Wavelet Methods in PDE Valuation of Financial Derivatives

M.A.H. Dempster, A. Eswaran & D.G. Richards

Centre for Financial Research

Judge Institute of Management

University of Cambridge

Email: {mahd2, ae206, dgr22}@cam.ac.uk

Web: www-cfr.jims.cam.ac.uk

Abstract. We investigate the application of a wavelet method of lines solution method to financial PDEs. We demonstrate the suitability of a numerical scheme based on biorthogonal interpolating wavelets to financial PDE problems where there are discontinuities or regions of sharp transitions in the solution. The examples treated are the Black Scholes PDE with discontinuous payoffs and a 3-dimensional cross currency swap PDE for which a speedup over standard finite difference methods of two orders of magnitude is reported.

1 Introduction

What are wavelets ? Wavelets are nonlinear functions which can be scaled and translated to form a basis for the Hilbert space $L^2(\mathbb{R})$ of square integrable functions. Thus wavelets generalize the trignometric functions given by $e^{ist}(s \in \mathbb{R})$ which generate the classical Fourier basis for L^2. It is therefore not surprising that wavelet and fast wavelet transforms exist which generalize the time to frequency map of the Fourier transform to pick up both the space and time behaviour of a function [8]. Wavelets have been used in the field of image compression and image analysis for quite some time. Indeed the main motivation behind the development of wavelets was the search for fast algorithms to compute compact representations of functions and data sets based on exploiting structure in the underlying functions. In the solution of PDE's using wavelets [4], [1], [15, 16], [5] functions and operators are expanded in a wavelet basis to allow a combination of the desirable features of finite-difference methods, spectral methods and front-tracking or adaptive grid approaches. The advantages of using wavelets to solve PDE's that arise in finance are that large classes of operators and functions which occur in this area are sparse, or sparse to some high accuracy, when transformed into the wavelet domain. Wavelets are also suitable for problems with multiple spatial scales (which occur frequently in financial problems) since they give an accurate representation of the solution in regions of sharp transitions and combine the advantages of both spectral and finite-difference methods.

In this paper we implement a wavelet method of lines scheme using biorthogonal wavelets to solve the Black Scholes PDE for option values with discontinuous payoff structures and a 3-dimensional cross currency swap PDE based on extended Vasicek interest rate models. We demonstrate numerically the advantages of using a wavelet based PDE method in solving these kind of problems. The paper is organized as follows. In Section 2 we give a brief introduction to wavelet theory. In Sections 3 and 4 we give an explanation of wavelet based PDE methods and explain the biorthogonal wavelet approach in more detail. Sections 5 and 6 contain respectively the problems and numerical results for the Black Scholes and cross currency swap PDEs and Section 7 concludes and describes research in progress.

2 Basic Wavelet Theory

We now give a brief introduction to wavelets for real valued functions of a real argument. Further detail can be found in the cited references and [6] and we shall extend the concepts needed for this paper to higher dimensions in the sequel.

Daubechies based wavelets

Consider two functions: the scaling function ϕ and the wavelet function ψ.

The *scaling function* is the solution of a *dilation equation*

$$\phi(x) = \sqrt{2}\sum_{k=0}^{\infty} h_k \phi(2x-k),$$

where ϕ is normalised so that $\int_{-\infty}^{\infty}\phi(x)dx = 1$ and the *wavelet function* is defined in terms of the scaling function as

$$\psi(x) = \sqrt{2}\sum_{k=0}^{\infty} g_k \phi(2x-k).$$

We can build up a *orthonormal basis* for the *Hilbert space* $L^2(\mathbb{R})$ of (equivalence classes of) square integrable functions from the functions ϕ and ψ by dilating and translating them to obtain the *basis functions*:

$$\phi_{j,k}(x) = 2^{-j/2}\phi(2^{-j}x-k) = 2^{-j/2}\phi\left(\frac{x-2^j k}{2^j}\right) \tag{1}$$

$$\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x-k) = 2^{-j/2}\psi\left(\frac{x-2^j k}{2^j}\right). \tag{2}$$

In the above equations j is the *dilation* or *scaling parameter* and k is the *translation parameter.* All wavelet properties are specified through the coefficients $H := \{h_k\}_{k=0}^{\infty}$ and $G := \{g_k\}_{k=0}^{\infty}$ which are chosen so that dilations and translations of the wavelet $\psi_{j,k}$ form an orthonormal basis of $L^2(\mathbb{R})$. In other words the $\psi_{j,k}$ will satisfy

$$\int_{-\infty}^{\infty}\psi_{j,k}(x)\psi_{l,m}(x)dx = \delta_{jl}\delta_{km} \qquad j,k,l,m \in \mathbb{Z}_+,$$

where $\mathbb{Z}_+ := \{0, 1, 2, \dots\}$ and δ_{jl} is the Kronecker delta function.

Under these conditions for any function $f \in L^2(\mathbf{R})$ there exists a set $\{d_{jk}\}$ such that

$$f(x) = \sum_{j\in\mathbb{Z}_+} \sum_{k\in\mathbb{Z}_+} d_{jk}\psi_{j,k}(x), \tag{3}$$

where

$$d_{jk} := \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx.$$

It is usual to denote the spaces spanned by $\phi_{j,k}$ and $\psi_{j,k}$ over the parameter k with j fixed by

$$\mathbf{V}_j := span_{k\in\mathbb{Z}_+}\phi_{j,k},$$

$$\mathbf{W}_j := span_{k\in\mathbb{Z}_+}\psi_{j,k}.$$

In the expansion (3) functions with arbitrary small scales can be represented, however in practice there is a limit on how small the smallest structure can be. (This could for example be dependent on a required grid size in a numerical computation as we shall see below.) To implement wavelet analysis on a computer, we need to have a bounded range and domain to generate approximations to functions $f \in L^2(\mathbb{R})$ and thus must limit H and G to *finite* sets termed *filters.* Approximation *accuracy* is specified by requiring that the wavelet function ψ satisfies

