International Journal of Theoretical and Applied Finance

ANALYTIC APPROXIMATION TO CONSTANT MATURITY SWAP CONVEXITY CORRECTION IN A MULTI-FACTOR SABR MODEL

BIN CHEN*

MAS-2, Center for Mathematics and Computer Science, Science Park 123,
1098 XG, Amsterdam, the Netherlands
B.Chen@cwi.nl

CORNELIS W. OOSTERLEE

MAS-2, Center for Mathematics and Computer Science, Science Park 123,
1098 XG, , Amsterdam, the Netherlands
C.W.Oosterlee@cwi.nl

SACHA VAN WEEREN

DR&V, Rabobank International, Croeselaan 18,
3521 CB, Utrecht, the Netherlands
Sacha.VanWeeren@rabobank.com

Received (17 January 2010)

We consider the convexity correction in a multi-factor SABR type stochastic volatility model, in which the volatility and the short-term forward rate are modeled as independent factors. In general, the convexity correction is not analytically tractable in a multi-factor model, but based on the assumption of linear swap rates an analytic solution is available. Linear swap rate models are popular among practitioners for their efficiency and their ability to capture the swaption volatility smile. For an efficient approximation of the solution, we adopt the small disturbance asymptotics technique and construct a stochastic Taylor series of the underlying process. Several numerical experiments compare the accuracy of the approximation with a Monte Carlo benchmark solution.

Keywords: Convexity Correction; SABR Model; Stochastic Taylor Expansion.

1. Introduction

The growing popularity of transactions of *constant maturity swap* (CMS) type in the fixed income market has increased the demand for accurate and efficient pricing methods. This research topic attracts efforts from academia and practitioners alike.

The main lines of research for pricing methods seem to go basically in two directions. In the first, one deals with the problem by setting up a term-structure model

*Corresponding author. Postal Address: P.O. Box 94079, 1090 GB, Amsterdam, the Netherlands. Email: B.Chen@cwi.nl

under the T-forward measure, where the pricing originally occurs. For example, Brigo and Mercurio [3] model the bond prices associated with the CMS swap and quanto CMS swap by a G2++ model (2-factor Gaussian short rate model). The papers by Lu and Neftci [24] and Henrard [7] express the CMS swap as a collection of forward LIBOR rates under the forward measure and compute numerically the CMS price in a full-factor LIBOR market model. These approaches result in black-box computational schemes in which the risk sensitivities, e.g. the Vega, cannot be derived directly.

In the second line of research the pricing problem is formulated under the so-called *swap measure* and the given implied swaption volatilities are considered as the 'market distribution of the swap rates'. Since CMS products are mainly hedged by forward swaps and swaptions, the advantage of the measure change approach is consistency between the CMS products and their hedging instruments. Because of the measure change, from the forward to the swap measure, the Radon-Nykodym derivatives need to be approximated. Hunt and Kennedy [9] and Pelsser [19] approximate this measure change ratio in terms of a linear function of the swap rate (assuming that the yield curve is mainly driven by the swap rate) and obtain an analytic solution to the CMS price. Hagan [4] and Mercurio [18] succeed in statically replicating the CMS swap/options by European swaptions. Because of the popularity of the static replication approach, an increasing volume of swaption transactions for hedging purposes has been observed in the market, resulting in a more pronounced smile. A problem is the assumption of a one-factor yield curve, as only *parallel shifts* in the yield curve can then be taken into account. However, a CMS structure depends significantly on the slope of the yield curve, but it is not very sensitive to parallel shifts [2].

In this paper, we adopt the Stochastic Alpha Beta Rho (SABR) model [5] to describe the dynamics of the underlying swap rate. The SABR model has the capability of generating rich skew/smile patterns and it is often used in the market [22]. We introduce an additional yield curve factor, next to the swap rate, in the measure change ratio, in order to take the dynamics at the short-end of the yield curve into account. Here, the CMS convexity correction is decomposed in two parts: A part driven by the variance of the swap rate, which is affected by the skew/smile in the implied swaption volatilities, and a second part related to the covariance between the swap and LIBOR rate, which is a result of the terminal decorrelation [a] between these two rates. One can view our pricing approach as a perturbation of the conventional CMS convexity correction away from the one-factor assumption.

[a]What influences the price of an exotic product, as Rebonato [21] states, is not the instantaneous correlation or volatility functions, but, the *terminal* (as opposed to instantaneous) decorrelation, $\bar{\rho}_{xy}(T)$, defined by

$$\bar{\rho}_{xy}(T) = \frac{\int_0^T \sigma_x(s)\sigma_y(s)\rho_{xy}(s)ds}{\sqrt{\int_0^T \sigma_x(s)^2 ds \int_0^T \sigma_y(s)^2 ds}}.$$

We obtain an analytic approximation formula for the covariance, based on the well-known stochastic Taylor expansion [15]. Deriving the stochastic Taylor expansion by a repeated use of Itô's lemma is somewhat cumbersome when higher orders terms are considered. We can simplify the derivation by adopting a *small disturbance asymptotics* technique (e.g. Yoshida [25] and Kunitomo [16]) to construct the Taylor series of the multi-factor SABR process.

There are a number of advantages to our approach. First of all, it models forward swap rates directly, and therefore achieves a very satisfactory agreement between the CMS contracts and their hedging instruments. Secondly, the SABR model can easily be calibrated to implied volatilities of the liquid swaptions. Thirdly, it reflects the CMS' price sensitivity to the yield curve forward correlation structure. And, finally, it provides an easy-to-implement approximation formula for the CMS convexity correction under the multi-factor model. Hence it can be used for a quick evaluation of the model risk resulting from the terminal decorrelation of the forward rates.

A less direct implication of our work is the following. The pricing of derivatives written on CMS contracts, such as CMS swaps and spread options, on the basis of underlying CMSs is impossible when the markets for the latter become illiquid, as in the recent financial distress. In such circumstances and as long as markets for plain interest rate swaps are still liquid, a conceptually sound and practically viable alternative is to price CMSs on the basis of the underlying swap prices and then price CMS derivative based on such 'synthetic CMS prices'.

The paper is organized as follows. In Section 2, the pricing problem is formulated in an arbitrage-free way so that CMS-based derivatives are consistently priced across measures. Section 3 presents the stochastic Taylor expansion formula to the covariance of a two-factor stochastic volatility model. Examples and corresponding results for the approximate model are summarized in Section 4, where the approximate solutions are compared against short time step Monte Carlo prices. In the last section, conclusions are made.

2. Problem Formulation

A feature which distinguishes CMS-type contracts from plain vanilla contracts is that they pay a swap rate of one maturity, say 10 years, at each resetting time, as opposed to a regular swap, which pays the same coupon rate throughout a whole period. Hence to compute the CMS rate an adjustment has to be made to the forward swap rate implied by the swap rate curve. This adjustment is convex in the swap rate as its 'official' name, *convexity correction*, suggests. The convexity (in the swap rate) is the result of positive correlation in the yield curve [b]. The

[b] Imagine the swap rate increases in value, then the discounting effect in the annuity, on which the forward swap is paid, will increase. As a result, the forward swap payoff may increase more slowly than the one-time-payment of the swap rate; On the other hand, when the swap rate decreases in value the discounting effect in the annuity may get smaller, and, consequently, the forward swap payoff may decrease faster than the one-time-payment swap rate.

one-time-payment of the swap rate is always greater than, or equal to, the regular forward swap rate.

The above description is merely heuristic; the mathematical set-up for the CMS contract will be described in more detail in Sec. 4.2.

Because of the existence of multiple admissible pricing measures [c], it is important to make sure that a product is consistently priced across measures without any *arbitrage* possibilities. The implication of this constraint is investigated in the following sections.

2.1. *Measure Change and Arbitrage-free Constraints*

Girsanov's transformation gives rise to a *convexity correction*, and pricing problems in general, due to the change of measure:

$$E^{Q*}[\phi(X_T)|\mathcal{F}_t] = \frac{N_t}{N_t^*} E^Q\Big[\frac{N_T^*}{N_T}\phi(X_T)|\mathcal{F}_t\Big], \tag{2.1}$$

where $\phi(X_T)$ is a payoff function; T is maturity time. We denote the value of X_t's natural numeraire at time t as N_t whereas N_t^* is the (unnatural) numeraire under which the payment terms are specified. Regarding the notation, we denote the filtration up to time t by a subscript t to the expectation sign, i.e. $E^Q[\cdot|\mathcal{F}_t] = E_t^Q[\cdot]$, whenever necessary. So, $E_0^Q[\cdot]$ indicates an expectation w.r.t the filtration up to current time point, $t = 0$.

In order to satisfy the no-arbitrage conditions, we make the following assumption:

Assumption 2.1. *All rates are priced in an arbitrage-free way under their own natural pricing measure. So, the rate X_t is a martingale process under the natural measure Q. Related to the RHS of Eq. (2.1), this assumption excludes the possibility of arbitrage in the rate X_t.*

By making use of the relation $E[XY] = E[X]E[Y] + \mathrm{Cov}[X,Y]$, one finds that the convexity correction originates from the covariance between two stochastic processes:

$$\begin{aligned} E_t^{Q*}[\phi(X_T)] &= \frac{N_t}{N_t^*} E_t^Q\Big[\frac{N_T^*}{N_T}\phi(X_T)\Big] \\ &= \frac{N_t}{N_t^*} E_t^Q\Big[\frac{N_T^*}{N_T}\Big] E_t^Q\Big[\phi(X_T)\Big] + \frac{N_t}{N_t^*}\mathrm{Cov}_t^Q\Big[\frac{N_T^*}{N_T}, \phi(X_T)\Big] \\ &= E_t^Q\Big[\phi(X_T)\Big] + \underbrace{\frac{N_t}{N_t^*}\mathrm{Cov}_t^Q\Big[\frac{N_T^*}{N_T}, \phi(X_T)\Big]}_{\text{Convexity correction } Cc_t}. \end{aligned} \tag{2.2}$$

The last equality in (2.2) is the result of the martingale property of the term N_T^*/N_T which is due to the fact that it is a ratio of two tradable assets and the martingale property of X_T under its natural measure Q.

[c] in this, as well as in many other interest rate derivative pricing problems.

Let us focus on the numeraire ratio N_T^*/N_T. Because the yield curve is highly correlated, changes in rate X_T give rise to proportional movements of the natural numeraire N_T. The numeraire N_T^* is, however, driven by another rate, which we denote by Y_T. So, the numeraire ratio is a function of two rates, i.e.

$$\frac{N_T^*}{N_T} = f(X_T, Y_T). \tag{2.3}$$

We further assume the following:

Assumption 2.2. *The function $f(X_s, Y_s)$ in Eq. (2.3) is smooth and twice differentiable w.r.t X_s and Y_s with $s \in (t, T]$.*

By Itô's lemma [10], we have

$$\frac{N_T^*}{N_T} - \frac{N_t^*}{N_t} = \int_t^T \{\frac{\partial f}{\partial X_s} dX_s + \frac{\partial f}{\partial Y_s} dY_s + O(ds)\}.$$

The CMS swap has a payoff which is linear in the swap rates, i.e. $\Phi(X_T) = X_T$ (we only consider this case in the present paper). Then, we have the payoff in stochastic integral form

$$\Phi(X_T) = X_T = X_t + \int_t^T dX_s, \tag{2.4}$$

and hence using (2.3) and (2.4) the convexity correction (2.2) simplifies:

$$\begin{aligned}
Cc_t &= \frac{N_t}{N_t^*}\mathrm{Cov}_t^Q\Big[\frac{N_T^*}{N_T}, \phi(X_T)\Big] = \frac{N_t}{N_t^*}E^Q\Big[\Big(\frac{N_T^*}{N_T} - \frac{N_t^*}{N_t}\Big)(X_T - X_t)\Big] \\
&= \frac{N_t}{N_t^*}E_t^Q\Big[\{\int_t^T \frac{\partial f}{\partial X_s} dX_s \cdot \int_t^T dX_s + \int_t^T \frac{\partial f}{\partial Y_s} dY_s \cdot \int_t^T dX_s + \int_t^T O(ds)ds \cdot \int_t^T dX_s\}\Big] \\
&\approx \frac{N_t}{N_t^*}\Big\{\frac{\partial f}{\partial X_t}E_t^Q[\int_t^T dX_s \cdot \int_t^T dX_s] + \frac{\partial f}{\partial Y_t}E_t^Q[\int_t^T dY_s \cdot \int_t^T dX_s] + \int_t^T O(ds)ds\, \underbrace{E_t^Q[\int_t^T dX_s]}_{=0}\Big\} \\
&= \frac{N_t}{N_t^*}\Big(\frac{\partial f}{\partial X_t}\mathrm{Var}_t^Q[X_T] + \frac{\partial f}{\partial Y_t}\mathrm{Cov}_t^Q[X_T, Y_T]\Big).
\end{aligned} \tag{2.5}$$

Usually the partial derivatives of the numeraire ratio are smooth and slow varying [8, 11]. A widely accepted approach is therefore to *freeze* them to their initial values, as demonstrated in Hull and White [8] and Jäckel and Rebonato [11]. This is also the approximation made in Eq. (2.5).

