

The resulting set of functions can be plugged into Eqs. (91–95), defining the coefficients in the corrections to option prices from the asymptotic expansion.

The result can be presented in a considerably more compact form simplifying a numerical implementation. We introduce the following tensor notations for the derivatives over initial values of LIBORs,

$$\mathcal{D}^{(1)}_k = \frac{\partial X_0}{\partial l_k}, \qquad \mathcal{D}^{(2)}_{nk} = \frac{\partial^2 X_0}{\partial l_n\,\partial l_k}, \qquad \mathcal{D}^{(3)}_{nmk} = \frac{\partial^3 X_0}{\partial l_n\,\partial l_m\,\partial l_k}, \tag{137}$$

$$\mathcal{M}_{nk} = \mathcal{D}^{(2)}_{nk} + \delta_{nk}\mathcal{D}^{(1)}_{k}\beta_k, \tag{138}$$

$$\mathcal{K}_{mnk} = \mathcal{D}^{(1)}_{m}\left(\frac{\partial^2 \ln M_{m+1,A}(0)}{\partial l_n\,\partial l_k} + \frac{\partial \ln M_{m+1,A}(0)}{\partial l_k}\,\beta_k\,\delta_{nk},\right), \tag{139}$$

along with the notations for the leading contribution to LIBORs and swap rate covariances,

$$V_{nm}(t) = \int_0^t d\tau \gamma_n(\tau)\cdot\gamma_m(\tau), \tag{140}$$

$$I_n(t) = \int_0^t d\tau \gamma_n(\tau)\cdot\lambda(\tau). \tag{141}$$

It is shown in Appendix B.2 that the coefficients (91–95) of the asymptotic expansion take the form

$$\begin{aligned}
C(T) &= \frac{1}{2} I^T(T)\,\mathcal{M}\,I(T), &(142)\\
\int_0^T dt|Q(t)|^2 &= I^T(T)\,\mathcal{M}V(T)\mathcal{M}\,I(T), &(143)\\
R(T) &= \frac{1}{2}\mathrm{Tr}\left(\mathcal{M}V(T)\mathcal{M}V(T)\right), &(144)\\
D(T) &= \frac{1}{6}\sum_{n,m,k}\mathcal{D}^{(3)}_{nmk}\,I_k(T)\,I_n(T)\,I_m(T) + \frac{1}{2}\sum_{m,k}\mathcal{D}^{(2)}_{mk}\,\beta_m I_m^2(T)\,I_k(T)\\
&+ \frac{1}{6}\sum_{k}\mathcal{D}^{(1)}_{k}\,\beta_k^2 I_k^3(T), &(145)\\
M(T) &= \sum_{n,m,k}\mathcal{K}_{m,n,k}\int_0^T dI_k(t)\,V_{nm}(t). &(146)
\end{aligned}$$

Eqs. (142–144) contain matrix products, with $\mathcal{M}_{nk}$ and V_{nk} interpreted as matrices, and I_n as a column vector. Despite the difference in appearance, the result is in agreement with [14] as shown in Appendix B.3.

5 Numerical Results

In this section we present numerical results for option prices in the time-dependent quadratic volatility model and LMM. We use a Monte Carlo simulation with a large number of paths as a baseline and test the quality of the approximations offered by the asymptotic expansion corrections and mappings to the time-independent quadratic volatility model solved analytically following [1].

5.1 Time-averaging for displaced diffusion and quadratic volatility models

We begin by checking numerical effectiveness of the results for parameter averaging, also including the displaced diffusion model as a particular case of the quadratic volatility model with vanishing curvature. For each call option included in the test, we report the implied Black volatility obtained using the following methods (with abbreviations used to distinguish the data in Tables and Figures given in parentheses).

- Baseline Monte Carlo (MC).

- Direct use of asymptotic expansion corrections (47–48) (AE).
- Asymptotic expansion-based mapping to the time-independent quadratic volatility model with strike-dependent curvature described in Sect. 3.2 (AE QV).
- Asymptotic expansion-based mapping to the time-independent quadratic volatility model with adjusted volatility and strike-independent curvature described in Sect. 3.3 (AE QVA).

In the case of a displaced-diffusion model, we also included an additional averaging method.

- Mapping to a displaced-diffusion model with the effective shift parameter given by Piterbarg's formula (55) without introducing a curvature (AE DD).

5.1.1 Displaced diffusion model with time-dependent parameters

The starting model is the displaced diffusion with time-dependent shift parameter $\beta(t)$,

$$dX = (1 + \beta(t)\,\Delta X(t))\,\lambda(t)\,dW(t), \quad X(0) = 1, \tag{147}$$

which is a special case of the quadratic volatility model without curvature. We chose a time-independent volatility function $\lambda(t) = 0.15$ and a strongly time-dependent skew,

$$\beta(t) = \begin{cases} 0 & \text{for} \quad 0 \le t \le 5, \\ 1 & \text{for} \quad 5 < t \le 10. \end{cases} \tag{148}$$

Implied Black volatilities for the option $E[(X(T) - K)^+]$ with $T = 10$ are given in Table 1 for a wide range of strikes. Deviations form the Monte Carlo baseline are reported in Table 2. The results are also presented graphically in Figs. 1,2.

We see that the time-dependent shift generates a slight smile in implied volatilities despite the absence of a quadratic term in the original model. The method of mapping on a strike-dependent quadratic volatility model (AE QV) slightly outperforms the method of mapping on a strike-independent model (AE QVA). The latter method is, of course, preferable from the point of view of performance since the curvature parameter does not have to be recalculated for every strike. We also see that a direct use of the asymptotic expansion corrections (AE) leads to large errors in the out-of-money region, while the projection on a displaced diffusion (AE DD) suffers from a large bias.

5.1.2 Quadratic volatility model with time-dependent parameters

The second numerical experiment uses a full quadratic volatility model,

$$dX(t) = \left(1 + \beta(t)\,\Delta X(t) + \frac{1}{2}\gamma(t)\,\Delta X^2(t)\right)\lambda(t)\,dW(t), \quad X(0) = 1 \tag{149}$$

with time-dependent parameters. We chose the same time-independent volatility function $\lambda(t) = 0.15$ and time-dependent shift $\beta(t)$ given by Eq. (148). The curvature $\gamma(t)$ is also strongly time-dependent,

$$\gamma(t) = \begin{cases} 0 & \text{for} \quad 0 \le t \le 5, \\ 3 & \text{for} \quad 5 < t \le 10. \end{cases} \tag{150}$$

Implied Black volatilities for the option $E[(X(T) - K)^+]$ with $T = 10$ are given in Table 3 for a wide range of strikes. Deviations form the Monte Carlo baseline are reported in Table 4. The results are presented graphically in Figs. 3,4.

We see that the method of mapping on a strike-independent quadratic volatility model (AE QVA) now slightly outperforms the method of mapping on a strike-independent model (AE QV) in the out-of-money region. We also confirm that a direct use of asymptotic expansion correction (AE) leads to unacceptably large errors.

Strike (%)	Implied volatility (%)				
	AE	AE DD	AE QV	AE QVA	MC
53.13	16.69	16.53	16.76	16.74	16.88
62.23	16.26	16.11	16.29	16.27	16.35
72.89	15.85	15.73	15.86	15.85	15.89
85.38	15.48	15.38	15.48	15.48	15.49
100.00	15.15	15.06	15.15	15.15	15.15
117.13	14.87	14.77	14.87	14.87	14.86
137.19	14.63	14.51	14.62	14.63	14.60
160.70	14.47	14.27	14.42	14.43	14.38
188.22	14.33	14.05	14.26	14.27	14.17

Table 1: Numerical results for different approximations of the call option price with maturity $T = 10$ in a displaced-diffusion model (147) with $\lambda(t) = 0.15$ and time-dependent shift function (148).

Strike (%)	Error in implied volatility (bps)			
	AE	AE DD	AE QV	AE QVA
53.13	-19.27	-35.31	-12.08	-14.1
62.23	-8.54	-23.53	-6.25	-7.5
72.89	-3.41	-15.71	-2.96	-3.55
85.38	-1.34	-11.23	-1.33	-1.47
100.00	-0.14	-9.06	-0.14	-0.18
117.13	1.05	-8.53	1.02	1.11
137.19	3.2	-9.24	2.22	2.84
160.70	9.88	-10.69	4.33	5.58
188.22	15.83	-12.33	8.17	10

Table 2: Deviations from the Monte Carlo baseline for different approximations to the option price reported in Table 1 for a displaced-diffusion model with time-dependent shift.

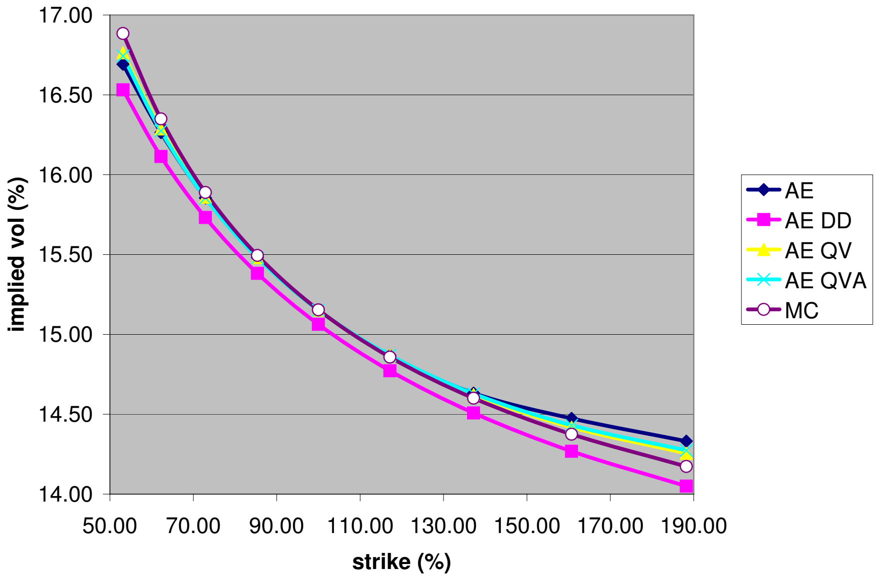

Figure 1: Numerical results for different approximations of the call option price with maturity $T = 10$ in a displaced-diffusion model (147) with $\lambda(t) = 0.15$ and time-dependent shift function (148).