$$\int_{-\infty}^{\infty} \psi(x)x^m dx = 0 \tag{4}$$

for $m = 0, \dots, M-1$, which implies exact approximation for polynomials of degree $M-1$. For *Daubechies wavelets* [6] the number of coefficients or the *length* L of the filters H and G is related to the number of vanishing moments M in (4) by $2M = L$. In addition elements of H and G are related by $g_k = (-1)^k h_{L-k}$ for $k = 0, \dots, L-1$ and the two finite sets of coefficients H and G are known in the signal processing literature as *quadrature mirror* filters. The coefficients H needed to define compactly supported wavelets with high degrees of regularity can be derived [6] and the usual notation to denote a Daubechies based wavelet defined by coefficients H of length L is D_L. Therefore on a computer an approximation subspace expansion would be in the form of a finite direct sum of finite dimensional vector spaces as

$$\mathbf{V_0} = \mathbf{W}_1 \oplus \mathbf{W}_2 \oplus \mathbf{W}_3 \oplus \cdots \oplus \mathbf{W}_J \oplus \mathbf{V}_J,$$

and the corresponding orthogonal wavelet series *approximation* to a continuous function f on a compact domain is given by

$$f(x) \approx \sum_k d_{J,k}\psi_{J,k}(x) + \sum_k d_{J-1,k}(x)\psi_{J-1,k}(x) + \cdots + \sum_k d_{1,k}\psi_{1,k}(x) + \sum_k s_{J,k}\phi_{J,k}, \tag{5}$$

where J is the number of *multiresolution components* (or *scales*) and k ranges from 1 to the number of coefficients in the specified component. The spaces W_j and V_j are termed *scaling function* and *approximation subspaces* respectively. The coefficients $s_{J,k}, d_{J,k}, \dots, d_{1,k}$ are termed the *wavelet transform coefficients* and the functions $\phi_{J,k}$ and $\psi_{J,k}$ are the *approximating wavelet functions.* Some examples of basic wavelets are the *Haar wavelet* which is just a square wave (the indicator function of the unit interval), the *Daubechies wavelets* [6] and *Coiflet wavelets* [2].

Biorthogonal wavelets

Biorthogonal wavelets are a generalization of orthogonal wavelets first introduced by Cohen, Daubechies and Feauveau [3]. Biorthogonal wavelets are symmetric and do not introduce phase shifts in the coefficients. In biorthogonal wavelet analysis we have four basic function types $\phi, \psi, \tilde{\phi}$ and $\tilde{\psi}$. The functions ϕ and ψ are termed *mother* and *father* wavelets and the functions $\tilde{\phi}$ and $\tilde{\psi}$ are the *dual wavelets.* The father and mother wavelets are used to compute the wavelet coefficients as in the orthogonal case, but now the *biorthogonal wavelet approximation* of a continuous function on a compact domain is expressed in terms of the dual wavelet functions as

$$f(x) \approx \sum_k d_{J,k}\tilde{\psi}_{J,k}(x) + \sum_k d_{J-1,k}(x)\tilde{\psi}_{J-1,k}(x) + \cdots + \sum_k d_{0,k}\tilde{\psi}_{0,k}(x) + \sum_k s_{J,k}\tilde{\phi}_{J,k}. \qquad (6)$$

In signal processing ϕ and ψ are used to *analyze* the signal and $\tilde{\phi}$ and $\tilde{\psi}$ are used to *synthesize* the signal. In general biorthogonal wavelets are not mutually orthogonal, but they do satisfy *biorthogonal* relationships of the form

$$\int \phi_{j,k}\tilde{\phi}_{j',k'}(x)dx = \delta_{j,j'}\delta_{k,k'}$$

$$\int \phi_{j,k}\tilde{\phi}_{j',k}(x)dx = 0$$

$$\int \psi_{j,k}\tilde{\phi}_{j,k'}(x)dx = 0$$

$$\int \psi_{j,k}\tilde{\psi}_{j',k'}(x)dx = \delta_{j,j'}\delta_{k,k'}.$$

3 Wavelets and PDE's

Wavelet based approaches to the solution of PDE's have been presented by Vasilyev *et al* [15, 16], Beylkin [1], Prosser and Cant [11], Cohen *et al* [4], Dahmen *et al* [5] and Xu and Shann [18]. There are two main approaches to the numerical solution of PDEs using wavelets. Consider the most general form for a system of parabolic PDEs given by

$$\frac{\partial u}{\partial t} = F(x, t, u, \nabla u)$$

$$\Phi(x, t, u, \nabla u) = 0, \qquad (7)$$

which describe the time evolution of a vector valued function u and the boundary conditions are possibly algebraic or differential constraints. The *wavelet-Galerkin* method assumes that the wavelet coefficients are functions of time. An appropriate wavelet decomposition for each component of the solution is substituted into (7) and a Galerkin projection is used to derive a nonlinear system of ordinary differential-algebraic equations which describe the time evolution of the wavelet coefficients. In a *wavelet-collocation* method (7) is evaluated at collocation points of the domain of u and a system of nonlinear ordinary differential-algebraic equations describing the evolution of the solution at these collocation points is obtained. If we want the numerical algorithm to be able to resolve all structures appearing in the solution and also to be efficient in terms of minimising the number of unknowns, the *basis* of active wavelets and consequently the *computational grid* for the wavelet-collocation algorithm should adapt dynamically in time to reflect local changes in the solution. This adaptation of the wavelet basis or computational grid is based on analysis of the wavelet coefficients. The contribution of a particular wavelet to the approximation is significant if and only if the nearby structures of the solution have a size comparable with the wavelet scale. Thus using a *thresholding* technique a large number of the fine scale wavelets may be dropped in regions where the solution is smooth. In the wavelet-collocation method every wavelet is uniquely associated with a collocation point. Hence a collocation point can be omitted from the grid if the associated wavelet is omitted from the approximation. This property of the *multilevel* wavelet approximation allows local grid refinement up to a prescribed small scale without a drastic increase in the number of collocation points. A fast adaptive wavelet collocation algorithm for two dimensional PDE's is presented in [15] and a spatial discretization scheme using bi-orthogonal wavelets is implemented in [9–11]. The wavelet scheme is used in the latter to solve the *reacting Navier-Stokes* equations and the main advantage of the approach is that when the solution is computed in wavelet space it is possible to exploit sparsity in order to reduce storage costs and speed up solution times. We will now explain the wavelet-collocation method in greater detail.