The form of Eq. (2.5) suggests that the convexity correction is driven by the terms $\mathrm{Var}_t^Q[X_T]$ and $\mathrm{Cov}_t^Q[X_T, Y_T]$. The covariance-based formulation naturally combines with multi-factor arbitrage-free interest rate modeling. This is a useful property of (2.5) because traders tend to have better formalized view about the correlation between two arbitrary interest rates than about the joint density of these two rates [21].

2.2. *Model Set-Up and Technical Issues*

The drawback of the covariance-based formulation (2.5) is of a technical nature, since the variance and covariance quantities are not easily computed, especially not when skew/smile features are taken into account. The main result of this paper is, therefore, an expansion formula for the covariance quantity in a multi-factor stochastic volatility model based on the Stochastic Alpha Beta Rho (SABR) model [5] under a reference measure, i.e. $\forall s \in (t, T]$

$$\begin{aligned} dX_s &= \qquad\qquad\quad \sigma_s (X_s)^{\beta_x} dW_s^{(1)} \\ dY_s &= \mu(s, Y_s) ds + \sigma_s (Y_s)^{\beta_y} dW_s^{(2)} \\ d\sigma_s &= \qquad\qquad\quad \eta \sigma_s dZ_s, \end{aligned} \tag{2.6}$$

with η the volatility-of-the-volatility. The three Brownian motions forms the following correlation matrix

$$\begin{bmatrix} 1 & \rho_{xy} & \rho_{xz} \\ \rho_{xy} & 1 & \rho_{yz} \\ \rho_{xz} & \rho_{yz} & 1 \end{bmatrix}$$

with $\langle dW_s^{(1)} dZ_s \rangle = \rho_{xz} ds$, $\langle dW_s^{(2)} dZ_s \rangle = \rho_{yz} ds$ and $\langle dW_s^{(1)} dW_s^{(2)} \rangle = \rho_{xy} ds$ ($\langle \cdot \rangle$ is a notation for inner products). Note that the model is defined under rate X_s's natural measure Q. Superscript Q is omitted for ease of presentation. In the rest part of the paper, the processes without specific superscripts are defined under the measure Q. Term $\mu(s, Y_s)$ is the arbitrage-free drift of the rate Y_s whose natural pricing measure is Q^*.

Practitioners often choose fewer volatility factors than the number of state variables, like a single volatility factor in, e.g., Piterbarg [20], Andersen and Andreasen [1]. In this paper, we also use a single volatility factor in (2.6). It serves as a first multi-factor academic model for the techniques proposed.

The convexity correction in a two-factor model has been derived for bi-variate log-normal models in [23]. In a bivariate log-normal distribution, the covariance can easily be computed by integrating over the *terminal* bi-variate distribution of the rate with respect to the *modified* [d] payoff function of the two rates involved, X_t and Y_t. However, when stochastic volatility is considered, the integration over the terminal bi-variate distribution does not result in the correct values, because the process $[X_t, Y_t]$ is not Markovian and the volatility realized along the path has a non-trivial impact on the convexity correction.

Even if we obtain the joint densities of the triplets $[X_t, Y_t, \sigma_t]$ correctly, there is no guarantee that we can directly integrate the terminal joint density over the payoff function when including the *arbitrage-free constraints*. In the next section, we therefore use a different method to approximate the covariance, based on the stochastic Taylor expansion.

[d] It is modified, because the relative numeraire is also included in the expected value of the payoff.

3. Stochastic Taylor Expansion to the Two-Factor SABR Model

Usually, multi-factor SABR prices are computed by a short time step Monte Carlo procedure, which is time consuming. In this section, we derive an approximation to the covariance quantity. The method used is based on the well-known Itô-Taylor expansion, described in Kloeden and Platen [15], and references therein. Instead of deriving the formula by directly applying Itô's lemma, we rely on the *small disturbance asymptotics* technique, described in Kunitomo and Takahashi [16], to construct the Taylor series of the processes X_t and Y_t. This technique has been applied to interest rate derivative pricing problems by Kawai [12,13] and Hagan [6]. Its theoretical validity was discussed in detail in [17].

3.1. *Stochastic Taylor Expansion to Asset Dynamics*

We express the solution in terms of *successive* terms with different orders of growth in time t. We first reformulate the system (2.6) by specifying a time rescaling $t = \epsilon^2\tau$, so that the processes $\epsilon W_\tau^{(\cdot)}$, ϵZ_τ have the same variances as $W_t^{(\cdot)}$ and Z_t, respectively.

$$\begin{aligned} dX_\tau^{(\epsilon)} &= && \epsilon\sigma_\tau^{(\epsilon)}(X_\tau^{(\epsilon)})^{\beta_x} dW_\tau^{(1)} \\ dY_\tau^{(\epsilon)} &= \epsilon^2\mu(\tau, Y_\tau^{(\epsilon)})d\tau + && \epsilon\sigma_\tau^{(\epsilon)}(Y_\tau^{(\epsilon)})^{\beta_y} dW_\tau^{(2)} \\ d\sigma_\tau^{(\epsilon)} &= && \epsilon\eta\sigma_\tau^{(\epsilon)} dZ_\tau \end{aligned} \tag{3.1}$$

The covariance of the time rescaling processes, $X_\tau^{(\epsilon)}$ and $Y_\tau^{(\epsilon)}$, does not change, i.e.

$$\langle dX_\tau^{(\epsilon)} dY_\tau^{(\epsilon)}\rangle \propto \epsilon^2\rho_{xy}d\tau = \rho_{xy}dt \propto \langle dX_t dY_t\rangle.$$

Since we do not know the distribution of system (3.1) explicitly, we consider the stochastic expansion around a deterministic process $[X_\tau^{(0)}, Y_\tau^{(0)}, \sigma_\tau^{(0)}]$ when the time rescaling parameter, ϵ, goes to zero, as required by the adopted small disturbance asymptotic technique. We substitute the time rescaling process in the definition of the covariance and truncate the solution up to the desired order of accuracy to obtain an approximation formula for the covariance.

Proposition 3.1. *The stochastic Taylor expansion of the volatility process, $\sigma_\tau^{(\epsilon)}$, up to fourth order, reads*

$$\sigma_\tau^{(\epsilon)} = \sigma_\tau^{(0)} + \epsilon\sigma_\tau^{(1)} + \epsilon^2\sigma_\tau^{(2)} + \epsilon^3\sigma_\tau^{(3)} + O(\epsilon^4),$$

where

$$\begin{aligned}
\sigma_\tau^{(0)} &= \sigma_0, \\
\sigma_\tau^{(1)} &:= \frac{\partial \sigma_\tau^{(\epsilon)}}{\partial \epsilon}\Big|_{\epsilon=0} = \int_0^\tau \eta \sigma_0 dZ_s, \\
\sigma_\tau^{(2)} &:= \frac{1}{2}\frac{\partial^2 \sigma_\tau^{(\epsilon)}}{\partial \epsilon^2}\Big|_{\epsilon=0} = \int_0^\tau \eta \int_0^{s_1} \eta \sigma_{s_2}^{(0)} dZ_{s_2} dZ_{s_1}, \\
\sigma_\tau^{(3)} &:= \frac{1}{6}\frac{\partial^3 \sigma_\tau^{(\epsilon)}}{\partial \epsilon^3}\Big|_{\epsilon=0} = \int_0^\tau \eta \int_0^{s_1} \eta \int_0^{s_2} \eta \sigma_{s_3}^{(0)} dZ_{s_3} dZ_{s_2} dZ_{s_1}.
\end{aligned}$$

Proof. This is a well-known result. The Taylor expansion of $\sigma_\tau^{(\epsilon)}$ around $\epsilon = 0$ gives

$$\sigma_\tau^{(\epsilon)} = \sigma_\tau^{(0)} + \epsilon \frac{\partial \sigma_\tau^{(\epsilon)}}{\partial \epsilon}\Big|_{\epsilon=0} + \frac{1}{2}\epsilon^2 \frac{\partial^2 \sigma_\tau^{(\epsilon)}}{\partial \epsilon^2}\Big|_{\epsilon=0} + \frac{1}{6}\epsilon^3 \frac{\partial^3 \sigma_\tau^{(\epsilon)}}{\partial \epsilon^3}\Big|_{\epsilon=0} + O(\epsilon^4).$$

From the volatility process in integral form,

$$\sigma_\tau^{(\epsilon)} = \sigma_0 + \epsilon \int_0^\tau \eta \sigma_s^{(\epsilon)} dZ_s,$$

one finds that $\sigma_\tau^{(0)} = \sigma_0$.

Then

$$\begin{aligned}
\frac{\partial \sigma_\tau^{(\epsilon)}}{\partial \epsilon}\Big|_{\epsilon=0} &= \Big[\int_0^\tau \eta \sigma_s^{(\epsilon)} dZ_s + \epsilon \int_0^\tau \eta \frac{\partial \sigma_s^{(\epsilon)}}{\partial \epsilon} dZ_s\Big]\Big|_{\epsilon=0} = \int_0^\tau \eta \sigma_0 dZ_s; \\
\frac{\partial^2 \sigma_\tau^{(\epsilon)}}{\partial \epsilon^2}\Big|_{\epsilon=0} &= \Big[2\int_0^\tau \eta \frac{\partial \sigma_s^{(\epsilon)}}{\partial \epsilon} dZ_s + \epsilon \int_0^\tau \eta \frac{\partial^2 \sigma_s^{(\epsilon)}}{\partial^2 \epsilon}\Big]\Big|_{\epsilon=0} \\
&= 2\int_0^\tau \eta \frac{\partial \sigma_s^{(0)}}{\partial \epsilon} dZ_s = 2\int_0^\tau \eta \int_0^{s_1} \eta \sigma_{s_2}^{(0)} dZ_{s_2} dZ_{s_1}.
\end{aligned}$$

Similarly, we find

$$\frac{\partial^3 \sigma_\tau^{(\epsilon)}}{\partial \epsilon^3}\Big|_{\epsilon=0} = 6\int_0^\tau \eta \int_0^{s_1} \eta \int_0^{s_2} \eta \sigma_{s_3}^{(0)} dZ_{s_3} dZ_{s_2} dZ_{s_1}. \qquad \square$$

