Efficient pricing and Greeks in the cross-currency LIBOR market model

Chris J. Beveridge, Mark S. Joshi and Will M. Wright
The University of Melbourne

October 14, 2010

Abstract

We discuss the issues involved in an efficient computation of the price and sensitivities of Bermudan exotic interest rate derivatives in the cross-currency displaced diffusion LIBOR market model. Improvements recently developed for an efficient implementation of the displaced diffusion LIBOR market model are extended to the cross-currency setting, including the adjoint-improved pathwise method for computing sensitivities and techniques used to handle Bermudan optionality. To demonstrate the application of this work, we provide extensive numerical results on two commonly traded cross-currency exotic interest rate derivatives: cross-currency swaps (CCS) and power reverse dual currency (PRDC) swaps.

1 Introduction

Long-dated callable notes with coupons linked to a foreign-exchange rate pose exceptional problems to a risk-manager. The callable power reverse dual is the most notorious such product. In order, to properly model the dynamics of the interest-rates in two different currencies and an exchange rate, a high-dimensional model is necessary. Whilst it is just about possible to use 3 factors, one per currency and the exchange rate, there is a clear need to benchmark against high-dimensional models for the assessment of model risk. Such models necessitate the use of Monte Carlo simulation.

However, each step of the Monte Carlo simulation for such complicated products requires much computational effort, and we will have many steps for a 30-year note. In addition, callability is tricky because of the non-availability of continuation values. Whilst many techniques have now been developed to handle early exercise features, these have not been robustly tested in the cross-currency setting. In addition, the rapid computation of Greeks is important for risk-assessment and hedging. In this paper, we therefore examine the problems of pricing and sensitivity computation for such notes using a high-dimensional cross-currency LIBOR market model. We demonstrate that our methods are robust and show how to compute both prices and first-order Greeks in a small amount of time.

Despite the popularity of cross-currency exotic interest rate products and the extensive literature on LIBOR market model, see for example (Brace, 2008; Brigo and Mercurio, 2006; Rebonato, 2002), there are surprisingly few articles in the literature which extend the LIBOR market model and associated techniques to the cross-currency setting. The first papers to develop cross-currency LIBOR market models were (Mikkelsen, 1999; Schlögl, 2002). The main result in the paper by Schlögl shows that it is not possible to model all of the forward rates in both the domestic and foreign currencies and all of the forward exchange rates using log-normal processes, this is discussed further in Section 2. It appears that the book (Brace, 2008, Chapter 14) has the most extensive discussion currently available on the cross-currency LIBOR market model. In (Musiela and Rutkowski, 2005, Chapter 14) cross-currency derivatives in the LIBOR market model are briefly discussed. The book (Fries, 2007, Chapter 26) also has a chapter on the cross currency LIBOR market model, but only discusses the formula for computing the drifts in the domestic forward rate, foreign forward rate and exchange rate SDEs.

In (Jarrow and Yildirim, 2003) it was shown that the cross-currency models can be extended to price inflation-linked derivatives. The approach they followed was to let the domestic rates represent the nominal rates (observed real-world interest rates) and the foreign rates represent the real rates (non-observed nominal minus inflation interest rates) and the exchange rate between the two currencies, represents the Consumer

Price Index (CPI). Inflation is then measured as a change in the CPI. In most cases, the nominal and real interest rates are modelled using a Hull–White model and geometric Brownian motion for the exchange between the nominal and real rates. One main disadvantages of this model is that calibration is difficult. In the book (Brigo and Mercurio, 2006, Part VI) two alternative approaches both based on the cross-currency LIBOR market model were discussed. In each case, closed form solutions were derived for the zero-coupon swaps and the year-to-year swaps, allowing more accurate calibrations. The reader is also referred to the book (Brace, 2008, Chapter 15), which describes an inflation model as an application of the cross-currency LIBOR market model. For an in depth discussion on inflation indexed securities we recommend the book (Deacon et al., 2004).

The computation of sensitivities is essential for hedging and risk management. For sufficiently smooth payoffs it is possible in the LIBOR market model to use the pathwise method, which was introduced by (Broadie and Glasserman, 1996), and improved upon in (Glasserman and Zhao, 1999). In (Giles and Glasserman, 2006) the adjoint method was used which is most effective when calculating large numbers of sensitivities for a small number of portfolios. This is exactly the situation we face in the cross-currency LIBOR market model. We extend the results given in (Giles and Glasserman, 2006) to the cross-currency setting.

Using Monte Carlo simulation is generally regarded as necessary in models with high dimensional state spaces, the cross-currency LIBOR market model being an indicative example. In recent times, the industry standard for pricing options with early-exercise features has been the regression approach developed by (Carrière, 1996; Tsitsiklis and van Roy, 2001; Longstaff and Schwartz, 2001). See also the books (Brace, 2008; Fries, 2007) for in depth discussions.

Recent progress, see (Broadie and Cao, 2008) and (Beveridge and Joshi, 2009), has been made on implementation issues relating to the pricing of early exercisable contracts using Monte Carlo simulation. Several improvements to the standard least squares regression were suggested in the single-currency setting: for observations which are considered to be close to the exercise boundary, a second regression is used to refine and improve the decision of whether to exercise; an adaptive approach to choosing basis functions which is generic and accurate; a delta hedge control variate, which uses the continuation value estimates of the least squares regression approach to obtain delta estimates, which are then used to form a delta hedge control variate portfolio. We extend these techniques to the cross-currency setting. These extensions make it possible to accurately and efficiently compute the price and sensitivities for exotic cross-currency options in the LIBOR market model providing an alternative to the method of choice namely the numerical solution of PDEs. In order to test them thoroughly, we introduce a comprehensive new batch of tests. These tests are more rigorous and thorough than those previously utilized in other settings since it is essential for a risk manager to be fully confident that the price is robust. All of the tests are passed convincingly.

The outline of the paper is as follows. In Section 2 we give a general formulation of the cross-currency LIBOR market model and derive expressions for the drifts in the domestic and foreign economies. The numerical methods used to evolve the forward rates in both economies and the exchange rate are discussed in Section 3. The efficient evaluation of the sensitivities is discussed in Section 4. We give a detailed description in Section 5 of the early exercise techniques included in our implementation. Calibration in the cross-currency LIBOR market model is discussed in Section 6. Finally, Section 7 gives various numerical results showing that the cross-currency LIBOR market model can be used to efficiently price cross-currency exotic interest rate derivatives.

2 Cross-currency LIBOR market model

We start the discussion by setting up the cross-currency LIBOR market model. We assume that there are two currencies, domestic and foreign. However, it is possible to extend everything in this section to include more than two currencies if necessary. Let $N \in \mathbf{N}$, and suppose we have a set of tenor dates $0 = T_0 < T_1 < \cdots < T_{N+1} \in \mathbf{R}$. Associated to these tenor dates are domestic and foreign forward rates, denoted by $f_i(t) \in \mathbf{R}$ in the domestic economy and $\widetilde{f}_i(t) \in \mathbf{R}$ in the foreign, for $i = 1, \ldots, N$. Define the tenors by $\tau_i = T_{i+1} - T_i$, and let $P_i(t) \in \mathbf{R}$ for $i = 1, \ldots, N+1$ denote the price at time t of the zero coupon bond paying D\$1 in the domestic currency at T_i. Likewise, let $\widetilde{P}_i(t) \in \mathbf{R}$ for $i = 1, \ldots, N+1$ denote the price at time t of the zero coupon bond paying F\$1 in the foreign currency at T_i. Note that $P_i(t)$ is the price process of a tradeable asset in the domestic currency but not in the foreign currency and $\widetilde{P}_i(t)$ is the price process of a tradeable asset in the foreign currency but not in the domestic currency. The forward rates in the domestic and foreign currencies are related to the bonds in the respective markets as follows,

$$f_i(t) = \frac{1}{\tau_i}\left(\frac{P_i(t)}{P_{i+1}(t)} - 1\right), \quad \widetilde{f}_i(t) = \frac{1}{\tau_i}\left(\frac{\widetilde{P}_i(t)}{\widetilde{P}_{i+1}(t)} - 1\right). \tag{1}$$

We also need a rate of exchange between the two currencies. Let $\mathrm{FX}(t) \in \mathbf{R}$ denote the time t value of 1 unit of foreign currency in domestic currency, that is F\$1 $\equiv$ D\$FX(t). We refer to $\mathrm{FX}(t)$ as the spot exchange rate. We also require the notion of forward exchange rates. Let $\mathrm{FFX}_i(t) \in \mathbf{R}$ denote the exchange rate we can lock in today for exchanging foreign currency to domestic at T_i. Accordingly, we obtain, using no arbitrage arguments, that the interest rate parity relationship for $i = 1, \ldots, N+1$ is

$$\mathrm{FFX}_i(t)P_i(t) = \mathrm{FX}(t)\widetilde{P}_i(t). \tag{2}$$

We will assume that the forward rates in both economies have displaced lognormal volatilities for $i = 1, \ldots, N$,

$$\mathrm{d}f_i(t) = \mu_i(t)\mathrm{d}t + (f_i(t) + \alpha_i)\lambda_i(t) \cdot \mathrm{d}W(t), \tag{3}$$

$$\mathrm{d}\widetilde{f}_i(t) = \widetilde{\mu}_i(t)\mathrm{d}t + (\widetilde{f}_i(t) + \widetilde{\alpha}_i)\widetilde{\lambda}_i(t) \cdot \mathrm{d}W(t), \tag{4}$$

with given initial conditions $f_i(T_0)$ and $\widetilde{f}_i(T_0)$, where $\lambda_i(t) \in \mathbf{R}^F$ and $\widetilde{\lambda}_i(t) \in \mathbf{R}^F$ are deterministic vectors, and $W(t) \in \mathbf{R}^F$ is an F-dimensional standard Brownian motion. The coefficients $\alpha_i \in \mathbf{R}$ and $\widetilde{\alpha}_i \in \mathbf{R}$ are used to fit to the skew observed in the market. However, they are unable to capture smile effects. We also assume that the forward exchange rates are modelled by the SDEs

$$\mathrm{dFFX}_i(t) = \widehat{\mu}_i(t)\mathrm{d}t + \mathrm{FFX}_i(t)\widehat{\lambda}_i(t) \cdot \mathrm{d}W(t), \tag{5}$$

where $i = 1, \ldots, N+1$, and $\widehat{\lambda}_i(t) \in \mathbf{R}^F$. We are interested in evolving the $3N$ quantities $f_i(t)$, $\widetilde{f}_i(t)$ and $\mathrm{FFX}_i(t)$ for $i = 1, \ldots, N$, however the relationship in Equation (2) implies that of $3N$ quantities $2N + 1$ need to be determined and the others are computed by satisfying Equation (2). Alternatively, if $\lambda_i(t)$ and $\widetilde{\lambda}_i(t)$ are chosen to be deterministic for all $i = 1, \ldots, N$, then we are only able to ensure that one of $\widehat{\lambda}_i(t)$ is deterministic and all others are then stochastic. This was first pointed out by (Schlögl, 2002). It is possible to include displaced diffusion into the model for the forward exchange rates Equation (5), but this leads to significant problems see (Brace, 2008, Remark 14.3) and we therefore avoid it.

We are now interested in computing the drift terms $\mu_i(t)$, $\widetilde{\mu}_i(t)$ and $\widehat{\mu}_i(t)$ and interrelationships between the volatilities $\lambda_i(t)$, $\widetilde{\lambda}_i(t)$ and $\widehat{\lambda}_i(t)$ given in Equations (3), (4) and (5). These results were first developed in (Schlögl, 2002) using alternative arguments, we rederive them now. Let $X(t)$ be given by the ratio of prices of tradeable assets, where the denominator is given by $Z(t)$ and $N(t)$ is the value of the numéraire asset, then we can compute the drift of $X(t)$, $\mu_X(t)$, using a result from (Joshi and Liesch, 2007) stating

$$\mu_X(t) = -\frac{N(t)}{Z(t)}\left\langle X(t), \frac{Z(t)}{N(t)}\right\rangle = \frac{Z(t)}{N(t)}\left\langle X(t), \frac{N(t)}{Z(t)}\right\rangle. \tag{6}$$

For two Itô processes

$$\mathrm{d}X(t) = \mu_X(t)\mathrm{d}t + \lambda_X(t) \cdot \mathrm{d}W(t),$$
$$\mathrm{d}Y(t) = \mu_Y(t)\mathrm{d}t + \lambda_Y(t) \cdot \mathrm{d}W(t),$$

the cross variation derivative from (Joshi and Liesch, 2007) is defined as

$$\langle X(t), Y(t)\rangle = \frac{\mathrm{d}X(t)\mathrm{d}Y(t)}{\mathrm{d}t} = \lambda_X(t) \cdot \lambda_Y(t).$$

An important point is that we do not need to know the form of $Z(t)$ and $N(t)$, we only need to know the process of the ratio $Z(t)/N(t)$ and that of $X(t)$.

Using this result, we can now easily calculate the drifts for each of our state variables. We calculate drifts using $P_j(t)$ as numéraire. Since we are considering a generalization of the standard LIBOR market model, we expect to recover the drift equations from that model for the domestic currency.

To determine the drifts of the domestic forward rates we note that $P_{i+1}(t)$ and $f_i(t)P_{i+1}(t) = (P_i(t) - P_{i+1}(t))/\tau_i$ are domestic tradeable assets. Using the result from Equation (6), the drifts of the domestic forward rates are therefore

$$\mu_i(t) = -\frac{P_j(t)}{P_{i+1}(t)}\left\langle f_i(t), \frac{P_{i+1}(t)}{P_j(t)}\right\rangle, \tag{7}$$

recovering the result from the standard LIBOR market model.

To determine the dynamics of the foreign forward rates, we note that $\widetilde{P}_{i+1}(t)\mathrm{FX}(t)$ and, from Equation (1), $\widetilde{f}_i(t)\widetilde{P}_{i+1}(t)\mathrm{FX}(t)$ are the values of domestic tradeable assets. Using Equation (2), we remove the dependence on the spot exchange rate in favour of the forward exchange rate, giving the tradeables

$$\frac{\widetilde{P}_{i+1}(t)P_j(t)\mathrm{FFX}_j(t)}{\widetilde{P}_j(t)}, \quad \frac{\widetilde{f}_i(t)P_j(t)\widetilde{P}_{i+1}(t)\mathrm{FFX}_j(t)}{\widetilde{P}_j(t)}.$$

Since $\widetilde{f}_i(t)$ is given by the ratio of the value of the two domestic tradeable assets above, again using the result from Equation (6) we get

$$\begin{aligned}\widetilde{\mu}_i(t) &= -\left(\frac{\widetilde{P}_{i+1}(t)}{\widetilde{P}_j(t)}\mathrm{FFX}_j(t)\right)^{-1}\left\langle \widetilde{f}_i(t), \frac{\widetilde{P}_{i+1}(t)}{\widetilde{P}_j(t)}\mathrm{FFX}_j(t)\right\rangle, \\ &= -\frac{\widetilde{P}_j(t)}{\widetilde{P}_{i+1}(t)}\left\langle \widetilde{f}_i(t), \frac{\widetilde{P}_{i+1}(t)}{\widetilde{P}_j(t)}\right\rangle - \left\langle \widetilde{f}_i(t), \mathrm{FFX}_j(t)\right\rangle \frac{1}{\mathrm{FFX}_j(t)}.\end{aligned} \tag{8}$$

The first term is the same as the regular drift term in LIBOR market model, but with $\widetilde{P}_j(t)$ as the numéraire. The second term is the necessary correction due to the fact that we are using the domestic bond $P_j(t)$ as the numéraire and not $\widetilde{P}_j(t)$.

To determine the drifts of the forward exchange rates we note that $P_i(t)$ and $\mathrm{FX}(t)\widetilde{P}_i(t)$ are the values of domestic tradeable assets, and $\mathrm{FFX}_i(t)$ can be recovered as their ratio. Once again, using Equation (6) the drifts of the forward exchange rates are

$$\widehat{\mu}_i(t) = -\frac{P_j(t)}{P_i(t)}\left\langle \mathrm{FFX}_i(t), \frac{P_i(t)}{P_j(t)}\right\rangle. \tag{9}$$

Equation (2) ensures that $\mathrm{FFX}_j(t)$ is a martingale when $P_j(t)$ is the numéraire since $\mathrm{FX}(t)\widetilde{P}_j(t)$ is a domestic tradeable. This is entirely consistent with Equation (9); when $i = j$, $\widehat{\mu}_i(t) = 0$. So when $P_i(t)$ is the numéraire the i:th forward exchange rate is given by

$$\mathrm{FFX}_i(t) = \mathrm{FFX}_i(0)\exp\left(-\frac{1}{2}\int_0^t \widehat{\lambda}_i(u)\cdot\widehat{\lambda}_i(u)\mathrm{d}t + \int_0^t \widehat{\lambda}_i(u)\cdot \mathrm{d}W(u)\right), \tag{10}$$

where the initial condition $\mathrm{FFX}_i(0)$ is known.