The biorthogonal wavelet approach

The main difference in using biorthogonal systems is that we have both *primal* and *dual* basis functions derived from primal and dual scaling and wavelet functions. Biorthogonal wavelet systems are derived from a *paired* hierarchy of approximation subspaces

$$\begin{aligned} &\mathbf{V}_{J-1} \subset \mathbf{V}_J \subset \mathbf{V}_{J+1} \\ &\tilde{\mathbf{V}}_{J-1} \subset \tilde{\mathbf{V}}_J \subset \tilde{\mathbf{V}}_{J+1}. \end{aligned} \tag{8}$$

(Note that here increasing j denotes refinement of the grid, although some authors in the wavelet literature use an increasing scale index j to indicate its coarsening.) For periodic discretizations $dim(\mathbf{V}_J) = 2^J$. The basis functions for these spaces are the *primal* scaling function ϕ and the *dual* scaling function $\tilde{\phi}$. Define two *innovation spaces* $\mathbf{W_J}$ and $\tilde{\mathbf{W}}_\mathbf{J}$ such that

$$
\begin{aligned}
\mathbf{V}_{J+1} &:= \mathbf{V}_J \oplus \mathbf{W}_J \\
\tilde{\mathbf{V}}_{J+1} &:= \tilde{\mathbf{V}}_J \oplus \tilde{\mathbf{W}}_J
\end{aligned} \tag{9}
$$

where $\tilde{\mathbf{V}}_{J+1} \perp \mathbf{W}_J$ and $\mathbf{V}_J \perp \tilde{\mathbf{W}}_J$. The innovation spaces so defined satisfy

$$
\bigoplus_{i=0}^{\infty} \mathbf{W}_i = \mathbf{L}^2(\mathbb{R}) = \bigoplus_{i=0}^{\infty} \tilde{\mathbf{W}}_i \tag{10}
$$

and the innovation space basis functions are ψ and $\tilde{\psi}$.

Interpolating wavelet transform

The *interpolating wavelet transform* has basis functions

$$
\begin{aligned}
\phi_{j,k}(x) &= \phi(2^j x - k) \\
\psi_{j,k}(x) &= \phi(2^{j+1} x - 2k - 1) \\
\tilde{\phi}_{j,k}(x) &= \delta(x - x_{j,k}),
\end{aligned} \tag{11}
$$

where $\delta(.)$ is the Dirac delta function. The wavelets are said to be *interpolating* because the primal scaling function ϕ to which they are related satisfies

$$
\phi(k) = \begin{cases} 1 & k = 0, \\ 0 & k \neq 0, \quad k \in \mathbb{Z}_+. \end{cases}
$$

The primal scaling function can be defined through the use of the *two scale* relation

$$
\phi(x) = \sum_{\xi \in \mathbb{Z}_+} \phi(\xi/2)\phi(2x - \xi). \tag{12}
$$

The smoothness of the primal scaling function is dictated by its $(M-1)^{st}$ degree polynomial span which in turn depends on the $M+1$ non-zero values of $\phi(\xi/2)$. Fast transform methods for the evaluation of the wavelet and scaling function coefficients are given in [12,14]. The *projection* of a function f onto a space of scaling functions $\mathbf{V}_J$ is given (discretizing $[0,1]$) by

$$
P_{\mathbf{V}_J} f(x) = \sum_{k=0}^{2^J} s^f_{J,k} \phi^{\square}_{J,k}(x), \tag{13}
$$

where $s^f_{J,k}$ is defined as $< f, \tilde{\phi}_{J,k} >$ in terms of a suitable inner product and $\phi^{\square}$ is used to denote a *boundary* or *internal wavelet* given by

$$
\phi^{\square}_{J,k}(x) = \begin{cases} \phi^L_{J,k}(x) & k = 0, \ldots, M-1 \\ \phi_{J,k}(x) & k = N, \ldots, 2^J - M \\ \phi^R_{J,k}(x) & k = 2^J - M + 1, 2^J. \end{cases} \tag{14}
$$

Fast biorthogonal wavelet transform algorithm

The projection of a function f onto a *finite dimensional* scaling function space $\mathbf{V}_J$ is given as above by

$$\begin{aligned} P_{\mathbf{V}_J} f(x) &= \sum_k < f(u), \tilde{\phi}_{\mathbf{J},k}(u) > \phi_{\mathbf{J},k}(x) \\ &= \sum_k f(k/2^J) \phi_{\mathbf{J},k}(x) \\ &= \sum_k s^f_{\mathbf{J},k} \phi_{\mathbf{J},k}(x), \end{aligned} \tag{15}$$

where $s^f_{\mathbf{J},k} = f(k/2^J)$. The coefficients at resolution level j must be derived using

$$\begin{aligned} P_{\mathbf{W}_j} f(x) &= P_{\mathbf{V}_{j+1}} f(x) - P_{\mathbf{V}_j} f(x) \\ \textstyle\sum_l d^f_{j,l} \psi_{j,l}(x) &= \textstyle\sum_m s^f_{j+1,m} \phi_{j+1,m}(x) - \sum_n s^f_{j,n} \phi_{j,n}(x). \end{aligned} \tag{16}$$

An arbitrary *wavelet coefficient* $d^f_{j,m}$ can be calculated from

$$d^f_{j,m} = s^f_{j+1,2m+1} - \sum_n s^f_{j,n} \phi(m-n+1/2) = s^f_{j+1,2m+1} - \sum_n \Gamma_{mn} s^f_{j,n}, \tag{17}$$

where Γ is a square matrix of size $2^J \times 2^J$ for periodic discretizations defined by

$$\Gamma_{mn} := \phi(m-n+1/2).$$

Because of the compact support of the primal scaling function this matrix has a band diagonal structure and as before each primal scaling function satisfies a *two scale* relation

$$\phi(x) = \sum_\xi \phi(\xi/2) \phi(2x-\xi).$$

The values $\phi(\xi/2)$ can be calculated using a explicit relation as in [10]. Irrespective of the choice of primal scaling function the transform vector that arises from the wavelet transform will have a structure of the form given below.

