A PDE Pricing Framework for Cross-Currency Interest Rate Derivatives

Duy Minh Dang[a,*], Christina C. Christara[a], Kenneth R. Jackson[a], Asif Lakhany[b]

[a]*Department of Computer Science, University of Toronto, Toronto, ON, M5S 3G4, Canada*
[b]*Algorithmics Inc., Toronto, ON, M5T 2C6, Canada*

Abstract

We propose a general framework for efficient pricing via a Partial Differential Equation (PDE) approach of cross-currency interest rate derivatives under the Hull-White model. In particular, we focus on pricing long-dated foreign exchange (FX) interest rate hybrids, namely Power Reverse Dual Currency (PRDC) swaps with Bermudan cancelable features. We formulate the problem in terms of three correlated processes that incorporate FX skew via a local volatility function. This formulation results in a time-dependent parabolic PDE in three spatial dimensions. Finite difference methods on uniform grids are used for the spatial discretization of the PDE. The Crank-Nicolson (CN) method and the Alternating Direction Implicit (ADI) method are considered for the time discretization. In the former case, the preconditioned Generalized Minimal Residual (GMRES) method is employed for the solution of the resulting block banded linear system at each time step, with the preconditioner solved by Fast Fourier Transform (FFT) techniques. Numerical results indicate that the numerical methods considered are second-order convergent, and, asymptotically, as the discretization granularity increases, almost optimal, with the ADI method being modestly more efficient than CN-GMRES-FFT. An analysis of the impact of the FX volatility skew on the PRDC swaps' prices is presented, showing that the FX volatility skew results in lower prices (i.e. profits) for the payer of PRDC coupons.

Keywords: Power Reverse Dual Currency (PRDC) swaps, Bermudan cancelable, Partial Differential Equation (PDE), Alternating Direction Implicit (ADI), Generalized Minimal Residual (GMRES), Fast Fourier Transform (FFT)

1. Introduction

We investigate the modeling and numerical valuation of cross-currency interest rate derivatives with strong emphasis on long-dated foreign exchange (FX) interest rate hybrids, namely Power Reverse Dual Currency (PRDC) swaps with Bermudan cancelable features using a partial differential equation (PDE) approach. The motivation for this work is that, although cross-currency interest rate derivatives in general, and FX interest rate hybrids in particular, are of enormous current practical importance, the valuation of these derivatives via a PDE approach is not well-developed in the literature. More specifically, the pricing of PRDC swaps has been a subject of great interest in practice, especially among financial institutions, yet PDE methods for doing so are not well-developed in the public domain. The popular choice for pricing PRDC swaps is Monte-Carlo (MC) simulation, but this approach has several major disadvantages, such as slow convergence and difficulty in computing hedging parameters.

Foreign exchange (FX) interest rate hybrids, such as PRDC swaps, are exposed to moves in both the spot FX rate and the interest rates in both currencies. The current standard modeling of such products consists of two one-factor Gaussian models for the term structures and a one-factor log-normal model for the spot

*Corresponding author
Email addresses: dmdang@cs.toronto.edu (Duy Minh Dang), ccc@cs.toronto.edu (Christina C. Christara), krj@cs.toronto.edu (Kenneth R. Jackson), asif.lakhany@algorithmics.com (Asif Lakhany)

FX rate [1]. These choices provide the benefits of (i) keeping the number of factors to the minimum (a total of three) and (ii) using very efficient, essentially closed-form, calibration for the spot FX rate model. However, FX options, especially long-dated ones, exhibit a significant skew, which cannot be well captured by the log-normal distribution. In addition, cross-currency interest rate derivatives with exotic features, such as Bermudan cancelable PRDC swaps, are particularly sensitive to the FX volatility smile/skew. As a result, the assumption of log-normality of spot FX rates is questionable in the modeling of such derivatives. One way to rectify this deficiency is to incorporate FX volatility smiles into the model via a local volatility function, as first suggested in [2]. By using a local volatility model, this approach avoids introducing more stochastic factors into the model. Hence it keeps the number of factors to the minimum, while providing better modeling for the skewness of the FX rate. Under a three-factor model, cross-currency interest rate derivatives are dependent on three stochastic state variables and thus the PDE which their value function must satisfy has three state variables in addition to the time variable. Furthermore, these products have additional complexity due to multiple cash flow dates and exotic features, such as Bermudan cancelability. As a result, pricing such derivatives via a PDE approach is computationally challenging.

The remainder of this paper is organized as follows. In Section 2, we present a three-factor cross-currency pricing model with FX skew and two short rates which follow the one-factor Hull-White model [3] and then derive the PDE for pricing cross-currency derivatives. Section 3 discusses discretization schemes for the pricing PDE and two numerical methods for solving the discretized problem. The first scheme employs Crank-Nicolson for the time discretization while the solution of the linear system at each timestep is handled by the preconditioned Generalized Minimal Residual (GMRES) method with a preconditioner solved by Fast Fourier Transform (FFT) techniques. The second scheme uses the Alternating-Direction Implicit (ADI) method. The pricing of PRDC swaps with Bermudan cancelable features is discussed in Section 4. Numerical results are provided in Section 5. Section 6 concludes the paper and outlines possible future work.