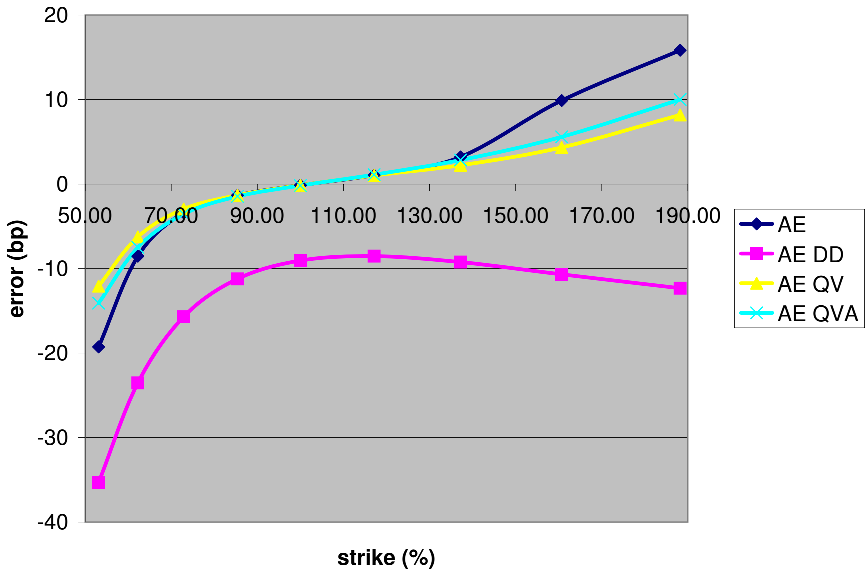

Figure 2: Deviations from the Monte Carlo baseline for different approximations to the option price reported in Table 1 for a displaced-diffusion model with time-dependent shift.

Strike (%)	Implied volatility (%)			
	AE	AE QV	AE QVA	MC
53.13	19.69	19.55	19.55	19.36
62.23	18.17	18.13	18.12	18.01
72.89	16.98	16.96	16.95	16.88
85.38	16.11	16.10	16.09	16.06
100.00	15.59	15.60	15.58	15.60
117.13	15.45	15.49	15.46	15.52
137.19	15.70	15.81	15.74	15.82
160.70	16.12	16.50	16.37	16.45
188.22	16.08	17.49	17.29	17.36

Table 3: Numerical results for different approximations of the call option price with maturity $T = 10$ in a quadratic volatility model (149) with $\lambda(t) = 0.15$, time-dependent shift function (148), and time-dependent curvature (150).

Strike (%)	Error in implied volatility (bps)		
	AE	AE QV	AE QVA
53.13	33.69	19.46	19.13
62.23	16.16	11.94	11.8
72.89	9.09	7.5	6.75
85.38	4.95	4.19	2.43
100.00	-0.88	0.29	-1.77
117.13	-7.14	-2.56	-5.5
137.19	-11.75	-0.88	-8.04
160.70	-33.16	4.73	-8.63
188.22	-127.81	12.83	-6.98

Table 4: Deviations from the Monte Carlo baseline for different approximations to the option price reported in Table 3 for a time-dependent quadratic volatility model.

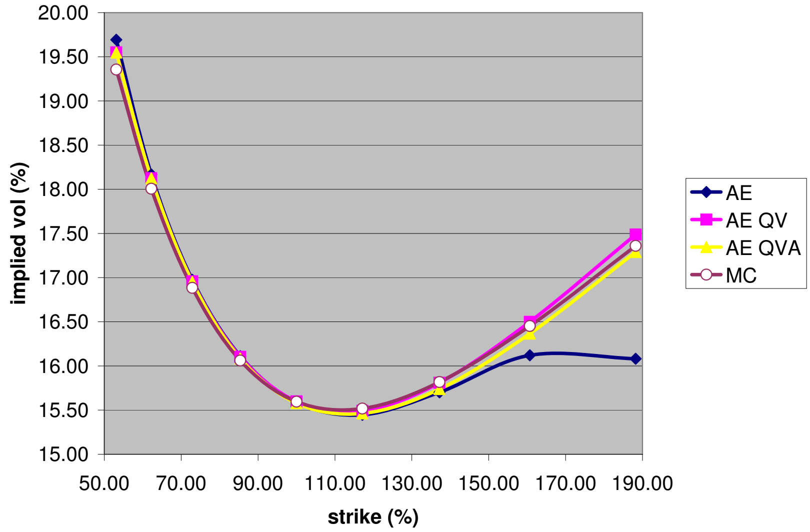

Figure 3: Numerical results for different approximations of the call option price with maturity $T = 10$ in a quadratic volatility model (149) with $\lambda(t) = 0.15$, time-dependent shift function (148), and time-dependent curvature (150).

Figure 4: Deviations from the Monte Carlo baseline for different approximations to the option price reported in Table 3 for a time-dependent quadratic volatility model.

5.2 Swaption in LMM

We proceed to numerical tests for the approximations obtained in Sect. 4 for swaption prices in LMM. We set up a 3-factor model in accordance with Sect. 4.3 with an annual tenor and maturities $T_n = n$ extended to $T_{N_{\max}} = 30$. The model has 29 stochastic LIBOR rates $L_n(t)$ spanning periods $[n, n+1]$. Initial LIBORs $l_n = L_n(0)$ and components of rescaled[2] volatility vectors $\tilde{\gamma}_n = \gamma_n / L_n(0)$ are presented in Table 5, corresponding to the current highly volatile situation. In a 3-factor model, the driving Brownian motion $W(t) = (W_1(t), W_2(t), W_3(t))$ is 3-dimensional, so each volatility vector has 3 components, $\left(\gamma_n^{(1)}, \gamma_n^{(2)}, \gamma_n^{(3)}\right)$, chosen to be time-independent. The correlation structure

$$C_{ij} = \frac{\gamma_i \cdot \gamma_j}{|\gamma_i| \, |\gamma_j|}$$

is a result of a reduction of a full-rank matrix $C_{ij}^F = e^{-0.1\,|i-j|}$ to 3 principal components.

We considered two cases, $\beta_n(t) \equiv 1/l_n$, corresponding to the log-normal LIBOR dynamics, and $\beta_n(t) \equiv 1/(2l_n)$, which is in-between normal and log-normal dynamics. In each case we computed prices of 10y10, 10y20, and 20y10 swaptions with different strikes. (Notation NyM means that the swaption is exercised at $t = N$ and spans the period from N to $N + M$, so that N is the exercise time and M is the length.) For each call option included in the test, we report the implied Black volatility obtained using the following methods (with abbreviations used to distinguish the data in Tables and Figures given in parentheses).

- Baseline Monte Carlo (MC).
- Direct use of asymptotic expansion corrections (47–48) (AE).
- Asymptotic expansion-based mapping to the time-independent quadratic volatility model with strike-dependent curvature described in Sect. 4.2 and Sect. 3.2 (AE QV).
- Asymptotic expansion-based mapping to the time-independent quadratic volatility model with adjusted volatility and strike-independent curvature described in Sect. 4.2 and Sect. 3.3 (AE QVA).
- Mapping to a displaced-diffusion model with effective shift parameter given Eq. (105) without introducing a curvature (AE DD).
- Mapping to an effective log-normal model proposed in [14], see also Appendix B.3 (Kawai).

Numerical results in the log-normal setup are given in Tables 6,7. Numerical results in the shifted setup are given in Tables 8,9. We also plotted the implied volatilities and the corresponding approximation errors for the 10y10 swaption in Figs. 5–8. We observe that the method of mapping on a strike-independent quadratic volatility model (AE QVA) consistently outperforms other methods. The mapping on a strike-dependent quadratic volatility model (AE QV) is not far behind in terms of accuracy but it is obviously less attractive from the point of view of performance since the curvature parameter has to be recalculated for every strike. Kawai's log-normal approximation also works very well in the log-normal case but breaks down in the shifted case with a mixture of normal and log-normal LIBOR dynamics. As with the results for parameter averaging, a direct use of the asymptotic expansion correction (AE) leads to unacceptably large errors in the wings. Finally, the projection on a displaced diffusion (AE DD) suffers from a large systematic bias.

[2]Absolute values of rescaled volatilities $\tilde{\gamma}_n$ are more intuitive numerically because they are equal to implied Black volatilities of LIBORs in the log-normal case $\beta_n(t) \equiv 1/l_n$.