$$\begin{array}{l}
\{s^f_{J,0}, s^f_{J,1}, s^f_{J,2} \cdots \qquad\qquad\qquad\qquad \cdots s^f_{J,2^J-1}\}^T \\
\qquad\qquad\qquad\qquad \downarrow \\
\{d^f_{J-1,0}, \cdots, d^f_{J-1,2^{J-1}-1} \qquad\qquad \mid s^f_{J-1,0}, s^f_{J-1,1} \cdots s^f_{J-1,2^{J-1}-1}\}^T \\
\qquad\qquad\qquad\qquad \downarrow \\
\{d^f_{J-1,0}, \cdots, d^f_{J-1,2^{J-1}-1}, \mid d^f_{J-2,0}, \qquad \cdots d^f_{J-2,2^{J-2}-1} \mid s^f_{J-2,0} \cdots s^f_{J-2,2^{J-2}-1}\}^T \\
\qquad\qquad\qquad\qquad \downarrow \\
\{d^f_{J-1,0}, \cdots, d^f_{J-1,2^{J-1}-1}, \mid \cdots d^f_{J-P-1,0}, \quad \cdots d^f_{J-P-1,2^{J-P-1}-1} \mid s^f_{J-P-1,0} \cdots s^f_{J-P-1,2^{J-P-1}-1}\}^T \\
\mathbf{W}_{J-1} \oplus \mathbf{W}_{J-2} \oplus \mathbf{W}_{J-3} \oplus \qquad\qquad \cdots \qquad\qquad \mathbf{V}_{J-P}
\end{array} \tag{18}$$

Algorithm complexity

The number of of floating point operations required for the fast biorthogonal wavelet transform algorithm for P *resolution* levels is $2M\sum_{i=J-P}^{J} 2^i = 2^{J-P}M\{2^{P+1}-1\}$. This comes from the fact that we require $2M$ filter coefficients to define the primal scaling function ϕ which spans the space of polynomials of degree less than $M-1$. The calculation of the wavelet coefficients $d^f_{j,k}$ for a given resolution j can be accomplished in $2(M-1)+1$ floating point operations. The sub sampling process for the scaling function coefficients $s^f_{j,k}$ requires a further 2^j operations and a total of $2^{j+1}M$ operations are required per resolution j. Thus for fixed J and P the complexity of the fast interpolating wavelet transform algorithm is $O(M)$ [12]. Since the finest resolution in a PDE spatial grid of N points is $J=\log_2 N$, for fixed M and P the complexity of the transform is $O(N)$.

Decomposition of differential operators

If we define $\partial_J^{(n)}$ by

$$\partial_J^{(n)} f(x) := P_{V_J}\frac{d^n}{dx^n}P_{V_J} f(x),$$

then repeated application of the approximation subspace decomposition gives us

$$\partial_J^{(n)} f(x) := \left(P_{V_{J-p}} + \sum_{i=J-p}^{J-1} P_{\mathbf{W}_i}\right)\frac{d^n}{dx^n}\left(P_{\mathbf{V}_{J-p}} + \sum_{i=J-p}^{J-1} P_{\mathbf{W}_i}\right) f(x). \tag{19}$$

For example, the decomposition of the first derivative operator $\frac{d}{dx}$ is given by

$$\partial_J = \left(P_{V_{J-p}} + \sum_{i=J-p}^{J-1} P_{\mathbf{W}_i}\right)\frac{d}{dx}\left(P_{\mathbf{V}_{J-p}} + \sum_{i=J-p}^{J-1} P_{\mathbf{W}_i}\right), \tag{20}$$

where $\partial_J := W\partial_J^1 W^{-1}$ and W and W^{-1} are matrices denoting the forward and inverse transforms with

$$\partial_J^1 = P_{\mathbf{V}_J}\frac{d}{dx}P_{\mathbf{V}_J}. \tag{21}$$

We can analyze ∂_J^1 instead of ∂_J without loss of generality because the forward and inverse transforms are exact up to machine precision. The matrix ∂_J^1 has a band diagonal structure and can be treated as a finite difference scheme for analysis. The biorthogonal expansion for $\frac{df}{dx}$ requires information on the interaction between the differentiated and undifferentiated scaling functions along with information about both the primal and dual basis functions. Using the sampling nature of the dual scaling function ∂_J^1 can be written as

$$\partial_J^1 = 2^J \sum_{\alpha,k} s^f_{J,k}\frac{d\phi^{\Box}}{dx}\,|_{x=\alpha-k} \tag{22}$$

and using equation (14) we get

$$\partial_J^1 = \begin{cases} 2^J \sum_{\alpha,k} s_{J,k}^f \frac{d\phi^L}{dx} \mid_{x=\alpha-k} \phi_{J,\alpha}^L(x) & k = 0, \dots, M-1 \\ 2^J \sum_{\alpha,k} s_{J,k}^f \frac{d\phi}{dx} \mid_{x=\alpha-k} \phi_{J,\alpha}(x) & k = M, \dots, 2^J - M \\ 2^J \sum_{\alpha,k} s_{J,k}^f \frac{d\phi^R}{dx} \mid_{x=\alpha-k} \phi_{J,\alpha}^R(x) & k = 2^J - M + 1, 2^J. \end{cases} \tag{23}$$

The entire operator ∂_J^1 can be determined provided the values of $r_{\alpha-k}^{(1)} = \frac{d\phi}{dx} \mid_{x=\alpha-k}$ can be obtained. An approach to determining filter coefficients for higher order derivatives is given in [10].

