Adaptive High-Order Finite-Difference Method for Nonlinear Wave Problems

I. Fatkullin
Department of Mathematical Sciences
Rensselaer Polytechnic Institute
Troy, NY 12180-3590, USA
email: FATKUI@RPI.EDU

and

J. S. Hesthaven
Division of Applied Mathematics
Brown University
Providence, RI 02912, USA
email: JAN.HESTHAVEN@BROWN.EDU

Abstract

We discuss a scheme for the numerical solution of one-dimensional initial value problems exhibiting strongly localized solutions or finite-time singularities. To accurately and efficiently model such phenomena we present a full space-time adaptive scheme, based on a variable order spatial finite-difference scheme and a 4th order temporal integration with adaptively chosen time step. A wavelet analysis is utilized at regular intervals to adaptively select the order and the grid in accordance with the local behavior of the solution. Through several examples, taken from gasdynamics and nonlinear optics, we illustrate the performance of the scheme, the use of which results in several orders of magnitude reduction in the required degrees of freedom to solve a problem to a particular fidelity.

1 Introduction

In a variety of physical systems, modeled by nonlinear wave equations, one often observes the development of highly localized solutions such as shocks and even finite-time singularities, e.g., numerous examples of the former can be found in gasdynamics (Whitham (1974)) while examples of the latter are found in nonlinear optics (Newell and Maloney (1992)).

The efficient and robust numerical modeling of such phenomena is made difficult by the high degree of localization which introduces severe stiffness and rules out the use of classical split-step FFT-based methods (Newell and Maloney (1992)) due to the inefficiency associated with the representation of a highly localized solution by a global expansion. The most natural

way to solve such problems would seem to involve the use of a highly adaptive solution technique, allowing for the use of sufficient resolution in regions where needed while using a sparse mesh in regions with little variation of the solution. Furthermore, as advocated Jameson (1998) in general and Hesthaven and Jameson (1998) in a multi-domain context, it seems advantageous to use a high-order accurate scheme on a coarse grid in regions of large regularity while a low-order scheme in combination with a very fine grid is preferable close to regions of very limited regularity. This is very much in the spirit of *hp*-finite element methods, see e.g. Babuška and Suri (1987) or Schwab (1998), used extensively in solid and fluid mechanics for solving elliptic problems.

Here we shall discuss the development of a scheme with a similar degree of flexibility and accuracy but tailored specifically to the solution of nonlinear wave problems exhibiting highly localized dynamics. We shall show that the use of such schemes has the potential to very significantly reduce the computational resources needed to solve a given problem to a specific accuracy. For simplicity we restrict our problems to be periodic in space, although there is nothing that prohibits the introduction of boundaries by one-sided stencils or even a multidomain formulation as discussed in detail by Hesthaven and Jameson (1998) for linear problems. The approach is entirely general and allows for solving for real and complex solutions and scalar problems as well as systems of equations.

The scheme is based on the approximation of the spatial derivatives by an explicit high-order finite-difference method of variable order while a 4th order Runge-Kutta method is employed in time. The central issue naturally becomes how to choose, in space-time, a suitable level of adaptation as the solution evolves. As we aim at a scheme applicable to a variety of problems we shall need an error estimator that is independent of the particular problem being considered, relying solely on the fast and accurate localization of regions of low regularity.

With this in mind it becomes natural to turn the attention to wavelets as tools for the localization of regions with high-frequency content, i.e., regions with steep gradients. Indeed, one could argue that a full wavelet based scheme would be best suited for the problems considered here and developments along such lines have indeed been attempted in the past, either in a Galerkin formulation (Bacry et al. (1992)) or using interpolating wavelets in a collocation formulation (Cai and Wang (1996), Vasilyev and Paolucci (1996), Beylkin and Keiser (1997)). The application to multidimensional linear problems is discussed by Cai and Zhang (1998) in the context of reaction-diffusion problems and by Katehi et al. (1998) for the solution of problems in electromagnetics.