2. The model and the associated PDE

We consider an economy with two currencies, "domestic" (d) and "foreign" (f). We denote by $s(t)$ the spot FX rate, the number of units of domestic currency per one unit of foreign currency. Let $r_i(t), i = d, f$, denote the domestic and foreign short rates respectively. Under the domestic risk-neutral measure, the dynamics of $s(t), r_d(t), r_f(t)$ are described by [2]

$$\begin{aligned} \frac{ds(t)}{s(t)} &= (r_d(t) - r_f(t))dt + \gamma(t, s(t))dW_s(t), \\ dr_d(t) &= (\theta_d(t) - \kappa_d(t)r_d(t))dt + \sigma_d(t)dW_d(t), \\ dr_f(t) &= (\theta_f(t) - \kappa_f(t)r_f(t) - \rho_{fs}(t)\sigma_f(t)\gamma(t, s(t)))dt + \sigma_f(t)dW_f(t), \end{aligned} \tag{1}$$

where $W_d(t), W_f(t)$, and $W_s(t)$ are correlated Brownian motions with $dW_d(t)dW_s(t) = \rho_{ds}dt$, $dW_f(t)dW_s(t) = \rho_{fs}dt$, and $dW_d(t)dW_f(t) = \rho_{df}dt$. The short rates follow the mean-reverting Hull-White model with the mean reversion rate and the volatility functions respectively denoted by $\kappa_i(t)$ and $\sigma_i(t)$, for $i = d, f$, while $\theta_i(t)$, $i = d, f$, captures the current term structures. The "quanto" drift adjustment, $-\rho_{fs}(t)\sigma_f(t)\gamma(t, s(t))$, for $dr_f(t)$ comes from changing the measure from the foreign risk-neutral measure to the domestic risk neutral one. The functions $\kappa_i(t), \sigma_i(t), \theta_i(t), i = d, f$, are all deterministic. The local volatility function $\gamma(t, s(t))$ for the spot FX rate has the functional form [2]

$$\gamma(t, s(t)) = \xi(t)\left(\frac{s(t)}{L(t)}\right)^{\varsigma(t)-1}, \tag{2}$$

where $\xi(t)$ is the relative volatility function, $\varsigma(t)$ is the time-dependent constant elasticity of variance (CEV) parameter and $L(t)$ is a time-dependent scaling constant which is usually set to the forward FX rate with expiry t, denoted by $F(0, t)$, for convenience in calibration.

Before any model can be used, calibration of the model parameters to specific market data is required. The parameters defining the volatility structures of interest rates in both currencies, i.e. the functions $\sigma_i(t), \kappa_i(t), i = d, f$, can be bootstrapped from European swaption market values for the respective currencies, while $\theta_i(t), i = d, f$, are determined by the current term structures of bond prices in the respective currency. Correlation parameters are typically chosen based on historical estimations. The calibration of the local volatility function can be done via forward FX options, as suggested in [2].

We now give a PDE that the price of any security whose payoff is a function of the domestic and foreign interest rates and the exchange rate must satisfy.

Theorem 1: Let $u \equiv u(s, r_d, r_f, t)$ denote the domestic value function of a security with a terminal payoff measurable with respect to the σ-algebra at maturity time T_{end} and without intermediate payments. Furthermore, assume that $u \in \mathcal{C}^{2,1}$ on $\mathbb{R}^3_+ \times [T_{start}, T_{end})$, i.e. u is at least twice differentiable with respect to the space variables and differentiable with respect to the time variable. Then on $\mathbb{R}^3_+ \times [T_{start}, T_{end})$, u satisfies the PDE

$$
\begin{aligned}
\frac{\partial u}{\partial t}+\mathcal{L}u \equiv \frac{\partial u}{\partial t}+(r_d-r_f)s\frac{\partial u}{\partial s}+\Big(\theta_d(t)-\kappa_d(t)r_d\Big)\frac{\partial u}{\partial r_d}+\Big(\theta_f(t)-\kappa_f(t)r_f-\rho_{fs}\sigma_f(t)\gamma(t,s(t))\Big)\frac{\partial u}{\partial r_f}& \\
+\frac{1}{2}\gamma^2(t,s(t))s^2\frac{\partial^2 u}{\partial s^2}+\frac{1}{2}\sigma_d^2(t)\frac{\partial^2 u}{\partial r_d^2}+\frac{1}{2}\sigma_f^2(t)\frac{\partial^2 u}{\partial r_f^2}+\rho_{ds}\sigma_d(t)\gamma(t,s(t))s\frac{\partial^2 u}{\partial s\partial r_d}& \\
+\rho_{fs}\sigma_f(t)\gamma(t,s(t))s\frac{\partial^2 u}{\partial s\partial r_f}+\rho_{df}\sigma_d(t)\sigma_f(t)\frac{\partial^2 u}{\partial r_d\partial r_f}-r_d u=0&
\end{aligned}
\tag{3}
$$

Proof: Under the domestic risk-neutral measure, the normalized price process of any security is a martingale. Since it is an Itô process, it must have zero drift. Calculating the drift term using Itô's formula and setting it to zero gives us the PDE (3). □

Since payoffs and fund flows are deal-specific, we defer specifying the terminal conditions until a later section. The difficulty with choosing boundary conditions is that, for an arbitrary payoff, they are not known. A detailed analysis of the boundary conditions is certainly beyond the scope of this short paper, and is a topic of future research. For this project, we only impose general approximate boundary conditions. We choose Dirichlet-type "stopped process" boundary conditions where we stop the processes $s(t), r_f(t), r_d(t)$ when any of the three hits the boundary. Thus, the value on the boundary is simply the discounted payoff for the current values of the state variables [4].

Note that the PDE (3) is in terms of forward time, but, since we solve the PDE backward in time, the change of variable $\tau = T_{end} - t$ is used. Under this change of variable, the PDE (3) becomes $\frac{\partial u}{\partial \tau} = \mathcal{L}u$. The pricing of cross-currency interest rate derivatives is defined in an unbounded domain $\{(s, r_d, r_f, \tau) | s \geq 0, r_d \geq 0, r_f \geq 0, \tau \in [0, T]\}$, where $T = T_{end} - T_{start}$. In order to use Finite Difference (FD) approximations for space variables, we truncate the unbounded domain into a finite-sized computational one $\{(s, r_d, r_f, \tau) \in [0, S] \times [0, R_d] \times [0, R_f] \times [0, T]\} \equiv \Omega \times [0, T]$, where S, R_f, R_d are sufficiently large [5].