Although it is possible to calculate the drifts of the spot exchange rate, this is something we avoid since the drifts depend on the difference between the domestic and foreign short rates; see, for example, (Fries, 2007, Chapter 26). The presence of the short rates in the dynamics complicates matters significantly, but we do not need to work with the spot exchange rate directly: from Equation (2), having a single forward exchange rate and the zero-coupon bond prices are enough to determine everything we need.

We focus our attention on the use of the spot measure, which is made up of an initial portfolio of one zero coupon bond expiring at T_1, with the proceeds being reinvested in bonds expiring at the next tenor date, up until T_{N+1}. The value of the numéraire portfolio at time t in the domestic currency is given by

$$N(t) = P_{\eta(t)}(t)\prod_{i=1}^{\eta(t)-1}\left(1 + \tau_i f_i(T_i)\right), \tag{11}$$

where $\eta(t)$ defines the index of the next forward rate to reset and is given by the unique integer satisfying

$$T_{\eta(t)-1} \leq t < T_{\eta(t)}.$$

We now have all the ingredients we need to express the drifts of the cross-currency LIBOR market model in the domestic spot measure. In particular, for $i = 1, \dots, N$,

$$\begin{aligned}\mu_i(t) &= \frac{P_{i+1}(t)}{P_{\eta(t)}(t)} \left\langle f_i(t), \frac{P_{\eta(t)}(t)}{P_{i+1}(t)} \right\rangle \\ &= \frac{P_{i+1}(t)}{P_{\eta(t)}(t)} \left\langle f_i(t), \prod_{k=\eta(t)}^{i} (1+\tau_k f_k(t) \right\rangle \\ &= \frac{P_{i+1}(t)}{P_{\eta(t)}(t)} \sum_{k=\eta(t)}^{i} \left[\langle f_i(t), 1+\tau_k f_k(t) \rangle \prod_{\substack{k=\eta(t)\\ j\neq k}}^{i} (1+\tau_j f_j(t)) \right] \\ &= \sum_{k=\eta(t)}^{i} \frac{\tau_k \langle f_i(t), f_k(t) \rangle}{1+\tau_k f_k(t)}.\end{aligned}$$

Using similar calculations, we obtain a complete set of drifts,

$$\mu_i(t) = (f_i(t) + \alpha_i) \sum_{k=\eta(t)}^{i} h_k(t)\, \lambda_i(t) \cdot \lambda_k(t), \tag{12}$$

$$\widetilde{\mu}_i(t) = (\widetilde{f}_i(t) + \widetilde{\alpha}_i) \left(\sum_{k=\eta(t)}^{i} \widetilde{h}_k(t)\, \widetilde{\lambda}_i(t) \cdot \widetilde{\lambda}_k(t) - \widetilde{\lambda}_i(t) \cdot \widehat{\lambda}_{\eta(t)}(t) \right), \tag{13}$$

$$\widehat{\mu}_i(t) = \mathrm{FFX}_i(t) \sum_{k=\eta(t)}^{i-1} h_k(t)\, \widehat{\lambda}_i(t) \cdot \lambda_k(t),$$

where $h_r(t)$ and $\widetilde{h}_r(t)$ are defined to be,

$$h_r(t) = \frac{\tau_r (f_r(t) + \alpha_r)}{1+\tau_r f_r(t)}, \quad \widetilde{h}_r(t) = \frac{\tau_r (\widetilde{f}_r(t) + \widetilde{\alpha}_r)}{1+\tau_r \widetilde{f}_r(t)}. \tag{14}$$

We have computed the drifts directly in the desired domestic measure. It is also possible to compute the foreign drifts in the desired foreign measure and use change of measure arguments. Comparing the drifts in Equations (12) and (13) yields the required measure change

$$\mathrm{d}W(t) = \mathrm{d}\widetilde{W}(t) + \widehat{\lambda}_{\eta(t)}(t)\mathrm{d}t,$$

that was given in (Schlögl, 2002, Equation (8)).

While only one forward exchange rate can have deterministic volatility, we need to determine the volatility of the other forward exchange rates for calibration purposes. Given that Equation (2) holds for all i, then for all k and ℓ

$$\mathrm{FFX}_k(t) \frac{P_k(t)}{\widetilde{P}_k(t)} = \mathrm{FFX}_\ell(t) \frac{P_\ell(t)}{\widetilde{P}_\ell(t)}.$$

Express the ratios of the domestic bonds $P_\ell(t)/P_k(t)$ and the foreign bonds $\widetilde{P}_k(t)/\widetilde{P}_\ell(t)$, in terms of the domestic and foreign forward rates respectively, and assuming without loss of generality that $k > \ell$, gives

$$\mathrm{FFX}_k(t) = \mathrm{FFX}_\ell(t) \prod_{r=\ell}^{k-1} \frac{(1+\tau_r f_r(t))}{(1+\tau_r \widetilde{f}_r(t))}.$$

Using Itô's Lemma and collecting only the stochastic terms, leads to the expression

$$\widehat{\lambda}_k(t) = \widehat{\lambda}_\ell(t) + \sum_{r=\ell}^{k-1} h_r(t)\lambda_r(t) - \sum_{r=\ell}^{k-1} \widetilde{h}_r(t)\widetilde{\lambda}_r(t), \tag{15}$$

This is a repeated application of (Schlögl, 2002, Equation (11)) which is useful for calibration purposes that we will discuss in Section 6.

3 Numerical methods

We now discuss how to evolve the state variables of our model. Without loss of generality, assume we are evolving from T_{k-1} to T_k. First, we need to evolve domestic forward rates and foreign forward rates using Equations (3) and (4), with drift terms given by Equations (12) and (13). Given that the drifts are state dependent it is not possible to find exact solutions to these SDEs, and we need to use approximations. To determine the spot exchange rate and all forward exchange rates, it is enough to evolve one forward exchange rate on top of the forward interest rates from (2). As such, we also evolve the forward exchange rate between each set of tenor dates that is a martingale and has deterministic volatility, see Equation (5), and can therefore be simulated exactly.

Let $X_i(t) = \log(f_i(t) + \alpha_i)$, which takes advantage of the fact that the displaced SDE has lognormal volatilities. Likewise, let $\widetilde{X}_i(t) = \log(\widetilde{f}_i(t) + \widetilde{\alpha}_i)$. Using Itô's lemma gives

$$X_i(T_k) = X_i(T_{k-1}) + \int_{T_{k-1}}^{T_k} \mu_i(u)\mathrm{d}u - \frac{1}{2}\int_{T_{k-1}}^{T_k} \lambda_i(u)\cdot\lambda_i(u)\mathrm{d}u + \int_{T_{k-1}}^{T_k} \lambda_i(u)\cdot \mathrm{d}W(u) \tag{16}$$

$$\widetilde{X}_i(T_k) = \widetilde{X}_i(T_{k-1}) + \int_{T_{k-1}}^{T_k} \widetilde{\mu}_i(u)\mathrm{d}u - \frac{1}{2}\int_{T_{k-1}}^{T_k} \widetilde{\lambda}_i(u)\cdot\widetilde{\lambda}_i(u)\mathrm{d}u + \int_{T_{k-1}}^{T_k} \widetilde{\lambda}_i(u)\cdot \mathrm{d}W(u). \tag{17}$$

Also consider $\widehat{X}_k(t) = \log(\mathrm{FFX}_k(t))$, which has SDE

$$\widehat{X}_k(T_k) = \widehat{X}_k(T_{k-1}) - \frac{1}{2}\int_{T_{k-1}}^{T_k} \widehat{\lambda}_k(u)\cdot\widehat{\lambda}_k(u)\mathrm{d}u + \int_{T_{k-1}}^{T_k} \widehat{\lambda}_k(u)\cdot \mathrm{d}W(u). \tag{18}$$

We can simulate the random integrals above exactly, since they are jointly normal with zero mean and have covariance matrix given by $C_k = C(T_k) - C(T_{k-1})$, where we can express the overall $(2N+1)\times(2N+1)$ cross-currency covariance matrix as follows

$$C(t) = \begin{bmatrix} C^D(t) & C^{DF}(t) & C^{DX}(t) \\ C^{DF}(t) & C^F(t) & C^{FX}(t) \\ C^{DX}(t) & C^{FX}(t) & C^X(t) \end{bmatrix}. \tag{19}$$

with the individual covariance terms,

$$\begin{aligned} C^D_{ij}(t) &= \int_0^t \lambda_i(u)\cdot\lambda_j(u)\mathrm{d}u, \\ C^F_{ij}(t) &= \int_0^t \widetilde{\lambda}_i(u)\cdot\widetilde{\lambda}_j(u)\mathrm{d}u, \\ C^{DF}_{ij}(t) &= \int_0^t \lambda_i(u)\cdot\widetilde{\lambda}_j(u)\mathrm{d}u, \\ C^{DX}_{i}(t) &= \int_0^t \lambda_i(u)\cdot\widehat{\lambda}_k(u)\mathrm{d}u, \\ C^{FX}_{i}(t) &= \int_0^t \widetilde{\lambda}_i(u)\cdot\widehat{\lambda}_k(u)\mathrm{d}u, \\ C^{X}(t) &= \int_0^t \widehat{\lambda}_k(u)\cdot\widehat{\lambda}_k(u)\mathrm{d}u. \end{aligned}$$

The main difficulty in evolving the state variables then revolves around approximating the integrals of the drifts in Equations (16) and (17). The simplest solution is to use a modified log Euler–Maruyama method where we only freeze state-dependence, approximating the drift integrals as

$$\int_{T_{k-1}}^{T_k} \mu_i(u)\mathrm{d}u \approx \mu_i^k = \sum_{j=k}^{i} h_j(T_{k-1}) C^D_{kij}, \tag{20}$$

$$\int_{T_{k-1}}^{T_k} \widetilde{\mu}_i(u)\mathrm{d}u \approx \tilde{\mu}_i^k = \sum_{j=k}^{i} \widetilde{h}_j(T_{k-1}) C^F_{kij} - C^{FX}_{ki}. \tag{21}$$

This constitutes one of the simplest numerical methods that can be used to compute approximations to the SDEs given in Equations (3) and (4). Given that only the drift terms need to be approximated, using more sophisticated approximations have been considered; see the paper Hunter et al. (2001) for a generalization of the trapezoidal rule, which has become known as the predictor-corrector method. Given that the shocks are constant over each integration interval (commonly referred to as SDEs with additive noise), it is likely that higher order methods can be computed with significantly less computational cost than higher order methods developed for the more general multiplicative noise; see Kloeden and Platen (1992); Burrage and Burrage (1998).

If our combined covariance matrix, C_k, is of rank F, it is possible to evolve our state variables with a computational complexity of $O(NF)$. This follows by the results from Joshi (2003), which extends directly to the cross-currency LIBOR market model. As such, to reduce computational time it is common to work in models where the number of factors is significantly less than the number of rates. This can be done by directly calibrating reduced factor models (as in Ametrano and Joshi (2008)), or by using a reduced-factor approximation once a calibration has been performed (as in Dun (2006)). In each case, we write

$$C_k = \begin{bmatrix} C_k^D & C_k^{DF} & C_k^{DX} \\ C_k^{DF} & C_k^F & C_k^{FX} \\ C_k^{DX} & C_k^{FX} & C_k^X \end{bmatrix} = A_k A_k^T, \quad A_k = \begin{bmatrix} A_k^D \\ A_k^F \\ A_k^X \end{bmatrix}, \tag{22}$$

where A_k is an $(2N+1) \times F$ matrix, and the first N rows A_k^D are used to evolve the domestic forward rates, the second N rows A_k^F are used to evolve the foreign forward rates and the last row A_k^X evolves the forward exchange rate. As was pointed out in (Brace, 2008, Pg. 133), generally $F \approx 7$ is needed to capture the dynamics of the domestic and foreign forward rates and the forward exchange rate.

Using this notation, we evolve our state variables according to

$$f_i(T_k) = (f_i(T_{k-1}) + \alpha_i) \exp\left(\mu_i^k - \frac{1}{2} C_{kii}^D + \sum_{f=1}^{F} A_{kif}^D Z_{kf}\right) - \alpha_i, \tag{23}$$

$$\widetilde{f}_i(T_k) = (\widetilde{f}_i(T_{k-1}) + \widetilde{\alpha}_i) \exp\left(\tilde{\mu}_i^k - \frac{1}{2} C_{kii}^F + \sum_{f=1}^{F} A_{kif}^F Z_{kf}\right) - \widetilde{\alpha}_i, \tag{24}$$

$$\mathrm{FX}(T_k) = \mathrm{FX}(T_{k-1}) \exp\left(-\frac{1}{2} C_k^X + \sum_{f=1}^{F} A_{kf}^X Z_{kf}\right) \frac{\widetilde{P}_k(T_{k-1})}{P_k(T_{k-1})}, \tag{25}$$

with the drift terms μ_i^k and $\tilde{\mu}_i^k$ given in Equations (20) and (21).

4 Greeks in the cross-currency LIBOR market model

Sensitivities are essential quantities in hedging and risk management. A naïve approach to calculating sensitivities is to use finite difference approximations. This typically involves re-computing the price of a financial contract with a small change in the initial input parameters. In the cross-currency LIBOR market model this would require us to re-compute the price for each forward rate to compute the relevant deltas, and can therefore be very time consuming. In a recent breakthrough, (Giles and Glasserman, 2006) propose an efficient method to compute the sensitivities on a path by path basis, using adjoints to extend the approaches derived in (Broadie and Glasserman, 1996; Glasserman and Zhao, 1999). The adjoint method leads to significant computational improvements, most noticeably when the sensitivities of a small number of products are required to a large number number of inputs, such as the initial forward rates in the cross-currency LIBOR market model. In (Giles and Glasserman, 2006), the necessary formulae for computing the deltas and vegas, for European contracts, were derived for the log Euler–Maruyama method under the spot LIBOR measure. This was extended to the predictor-corrector method under the spot and terminal measures in (Denson and Joshi, 2009a). In (Denson and Joshi, 2009b), it was shown that evolving the log of the forward rates rather than the forward rates led to computational savings of around 20%. Here we extend the case of computing the sensitivities using the log Euler–Maruyama method under the spot LIBOR measure to the cross-currency LIBOR market model.

For simplicity, we assume that we are calculating the sensitivities of the coupon paid at T_i. In the situation where we have multiple coupon payments during each integration step; we view the product as a portfolio of individual coupon payments. To extend the pathwise method to Bermudan contracts we use a similar approach to that introduced by (Piterbarg, 2003). The first pass simulation and the construction of the of the exercise strategy, see Section 5, are left untouched. The second pass simulation, also see Section 5, now uses a modified product, which effectively contains several products, the underlying product and a product for each sensitivity which needs to be computed. For each path, the modified product evaluates the underlying product and stores the domestic and foreign forward rates along with the exchange rate at each step. We are also required to store, at each step, the partial derivatives of the product with respect to the domestic and foreign forward rates and the exchange rate. Once the decision to exercise has been made the stored data is used to compute the corresponding sensitivities as outlined below. Finally, the stored information is removed and a new path starts. What is surprising is that this can be achieved at little more than the computational time needed to compute one delta using finite differences, see Experiment 7 in Section 7 for relative timings.

In the cross-currency setting we are evolving the forward rates in the domestic and foreign economies and the exchange rate (via the forward exchange rate) using the Equations (23), (24) and (25). In this case, define the vector

$$X(t) = [f_1(t), f_2(t), \ldots, f_N(t), \widetilde{f}_1(t), \widetilde{f}_2(t), \ldots, \widetilde{f}_N(t), \mathrm{FX}(t)]^T \in \mathbf{R}^{2N+1},$$

and let $g : \mathbf{R}^{2N+1} \to \mathbf{R}$ be the discounted payoff of the financial contract. We start by discussing how to calculate deltas, before moving on to the more complicated case of vegas.