In these previous works the solution itself is represented in terms of the wavelet expansion and all operations, e.g., differentiation and integration,

are performed directly on this basis. Here we take a different route, inspired by some deep relations, realized independently by Beylkin (1992), and Jameson (1993, 1996), between classical finite-difference stencils and differentiation matrices based on wavelets. Indeed, on the finest level of refinement these discrete operators turn out to be equivalent or almost equivalent depending on the wavelet family being considered. This suggests, as has also been exploited by Jameson (1998), and Hesthaven and Jameson (1998) for solving linear problems, that one can solve the initial value problem using a variable order classical finite difference scheme, but use a wavelet analysis at periodic intervals to facilitate an adaptive spatial refinement of the grid and, subsequently, in time to ensure stability. This is exactly the approach we shall take in the following.

What remains of the paper is organized as follows. In Sec. 2 we discuss the details of the wavelet optimized high-order finite-difference method with an emphasis on the adaptive construction of the finite-difference stencil and the associated wavelet analysis that allows for correctly adapting the local resolution and the order of the scheme. Guidelines for wavelet truncations and the balance between the order and the resolution is also discussed. This sets the stage for Sec. 3 where we present a number of applications to the solution of problems derived primarily from the area of nonlinear fiber optics. Section 4 contains a few concluding remarks and guidelines for future work.

2 A Wavelet Optimized Finite-Difference Method

We consider the general well posed initial value problem

$$\begin{aligned} \frac{\partial \boldsymbol{u}}{\partial t} &= \mathcal{F}(\boldsymbol{u}, t, \boldsymbol{x}) \ , \\ \boldsymbol{u}(\boldsymbol{x}, t) &= \boldsymbol{u}_0(\boldsymbol{x}) \ , \end{aligned} \tag{1}$$

where $\boldsymbol{x} \in \mathsf{D}$ refers to the spatial domain and $\boldsymbol{u}(\boldsymbol{x}, t) : \mathsf{D} \times \mathsf{R}^+ \rightarrow \mathsf{C}^n$ represents the n-dimensional state vector defined on the complex field, C. The general nonlinear flux, $\mathcal{F}(\boldsymbol{u}, t, \boldsymbol{x}) : \mathsf{C}^n \times \mathsf{D} \times \mathsf{R}^+ \rightarrow \mathsf{C}^n$ may depend on space-time as well as $\boldsymbol{u}$ and operations on $\boldsymbol{u}$, e.g., differentiation and integration.

For simplicity we restrict ourselves to a detailed discussion of the one-dimensional case, furthermore assuming that the solutions are of a periodic nature. There is, however, nothing intrinsic in the scheme that utilizes this assumption and it can be relaxed to include non-periodic initial boundary value problems and multi-dimensional problems.

In formulating the scheme for solving Eq.(1) we shall, as is the basis of finite difference methods, assume that the solution, $\boldsymbol{u}$, is well represented

locally by a interpolation polynomial and utilize this to approximate derivatives etc. By requiring that the equation be satisfied in a collocation way at the grid-points supporting the local polynomials we recover a method-of-lines formulation with a large system of ordinary differential equations to be solved for advancing in time.

In the following we shall discuss in some detail the construction of the high-order finite-difference stencils on arbitrary grids and, subsequently, how this interacts with a wavelet analysis of the solution to enable order and grid adaptivity.

2.1 Polynomial Interpolation and General Finite-Difference Stencils

A central element of the scheme is the ability to construct the arbitrary order finite-difference approximations to the spatial derivatives. We recall that, due to the fully adaptive nature of the scheme, the order of the approximation as well as the grid structure is arbitrary.

Let us hence briefly review this construction, referring the reader to Fornberg (1998) for details. The classic approach is to construct Lagrange interpolation polynomials of the specified order and, subsequently, evaluate the derivative of these polynomials at the grid points to obtain the coefficients in the finite-difference stencil.