Extension to multiple dimensions

The entire wavelet multiresolution framework presented so far can be extended to several spatial dimensions by taking straightforward tensor products of the appropriate 1D wavelet bases. The imposition of boundary conditions on nonlinearly bounded domains is nontrivial, but these are fortunately rare in derivative valuation PDE problems which are usually Cauchy problems on a strip.

The fast biorthogonal interpolating wavelet transform used with wavelet collocation methods for problems posed over d-dimensional domains exhibits better complexity than its alternatives. Indeed, since one basis function is needed for each collocation point, using a spatial grid of n points in each dimension there are $N := n^d$ points in the spatial domain to result in transform complexity $O(n^d)$ – versus $O(n^d \log_2 n)$ for the *Fast Fourier Transform* (where applicable), $O(n^{2d})$ for an *explicit* finite difference scheme and $O(n^{3d})$ for a *Crank-Nicholson* or *implicit* scheme (which makes these methods impractical for $d > 2$, *cf.* [7]).

4 Wavelet Method of Lines

In a traditional finite difference scheme partial derivatives are replaced with algebraic approximations at grid points and the resulting system of algebraic equations is solved to obtain the numerical solution of the PDE. In the *wavelet method of lines* we transform the PDE into a vector system of ODEs by replacing the spatial derivatives with their wavelet transform approximations but retain the time derivatives. We then solve this vector system of ODEs using a suitable stiff ODE solver. We have implemented both a fourth order Runge Kutta method and a method based on the backward differentiation formula (LSODE) developed at the Lawrence Livermore Laboratories [13]. The fundamental complexity of this method is $O(\tau n^d)$ for space and time discretizations of size n and τ respectively over domains of dimension d (*cf.* §3, [13]).

An example

Consider a first order nonlinear hyperbolic *transport* PDE defined over an interval $\Omega = [x_l, x_r]$:

$$\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{\partial u}{\partial x} + S^{u/\rho} \qquad x \notin \partial\Omega \\
\frac{\partial u}{\partial t} &= -\chi^L(t) \qquad\qquad\qquad x = x_l \\
\frac{\partial u}{\partial t} &= -\chi^R(t) \qquad\qquad\qquad x = x_r.
\end{aligned}$$

The numerical scheme is applied to the wavelet transformed counterpart of the above equations

$$\frac{\partial}{\partial t}\wp_{J-P}^{J-1}(u) = -\partial_J^{(1)} u + \wp_{J-P}^{J-1} S^{u/\rho} \qquad x \notin \partial\Omega,$$

where $\wp_{J-P}^{J-1} := \left(P_{\mathbf{V}_{J-P}} + \sum_{i=J-P}^{J-1} P_{\mathbf{W}_i}\right)$ and $\partial_J^{(1)}$ is the standard decomposition of $\frac{d}{dx}$ defined as $\wp_{J-P}^{J-1}\frac{d}{dx}\wp_{J-P}^{J-1}$. In using the multiresolution strategy to discretize the problem we represent the domain $P+1$ times, where P is the number of different resolutions in the discretization, because of the P wavelet spaces and the coarse resolution scaling function space $\mathbf{V}_{J-P}, P \geq 1$. In the transform domain each representation of the solution defined at some resolution p should be supplemented by boundary conditions and [12] shows how to impose boundary conditions in the both the scaling function spaces and the wavelet spaces.

5 Financial Derivative Valuation PDEs

In this section we introduce briefly the PDEs for financial derivative valuation and the products we have valued using the wavelet method of lines described above. More details may be found in [7,17].

Black Scholes products

We have applied wavelet methods to solve the Black Scholes PDE for a vanilla European call option and two binary options. The *Black Scholes* quasilinear parabolic *PDE* is given by

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0 \tag{24}$$

where S is the *stock price*, σ is *volatility*, r is the *risk free rate* of interest. We transform (24) to the *heat diffusion* equation

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2} \quad \text{for} \quad -\infty < x < \infty, \quad \tau > 0$$

with the transformations

$$S := Ke^x, \quad t := T - \frac{2\tau}{\sigma^2}$$

$$C := e^{-1/2(k-1)x - 1/4(k+1)^2\tau} K u(x,\tau)$$

where $k = 2r/\sigma^2$, K is the *exercise price* and T is the *time to maturity* of the option to be valued. The boundary conditions for the PDE depend on the specific type of option. For a vanilla *European call option* the boundary conditions are:

$$C(0,t) = 0, \qquad C(S,t) \sim S \qquad \text{as} \quad S \to \infty$$

$$C(S,T) = \max(S-K, 0).$$

The boundary conditions for the transformed PDE are:

$$u(x,\tau) = 0 \quad \text{as} \quad x \to -\infty,$$

$$u(x,\tau) = e^{1/2(k+1)x + 1/4(k+1)^2\tau} \quad \text{as} \quad x \to \infty,$$

$$u(x,0) = \max(e^{1/2(k+1)x} - e^{1/2(k-1)x}, 0).$$

The first type of binary option that we solved was the *cash-or-nothing call* option with a payoff given by

$$\Pi(S) = B\mathcal{H}(S-K),$$

where $\mathcal{H}$ is the Heaviside function, i.e the payoff is B if at expiry the stock price $S > K$. The boundary conditions for this option in the transformed domain are

$$u(x,\tau) = 0 \quad \text{as} \quad x \to -\infty,$$

$$u(x,\tau) = \frac{B}{K} e^{\frac{1}{2}(k-1)x + \frac{1}{4}(k+1)^2\tau} \quad \text{as} \quad x \to \infty,$$

$$u(x,0) = e^{\frac{1}{2}(k-1)x} \frac{B}{K} \mathcal{H}(Ke^x - K).$$

The second binary option we solved was a *supershare call* [17] option that pays an amount $1/d$ if the stock price lies between K and $K+d$ at expiry. Its payoff is thus

$$\Pi(S) = \frac{1}{d}(\mathcal{H}(S-K) - \mathcal{H}(S-K-d))$$

which becomes a *delta function* in the limit $d \to 0$. The initial boundary condition for this option is

$$u(x,0) = \frac{1}{dK} e^{\frac{1}{2}(k-1)x} (\mathcal{H}(Ke^x - K) - \mathcal{H}(Ke^x - K - d)).$$