Proposition 3.2. *The stochastic Taylor expansion of $X_\tau^{(\epsilon)}$, up to fourth order, can be expressed as follows:*

$$X_\tau^{(\epsilon)} = X_\tau^{(0)} + \epsilon X_\tau^{(1)} + \epsilon^2 X_\tau^{(2)} + \epsilon^3 X_\tau^{(3)} + O(\epsilon^4), \tag{3.2}$$

where

$$
\begin{aligned}
X_\tau^{(0)} =& X_0,\\
X_\tau^{(1)} :=& \frac{\partial X_\tau^{(\epsilon)}}{\partial \epsilon}\Big|_{\epsilon=0} = \int_0^\tau \sigma_0 (X_0)^{\beta_x} dW_s^{(1)},\\
X_\tau^{(2)} :=& \frac{1}{2}\frac{\partial^2 X_\tau^{(\epsilon)}}{\partial \epsilon^2}\Big|_{\epsilon=0} = \int_0^\tau \sigma_0 \int_0^{s_1} \sigma_0 \beta_x (X_0)^{\beta_x-1} dW_{s_2}^{(1)} dW_{s_1}^{(1)} + \int_0^\tau (X_0)^{\beta_x} \int_0^{s_1} \eta\sigma_0 dZ_{s_2} dW_{s_1}^{(1)},\\
X_\tau^{(3)} :=& \frac{1}{6}\frac{\partial^3 X_\tau^{(\epsilon)}}{\partial \epsilon^3}\Big|_{\epsilon=0} = \int_0^\tau \sigma_0 \int_0^{s_1} \sigma_0 \int_0^{s_2} \sigma_0 \beta_x(\beta_x-1)(X_0)^{\beta_x-2} dW_{s_3}^{(1)} dW_{s_2}^{(1)} dW_{s_1}^{(1)}\\
&+ \int_0^\tau \sigma_0 \int_0^{s_1} \beta_x (X_0)^{\beta_x-1} \int_0^{s_2} \eta\sigma_0 dZ_{s_3} dW_{s_2}^{(1)} dW_{s_1}^{(1)}\\
&+ \int_0^\tau (X_0)^{\beta_x} \int_0^{s_1} \eta \int_0^{s_2} \eta\sigma_0 dZ_{s_3} dZ_{s_2} dW_{s_1}^{(1)}\\
&+ \int_0^\tau \int_0^{s_1} \eta\big(\sigma_0\big)^2 \beta_x (X_0)^{\beta_x-1} Z_{s_1} W_{s_1}^{(1)} dW_{s_1}^{(1)}.
\end{aligned}
$$

Proof. We make a Taylor expansion of the process $X_\tau^{(\epsilon)}$ around $\epsilon = 0$:

$$
X_\tau^{(\epsilon)} = X_\tau^{(0)} + \epsilon \frac{\partial X_\tau^{(\epsilon)}}{\partial \epsilon}\Big|_{\epsilon=0} + \frac{1}{2}\epsilon^2 \frac{\partial^2 X_\tau^{(\epsilon)}}{\partial \epsilon^2}\Big|_{\epsilon=0} + \frac{1}{6}\epsilon^3 \frac{\partial^3 X_\tau^{(\epsilon)}}{\partial \epsilon^3}\Big|_{\epsilon=0} + O(\epsilon^4).
$$

It is again easy to see that $X_\tau^{(0)} = X_0$. Following the arguments in Kunitomo [16], we have

$$
\begin{aligned}
\frac{\partial X_\tau^{(\epsilon)}}{\partial \epsilon}\Big|_{\epsilon=0} =& \Big[\int_0^\tau \sigma_s^{(\epsilon)} (X_s^{(\epsilon)})^{\beta_x} dW_s^{(1)} + \epsilon \int_0^\tau \sigma_s^{(\epsilon)} \frac{\partial (X_s^{(\epsilon)})^{\beta_x}}{\partial \epsilon} dW_s^{(1)} + \epsilon \int_0^\tau (X_s^{(\epsilon)})^{\beta_x} \frac{\partial \sigma_s^{(\epsilon)}}{\partial \epsilon} dW_s^{(1)}\Big]\Big|_{\epsilon=0}\\
=& \int_0^\tau \sigma_s^{(0)} (X_s^{(0)})^{\beta_x} dW_s^{(1)} = \int_0^\tau \sigma_0 (X_0)^{\beta_x} dW_s^{(1)};\\
\frac{\partial^2 X_\tau^{(\epsilon)}}{\partial \epsilon^2}\Big|_{\epsilon=0} =& \Big[2\int_0^\tau \sigma_s^{(\epsilon)} \frac{\partial (X_s^{(\epsilon)})^{\beta_x}}{\partial \epsilon} dW_s^{(1)} + 2\int_0^\tau (X_s^{(\epsilon)})^{\beta_x} \frac{\partial \sigma_s^{(\epsilon)}}{\partial \epsilon} dW_s^{(1)} + \epsilon \int_0^\tau \sigma_s^{(\epsilon)} \frac{\partial^2 (X_s^{(\epsilon)})^{\beta_x}}{\partial \epsilon^2} dW_s^{(1)}\\
&+ \epsilon \int_0^\tau (X_s^{(\epsilon)})^{\beta_x} \frac{\partial^2 \sigma_s^{(\epsilon)}}{\partial \epsilon^2} dW_s^{(1)} + 2\epsilon \int_0^\tau \frac{\partial (X_s^{(\epsilon)})^{\beta_x}}{\partial \epsilon} \frac{\partial \sigma_s^{(\epsilon)}}{\partial \epsilon} dW_s^{(1)}\Big]\Big|_{\epsilon=0}\\
=& 2\int_0^\tau \sigma_0 \int_0^{s_1} \sigma_0 \beta_x (X_0)^{\beta_x-1} dW_{s_2}^{(1)} dW_{s_1}^{(1)} + 2\int_0^\tau (X_0)^{\beta_x} \int_0^{s_1} \eta\sigma_0 dZ_{s_2} dW_{s_1}^{(1)}.
\end{aligned}
$$

Recursive application of this scheme gives

$$
\begin{aligned}
\frac{\partial^3 X_\tau^{(\epsilon)}}{\partial \epsilon^3}\Big|_{\epsilon=0} =& \Big[3\int_0^\tau \sigma_s^{(\epsilon)} \frac{\partial^2 (X_s^{(\epsilon)})^{\beta_x}}{\partial \epsilon^2} dW_s^{(1)} + 3\int_0^\tau (X_s^{(\epsilon)})^{\beta_x} \frac{\partial^2 \sigma_s^{(\epsilon)}}{\partial \epsilon^2} dW_s^{(1)} \\
&+ 6\int_0^\tau \frac{\partial (X_s^{(\epsilon)})^{\beta_x}}{\partial \epsilon} \frac{\partial \sigma_s^{(\epsilon)}}{\partial \epsilon} dW_s^{(1)}\Big]\Big|_{\epsilon=0} \\
=& 3\int_0^\tau \sigma_s^{(0)} \frac{\partial^2 (X_s^{(0)})^{\beta_x}}{\partial \epsilon^2} dW_s^{(1)} + 3\int_0^\tau (X_s^{(0)})^{\beta_x} \frac{\partial^2 \sigma_s^{(0)}}{\partial \epsilon^2} dW_s^{(1)} \\
&+ 6\int_0^\tau \frac{\partial (X_s^{(0)})^{\beta_x}}{\partial \epsilon} \frac{\partial \sigma_s^{(0)}}{\partial \epsilon} dW_s^{(1)} \\
=& 6\int_0^\tau \sigma_0 \int_0^{s_1} \sigma_0 \int_0^{s_2} \sigma_0 \beta_x(\beta_x - 1)(X_0)^{\beta_x - 2} dW_{s_3}^{(1)} dW_{s_2}^{(1)} dW_{s_1}^{(1)} \\
&+ 6\int_0^\tau \sigma_0 \int_0^{s_1} \beta_x (X_0)^{\beta_x - 1} \int_0^{s_2} \eta\sigma_0 dZ_{s_3} dW_{s_2}^{(1)} dW_{s_1}^{(1)} \\
&+ 6\int_0^\tau (X_0)^{\beta_x} \int_0^{s_1} \eta \int_0^{s_2} \eta\sigma_0 dZ_{s_3} dZ_{s_2} dW_{s_1}^{(1)} \\
&+ 6\int_0^\tau \int_0^{s_1} \eta\big(\sigma_0\big)^2 \beta_x (X_0)^{\beta_x - 1} Z_{s_1} W_{s_1}^{(1)} dW_{s_1}^{(1)}. \qquad \square
\end{aligned}
$$

Proposition 3.3. *The stochastic Taylor expansion of $Y_\tau^{(\epsilon)}$, up to fourth order, gives:*

$$
Y_\tau^{(\epsilon)} = Y_\tau^{(0)} + \epsilon Y_\tau^{(1)} + \epsilon^2 Y_\tau^{(2)} + \epsilon^3 Y_\tau^{(3)} + O(\epsilon^4),
$$

where

$$
\begin{aligned}
Y_\tau^{(0)} =& Y_0, \\
Y_\tau^{(1)} :=& \frac{\partial Y_\tau^{(\epsilon)}}{\partial \epsilon}\Big|_{\epsilon=0} = \int_0^\tau \sigma_0 (Y_0)^{\beta_y} dW_s^{(2)}, \\
Y_\tau^{(2)} :=& \frac{1}{2}\frac{\partial^2 Y_\tau^{(\epsilon)}}{\partial \epsilon^2}\Big|_{\epsilon=0} = \int_0^\tau \mu(0, Y_0) d\tau + \int_0^\tau \sigma_0 \int_0^{s_1} \sigma_0 \beta_y (Y_0)^{\beta_y - 1} dW_{s_2}^{(2)} dW_{s_1}^{(2)} + \int_0^\tau (Y_0)^{\beta_y} \int_0^{s_1} \eta\sigma_0 dZ_{s_2} dW_{s_1}^{(2)}, \\
Y_\tau^{(3)} :=& \frac{1}{6}\frac{\partial^3 Y_\tau^{(\epsilon)}}{\partial \epsilon^3}\Big|_{\epsilon=0} = \int_0^\tau \sigma_0 \int_0^{s_1} \sigma_0 \int_0^{s_2} \sigma_0 \beta_y (\beta_y - 1)(Y_0)^{\beta_y - 2} dW_{s_3}^{(2)} dW_{s_2}^{(2)} dW_{s_1}^{(2)} \\
&+ \int_0^\tau \sigma_0 \int_0^{s_1} \beta_y (Y_0)^{\beta_y - 1} \int_0^{s_2} \eta\sigma_0 dZ_{s_3} dW_{s_2}^{(2)} dW_{s_1}^{(2)} \\
&+ \int_0^\tau (Y_0)^{\beta_y} \int_0^{s_1} \eta \int_0^{s_2} \eta\sigma_0 dZ_{s_3} dZ_{s_2} dW_{s_1}^{(2)} \\
&+ \int_0^\tau \int_0^{s_1} \eta\big(\sigma_0\big)^2 \beta_y (Y_0)^{\beta_y - 1} Z_{s_1} W_{s_1}^{(2)} dW_{s_1}^{(2)}.
\end{aligned}
$$

The proof is similar to that of Proposition 3.2.

3.2. *Expansion Solution to the Covariance*

We recall the definition of the covariance and substitute the expansions $X_\tau^{(\epsilon)}$ and $Y_\tau^{(\epsilon)}$ of X_τ and Y_τ, respectively. This way we obtain a stochastic Taylor expansion formula for the covariance, which is given by the following lemma.