We assume that the function, $u_i = u(x_i)$, is given on an arbitrary N-element grid x_i and construct the N'th order Lagrange interpolation polynomials, $l_{i,N}(x)$. This allows us to recover the local N'th order interpolation polynomial, $p_N(x)$, on the form

$$p_N(x) = \sum_{i=0}^{N} u_i l_{i,N}(x) \quad , \tag{2}$$

where

$$l_{i,N}(x) = \frac{(x-x_0)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_N)}{(x_i-x_0)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_N)} \quad . \tag{3}$$

The k'th order derivative of the interpolation polynomial, p_N, will thus serve as the N'th order approximation to the k'th derivative of $u(x)$, i.e.,

$$\frac{\mathrm{d}^k u(x)}{\mathrm{d}x^k} \approx \frac{\mathrm{d}^k p_N(x)}{\mathrm{d}x^k} = \sum_{i=0}^{N} u_i \frac{\mathrm{d}^k l_{i,N}(x)}{\mathrm{d}x^k} = \sum_{i=0}^{N} u_i c^k_{i,N}(x) \quad . \tag{4}$$

To obtain the k'th derivative of u at x_α we thus need to evaluate $c^k_{i,N}(x_\alpha)$. To recover these, consider

$$l_{i,N}(x) = \sum_{k=0}^{N} \left. \frac{\mathrm{d}^k l_{i,N}(x)}{\mathrm{d}x^k} \right|_{x_\alpha} \frac{(x - x_\alpha)^k}{k!} = \sum_{k=0}^{N} c_{i,N}^k(x_\alpha) \frac{(x - x_\alpha)^k}{k!}, \qquad (5)$$

i.e., they are nothing but the Taylor expansion coefficients of $l_{i,N}(x)$ around x_α.

Directly from Eq.(3) we can recover the relations

$$i < N \;\; : \;\; l_{i,N}(x) = \frac{x - x_N}{x_i - x_N} l_{i,N-1}(x) \;\; , \qquad (6)$$

and

$$i = N \;\; : \;\; l_{N,N}(x) = \frac{\prod_{i=0}^{N-2} (x_{N-1} - x_i)}{\prod_{i=0}^{N-1} (x_N - x_i)} (x - x_{N-1})\, l_{N-1,N-1}(x) \;\; . \qquad (7)$$

Combining these with the Taylor series, Eq.(5), yields the recurrence relations (Fornberg (1998)) ($k \leq N$)

$$i < N \;\; : \;\; c_{i,N}^k(x_\alpha) = \frac{1}{x_i - x_N} \left[(x_\alpha - x_N)\, c_{i,N-1}^k - k\, c_{i,N-1}^{k-1} \right] \;\; , \qquad (8a)$$

$$c_{N,N}^k(x_\alpha) = \frac{\prod_{i=0}^{N-2} (x_{N-1} - x_i)}{\prod_{i=0}^{N-1} (x_N - x_i)} \left[k\, c_{N-1\;,N-1}^{k-1} + (x_\alpha - x_{N-1})\, c_{N-1\;,N-1}^{k} \right] \;\; . \qquad (8b)$$

Using as starting point that $c_{0,0}^0 = 1$ due to the nature of Lagrange polynomial and taking all undefined coefficients to be zero, one now computes all necessary weights $c_{i,N}^k$ in $\mathcal{O}(kN^2)$ operations and, hence, evaluates the derivative up to k'th order at some point x_α. Note that taking $k = 0$ simply yields the scheme for interpolation.

One should realize that the above construction depends only on the grid points, x_i, i.e., only when the grid is changed due to adaptation shall we need to recompute the weights. As this is done at periodic intervals only and not at every time step, the computation of the weights and, thus, the finite difference stencils, represents a negligible computational expense.