LIBOR span	Initial value $L_n(0)$ (%)	Volatility (%)			
		$\tilde{\gamma}_n^{(1)}$	$\tilde{\gamma}_n^{(2)}$	$\tilde{\gamma}_n^{(3)}$	$\lvert\tilde{\gamma}_n\rvert$
$[1, 2]$	1.65	49.71	-5.31	0.57	50.00
$[2, 3]$	1.79	47.87	-9.72	-2.79	48.93
$[3, 4]$	1.93	45.87	-12.92	-4.41	47.86
$[4, 5]$	2.07	44.23	-14.68	-4.16	46.79
$[5, 6]$	2.21	43.12	-15.03	-2.23	45.71
$[6, 7]$	2.35	42.29	-14.26	0.86	44.64
$[7, 8]$	2.50	41.40	-12.82	4.47	43.57
$[8, 9]$	2.64	40.24	-11.05	8.03	42.50
$[9, 10]$	2.78	38.81	-9.19	11.21	41.43
$[10, 11]$	2.92	37.20	-7.31	13.82	40.36
$[11, 12]$	3.06	35.55	-5.42	15.81	39.29
$[12, 13]$	3.21	33.96	-3.43	17.18	38.21
$[13, 14]$	3.35	32.50	-1.27	17.94	37.14
$[14, 15]$	3.49	31.18	1.10	18.10	36.07
$[15, 16]$	3.63	30.00	3.65	17.66	35.00
$[16, 17]$	3.78	28.90	6.29	16.63	33.93
$[17, 18]$	3.92	27.83	8.86	15.05	32.86
$[18, 19]$	4.06	26.76	11.20	12.99	31.79
$[19, 20]$	4.21	25.64	13.21	10.55	30.71
$[20, 21]$	4.35	24.45	14.80	7.85	29.64
$[21, 22]$	4.49	23.17	15.95	4.99	28.57
$[22, 23]$	4.64	21.78	16.65	2.11	27.50
$[23, 24]$	4.78	20.29	16.92	-0.61	26.43
$[24, 25]$	4.93	18.74	16.82	-2.97	25.36
$[25, 26]$	5.07	17.24	16.44	-4.72	24.29
$[26, 27]$	5.22	15.91	15.94	-5.60	23.21
$[27, 28]$	5.36	14.88	15.44	-5.50	22.14
$[28, 29]$	5.51	14.14	14.96	-4.52	21.07
$[29, 30]$	5.65	13.55	14.42	-2.94	20.00

Table 5: LMM setup for numerical experiments, including initial values of LIBOR rates, $L_n(0) = l_n$, and components and absolute values of LIBOR volatilities, γ_n, chosen to be time-independent.

Exercise (years)	Length (years)	Strike (%)	Swaption implied volatility (%)					
			MC	AE	AE DD	Kawai	AE QV	AE QVA
10	10	1.87	33.77	32.45	34.02	33.67	33.49	33.66
10	10	2.19	33.76	32.84	34.09	33.70	33.59	33.70
10	10	2.57	33.75	33.14	34.15	33.73	33.68	33.73
10	10	3.01	33.74	33.40	34.21	33.77	33.74	33.77
10	10	**3.52**	**33.73**	**33.65**	**34.26**	**33.81**	**33.79**	**33.82**
10	10	4.12	33.71	33.90	34.31	33.86	33.85	33.87
10	10	4.83	33.70	34.14	34.35	33.91	33.93	33.92
10	10	5.66	33.68	34.44	34.39	33.97	34.05	33.99
10	10	6.63	33.65	35.07	34.42	34.04	34.19	34.06
10	20	2.17	24.92	22.95	24.77	24.39	23.90	24.45
10	20	2.54	24.97	23.71	25.20	24.56	24.33	24.60
10	20	2.98	25.05	24.25	25.58	24.76	24.70	24.78
10	20	3.49	25.13	24.71	25.92	24.99	25.01	24.99
10	20	**4.08**	**25.23**	**25.15**	**26.22**	**25.25**	**25.28**	**25.24**
10	20	4.78	25.35	25.63	26.49	25.55	25.59	25.54
10	20	5.60	25.49	26.18	26.73	25.91	26.02	25.87
10	20	6.56	25.65	26.94	26.95	26.31	26.59	26.25
10	20	7.69	25.84	28.17	27.14	26.78	27.24	26.68
20	10	2.02	23.53	21.30	23.54	23.32	22.87	23.31
20	10	2.53	23.54	22.21	23.72	23.37	23.13	23.37
20	10	3.16	23.55	22.74	23.88	23.44	23.33	23.43
20	10	3.95	23.55	23.12	24.01	23.51	23.48	23.51
20	10	**4.94**	**23.55**	**23.46**	**24.12**	**23.60**	**23.59**	**23.61**
20	10	6.18	23.54	23.80	24.22	23.71	23.71	23.72
20	10	7.73	23.53	24.18	24.30	23.83	23.91	23.85
20	10	9.67	23.52	25.00	24.38	23.96	24.19	24.00
20	10	12.09	23.50	26.90	24.44	24.10	24.50	24.17

Table 6: Numerical results for implied Black volatilities of different swaptions in a LMM defined by Eq. (109) with the inputs given in Table 5 and log-normal setup, $\beta_n(t) \equiv 1/l_n$. Results for the ATM strike are boldfaced.

Exercise (years)	Length (years)	Strike (%)	Error in swaptions implied volatility (bps)				
			AE	AE DD	Kawai	AE QV	AE QVA
10	10	1.87	-132	25	-10	-28	-10
10	10	2.19	-92	33	-6	-16	-6
10	10	2.57	-61	41	-2	-7	-2
10	10	3.01	-34	47	3	1	3
10	10	**3.52**	**-7**	**53**	**8**	**6**	**9**
10	10	4.12	18	59	14	13	15
10	10	4.83	44	65	22	23	23
10	10	5.66	76	71	30	37	31
10	10	6.63	142	77	38	54	41
10	20	2.17	-196	-14	-52	-102	-46
10	20	2.54	-127	23	-41	-65	-38
10	20	2.98	-80	53	-29	-35	-27
10	20	3.49	-43	79	-15	-12	-14
10	20	**4.08**	**-8**	**99**	**2**	**5**	**1**
10	20	4.78	28	114	21	24	19
10	20	5.60	70	124	42	53	38
10	20	6.56	129	129	66	94	60
10	20	7.69	233	130	93	140	84
20	10	2.02	-223	1	-21	-66	-22
20	10	2.53	-133	18	-17	-41	-18
20	10	3.16	-81	33	-11	-22	-12
20	10	3.95	-43	46	-4	-7	-4
20	10	**4.94**	**-9**	**57**	**5**	**4**	**6**
20	10	6.18	25	68	17	17	18
20	10	7.73	65	77	30	38	32
20	10	9.67	148	86	45	67	48
20	10	12.09	340	94	60	100	68

Table 7: Deviations from the Monte Carlo baseline for different approximations to the swaption prices reported in Table 6 for a LMM in the log-normal setup. Results for the ATM strike are boldfaced.

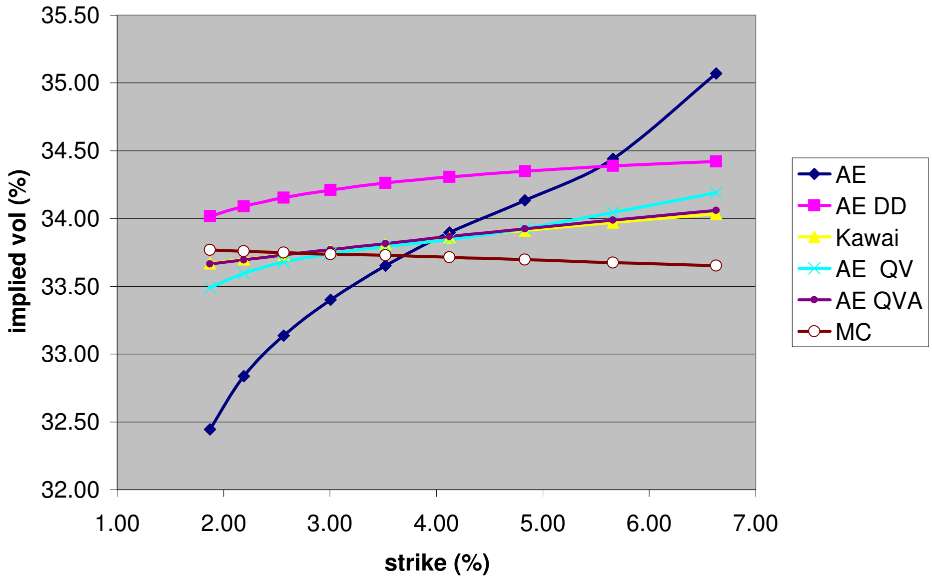

Figure 5: Numerical results for implied Black volatilities of 10y10 swaptions in a LMM defined by Eq. (109) with the inputs given in Table 5 and log-normal setup, $\beta_n(t) \equiv 1/l_n$.