3. Discretization of the PDE

Let the number of subintervals be $n+1$, $p+1$, $q+1$ and $l+1$ in the s-, r_d-, r_f-, and τ-directions, respectively. The uniform grid stepsizes in the respective directions are denoted by $\Delta s = \frac{S}{n+1}$, $\Delta r_d = \frac{R_d}{p+1}$, $\Delta r_f = \frac{R_f}{q+1}$, and $\Delta \tau = \frac{T}{l+1}$. The grid point values of a FD approximation are denoted by $u^m_{i,j,k} \approx u(s_i, r_{dj}, r_{fk}, \tau_m) = u(i\Delta s, j\Delta r_d, k\Delta r_f, m\Delta\tau)$, where $i = 1\ldots, n$, $j = 1\ldots, p$, $k = 1\ldots, q$, $m = 1\ldots, l+1$. Second-order FD approximations to the first and second partial derivatives of the space variables in (3) are obtained by *central* schemes, while the cross-derivatives are approximated by a four-point FD stencil. For example, at the reference point $(s_i, r_{dj}, r_{fk}, \tau_m)$, the first and second derivatives with respect to the spot FX rate s, i.e. $\frac{\partial u}{\partial s}$ and $\frac{\partial^2 u}{\partial s^2}$, are approximated by

$$\frac{\partial u}{\partial s} \approx \frac{u^m_{i+1,j,k} - u^m_{i-1,j,k}}{2\Delta s}, \quad \frac{\partial^2 u}{\partial s^2} \approx \frac{u^m_{i+1,j,k} - 2u^m_{i,j,k} + u^m_{i-1,j,k}}{(\Delta s)^2}, \tag{4}$$

while the cross-derivative $\frac{\partial^2 u}{\partial s \partial r_d}$ is approximated by

$$\frac{\partial^2 u}{\partial s \partial r_d} \approx \frac{u^m_{i+1,j+1,k} + u^m_{i-1,j-1,k} - u^m_{i-1,j+1,k} - u^m_{i+1,j-1,k}}{4\Delta s \Delta r_d}. \tag{5}$$

Similar approximations can be obtained for the remaining spatial derivatives. For brevity, we omit the derivations of (4) and (5), but, using Taylor expansions, it can be verified that each of these formulas has a second-order truncation error, provided that the function u is sufficiently smooth. The FD discretization of the spatial differential operator $\mathcal{L}$ of (3) is performed as follows. At the spatial grid Ω, each spatial derivative appearing in the operator $\mathcal{L}$ is replaced by its corresponding FD scheme (as in (4) and (5)). We denote by $\mathcal{L}u^m_{i,j,k}$ the FD discretization of $\mathcal{L}$ at $(s_i, r_{dj}, r_{fk}, \tau_m)$.

3.1. The Crank-Nicolson scheme

To step from time τ_{m-1} to τ_m, we apply the Crank-Nicolson scheme

$$\frac{u^m_{i,j,k} - u^{m-1}_{i,j,k}}{\Delta \tau} = \frac{1}{2}\mathcal{L}u^m_{i,j,k} + \frac{1}{2}\mathcal{L}u^{m-1}_{i,j,k},$$

where $i = 1, \dots, n, j = 1, \dots, p, k = 1, \dots, q$. Unless otherwise stated, assume that the mesh points are ordered in the s-, r_d- , then r_f- directions. Let $\mathbf{u}^m$ denote the vector of values at time τ_m on the mesh Ω that approximates the exact solution $u^m = u(s, r_d, r_f, \tau_m)$. The Crank-Nicolson method defines approximations $\mathbf{u}^m$ successively for $m = 1, 2, \dots, l+1$, by

$$(\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{A}^m)\mathbf{u}^m = (\mathbf{I} + \frac{1}{2}\Delta\tau\mathbf{A}^{m-1})\mathbf{u}^{m-1} + \frac{1}{2}\Delta\tau(\mathbf{g}^m + \mathbf{g}^{m-1}), \tag{6}$$

where $\mathbf{I}$ denotes the $npq \times npq$ identity matrix and $\mathbf{A}^m$ is the matrix of the same size as $\mathbf{I}$ arising from the FD discretization of the differential operator $\mathcal{L}$. For brevity, we omit the explicit formula for $\mathbf{A}^m$. However, it can be written in a compact format using tensor products, similar to (7). The vectors $\mathbf{g}^{m-1}$ and $\mathbf{g}^m$ are obtained from the boundary conditions. Applying direct methods, such as LU factorization, to solve this linear system can be very computationally expensive for several reasons: (i) the matrix $\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{A}^m$ possesses a bandwidth proportional to $\min\{np, nq, pq\}$, depending on the ordering of the grid points, (ii) sparse solvers suffer considerable fill-in when solving systems derived from PDEs of the form (3), and (iii) this matrix needs to be factored at each timestep because of its dependence on the timestep index m of the local volatility function.

3.1.1. GMRES with a preconditioner solved by FFT techniques

To avoid the high computational cost of direct methods, we choose to apply an iterative method to solve (6), namely GMRES. We choose GMRES because $\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{A}^m$ is neither symmetric nor positive (semi-) definite in general. Thus, commonly used iterative schemes, such as the conjugate gradient method, designed primarily for symmetric positive-definite systems, are not likely to converge. A detailed description of the GMRES method, and the "restarted" version of the method, can be found in [6]. In our implementation, the initial guess for GMRES is based on linear extrapolation of the numerical solution from the two previous timesteps, except for the first timestep. It is important to mention that, due to the use of this initial guess and the use of a preconditioner which is described below, only a few iterations (usually 5 or 6) are required for the GMRES method to converge, hence no restarting is needed.