2.2 Mesh and Order Adaptation

Having laid the foundation for the computations of arbitrary order stencils on arbitrary grids, let us now turn the attention to the use of wavelets and wavelet expansions to determine where to refine/coarsen the grid and/or adjust the order of the scheme.

Prior to that, however, we shall first recall a few basic properties of wavelets and wavelet expansions of relevance to the present discussion.

2.2.1 Briefly on Wavelets and Wavelet Expansions

The term wavelet is used to describe a spatially localized function, i.e., the wavelet is assumed to have most of its energy contained in a very narrow region of the physical space. We shall restrict ourselves to wavelets with compact support and focus on the family defined by Daubechies (1988).

To define the Daubechies wavelets, consider the two functions, $\phi(x)$ and $\psi(x)$, as solutions to the equations

$$\phi(x) = \sqrt{2}\sum_{k=0}^{L-1} h_k \phi(2x-k), \tag{9}$$

$$\psi(x) = \sqrt{2}\sum_{k=0}^{L-1} g_k \phi(2x-k), \tag{10}$$

with $\phi(x)$ normalized as

$$\int_{-\infty}^{\infty} \phi(x)dx = 1 \ .$$

Let also

$$\phi_k^j(x) = 2^{-\frac{j}{2}}\phi(2^{-j}x-k) \ , \ \psi_k^j(x) = 2^{-\frac{j}{2}}\psi(2^{-j}x-k) \ ,$$

where $j,k \in \mathsf{Z}$, denote the dilations and translations of the scaling function, $\phi_k^j(x)$, and the wavelet, $\psi_k^j(x)$, respectively.

The sets, $H = \{h_k\}_{k=0}^{L-1}$ and $G = \{g_k\}_{k=0}^{L-1}$, are related as $g_k = (-1)^k h_{L-1-k}$ for $k = 0, ..., L-1$. Furthermore, H and G are chosen so that dilations and translations of the wavelet, $\psi_k^j(x)$, form an orthonormal basis on $L^2(\mathsf{R})$ and such that the mother wavelet, $\psi(x)$, has $M = L/2$ vanishing moments. In other words, $\psi_k^j(x)$, satisfies

$$\delta_{kl}\delta_{jm} = \int_{-\infty}^{\infty} \psi_k^j(x)\psi_l^m(x)dx \ , \tag{11}$$

where δ_{kl} is the Kronecker delta function, and the mother wavelet, $\psi(x) = \psi_0^0(x)$, is defined by

$$\forall m \in [0, \ldots, M-1] : \int_{-\infty}^{\infty} \psi(x) x^m dx = 0 \ . \tag{12}$$

It is usual to let the spaces spanned by $\phi_k^j(x)$ and $\psi_k^j(x)$ over the parameter k, with j fixed, be denoted by V_j and W_j, i.e.

$$V_j = \text{span}\left\{\phi_k^j(x)\right\}_{k\in Z} \ , \ W_j = \text{span}\left\{\psi_k^j(x)\right\}_{k\in Z} \ .$$

These spaces, V_j and W_j, are related as

$$... \subset V_1 \subset V_0 \subset V_{-1} \subset ...,$$

and

$$V_j = V_{j+1} \oplus W_{j+1} \ ,$$

i.e., W_{j+1} is the orthogonal complement of V_{j+1} in V_j. Utilizing orthonormality of the wavelets, ψ_k^j, one recovers

$$L^2(\mathsf{R}) = \bigoplus_{j\in Z} W_j \ , \tag{13}$$

reflecting the completeness of the wavelet basis. Hence, any $u(x) \in L^2(\mathsf{R})$ can be written as,

$$u(x) = \sum_{j\in Z} \sum_{k\in Z} d_k^j \psi_k^j(x) \ . \tag{14}$$

The expansion coefficients, d_k^j, appear directly through orthogonality as

$$d_k^j = \int_{-\infty}^{\infty} u(x) \psi_k^j(x) dx \ , \tag{15}$$

where the decay of d_k^j depends on the local regularity of $u(x)$ as

$$|d_k^j| \leq C 2^{-j\frac{L+1}{2}} \max_{\xi\in[k2^{-j},(k+M-1)2^{-j}]} \left|u^{(M)}(\xi)\right| \ . \tag{16}$$