4.1 Deltas

We assume that we want to compute the deltas for all the forward rates, both domestic and foreign, and the spot exchange rate. That is, we want to calculate, for all j

$$\frac{\mathrm{dE}\left[g(X(T_i))\right]}{\mathrm{d}X_j(T_0)} = \mathrm{E}\left[\frac{\mathrm{d}g(X(T_i))}{\mathrm{d}X_j(T_0)}\right].$$

The pathwise method requires that the above equality holds. The conditions for this to be the case can be found in (Glasserman, 2004, Pg. 393-395). To compute the deltas we need to calculate

$$\frac{\mathrm{d}g(X(T_i))}{\mathrm{d}X_j(T_0)} = \sum_{i=1}^{2N+1} \frac{\partial g(X(T_i))}{\partial X_i(T_i)} \frac{\mathrm{d}X_i(T_i)}{\mathrm{d}X_j(T_0)}, \tag{26}$$

where

$$\begin{aligned} \frac{\mathrm{d}g(X(T_i))}{\mathrm{d}X(T_0)} &= \frac{\partial g(X(T_i))}{\partial X(T_i)} \frac{\mathrm{d}X(T_i)}{\mathrm{d}X(T_0)}, \\ &= \frac{\partial g(X(T_i))}{\partial X(T_i)} \frac{\partial X(T_i)}{\partial X(T_{i-1})} \frac{\partial X(T_{i-1})}{\partial X(T_{i-2})} \cdots \frac{\partial X(T_2)}{\partial X(T_1)} \frac{\mathrm{d}X(T_1)}{\mathrm{d}X(T_0)}. \end{aligned} \tag{27}$$

We note that in the spot LIBOR measure, the domestic forward rates at time T_k, see Equation (23), only depend on the domestic forward rates at time T_{k-1} and not the foreign forward rates or exchange rate at time T_{k-1}. This is also the case for the foreign forward rates, see Equation (24). This results in a significant simplification: the resulting Jacobians, excluding the last row, are block diagonal with blocks of size $N \times N$ and $N \times N$, that is

$$\frac{\partial X(T_k)}{\partial X(T_{k-1})} = \begin{bmatrix} \frac{\partial f(T_k)}{\partial f(T_{k-1})} & 0 & 0 \\ 0 & \frac{\partial \widetilde{f}(T_k)}{\partial \widetilde{f}(T_{k-1})} & 0 \\ \frac{\partial \mathrm{FX}(T_k)}{\partial f(T_{k-1})} & \frac{\partial \mathrm{FX}(T_k)}{\partial \widetilde{f}(T_{k-1})} & \frac{\partial \mathrm{FX}(T_k)}{\partial \mathrm{FX}(T_{k-1})} \end{bmatrix}. \tag{28}$$

The elements of the first block are then just the quantities from the single-currency LIBOR market model, and can be computed using the results from (Giles and Glasserman, 2006).

The elements of the Jacobian in the second block are found by differentiating Equation (24) with respect to $\widetilde{f}_j(T_{k-1})$, giving similar results as for the single-currency LIBOR market model. In particular,

$$\frac{\partial \widetilde{f}_i(T_k)}{\partial \widetilde{f}_j(T_{k-1})} = \begin{cases} 1, & i = j < k, \\ I_{i=j}\dfrac{\widetilde{f}_i(T_k) + \widetilde{\alpha}_i}{\widetilde{f}_i(T_{k-1}) + \widetilde{\alpha}_i} + \left(\widetilde{f}_i(T_k) + \widetilde{\alpha}_i\right)\dfrac{\partial \widetilde{\mu}_i^k}{\partial \widetilde{f}_j(T_{k-1})}, & i \geq j \geq k, \\ 0, & \text{otherwise}, \end{cases} \tag{29}$$

where the delta function $I_x = 1$ if x is true and 0 otherwise. The derivative of the foreign drifts for $i \geq j \geq k$ are

$$\frac{\partial \widetilde{\mu}_i^k}{\partial \widetilde{f}_j(T_{k-1})} = \frac{\tau_j(1 - \widetilde{\alpha}_j\tau_j)C^F_{kij}}{\left(1 + \tau_j\widetilde{f}_j(T_{k-1})\right)^2}.$$

Note that this structure implies that the sub-Jacobians are lower triangular. The elements in the last row of the Jacobian matrix are found by differentiating Equation (25) first with respect to $f(T_{k-1})$ and $\widetilde{f}(T_{k-1})$, giving

$$\frac{\partial \mathrm{FX}(T_k)}{\partial f_j(T_{k-1})} = I_{j=k-1}\frac{\tau_j \mathrm{FX}(T_k)}{1 + \tau_j f_j(T_{k-1})}, \quad \frac{\partial \mathrm{FX}(T_k)}{\partial \widetilde{f}_j(T_{k-1})} = I_{j=k-1}\frac{\tau_j \mathrm{FX}(T_k)}{1 + \tau_j f_j(T_{k-1})}, \tag{30}$$

then by differentiating Equation (25) with respect to $\mathrm{FX}(T_{k-1})$ giving

$$\frac{\partial \mathrm{FX}(T_k)}{\partial \mathrm{FX}(T_{k-1})} = \frac{\mathrm{FX}(T_k)}{\mathrm{FX}(T_{k-1})}. \tag{31}$$

A direct evaluation (from right to left) of Equation (27), requires calculating the product of i Jacobians. Given that the Jacobians are block diagonal, see Equation (28), this is an $O(N^3)$ operation. However, as was pointed out by (Giles and Glasserman, 2006), if the computations are performed from left to right, then we need to take the product of, in the cross-currency case, a $2N + 1$ vector and a Jacobian i times. Again using the structure of the Jacobians, see Equation (28), this is an $O(N^2)$ operation. So, rather than calculate Equation (27) from right to left as we move forward in our simulation, we can store the Jacobians on a path-by-path basis, and then evaluate Equation (28) from left to right once we reach T_i. In fact given the results of Equations (29), (30) and (31) we only need to store the domestic and foreign forward rates and the exchange rates for a single path, from which we can compute all the required Jacobians. This only requires modest storage space.

Following (Giles and Glasserman, 2006), we can do even better using the special structure of the equations in the cross-currency LIBOR market model, needing only $O(NF)$ operations per step. Set the $2N + 1$ vector

$$D(T_i) = \frac{\partial g(X(T_i))}{\partial X(T_i)},$$

and define

$$D(T_k) = D(T_i)\frac{\partial X(T_i)}{\partial X(T_{i-1})} \cdots \frac{\partial X(T_{k+1})}{\partial X(T_k)}.$$

The cross-currency adjoint method requires us to compute

$$D(T_{k-1}) = D(T_k)\frac{\partial X(T_k)}{\partial X(T_{k-1})}.$$

Using the fact that the Jacobian is block diagonal (see Equation (28)) and the expressions for these sub Jacobians given in (Giles and Glasserman, 2006), Equations (29), (30) and (31), the cross-currency adjoint method requires us to compute, for $i = k+1, \ldots, N$,

$$\begin{aligned}
D_i(T_{k-1}) &= D_i(T_k)\frac{f_i(T_k)+\alpha_i}{f_i(T_{k-1})+\alpha_i} + D_{2N+1}(T_k)I_{i=k-1}\frac{\tau_{k-1}\mathrm{FX}(T_k)}{1+\tau_{k-1}f_{k-1}(T_{k-1})} \\
&\quad + \frac{\tau_i(1-\alpha_i\tau_i)}{(1+\tau_i f_i(T_{k-1}))^2}\sum_{j=i}^{N} D_j(T_k)\left(f_j(T_k)+\alpha_j\right)C_{kij}, \\
D_{N+i}(T_{k-1}) &= D_{N+i}(T_k)\frac{\widetilde{f}_i(T_k)+\widetilde{\alpha}_i}{\widetilde{f}_i(T_{k-1})+\widetilde{\alpha}_i} + D_{2N+1}(T_k)I_{i=k-1}\frac{\tau_{k-1}\mathrm{FX}(T_k)}{1+\tau_{k-1}\widetilde{f}_{k-1}(T_{k-1})} \\
&\quad + \frac{\tau_i(1-\widetilde{\alpha}_i\tau_i)}{\left(1+\tau_i\widetilde{f}_i(T_{k-1})\right)^2}\sum_{j=i}^{N} D_{N+j}(T_k)\left(\widetilde{f}_j(T_k)+\widetilde{\alpha}_j\right)C_{kij}, \\
D_{2N+1}(T_{k-1}) &= D_{2N+1}(T_k)\frac{\mathrm{FX}(T_k)}{\mathrm{FX}(T_{k-1})}.
\end{aligned}$$

By writing $C_{kij} = \sum_{n=1}^{F} A_{kin}A_{kjn}$, and swapping the order of summation, we get

$$\begin{aligned}
D_i(T_{k-1}) &= D_i(T_k)\frac{f_i(T_k)+\alpha_i}{f_i(T_{k-1})+\alpha_i} + D_{2N+1}(T_k)I_{i=k-1}\frac{\tau_{k-1}\mathrm{FX}(T_k)}{1+\tau_{k-1}f_{k-1}(T_{k-1})} \\
&\quad + \frac{\tau_i(1-\alpha_i\tau_i)}{(1+\tau_i f_i(T_{k-1}))^2}\sum_{n=1}^{F} A_{kin}\sum_{j=i}^{N} D_j(T_k)\left(f_j(T_k)+\alpha_j\right)A_{kjn}, \\
D_{N+i}(T_{k-1}) &= D_{N+i}(T_k)\frac{\widetilde{f}_i(T_k)+\widetilde{\alpha}_i}{\widetilde{f}_i(T_{k-1})+\widetilde{\alpha}_i} + D_{2N+1}(T_k)I_{i=k-1}\frac{\tau_{k-1}\mathrm{FX}(T_k)}{1+\tau_{k-1}\widetilde{f}_{k-1}(T_{k-1})} \\
&\quad + \frac{\tau_i(1-\widetilde{\alpha}_i\tau_i)}{\left(1+\tau_i\widetilde{f}_i(T_{k-1})\right)^2}\sum_{n=1}^{F} A_{kin}\sum_{j=i}^{N} D_{N+j}(T_k)\left(\widetilde{f}_j(T_k)+\widetilde{\alpha}_j\right)A_{kjn}, \\
D_{2N+1}(T_{k-1}) &= D_{2N+1}(T_k)\frac{\mathrm{FX}(T_k)}{\mathrm{FX}(T_{k-1})}.
\end{aligned}$$

Since the inner summations in the above expressions do not depend on i, it is possible to calculate $D(T_{k-1})$ with $O(NF)$ computations. We note that it is possible to extend these results to methods other than the log Euler–Maruyama method following the results presented for the single currency case in (Denson and Joshi, 2009a). In the theory of automatic differentiation it was shown, see (Griewank and Walther, 2008), that the adjoint calculations can be computed with no more than four times the number of operations used to compute the original algorithm.

4.2 Vegas

We now turn our attention to the calculation of vegas. To compute the vegas we need to simulate

$$\frac{\mathrm{d}g(X(T_i))}{\mathrm{d}\theta} = \frac{\partial g(X(T_i))}{\partial X(T_i)}\frac{\mathrm{d}X(T_i)}{\mathrm{d}\theta}, \tag{32}$$

for some volatility parameter θ. It is advantageous to break the vegas up into smaller parts, commonly referred to as elementary vegas. Elementary vegas are sensitivities with respect to a single element of a pseudo square root matrix used in the simulation, denoted by A_{klm}. To avoid trivialities, we assume $i \geq k$. Similarly to deltas,

$$\frac{\mathrm{d}g(X(T_i))}{\mathrm{d}A_{klm}} = \frac{\partial g(X(T_i))}{\partial X(T_i)}\frac{\mathrm{d}X(T_i)}{\mathrm{d}A_{klm}},$$

where,

$$\frac{\mathrm{d}X(T_i)}{\mathrm{d}A_{klm}} = \frac{\partial X(T_i)}{\partial X(T_{i-1})}\frac{\partial X(T_{i-1})}{\partial X(T_{i-2})}\cdots\frac{\partial X(T_{k+1})}{\partial X(T_k)}\frac{\mathrm{d}X(T_k)}{\mathrm{d}A_{klm}}.$$

However, we only need the Jacobians up to the relevant integration interval, in this case $[T_{k-1}, T_k]$, since dependence on the particular pseudo square root element does not enter before this point. Given that we have already computed the Jacobians for evaluating the deltas, the additional work revolves around calculating

$$\frac{\mathrm{d}X(T_k)}{\mathrm{d}A_{klm}}.$$

We can split this into three possible cases, depending on the row of the particular pseudo square root element. First, if $l \leq N$, then the pseudo square-root element comes from the domestic portion of the matrix and

$$\frac{\mathrm{d}X_j(T_k)}{\mathrm{d}A_{klm}} = 0,$$

for $j > N$. For $j \leq N$, we obtain the same equations that apply to the single-currency LIBOR market model given in (Giles and Glasserman, 2006; Denson and Joshi, 2009a).

Second, if $N < l \leq 2N$, then the pseudo square-root element comes from the foreign forward rate portion of the matrix and

$$\frac{\mathrm{d}X_j(T_k)}{\mathrm{d}A_{klm}} = 0,$$

for $j \leq N$ and $j = 2N + 1$. Otherwise, we obtain equivalent equations to the single-currency case. In particular,

$$\frac{\mathrm{d}\widetilde{f}_j(T_k)}{\mathrm{d}A^F_{klm}} = (\widetilde{f}_j(T_k) + \widetilde{\alpha}_j)\frac{\mathrm{d}}{\mathrm{d}A^F_{klm}}\left(\widetilde{\mu}^k_j - \frac{1}{2}C^F_{kjj} + \sum_{f=1}^{F} A^F_{kjf}Z_{kf}\right).$$

where,

$$\frac{\mathrm{d}\widetilde{\mu}^k_j}{\mathrm{d}A^F_{klm}} = \begin{cases} A^F_{kjm}\widetilde{h}_l(T_{k-1}), & j > l, \\ \sum_{f=k}^{i} A^F_{kfm}\widetilde{h}_f(T_{k-1}) + A_{kjm}\widetilde{h}_j(T_{k-1}) + A^X_{km}, & j = l, \\ 0, & \text{otherwise,} \end{cases}$$

and,

$$\frac{\mathrm{d}C^F_{kjj}}{\mathrm{d}A^F_{klm}} = \frac{\mathrm{d}}{\mathrm{d}A^F_{klm}}\sum_{f=1}^{F} A^F_{kjf}A^F_{kjf} = \begin{cases} 2A^F_{kjm}, & j = l, \\ 0 & \text{otherwise,} \end{cases}$$

and, finally,

$$\frac{\mathrm{d}}{\mathrm{d}A^F_{klm}}\sum_{f=1}^{F} A^F_{kjf}Z_{kf} = \begin{cases} Z_{km}, & j = l, \\ 0, & \text{otherwise.} \end{cases}$$

Finally, for the case where $l = 2N + 1$, we have

$$\frac{\mathrm{d}X_j(T_k)}{\mathrm{d}A_{klm}} = 0,$$

for $j \leq N$. The derivative of the foreign forward rates with respect to an element of the exchange pseudo root vector A^X_k is

$$\frac{\mathrm{d}\widetilde{f}_j(T_k)}{\mathrm{d}A^X_{kl}} = -(\widetilde{f}_j(T_k) + \alpha_j)\begin{cases} A^F_{kjl}, & j = l, \\ 0, & \text{otherwise.} \end{cases}$$

The final quantity that we need to compute is the derivative of the exchange rate with respect to an element of the exchange pseudo root matrix A^X_k, which is given by

$$\frac{\mathrm{dFFX}(T_k)}{\mathrm{d}A^X_{kl}} = \mathrm{FX}(T_k)\left(Z_{kl} - A^X_{kl}\right).$$

Given the adjoint method for calculating deltas, the corresponding adjoint method for computing the elementary vegas requires an inner product of the delta $Y(T_k)$ and $\frac{\mathrm{d}X(T_k)}{\mathrm{d}A_{klm}}$. Evaluating this inner product is straight forward.

We now briefly discuss some implementation details concerning the Greek calculations. In our experiments we have implemented the Greek calculations as if they were an actual product, given that they share many of the features of the underlying product. At each step we do not expect a single cashflow but a cashflow for each Greek and the price. To compute the cashflows for each Greek we must store the domestic forward rates, foreign forward rates and the exchange rate at each timestep so that they can be accessed to compute the updated $D(T_{k-1})$. The derivative of the underlying products discounted cashflow with respect to domestic forward rates, foreign forward rates and the exchange rate, must also be computed.

Once we have calculated the elementary vegas, we can use the procedure recently introduced by (Joshi and Kwon, 2010) to calculate sensitivities with respect to market observable volatility parameters. The most significant difference compared to the single-currency LIBOR market model is that there are significantly more Greeks in the cross-currency LIBOR market model.

5 Improving lower and upper bounds

A lot of recent work has been undertaken to produce tight unbiased lower bounds for Bermudan style derivatives using Monte Carlo simulation. The results from this section are based on the article (Beveridge and Joshi, 2009). We describe them here for completeness and outline any modifications needed in the cross-currency setting.

We will consider a cancellable product, which involves a series of cashflows $\mathrm{CF}(T_i)$ at each tenor date T_i, until the time of exercise. For ease of exposition, we assume that the product can be exercised at each tenor date, with the extension to more complicated exercise structures simple. We also assume that exercise does not incur a penalty, that is no rebate is paid upon cancellation. Let $C(T_i)$ denote the cashflows paid at tenor date T_i transformed into amounts of the numéraire, that is $C(T_i) = \mathrm{CF}(T_i)/N(T_i)$.