We use the matrix $\mathbf{P}$ arising from the discretization of

$$\frac{\partial^2 u}{\partial s^2} + \frac{\partial^2 u}{\partial r_d^2} + \frac{\partial^2 u}{\partial r_f^2} + u$$

as the preconditioner. In the rest of the subsection, we describe an efficient algorithm for solving linear systems of the form $\mathbf{P}\mathbf{v} = \mathbf{b}$. The matrix $\mathbf{P}$ can be written in the following form using tensor products:

$$\mathbf{P} = \frac{1}{(\Delta s)^2}(\mathbf{I}_q \otimes \mathbf{I}_p \otimes \mathbf{T}_n) + \frac{1}{(\Delta r_d)^2}(\mathbf{I}_q \otimes \mathbf{T}_p \otimes \mathbf{I}_n) + \frac{1}{(\Delta r_f)^2}(\mathbf{T}_q \otimes \mathbf{I}_p \otimes \mathbf{I}_n) + \mathbf{I}. \tag{7}$$

Here, $\mathbf{I}_n$ denotes the identity matrix of size $n \times n$ and $\mathbf{T}_n$ denotes the tridiagonal matrix representation of the classic second-order difference operator on n points. Note that the matrix $\mathbf{T}_n$ is a tridiagonal matrix with entries $\{1, -2, 1\}$ on each row except the first and last rows, which are $\{-2, 1\}$ and $\{1, -2\}$, respectively, due to Dirichlet boundary conditions. Matrices $\mathbf{T}_p$ and $\mathbf{T}_q$ are similar to $\mathbf{T}_n$ but of size $p \times p$ and $q \times q$, respectively. It is known [7] that, if we let $\mathbf{V}_n$ be the matrix of normalized eigenvectors of $\mathbf{T}_n$ and $\mathbf{F}_n$ be the Discrete Sine Transform (DST) matrix of order n, then we have $\mathbf{V}_n^{-1} = \mathbf{F}_n$. Matrices $\mathbf{V}_p$ and $\mathbf{V}_q$ are similar to $\mathbf{V}_n$, and matrices $\mathbf{F}_p$ and $\mathbf{F}_q$ are similar to $\mathbf{F}_n$. A useful observation is that $\mathbf{P}$ can be block-diagonalized via

$$\begin{aligned}\mathbf{B} &= (\mathbf{V}_q^{-1} \otimes \mathbf{V}_p^{-1} \otimes \mathbf{I}_n)\mathbf{P}(\mathbf{V}_q \otimes \mathbf{V}_p \otimes \mathbf{I}_n) \\ &= \frac{1}{(\Delta s)^2}(\mathbf{I}_q \otimes \mathbf{I}_p \otimes \mathbf{T}_n) + \frac{1}{(\Delta r_d)^2}(\mathbf{I}_q \otimes \mathbf{\Lambda}_p \otimes \mathbf{I}_n) + \frac{1}{(\Delta r_f)^2}(\mathbf{\Lambda}_q \otimes \mathbf{I}_p \otimes \mathbf{I}_n) + \mathbf{I},\end{aligned} \tag{8}$$

where $\mathbf{\Lambda}_n = \mathbf{V}_n^{-1}\mathbf{T}_n\mathbf{V}_n$ is a diagonal matrix with the eigenvalues of $\mathbf{T}_n$ on its diagonal. The matrix $\mathbf{B}$ has pq blocks, each of size $n \times n$. These observations give rise to the following fast solver for the preconditioner using FFT techniques, more specifically, using Fast Sine Transforms (FSTs). (Note that the main computational requirement of an FST is that of an FFT.) Consider the linear system $\mathbf{P}\mathbf{v} = \mathbf{b}$. Note that $\mathbf{b} \in \mathbb{R}^{npq}$ and for the sake of presentation, denote by $\mathbf{b}_{q \times pn}$ the $q \times pn$ matrix with entries being the components of $\mathbf{b}$ laid out in q rows and pn columns, column by column. Taking (8) into account, the solution $\mathbf{v}$ to the system $\mathbf{P}\mathbf{v} = \mathbf{b}$ can be written as

$$\mathbf{v} = \mathbf{P}^{-1}\mathbf{b} = (\mathbf{V}_q \otimes \mathbf{V}_p \otimes \mathbf{I}_n)\mathbf{B}^{-1}(\mathbf{V}_q^{-1} \otimes \mathbf{V}_p^{-1} \otimes \mathbf{I}_n)\mathbf{b} = (\mathbf{F}_q^{-1} \otimes \mathbf{F}_p^{-1} \otimes \mathbf{I}_n)\mathbf{B}^{-1}(\mathbf{F}_q \otimes \mathbf{F}_p \otimes \mathbf{I}_n)\mathbf{b}. \tag{9}$$

The FST algorithm for performing the computation in (9) consists of the following steps:
1. Perform the FST on each of the pn columns of $(\mathbf{b}_{pn \times q})^T$ to obtain $(\mathbf{b}^{(1)})_{q \times pn} = \mathbf{F}_q(\mathbf{b}_{pn \times q})^T$.
2. Perform the FST on each of the qn columns of $((\mathbf{b}^{(1)})_{qn \times p})^T$ to obtain $(\mathbf{b}^{(2)})_{p \times qn} = \mathbf{F}_p(\mathbf{b}^{(1)}_{qn \times p})^T$, or, equivalently $\mathbf{b}^{(2)} = (\mathbf{F}_q \otimes \mathbf{F}_p \otimes \mathbf{I}_n)\mathbf{b}$.
3. Solve the block-diagonal system $\mathbf{B}\mathbf{b}^{(3)} = \mathbf{b}^{(2)}$.
4. Perform the inverse FST on each of the pn columns of $((\mathbf{b}^{(3)})_{pn \times q})^T$ to obtain $(\mathbf{b}^{(4)})_{q \times pn} = \mathbf{F}_q^{-1}((\mathbf{b}^{(3)})_{pn \times q})^T$.
5. Perform the inverse FST on each of the qn columns of $(\mathbf{b}^{(4)}_{qn \times p})^T$ to obtain $\mathbf{v}_{p \times qn} = \mathbf{F}_p^{-1}(\mathbf{b}^{(4)}_{qn \times p})^T$, or equivalently $\mathbf{v} = (\mathbf{F}_q^{-1} \otimes \mathbf{F}_p^{-1} \otimes \mathbf{I}_n)\mathbf{B}^{-1}(\mathbf{F}_q \otimes \mathbf{F}_p \otimes \mathbf{I}_n)\mathbf{b}$.

The above five steps form an FFT technique for solving the linear system $\mathbf{P}\mathbf{v} = \mathbf{b}$. Clearly, the five steps involve $\mathcal{O}(npq \log(npq))$ flops.