From Eq.(16) we find that if $u(x)$ behaves like a polynomial of order less than M inside the small interval, then d_k^j vanishes exactly, i.e., one can view the magnitude of d_k^j as a direct measure of how well the assumption of local polynomial behavior, underlying the finite difference scheme, is satisfied. This makes the use of wavelets as local error estimators in connection with finite-difference methods a natural and efficient approach.

Indeed, even if $u^{(M)}$ differs from zero, it will nevertheless decay exponentially with the scale parameter, j, and the information given by Eq.(16)

provides a local measure of the regularity of the function or rather the closeness to local polynomial behavior. This is exactly what we utilize to determine the need for order and/or mesh adjustments.

Naturally, infinite sums and integrals are meaningless when one begins to implement a wavelet expansion on a computer. However, j is related directly to the different scales, i.e., level of meshes in this context, with $j = 0$ reflecting the finest possible grid. Furthermore, as k reflects translation, these parameters are set naturally by the problems considered and levels of refinement allowed.

Hence, to obtain the wavelet expansion of a function, $u(x)$, suppose that it is sampled at a uniform mesh at level j. We can then recover the scaling function and wavelet coefficients on the next coarser scale as

$$u_k^{j+1} = \sum_{n=1}^{2M} h_n u_{n+2k-M}^j \quad ,$$

$$d_k^{j+1} = \sum_{n=1}^{2M} g_n u_{n+2k-M}^j \quad ,$$

where $M = L/2$, and H and G basically reflects low- and high-pass filtering of the function. We can continue this decomposition to recover the wavelet coefficients on a number of scales, $\boldsymbol{d}^0, \boldsymbol{d}^1, \ldots$, and obtain information about the local regularity of the function using Eq.(16). This provides a basis for deciding whether the solution is sufficiently well resolved. We note in particular that this error estimator is independent of the particular problem being consider and, hence, is applicable to the adaptive solution of a variety of nonlinear wave problems.

The number of vanishing moments, M, of the wavelet, $\psi(x)$, is related to the order of the wavelet expansion. For Daubechies wavelets, D_L, the number of elements in H and G, or the length of the filters H and G, denoted by L, is related to the number of vanishing moments M by $2M = L$. Moreover, L also reflects the support of the wavelet, i.e. small L implies narrow local support.

In this work we shall only consider the use of the D_4 wavelet with the filter coefficients given in Table 4. In Jameson (1998), Hesthaven and Jameson (1998) this was found to provide a good alternative between compactness and regularity. It is worth emphasizing, though, that numerous other choices may well yield similar or better results and some experimentation and analysis may well be warranted to establish a suitable filter for a particular application.

The filter coefficients, H, needed to define compactly supported wavelets with a higher degree of regularity can be found in Daubechies (1988). As expected, the support increases with the regularity of the wavelet.

2.2.2 Wavelet Optimized Mesh Generation

With the ability to localize regions with little regularity, e.g., $|d_k^j| > \varepsilon_R^j$ where ε_R^j is some threshold set for refinement at level j, as well as with a high degree of regularity, e.g., $|d_k^j| < \varepsilon_C^j$ with ε_C^j being some threshold set for coarsening at level j, we are able to optimize the mesh distribution as guided by the wavelet decomposition.

We shall adjust the grid by bisection, i.e., by dividing existing intervals by half or removing central mesh points. As illustrated in Fig. 4, we associate a special index with each point in the grid, reflecting the level of the refinement required to activate the point in the specified position. Hence, the domain boundaries have index zero, the one in the center of the interval corresponds has index one, those nodes specifying 1/4 and 3/4 of the interval have index two and so on. Clearly this index is directly related to the local resolution level, i.e., $j = 0$ corresponds to all indices 0-4 while $j = 4$ contains only index zero mesh points in Fig. 4.