For all of the above options the solution is transformed back to real variables using the transformation

$$C(S,t) = K^{\frac{1}{2}(k+1)} S^{\frac{1}{2}(1-k)} e^{\frac{1}{8}(k+1)^2\sigma^2(T-t)} u(\log(S/K), 1/2\sigma^2(T-t)),$$

where $k = 2r/\sigma^2$. There are closed form solutions for all the above options (see for example [17]). The Black Scholes solution for the vanilla European call option is

$$C(S,t) = SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

$$d_1 := \frac{\log(S/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}}$$

$$d_2 := \frac{\log(S/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}}.$$

The solution for a cash or nothing call is

$$C(S,t) = Be^{-r(T-t)}N(d_2).$$

The solution for the supershare option is

$$C(S,t) = \frac{1}{d}e^{-r(T-t)}(N(d_2) - N(d_3))$$

$$d_3 := d_2 - \frac{\log(1 + \frac{d}{K})}{\sigma\sqrt{T-t}}\ .$$

Cross currency swap products

A *cross currency swap* is a derivative contract between two counterparties to exchange cash flows in their respective domestic currencies. Such contracts are an increasing share of the global swap markets and are individually structured products with many complex valuations. With two economies, i.e one domestic and one foreign, there are different term structure processes and risk preferences in each economy and a rate of currency exchange between them. We will model the interest rates in single factor a extended Vasicek framework.

To value any European-style derivative security whose payoff is a measurable function with respect to a filtration $\mathcal{F}_T$ we may derive a PDE for its value. The *domestic* and *foreign bond prices* and *exchange rate* are specified in terms of the driftless Gaussian state variables X_d, X_f and X_S whose corresponding processes $\mathbf{X}_d$, $\mathbf{X}_f$ and $\mathbf{X}_S$ are sufficient statistics for movements in the term structure dynamics. Let $V = V(X_d, X_f, X_S, t)$ be the *domestic value function* of a security with a terminal payoff measurable with respect to $\mathcal{F}_T$ and no intermediate payments, and assume that $V \in C^{2,1}\left(\mathbb{R}^3 \times [0,T)\right)$. Then the *normalised domestic value* process, defined by

$$V^*(t) := \frac{V(t)}{P_d(t,T)}, \tag{25}$$

satisfies the quasilinear parabolic PDE with time dependent coefficients given by

$$\frac{1}{2}\lambda_d^2\frac{\partial^2 V^*}{\partial X_d^2} + \frac{1}{2}\lambda_f^2\frac{\partial^2 V^*}{\partial X_f^2} + \frac{1}{2}H^{SS}\frac{\partial^2 V^*}{\partial X_S^2} + H^{df}\frac{\partial^2 V^*}{\partial X_d \partial X_f} + H^{dS}\frac{\partial^2 V^*}{\partial X_d \partial X_S} + H^{fS}\frac{\partial^2 V^*}{\partial X_f \partial X_S} + \frac{\partial V^*}{\partial t} = 0, \tag{26}$$

on $\mathbb{R}^3 \times [0, T)$. Here the functions H^{SS}, H^{df}, H^{dS} and H^{fS} are defined by

$$\begin{aligned}
H^{SS}(s) &:= G_d^2(s)\lambda_d^2(s) + G_f^2(s)\lambda_f^2(s) + \sigma_S^2(s) - 2\rho_{df}(s)G_d(s)\lambda_d(s)G_f(s)\lambda_f(s) \\
&\quad +2\rho_{dS}(s)G_d(s)\lambda_d(s)\sigma_S(s) - 2\rho_{fS}(s)G_f(s)\lambda_f(s)\sigma_S(s) \\
H^{df}(s) &:= \rho_{df}(s)\lambda_d(s)\lambda_f(s) \\
H^{dS}(s) &:= \lambda_d(s)\left[G_d(s)\lambda_d(s) - \rho_{df}(s)G_f(s)\lambda_f(s) + \rho_{dS}(s)\sigma_S(s)\right] \\
H^{fS}(s) &:= \lambda_f(s)\left[\rho_{df}(s)G_d(s)\lambda_d(s) - G_f(s)\lambda_f(s) + \rho_{fS}(s)\sigma_S(s)\right] \qquad (27)
\end{aligned}$$

and the volatility is of the form

$$\sigma_k(t, T) = [G_k(T) - G_k(t)]\lambda_k(t) \quad k = d, f. \qquad (28)$$

$$G_k(t) := \frac{1 - e^{-\xi_k t}}{\xi_k} \qquad k = d, f, \qquad (29)$$

for some *mean reversion rates* ξ_d and ξ_f and

$$\lambda_k(t) = e^{\xi_k t}\kappa_k(t) \qquad k = d, f, \qquad (30)$$

where $\kappa_k(t)$ is the *prospective variability* of the short rate. For the derivation of the PDE and further details of the extended Vasicek model see [7]. For a standard European-style derivative security we solve the PDE with the appropriate boundary conditions.

The most common type of cross-currency swap is the exchange of floating or fixed rate interest payments on *notional principals* Z_d and Z_f in the domestic and foreign currencies respectively. We can also have a short rate or *diff swap* where payments are swapped over $[0, T]$ on a domestic principal, with the floating rates based on the short rates in each country. A *LIBOR currency swap* is a swap of interest rate payments on two notional principals where the interest rates are based on the LIBOR for each country. The swap period $[0, T]$ is divided into N periods and payments are denoted by p_j. Now we describe precisely the deal that we are going to value which differs from that of [7].