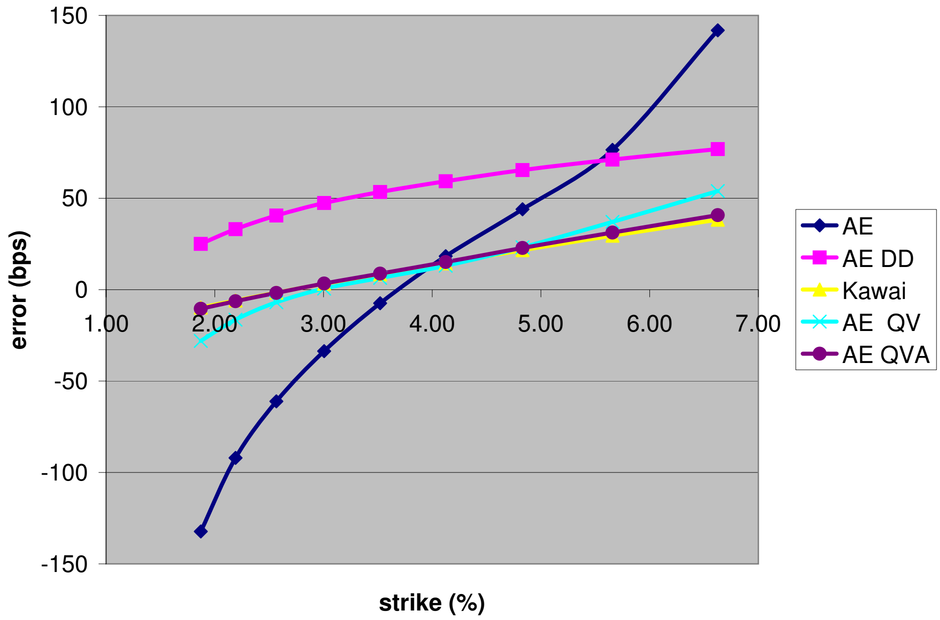

Figure 6: Deviations from the Monte Carlo baseline for different approximations to the 10y10 swaption prices reported in Table 6 for a LMM in the log-normal setup.

Exercise (years)	Length (years)	Strike (%)	Swaption implied volatility (%)					
			MC	AE	AE DD	Kawai	AE QV	AE QVA
10	10	1.87	42.86	42.58	43.19	40.35	42.73	42.78
10	10	2.19	40.71	40.52	41.06	39.19	40.63	40.66
10	10	2.57	38.75	38.63	39.11	37.92	38.71	38.72
10	10	3.01	36.96	36.90	37.33	36.59	36.95	36.96
10	10	**3.52**	**35.33**	**35.33**	**35.70**	**35.25**	**35.35**	**35.35**
10	10	4.12	33.84	33.90	34.22	34.03	33.89	33.89
10	10	4.83	32.49	32.61	32.87	33.12	32.57	32.56
10	10	5.66	31.27	31.46	31.64	32.85	31.39	31.36
10	10	6.63	30.16	30.48	30.53	33.61	30.34	30.28
10	20	2.17	31.11	30.57	31.58	29.86	30.74	30.88
10	20	2.54	29.70	29.33	30.28	29.01	29.46	29.53
10	20	2.98	28.42	28.18	29.08	28.11	28.29	28.30
10	20	3.49	27.25	27.13	27.99	27.16	27.21	27.19
10	20	**4.08**	**26.21**	**26.19**	**26.98**	**26.23**	**26.23**	**26.20**
10	20	4.78	25.28	25.37	26.06	25.41	25.36	25.32
10	20	5.60	24.47	24.67	25.22	24.86	24.63	24.56
10	20	6.56	23.76	24.12	24.45	24.81	24.05	23.91
10	20	7.69	23.16	23.77	23.75	25.56	23.60	23.37
20	10	2.02	32.64	32.14	32.96	29.38	32.35	32.47
20	10	2.53	30.27	29.93	30.63	28.40	30.09	30.15
20	10	3.16	28.19	27.97	28.57	27.26	28.08	28.10
20	10	3.95	26.35	26.24	26.75	26.00	26.30	26.31
20	10	**4.94**	**24.75**	**24.72**	**25.16**	**24.69**	**24.74**	**24.74**
20	10	6.18	23.34	23.40	23.76	23.57	23.38	23.37
20	10	7.73	22.12	22.28	22.53	23.06	22.23	22.20
20	10	9.67	21.06	21.42	21.47	23.90	21.28	21.19
20	10	12.09	20.16	20.91	20.54	27.01	20.48	20.34

Table 8: Numerical results for implied Black volatilities of different swaptions in a LMM defined by Eq. (109) with the inputs given in Table 5 and shifted setup, $\beta_n(t) \equiv 1/(2l_n)$. Results for the ATM strike are boldfaced.

Exercise (years)	Length (years)	Strike (%)	Error in swaption implied volatility (bps)				
			AE	AE DD	Kawai	AE QV	AE QVA
10	10	1.87	-28	34	-250	-13	-8
10	10	2.19	-20	35	-153	-8	-5
10	10	2.57	-12	36	-83	-4	-3
10	10	3.01	-6	37	-37	0	0
10	10	**3.52**	**0**	**38**	**-8**	**2**	**2**
10	10	4.12	7	38	19	5	5
10	10	4.83	12	38	63	8	7
10	10	5.66	19	38	158	12	9
10	10	6.63	32	37	345	18	12
10	20	2.17	-54	47	-126	-37	-23
10	20	2.54	-38	58	-69	-24	-18
10	20	2.98	-24	67	-31	-13	-12
10	20	3.49	-12	73	-9	-4	-7
10	20	**4.08**	**-2**	**77**	**2**	**2**	**-1**
10	20	4.78	9	78	13	8	4
10	20	5.60	20	75	40	16	10
10	20	6.56	36	69	105	29	15
10	20	7.69	61	59	240	43	21
20	10	2.02	-50	33	-326	-28	-16
20	10	2.53	-34	36	-187	-18	-12
20	10	3.16	-21	38	-92	-11	-8
20	10	3.95	-12	40	-36	-5	-5
20	10	**4.94**	**-3**	**41**	**-5**	**-1**	**-1**
20	10	6.18	6	42	23	4	3
20	10	7.73	16	41	94	11	8
20	10	9.67	35	40	284	21	13
20	10	12.09	75	37	685	32	18

Table 9: Deviations from the Monte Carlo baseline for different approximations to the swaption prices reported in Table 8 for a LMM in the shifted setup. Results for the ATM strike are boldfaced.

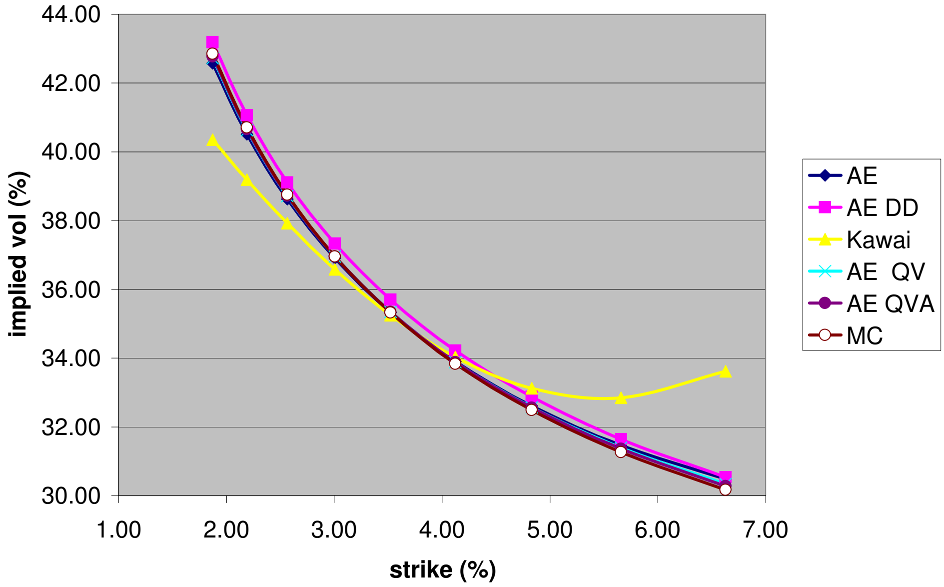

Figure 7: Numerical results for implied Black volatilities of 10y10 swaptions in a LMM defined by Eq. (109) with the inputs given in Table 5 and shifted setup, $\beta_n(t) \equiv 1/(2l_n)$.

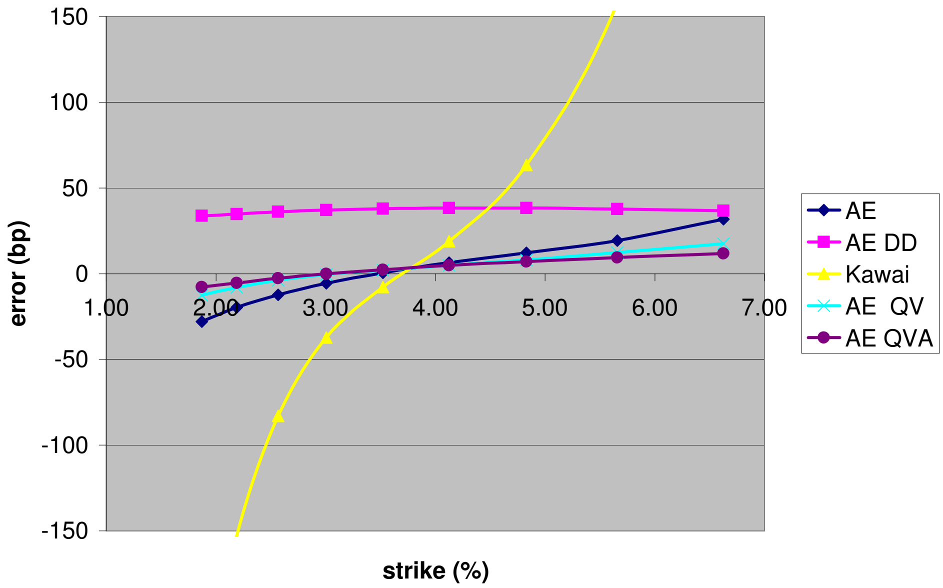

Figure 8: Deviations from the Monte Carlo baseline for different approximations to the 10y10 swaption prices reported in Table 8 for a LMM in the shifted setup.