3.2. The ADI scheme

We decompose the matrix $\mathbf{A}^m$ into four submatrices: $\mathbf{A}^m = \sum_{i=0}^{3} \mathbf{A}_i^m$. Here, we choose the matrix $\mathbf{A}_0^m$ as the part of $\mathbf{A}$ that comes from the FD discretization of the mixed derivative terms in (3), while the matrices $\mathbf{A}_1^m$, $\mathbf{A}_2^m$ and $\mathbf{A}_3^m$ are the three parts of $\mathbf{A}^m$ that correspond to the spatial derivatives in the s-, r_d-, and r_f-directions, respectively. The term $r_d u$ in (3) is distributed evenly over $\mathbf{A}_1^m$, $\mathbf{A}_2^m$ and $\mathbf{A}_3^m$. The FD discretization for the spatial variable described in (4) implies that if the grid points are ordered appropriately, then $\mathbf{A}_1^m$, $\mathbf{A}_2^m$ and $\mathbf{A}_3^m$ are tridiagonal. (There is a different ordering for each of $\mathbf{A}_1^m$, $\mathbf{A}_2^m$ and $\mathbf{A}_3^m$.) The following splitting scheme based on the Hundsdorfer and Verwer (HV) approach [8] generates an approximation $\mathbf{u}^m$ to the exact solution u^m successively for $m = 1, 2, \ldots, l+1$:[1]

$$\begin{cases} \mathbf{v}_0 = \mathbf{u}^{m-1} + \Delta\tau(\mathbf{A}^{m-1}\mathbf{u}^{m-1} + \mathbf{g}^{m-1}), \\ (\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{A}_i^m)\mathbf{v}_i = \mathbf{v}_{i-1} - \frac{1}{2}\Delta\tau\mathbf{A}_i^{m-1}\mathbf{u}^{m-1} + \frac{1}{2}\Delta\tau(\mathbf{g}_i^m - \mathbf{g}_i^{m-1}), \quad i = 1, 2, 3, \\ \widetilde{\mathbf{v}}_0 = \mathbf{v}_0 + \frac{1}{2}\Delta\tau(\mathbf{A}^m\mathbf{v}_3 - \mathbf{A}^{m-1}\mathbf{u}^{m-1}) + \frac{1}{2}\Delta\tau(\mathbf{g}^m - \mathbf{g}^{m-1}), \\ (\mathbf{I} - \frac{1}{2}\Delta\tau\mathbf{A}_i^m)\widetilde{\mathbf{v}}_i = \widetilde{\mathbf{v}}_{i-1} - \frac{1}{2}\Delta\tau\mathbf{A}_i^m\mathbf{v}_3, \quad i = 1, 2, 3, \\ \mathbf{u}^m = \widetilde{\mathbf{v}}_3. \end{cases}$$

Here, the vector $\mathbf{g}^m$ is given by $\mathbf{g}^m = \sum_{i=0}^{3}\mathbf{g}_i^m$ where $\mathbf{g}_i^m$ are obtained from boundary conditions corresponding to the respective derivative terms. The above splitting scheme treats the mixed derivative part $\mathbf{A}_0^m$ in a fully explicit way while the $\mathbf{A}_i^m$ parts, $i = 1, 2, 3$, are treated implicitly. Since the matrices $\mathbf{A}_i^m$, $i = 1, 2, 3$ are tridiagonal, the number of floating point operations per time step is directly proportional to npq, which yields a big reduction in computational cost compared to solving (6) by a direct method. The HV scheme has been proved to be unconditionally stable for arbitrary spatial dimensions [8].

4. Pricing PRDC swaps

For the past few years, PRDC swaps have been one of the most widely traded and liquid cross-currency exotics. Readers are referred to [1] for a detailed discussion on the dynamics of PRDC swaps. PRDC swaps are essentially long-dated (usually 30 years or more) swaps which pay FX-linked coupons in exchange for LIBOR floating-rate payments. We investigate PRDC swaps from the perspective of the *payer* of PRDC coupons (the receiver of the floating-rate payments).[2] The other party of the deal is the *investor* (the receiver of the PRDC coupons). The floating-rate payments are termed the *funding leg*. Both the coupon rate and the floating rate are applied on the domestic currency principal N_d. More specifically, suppose that we have the following tenor structure:

$$T_0 = 0 < T_1 < \cdots < T_{\beta-1} < T_\beta = T, \quad \nu_\alpha = \nu(T_{\alpha-1}, T_\alpha) = T_\alpha - T_{\alpha-1}, \quad \alpha = 1, 2, \ldots, \beta - 1.$$

Here, ν_α represents the year fraction between $T_{\alpha-1}$ and T_α using the Actual/365 day counting convention. The PRDC coupon rate C_α of the coupon amount $\nu_\alpha C_\alpha N_d$ issued at time T_α for the period $[T_\alpha, T_{\alpha+1}]$, $\alpha = 1, 2, \ldots, \beta - 1$, has the following structure:

$$C_\alpha = \min\Big(\max\Big(c_f\frac{s(T_\alpha)}{f_\alpha} - c_d, b_f\Big), b_c\Big), \qquad \alpha = 1, \ldots, \beta - 1. \tag{10}$$

Here, $s(T_\alpha)$ is the spot FX rate at time T_α; c_d and c_f are domestic and foreign coupon rates; b_f and b_c are the floor and cap of the payoff. The scaling factor f_α usually is set to the forward FX rate $F(T_0, T_\alpha)$. All parameters can vary from coupon to coupon (i.e. they may depend on T_α, $\alpha = 1, \ldots, \beta - 1$). In the standard structure, in which $b_f = 0$ and $b_c = \infty$, by letting $h_\alpha = \frac{c_f}{f_\alpha}$ and $k_\alpha = \frac{f_\alpha c_d}{c_f}$, the coupon rate C_α can be viewed as a call option on spot FX rates, since

$$C_\alpha = h_\alpha \max(s(T_\alpha) - k_\alpha, 0). \tag{11}$$

[1] This is the scheme (1.4) in [8] with $\theta = \mu = \frac{1}{2}$.
[2] Usually, the payer of PRDC coupons is a bank.