At time T_i, the continuation value of the cancellable product, $V(T_i)$, is given by

$$\frac{V(T_i)}{N(T_i)} = \sup_{\gamma \in \Gamma_{i+1}} \mathrm{E}_i \left[\sum_{j=i+1}^{\gamma} C(T_j) \right], \tag{33}$$

where Γ_i denotes the set of stopping times taking values in $\{i, \ldots, N+1\}$ and $\mathrm{E}_i[\cdot] = \mathrm{E}[\cdot|\mathcal{F}_{T_i}]$ is shorthand for conditional expectations taken in the equivalent martingale measure associated with using $N(T_i)$ as numéraire. This expression gives the time zero value by setting $i = 0$. Note that we assume if the product has not been exercise previously, it must be exercised at T_{N+1}, that is, after all cash-flows have been received. This ensures that we have finite stopping times as exercise strategies.

We consider the position of the issuer of the product, who we assume holds the right to cancel the product. The issuer receives the floating LIBOR rate and pays some complicated coupon. For example, under PRDC swaps, each coupon paid by the issuer is a call option on the exchange rate.

5.1 Least squares regression

In the papers (Carrière, 1996; Tsitsiklis and van Roy, 2001; Longstaff and Schwartz, 2001) simple, but elegant, regression arguments were developed for pricing callable early exercisable options. In our discussion we follow the most widely used approach of Longstaff–Schwartz which has become the industry standard. We now briefly describe their approach, applied directly to the cancellable product case as suggested by (Amin, 2003). To obtain an unbiased lower bound estimate a three step process has become the standard practice.

The first step, often called the first pass, is where the necessary data to develop an approximate exercise strategy is collected. In the first pass, a Monte Carlo simulation is performed with a certain number of first pass paths, say $N_1 \in \mathbf{N}$. The forward rates and exchange rate are evolved to every tenor date in the term structure. We store the discounted cashflows realized on each path of the simulation at T_i, which we denote by $C_j(T_i)$ for $j = 1, \ldots, N_1$. In the recording of the cashflows, it is convenient to identify whether the cashflows are paid on the date they are realized or more commonly paid at the next tenor date. Along with the cashflows received at T_i, we also record a column vector of explanatory variables, $B_j(T_i)$, for $j = 1, \ldots, N_1$. The explanatory variables that we record at T_i will be used to estimate the continuation value at that time. Note that in the most common situation where there is a timelag in the payment of the cashflows, the explanatory variables, $B_j(T_i)$, and the cashflows, $C_j(T_{i+1})$, are known at time T_i. The explanatory variables that we will record at each tenor date T_i, for $i = 0, \dots, N-1$, are

$$B_j(T_i) = \left[\begin{array}{ccccc} f_i(T_i) & \widetilde{f}_i(T_i) & S_{i+1}(T_i) & \widetilde{S}_{i+1}(T_i) & \mathrm{FX}(T_i) \end{array}\right]^T, \tag{34}$$

where the domestic and foreign swap rates are defined, for $j \geq i$, as

$$S_j(T_i) = \frac{P_j(T_i) - P_{N+1}(T_i)}{A_j(T_i)}, \quad \widetilde{S}_j(T_i) = \frac{\widetilde{P}_j(T_i) - \widetilde{P}_{N+1}(T_i)}{\widetilde{A}_j(T_i)}, \tag{35}$$

with the domestic and foreign annuities given by

$$A_j(T_i) = \sum_{k=j}^{N} \tau_k P_{k+1}(T_i), \quad \widetilde{A}_j(T_i) = \sum_{k=j}^{N} \tau_k \widetilde{P}_{k+1}(T_i).$$

The explanatory variables that we record at the final tenor date, T_N, are

$$B_j(T_N) = \left[\begin{array}{ccc} f_N(T_N) & \widetilde{f}_N(T_N) & \mathrm{FX}(T_N) \end{array}\right]^T. \tag{36}$$

In the second step we use the data collected in the first pass to build an exercise strategy. To make the explanation clear we will assume that we have a natural timelag; that is, the cashflows paid at time T_{i+1} are known at time T_i. The case where there is no timelag follows with small changes in the details that follow. Assume that we are at time T_N. If we do not exercise at time T_N, then at time T_{N+1} we will receive, for each path, a cashflow $C_j(T_{N+1})$. We define our pathwise observations of the discounted continuation value at time T_N as $V_j(T_N) = C_j(T_{N+1})$. At T_N, we have N_1 stored realizations of the explanatory variables, $B_j(T_N)$, given by Equation (36). We now want to construct a quadratic polynomial with the variables given by the basis functions, which can be used to approximate the continuation value. That is, we want to compute the regression coefficients, $\alpha(T_N)$, which minimize

$$\sum_{j=1}^{N_1} (V_j(T_N) - \alpha(T_N) \cdot q(B_j(T_N)))^2,$$

where $q : \mathbf{R}^n \to \mathbf{R}^{\frac{(n+1)(n+2)}{2}}$, is the vector formed with elements 1, the n basis variables and the $n(n+1)/2$ products of the basis variables. We now collect the appropriate information needed to compute the regression coefficients at T_N as follows,

$$Q(T_N) = \begin{bmatrix} q(B_1(T_N))^T \\ \vdots \\ q(B_{N_1}(T_N))^T \end{bmatrix}, \qquad U(T_N) = \begin{bmatrix} V_1(T_N) \\ \vdots \\ V_{N_1}(T_N) \end{bmatrix}.$$

The regression coefficients at T_N satisfying the above equation are well known to have the form

$$\alpha(T_N) = \left(Q(T_N)Q(T_N)^T\right)^{-1} Q(T_N)^T U(T_N). \tag{37}$$

Now that we have the regression coefficients at time T_N, we can determine our approximate exercise strategy on each path. To do so we compute the estimated continuation value by taking the inner product of the regression coefficients and the relevant basis functions via

$$E_j(T_N) = \alpha(T_N) \cdot q(B_j(T_N)). \tag{38}$$

Now, for each path we update our pathwise observations of the continuation value, to give observations of the continuation value from the next point backwards in time. In particular,

$$V_j(T_{N-1}) = \begin{cases} C_j(T_N), & \text{if } E_j(T_N) < 0, \\ V_j(T_N) + C_j(T_N), & \text{if } E_j(T_N) > 0. \end{cases} \tag{39}$$

We now cycle through the above procedure for each tenor date from T_k for $k = N-1, \ldots, 0$. At T_0 we stop once we compute the regression coefficients. An initial biased estimate of the products price is then given by

$$\frac{V(T_0)}{N(T_0)} \approx \frac{1}{N_1} \sum_{j=1}^{N_1} V_j(T_0). \tag{40}$$

The bias is due to the fact that information ahead of the current exercise time is implicitly used to evaluate the exercise decision, introducing foresight bias, see the article by (Fries, 2008) for an in-depth look at foresight bias. Note that when products have cash-flows with natural time lags, we can calculate the continuation value exactly at T_N, and therefore do not need to perform a regression. However, we have described performing a regression at T_N to keep the treatment general.

The third and final step, often called the second pass or pricing pass, is the most straight forward of the three steps. In the second pass a Monte Carlo simulation is used with $N_2 \in \mathbf{N}$ paths. As a rough rule of thumb, the number of second pass paths should be double the number of first pass paths. The estimated continuation value at time T_k is then computed by taking the inner product of the regression coefficients and the relevant basis functions via

$$E_j(T_k) = \alpha(T_k) \cdot q(B_j(T_k)). \tag{41}$$

The forward rates are evolved up until the first exercise date T_1, the estimated continuation value $E_j(T_1)$ is then compared to the amount received upon exercise, which is zero in our case. So if the estimated continuation value is negative the product is exercised and we start evolving the next path, otherwise we evolve the forward rates to the next exercise date T_2 and compare whether the estimated continuation value $E_j(T_2)$ is positive or negative. This process is repeated until the product is exercised or the last exercise date is reached, which we have denoted as T_{M_j} for the j:th path. The product value is then estimated as

$$\frac{V(T_0)}{N(T_0)} \approx \frac{1}{N_2} \sum_{j=1}^{N_2} \sum_{k=1}^{M_j} C_j(T_k). \tag{42}$$

5.2 Double regression

In the least squares regression procedure outlined above, a vector of regression coefficients is generated at each tenor date. In the second pass, at each exercise date we have to make the decision whether to exercise or not. Generally, this decision is correct if the estimated continuation value is not close to zero. That is, the option seems to be deeply in or out of the money. The decision is less clear if the estimated continuation value is close to zero. In (Broadie and Cao, 2008; Beveridge and Joshi, 2009) it was suggested that a second regression be performed only including points were the estimated continuation value is close to zero. This was done by modifying the procedure described in Section 5.1. At each tenor date T_i, compute the regression coefficients using Equation (37), then use these regression coefficients to estimate the continuation value using Equation (38). Now for each path $j \in \{1, \ldots, N_1\}$ where $|E_j(T_i)| < \delta$, record the corresponding basis functions in the matrices $\widetilde{Q}(T_i)$ and the observed discounted continuation values in the matrix $\tilde{U}(T_i)$, that is

$$\widetilde{Q}(T_i) = \begin{bmatrix} q(B_1(T_i))^T \\ \vdots \\ q(B_{\widetilde{N}_1}(T_i))^T \end{bmatrix}, \qquad \tilde{U}(T_i) = \begin{bmatrix} V_1(T_i) \\ \vdots \\ V_{\widetilde{N}_1}(T_i) \end{bmatrix}.$$

Here $\widetilde{N}_1 \leq N_1$ is the number of paths satisfying $|E_j(T_i)| < \delta$. The second set of regression coefficients at time T_i are then computed using

$$\widetilde{\alpha}(T_i) = \left(\widetilde{Q}(T_i)\widetilde{Q}(T_i)^T\right)^{-1} \widetilde{Q}(T_i)^T \tilde{U}(T_i). \tag{43}$$

Now recompute the estimated continuation value at time T_i as

$$\widetilde{E}_j(T_i) = \widetilde{\alpha}(T_i) \cdot q(B_j(T_i)). \tag{44}$$

The observed continuation value at time T_{i-1} is then updated using the estimated continuation values computed in Equations (38) and (44), via

$$V_j(T_{i-1}) = \begin{cases} C_j(T_i), & \text{if } E_j(T_N) < -\delta, \\ V_j(T_i) + C_j(T_i), & \text{if } E_j(T_N) > \delta, \\ C_j(T_i), & \text{if } \widetilde{E}_j(T_i) < 0, \\ V_j(T_i) + C_j(T_i), & \text{if } \widetilde{E}_j(T_i) > 0. \end{cases} \tag{45}$$

5.3 Exclusion of suboptimal points

The exclusion of suboptimal points was first suggested in the paper by (Longstaff and Schwartz, 2001) and was extended to apply to cancellable exotic interest products by (Beveridge and Joshi, 2008). The continuation value at exercise time T_i, is given by

$$\begin{aligned} \frac{V(T_i)}{N(T_i)} &= \sup_{\gamma \in \Gamma_{i+1}} \mathrm{E}_i \left[\sum_{j=i+1}^{\gamma} C(T_j) \right], \\ &\geq \max_{r \in \{i+1,\ldots,N+1\}} \mathrm{E}_i \left[\sum_{j=i+1}^{r} C(T_j) \right] \geq \mathrm{E}_i \left[C(T_{i+1}) \right]. \end{aligned}$$

Recall that we are assuming that upon exercise we do not receive a rebate. In this case, we should never exercise if the expected value of the discounted cashflows at time T_{i+1}, given our information at time T_i, is positive, that is $\mathrm{E}_i[C(T_{i+1})] > 0$. This follows from the above equation, since we know the true continuation value is at least as big as $\mathrm{E}_i[C(T_{i+1})]$. In practice, it is often the case that the coupons are paid with a natural timelag, that is the cashflows are determined at T_i but paid at time T_{i+1}. In this case, we know that if cashflows at T_{i+1} will be positive it is suboptimal to exercise at time T_i and the optimal strategy would never exercise at this point.

Excluding suboptimal points affects all three steps in the least squares regression. In the first pass, we need only to record the explanatory variables for points which are not suboptimal. In the implementation we also include a boolean variable which determines whether the point is suboptimal or not. During the second step, the regression coefficients are calculated using Equation (37), where $Q(T_i)$ and $U(T_i)$ only include basis functions and continuation values respectively for points which are not suboptimal. During the second pass, the product is never exercised at suboptimal points.

The use of suboptimal points can lead to improvements in accuracy, since we limit the domain over where we need to fit the continuation value, and in computational time by avoiding unnecessary calculations.

5.4 Adaptive basis functions

Until recently, the choice of basis functions that are used in the least squares regression process generally had to be tailored to the particular product being priced, and usually required a significant amount of testing before a reasonable choice of basis functions could be finalized. In the recent article (Beveridge and Joshi, 2009), an approach is proposed to remove a lot of the hand-crafting needed in constructing the basis functions. To use the approach, a base set of explanatory variables must be specified: our base set will be the basis functions given in Equations (34) and (36). In addition, an extra set of explanatory variables must also be specified: we include all the remaining domestic bonds. That is, at time T_i, for $i = 0, \ldots, N$, our additional explanatory variables are

$$A_j(T_i) = \begin{bmatrix} P_{i+1}(T_i) & \cdots & P_{N+1}(T_i) \end{bmatrix}^T.$$

Now we must choose how many of the additional basis variables can be used in the regression: let's say that we allow at most m. The value of m could potentially vary depending on at what time T_i we are at. However, we have found that it is generally sufficient to choose $m = 1$, which we do in the experiments reported in Section 7. Let $A_j^{\ell}(T_i)$ be a sub-vector of $A_j(T_i)$ with at most m entries. There are $\binom{N-i+1}{m}$ such subvectors. At each tenor date T_i, perform the calculations given in Equation (37) $\binom{N-i+1}{m}$ times, where the matrix $Q(T_i)$

is replaced by

$$Q^{\ell}(T_k) = \begin{bmatrix} q(B_1(T_k), A^{\ell}(T_k))^T \\ \vdots \\ q(B_{N_1}(T_k), A^{\ell}(T_k))^T \end{bmatrix}.$$

For each choice of regression coefficients, compute the adjusted R^2 value, given by

$$R^2 = 1 - \left(\frac{\text{SSE}}{\text{SST}}\right)\frac{k-1}{k-\ell-1},$$

where SSE is the sum of the squared errors and SST is the total sum of the squares. Also, k is the number of points included in the regression. If suboptimal points are not excluded then $k \in \{N_1, \widetilde{N}_1\}$ and ℓ is the number of basis functions. Choose the set of basis functions at each tenor date T_k, which maximizes the adjusted R^2 value.

5.5 Delta hedge control variate

It is well known that in complete markets, such as the cross-currency LIBOR market model, an option's payoff can be perfectly replicated using delta hedging if we can compute the deltas exactly and trade continuously. During a simulation of the cross-currency LIBOR market model, neither of these are possible. However, by following the approach outlined by (Beveridge and Joshi, 2009) in the single currency setting, it is possible to estimate deltas quickly and easily using the regression based continuation value estimates, and these can be used to obtain significant variance reduction. In particular, the estimated continuation value at time T_i can be expressed as a function of the fundamental tradeable assets, the domestic and foreign zero coupon bonds yet to reset. That is,

$$E_j(T_i) = f\left(P_{i+1}(T_i), \ldots, P_{N+1}(T_i), \widetilde{P}_{i+1}(T_i), \ldots, \widetilde{P}_{N+1}(T_i)\right).$$

We can then compute the partial derivatives of the estimated continuation value with respect to the domestic and foreign zero coupon bonds,

$$\frac{\partial E_j(T_i)}{\partial P_k(T_i)}, \quad \frac{\partial E_j(T_i)}{\partial \widetilde{P}_k(T_i)},$$

for $k = i+1, \ldots, N+1$. We now discuss how to compute an approximation to the replication portfolio which we will use as a control variate. Note that the replication is more accurate the more frequently the control variate portfolio is updated. We therefore choose to evaluate the portfolio at each tenor date, rather than at the exercise dates exclusively. We set up our replicating portfolio so that we hold delta units of each asset across each step in the simulation

$$R_B^{\ell}(t) = \sum_{k=\ell+1}^{N+1}\left(\frac{\partial E_j(T_\ell)}{\partial P_k(T_\ell)}P_k(t) + \frac{\partial E_j(T_\ell)}{\partial \widetilde{P}_k(T_\ell)}\text{FX}(t)\widetilde{P}_k(t)\right).$$

with all additional cash invested in the numéraire asset

$$R_N^{\ell}(t) = -\frac{R_B^{\ell}(T_\ell)}{N(T_\ell)}N(t).$$

Then the overall portfolio, at time T_i, is given by

$$R(T_i) = R(T_{i-1}) + R_B^{i-1}(T_i) + R_N^{i-1}(T_i) = \sum_{j=0}^{i-1}(R_B^j(T_{j+1}) + R_N^j(T_{j+1})).$$

Note that, at time T_0, the overall portfolio $R(T_0) = 0$. Substituting the expressions for the replicating portfolio and the numéraire portfolio into the expression for the overall portfolio gives

$$R(T_i) = R(T_{i-1}) + \sum_{k=i}^{N+1}\left(\frac{\partial E_j(T_{i-1})}{\partial P_k(T_{i-1})}X_k(T_i) + \frac{\partial E_j(T_{i-1})}{\partial \widetilde{P}_k(T_{i-1})}\widetilde{X}_k(T_i)\right),$$

where

$$X_k(T_i) = \frac{P_k(T_i)}{N(T_i)} - \frac{P_k(T_{i-1})}{N(T_{i-1})}, \quad \widetilde{X}_k(T_i) = \frac{\widetilde{P}_k(T_i)}{N(T_i)}\mathrm{FX}(T_i) - \frac{\widetilde{P}_k(T_{i-1})}{N(T_{i-1})}\mathrm{FX}(T_{i-1}).$$

In our case where the numéraire is given by Equation (11), it follows using Equation (1) that $X_i(T_i) = 0$ for $i = 1, \ldots, N+1$. This is important since at time T_i there are $2(N-i)+1$ sources of risk, that is the $N-i$ domestic forward rates and $N-i$ foreign forward rates yet to reset and the forward exchange rate. To hedge this risk there is exactly $2(N-i)+1$ non-numéraire bonds.