6 Conclusions

Building on earlier developments in the asymptotic expansion for local volatility models and LMM, most notably by Kawai [14], we worked out first three terms in the asymptotic expansion for the option price in a generic form for a large class of multifactor models, including LMM. Numerical experiments showed that using the corrections to the option price directly is not the best approach. Instead much better accuracy for a range of option strikes is achieved if the original (more complicated) model is replaced by an approximating (simpler) model in such a way that the asymptotic expansion corrections in both models agree. The process of going from the original model to the approximating model can be regarded as time averaging of model parameters (if the approximating model differs from the original one only by the presence of time-dependent parameters) or as Markovian projection (if the original model is not Markovian while the approximating model is). In this work we used the quadratic volatility model as the approximating model and studied the cases of both parameter averaging and Markovian projection.

We found that the effective linear (skew) parameter of the approximating model is always independent of the option strike and agrees with the expressions obtained earlier by Piterbarg [17] in the case of parameter averaging and by the authors [3] in the case of LMM swaption. On the other hand, the curvature parameter can be made strike-independent only with an additional adjustment of the local volatility, in agreement with a recent result by Andersen and Hutchings [2]. Numerical comparison of the accuracy of the approximations with strike-independent and strike-dependent curvature showed that the former usually outperforms the latter and also generally outperforms simpler approximating models.

An asymptotic expansion starts as a series in iterated Itô integrals and can become cumbersome. To control the complexity of resulting expressions, we systematically explored matrix and tensor structures of intermediate quantities, leading to condensed notations and efficient numerical implementations. One notable technical simplification in comparison with previous works on asymptotic expansions is the use of the product formula for iterated Itô integrals, avoiding an unnecessary step of the calculation of expectations conditional on the density in the Gaussian approximation. We expect that the asymptotic expansions will be regarded as a much more accessible technique once the product formula takes its due place on par with Itô's lemma in the toolkit of a quantitative finance practitioner.

Further prospective applications of the described approach include Heston model and stochastic volatility LMM, in particular with a view to establishing links with the projection on a Heston model [4].

We are grateful to Leif Andersen and Vladimir Piterbarg for stimulating discussions. AA is indebted to his colleagues at Numerix, especially, to Gregory Whitten, Serguei Mechkov and Serguei Issakov for supporting this work.

References

[1] L. Andersen (2008) "Option Pricing with Quadratic Volatility: A Revisit," available at SSRN: http://ssrn.com/abstract=1118399.

[2] L. Andersen and N. Hutchings (2008) "Parameter Averaging of Quadratic SDEs with Stochastic Volatility," available at SSRN: http://ssrn.com/abstract=1339971.

[3] A. Antonov and T. Misirpashaev (2006) "Markovian Projection onto a Displaced Diffusion: Generic Formulas with Applications," available at SSRN: http://ssrn.com/abstract=937860.

[4] A. Antonov, T. Misirpashaev, and V. Piterbarg (2007) "Markovian Projection onto a Heston Model," available at SSRN: http://ssrn.com/abstract=997001.

[5] M. Avellaneda, D. Boyer-Olson, J. Busca, and P. Friz (2002) "Reconstructing Volatility," *Risk*, October, 91–95.

[6] A. Brace, D. Gatarek, and M. Musiela (1997) "The Market Model of Interest Rate Dynamics," *Mathematical Finance*, **7**, 127–155.

[7] G. Di Nunno, B. Øksendal, and F. Proske (2009) "Malliavin Calculus for Lévy Processes with Applications to Finance," Springer.

[8] E. Fournié, F. Lasry, J. Lebuchoux, and J. Lions (2001), "Applications of Malliavin Calculus to Monte Carlo Methods in Finance II," *Finance and Stochastics*, **5**, 201–236.

[9] I. Gyöngy (1986) "Mimicking the One-Dimensional Marginal Distributions of Processes Having an Itô Differential," *Prob. Theory and Rel. Fields*, **71**, 501–516.

[10] P. Henry-Labordère (2008) "Analysis, Geometry, and Modeling in Finance: Advanced Methods in Option Pricing," Chapman & Hall/CRC.

[11] C. Houdré, V. Pérez-Abreu, and A. S. Üstünel (1994) "Multiple Wiener-Itô Integrals: An Introductory Survey," in *Chaos Expansions, Multiple Wiener-Itô Integrals, and Their Applications*, CRC Press, Inc., Eds. C. Houdré and V. Pérez-Abreu, 1–34.

[12] K. Itô (1951) "Multiple Wiener Integral," *J. Math. Soc. Japan*, **3**, 157–169.

[13] F. Jamshidian (1997) "LIBOR and Swap Market Models and Measures," *Finance and Stochastics*, **1**, 293–330.

[14] A. Kawai (2003) "A New Approximate Swaption Formula in the Libor Market Model: An Asymptotic Expansion Approach," *Appl. Math. Finance*, **10**, 49–74.

[15] A. Kawai and P. Jäckel (2007) "An Asymptotic FX Option Formula in the Cross Currency Libor Market Model," *Wilmott*, March, 74–84.

[16] N. Kunimoto and A. Takahashi (2001), "The Asymptotic Expansion Approach to the Valuation of Interest Rate Contigent Claims," *Mathematical Finance*, **11**, 117–151.

[17] V. Piterbarg (2005) "Time to Smile,"*Risk*, May, 71–75.

[18] V. Piterbarg (2007) "Markovian Projection for Volatility Calibration," *Risk*, April, 84–89.

[19] A. Takahashi and K. Takehara (2008) "A Hybrid Asymptotic Expansion Scheme: an Application to Long-Term Currency Options", working paper

[20] S. Watanabe (1987) "Analysis of Wiener Functionals (Malliavin Calculus) and its Applications to Heat Kernels," *Ann. Probab.*, **15**, 1–39.

[21] N. Yoshida (1992) "Asymptotic Expansions of Maximum Likelihood Estimator for Small Diffusions via the Theory of Malliavin-Watanabe," *Prob. Theory and Rel. Fields*, **92**, 275–311.

A Calculation of expectations using Wiener chaos decomposition

The calculation of the expectations of the form

$$E\left[\exp\left(\xi\int_0^T dW(t)\lambda(t)\right)P\right], \tag{151}$$

where P is a polynomial in one or several iterated Itô integrals with respect to $dW(t)$ is conveniently handled using the technique of Wiener chaos decomposition, introduced by Itô [12] (see [11], [7] for a summary of more recent work).

We use the standard notation for the basic integral of order n associated with a symmetric function of n arguments $f(t_1,\ldots,t_n)$,

$$I_n(f_n) = n!\int_0^T dW(t_1)\int_0^{t_1} dW(t_2)\ldots\int_0^{t_{n-1}} dW(t_n)f(t_1,\ldots,t_n). \tag{152}$$

By convention, $I_0(f_0) = f_0$, where f_0 is a number considered as a function with no arguments. Two facts that make the standard integrals an extremely versatile tool in the calculations are the *orthogonality relation*,

$$E\left[I_m(f_m)I_n(g_n)\right] = \delta_{mn}n!\int_0^T dt_1\ldots\int_0^T dt_n f(t_1,\ldots,t_n)g(t_1,\ldots,t_n), \tag{153}$$

and the *product formula*,

$$I_m(f_m)I_n(g_n) = \sum_{r=0}^{\min(m,n)} r!\binom{m}{r}\binom{n}{r}I_{m+n-2r}(f_m\hat{\otimes}_r g_n), \tag{154}$$

where $\hat{\otimes}_r$ denotes the partially contracted symmetrized tensor product,

$$(f_m\hat{\otimes}_r g_n) = \mathrm{Sym}(f_m\otimes_r g_n), \tag{155}$$

that is the result of the ordinary partial contraction of two symmetric functions,

$$(f_m\otimes_r g_n)(t_1,\ldots,t_{m-r},s_1,\ldots,s_{n-r}) = \\ \int_0^T du_1\ldots\int_0^T du_r f_m(t_1,\ldots,t_{m-r},u_1,\ldots,u_r)g_n(s_1,\ldots,s_{n-r},u_1,\ldots,u_r), \tag{156}$$

followed by a symmetrization over the remaining free variables $t_1,\ldots,t_{m-r},s_1,\ldots,s_{n-r}$. The symmetrization for a function of n arguments is defined as an average over the group S_n of all n-element permutations π,

$$(\mathrm{Sym}f_n)(t_1,\ldots,t_n) = \frac{1}{n!}\sum_{\pi\in S_n} f_n(t_{\pi_1},\ldots,t_{\pi_n}). \tag{157}$$

The product formula is a powerful generalization of the well-known elementary result

$$\left(\int_0^T dW(t)\lambda(t)\right)^2 = \int_0^T dt\,\lambda^2(t) + 2\int_0^T dW(t_1)\lambda(t_1)\int_0^{t_1} dW(t_2)\lambda(t_2). \tag{158}$$

We also note the following equation for non-stochastic iterated integrals, which is often used (in both directions) to simplify final integral expressions,

$$\left(\int_0^T dt\,f(t)\right)^n = n!\int_0^T dt_1 f(t_1)\int_0^{t_1} dt_2 f(t_2)\ldots\int_0^{t_{n-1}} dt_n f(t_n). \tag{159}$$

Of course, the orthogonality relation follows from the product formula if we take into account another fundamental fact,

$$E\left[I_n(f_n)\right] = 0, \quad \text{for} \quad n > 0. \tag{160}$$

The first step in the calculation of the expectation (151) consists of reducing P into a linear combination $\sum_k I_k(f_k)$ by repeated use of the product formula. After that the calculation is completed using the orthogonality relation and the decomposition of the exponential,

$$\exp(\xi I_1(\lambda)) = \exp\left(\frac{\xi^2}{2}\int_0^T dt\lambda^2(t)\right)\sum_{n=0}^{\infty}\frac{\xi^n}{n!}I_n(\lambda^{\otimes n}), \tag{161}$$

where $\lambda^{\otimes n}(t_1,\ldots,t_n) = \lambda(t_1)\ldots\lambda(t_n)$ is the nth tensor power of the function $\lambda(t)$.