Lemma 3.1. *For a multi-factor parametric stochastic volatility model of the form (3.1), the stochastic Taylor expansion of the covariance, $Cov_0[X_t^{(\epsilon)}, Y_t^{(\epsilon)}]$, is given by*

$$Cov_0[X_t^{(\epsilon)}, Y_t^{(\epsilon)}] = \nu^2 t + (\Lambda + \Gamma + \Sigma)\frac{t^2}{2} + O(t^6), \tag{3.3}$$

where

$$\begin{aligned}
\nu^2 =& \big(\sigma_0\big)^2 \big(X_0\big)^{\beta_x} \big(Y_0\big)^{\beta_y} \rho_{xy},\\
\Lambda =& \big(\sigma_0\big)^4 \beta_x \big(X_0\big)^{\beta_x - 1} \beta_y \big(Y_0\big)^{\beta_y - 1} (\rho_{xy})^2,\\
\Gamma =& \eta\big(\sigma_0\big)^3 \beta_x \big(X_0\big)^{\beta_x-1} \big(Y_0\big)^{\beta_y} \rho_{xy}\rho_{xz} + \eta\big(\sigma_0\big)^3 \big(X_0\big)^{\beta_x} \beta_y \big(Y_0\big)^{\beta_y-1} \rho_{xy}\rho_{yz}\\
&+ \eta\big(\sigma_0\big)^3 (X_0)^{\beta_x} \beta_y (Y_0)^{\beta_y-1} \rho_{yz}\rho_{xy} + \eta\big(\sigma_0\big)^3 \beta_x (X_0)^{\beta_x-1} (Y_0)^{\beta_y} \rho_{xz}\rho_{xy},\\
\Sigma =& \eta^2 \big(\sigma_0\big)^2 \big(X_0\big)^{\beta_x} \big(Y_0\big)^{\beta_y} \rho_{xy}.
\end{aligned}$$

Proof. To facilitate the proof, we firstly recall the product formula for two Itô integrals:

$$E\big[\int_0^t f dW_s^{(1)} \cdot \int_0^t g dW_s^{(2)}\big] = \int_0^t (f \cdot g)\rho ds, \tag{3.4}$$

with $\langle W_s^{(1)} W_s^{(2)} \rangle = \rho dt$.

Since the process X_τ is a martingale, we have $E_0[X_\tau] = X_0$. For process Y_τ we have

$$E_0[Y_\tau] = Y_0 + \int_0^\tau \mu(s_1, Y_{s_1}) ds_1.$$

We substitute these two expectations and the expansions for X_τ and Y_τ in the definition of the covariance:

$$\begin{aligned}
\mathrm{Cov}_0^{(\epsilon)}[X_\tau^{(\epsilon)}, Y_\tau^{(\epsilon)}] =& E\big[\big(X_\tau^{(\epsilon)} - E_0[X_\tau^{(\epsilon)}]\big)\big(Y_\tau^{(\epsilon)} - E_0[Y_\tau^{(\epsilon)}]\big)\big]\\
\approx& E\big[\big(X_\tau^{(0)} + \epsilon X_\tau^{(1)} + \epsilon^2 X_\tau^{(2)} + \epsilon^3 X_\tau^{(3)} - X_0\big)\cdot\\
&\big(Y_\tau^{(0)} + \epsilon Y_\tau^{(1)} + \epsilon^2 Y_\tau^{(2)} + \epsilon Y_\tau^{(3)} - Y_0 - \epsilon^2 \int_0^t \mu(s_1, Y_{s_1}^{(\epsilon)}) ds_1\big)\big]\\
=& E\big[\big(\epsilon^2 X_\tau^{(1)} Y_\tau^{(1)} + \epsilon^3 X_\tau^{(1)} Y_\tau^{(2*)} + \epsilon^4 X_\tau^{(1)} Y_\tau^{(3)} + \epsilon^3 X_\tau^{(2)} Y_\tau^{(1)} +\\
&\epsilon^4 X_\tau^{(2)} Y_\tau^{(2*)} + \epsilon^4 X_\tau^{(3)} Y_\tau^{(1)} + O(\epsilon^5)\big)\big],
\end{aligned} \tag{3.5}$$

where we have eliminated the drift term from $Y_\tau^{(2)}$ and defined

$$Y_\tau^{(2*)} := \int_0^\tau \sigma_0 \int_0^{s_1} \sigma_0 (Y_0)^{\beta_y} dW_{s_2}^{(2)} dW_{s_1}^{(2)} + \int_0^\tau (Y_0)^{\beta_y} \int_0^{s_1} \eta\sigma_0 dZ_{s_2} dW_{s_1}^{(2)}.$$

By Itô product formula (3.4), we find that

$$X_\tau^{(1)}Y_\tau^{(1)} = \int_0^\tau (\sigma_0)^2 (X_0)^{\beta_x} (Y_0)^{\beta_y} \rho_{xy} ds = (\sigma_0)^2 (X_0)^{\beta_x} (Y_0)^{\beta_y} \rho_{xy}\tau, \tag{3.6}$$

$$\begin{aligned} X_\tau^{(1)}Y_\tau^{(2*)} &= \int_0^\tau (\sigma_0)^2 (X_0)^{\beta_x} \int_0^{s_1} \sigma_0 (Y_0)^{\beta_y} dW_{s_2}^{(2)} \rho_{xy} ds_1 + \int_0^\tau \sigma_0 (X_0)^{2\beta_x} (Y_0)^{\beta_y} \int_0^{s_1} \eta\sigma_0 dZ_2 \rho_{xy} ds_1 \\ &= (\sigma_0)^3 (X_0)^{\beta_x} (Y_0)^{\beta_y} \rho_{xy} \int_0^\tau \int_0^{s_1} dW_{s_2}^{(2)} ds_1 + \eta(\sigma_0)^2 (X_0)^{2\beta_x} (Y_0)^{\beta_y} \rho_{xy} \int_0^\tau \int_0^{s_1} dZ_{s_2} ds_1. \end{aligned} \tag{3.7}$$

It is not difficult to see that the two terms in (3.7) are Itô integrals with mean zero. After taking the expectation, the term $X_\tau^{(1)}Y_\tau^{(2*)}$ disappears, i.e., $E_0[X_\tau^{(1)}Y_\tau^{(2*)}] = 0$. The same holds for all terms of odd order, e.g. for $\epsilon^3, \epsilon^5, \ldots$. Hence we only have to consider the even terms in the expansion of the covariance, for example:

$$\begin{aligned} X_\tau^{(1)}Y_\tau^{(3)} &= \int_0^\tau (\sigma_0)^2 (X_0)^{\beta_x} \int_0^{s_1} \sigma_0 \int_0^{s_2} \sigma_0 (Y_0)^{\beta_y} dW_{s_3}^{(2)} dW_{s_2}^{(2)} \rho_{xy} ds_1 \\ &+ \int_0^\tau (\sigma_0)^2 (X_0)^{\beta_x} \int_0^{s_1} (Y_0)^{\beta_y} \int_0^{s_2} \eta\sigma_0 dZ_{s_3} dW_{s_2}^{(2)} \rho_{xy} ds_1 \\ &+ \int_0^\tau \sigma_0 (X_0)^{\beta_x} (Y_0)^{\beta_y} \int_0^{s_1} \eta \int_0^{s_2} \eta\sigma_0 dZ_{s_3} dZ_{s_2} \rho_{xy} ds_1 \\ &+ \int_0^\tau \sigma_0 (X_0)^{\beta_x} \int_0^{s_1} \eta (\sigma_0)^2 (Y_0)^{\beta_y} \rho_{yz} ds_2 \rho_{xy} ds_1 \\ &= (\sigma_0)^4 (X_0)^{\beta_x} (Y_0)^{\beta_y} \rho_{xy} \int_0^\tau \int_0^{s_1} \int_0^{s_2} dW_{s_3}^{(2)} dW_{s_2}^{(2)} ds_1 \\ &+ (\sigma_0)^3 (X_0)^{\beta_x} (Y_0)^{\beta_y} \eta\rho_{xy} \int_0^\tau \int_0^{s_1} \int_0^{s_2} dZ_{s_3} dW_{s_2}^{(2)} ds_1 \\ &+ (\sigma_0)^2 (X_0)^{\beta_x} (Y_0)^{\beta_y} \eta^2 \rho_{xy} \int_0^\tau \int_0^{s_1} \int_0^{s_2} dZ_{s_3} dZ_{s_2} ds_1 \\ &+ \eta(\sigma_0)^3 (X_0)^{\beta_x} (Y_0)^{\beta_y} \rho_{xy}\rho_{yz} \frac{\tau^2}{2}. \end{aligned} \tag{3.8}$$

The computation of the expectation of $X_\tau^{(1)}Y_\tau^{(3)}$ requires that we deal with three double Itô integrals, i.e. $E\Big[\int_0^{s_1}\int_0^{s_2} dW_{s_3}^{(2)} dW_{s_2}^{(2)}\Big]$, $E\Big[\int_0^{s_1}\int_0^{s_2} dZ_{s_3} dW_{s_2}^{(2)}\Big]$ and $E\Big[\int_0^{s_1}\int_0^{s_2} dZ_{s_3} dZ_{s_2}\Big]$. For the first integral in (3.8), we can compute its expectation by applying Itô's lemma:

$$W_{s_2}^{(2)} dW_{s_2}^{(2)} = \frac{1}{2} d\big(W_{s_2}^{(2)}\big)^2 + \frac{1}{2} ds_2, \text{ or,}$$

$$\int_0^{s_1} W_{s_2}^{(2)} dW_{s_2}^{(2)} = \frac{1}{2}\int_0^{s_1} d\big(W_{s_2}^{(2)}\big)^2 + \int_0^{s_1} \frac{1}{2} ds_2 = \frac{1}{2}\big(W_{s_1}^{(2)}\big)^2 + \frac{1}{2}s_1$$

$$\Rightarrow E\Big[\int_0^{s_1} \underbrace{\int_0^{s_2} dW_{s_3}^{(2)}}_{=W_{s_2}^{(2)} - W_0^{(2)}} dW_{s_2}^{(2)}\Big] = E\big[\frac{1}{2}\big(W_{s_1}^{(2)}\big)^2 + \frac{1}{2}s_1\big] = 0.$$

Similarly, we find that $E\Big[\int_0^{s_1}\int_0^{s_2} dZ_{s_3}dZ_{s_2}\Big] = 0$ and $E\big[\int_0^{s_1}\int_0^{s_2} dZ_{s_3}dW_{s_2}^{(2)}\big] = 0$.

The computations of $E\big[X_\tau^{(2)}Y_\tau^{(2)}\big]$ and $E\big[X_\tau^{(3)}Y_\tau^{(1)}\big]$ are performed in the same manner. Finally, we substitute the result in Eq. (3.5) and obtain

$$\begin{aligned}
\mathrm{Cov}_0[X_\tau, Y_\tau] =& \epsilon^2\big(\sigma_0\big)^2\tau + \epsilon^4(\big(\sigma_0\big)^4\big(X_0\big)^{\beta_x}\big(Y_0\big)^{\beta_y}(\rho_{xy})^2\frac{\tau^2}{2} \\
&+ 2\epsilon^4\eta\big(\sigma_0\big)^3\big(X_0\big)^{\beta_x}\big(Y_0\big)^{\beta_y}\rho_{xy}(\rho_{xz}+\rho_{yz})\frac{\tau^2}{2} + \epsilon^4\eta^2\big(\sigma_0\big)^2\big(X_0\big)^{\beta_x}\big(Y_0\big)^{\beta_y}\rho_{xy}\frac{\tau^2}{2} + O(\epsilon^6) \\
=& \big(\sigma_0\big)^2 t + (\big(\sigma_0\big)^4\big(X_0\big)^{\beta_x}\big(Y_0\big)^{\beta_y}(\rho_{xy})^2\frac{t^2}{2} \\
&+ 2\eta\big(\sigma_0\big)^3\big(X_0\big)^{\beta_x}\big(Y_0\big)^{\beta_y}\rho_{xy}(\rho_{xz}+\rho_{yz})\frac{t^2}{2} + \eta^2\big(\sigma_0\big)^2\big(X_0\big)^{\beta_x}\big(Y_0\big)^{\beta_y}\rho_{xy}\frac{t^2}{2} + O(t^6),
\end{aligned}$$

where the last equality is a consequence of the time rescaling we defined earlier, i.e. $\epsilon^2\tau = t$. As stated earlier, the rescaled system $[X_t^{(\epsilon)}, X_t^{(\epsilon)}, \sigma_t^{(\epsilon)}]$ preserves the variance and covariance. Therefore, we obtain the desired approximation for the covariance in the original time scale. □

The terms in Eq. (3.3) can be interpreted as follows:

- The first term, ν^2, is the leading term of the covariance which grows linearly with time. It is also the solution we would obtain by assuming a constant volatility;
- Correction term Λ is due to the first order sensitivity of the covariance w.r.t to the forward rate dynamics;
- Term Σ quantifies the impact of stochastic volatility. It is positive, hence it adds a positive contribution to the covariance;
- Finally, term Γ is related to the interaction between the forward rate and the volatility dynamics [6] and hence it can be of either sign, depending on the correlation parameters, ρ_{xz}, ρ_{yz}, in the model.