In (11), the *option notional* h_α determines the overall level of the coupon payment, while the strike k_α determines the likelihood of the positiveness of the coupon. It is important to emphasize that, if the strike k_α is low, the coupon has a relatively high chance of paying a positive amount. However, in this case, the option notional h_α is typically chosen to be low and thus the overall level of a coupon payment is small. This is a *low-leverage* situation. On the other hand, if both k_α and h_α are high, then we have a *high-leverage* situation.

The funding leg pays the amount $\nu_\alpha L_d(T_{\alpha-1},T_\alpha)N_d$ at time $\{T_\alpha\}_{\alpha=1}^{\beta-1}$ for the period $[T_{\alpha-1},T_\alpha]$, where $L_d(T_{\alpha-1},T_\alpha)$ denotes the domestic LIBOR rates for the period $[T_{\alpha-1},T_\alpha]$, as observed at time $T_{\alpha-1}$. Note that $L_d(T_{\alpha-1},T_\alpha)$ is set at time $T_{\alpha-1}$, but the actual floating leg payment for the period $[T_{\alpha-1},T_\alpha]$ does not occur until time T_α, i.e. "in arrears". There is an initial fixed-rate coupon paid to the investor that is not included in the above definition as its valuation is straightforward.

Since a "vanilla" PRDC swap can be seen as a collection of simple FX options with different maturities (see (11)), its valuation is relatively simple. Let $u^c_\alpha(t)$ and $u^f_\alpha(t)$ be the value at time t of all PRDC coupons and floating payments, respectively, of the "vanilla" PRDC swap scheduled on or after $T_{\alpha+1}$. The value $u^f_\alpha(t)$ can be obtained using the "fixed notional" method and not by solving the PDE. The payoff of the coupon part at each $\{T_\alpha\}_{\alpha=1}^{\beta-2}$ is

$$u^c_\alpha(T_\alpha)+\nu_\alpha C_\alpha N_d. \tag{12}$$

The value of the payoff (12) at time $T_{\alpha-1}$ can be obtained by solving backward in time the PDE (3) from T_α to $T_{\alpha-1}$, with terminal condition (12). The time iteration for the coupon part starts at the time $T_{\beta-1}$ with

$$u^c_{\beta-1}(T_{\beta-1})=\nu_{\beta-1}C_{\beta-1}N_d,$$

and by progressing backward in time to T_0, we obtain $u^c_0(T_0)$. The value of the "vanilla" PRDC swap is $u^f_0(T_0)-u^c_0(T_0)$.

It is important to emphasize that variations of PRDC swaps with exotic features, such as *Bermudan cancelable* (or *cancelable* for short), are much more popular than "vanilla" PRDC swaps. Cancelable PRDC swaps give the payer of the coupons the right to cancel the swap at any of the dates $\{T_\alpha\}_{\alpha=1}^{\beta-1}$.[3] Such features are designed to limit the downside risk arising from excessive movements in spot FX rate for the payer.

The key observation in valuing cancelable swaps is that terminating a swap is the same as (i) continuing the original swap and (ii) entering into the offsetting swap. Since the payer has the option to cancel the PRDC swap on any of the dates $\{T_\alpha\}_{\alpha=1}^{\beta-1}$, we can regard a cancelable PRDC swap as a "vanilla" PRDC swap with the same tenor structure, referred to as the *underlying swap*, plus a long position in a Bermudan swaption, the underlying swap of which is a "vanilla" swap with the same tenor structure, but involves coupons being received and domestic floating payments being paid. We refer to this Bermudan swaption and its underlying swap as the *offsetting Bermudan swaption* and the *offsetting swap*, respectively. Denote $u^e_\alpha(t)$ the value at time t of all fund flows in the offsetting swap scheduled on or after $T_{\alpha+1}$. Let $u^h_\alpha(t)$ be the value at time t of the offsetting Bermudan swaption that has only the dates $\{T_{\alpha+1},\ldots,T_{\beta-1}\}$ as exercise opportunities. Assume optimal exercise at each of $\{T_\alpha\}_{\alpha=1}^{\beta-1}$, i.e. the coupon payer will exercise the offsetting Bermudan swaption at T_α if and only if the value $u^e_\alpha(T_\alpha)$ (the "exercise value") exceeds the value $u^h_\alpha(T_\alpha)$ (the "hold value") of the option. Thus the payoff of the offsetting Bermudan swaption at each T_α is

$$\max(u^h_\alpha(T_\alpha),u^e_\alpha(T_\alpha)).$$

We can progress backward in time in the same fashion as described above for a "vanilla" PRDC swap, starting with

$$u^h_{\beta-1}(T_{\beta-1})=u^e_{\beta-1}(T_{\beta-1})=0.$$

[3]Another popular exotic feature is the *knock-out* provision, which stipulates that the swap terminates if the FX rate on any of the dates $\{T_\alpha\}_{\alpha=1}^{\beta-1}$ exceeds a specified level.

Note that the value $u^e_\alpha(T_\alpha)$ can be computed by $u^e_\alpha(T_\alpha) = -(u^f_\alpha(T_\alpha) - u^c_\alpha(T_\alpha))$. The value of the cancelable PRDC swap is $u^h_0(T_0) + (u^f_0(T_0) - u^c_0(T_0))$.