Now to value the product we use, instead of Equation (42), the adjusted value

$$\frac{V(T_0)}{N(T_0)} \approx \frac{1}{N_2} \sum_{j=1}^{N_2} \left(\sum_{k=1}^{M} C_j(T_k) - R_j(T_M) \right), \tag{46}$$

where $R_j(T_M)$ is the j:th realization of the portfolio $R(T_M)$. Given that the deflated portfolio $R(T_M)$ is a martingale the control variate is unbiased. One point of practical interest is that we need an estimate of the continuation value at time T_0 as a function of the domestic and foreign bonds at time T_0. Given that the forward rates are inputs the bonds are deterministic, therefore it is not possible to compute an estimate of the continuation value using least squares regression. To get around this problem we randomize the initial forward rates as follows

$$f_{ji}(T_0) = f_i(T_0) \exp\left(-\frac{1}{2}a^2 + aZ_{ji} \right),$$

where $Z_{ij} \sim N(0,1)$ and $f_{ji}(T_0)$ is the i:th forward rate on j:th path. The estimated continuation value, at time T_0, can now be computed using least squares regression. The deltas can now be computed by differentiating the estimated continuation value with respect to the domestic and foreign bonds at time T_0. In our experiments we have chosen the parameter $a = 0.3$.

5.6 Upper bounds

The methods discussed above relate to calculating lower bounds for Bermudan derivatives. However, one can never be sure of the accuracy of a particular lower bound methodology until we know an upper bound for the price of the option. Here, we discuss how upper bounds can be calculated, focusing on the duality result of (Rogers, 2002; Haugh and Kogan, 2004), and the generic and accurate method from (Andersen and Broadie, 2004) for implementation. See also (Joshi, 2006) for the extension to cancellable contracts, and (Joshi, 2007) for improvements.

First, consider the duality result of (Rogers, 2002; Haugh and Kogan, 2004). Suppose we have a martingale, $M(t)$, with $M(T_0) = 0$. Then, from Equation (33),

$$\begin{aligned}
\frac{V(T_0)}{N(T_0)} &= \sup_{\gamma \in \Gamma_1} \mathrm{E}_0 \left[\sum_{j=1}^{\gamma} C(T_j) \right], \\
&= \sup_{\gamma \in \Gamma_1} \mathrm{E}_0 \left[\sum_{j=1}^{\gamma} C(T_j) - M(T_\gamma) \right], \\
&\leq \sup_{\gamma \in \Omega_1} \mathrm{E}_0 \left[\sum_{j=1}^{\gamma} C(T_j) - M(T_\gamma) \right], \\
&= \mathrm{E}_0 \left[\max_{k \in \{1,\ldots,N+1\}} \left(\sum_{j=1}^{k} C(T_j) - M(T_k) \right) \right],
\end{aligned} \tag{47}$$

where Ω_j denotes the set of random times taking values in $\{j, j+1, \ldots, N+1\}$. In addition, it can be shown that when M is chosen optimally, it is possible to obtain equality in the above expression; see, for example, (Rogers, 2002; Joshi, 2006).

Equation (47) gives us a way of calculating upper bounds. However, to use this result we have to develop an approximation to the optimal martingale. An attractive way of doing this is introduced by (Andersen and

Broadie, 2004). Financially, the optimal martingale is given by the optimal hedge to Bermudan derivatives, which is obtained by holding the underlying product where, if the exercise strategy says exercise, the product is exercised and re-purchased with one less exercise date and any residual cash is pocketed; see, for example, (Joshi, 2006). In particular, the Andersen–Broadie approach is to use the underlying product exercised according to an approximate exercise strategy in place of the optimal hedge.

This has the significant advantage that no further optimizations beyond those required for the least squares regression method are necessary to develop the optimal martingale; once we have our least squares regression exercise strategy, we immediately have the associated approximation to the optimal martingale. The downside is that to value the hedge portfolio, we need to value the underlying product exercised according to the given strategy, and this generally requires sub-Monte Carlo simulations.

Mathematically, if $\widehat{M}_i$ denotes the value of the approximate hedge, described in (Andersen and Broadie, 2004), at T_i, then

$$\widehat{M}_i = \sum_{j=1}^{i} (\widehat{Y}_j + C_j - Z_{j-1}),$$

where

$$\begin{aligned} Z_{j-1} &= \widehat{\mathrm{E}}_{j-1}(\widetilde{Y}_j + C_j), \\ \widetilde{Y}_j &= (\widetilde{Y}_{j+1} + C_{j+1}) I_{A_j^{\mathrm{C}}}, \\ \widehat{Y}_j &= Z_j I_{A_j^{\mathrm{C}}}, \end{aligned}$$

with A_j denoting the event that exercise occurs according to the least squares regression strategy at T_j and Z_j denoting the Monte Carlo estimate of the T_j continuation value using the least squares regression strategy.

It is shown in (Andersen and Broadie, 2004; Joshi, 2006) that the Monte Carlo error of the sub-simulations adds to the upward bias, so the Andersen–Broadie method can be used to successfully obtain upper bounds. We will use this method in the next section to show that it is possible to obtain very tight lower bounds for exotic cross-currency interest rate derivatives.

6 Cross-currency correlation, covariance and calibration

In this section we will discuss the correlation, covariance and calibration that we have used in our experiments in Section 7. We also refer the reader to (Brace, 2008, §14), which appears to be the only place in the literature where these issues in the cross-currency setting have been discussed. In Section 3 we developed a numerical method for the cross-currency LIBOR market model, which relied on obtaining the correct covariance structure. The covariance structure was completely determined by the deterministic functions $\lambda_i(t)$, $\widetilde{\lambda}_i(t)$ and $\widehat{\lambda}_\eta(t)$. To ensure consistency with standard single currency approaches, we choose to define the deterministic functions, $\sigma_i(t), \widetilde{\sigma}_i(t) \in \mathbf{R}$, and the deterministic vector functions $\beta_i(t), \widetilde{\beta}_i(t) \in \mathbf{R}^F$ for $0 \le t \le T_i$, such that $\lambda_i(t) = \sigma_i(t)\beta_i(t)$ and $\widetilde{\lambda}_i(t) = \widetilde{\sigma}_i(t)\widetilde{\beta}_i(t)$. This formulation is useful as it decomposes the shocks into two parts. The first part $\sigma_i(t)$ or $\widetilde{\sigma}_i(t)$ is the deterministic volatility associated with the i:th forward rate, which must be chosen specifically to ensure that the caplets and or swaptions are priced correctly in the domestic or foreign currency. See the discussion at the end of this section, where we show how to do this. The second part $\beta_i(t)$ or $\widetilde{\beta}_i(t)$ links the correlation structure with the Weiner processes. That is $\beta_i(t) \cdot \beta_j(t) = \rho_{ij}^D(t)$ and $\widetilde{\beta}_i(t) \cdot \widetilde{\beta}_j(t) = \rho_{ij}^F(t)$. Given that only one of the forward exchange rates can have deterministic volatility, we have chosen to model the forward exchange rate which will reset next. In this case we assume, similar to the arguments above that $\widehat{\lambda}_k(t) = \widehat{\sigma}_k(t)\beta_k(t)$, for $t \in [T_{k-1}, T_k]$. In this case $\widehat{\sigma}_k(t)$ will be chosen to ensure that the implied volatilities of the forward foreign exchange option are recovered. This is discussed in further detail later in the section see Equation (50). To summarize we have made the following assumptions

$$\begin{aligned} \lambda_i(t) \cdot \mathrm{d}W(t) &= \sigma_i(t)\beta_i(t) \cdot \mathrm{d}W(t) = \sigma_i(t)\mathrm{d}B_i(t), \\ \widetilde{\lambda}_i(t) \cdot \mathrm{d}W(t) &= \widetilde{\sigma}_i(t)\widetilde{\beta}_i(t) \cdot \mathrm{d}W(t) = \widetilde{\sigma}_i(t)\mathrm{d}\widetilde{B}_i(t), \\ \widehat{\lambda}_k(t) \cdot \mathrm{d}W(t) &= \widehat{\sigma}_k(t)\widehat{\beta}_k(t) \cdot \mathrm{d}W(t) = \widehat{\sigma}_k(t)\mathrm{d}\widehat{B}_k(t). \end{aligned}$$

We can now define the correlations between the $2N+1$ different Brownian motions, $B_i(t), \widetilde{B}_i(t), \widehat{B}_k(t) \in \mathbf{R}$, as follows

$$\begin{aligned}
\mathrm{d}B_i(t)\mathrm{d}B_j(t) &= \beta_i(t)\cdot\beta_j(t)\mathrm{d}t = \rho^D_{ij}(t)\mathrm{d}t,\\
\mathrm{d}\widetilde{B}_i(t)\mathrm{d}\widetilde{B}_j(t) &= \widetilde{\beta}_i(t)\cdot\widetilde{\beta}_j(t)\mathrm{d}t = \rho^F_{ij}(t)\mathrm{d}t,\\
\mathrm{d}B_i(t)\mathrm{d}\widetilde{B}_j(t) &= \beta_i(t)\cdot\widetilde{\beta}_j(t)\mathrm{d}t = \rho^{DF}_{ij}(t)\mathrm{d}t,\\
\mathrm{d}B_i(t)\mathrm{d}\widehat{B}_k(t) &= \beta_i(t)\cdot\widehat{\beta}_k(t)\mathrm{d}t = \rho^{DX}_{i}(t)\mathrm{d}t,\\
\mathrm{d}\widetilde{B}_i(t)\mathrm{d}\widehat{B}_k(t) &= \widetilde{\beta}_i(t)\cdot\widehat{\beta}_k(t)\mathrm{d}t = \rho^{FX}_{i}(t)\mathrm{d}t.
\end{aligned}$$

We can express the overall $(2N+1)\times(2N+1)$ cross-currency correlation matrix as

$$\rho(t) = \begin{bmatrix} \rho^D(t) & \rho^{DF}(t) & \rho^{DX}(t) \\ \rho^{DF}(t) & \rho^F(t) & \rho^{FX}(t) \\ \rho^{DX}(t) & \rho^{FX}(t) & 1 \end{bmatrix}. \tag{48}$$

This correlation matrix provides significant freedom for capturing the correlation present in the cross-currency market. In our experiments we have chosen to simplify the correlation structure significantly by requiring the matrix $\rho^{DF}(t) = \rho^{DF}\mathbf{1}\mathbf{1}^T$ and the vectors $\rho^{DX}(t) = \rho^{DX}\mathbf{1}$ and $\rho^{FX}(t) = \rho^{FX}\mathbf{1}$, where $\mathbf{1}\in\mathbf{R}^N$ is a vector of ones. We also assume that the domestic $\rho^D(t)$ and foreign $\rho^F(t)$ correlation matrices are time independent and $\eta, \gamma, \widetilde{\eta}, \widetilde{\gamma} \in \mathbf{R}$, have the structure

$$\rho^D_{ij} = \eta + (1-\eta)\mathrm{e}^{-\gamma|T_i-T_j|}, \quad \rho^F_{ij} = \widetilde{\eta} + (1-\widetilde{\eta})\mathrm{e}^{-\widetilde{\gamma}|T_i-T_j|}.$$

This correlation structure is a dramatic simplification of the overall correlation that can be captured from (48), however similar simplifications are employed in (Fries, 2007, Chapter 26), where $\rho^{DF}(t) = \rho^{DF}\rho^D(t)\rho^F(t)$ and the vectors $\rho^{DX}(t) = \rho^{DX}\rho^D_k$ and $\rho^{FX}(t) = \rho^{FX}\rho^F_k(t)$, where ρ_k represents the k:th column of ρ. Now that we have computed the required correlation matrix, we turn our attention to evaluating the $(2N+1)\times(2N+1)$ cross-currency covariance matrix (19), substituting the expressions for $\lambda_i(t)$, $\lambda_i(t)$ and $\widehat{\lambda}_k(t)$, gives the sub-matrices of the covariance matrices C_k

$$\begin{aligned}
C^D_{kij} &= \int_{T_{k-1}}^{T_k} \sigma_i(u)\sigma_j(u)\rho^D_{ij}\mathrm{d}u,\\
C^F_{kij} &= \int_{T_{k-1}}^{T_k} \widetilde{\sigma}_i(u)\widetilde{\sigma}_j(u)\rho^F_{ij}\mathrm{d}u,\\
C^{DF}_{kij} &= \int_{T_{k-1}}^{T_k} \sigma_i(u)\widetilde{\sigma}_j(u)\rho^{DF}_{ij}\mathrm{d}u,\\
C^{DX}_{ki} &= \int_{T_{k-1}}^{T_k} \sigma_i(u)\widehat{\sigma}_k(u)\rho^{DX}_{i}\mathrm{d}u,\\
C^{FX}_{ki} &= \int_{T_{k-1}}^{T_k} \widetilde{\sigma}_i(u)\widehat{\sigma}_k(u)\rho^{FX}_{i}\mathrm{d}u,\\
C^X_k &= \int_{T_{k-1}}^{T_k} \widehat{\sigma}_k(u)\widehat{\sigma}_k(u)\mathrm{d}u.
\end{aligned}$$

One of the advantages of certain choices of the deterministic volatility functions $\sigma_i(t)$ and $\widetilde{\sigma}_i(t)$ and time independent correlation matrices is that the covariance matrices C^D_k, C^F_k and C^{DF}_k, for all k, can be computed exactly. The deterministic volatility functions for the domestic and foreign economies, that we have used in our experiments are those suggested in the book by (Rebonato, 2004, Pg. 167), which are non-zero for $t < T_i$

$$\sigma_i(t) = k_i\left[(a + b(T_i - t))\mathrm{e}^{-c(T_i-t)} + d\right], \quad \widetilde{\sigma}_i(t) = \widetilde{k}_i\left[(\widetilde{a} + \widetilde{b}(T_i - t))\mathrm{e}^{-\widetilde{c}(T_i-t)} + \widetilde{d}\right],$$

where $a, b, c, d, \widetilde{a}, \widetilde{b}, \widetilde{c}, \widetilde{d}, k_i, \widetilde{k}_i \in \mathbf{R}$. We are now left with the choice of the form of $\widehat{\sigma}_k(t)$ which we assume to be piecewise constant within each tenor date, that is $\widehat{\sigma}_k(t) = \widehat{\sigma}_k$, for $t \in [T_{k-1}, T_k]$. This choice of volatility structure in the forward exchange rate ensures that the covariance matrices C_k^{DX}, C_k^{FX} and C_k^{X}, for all k, can also be computed exactly.