Fixed-for-fixed cross-currency swap with a Bermudan option to cancel

The cross-currency swap tenor is divided into N_{cpn} *coupon periods.* The start and end dates for these periods are given by $T_0, \ldots, T_{N_{cpn}}$ and cashflows are exchanged at coupon period end dates $T_1, \ldots, T_{N_{cpn}}$. Typically, the swap cashflows consist of coupon payments at annualized rates on notional amounts Z_d for the first currency Z_f for the second currency. In addition, notional amounts Z_d and Z_f may be exchanged at the swap start and/or end dates. The size of a coupon payment is given by: *coupon rate* $\times$ *notional amount* $\times$ *coupon period day count.* Both interest rates R_f and R_d are *fixed* at the outset of the contract, as opposed to those for a LIBOR swap where they are floating [7]. There is no path-dependence in the payoffs, i.e. the path taken is not relevant because the payoff is fully determined by component values at the payment date. Payments p_j are made at at the end of each period at time t_j^- of size

$$p_j = \delta_j \left(S(t_j^-) R_f Z_f - m - Z_d R_d \right),$$

where m is the *margin* to the issuing counterparty. This is the terminal condition for the period $[t_{j-1}, t_j)$. The *value* of the deal is the sum of the present values of all payments.

When the contract has a *Bermudan option to cancel*, one of the counterparties is given an option to cancel all the future payments at times $t_1, \ldots, t_n$. Typically $t_1, \ldots, t_n$ are set a fixed number of calendar days before the start date of each period, i.e $t_1 = T_{N_{cpn}-1-n} - \Delta, \ldots, T_{N_{cpn}-1} - \Delta$, where Δ is the *notification period*. We assume that net principal amounts $(Z_d - S(0)Z_f))$ are paid at time 0 and at time t_N if the option is *not* cancelled, or at time t_{k+1} if the option is cancelled. The *terminal condition* at t_N^- is given by

$$V(t_N^-) = \delta_N \left(S(t_N^-) Z_f (1 + R_f) - m - Z_d (R_d + 1) \right). \tag{31}$$

When the option to cancel is exercised at t_k, we exchange coupon payments due on $T_{N_{cpn}-2-n+k}$ and notional amounts. The holder of the option will terminate the deal if the expected future value of the deal is less than the termination cost. Thus the decision at time $t_{k+1-\Delta}$ is to continue if

$$P_d(t_{k+1} - \Delta, t_{k+1}^-) V(t_{k+1}) < (S(t_{k+1} - \Delta) Z_f P_f(t_{k+1} - \Delta, t_{k+1}^-) - P_d(t_{k+1} - \Delta, t_{k+1}^-) Z_d).$$

This yields the boundary condition

$$V(t_{k+1} - \Delta) =$$

$$\min \left\{ P_d(t_{k+1} - \Delta, t_{k+1}^-) V(t_{k+1}), (S(t_{k+1} - \Delta) Z_f P_f(t_{k+1} - \Delta, t_{k+1}^-) - P_d(t_{k+1} - \Delta, t_{k+1}^-) Z_d) \right\} \tag{32}$$

This deal is valued by solving the PDE for the last period using the terminal condition (31) and stepping backwards in time using the termination condition

$$V(t_{k+1} - \Delta) =$$

$$\min \left\{ P_d(t_{k+1} - \Delta, t_{k+1}^-) V(t_{k+1}), (S(t_{k+1} - \Delta) Z_f P_f(t_{k+1} - \Delta, t_{k+1}^-) - P_d(t_{k+1} - \Delta, t_{k+1}^-) Z_d) \right\}$$

for earlier periods. We then add on the exchange of principals at time 0.

6 Numerical Results

The numerical results using a 1D and 3D implementation of the wavelet method of lines algorithm and the LSODE stiff vector ODE solver [13] are given below. In each case the numerical deal values are compared with a standard PDE solution technique and the known exact solution. Practical speed-up factors are reported which increase with both boundary condition discontinuities and spatial dimension.

European call option

Stock price: 10 *Strike price*: 10 *Interest rate*: 5% *Volatility*: 20% *Time to maturity*: 1 Year

The exact value of this option is: **1.04505**.

Comparing tables 1 and 2 shows a speedup of 1.9.

Table 1. Wavelet Method of Lines Solution

Space Steps	Time Steps	Value	Solution Time in Seconds
64	60	1.03515	.05
128	100	1.04220	.10
256	200	1.04502	.13
512	200	**1.04505**	**.30**
1024	200	1.04505	.90

Table 2. Crank-Nicolson Finite Difference Method

Space Steps	Time Steps	Value	Solution Time in Seconds
64	60	1.03184	.02
128	100	1.04184	.04
256	200	1.04426	.09
512	200	1.04486	.16
1024	200	1.04501	.30
2000	200	**1.04505**	**.57**

Cash-or-nothing call

The option with the same parameters as the European call has a payoff $B * \mathcal{H}(S - K)$, where $B := 3$ is the cash given, with a single discontinuity.

The exact value of this option is: **1.59297.**

Comparing tables 3 and 4 shows a speed up of 2.5.

Table 3. Wavelet Method of Lines Solution

Space Steps	Time Steps	Value	Solution Time in Seconds
128	100	1.49683	.10
256	200	1.54904	.13
512	200	1.59216	.30
1024	400	**1.59288**	**1.02**

Table 4. Crank-Nicolson Finite Difference Scheme

Space Steps	Time Steps	Value	Solution Time in Seconds
128	200	1.46296	.04
256	400	1.53061	.10
512	400	1.56391	.18
1024	400	1.58046	.31
2048	800	1.58872	1.35
4096	800	**1.59285**	**2.56**

Supershare call

Stock price: 10 *Strike price*: 10 *Parameter* d: 3 *Interest rate*: 5% *Volatility*: 20%

Time to maturity: 1 Year

The option pays an amount $1/d$ if the stock price lies between K and $K+D$ i.e. the option has a payoff $1/d * (\mathcal{H}(S-K) - \mathcal{H}(S-K-D))$ with two discontinuities.

The exact value of this option is: **0.13855.**

Comparing tables 5 and 6 shows a speed up of 4.9.