Consider, for example, the case encountered in the calculation of the asymptotic expansion for the quadratic volatility model,

$$P = \left(2\int_0^T dW(t_1)a(t_1)\int_0^{t_1} dW(t_2)b(t_2)\right)^n, \tag{162}$$

with $n = 1, 2$. We notice first that $P = (I_2(f))^n$, where $f = f(t_1, t_2)$ is the integrand written in a symmetric form[3]

$$f(t_1,t_2) = a(t_1)b(t_2)\mathbf{1}_{t_1>t_2} + a(t_2)b(t_1)\mathbf{1}_{t_2>t_1}. \tag{163}$$

With $n = 1$, we can write down the answer right away,

$$\begin{aligned} e^{-\frac{\xi^2}{2}\int_0^T dt\,\lambda^2(t)}E[\exp(\xi I_1(\lambda))I_2(f)] &= (\xi^2/2)E[I_2(\lambda\otimes\lambda)I_2(f)] \\ &= 2\xi^2\int_0^T dt_1\lambda(t_1)a(t_1)\int_0^{t_1} dt_2\lambda(t_2)b(t_2). \end{aligned} \tag{164}$$

With $n = 2$, we first need to use the product formula to obtain

$$I_2(f)I_2(f) = h_0 + I_2(h_2) + I_4(h_4), \tag{165}$$

where

$$h_0 = 2f\hat{\otimes}_4 f = 2\int_0^T\int_0^T f^2(t_1,t_2)dt_1dt_2, \tag{166}$$

$$h_2(t_1,t_2) = 4f\hat{\otimes}_2 f = 4\int_0^T f(t_1,t)f(t_2,t)dt, \tag{167}$$

$$h_4(t_1,t_2,t_3,t_4) = f\hat{\otimes} f = \frac{1}{3}(f(t_1,t_2)f(t_3,t_4) + f(t_1,t_3)f(t_2,t_4) + f(t_1,t_4)f(t_2,t_3)). \tag{168}$$

Now we are in a position to compute the desired expectation,

$$\begin{aligned} e^{-\frac{\xi^2}{2}\int_0^T dt\,\lambda^2(t)}E\left[\exp(\xi I_1(\lambda))(I_2(f))^2\right] &= h_0 + 2\xi^2\int_0^T dt_1\int_0^{t_1} dt_2 h_2(t_1,t_2)\lambda(t_1)\lambda(t_2) \\ &+ 24\xi^4\int_0^T dt_1\int_0^{t_1} dt_2\int_0^{t_2} dt_3\int_0^{t_3} dt_4 h_4(t_1,t_2,t_3,t_4)\lambda(t_1)\lambda(t_2)\lambda(t_3)\lambda(t_4) \\ &= 4\int_0^T dt_1 a^2(t_1)\int_0^{t_1} dt_2 b^2(t_2) + 4\xi^2\int_0^T dt_1\left(a(t_1)\int_0^{t_1} dt_2\lambda(t_2)b(t_2) + b(t_1)\int_{t_1}^T dt_3\lambda(t_3)a(t_3)\right)^2 \\ &+ 4\xi^4\Big(\int_0^T dt_1\lambda(t_1)a(t_1)\int_0^{t_1} dt_2\lambda(t_2)b(t_2)\Big)^2. \end{aligned} \tag{169}$$

[3]Note that it would be a mistake to use a full symmetrization $(a\hat{\otimes}b)(t_1,t_2) = (a(t_1)b(t_2) + a(t_2)b(t_1))/2$ because this function is not equal to $a(t_1)b(t_2)$ in the domain $0 < t_2 < t_1 < T$. Note also that we do not need to be concerned about the behavior on the diagonal $t_1 = t_2$ as long as we are dealing with non-singular integrands.

Applications of the method to multifactor problems of mathematical finance require a generalization to the case of a multidimensional Brownian motion, $W(t) = (W_1(t), \ldots, W_F(t))$ with independent components. The definition of the basic iterated integral $I_n(f_n)$ is generalized to be based on a tensor of functions, $f_n^{\alpha_1 \ldots \alpha_n}(t_1, \ldots, t_n)$, symmetric with respect to a simultaneous permutation of arguments and corresponding factor indexes,

$$I_n(f_n^{\alpha_1 \ldots \alpha_n}) = n! \sum_{\alpha_1,\ldots,\alpha_n} \int_0^T dW_{\alpha_1}(t_1) \int_0^{t_1} dW_{\alpha_2}(t_2) \ldots \int_0^{t_{n-1}} dW_{\alpha_n}(t_n) f_n^{\alpha_1 \ldots \alpha_n}(t_1, \ldots, t_n). \tag{170}$$

with tensor indexes running over the factors. The orthogonality relation (153) remains valid with the understanding that contractions are done both over time arguments and corresponding factor indexes,

$$E\left[I_m(f_m^{\alpha_1 \ldots \alpha_m}) I_n(g_n^{\beta_1 \ldots \beta_n}) \right] = \delta_{mn} n! \sum_{\alpha_1,\ldots,\alpha_n} \int_0^T dt_1 \ldots \int_0^T dt_n f^{\alpha_1 \ldots \alpha_n}(t_1, \ldots, t_n) g^{\alpha_1 \ldots \alpha_n}(t_1, \ldots, t_n). \tag{171}$$

The product formula (154) also remains valid with a similar extension of the definition of partial contraction,

$$\begin{aligned}(f_m \otimes_r g_n)^{\alpha_1 \ldots \alpha_{m-r} \beta_1 \ldots \beta_{n-r}}(t_1, \ldots, t_{m-r}, s_1, \ldots, s_{n-r}) = \sum_{\gamma_1,\ldots,\gamma_r} \int_0^T du_1 \ldots \int_0^T du_r \\ f_m^{\alpha_1 \ldots \alpha_{m-r} \gamma_1, \ldots, \gamma_r}(t_1, \ldots, t_{m-r}, u_1, \ldots, u_r) g_n^{\beta_1 \ldots \beta_{n-r} \gamma_1, \ldots, \gamma_r}(s_1, \ldots, s_{n-r}, u_1, \ldots, u_r),\end{aligned} \tag{172}$$

followed by symmetrization with respect to the group S_n of simultaneous permutations π of arguments and factor indexes,

$$(\mathrm{Sym} f_n)^{\alpha_1 \ldots \alpha_n}(t_1, \ldots, t_n) = \frac{1}{n!} \sum_{\pi \in S_n} f_n^{\alpha_{\pi_1} \ldots \alpha_{\pi_n}}(t_{\pi_1}, \ldots, t_{\pi_n}). \tag{173}$$

Consequently, results derived for the 1-factor case typically extend to the multifactor case with the replacement of scalar coefficient functions, such as $\lambda(t)$, $a(t)$, $b(t)$ in the example considered above by vectors (or possibly higher order tensors in more complicated examples) in the factor space. In particular, the derivation of the averages (88) and (90) mirrors the derivation given above for the scalar case, starting with the basic integral I_2 of a tensor function

$$f^{\alpha_1 \alpha_2}(t_1, t_2) = \sum_n \left(a_n^{\alpha_1}(t_1) b_n^{\alpha_2}(t_2) \mathbf{1}_{t_1 > t_2} + a_n^{\alpha_2}(t_2) b_n^{\alpha_1}(t_1) \mathbf{1}_{t_2 > t_1} \right). \tag{174}$$

Additional contraction (scalar product) over the factor indexes is understood in all cases of integration of products of functions taken at equal times. In this paper, factor indexes are usually omitted, scalar products are denoted by a dot, and component-wise products are denoted using the standard tensor product sign $\otimes$ where it is helpful for the understanding of the formulas.

We conclude this brief excursus into the Wiener chaos techniques by the formula for the average of a pure multiple Itô integral, which has a particularly simple structure,

$$\begin{aligned} e^{-\frac{\xi^2}{2} \int_0^T dt\, |\lambda(t)|^2} E\left[e^{\xi \int_0^T dW(t) \cdot \lambda(t)} \int_0^T dW(t_1) \cdot a_1(t_1) \int_0^{t_1} dW(t_2) \cdot a_2(t_2) \ldots \int_0^{t_{n-1}} dW(t_n) \cdot a_n(t_n) \right] \\ = \xi^n \int_0^T dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{n-1}} dt_n (\lambda \otimes \lambda \otimes \ldots \otimes \lambda) \cdot (a_1 \otimes a_2 \otimes \ldots \otimes a_n)(t_1, t_2, \ldots, t_n) \\ = \xi^n \int_0^T dt_1 \lambda(t_1) \cdot a_1(t_1) \int_0^{t_1} dt_2 \lambda(t_2) \cdot a_2(t_2) \ldots \int_0^{t_{n-1}} dt_n \lambda(t_n) \cdot a_n(t_n). \end{aligned} \tag{175}$$

B Details of LMM swaption calculation

Here we provide additional details related to the study of the LMM swaption pricing using asymptotic expansion in Sect. 4.3.