We are now left with the decision how do we choose the coefficients $\widehat{\sigma}_k$ so that we ensure that we obtain the correct implied volatilities of the forward foreign exchange option. The implied volatility of the forward foreign exchange option is expressed, using Equation (15), as

$$\alpha^2(T_k)T_k = \int_{T_0}^{T_k} \left|\widehat{\lambda}_k(t)\right|^2 \mathrm{d}t = \sum_{l=1}^{k} \int_{T_{\ell-1}}^{T_\ell} \left|\widehat{\lambda}_k(t)\right|^2 \mathrm{d}t$$

$$= \sum_{\ell=1}^{k} \int_{T_{\ell-1}}^{T_l} \left|\widehat{\lambda}_\ell(t) + \sum_{r=\ell}^{k-1} h_r(t)\lambda_r(t) - \sum_{r=\ell}^{k-1} \widetilde{h}_r(t)\widetilde{\lambda}_r(t)\right|^2 \mathrm{d}t.$$

Unfortunately, this expression is stochastic because the terms $h_r(t)$ and $\widetilde{h}_r(t)$ are stochastic, see Equation (14). Given that these terms are low variance martingales, see (Brace, 2008, Equation (3.4)), we can approximate the implied volatility as

$$\alpha^2(T_k)T_k \approx \sum_{\ell=1}^{k} \int_{T_{\ell-1}}^{T_\ell} \left|\widehat{\lambda}_\ell(t) + \sum_{r=\ell}^{k-1} h_r(0)\lambda_r(t) - \sum_{r=\ell}^{k-1} \widetilde{h}_r(0)\widetilde{\lambda}_r(t)\right|^2 \mathrm{d}t. \tag{49}$$

The first parameter $\widehat{\sigma}_1 = \alpha(T_1)$ follows immediately, the others follow once we notice that

$$X_{\ell k} = \int_{T_{\ell-1}}^{T_\ell} \left|\widehat{\lambda}_\ell(t) + \sum_{r=\ell}^{k-1} h_r(0)\lambda_r(t) - \sum_{r=\ell}^{k-1} \widetilde{h}_r(0)\widetilde{\lambda}_r(t)\right|^2 \mathrm{d}t$$

$$= \left|A_\ell^X + \sum_{r=\ell}^{k-1} h_r(0)A_{\ell r}^D + \sum_{r=\ell}^{k-1} \widetilde{h}_r(0)A_{\ell r}^F\right|^2,$$

where $A_{\ell r}^D$ and $A_{\ell r}^F$ are the r:th rows of the matrices A_ℓ^D and A_ℓ^F. The above expression can be computed before the integration begins. It now follows that the remaining parameters are chosen as

$$\widehat{\sigma}_k = \sqrt{\frac{1}{T_k - T_{k-1}} \left(\alpha^2(T_k)T_k - \sum_{i=1}^{k-1} X_{ik}\right)}. \tag{50}$$

It was pointed out by (Brace, 2008, Pg. 135), that for long-dated products the implied volatilities of the forward exchange option are often unreliable and approximations to them need to be computed. It is possible that approximating the expressions $h_r(t)$ and $\widetilde{h}_r(t)$, given in Equation (14) by $h_r(0)$ and $\widetilde{h}_r(0)$, is not sufficiently accurate, to improve this approximation (Kawai and Jäckel, 2007), suggested using asymptotic expansions to obtain more accurate approximations.

In the cross-currency LIBOR market model once we have determined the pseudo square root matrices A_k we have a calibration from which we can easily compute the corresponding covariance and correlation.

7 Numerical results

In this section we perform seven numerical experiments on two commonly traded cross-currency exotic interest rate derivatives, hoping to convince the reader that the cross-currency LIBOR market model, with the improvements suggested in Section 5 can be used to efficiently compute the price and sensitivities of such exotics.

Generally, cross-currency exotics are valued using PDE based methods. The stand out paper using PDE techniques is (Piterbarg, 2005), where foreign exchange volatility skew is introduced. We take the following quote from that paper: "Low factor models which allow partial differential equation (PDE) based methods for evaluation, is an essential requirement for large books." Basically this quote states that the cross-currency LIBOR market model is computationally too expensive to be considered for the valuation of cross-currency exotics. We aim to address this issue.

The first two experiments aim to clearly illustrate the effects that the improvements suggested in Section 5 have on the overall price of the two exotics. In the third experiment we aim to show the significant effects that the choice of seed in the first pass simulation can have on the unbiased second pass price. The fourth experiment shows the variance reductions that can be expected by using the delta hedge control variate from Section 5. Experiment five shows the average exercise distributions for the two exotics. The effects of the improvements suggested Section 5 in computing tight upper bounds are discussed in experiment six. The final experiment compares the relative timings of including the improvements suggested in Section 5 in computing the price and the sensitivities derived in Section 4. The two exotics that we will consider are cancellable cross-currency swaps and power-reverse dual-currency swaps. These are two of the most highly traded and liquid cross-currency exotic interest rate derivatives.

Power reverse dual currency swaps

Power-reverse-dual-currency (PRDC) swaps are very popular with Japanese investors. They derive their value from movements in the FX rate and the interest rates in both the domestic and foreign currencies. Typically, PRDC swaps are long-dated, with expiries of 30 or 40 years. Important papers on the PRDC swap contracts include (Sippel and Ohkoshi, 2002; Cover Story, 2003; Piterbarg, 2005). We follow the contract structure and leverage types detailed in (Piterbarg, 2005). The PRDC swap pays

$$\mathrm{CF}(T_{i+1}) = \tau_i(f_i(T_i) - Y_i(T_i)),$$

at time T_{i+1} for the period $[T_i, T_{i+1})$, where the PRDC coupons are call options on the FX rate, defined as

$$Y_i(T_i) = N_i \max(\mathrm{FX}(T_i) - K_i, 0), \quad K_i = \frac{\mathrm{FFX}_i(T_0)c_f}{c_d}, \quad N_i = \frac{c_f}{\mathrm{FFX}_i(T_0)}.$$

The constants c_d and c_f are called the domestic and foreign coupons and can vary over each tenor date, but in our examples they are kept constant. The option notional N_i determines the level of coupon payment, whereas the option strike K_i determines the likelihood that the option pays a non-zero amount. To compute the sensitivities, discussed in Section 4, we need to differentiate the discounted payoff with respect to $X(T_i)$. To do this in a way that can be easily extended to other products, we define the time T_i value of the cashflow received at time T_{i+1} as

$$h(T_i) = \frac{\tau_i(f_i(T_i) - Y_i(T_i))}{1 + \tau_i f_i(T_i)}.$$

If a product has a cashflow that is not in arrears then the denominator will vanish. The discounted payoff $g(T_i) = g(X(T_i))$ and its derivative $\frac{\partial g(T_i)}{\partial X(T_i)}$ are

$$g(T_i) = \frac{h(T_i)}{N(T_i)}, \quad \frac{\partial g(T_i)}{\partial X(T_i)} = \frac{1}{N(T_i)} \frac{\partial h(T_i)}{\partial X(T_i)} + h(T_i) \frac{\partial}{\partial X(T_i)} \left(\frac{1}{N(T_i)} \right).$$

From Equation (11) we see that the numéraire depends only on the domestic forward rates and not on the foreign forward rates or the exchange rate. Given that forward rates which have reset satisfy $f_k(T_i) = f_k(T_k)$ for $i \geq k$, then

$$\frac{\partial}{\partial f_k(T_i)} \left(\frac{1}{N(T_i)} \right) = \frac{\partial}{\partial f_k(T_k)} \left(\frac{1}{N(T_i)} \right) = -\frac{1}{N(T_i)(1 + \tau_k f_k(T_k))}.$$

Now we are only left with computing $\frac{\partial h(T_i)}{\partial X(T_i)}$, which is non-zero for

$$\frac{\partial h(T_i)}{\partial f_i(T_i)} = \frac{\tau_i(1 - h(T_i))}{1 + \tau_i f_i(T_i)}, \quad \frac{\partial h(T_i)}{\partial \mathrm{FX}(T_i)} = \frac{-\tau_i N_i}{1 + \tau_i f_i(T_i)} I_{\mathrm{FX}(T_i) > K(T_i)}.$$

Following (Piterbarg, 2005), the domestic (Japanese yen) and the foreign (US dollars) bonds satisfy the following

$$P_i(T_0) = \exp(-0.02\, T_i), \quad \widetilde{P}_i(T_0) = \exp(-0.05\, T_i).$$

The initial forward rates in the domestic and foreign currencies can be computed using Equation (1). The correlation linking the domestic and foreign currencies and the exchange rate are given by

$$\rho^{DF} = 0.25, \quad \rho^{DX} = -0.15, \quad \rho^{FX} = -0.15.$$

The initial spot FX rate is set at 105.0. In (Piterbarg, 2005) three choices of the domestic and foreign coupons were suggested representing low, medium and high leverage situations, we have chosen the domestic and foreign coupons to have value $c_d = 0.0225$ and $c_f = 0.0450$, which is the low leverage situation given in (Piterbarg, 2005). The results presented in this section are indicative of those obtained for the medium and high leverage situations, and we therefore do not report these here. We choose the simple volatility functions for the domestic, foreign and deterministic volatility forward foreign exchange rate as

$$\sigma_i(t) = 0.007, \quad \widetilde{\sigma}_i(t) = 0.012, \quad \widehat{\sigma}_k(t) = 0.150.$$

This choice gives an implied volatility of around 0.20. The domestic and foreign correlation matrices are constructed using Equation (48), with $\eta = \widetilde{\eta} = 0$ and $\gamma = 0.06$ and $\widetilde{\gamma} = 0.02$. The domestic and foreign displacements are constant for each forward rate with values $\alpha = \widetilde{\alpha} = 0.0$. We consider two examples, the first with a 10 year expiry and the second with a 30 year expiry, both with yearly tenor dates. We are using $F = 7$ factors in our experiments.

Cross-currency swaps

A cross-currency swap (CCS) is a foreign-exchange agreement between two parties to exchange principal and/or interest payments of a loan in one currency for principal and/or interest payments of a loan in another currency, where the two loans have the same net present value. CCS's are very popular products for investors who have cashflows in more than one currency as they provide an effective way to manage the exchange rate risk. The CCS comes in various types. An example of such a swap is the exchange of fixed-rate US Dollar interest payments for floating-rate interest payments in Euros. In our experiments we will consider a similar example, where both interest rates are variable. As such, at time T_{i+1} the CCS pays

$$\mathrm{CF}(T_{i+1}) = \tau_i(f_i(T_i) - \widetilde{f}_i(T_i)),$$

for the period $[T_i, T_{i+1})$. Similar expressions for the derivative of the payoff, as those computed for the PRDC swaps, can be easily computed.

We assume that the right to cancel the contract lies with the domestic investor, however other possibilities exist. Generally, when the swap is cancelled there is a final payment, which was decided upon at the start of the contract. This final payment can be dealt with easily by introducing a final rebate. However, for simplicity we assume that the rebate is zero. For our experiments, we have chosen the following initial input. The domestic and the foreign bonds satisfy

$$P_i(T_0) = \exp(-0.042\,T_i), \quad \widetilde{P}_i(T_0) = \exp(-0.036\,T_i).$$

The initial forward rates in the domestic and foreign currencies can be computed using Equation (1). The correlation linking the domestic and foreign currencies and the exchange rate are given by

$$\rho^{DF} = 0.75, \quad \rho^{DX} = -0.75, \quad \rho^{FX} = -0.55.$$

The initial spot FX rate is set at 105.0. We choose the simple volatility functions for the domestic, foreign and next forward exchange rate as

$$\begin{aligned}
\sigma_i(t) &= (0.05 + 0.09(T_i - t))\mathrm{e}^{-0.44(T_i - t)} + 0.20, \\
\widetilde{\sigma}_i(t) &= (0.01 + 0.05(T_i - t))\mathrm{e}^{-0.32(T_i - t)} + 0.25, \\
\widehat{\sigma}_k(t) &= 0.150.
\end{aligned}$$

This choice gives an implied volatility of around 0.20. The domestic and foreign correlation matrices are constructed using Equation (48), with $\eta = \widetilde{\eta} = 0$ and $\gamma = 0.06$ and $\widetilde{\gamma} = 0.04$. The domestic and foreign displacements are constant for each forward rate with values $\alpha = 0.015$ and $\widetilde{\alpha} = 0.020$ respectively. We consider two examples, the first with a 5 year expiry and the second with a 15 year expiry, both with half yearly tenor dates. Again we are using $F = 7$ factors in our experiments.

Experiment 1: Performance profiles in early exercise improvements

This experiment aims to clearly illustrate the improvements in the lower bound price, compared to the standard least squares regression, that result from excluding suboptimal points in both the first and second pass, and using the double regression enhancement and adaptive basis functions in the construction of the exercise strategy. In order to best eliminate the effects that the random number generation have on prices we use performance profiles see (Dolan and Moré, 2002) for a detailed description.

From preliminary testing we noticed that for the PRDC swaps the basis functions are almost exclusively driven by the inclusion of the exchange rate. This means that as long as the exchange rate is included in the set of basis functions all other basis functions have very little additional effect. Therefore, in our experiments on PRDC swaps we only compare the efficiency improvements gained from the use of the double regression enhancement and the exclusion of suboptimal points. We compare each of the four possible cases, setting the use of exclusion of suboptimal points to true or false and/or the use of double regression to true or false, that is m ={FF, FT, TF, TT}. For each case, we simulate 2^{12} first pass paths using the Mersenne twister pseudo random number generator, evaluate the exercise strategy and simulate 2^{13} second pass paths with the Sobol quasi random number generator, to compute an unbiased approximation to the PRDC swaps value. For each case this is repeated 100 times, with a different seed used in the Mersenne twister pseudo random number generator. This removes the often influential choice of seed in the overall comparisons. In Figure 1 we plot the performance profile of the "approximate error" (see the end of this experiment for detailed description) in the second pass price for the four different cases of dealing with early exercise, for the 10 year (left) and 30 year (right) expiries. For a given x on the horizontal axis, we plot on the vertical axis the number of seeds for which the particular choice of m has a relative error within a factor of x of the smallest error in the set m. The vertical axis intercept denotes the number of seeds for which a particular choice of m is the most accurate. For a choice of m to be better than all others, its envelope will lie above the envelopes of all other choices of m. Algorithmically,

- For 100 different first pass seeds:
  - Simulate 2^{12} first pass paths, using the Mersenne twister and the given seed;
  - For each of the four cases m ={FF, FT, TF, TT}:
    - Evaluate the exercise strategy strategy;
    - Perform the second pass simulation with 2^{13} Sobol paths;
- Find the maximum lower bound over all 400 tests, let this be the "exact solution";
- For each of the 400 tests compute the relative error using the "exact solution", one experiment with have zero relative error;
- For each of the 100 seeds record the smallest relative error;
- On the horizontal axis x plot values from 1 to X.
- On the vertical axis, for each of the four cases in m, plot the number of seeds that have a relative error smaller than x times the smallest relative error amongst the four choices of m.

From Figure 1, we can see that the double regression enhancement and the exclusion of suboptimal points on their own, that is FT and TF respectively, give significant improvement over the standard least squares regression, the FF case. In addition, their combination, represented by TT, has an even larger effect, making these improvements a significant advance in determining the correct exercise decision. We can see that over 80% of the time the combined effect of the exclusion of suboptimal points and the double regression enhancement give the largest lower bound and 100% of the time it is within a factor of 2.5 times the smallest relative error.

We now perform a similar performance profile on the CCS. However, in this case we assume that suboptimal points are not excluded. We compare each of the four possible cases, setting the use adaptive basis functions to true or false or the use of double regression to true or false, that is m ={FF, FT, TF, TT}.

From Figure 2 we can see, similar to the case for the PRDC swap, the combined effect of adaptive basis functions and the double regression enhancement dramatically improves the lower bound estimate. It is interesting that in some situations (20% of the time an unlucky choice of seed occurs) the TT method is not the best.

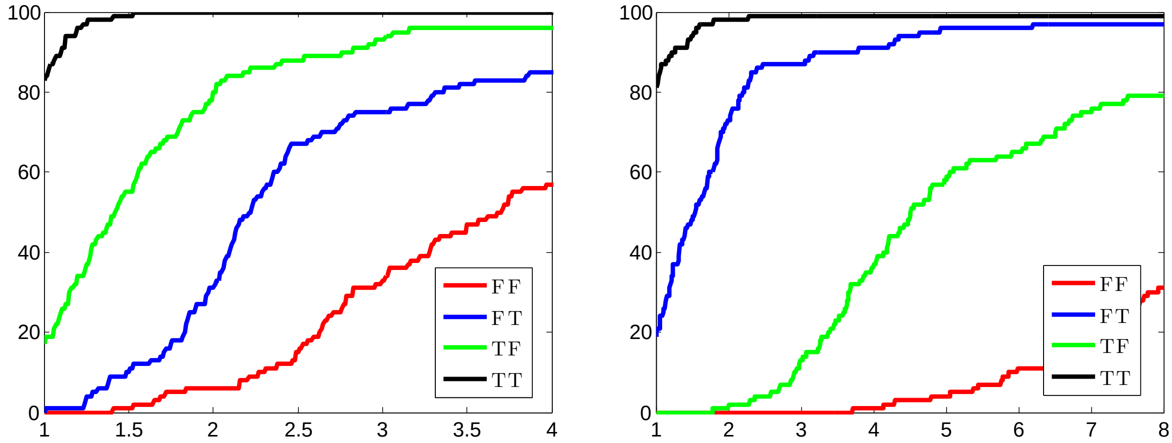

Figure 1: Simulate 2^{12} first pass paths 100 different times each with a different seed in the Mersenne twister random number generator, for each first pass run an independent second pass with 2^{13} paths using the Sobol random number generator. Repeat this experiment for each of the four methods {FF, FT, TF, TT} now plot the number of times that each method provides the highest lower bound on the vertical axis with a certain factor on the horizontal axis for the 10 year (left) and 30 year (right) cancellable PRDC swaps.