4. Examples

In this section, we present some examples of the expansion formula for the covariance.

4.1. *Two-Factor Log-Normal Model*

We begin by analyzing the accuracy of the stochastic Taylor expansion formula for the two-factor log-normal model. This is a special case of the two-factor SABR model, with $\beta_x = \beta_y = 1$ and volatility-of-volatility parameter, η, equal to zero, i.e.

$$\begin{aligned}
dX_t &= \sigma_x X_t dW_t^{(1)}, \\
dY_t &= \sigma_y Y_t dW_t^{(2)}.
\end{aligned} \tag{4.1}$$

Since an analytic solution for the covariance is available for this model, the stochastic Taylor expansion solution (4.5) is compared to the exact solution. It is shown that the stochastic Taylor expansion of the covariance between two log-normally distributed variables agrees well with the Taylor expansion of the exact solution of the same quantity.

We reformulate system (4.1) by making the time rescaling $t = \epsilon^2\tau$, so that the processes $\epsilon W_\tau^{(1)}$, $\epsilon W_\tau^{(2)}$ have the same variances as $W_t^{(1)}$ and $W_t^{(2)}$, respectively,

$$\begin{aligned} dX_\tau^{(\epsilon)} &= \epsilon\sigma_x X_\tau dW_\tau^{(1)}, \\ dY_\tau^{(\epsilon)} &= \epsilon\sigma_y Y_\tau dW_\tau^{(2)}, \end{aligned} \tag{4.2}$$

with $E[dW_\tau^{(1)} dW_\tau^{(2)}] = \rho$.

The asymptotic expansion of $X_\tau^{(\epsilon)}$, up to fourth order, reads:

$$\begin{aligned} X_\tau^{(\epsilon)} = X_0 + \epsilon\sigma_x X_0 \int_0^\tau dW_s^{(1)} + \frac{1}{2}\epsilon^2 2\sigma_x^2 X_0 \int_0^\tau \int_0^{s_1} dW_{s_2}^{(1)} dW_{s_1}^{(1)}, \\ + \frac{1}{6}\epsilon^3 6\sigma_x^3 X_0 \int_0^\tau \int_0^{s_1} \int_0^{s_2} dW_{s_3}^{(1)} dW_{s_2}^{(1)} dW_{s_1}^{(1)} + O(\epsilon^4). \end{aligned} \tag{4.3}$$

Similarly, we have

$$Y_\tau^{(\epsilon)} = Y_0 + \epsilon\sigma_y Y_0 \int_0^\tau dW_s^{(1)} + \epsilon^2\sigma_y^2 Y_0 \int_0^\tau \int_0^{s_1} dW_{s_2}^{(1)} dW_{s_1}^{(1)} + \epsilon^3\sigma_y^3 Y_0 \int_0^\tau \int_0^{s_1} \int_0^{s_2} dW_{s_3}^{(1)} dW_{s_2}^{(1)} dW_{s_1}^{(1)} + O(\epsilon^4). \tag{4.4}$$

Substituting Eqs. (4.3) and (4.4) in the definition of the covariance, we find

$$\begin{aligned} \mathrm{Cov}_0\big[X_\tau^{(\epsilon)}, Y_\tau^{(\epsilon)}\big] =& E\big[\big(X_\tau^{(\epsilon)} - E[X_\tau^{(\epsilon)}]\big)\big(Y_\tau^{(\epsilon)} - E[Y_\tau^{(\epsilon)}]\big)\big] \\ =& E\Big[\Big(\epsilon^2 X_\tau^{(1)} Y_\tau^{(1)} + \frac{1}{4}\epsilon^4 X_\tau^{(2)} Y_\tau^{(2)} + \frac{1}{6}\epsilon^4 X_\tau^{(1)} Y_\tau^{(3)} + \frac{1}{6}\epsilon^4 X_\tau^{(3)} Y_\tau^{(1)}\Big)\Big] + O(\epsilon^6), \end{aligned}$$

where

$$\begin{aligned} X_\tau^{(1)} Y_\tau^{(1)} &= \sigma_x\sigma_y X_0 Y_0 \rho\tau, \\ X_\tau^{(2)} Y_\tau^{(2)} &= 4\sigma_x^2\sigma_y^2 X_0 Y_0 \rho^2 \frac{\tau^2}{2}, \\ X_\tau^{(1)} Y_\tau^{(3)} &= 6\sigma_x\sigma_y^3 X_0 Y_0 \rho\tau \int_0^{s_1} \int_0^{s_2} dW_{s_3}^{(2)} dW_{s_2}^{(2)}, \\ X_\tau^{(3)} Y_\tau^{(1)} &= 6\sigma_y\sigma_x^3 X_0 Y_0 \rho\tau \int_0^{s_1} \int_0^{s_2} dW_{s_3}^{(1)} dW_{s_2}^{(1)}. \end{aligned}$$

Since $\int_0^{s_1}\int_0^{s_2} dW_{s_3}^{(1)} dW_{s_2}^{(1)}$ and $\int_0^{s_1}\int_0^{s_2} dW_{s_3}^{(2)} dW_{s_2}^{(2)}$ have zero expectation, we find

$$\begin{aligned} \mathrm{Cov}_0\big[X_\tau^{(\epsilon)}, Y_\tau^{(\epsilon)}\big] &= \frac{1}{2}\epsilon^2 X_\tau^{(1)} Y_\tau^{(1)} + \frac{1}{4}\epsilon^4 X_\tau^{(2)} Y_\tau^{(2)} \\ &= X_0 Y_0 \Big(\epsilon^2\sigma_x\sigma_y\rho\tau + \epsilon^4\sigma_x^2\sigma_y^2\rho^2\frac{\tau^2}{2}\Big) \\ &= X_0 Y_0 \Big(\sigma_x\sigma_y\rho t + \sigma_x^2\sigma_y^2\rho^2\frac{t^2}{2}\Big). \end{aligned} \tag{4.5}$$

The first term in (4.5) is the Gaussian approximation which grows linearly with time t. The second term acts as the convexity correction and accounts for the non-Gaussian part of the distribution.

Due to the tractability of log-normally distributed random variables, their covariance can be computed directly by:

$$
\begin{aligned}
\mathrm{Cov}_0\big[X_t, Y_t\big] =& E\big[(X_t - E[X_t])(Y_t - E[Y_t])\big] = E\big[X_t Y_t\big] - E[X_t]E[Y_t] \\
=& X_0 Y_0 e^{-\frac{1}{2}\sigma_x^2 t - \frac{1}{2}\sigma_y^2 t} E\big[e^{\sigma_x Z_t^{(1)} + \sigma_y(\rho Z_t^{(1)} + \sqrt{1-\rho^2}) Z_t^{(2)}}\big] - X_0 Y_0 \\
=& X_0 Y_0 e^{-\frac{1}{2}\sigma_x^2 t - \frac{1}{2}\sigma_y^2 t} E\big[e^{(\sigma_x + \sigma_y \rho) Z_t^{(1)}} e^{\sigma_y \sqrt{1-\rho^2} Z_t^{(2)}}\big] - X_0 Y_0 \\
=& X_0 Y_0 e^{-\frac{1}{2}\sigma_x^2 t - \frac{1}{2}\sigma_y^2 t} e^{\frac{1}{2}(\sigma_x + \sigma_y \rho)^2 t + \frac{1}{2}\left(\sigma_y \sqrt{1-\rho^2}\right)^2 t} - X_0 Y_0 \\
=& X_0 Y_0 e^{\rho \sigma_x \sigma_y t} - X_0 Y_0,
\end{aligned}
\tag{4.6}
$$

since $E\Big[e^{(\sigma_x+\sigma_y\rho)Z_t^{(1)} - \frac{1}{2}(\sigma_x+\sigma_y\rho)^2 t} e^{\sigma_y\sqrt{1-\rho^2}Z_t^{(2)} - \frac{1}{2}\left(\sigma_y\sqrt{1-\rho^2}\right)^2 t}\Big] = 1$, and $Z_t^{(1)}$ and $Z_t^{(2)}$ are independent Brownian motions.

The expansion solution (4.5) for the variables X_t and Y_t in (4.1) agrees with the first two terms of the Taylor expansion of the solution in Eq. (4.6). We denote the term $\rho\sigma_x\sigma_y t$ in (4.5) by "Expn. O1" (the first order term), the term $\frac{1}{2}\big(\rho\sigma_x\sigma_y t\big)^2$ in (4.5) by "Expn. O2" (the second order term), etc.

4.1.1. *Numerical Experiment*

In this section, we evaluate numerically the accuracy of the expansion solution (4.5) for two different sets of parameters, given in Table 1. The correlation between the two Brownian motions is set to $\rho_{xy} = 0.6$, for both experiments.

	X_0	σ_x	Y_0	σ_y
High vol.	1	40%	1	45%
Low vol.	1	20%	1	25%

Table 1. Two parameter sets for the evaluation of the expansion in the log-normal case.

The results of the Taylor approximation are presented in Table 2. The expansion is accurate, especially when the volatilities are small. In order to obtain a satisfactory accuracy in the case of high volatility, one has to expand up to terms of higher order. In the interest rate derivative pricing problems, however, the value of an underlying is typically only a few percentage points. Figure 1 displays the accuracy results graphically. The results of the left-hand side in Table 2 suggest that the expansion up to order fourth order is sufficient even for long time to maturity.

	Expn. O2	Expn. O3	Expn. O4	Exact	Expn. O2	Expn. O3	Expn. O4	Exact
	Low volatility				High volatility			
2 yr	0.0618	0.0618	0.0618	0.0618	0.2393	0.2410	0.2411	0.2411
5 yr	0.1613	0.1618	0.1618	0.1618	0.6858	0.7120	0.7156	0.7160
10 yr	0.3450	0.3495	0.3498	0.3499	1.6632	1.8732	1.9298	1.9447
15 yr	0.5513	0.5664	0.5681	0.5683	2.9322	3.6408	3.9278	4.0531
20 yr	0.7800	0.8160	0.8214	0.8221	4.4928	6.1724	7.0794	7.6711
25 yr	1.0313	1.1016	1.1147	1.1170	6.3450	9.6255	11.8398	13.8797
30 yr	1.3050	1.4265	1.4538	1.4596	8.4888	14.1575	18.7492	24.5337

(Unit: percentage points)

Table 2. Comparison of the accuracy of the expansion for the covariance against the exact covariance, for the log-normal case.

4.2. *Constant Maturity Swap*

We will now analyze the approximation for the CMS contract, which we first describe in some more detail.

Recall that the computation of the CMS convexity correction is reduced to the approximation:

$$Cc_t \approx \frac{N_t}{N_t^*}\Big(\frac{\partial f}{\partial X_t}\mathrm{Var}_t^Q[X_T] + \frac{\partial f}{\partial Y_t}\mathrm{Cov}_t^Q[X_T, Y_T]\Big), \tag{4.7}$$

where the numeraire ratio is a function of the rates X_s and Y_s, i.e., $\frac{N_s^*}{N_s} = f(X_s, Y_s)$, $s \in (t, T]$.