B.1 LMM with time-dependent shift functions

In Sect. 4.3 we restricted the derivation of the asymptotic expansion of the swaption price in LMM to the case of time-independent shift functions $\beta_n(t)$. Here we show that the method of embedding the rate process into a separable model can also be used with time-dependent shifts, although the final expressions are more cumbersome. We split the right-hand side of Eq. (125) into two separate summations,

$$dX(t) = \sum_n \frac{\partial X}{\partial L_n}\, \gamma_n(t) \cdot dW^A(t) + \sum_n \frac{\partial X}{\partial L_n}\, \beta_n(t)\, \Delta L_n(t) \gamma_n(t) \cdot dW^A(t). \tag{176}$$

We now need twice as many quantities $\{S_m(t)\}$ as the number of LIBORs $\{L_n(t)\}$ to complete the immersion of the swap rate $X(t)$ into a separable model,

$$\begin{aligned} S_{2k+1}(t) &= \frac{\partial X}{\partial L_k}, & a_{2k+1}(t) = \gamma_k(t), \quad (177)\\ S_{2k}(t) &= \frac{\partial X}{\partial L_k} \Delta L_k(t), & a_{2k}(t) = \beta_k(t) \gamma_k(t). \quad (178) \end{aligned}$$

Using Itô's lemma, we obtain $dS_n(t) = \mu_n(t)\, dt + \sigma_n(t) \cdot dW^A(t)$ with

$$\begin{aligned} \mu_{2k+1}(t) &= \sum_n \frac{\partial^2 X}{\partial L_n\, \partial L_k} \nu_n(t) + \frac{1}{2} \sum_{n,m} \frac{\partial^3 X}{\partial L_n\, \partial L_m\, \partial L_k} \lambda_n(t) \cdot \lambda_m(t), & (179)\\ \sigma_{2k+1}(t) &= \sum_n \frac{\partial^2 X}{\partial L_n\, \partial L_k} (1 + \beta_n(t)\, \Delta L_n)\, \gamma_n(t), & (180)\\ \mu_{2k}(t) &= \Delta L_k(t)\, \mu_{2k+1}(t) + S_{2k+1}(t) \nu_k(t) + \sigma_{2k+1}(t) \cdot \lambda_k(t), & (181)\\ \sigma_{2k}(t) &= \Delta L_k(t)\, \sigma_{2k+1}(t) + S_{2k+1}(t)\, \lambda_k(t). & (182) \end{aligned}$$

We also need the following set of derivatives over LIBORs,

$$\begin{aligned} \frac{\partial \sigma_{2k+1}(t)}{\partial L_m} &= \sum_n \frac{\partial^3 X}{\partial L_n\, \partial L_k\, \partial L_m} (1 + \beta_n(t)\, \Delta L_n(t))\, \gamma_n(t) + \frac{\partial^2 X}{\partial L_m\, \partial L_k} \beta_m(t)\, \gamma_m(t), & (183)\\ \frac{\partial \sigma_{2k}(t)}{\partial L_m} &= \delta_{km} \left(\sigma_{2k+1}(t) + \frac{\partial X}{\partial L_k}\, \beta_k(t)\, \gamma_k(t) \right) & \\ &+ \Delta L_k(t) \frac{\partial \sigma_{2k+1}(t)}{\partial L_m} + \frac{\partial^2 X}{\partial L_m\, \partial L_k} (1 + \beta_k(t)\, \Delta L_k(t))\, \gamma_k(t). & (184) \end{aligned}$$

Finally we "freeze" stochastic quantities, that is evaluate them for the initial LIBOR values, $L_m(0) = l_m$, keeping the time dependence for deterministic parameter functions of the model (volatilities and shifts). We get

$$
\begin{aligned}
a_{2k+1}(t) &= \gamma_k(t), && (185)\\
a_{2k}(t) &= \beta_k(t)\gamma_k(t), && (186)\\
S_{2k+1}(0) &= \frac{\partial X_0}{\partial l_k}, && (187)\\
S_{2k}(0) &= 0, && (188)\\
\lambda(t) &= \sum_n \frac{\partial X_0}{\partial l_n}\gamma_n(t), && (189)\\
\hat{\sigma}_{2k+1}(t) &= \sum_n \frac{\partial^2 X_0}{\partial l_n\,\partial l_k}\gamma_n(t), && (190)\\
\hat{\sigma}_{2k}(t) &= \frac{\partial X_0}{\partial l_k}\,\gamma_k(t), && (191)\\
\hat{\lambda}_k(t) &= \gamma_k(t) && (192)\\
\hat{\nu}_k(t) &= -\sum_m \frac{\partial \ln M_{k+1,A}(0)}{\partial l_m}\,\gamma_m(t)\cdot\gamma_k(t), && (193)\\
\hat{\mu}_{2k+1}(t) &= \sum_n \frac{\partial^2 X_0}{\partial l_n\,\partial l_k}\hat{\nu}_n(t) + \frac{1}{2}\sum_{n,m}\frac{\partial^3 X_0}{\partial l_n\,\partial l_m\,\partial l_k}\gamma_n(t)\cdot\gamma_m(t), && (194)\\
\hat{\mu}_{2k}(t) &= \frac{\partial X_0}{\partial l_k}\,\hat{\nu}_k(t) + \sum_n \frac{\partial^2 X_0}{\partial l_n\,\partial l_k}\gamma_n(t)\cdot\gamma_k(t), && (195)\\
\frac{\partial\hat{\sigma}_{2k+1}}{\partial l_m}(t) &= \sum_n \frac{\partial^3 X_0}{\partial l_n\,\partial l_k\,\partial l_m}\,\gamma_n(t) + \frac{\partial^2 X_0}{\partial l_m\,\partial l_k}\beta_m(t)\,\gamma_m(t), && (196)\\
\frac{\partial\hat{\sigma}_{2k}}{\partial L_m}(t) &= \delta_{km}\left(\sum_n \frac{\partial^2 X_0}{\partial l_n\,\partial l_k}\gamma_n(t) + \frac{\partial X_0}{\partial l_k}\,\beta_k(t)\,\gamma_k(t)\right) + \frac{\partial^2 X_0}{\partial l_m\,\partial l_k}\,\gamma_k(t). && (197)
\end{aligned}
$$

The resulting set of functions can be plugged into Eqs. (91–95), defining the coefficients in the corrections to option prices from the asymptotic expansion.

B.2 Derivation of Eqs. (142–146)

First we note that Eq. (132) can be written as

$$\hat{\sigma}_n(t) = \sum_k \mathcal{M}_{nk}\gamma_k(t), \tag{198}$$

in terms of a symmetric matrix $\mathcal{M}$ defined by Eq. (138). The quantities $C(T)$, $Q(t)$, and $R(T)$ defined by Eqs. (91),(94), and (95) are based on a kernel

$$\xi(t_1,t_2) = \sum_n a_n(t_1)\otimes\hat{\sigma}_n(t_2) = \sum_{n,k}\mathcal{M}_{nk}\gamma_n(t_1)\gamma_k(t_2), \tag{199}$$

which is a tensor in implicitly present indexes[4] corresponding to the components of the driving Brownian motion $W(t) = (W_1(t), \ldots, W_F(t))$. Indeed,

$$
\begin{aligned}
C(T) &= \sum_n \int_0^T dt_1\, \lambda(t_1)\cdot a_n(t) \int_0^{t_1} dt_2\, \lambda(t_2)\cdot\hat{\sigma}_n(t_2) \\
&= \int_0^T dt_1 \int_0^{t_1} dt_2\, \xi(t_1, t_2)\cdot(\lambda(t_1)\otimes\lambda(t_2)) \\
&= \int_0^T dt_1 \int_0^{t_1} dt_2 \sum_{n,k} \mathcal{M}_{nk}\gamma_k(t_1)\cdot\lambda(t_1)\gamma_n(t_2)\lambda(t_2) = \frac{1}{2} I^T(T)\,\mathcal{M}\,I(T),
\end{aligned} \tag{200}
$$

$$
\begin{aligned}
Q(t) &= \sum_n a_n(t) \int_0^t dt_1\, \lambda(t_1)\cdot\hat{\sigma}_n(t_1) + \sum_n \hat{\sigma}_n(t) \int_t^T dt_1\, \lambda(t_1)\cdot a_n(t_1) \\
&= \int_0^t dt_1\, \xi(t, t_1)\cdot(1\otimes\lambda(t_1)) + \int_t^T dt_1\, \xi^T(t_1, t)\cdot(1\otimes\lambda(t_1)) \\
&= \sum_{n,k} \mathcal{M}_{nk}\gamma_k(t) \int_0^T dt_1 \gamma_n(t_1)\cdot\lambda(t_1) = \sum_{n,k} \mathcal{M}_{nk}\gamma_k(t) I_n(T),
\end{aligned} \tag{201}
$$

$$
\begin{aligned}
R(T) &= \sum_{n,k} \int_0^T dt_1\, a_k(t_1)\cdot a_n(t_1) \int_0^{t_1} dt_2\, \hat{\sigma}_k(t_2)\cdot\hat{\sigma}_n(t_2) \\
&= \int_0^T dt_1 \int_0^{t_1} dt_2\, \xi(t_1, t_2)\cdot\xi(t_1, t_2) \\
&= \sum_{n,k,m,s} \mathcal{M}_{nk}\,\mathcal{M}_{mi} \int_0^T dt_1 \int_0^{t_1} dt_2\, \gamma_k(t_1)\cdot\gamma_m(t_1)\;\gamma_n(t_2)\cdot\gamma_s(t_2) \\
&= \frac{1}{2} \sum_{n,k,m,s} \mathcal{M}_{nk}\,\mathcal{M}_{ms}\, V_{km}(T)\, V_{sn}(T) = \frac{1}{2}\mathrm{Tr}\left(\mathcal{M}V(T)\mathcal{M}V(T)\right).
\end{aligned} \tag{202}
$$

Using Eq. (201), the integral $\int_0^T dt|Q(t)|^2$ simplifies to

$$
\int_0^T dt|Q(t)|^2 = \sum_{n,k,m,s} \mathcal{M}_{nk}\mathcal{M}_{ms} I_n(T)\, I_m(T) \int_0^T dt\gamma_k(t)\cdot\gamma_s(t) = I^T(T)\;\mathcal{M}V(T)\mathcal{M}\;I(T). \tag{203}
$$

This completes the proof of Eqs. (142–144).