5. Numerical results

We consider the same short rate models, correlation parameters, and local volatility function for the spot FX rate as given in [2]. In this cross-currency example, the Japanese yen (JPY) and the U.S. dollar (USD) are the domestic and foreign currencies, respectively. Their interest rate curves given by $P_d(0,T) = \exp(-0.02\times T)$ and $P_f(0,T) = \exp(-0.05 \times T)$, where $P_d(0,T)$ and $P_f(0,T)$ denote the current market values in domestic and foreign currency of a bond with maturity T, respectively. The volatility parameters for the short rates and correlations are given by $\sigma_d(t) = 0.7\%$, $\kappa_d(t) = 0.0\%$, $\sigma_f(t) = 1.2\%$, $\kappa_f(t) = 5.0\%$, $\rho_{df} = 25.0\%$, $\rho_{ds} = -15.0\%$, $\rho_{fs} = -15.0\%$. The initial spot FX rate is set to $s(0) = 105.00$. The parameters $\xi(t)$ and $\varsigma(t)$ of the local volatility function are assumed to be piecewise constant and given in the following table.

period (years)	$(\xi(t))$	$(\varsigma(t))$	period (years)	$(\xi(t))$	$(\varsigma(t))$
(0 0.5]	9.03%	-200%	(7 10]	13.30%	-24%
(0.5 1]	8.87%	-172%	(10 15]	18.18%	10%
(1 3]	8.42%	-115%	(15 20]	16.73%	38%
(3 5]	8.99%	-65%	(20 25]	13.51%	38%
(5 7]	10.18%	-50%	(25 30]	13.51%	38%

Note the parameters $\theta_i(t)$, $i = d, f$, associated the domestic and foreign short rates are fully determined by the above information [9]. We consider a PRDC swap that has following features:

- Tenor structure: $\nu_\alpha = T_\alpha - T_{\alpha-1} = 1$ (year), $\alpha = 1, \ldots, \beta - 1$ and $\beta = 30$ (years).
- Pay annual PRDC coupons and receive annual domestic LIBOR payments.
- Standard structure, i.e. $b_f = 0$, $b_c = +\infty$. The scaling factor $\{f_\alpha\}_{\alpha=1}^{\beta-1}$ is set to $F(0, T_\alpha)$.
- Bermudan cancelable, which allows the payer to cancel the swap on each of $\{T_\alpha\}_{\alpha=1}^{\beta-1}$.
- The domestic and foreign coupons are chosen to provide different levels of leverage: low ($c_d = 2.25\%, c_f = 4.50\%$), medium ($c_d = 4.36\%, c_f = 6.25\%$), high ($c_d = 8.1\%, c_f = 9.00\%$).

Note that the forward FX rate $F(0, T_\alpha)$ can be obtained from the domestic and foreign interest rate curves via the well-known formula

$$F(0, T_\alpha) = \frac{P_f(0, T_\alpha)}{P_d(0, T_\alpha)} s(0),$$

which follows from no-arbitrage arguments. The domestic LIBOR rate $L_d(T_{\alpha-1}, T_\alpha)$, as observed at time $T_{\alpha-1}$ for the maturity T_α, can be computed by

$$L_d(T_{\alpha-1}, T_\alpha) = \frac{1 - P_d(T_{\alpha-1}, T_\alpha)}{\nu(T_{\alpha-1}, T_\alpha) P_d(T_{\alpha-1}, T_\alpha)}.$$

The truncated computational domain Ω is defined by setting $S = 3s(0) = 315$, $R_d = 3r_d(0) = 0.06$, and $R_f = 3r_f(0) = 0.15$. For the GMRES method, the tolerance is 10^{-5}. Selected numerical results are presented in Table 1. Grid sizes indicated in Table 1 are for each period $[T_{\alpha-1}, T_\alpha]$, $\alpha = 1, \ldots, \beta - 1$. The values of swaps are expressed as a percentage of the notional N_d. In terms of accuracy, both the ADI and the preconditioned GMRES methods give identical prices to four digits of accuracy for the underlying PRDC swap and the cancelable PRDC swap. (Hence, we do not present prices obtained by the two methods separately to save space in the table.) As expected from the discretization methods, second-order accuracy is obtained. Since for different leverage levels, the computation times of the methods considered are virtually the same, we report only the selected computation statistics for the low leverage case in the last three columns of Table 1 obtained from pricing the underlying PRDC swap. We note that, asymptotically, for each doubling

leverage	m	n	p	q	underlying swap			cancelable swap			performance		
					ADI – GMRES						ADI	GMRES	
					value (%)	change	ratio	value(%)	change	ratio	time (s)	time (s)	avg. iter.
low	4	12	6	6	-11.41			11.39			0.78	1.19	5
	8	24	12	12	-11.16	2.5e-3		11.30	8.6e-4		8.59	12.27	6
	16	48	24	24	-11.11	5.0e-4	5.0	11.28	1.7e-4	5.0	166.28	253.35	6
	32	96	48	48	-11.10	1.0e-4	5.0	11.28	4.1e-5	4.1	3174.20	4882.46	6
medium	4	12	6	6	-13.87			13.42					
	8	24	12	12	-12.94	9.3e-3		13.76	3.3e-3				
	16	48	24	24	-12.75	1.9e-3	4.7	13.85	9.5e-4	3.5			
	32	96	48	48	-12.70	5.0e-4	3.9	13.88	2.6e-4	3.6			
high	4	12	6	6	-13.39			18.50					
	8	24	12	12	-11.54	1.8e-2		19.31	8.1e-3				
	16	48	24	24	-11.19	3.5e-3	5.2	19.56	2.5e-3	3.2			
	32	96	48	48	-11.12	8.0e-4	4.3	19.62	5.4e-4	4.6			

Table 1: Values of the underlying PRDC swap and cancelable PRDC swap with FX skew for various leverage levels; "change" is the difference in the solution from the coarser grid; "ratio" is the ratio of the changes on successive grids; "avg. iter." is the average number of iterations.

of the number of timesteps and gridpoints in all directions, both the ADI and GMRES computation times increase by a factor of about 19, which is close to the optimal factor of 16. It is also evident the ADI method is modestly more efficient than the GMRES method, in absolute terms. It is worth noting that the average number of iterations required by the GMRES method per timestep is quite small, and more importantly, is independent of the size of discretized problem. These results show the combined effect of using an effective preconditioner and a good initial guess based on linear extrapolation.