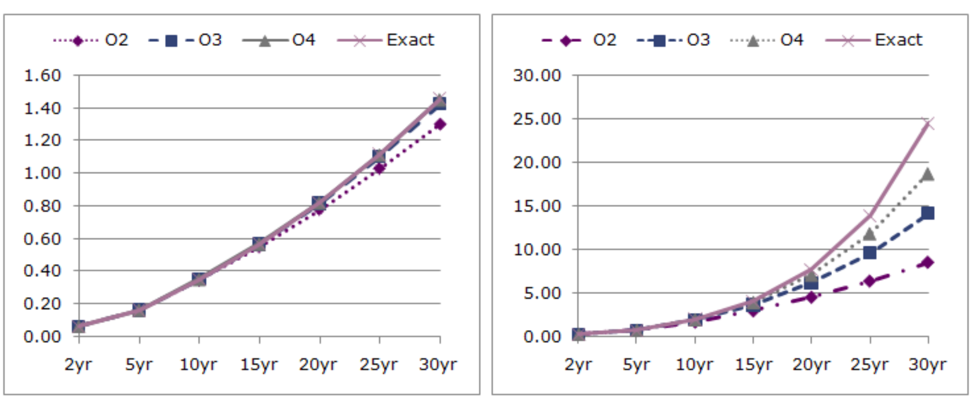


Fig. 1. Accuracy of the expansion solution (4.5) for two-factor log-normal model. Left-hand side: low volatility, i.e. $\sigma_x = 20\%$ and $\sigma_y = 45\%$; Right-hand side: high volatility with $\sigma_x = 40\%$ and $\sigma_y = 45\%$.

The CMS pricing formula reads:

$$\mathrm{CMS}(t) = P(t, T_{\mathrm{pay}}) E^{Q^{T_{\mathrm{pay}}}} [\mathrm{SR}(T_0)] = A(t) E^{Q^A} \left[\frac{\mathrm{SR}(T_0) P(T_0, T_{\mathrm{pay}})}{A(T_0)} \Big| \mathcal{F}_t \right], \quad (4.8)$$

where

- t denotes the current time point,
- T_0 is the starting (or expiry) time of the CMS contract,
- T_{pay} is the delayed payment time of the CMS contract, i.e. $T_{\mathrm{pay}} = T_0 + \tau$ where τ is the time fraction of the payment delay,
- $A(T_0)$ is the T_0-value of the annuity of the reference swap $\mathrm{SR}(T_0)$, i.e. $A(T_0) = \sum_{i=1}^{m} \delta_i P(T_0, T_i)$ with accrual factors δ_i,
- T_i $(i = 1, 2, 3, \ldots, m)$ represents a series of m resetting dates for the underlying reference swap,
- $\mathrm{SR}(T_0)$ stands for the T_0-value of a swap starting from T_0 with maturity T_m, i.e. $\mathrm{SR}(T_0) = \mathrm{SR}(T_0, T_0, T_m)$,

Two measures are involved in the CMS pricing problem:

- The *T-forward measure*, which is denoted by $Q^{T_{\mathrm{pay}}}$, is associated to zero coupon bonds with some maturity T;
- The *annuity measure*, denoted by Q^A, is the natural martingale measure for (forward starting) swaps and swaptions. The annuity pays 1 Euro at each coupon day of the swap, accrued according to the swap day count conventions.

Note that the swap rate, $\mathrm{SR}(t)$, corresponds to the rate X_t in Eq. (4.7). There is no drift term under the annuity measure. Since the LIBOR rate on the payment date, $L(t, T_{\mathrm{pay}})$, corresponds to the rate Y_t in our problem formulation, it is, in general, not a martingale process under the swap measure.

The variance/covariance quantity in (4.7) can be approximated by the Taylor expansion formula once the parameters are calibrated. The numeraire ratio is problem-specific and the partial derivatives in (4.7) have to be determined according to the payment features. The numeraire ratio is here given by: $N_t/N_t^* = P(t, T_{\mathrm{pay}})/A(t)$.

$P(t, T_{\mathrm{pay}})$ is driven by a LIBOR rate, so that:

$$P(t, T_{\mathrm{pay}}) = P(t, T_0) \frac{1}{1 + \tau L(t, T_{\mathrm{pay}})}.$$

The swap annuity is defined by $A(t) := \sum_{i=1}^{m} \delta_i P(t, T_i)$. This expression is approximated by the following relation

$$P(t, T_i) \approx P(t, T_0) \prod_{j=1}^{i} \frac{1}{1 + \delta_j \mathrm{SR}(t)}, \qquad i = 1, \ldots, m.$$

Then, the annuity reads

$$A(t) = \sum_{i=1}^{m} \delta_i P(t, T_i) \approx P(t, T_0) \sum_{i=1}^{m} \Big(\delta_i \prod_{j=1}^{i} \frac{1}{1 + \delta_j \mathrm{SR}(t)} \Big).$$

So, the numeraire ratio considered in the CMS pricing problem reads

$$\begin{aligned} \frac{N_t^*}{N_t} =& f\Big(\mathrm{SR}(t), L(t, T_{\mathrm{pay}})\Big) \\ =& \frac{P(t, T_{\mathrm{pay}})}{A(t)} = \frac{1}{\big(1 + \tau L(t, T_{\mathrm{pay}})\big) \sum_{i=1}^{m} \Big(\delta_i \prod_{j=1}^{i} \frac{1}{1+\delta_j \mathrm{SR}(t)} \Big)}. \end{aligned} \tag{4.9}$$

The partial derivatives to the numeraire ratio (4.9) w.r.t the swap rate and the LIBOR rate then read:

$$\frac{\partial f}{\partial \mathrm{SR}(t)} = \sum_{i=1}^{m} \delta_i \sum_{j=1}^{i} \frac{\delta_j}{1 + \delta_j \mathrm{SR}(t)} \cdot \frac{1}{\prod_{j=1}^{i} \big(1 + \delta_j \mathrm{SR}(t)\big)} \cdot \Big(\frac{1}{\sum_{i=1}^{m} \delta_i \prod_{j=1}^{i} \frac{1}{1+\delta_j \mathrm{SR}(t)}} \Big)^2, \tag{4.10}$$

$$\frac{\partial f}{\partial L(t, T_{\mathrm{pay}})} = - \frac{\tau}{\big(1 + \tau L(t, T_{\mathrm{pay}})\big)^2} \frac{1}{\sum_{i=1}^{m} \Big(\delta_i \prod_{j=1}^{i} \frac{1}{1+\delta_j \mathrm{SR}(t)} \Big)}. \tag{4.11}$$

4.2.1. *Numerical Experiment for CMS with SABR Model*

Here, we compare numerically the accuracy of the approximation for the CMS convexity correction with a reference solution, generated by the Monte Carlo method, and also with other approximations available in the literature. We are also interested in the price impact of the factor decorrelation. The SABR model is popular among practitioners and often used as an "implied volatility interpolation tool" for swaptions and caplets.

The CMS contract priced here pays a 10 years, annually reset, par swap rate with a 6 month payment delay.

For the multi-factor SABR model described in Eq. (2.6), the parameters chosen for the processes X_t and Y_t are given in Table 3.

X_0	vol-of-vol	Corr. (ρ_{xz})	skew (β)
3.4%	0.2	-0.4	1 and 0.6
Y_0	vol-of-vol	Corr. (ρ_{xz})	skew (β)
3.0%	0.2	-0.5	1 and 0.6

Table 3. Parameters for the CMS experiments.

In the Monte Carlo method for the benchmark prices for the convexity correction, we choose the Euler time discretization scheme and use a grid of 10 steps per year.

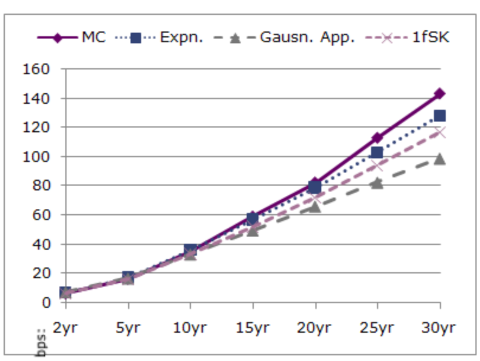
The following methods are compared in this section:

(1) *MC* represents the short time step Monte Carlo method for the two-factor SABR model;
(2) *Expn.* is the expansion solution derived in this paper. The features captured by this method are the skew/smile surface and the factor decorrelation;
(3) *Gausn. App.* denotes the Gaussian approximation method, obtained by assuming that the underlying diffusion processes are Gaussian; This approximation can model a terminal decorrelation but not a smile or skew.
(4) *1fSK* is a one-factor model with skew/smile features. More specifically, we consider Mercurio's method [18] in this experiment. The skew/smile is captured, but the terminal decorrelation is not modeled.

In order to investigate the price impact of factor decorrelation, we set up two experiments with different correlations between the swap rate and the LIBOR rate, ρ_{xy}: One with a *positive* correlation, $\rho_{xy} = 0.3$, and another in which a *negative* correlation is chosen, $\rho_{xy} = -0.3$. The numerical results obtained are summarized in Figure 2 and Table 4 for $\beta = 1$, and in Table 5 for $\beta = 0.6$.

The expansion solutions derived in this paper, Expn, agree well with the Monte Carlo benchmark prices for these two experiments, see Tables 4 and 5. The one-factor-with-skew model, 1fSK, returns, by construction, the same price in the two experiments, which is an obvious drawback of that model. For short expiry times, e.g. smaller than 10 years, all four methods provide more-or-less the same level of accuracy for $\beta = 1$. This is different for $\beta = 0.6$, for which only approximation "Expn" agrees well with the benchmark prices. For longer expiry times, the differences between the Monte Carlo prices and the one-factor, as well as the Gaussian, approximation increase. Our approximation, Expn, resembles the benchmark prices rather well, even up to expiration times of 30 years.

When $\beta \approx 1$, the skew/smile feature has a more significant impact than the



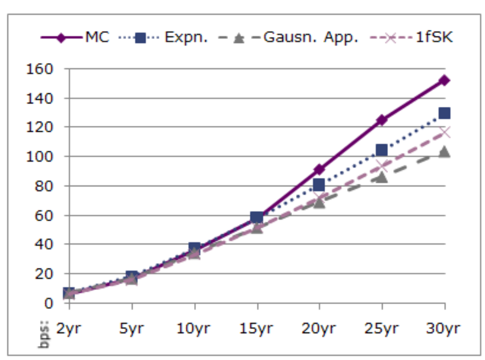


Fig. 2. Comparison of several convexity correction methods with $\beta_x = \beta_y = 1$ and two different swap-LIBOR correlations. Left-hand side: the convexity correction in time for $\rho_{xy} > 0$. Right-hand side: convexity correction for $\rho_{xy} < 0$.

	MC	*Expn.*	Gausn. App.	*1fSk*	*MC*	*Expn.*	Gausn. App.	*1fSk*
	Positive Correlation $\rho_{xy} = 0.3$				Negative Correlation $\rho_{xy} = -0.3$			
5 yr	16.75	17.30	16.49	15.98	17.33	18.07	17.36	15.98
10 yr	35.36	36.23	32.98	33.36	37.54	37.05	34.73	33.36
15 yr	59.18	56.78	49.47	52.15	59.97	58.41	52.09	52.15
20 yr	82.38	78.95	65.96	72.34	82.73	80.70	69.45	72.34
25 yr	102.88	102.75	82.45	93.94	122.24	104.38	86.81	93.94
30 yr	140.71	128.17	98.94	116.94	152.49	129.48	104.18	116.94

(Unit: basis points)

Table 4. The CMS convexity corrections with $\beta_x = \beta_y = 1$ and two different swap-LIBOR correlations. Left-hand side: $\rho_{xy} > 0$; Right-hand side: $\rho_{xy} < 0$

	MC	*Expn.*	Gausn. App.	*1fSk*	*MC*	*Expn.*	Gausn. App.	*1fSk*
	Positive Correlation $\rho_{xy} = 0.3$				Negative Correlation $\rho_{xy} = -0.3$			
5 yr	46.40	45.93	58.15	42.11	47.17	46.91	59.89	42.11
10 yr	119.00	121.10	169.98	110.34	122.20	121.54	173.47	110.34
15 yr	237.53	225.52	335.49	204.70	214.52	223.89	340.72	204.70
20 yr	371.47	359.18	554.68	325.18	345.88	353.96	561.66	325.18
25 yr	535.45	522.08	827.55	471.79	529.24	511.74	836.27	471.79
30 yr	740.33	714.23	1154.10	644.53	728.20	697.25	1164.57	644.53

(Unit: basis points)

Table 5. The CMS convexity corrections with $\beta_x = \beta_y = 0.6$ and two different swap-LIBOR correlations. Left-hand side: $\rho_{xy} > 0$; Right-hand side: $\rho_{xy} < 0$

terminal decorrelation. Figure 2 shows that the one-factor model with skew is preferred over the two-factor model, Gausn. App., for all expiry times farther than 5 years.