The coefficient $D(T)$ for the separable model, defined by Eq. (92),

$$
D(T) = \sum_{n,m} \int_0^T dt_1\, \lambda(t_1)\cdot a_n(t_1) \int_0^{t_1} dt_2\, \lambda(t_2)\cdot\frac{\partial\hat{\sigma}_n}{\partial l_m}(t_2) \int_0^{t_2} t_3\, \lambda(t_3)\cdot\hat{\lambda}_m(t_3),
$$

is based on another tensor kernel,

$$
\mathcal{O}_D(t_1, t_2, t_3) = \sum_{n,m} a_n(t_1)\otimes\frac{\partial\hat{\sigma}_n}{\partial l_m}(t_2)\otimes\hat{\lambda}_m(t_3), \tag{204}
$$

[4]Note that we take advantage of a matrix structure in the space of components of the separable model (with indexes written explicitly) and of a tensor structure in the space of components of the driving Brownian motion (with indexes omitted) simultaneously. This gives rise to concise expressions but requires care.

so that

$$D(T) = \int_0^T dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3\, \mathcal{O}_D(t_1,t_2,t_3)\cdot(\lambda(t_1)\otimes\lambda(t_2)\otimes\lambda(t_3)). \tag{205}$$

From Eq. (136), the tensor $\mathcal{O}_D(t_1,t_2,t_3)$ can be represented as

$$\begin{aligned}
\mathcal{O}_D(t_1,t_2,t_3) &= \sum_{n,m,k} \mathcal{D}^{(3)}_{nmk}\, \gamma_n(t_1)\otimes\gamma_k(t_2)\otimes\gamma_m(t_3) \\
&+ \sum_{n,m} \mathcal{D}^{(2)}_{nm}\, \beta_m\gamma_n(t_1)\otimes\gamma_m(t_2)\otimes\gamma_m(t_3) \\
&+ \sum_{n,m} \mathcal{D}^{(2)}_{nm}\, \beta_n\gamma_n(t_1)\otimes\gamma_n(t_2)\otimes\gamma_m(t_3) \\
&+ \sum_{k,n} \mathcal{D}^{(2)}_{kn}\, \beta_n\gamma_n(t_1)\otimes\gamma_k(t_2)\otimes\gamma_n(t_3) \\
&+ \sum_{n} \mathcal{D}^{(1)}_{n}\, \beta_n^2\gamma_n(t_1)\otimes\gamma_n(t_2)\otimes\gamma_n(t_3),
\end{aligned} \tag{206}$$

leading to Eq. (145).

Finally, the coefficient $M(T)$ for the separable model (Eq. (93)) is given by

$$\begin{aligned}
M(T) &= \sum_n \int_0^T dt_1\, \lambda(t_1)\cdot a_n(t) \int_0^{t_1} dt_2\, \hat{\mu}_n(t_2) \\
&= \int_0^T dt_1 \int_0^{t_1} dt_2\, \mathcal{O}_M(t_1,t_2)\cdot\lambda(t_1).
\end{aligned} \tag{207}$$

We again introduce an appropriate kernel,

$$\mathcal{O}_M(t_1,t_2) = \sum_n a_n(t_1)\, \hat{\mu}_n(t_2), \tag{208}$$

which is a vector in a space of factors with the factor index implicitly carried by $a_n(t_1) \equiv \gamma_n(t_1)$. Plugging the expressions (134–135) into Eq. (208), we get

$$\begin{aligned}
\mathcal{O}_M(t_1,t_2) &= \sum_{n,m,k} \left(\frac{1}{2}\mathcal{D}^{(3)}_{nmk} - \mathcal{D}^{(2)}_{nk}\frac{\partial \ln M_{n+1,A}}{\partial l_m}\right) \gamma_k(t_1)\, \gamma_n(t_2)\cdot\gamma_m(t_2) \\
&+ \sum_{m,k} \left(\mathcal{D}^{(2)}_{mk} - \mathcal{D}^{(1)}_{k}\frac{\partial \ln M_{k+1,A}}{\partial l_m}\right) \beta_k\, \gamma_k(t_1)\, \gamma_k(t_2)\cdot\gamma_m(t_2).
\end{aligned} \tag{209}$$

Further simplification is possible using the drift cancelation condition (122) taken for initial LIBOR values,

$$\mathcal{D}^{(1)}_{n}\frac{\partial \ln M_{n+1,A}(0)}{\partial l_m} + \mathcal{D}^{(1)}_{m}\frac{\partial \ln M_{m+1,A}(0)}{\partial l_n} = \mathcal{D}^{(2)}_{nm}. \tag{210}$$

This leads to

$$\begin{aligned}
\mathcal{O}_M(t_1,t_2) &= \sum_{n,m,k} \mathcal{D}^{(1)}_{m}\frac{\partial^2 \ln M_{m+1,A}}{\partial l_n\, \partial l_k}\, \gamma_k(t_1)\, \gamma_n(t_2)\cdot\gamma_m(t_2) \\
&+ \sum_{m,k} \mathcal{D}^{(1)}_{m}\frac{\partial \ln M_{m+1,A}}{\partial l_k}\, \beta_k\, \gamma_k(t_1)\, \gamma_k(t_2)\cdot\gamma_m(t_2),
\end{aligned} \tag{211}$$

and the final result (146).

B.3 Comparison with [14]

The asymptotic expansion for the swaption price in LMM was first derived by Kawai [14] with a generic local volatility function $\zeta_n(L_n)$ in the dynamics of the LIBOR rates in their martingale measures[5],

$$dL_n(t) = \zeta_n(L_n)\gamma_n(t) \cdot dW^{(n+1)}(t). \tag{212}$$

In our work, we set $\zeta_n(L_n) = 1 + \beta_n \Delta L_n$ and fully reproduced the results of [14] for the three leading terms in the asymptotic expansion of the swaption price. Indeed, we verified the correspondence between the coefficients in Lemma 4.6 of [14] in the left-hand side and our coefficient in the right-hand side of the following equations[6],

$$\Sigma = v(T), \tag{213}$$

$$c_1 = C(T)/v^2(T), \tag{214}$$

$$d_1 = \int_0^T dt |Q(t)|^2 / v^3(T), \tag{215}$$

$$d_2 = R(T)/v^2(T), \tag{216}$$

$$f_1 + f_3 = D(T)/v^3(T), \tag{217}$$

$$f_2 = M(T)/v^2(T). \tag{218}$$

After deriving the asymptotic expansion, Kawai observes that the correction to the option price can be expressed as the correction to the normal variance. In our notations,

$$C_0(T,k) + \varepsilon C_1(T,k) + \varepsilon^2 C_2(T,k) = \tilde{C}_0(T,k), \tag{219}$$

where

$$\tilde{C}_0(T,k) = \sqrt{\frac{\tilde{v}(T)}{2\pi}} e^{-k^2/2\tilde{v}(T)} - k\mathcal{N}(-k/\sqrt{\tilde{v}(T)}), \tag{220}$$

$$\begin{aligned} \tilde{v}(T) &= v(T) + 2\varepsilon\phi_{13} \frac{k}{v(T)} \\ &+ 2\varepsilon^2 \left[\phi_{22} - \frac{\phi_{24}}{v(T)} + \frac{3}{2}\frac{\phi_{13}^2}{v^2(T)} + \frac{k^2}{v(T)} \left(\frac{\phi_{24}}{v(T)} - \frac{5}{2}\frac{\phi_{13}^2}{v^2(T)} \right) \right] + O(\varepsilon^3). \end{aligned} \tag{221}$$

Matching of the effective normal variance $\tilde{v}(T)$ can be used as a condition for model mapping. One possibility explored in [14] is to use this as a condition for a mapping on a log-normal (that is Black's) model. Generalizing this procedure to the case where skew is present, we consider the target model

$$dX(t) = \varepsilon(1 + \beta \, \Delta X(t)) \sqrt{v_D/T} \, dW(t), \tag{222}$$

where $\beta = 1/X(0)$ would correspond to the log-normal case. Effective normal variance in this model is given by

$$\tilde{v}_D = v_D \left(1 + \varepsilon\beta \, k + \frac{1}{12} \varepsilon^2 \, v_D \, \beta^2 \left(\frac{k^2}{v_D} - 1 \right) \right) + O(\varepsilon^3). \tag{223}$$

It is not possible to match the terms corresponding to different powers of ε in Eqs. (221) and (223) separately with only one free parameter, v_D. The procedure proposed in [14] (and also

[5] Our quantities $\gamma_n(t)$ correspond to $\lambda_n(t)$ in the notations of [14].

[6] We believe there is a misprint in the expression for d_1 in Appendix A.2 in the copy of [14] available to us where the denominators in the right-hand side should be Σ^3 rather than Σ^2.

used in a similar setting in [2]) is to solve the equation $\tilde{v}(T) = \tilde{v}_D$ for v_D and set $\varepsilon = 1$ in the final result. Letting v_D be a function of ε is not consistent with the scheme of asymptotic expansion, however, the resulting approximation may perform reasonably well in practice. In fact, we confirmed in Sect. 5.2 that it does work well in the log-normal case but not in the case with a significant skew.