To investigate the effects of the FX skew, we compare our numerical results with those obtained under the log-normal model, where the local volatility function is a deterministic function of only the time variable. To this end, we used the parametrization as in (2) but independent of $s(t)$ for the log-normal local volatility function, and calibrated it to the same at-the-money FX option data (Table A of [2]) that was used for the calibration of the skew model. Our experiments show that the values under the log-normal model corresponding to the finest mesh in Table 1 for the three leverage levels are -9.01, -9.67, and -9.85, respectively, for the underlying PRDC swap; for the cancelable PRDC swap, the values are 13.31, 16.89, and 22.95, respectively. First, we consider the effect of the FX skew on the underlying PRDC swap. Negative values of the underlying swap indicate the price that the investor has to pay to the coupon payer to enter into a "vanilla" PRDC swap. It is important to emphasize that due to the rate differential between JPY and USD, the forward FX curve is strongly downward sloping, hence in (11), f_α is considerably smaller than $s(T_\alpha)$. Thus the coupon payer essentially shorts a collection of FX call options with low strikes. (For the low, medium, and high leverage cases, the strike k_α is set to 50%, 70% and 90% of f_α, respectively.) The numerical results indicate that the prices of the underlying swap under the skew model are more negative than the prices under the log-normal model (for example, -11.10 versus -9.01). These results are expected, since, in a skew model, the implied volatility increases for low-strike options, resulting in higher prices for the options and hence pushes down the value of the underlying swap for the payer. However, the effect of the skew is not uniform across the leverage levels. The effect seems most pronounced for the medium-leverage swaps. An explanation for this observation is that the total effect is a combination of the change in implied volatility and the sensitivities (the Vega) of the options' prices to that change. Due to the skew, the lower the strikes are, the higher the implied volatility changes are. Thus, among the three leverage levels, the volatilities change the most for the low-leverage swaps, since the strikes of the coupon rates are the lowest in this case. However, it is important to note that the Vega of an option is an increasing function of the strike [10]. Thus, the Vega of low-leverage options is the smallest, since the strikes for coupon rates are the lowest. As a result, the combined effect is limited. The situation is reversed for high-leverage swaps, while the combined effect is the most pronounced for medium-leverage swaps. For cancelable PRDC swaps, the impact of the FX skew is increasing across the leverage levels. The positive values of the cancelable swap indicate the level of the initial fixed coupon that the payer is willing to pay to the investor to enter the cancelable PRDC swap. Under both skew and log-normal models, a high-leverage cancelable PRDC swap provides more attractive initial coupons to the investor.

It is important to note that the prices of the underlying PRDC swap and the cancelable PRDC swap are pushed down under the FX skew model as compared to those obtained by the log-normal model. These changes in values are quite significant and are considered as profits for the payer when a FX skew is incorporated. In other words, not accounting for the FX skew can result in a loss, a fact that indicates the importance of having a proper skew model for pricing and risk managing PRDC swaps.

6. Summary and future work

We have proposed a general PDE pricing framework for exotic cross-currency interest rate derivatives under a FX skew model, with strong emphasis on Bermudan cancelable PRDC swaps. Over each period of the tenor structure, we partition the pricing of Bermudan cancelable PRDC swaps into two entirely independent pricing subproblems: (i) the pricing of the underlying PRDC swap and (ii) the pricing of an associated Bermudan swaption, each of which can be solved efficiently. We consider two numerical methods for the solution of each of the subproblems. Both methods are built upon second-order central FD on uniform grids for the discretization of the space variables and differ primarily in the temporal discretization, with one method using the Crank-Nicolson discretization and the other using the level-splitting ADI method. In the former case, the GMRES method is employed for the solution of the resulting block banded linear system at each time step, with the preconditioner solved by FFT techniques. Experimental results verify that our methods are second-order, and that ADI scheme is modestly faster than the other method. Preliminary analysis shows that cancelable PRDC swaps are very sensitive to the FX volatility skew, which highlights the importance of having a realistic FX skew model for pricing and risk managing PRDC swaps.

A possible extension of this work is to use non-uniform meshes refined around the initial FX and the initial short rates to improve the performance of the numerical methods. In this case, a different fast solver for the preconditioner needs to be developed and the stability of the ADI method on a non-uniform mesh needs to be studied.

From a modeling perspective, due to the sensitivity to the FX volatility skew of the PRDC swaps with exotic features, it is highly desirable to have a mechanism that more accurately approximates the observed FX volatility skew. This could possibly be achieved by incorporating stochastic volatility into the spot FX rate model so that the market-observed FX volatility smiles are more accurately simulated. This enrichment to the current model leads to a time-dependent PDE in four state variables — the spot FX rate, domestic and foreign short rates, and volatility. In such a application, numerical methods presented in this paper could be easily extended to cope with an extra spatial dimension, but the computational work required to solve the extended problem would likely rise significantly. Possibly an effective parallel numerical method could be developed to solve the extended problem in an acceptable amount of time.

Acknowledgment

This research was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada. The first author would like to thank Vicky Huang, formerly at Algorithmics Inc., Toronto, for helping with the spreadsheet that illustrates the dynamics of PRDC swaps.

References

[1] J. Sippel, S. Ohkoshi, All power to PRDC notes, Risk 15 (11) (2002) 31–33.

[2] V. Piterbarg, Smiling hybrids, Risk 19 (2) (2005) 66–70.

[3] J. Hull, A. White, One factor interest rate models and the valuation of interest rate derivative securities, Journal of Financial and Quantitative Analysis 28 (2) (1993) 235–254.

[4] M. Dempster, J. Hutton, Numerical valuation of cross-currency swaps and swaptions, Cambridge University Press, 1997.

[5] H. Windcliff, P. A. Forsyth, K. R. Vetzal, Analysis of the stability of the linear boundary condition for the Black-Scholes equation, Journal of Computational Finance 8 (2004) 65–92.

[6] W. Hackbush, Iterative Solution of Large Sparse Systems of Equations, Springer, 1993.

[7] C. V. Loan, Computational Frameworks for the Fast Fourier Transform, Society for Industrial and Applied Mathematics, 1992.

[8] K. in't Hout, B. Welfert, Unconditional stability of second-order ADI schemes applied to multi-dimensional diffusion equations with mixed derivative terms, Applied Numerical Mathematics 59 (2009) 677–692.

[9] D. Brigo, F. Mercurio, Interest Rate Models - Theory and Practice, 2nd Edition, Springer, 2006.

[10] J. C. Hull, Options, Futures, and Other Derivatives, sixth Edition, Prentice Hall, 2006.