The reason is that a significant part of the change in $f(X_t, Y_t) = P(t, T_{\text{pay}})/A(t)$ is related to changes in the level of the annuity. The changes in the numeraire ratio are therefore highly *correlated* to the movements in the swap rate. In the present experiment, the partial derivative $\partial f/\partial X$ is significantly larger than $\partial f/\partial Y$ ($\partial f/\partial X = 0.595602318 >> \partial f/\partial Y = -0.058036757$). So, even if the covariance between the two rates is low by specifying a strongly negative correlation between the swap and the LIBOR rate, the terminal decorrelation of the convexity correction, $\frac{\partial f}{\partial Y}\mathrm{Cov}[X, Y]$, is still very small compared to the overall convexity correction.

However, when $\beta \leq 0.5$ the terminal decorrelation has a more significant impact and the accuracy of the expansion reduces quickly with increasing time. Small values of β imply a stronger variability in the underlying process, and the approximation obtained is not sufficiently accurate then.

4.2.2. *Decomposition of the Error*

Two approximations that we made that may have an impact on the accuracy of our pricing formula are the following: We keep the partial derivatives $\frac{\partial f}{\partial X_s}, \frac{\partial f}{\partial Y_s}$ constant, at their initial values, and we approximate the variance/covariance term by the stochastic Taylor expansion. Here, we consider the impact of each of these simplifications separately, so that we can indicate directions for further improvement. We consider here the case $\beta = 1$.

First of all, we focus on the error which originates from the assumption of constant partial derivatives. We compare a CMS convexity correction *with frozen partial derivatives* sampled with the Monte Carlo method with the Monte Carlo results for the true convexity correction. More specifically, given the same set of Monte Carlo paths, we compute the convexity correction, with constant partial derivatives, by

$$\begin{aligned} Cc_t^{(1)} =& \frac{1}{M}\frac{\partial f}{\partial X_t}\sum_{k=1}^{M}\Big[\sum_{i=1}^{N}\Big(\big(\sigma^{(k)}(s_{i-1})\big)^2(X_{s_{i-1}}^{(k)})^{2\beta_x}\Delta t\Big)\Big] \\ &+\frac{1}{M}\frac{\partial f}{\partial Y_t}\sum_{k=1}^{M}\Big[\sum_{i=1}^{N}\Big(\sigma^{(k)}(s_{i-1})(X_{s_{i-1}}^{(k)})^{\beta_x}\sigma^{(k)}(s_{i-1})(Y_{s_{i-1}}^{(k)})^{\beta_y}\rho_1\rho_2\Delta t\Big)\Big], \end{aligned}$$

where $i = 1, 2, \ldots, N$ represents the index for the time steps and k denotes for the number of trails.

We then compute the convexity correction by a step-wise approximation of the time varying partial derivatives, i.e.,

$$\begin{aligned} Cc_t^{(2)} =& \frac{1}{M}\sum_{k=1}^{M}\Big[\sum_{i=1}^{N}\frac{\partial f}{\partial X_{s_{i-1}}}\Big(\big(\sigma^{(k)}(s_{i-1})\big)^2(X_{s_{i-1}}^{(k)})^{2\beta_x}\Delta t\Big)\Big] \\ &+\frac{1}{M}\sum_{k=1}^{M}\Big[\sum_{i=1}^{N}\frac{\partial f}{\partial Y_{s_{i-1}}}\Big(\sigma^{(k)}(s_{i-1})(X_{s_{i-1}}^{(k)})^{\beta_x}\sigma^{(k)}(s_{i-1})(Y_{s_{i-1}}^{(k)})^{\beta_y}\rho_1\rho_2\Delta t\Big)\Big]. \end{aligned}$$

Since we choose the same set of paths, the distributional statistics for the two formulas above are exactly the same. The difference in the prices, $Cc_t^{(1)}$ and $Cc_t^{(2)}$, therefore comes from the assumption of constant partial derivatives. The results for $\beta = 1$ are summarized in Table 6.

The approximation with constant partial derivatives consistently undervalues the convexity correction. The error grows almost linearly in time (see Figure 3). This error can be explained as follows: In the derivation of the convexity correction formula (2.5), constant values for the partial derivatives are set, so that

$$E_t^Q\Big[\int_t^T\frac{\partial f}{\partial X_s}dX_s\cdot\int_t^T dX_s\Big]\approx\frac{\partial f}{\partial X_t}E_t^Q\Big[\int_t^T dX_s\cdot\int_t^T dX_s\Big].$$

This can be interpreted as a first order approximation to the stochastic integral

$\int_t^T \frac{\partial f}{\partial X_s} dX_s$. The accuracy will improve if we include higher order terms, i.e.,

$$\begin{aligned}\int_t^T \frac{\partial f}{\partial X_s} dX_s &= \int_t^T \Big(\frac{\partial f}{\partial X_t} + \int_t^s \big(\frac{\partial^2 f}{\partial X_{s_1}^2}\big) dX_{s_1} + O(s) \Big) dX_s \\ &= \frac{\partial f}{\partial X_t} \int_t^T dX_s + \int_t^T \int_t^s \big(\frac{\partial^2 f}{\partial X_{s_1}^2}\big) dX_{s_1} dX_s + o(T) \\ &= \frac{\partial f}{\partial X_t} \int_t^T dX_s + R,\end{aligned}$$

where R is the remainder resulting from the approximation with constant partial derivatives. Thus, the integral related to the second order partial derivatives, which is the leading term of the remainder R, forms the basis of the pricing error in Table 6. Under the assumption in Sec. 2.1 that the numeraire ratio is varying slowly and thus that the second order partial derivatives in the swap rate are small, these second and higher order partial derivative terms can be neglected.

	Convexity corrections		
	Reference Values	Constant Approx.	Error
	Positive Correlation $\rho_{xy} = 0.3$		
2 yr	6.59	6.55	-0.03
5 yr	16.56	16.22	-0.34
10 yr	35.36	32.67	-2.69
15 yr	59.18	51.72	-7.46
20 yr	82.38	69.80	-12.58
25 yr	112.92	91.64	-21.28
30 yr	143.16	104.89	-38.26
	Negative Correlation $\rho_{xy} = -0.3$		
2 yr	6.5878	6.87	-0.04
5 yr	17.33	16.90	-0.43
10 yr	36.35	34.16	-2.19
15 yr	58.74	52.48	-6.26
20 yr	91.59	76.05	-15.54
25 yr	125.42	99.68	-25.74
30 yr	152.64	117.97	-34.67

(Unit: basis points)

Table 6. Approximation error due to constant partial derivatives, $\beta = 1$.

We also consider the accuracy of the variance/covariance approximation by com-

	Variance approximation			Covariance approximation		
	MC	Expn.	Err.	MC	Expn.	Err.
	Positive Correlation $\rho_{xy} = 0.3$					
2 yr	1.43	1.47	0.04	0.36	0.37	0.01
5 yr	3.55	3.78	0.23	0.80	0.88	0.08
10 yr	7.10	7.89	0.79	1.31	1.61	0.30
15 yr	11.00	12.33	1.33	1.66	2.20	0.53
20 yr	15.00	17.11	2.11	1.90	2.63	0.73
25 yr	20.00	22.21	2.21	2.05	2.91	0.86
30 yr	24.00	27.65	3.65	2.18	3.05	0.87
	Negative Correlation $\rho_{xy} = -0.3$					
2 yr	1.44	1.47	0.03	-0.34	-0.36	-0.01
5 yr	3.54	3.78	0.24	-0.73	-0.79	-0.07
10 yr	7.18	7.89	0.71	-1.12	-1.26	-0.13
15 yr	12.00	12.33	0.33	-1.34	-1.39	-0.05
20 yr	16.00	17.11	1.11	-1.48	-1.20	0.28
25 yr	21.00	22.21	1.21	-1.56	-0.67	0.89
30 yr	25.00	27.65	2.65	-1.63	0.18	1.81

(Unit: basis points)

Table 7. Error in the variance and in the covariance approximation for positive and negative correlations, $\beta = 1$.

paring the expansion solution (3.3) with the following Monte Carlo statistics:

$$\mathrm{Var} \approx \frac{1}{M}\sum_{k=1}^{M}\Big[\sum_{i=1}^{N}\Big(\big(\sigma^{(k)}(s_{i-1})\big)^2 (X^{(k)}_{s_{i-1}})^{2\beta_x}\Delta t\Big)\Big],$$

$$\mathrm{Cov} \approx \frac{1}{M}\sum_{k=1}^{M}\Big[\sum_{i=1}^{N}\Big(\sigma^{(k)}(s_{i-1})(X^{(k)}_{s_{i-1}})^{\beta_x}\sigma^{(k)}(s_{i-1})(Y^{(k)}_{s_{i-1}})^{\beta_y}\rho_1\rho_2\Delta t\Big)\Big].$$

We find that the approximation of the variance by the expansion over-estimates the true variance (see Figure 4). In the current model, the variance component dominates the convexity correction as the partial derivatives of the numeraire ratio w.r.t the swap rate are approximately a factor of 10 larger in size than the partial derivatives of the numeraire ratio w.r.t the LIBOR rate. Hence the approximation errors in the variance component and the corresponding partial derivatives represent the dominant part of the error.

This over-estimation of the variance cancels out, to a large extent, the undervaluation due to the constant approximation of the partial derivatives, and, therefore, the overall error is smaller than the individual errors generated by each approximation. Table 7 presents the details of this test for $\beta = 1$.

Finally, we found that the approximation for the covariance is not stable for contracts with long expiry times and that the error increases with cubic order in time (see Figure 5). This suggests that an approximation of quadratic polynomial (in time) is not sufficient for the estimation of the covariance and that higher order terms (e.g. $O(t^6)$ terms) need to be included for a more accurate approximation.

5. Conclusion

In this paper we have focused on the convexity correction for CMS products under a two-factor SABR model. We derived an approximation for the convexity correction by applying the *small time asymptotics* technique to the Wiener processes involved. An efficient and easy-to-implement approximation formula for the CMS convexity correction with standard measure change is the result of this work.

By numerical experiments, comparing with the corresponding Monte Carlo

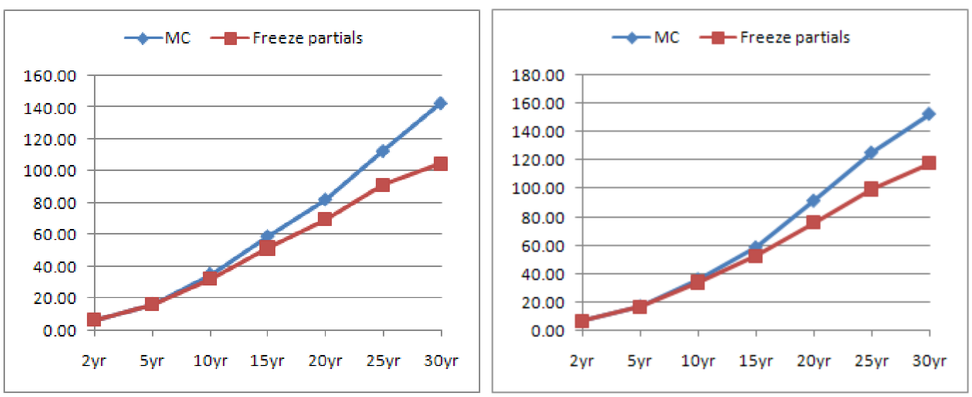


Fig. 3. Error due to constant partial derivatives, with $\beta = 1$. Left-hand side: positive correlation between swap rate and the LIBOR rate. Right-hand side: negative correlation between these two rates.