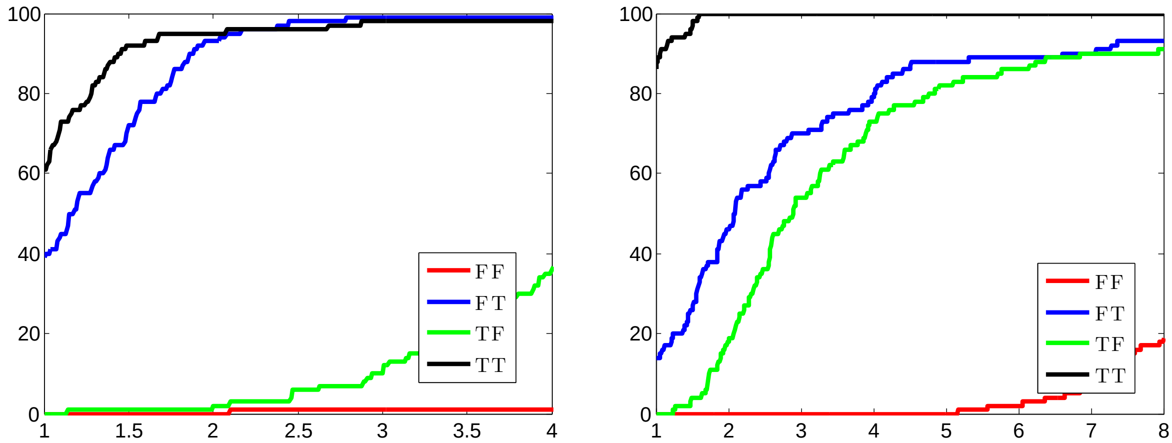

Figure 2: Simulate 2^{12} first pass paths 100 different times each with a different seed in the Mersenne twister random number generator, for each first pass run an independent second pass with 2^{13} paths using the Sobol random number generator. Repeat this experiment for each of the four methods {FF, FT, TF, TT} now plot the number of times that each method provides the highest lower bound on the vertical axis with a certain factor on the horizontal axis for the 5 year (left) and 15 year (right) cancellable CCS.

By approximate error we mean that of the 400 tests, we find the test, which gives the largest lower bound price and use that as an approximation to the exact price, then compare all other prices to this approximate exact price. This implies that one observation will have zero approximate error. Given that the price that we chosen to be the approximate exact price is a lower bound price, the true error is larger, however all approximate errors are shifted by the same amount. Given that in Experiment 7 the duality gap is shown to be very small this approximation of the exact price is very close to the true price.

Experiment 2: Comparing relative errors in early exercise improvements

This experiment provides an alternative illustration of the improvements in the lower bound price that the enhancements have over the standard least squares regression by comparing the approximate errors of the various methods.

We use the same data generated in Experiment 1. In Figure 3, we plot the approximate error in the second pass price for the PRDC with 10 year expiry (left) and 30 year expiry (right). The relative errors for the base case with all enhancements set to false, that is m =FF, are ordered from smallest to largest moving from left to right across the horizontal axis. The three possible enhancements to the base case have the same ordering as the base case. The horizontal axis plots the different seeds and the vertical axis plots the corresponding relative error. For a particular enhancement, the more points that lie below the base case points, the more successful the enhancement is.

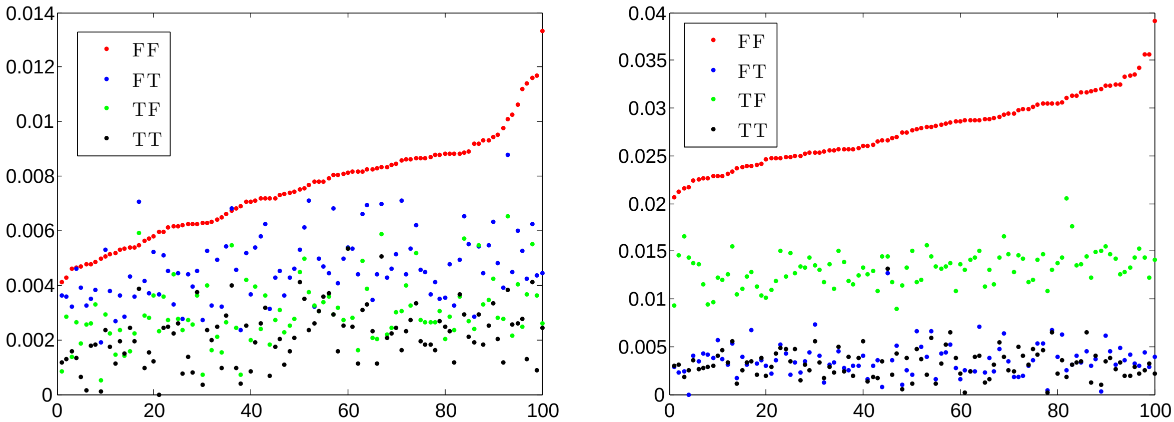

Figure 3: Plots the relative error of the four early exercise methods {FF, FT, TF, TT} for the 10 year (left) and 30 year (right) cancellable PRDC swaps. On the vertical axis is the relative error and the horizontal axis indicates each of the 100 different simulations. The results are ordered so that the FF methods has increasing relative errors.

In the left plot of Figure 3, almost all points from the regression with exclusion of suboptimal points and/or the use of the double regression enhancement lie well below the standard least squares regression case. In the right plot of Figure 3, all the points with the enhancements lie significantly below the base case.

In Figure 4 we again plot the relative error in the second pass price, for the CCS with 5 year expiry (left) and 15 year expiry (right). In Figure 4 both plots show that the enhancements provide significant

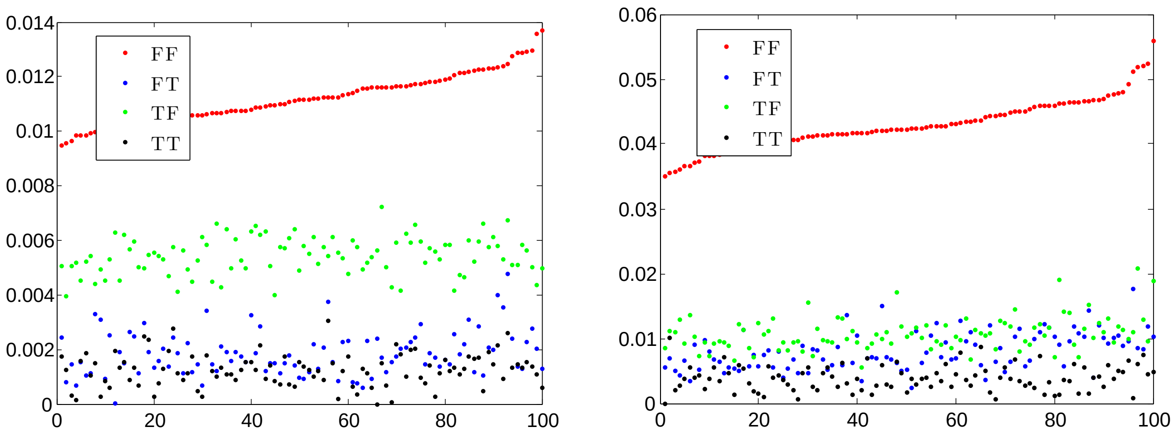

Figure 4: Plots the relative error of the four early exercise methods {FF, FT, TF, TT} for the 5 year (left) and 15 year (right) cancellable CCS. On the vertical axis is the relative error and the horizontal axis indicates each of the 100 different simulations. The results are ordered so that the FF methods has increasing relative errors.

improvements over the base case, with double regression providing the largest improvement.

Experiment 3: Convergence plots

In this experiment we want to clearly illustrate the significant effects that the first pass simulation, in particular the choice of seed, has on the overall behaviour of the final price. To do this we will perform two similar tests. In the first test we first fix the number of first pass simulations and vary the number of second pass simulations, while in the second test we will fix the number of second pass simulations and vary the number of first pass simulations.

In the first test we simulate 2^{10} and 2^{16} (left and right plots in Figure 5 respectively) first pass paths using the Mersenne Twister pseudo random number generator, evaluate the exercise strategy, and simulate $2^8, \ldots, 2^{19}$ second pass paths with the Sobol quasi random number generator. This is repeated 20 times, with a different seed used in the Mersenne Twister in the first pass. For each of the 20 seeds we assume that the price computed using 2^{19} second pass paths is an approximation to the exact solution. We then compute approximations to the absolute error as the absolute value of the difference between the "exact" solution and each of the other ten choices of the second pass paths. We then compare the approximate absolute errors against the number of simulations used in the second pass. Figure 5 plots the convergence for PRDC swaps with the standard regression (blue dots), that is without adaptive basis functions, without the exclusion of suboptimal points and without the double regression enhancement and the enhanced regression (green dots), that is without adaptive basis functions, with the exclusion of suboptimal points and with the double regression enhancement. The red line has slope 0.8 in both cases.

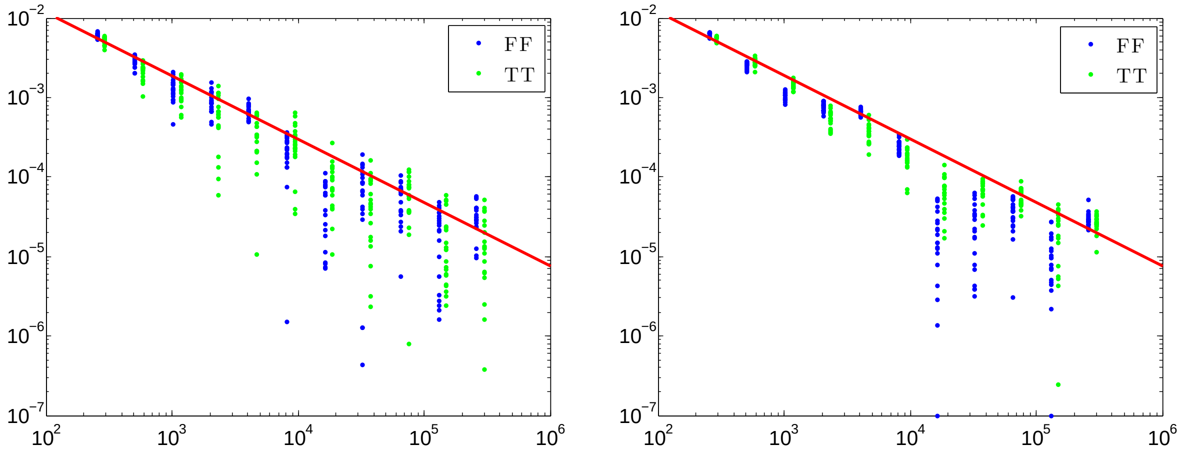

Figure 5: Plots of the approximate absolute error in the 30 year PRDC swaps for the standard regression (blue dots) and with the double regression enhancement and exclusion of suboptimal points (green dots). The left plot uses 2^{10} first pass paths and the right plot uses 2^{16}. On the vertical axis is the approximate absolute error and on the horizontal axis is the number of simulations in the second pass. The "exact" solution is computed using 2^{19} second pass paths.

Both plots seem to indicate a convergence order of at least 0.8, indicating the significant performance improvement using the Sobol quasi random number generator in the second pass path gives over the pseudo random number generator which has a convergence order of 0.5. What is significant in this plot is the dramatic effect that the particular choice of seed can have. Even more surprisingly, the number of simulations performed in the second pass can have a significant effect. For example, in the right hand plot in Figure 5, for around 8 of the 20 seeds, 2^{14} second pass paths are more accurate than 2^{18} second pass paths. It can also be argued that the use of the regression enhancements reduce the variability caused by the seed and the number of second pass paths chosen, over and above the accuracy improvements found in Experiments 1 and 2. That is, the spread of the blue dots from smallest to largest is larger than the corresponding spread for the green dots. Thus papers which provide tables of results from simulations must clearly state what random number generator is used and with what initial information was supplied, so that the experiments can be reproduced.

In the second test we will fix the number of second pass paths to be 2^{18} and vary the number of first pass paths to be $2^8, \ldots, 2^{18}$. For each number of first pass paths we will perform the pricing simulation 20 times, each with a different seed in the first pass. We repeat this test for the PRDC swap with 30 year expiry with standard regression and with the double regression enhancement and exclusion of suboptimal points. We will approximate the exact price to be the maximum price over all 400 simulations. In Figure 6 we plot on the horizontal axis the number of first pass simulations and on the vertical axis the relative error for each of the 20 different seeds. The three black lines indicate the mean relative error and one standard deviation either side of the mean. The large effect that the choice of seed has on the relative error is especially evident for a

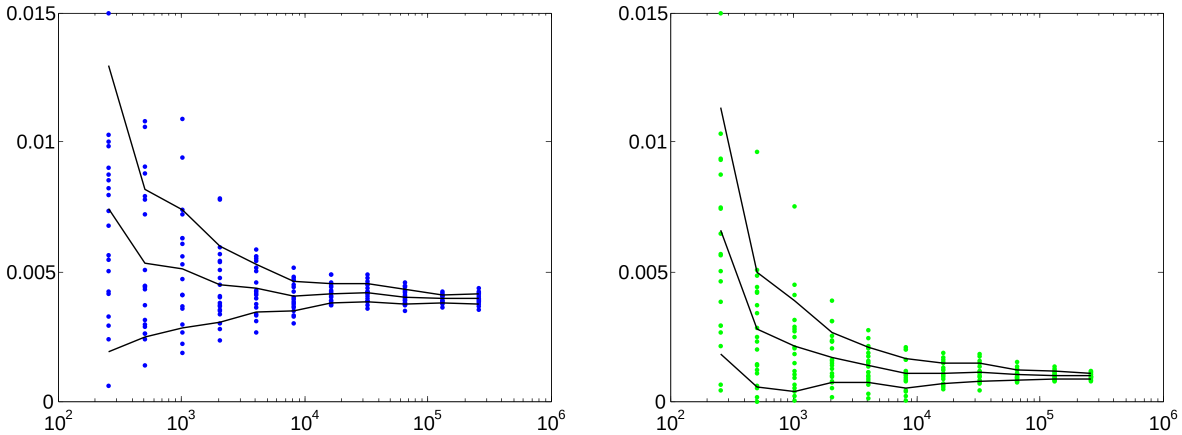

Figure 6: Plots of the approximate relative error in the 10 year PRDC swaps for the standard regression (blue dots) and with the double regression enhancement and exclusion of suboptimal points (green dots). Both plots use 2^{18} second pass paths and the number of first pass paths is indicated on the horizontal axis. On the vertical axis the relative error is given, where the "exact" solution is the maximum over all simulations

low number of first pass simulations, however even for 2^{12} first pass simulations a common choice in literature the effect of seed is significant. We can also clearly see that on average the double regression enhancement and the exclusion of suboptimal points reduces the relative error for all numbers of first pass simulations and that the spread between the mean and one standard deviation is slightly reduced.

Experiment 4: Exercise distribution

We again use the same data that was collected in Experiment 1 for both the CCS and the PRDC swap. We want to compare for both options with 10 year expiries the average distribution of the exercise time. To do this, for each of the 100 seeds we collect the proportion of time that the option is exercised at each tenor date. We then average these exercise decisions over the 100 seeds. In Figure 7, on the left/right we plot the exercise time distributions of the 10/30 year PRDC with the double regression enhancement and the exclusion of suboptimal points and the 5/15 year CCS with the double regression enhancement and adaptive basis functions.

This sort of exercise distribution is common for exotic interest rate derivatives. For long dated contracts, such as the PRDC swaps, the banks prefer reasonably liquid markets.

Experiment 5: Delta hedge control variate

In this experiment we aim to show the effects that the delta hedge control variate has on reducing the standard error in the computation of the second pass price.

To do this we again use 100 first pass simulations using the Mersenne Twister pseudo random number generator, each with 2^{12} first pass simulations and build the exercise strategy using the standard least squares regression. For each of the 100 experiments we compare the ratio of standard errors with and without the delta hedge control variate using 2^s, with $s = \{12, \ldots, 18\}$ second pass paths, for the 10 and 30 year PRDC swaps.