Table 5. Wavelet Method of Lines Solution

Space Steps	Time Steps	Value	Solution Time in Seconds
128	100	0.12796	.10
256	200	0.13310	.14
512	200	0.13808	.30
1024	400	**0.13848**	**1.04**

Table 6. Crank-Nicolson Finite Difference Scheme

Space Steps	Time Steps	Value	Solution Time in Seconds
128	200	0.12369	.04
256	400	0.13290	.09
512	400	0.13435	.16
1024	400	0.13666	.34
2048	800	0.13787	1.35
4096	800	0.13800	2.56
8000	800	**0.13835**	**5.11**

Cross Currency Swap

Domestic fixed rate: 10%, *Foreign fixed rate:* 10%

The exact value of this option is: **0.0**

Comparing tables 7 and 8 shows a speed up exceeding 81.

Table 7. Wavelet Method of Lines Solution

Discretization	Value	Solution Time in Seconds
20 X 8 X 8 X 8	-0.00082	1.2
20 X 16 X 16 X 16	-0.00052	6.54
20 X 32 X 32 X 32	-0.00047	40.40
40 X 64 X 64 X 64	**-0.00034**	**410.10**
100 X 128 X 128 X 128	-0.00028	4240.30
160 X 256 X 256 X 256	-0.00025	53348.10

Table 8. Explicit Finite Difference Scheme

Discretization	Value	Solution Time in Seconds
20 X 8 X 8 X 8	-0.00109	0.28
20 X 16 X 16 X 16	-0.00101	1.70
20 X 32 X 32 X 32	-0.00074	16.82
40 X 64 X 64 X 64	-0.00058	188.10
100 X 128 X 128 X 128	-0.00046	2421.6
160 X 256 X 256 X 256	**-0.00038**	**33341.8**

7 Conclusions and Future Directions

The wavelet method of lines performs well on problems with one spatial dimension and discontinuities or spikes in the payoff. For example in the supershare option the wavelet method requires a lower discretization than the Crank Nicolson finite difference scheme for equivalent accuracy (3 decimal places), as the discontinuities in the payoff can be resolved better in wavelet space. We also see that for the (prototype) cross currency swap PDE in 3 spatial dimensions the wavelet method outperforms the (tuned) explicit finite difference scheme by approximately two orders of magnitude – a very promising result. One of the important things to note is that $O(N)$ wavelet based PDE methods generalize $O(N \log N)$ spectral methods without their drawbacks. This lower basic complexity feature of the wavelet PDE method makes it suitable to solve higher dimensional PDEs. Further, to improve basic efficiency of the method we are currently implementing an adaptive wavelet technique in which the wavelet coefficients are thresholded at each time step (*cf.* §3). This should result in an improvement in both speed and memory usage because of sparse wavelet representation. Such a technique has resulted in a further order magnitude speedup in other applications [15]. Future work will thus involve applying the wavelet technique to solve cross currency swap problems with two and three factor interest rate models for each currency to result in solving respectively 5 and 7 spatial dimension parabolic PDEs.

References

[1] BEYLKIN, G. (1993). Wavelets and Fast Numerical Algorithms. Lecture Notes for Short Course, AMS 93.

[2] BEYLKIN, G., R. COIFMAN AND V. ROKHLIN (1991). Fast Wavelet Transforms and Numerical Algorithms. *Comm. Pure and Appl. Math* **44** 141–183.

[3] COHEN, A., I. DAUBECHIES AND J. FEAUVEAU (1992). Biorthogonal Bases of Compactly Supported Wavelets. *Comm. Pure Appl Math* **45** 485–560.

[4] COHEN, A., S. KABER, S. MULLER AND M. POSTEL (2000). Accurate adaptive multiresolution scheme for scalar conservation laws. Preprint, LAN University Paris.

[5] DAHMEN, W., S. MULLER AND T. SCHLINKMANN (1999). On a Robust Adaptive Multigrid Solver for Convection-Dominated Problems. Technical report, RWTH Aachen. IGPM Report No 171.

[6] DAUBECHIES, I. (1992). *Ten Lectures on Wavelets.* SIAM, Philadelphia.

[7] DEMPSTER, M. A. H. AND J. P. HUTTON (1997). Numerical Valuation of Cross-Currency Swaps and Swaptions. In *Mathematics of Derivative Securities*, eds. M. A. H. Dempster and S. R. Pliska. Cambridge University Press, 473 – 503.

[8] POLIKAR, R. The Wavelet Tutorial. http://www.public.iastate.edu/ rpolikar/wavelet.html.

[9] PROSSER, R. AND R. CANT (1998). Evaluation of Nonlinear Terms Using Interpolating Wavelets. Working paper, CFD laboratory, Department of Engineering, University of Cambridge.

[10] PROSSER, R. AND R. CANT (1998). On the Representation of Derivatives Using Interpolating Wavelets. Working paper, CFD laboratory, Department of Engineering, University of Cambridge.

[11] PROSSER, R. AND R. CANT (1998). On the Use of Wavelets in Computational Combustion. Working paper, CFD laboratory, Department of Engineering, University of Cambridge.

[12] PROSSER, R. AND R. CANT (1998). A Wavelet-Based Method for the Efficient Simulation of Combustion. *J. Comp. Phys* **147** (2) 337–361.

[13] SCHIESSER, W. (1991). *The Numerical Method of Lines.* Academic Press, Inc.

[14] SWELDENS, W. (1996). Building Your Own Wavelets at Home. ACM SIGGRAPH Course Notes.

[15] VASILYEV, O., D. YUEN AND S. PAOLUCCI (1996). A Fast Adaptive Wavelet Collocation Algorithm for Multi-Dimensional PDEs. *J. Comp. Phys* **125** 498–512.

[16] VASILYEV, O., D. YUEN AND S. PAOLUCCI (1997). Wavelets: An Alternative Approach to Solving PDEs. Supercomputer Institute Research Report, University of Minnesota.

[17] WILMOTT, P., S. HOWISON AND J. DEWYNNE (1995). *The Mathematics of Financial Derivatives.* Cambridge University Press.

[18] XU, J. AND W. SHANN (1992). Galerkin-Wavelet Methods for Two-Point Boundary Value Problems. *Numerische Mathematik* **63** 123–144.