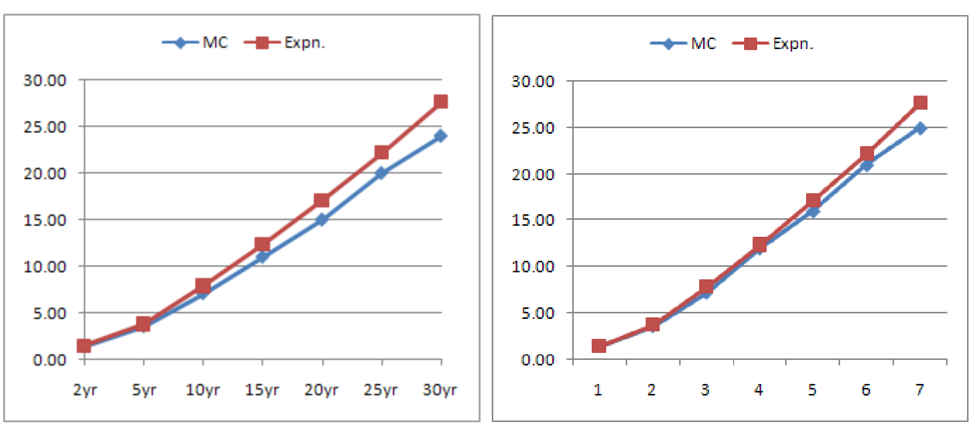


Fig. 4. Error due to the expansion for the variance, $\beta = 1$. Left-hand side: positive correlation between swap rate and the LIBOR rate; Right-hand side: negative correlation between these two rates.

prices, we find that the approximations result in satisfactory accuracy for β-values larger than one-half. In order to detail the impact of the various approximations in our pricing approach, we have set up numerical experiments to determine the individual errors of each approximation. Apparently, two significant parts of the approximation error cancel out, to a large extent. However, a fourth order expansion formula for the covariance as presented here does not appear to be fully sufficient for contracts with very long expiration times, like thirty years.

The approximation has been derived for payoffs that are linear in the swap rate. For more general contracts, the constant partial derivatives approximation requires improvement. Furthermore, only a small amount of terminal decorrelation can be captured by adding an additional factor on the payment leg, since the annuity is still driven by one factor only. A two-factor model which has the two factors that are principal components of the empirical covariance matrix could in principle describe the terminal decorrelation of forward and swap rates in a more realistic way. This will be covered in near-future research.

Appendix A. Description of the Monte Carlo Scheme

We explain the set up of the Monte Carlo simulation of the covariance. The implementation is done in a short time step procedure.

Given the SABR dynamics of X_t and Y_t described in Eq. (2.6), where the stochastic volatility process of both rates is driven by Brownian motion Z_t.

By the Euler discretization of the stochastic differential equation system, we

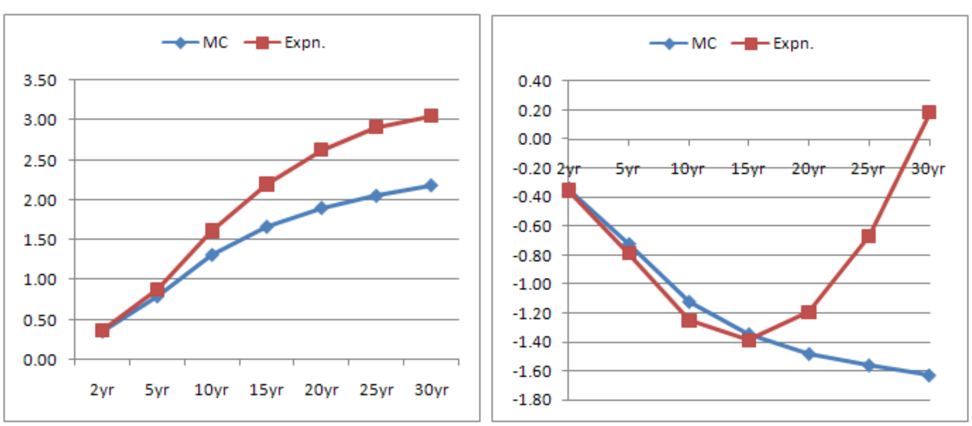


Fig. 5. Error due to the approximation for the covariance, $\beta = 1$. Left-hand side: positive correlation between swap rate and the LIBOR rate; Right-hand side: negative correlation between these two rates.

have the following discretization for X_t in time:

$$X_{s_i} = X_{s_{i-1}} + \sigma(s_{i-1})(X_{s_{i-1}})^{\beta_x}\big(\lambda_x Z_i + \sqrt{1-\lambda_x^2}U_i\big)\sqrt{\Delta t}, \text{ and}$$
$$\sigma(s_i) = \sigma(s_{i-1}) + \eta_x\sigma(s_{i-1})Z_i\sqrt{\Delta t}, \qquad \forall s_i = t + i\Delta t \leq T,$$

where Z_i and U_i are independent Gaussian pseudo-random numbers (with zero mean and unit variance). Similarly, the process Y_t, in discrete time, reads

$$Y_{s_i} = Y_{s_{i-1}} + \Big(\frac{A(s_{i-1})}{P(s_{i-1},T_{\text{pay}})} - \frac{A(s_{i-2})}{P(s_{i-2},T_{\text{pay}})}\Big)\Big(Y_{s_{i-1}} - Y_{s_{i-2}}\Big) + \sigma(s_{i-1})(Y_{s_{i-1}})^{\beta_y}\big(\lambda_y Z_i + \sqrt{1-\lambda_y^2}V_i\big)\sqrt{\Delta t},$$

and

$$\sigma(s_i) = \sigma(s_{i-1}) + \eta_y\sigma(s_{i-1})Z_i\sqrt{\Delta t}, \qquad \forall s_i = t + i\Delta t \leq T,$$

where Z_i represents the same set of random numbers used before, and the V_i represents another set of random numbers. As a result, for the variance and covariance quantities in Eq. (4.7), in discrete time, we find:

$$\begin{aligned}
\frac{\partial f}{\partial X_t}\text{Var}_t^Q[X_T] + \frac{\partial f}{\partial Y_t}\text{Cov}_t^Q[X_T,Y_T] \approx& \frac{\partial f}{\partial X_t}E\Big[\sum_{i=1}^{N}\Big(\sigma(s_{i-1})^2(X_{s_{i-1}})^{2\beta_x}\Delta t\Big)\Big]\\
&+ \frac{\partial f}{\partial Y_t}E\big[\sum_{i=1}^{N}\Big(\sigma(s_{i-1})(X_{s_{i-1}})^{\beta_x}\sigma(s_{i-1})(Y_{s_{i-1}})^{\beta_y}\rho_1\rho_2\Delta t\Big)\big]\\
\approx& \frac{1}{M}\frac{\partial f}{\partial X_t}\sum_{k=1}^{M}\Big[\sum_{i=1}^{N}\Big(\big(\sigma^{(k)}(s_{i-1})\big)^2(X_{s_{i-1}}^{(k)})^{2\beta_x}\Delta t\Big)\Big]\\
&+ \frac{1}{M}\frac{\partial f}{\partial Y_t}\sum_{k=1}^{M}\Big[\sum_{i=1}^{N}\Big(\sigma^{(k)}(s_{i-1})(X_{s_{i-1}}^{(k)})^{\beta_x}\sigma^{(k)}(s_{i-1})(Y_{s_{i-1}}^{(k)})^{\beta_y}\rho_1\rho_2\Delta t\Big)\Big],
\end{aligned}$$

where the superscript denotes the k-th path of the Monte Carlo simulation. The expectation is approximated by the average of a large number of paths. Note that we have applied the Itô product formula to obtain the time discretization scheme above.

References

[1] L. Andersen and J. Andreasen, Volatile Volatilities, *Risk Magazine* (2002).
[2] E. Benhamou, Swaps: Constant maturity swaps (CMS) and Constant maturity Treasury (CMT) swaps, *Swap Strategy Note, FICC, Goldman Sachs International* (2003).
[3] D. Brigo and F. Mercurio, Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit, *Springer Finance* (2007) Second Edition.
[4] P. Hagan, Convexity Conundrums: Pricing CMS swaps, Caps and Floors, *Wilmott Magazine* (2000) 38–44.
[5] P.S. Hagan, D. Kumar, A.S. Lesniewski and D.E. Woodward, Managing smile risk, *Wilmott Magazine* (2002) 84–108.
[6] P. Hagan and A. Lesniewski, LIBOR Market Model with SABR style stochastic volatility, *Working Paper* (2008).

[7] M. Henrard, CMS swaps in separable one-factor Gaussian LLM and HJM model, *Bank for International Settlements Working Paper* (2007).

[8] J. Hull and A. White, Forward Rate Volatilities, Swap Rate Volatilities and the Implementation of the LIBOR Market Model, *Journal of Fixed Income* **10**:2 (2000) 46–62.

[9] P.J. Hunt and J.E. Kennedy, Financial Derivatives in Theory and Practice, *WILEY, UK* (2000).

[10] K. Itô, Multiple Wiener integral, *The Mathematical Society of Japan* **3** (1951) 157–169.

[11] P. Jäckel and R. Rebonato, Linking Caplet and Swaption Volatilities in a BGM/J framework: approximate Solutions, *Journal of Computational Finance*, **6**:4 (2003).

[12] A. Kawai, A New Approximate Swaption Formula in the LIBOR Market Model: An Asymptotic Expansion Approach, *Applied Mathematical Finance* **10** (2003) 49–74.

[13] A. Kawai and P. Jäckel, An asymptotic FX option formula in the cross currency LIBOR Market Model, *Wilmott Magazine* (2007) 74–84.

[14] P. Jäckel and A. Kawai, The Future is Convex, *Wilmott* (2005) 44 – 55.

[15] P.E. Klöden and E. Platen, Numerical Solution of Stochastic Differential Equations, *Springer Heidelberg* (1992)

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

[17] N. Kunitomo and A. Takahashi, On Validity of the Asymptotic Expansion Approach in Contingent Claim Analysis, *Annals of Applied Probabilities* **13**:3 (2003) 914 –952.

[18] F. Mercurio and A. Pallavicini, Smiling at Convexity: Bridging Swaption Skews and CMS Adjustments, *Available at SSRN: http://ssrn.com/abstract=892287 (2006)* (2006).

[19] A. Pelsser, Mathematical Foundation of Convexity Correction, *Quantitative Finance* **3**:1 (2003) 59–65.

[20] V. Piterbarg, Stochastic Volatility Model with Time dependent Skew, *Applied Mathematical Finance* **12**:2 (2005) 147–185.

[21] R. Rebonato, Modern Pricing of Interest-Rate Derivatives, *Princeton University Press* (2002) First Edition.

[22] R. Rebonato, A Time-Homogeneous, SABR-Consistent Extension of the LMM, *Risk* **20**:November (2007) 92–97.

[23] S. van Weeren and E. van Dijk, Convexity Correction: a General Approach, *Rabobank Derivative Research & Validation Working Paper* (1999).

[24] L. Yinqiu and N. Salih, Convexity Adjustments and Forward LIBOR Model: Case of Constant Maturity Swaps, *Switzerland National Centre of Competence in Research Financial Valuation and Risk Management Working Paper No. 115* (2003).

[25] N. Yoshida, Asymptotic Expansion for Small Diffusion via the Theory of Malliavin-Watanabe, *Probability Theory and Related Fields* **92** (1992) 275 – 311.