The standard error reduction is approximately independent of the number of second pass paths used. The standard error reduction is reduced as the number of forward rates increases. However, given that the computational time needed to compute the delta hedge control variate is approximately 15% more than the cost of computing the second pass product, see Experiment 7, the standard error reductions that result are worth the extra computational cost.

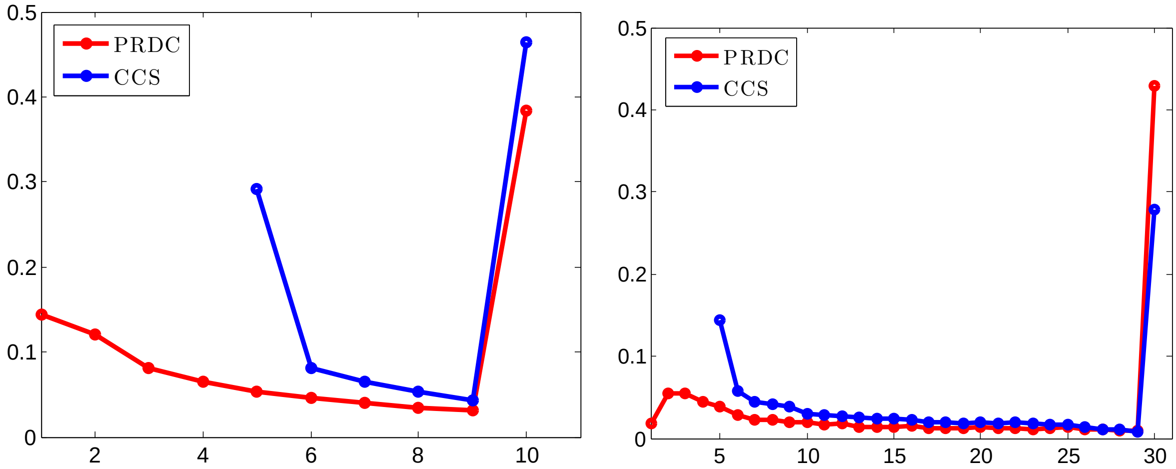

Figure 7: Exercise distributions for PRDC swaps and CCS with 10 and 5 year expiries respectively on the left and 30 and 15 year expiries on the right.

Table 1: Relative standard error reductions exhibited from using the delta hedge control variate.

Expiry	2^{12}	2^{13}	2^{14}	2^{15}	2^{16}	2^{17}	2^{18}
10	3.03	3.03	3.03	2.81	2.90	2.93	2.75
30	1.40	1.38	1.38	1.38	1.38	1.39	1.40

Experiment 6: Duality gaps

One can never rest easy with a particular lower bound methodology until it is confirmed the lower bounds are indeed close to a corresponding upper bound.

In this experiment we compare the use of the early exercise enhancements in reducing the duality gap, that is, the difference between the upper and lower bounds for the price. The duality gaps were computed using 2^{16} first pass paths with the Mersenne Twister pseudo random number generator and 2^{18} second pass paths with the Sobol quasi random number generator. We used 3500 inner paths and 2500 outer paths in the simulation calculating upper bounds, both of which used the Mersenne Twister pseudo random number generator. In Table 2, we compare results for PRDC swaps with 10 and 30 years to expiry using two types of regression, the base regression labelled FF and the enhanced regression labelled TT. The enhanced regression uses the double regression enhancement and the exclusion of suboptimal points.

Table 2: First pass, second pass and upper bound prices with corresponding duality gaps for PRDC swaps with 10 and 30 years to expiry. Results are provided with and without the double regression enhancement and the exclusion of suboptimal points.

Expiry	Type	First pass	Second pass	Upper bound	Duality gap
10	FF	0.030059	0.030212	0.030571	3.595292×10^{-4}
10	TT	0.030186	0.030345	0.030383	3.757620×10^{-5}
30	FF	0.108739	0.108794	0.113110	4.315761×10^{-3}
30	TT	0.111449	0.111527	0.111805	2.777933×10^{-4}

First we notice that in the 10 year case we are able to obtain very tight bounds, even without the regression enhancements. In particular, the largest duality gap is only 3.6 basis points. However, in the 30 year case without the regression enhancements the duality gap is 43.2 basis points where as with the regression enhancements the duality gap is reduced to 2.8 basis points. The use of the regression enhancements has a significant effect on the duality gap, reducing its size by an order of magnitude in both the 10 year and 30 year examples.

In Table 3, we compare the results for CCS's with 5 and 15 years to expiry. Again we use two types of regression, the base regression labelled FFF and the enhanced regression labelled TTT, which uses the double regression, the exclusion of suboptimal points and adaptive basis functions.

Table 3: First pass, second pass and upper bound prices with corresponding duality gaps for CCS's with 5 and 15 years to expiry. Results are provided with and without the double regression enhancement, the exclusion of suboptimal points and adaptive basis functions.

Expiry	Type	First pass	Second pass	Upper bound	Duality gap
5	FFF	0.031246	0.031257	0.031611	3.536043×10^{-4}
5	TTT	0.031646	0.031646	0.031657	1.023720×10^{-5}
15	FFF	0.087515	0.087804	0.096081	8.277386×10^{-3}
15	TTT	0.094442	0.094641	0.094890	2.486480×10^{-4}

The use of the regression enhancements has again had a significant effect on the duality gap, reducing it by over an order of magnitude in both the 5 year and 15 year examples. It is an important point to note that the upper bounds were computed using the regression enhancements, this results in significantly lower upper bounds than if the standard least squares regression was used. Therefore, the enhancements provide a twofold improvement, they increase the lower bounds and reduce the upper bounds resulting in significantly tighter bounds on the price of the cross-currency exotic interest rate derivatives.

Experiment 7: Relative timings

In this experiment we aim to compare the relative computational speeds of the enhancements that we have discussed in this paper. In particular we compare the effects of removing suboptimal points, allowing double regression and adaptive basis functions in building the exercise strategy, computing Greeks and reducing the variance using the delta hedge control variate. From preliminary testing, we have noticed significant differences in the computational times needed to compute prices and Greeks depending on the structure of the implementation. We have not attempted to optimize our runtime for each of the following experiments, but rather have one implementation with flags for each of the enhancements discussed.

We will measure the computational times relative to the basic implementation of the cross-currency LIBOR market model, that is with standard least squares regression used to compute the exercise strategy. The computational time to simulate the second pass of a twenty rate PRDC was approximately 0.73 seconds on a MacBookPro running windows and Visual Studio 2008. As a point of reference we have also compared the cross-currency LIBOR market model, without the Bermudan feature, with fourty rates and seven factors to a LIBOR market model implementation with the same number of rates and factors. The cross-currency implementation runs in less than twice the computational time than the single currency analog. When considering the exclusion of suboptimal points, it is possible to remove points in the the first pass, this results in a slight reduction in computational time given that storage requirements are reduced. We have chosen to collect all information in the first pass and do the processing while building the exercise strategy, therefore for each test the time taken in the first pass implementation is identical.

We will run twelve different tests on the Bermudan PRDC swap, six with 10 year and six with 30 year expiry. For each test, we simulate 2^{12} first pass paths using the Mersenne twister pseudo random number generator, evaluate the exercise strategy and simulate 2^{16} second pass paths with the Sobol quasi random number generator, to compute an unbiased approximation to the PRDC swaps value. Each test is repeated 100 times, with a different seed used in the Mersenne twister pseudo random number generator and the average time taken to do the test is recorded. The six tests are $m =$ {FFFF,TFFF,FTFF,FFTF,FFFT,TTFF,TTTF}, where in this order double regression, exclusion of suboptimal points, delta hedge control variate and Greeks are set to true or false. The ratio of the computational time of the last five elements of m divided by the computational time of the first is recorded and broken up into its three main parts the first pass, the construction of the exercise strategy and the second pass.

The first pass takes slightly less than 10% of the time taken for the second pass, whereas building the exercise strategy takes slightly more than 10% of the time taken for the second pass. Apart from the computations of the Greeks the other enhancements only result in modest computational time increases, in many cases there is in fact a reduction in computational time as a result of the improved exercise strategy. One point of interest is the difference in time required to build the strategy in the 10 and 30 year PRDC swap contracts, when double regression is included. From Section 5.2 recall that a single number δ is used

Table 4: Relative time changes compared to the standard least squares regression for the Bermudan PRDC with expiries of 10 and 30 years. The type of experiment is chosen by deciding whether, in the following order, the: double regression enhancement, exclusion of suboptimal points, delta hedge control variate and Greeks are set to true or false.

Expiry	Type	Strategy time	Second pass time
10	TFFF	1.45	1.02
10	FTFF	0.52	0.93
10	FFTF	1.00	1.14
10	FFFT	1.00	2.36
10	TTFF	0.92	0.94
10	TTTF	0.92	1.13
30	TFFF	1.19	0.98
30	FTFF	0.44	0.96
30	FFTF	1.00	1.18
30	FFFT	1.00	2.05
30	TTFF	0.67	0.92
30	TTTF	0.67	1.10

to decide whether an observation should be included in the second regression, an alternative approach would be to include a certain percentage, say 20%, of the observations in the second regression, this would result in a more uniform increase in computational time.

The timings, presented in Table 4, should be compared with Experiment 1, where the relative advantages of including them are given. For example, in the 30 year PRDC swap, including the double regression and suboptimal enhancements results in a reduction in the computational time spent in the second pass by 8%, but results for all 100 tests in at least a 5 times improvement in overall accuracy. For the PRDC swaps the adaptive basis functions do not provide any significant improvement, however this is generally not the case. The adaptive basis functions, if implemented in such a way, as to take advantage of previously computed information, result in a two fold increase in the computational time building the exercise strategy and only a modest increase in the second pass time. For the computation of the Greeks we recall that if finite differences are used to compute one delta this results in an increase in the second pass time by a factor of two. Using the adjoint pathwise method we can compute all the deltas a vega for each step and the exchange vega, for a very slight increase in the computational time required to compute one delta using finite differences.

8 Conclusion

We have seen that despite the challenging nature of callable cross-currency notes, it is possible to develop a cross-currency LIBOR market model which computes both prices and Greeks in a small amount of time. We have also carried out extensive testing of various methodologies for estimating the exercise strategy and we have comprehensively demonstrated that our methodologies are robust and lead to very good estimates of the true price. In conclusion, the cross-currency LIBOR market model can be effectively used for marking to market and risk computations, as well as for benchmarking lattice models.

References

Ametrano, F. and Joshi, M. S. 2008. Smooth simultaneous calibration of the LMM to caplets and coterminal swaptions. *SSRN eLibrary*, 1–20.

Amin, A. 2003. Multi-factor cross currency LIBOR market model: implementation, calibration and examples. *SSRN eLibrary*, 1–18.

Andersen, L. and Broadie, M. 2004. Primal-dual simulation algorithm for pricing multidimensional American options. *Management Science 50,* 9, 1222–1234.

Beveridge, C. and Joshi, M. S. 2008. Juggling snowballs. *SSRN eLibrary*, 1–16.

Beveridge, C. and Joshi, M. S. 2009. Practical policy iteration: generic methods for obtaining rapid and tight bounds for Bermudan derivativs using Monte Carlo simulation. *SSRN eLibrary*, 1–36.

Brace, A. 2008. *Engineering BGM.* Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL.

Brigo, D. and Mercurio, F. 2006. *Interest rate models—theory and practice*, Second ed. Springer Finance. Springer-Verlag, Berlin. With smile, inflation and credit.

Broadie, M. and Cao, M. 2008. Improved lower and upper bound algorithms for pricing American options by simulation. *Quant. Finance 8,* 8, 845–861.

Broadie, M. and Glasserman, P. 1996. Estimating security price derivatives using simulation. *Management Science 42,* 2, 269–285.

Burrage, K. and Burrage, P. M. 1998. General order conditions for stochastic Runge–Kutta methods for both commuting and non-commuting stochastic ordinary differential equation systems. *Appl. Numer. Math. 28,* 2-4, 161–177. Eighth Conference on the Numerical Treatment of Differential Equations (Alexisbad, 1997).

Carrière, J. F. 1996. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance Math. Econom. 19,* 1, 19–30.

Cover Story. 2003. The problem with power-reverse duals. *Risk 16,* 10.

Deacon, M., Derry, A., and Mirfendereski, D. 2004. *Inflation-indexed securities*, Second ed. John Wiley & Sons, Ltd, Chichester.

Denson, N. and Joshi, M. 2009a. Fast and accurate greeks for the LIBOR market model. *SSRN eLibrary*, 1–20.

Denson, N. and Joshi, M. S. 2009b. Flaming logs. *Wilmott J. 1*, 259–262.

Dolan, E. D. and Moré, J. J. 2002. Benchmarking optimization software with performance profiles. *Math. Program. 91,* 2, Ser. A, 201–213.

Dun, T. 2006. Calibration of a cross-currency LIBOR market model.

Fries, C. 2007. *Mathematical finance.* Wiley-Interscience [John Wiley & Sons], Hoboken, NJ. Theory, modeling, implementation.

Fries, C. P. 2008. Foresight bias and suboptimality correction in Monte-Carlo pricing of options with early exercise. In *Progress in industrial mathematics at ECMI 2006.* Math. Ind., vol. 12. Springer, Berlin, 645–649.

Giles, M. and Glasserman, P. 2006. Smoking adjoints: fast Monte Carlo Greeks. *Risk January*, 92–96.

Glasserman, P. 2004. *Monte Carlo methods in financial engineering.* Applications of Mathematics (New York), vol. 53. Springer-Verlag, New York. Stochastic Modelling and Applied Probability.

Glasserman, P. and Zhao, X. 1999. Fast Greeks by simulation in forward LIBOR models. *J. Comput. Finance 3*, 5–39.

Griewank, A. and Walther, A. 2008. *Evaluating derivatives*, Second ed. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. Principles and techniques of algorithmic differentiation.

Haugh, M. B. and Kogan, L. 2004. Pricing American options: a duality approach. *Oper. Res. 52,* 2, 258–270.

Hunter, C. J., Jäckel, P., and Joshi, M. S. 2001. Getting the drift. *Risk 14,* 7, 81–84.

Jarrow, R. and Yildirim, Y. 2003. Pricing treasury inflation protected securities and related derivatives using an HJM model. *J. Finan. Quant. Anal. 38,* 2, 337–358.

Joshi, M. S. 2003. Rapid computation of drifts in a reduced factor LIBOR market model. *Wilmott Magazine May*, 84–85.

Joshi, M. S. 2006. Monte Carlo bounds for callable products with non-analytic break costs. *SSRN eLibrary*, 1–10.

Joshi, M. S. 2007. A simple derivation of and improvements to Jamshidian's and Roger's upper bound methods for Bermudan options. *Appl. Math. Finance 14,* 3, 197–205.

Joshi, M. S. and Kwon, O. K. 2010. Monte carlo market Greeks in the displaced diffusion LIBOR market model. *SSRN eLibrary*, 1–15.

Joshi, M. S. and Liesch, L. 2007. Effective implementation of generic market models. *Astin Bull. 37,* 2, 453–473.

Kawai, A. and Jäckel, P. 2007. An asymptotic FX option formula in the cross-currency LIBOR market model.

Kloeden, P. E. and Platen, E. 1992. *Numerical solution of stochastic differential equations.* Applications of Mathematics (New York), vol. 23. Springer-Verlag, Berlin.

Longstaff, F. A. and Schwartz, E. S. 2001. Valuing American options by simulation: a simple least squares approach. *Rev. Financ. Stud. 14*, 113–147.

Mikkelsen, P. 1999. Cross currency LIBOR market models. Tech. rep., The Aarhus School of Business.

Musiela, M. and Rutkowski, M. 2005. *Martingale methods in financial modelling*, Second ed. Stochastic Modelling and Applied Probability, vol. 36. Springer-Verlag, Berlin.

Piterbarg, V. V. 2003. Computing deltas of callable LIBOR exotics in forward LIBOR models. *SSRN eLibrary*, 1–32.

Piterbarg, V. V. 2005. A multi-currency model with FX volatility skew. *SSRN eLibrary*, 1–25.

Rebonato, R. 2002. *Modern Pricing of Interest-Rate Derivatives: The LIBOR Market and Beyond.* Princeton University Press, New Jersey.

Rebonato, R. 2004. Interest-rate term-structure pricing models: a review. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 460,* 2043, 667–728.

Rogers, L. C. G. 2002. Monte Carlo valuation of American options. *Math. Finance 12,* 3, 271–286.

Schlögl, E. 2002. A multicurrency extension of the lognormal interest rate market models. *Finance Stoch. 6,* 2, 173–196.

Sippel, J. and Ohkoshi, S. 2002. All power to PRDC notes. *Risk 15,* 11, 31–33.

Tsitsiklis, J. N. and van Roy, B. 2001. Regression methods for pricing complex American-style options. *IEEE Trans. Neural Networks 12,* 4, 694–703.