The procedure for determining the wavelet optimized grid, given a particular solution, at a generally nonuniform grid, involves interpolation to recover a solution at a uniform grid, corresponding to a level of refinement one level finer than is currently considered unless the finest level is already in use.

Decomposing the solution, $u(x)$, into the wavelet expansion now gives a direct measure of where the regions of smoothness as well as high gradients are located. To decide whether to refine or coarsen the grid locally we shall use simple thresholding of the wavelet coefficients inspired by the simple error estimate in Eq.(16), i.e., we use the following threshold values

$$\varepsilon_R^j = C_R 2^{-\frac{5}{2}j} \ , \ \varepsilon_C^j = C_C 2^{-\frac{5}{2}j} \ ,$$

and refine, or rather retain the level of resolution, if $|d_k^j| > \varepsilon_R^j$ while we coarsen if $|d_k^j| < \varepsilon_C^j$, corresponding to dropping that particular level of resolution locally and consider the wavelet coefficients for level $j + 1$.

In Fig. 4 we illustrate this procedure to recover wavelet optimized meshes for functions with different degrees of regularity. We take $2^{5/2} C_R < C_C = 6C_R$ for simplicity although this is by no means unique. The functions are initially sampled at 1024 points on [0..1] and then analyzed by the wavelet analysis outlined in the above using different values of the threshold, C_R. The grid points are clearly clustered in regions with a high degree of variation as one would intuitively prefer.

It is worth pointing our that a number of additional strategies may be worth considering in the development of a robust and effective mesh generation scheme, e.g., it would seem natural to avoid too large jumps in the local resolution by requiring that grid-points around areas with high

wavelet coefficients be included to ensure that the solution remains well resolved even after a period of evolution. The degree of this, however, is clearly related to how often the grid is adapted and some compromise between work and accuracy is required, depending on the requirements of the particular problem being considered.

2.2.3 Order Adjustment

A careful inspection of the approximation to the discontinuous function in Fig. 4 reveals that two points do not fall exactly on the function. This is a result of using a 4'th order polynomial to interpolate a discontinuous function, resulting in artificial oscillations much as the well known Gibbs phenomenon for Fourier series and expansions utilizing orthogonal polynomials (Gottlieb and Shu (1997)). While one can choose to ignore these oscillations for the approximation of functions, they may render the numerical scheme unstable due to nonlinear mixing of these high-frequency components with the solution as is well known when solving nonlinear problems using spectral methods, see e.g. Gottlieb and Hesthaven (2001).

There are essentially two ways to deal with these concerns for general nonlinear problems. One can choose to compute with the oscillations and stabilize the numerical scheme by introducing additional dissipation by the use of a filter (Gottlieb and Hesthaven (2001)). For spectral methods, where the ability to change the order locally is very limited, this is the method of choice and one could attempt a similar technique within the present context. However, as we also have the ability to locally modify the order of the finite-difference scheme, using the magnitude of the wavelet coefficients as guides, we can simply choose to lower the order of the scheme locally to a 1st order scheme if the wavelet expansion indicates an inability to resolve the solution even at the finest level of mesh refinement. Such an approach allows us to take maximum advantage of the flexibility of the scheme presented here by using a low order nonoscillatory scheme in regions with very low regularity or severely underresolved solutions while we can use a high-order scheme in regions with a high degree of smoothness.

For any interval we introduce a number, n_j, containing the maximum index of the points limiting this interval, i.e., it relates directly to level of local refinement at the j'th interval. Similarly for all grid point we introduce m_j as the maximum index of the point itself and the two nearest grid points.

As the wavelet optimized mesh analysis will refine the mesh in regions of less regularity, we can utilize n_j and m_j to control the order of the spatial interpolation. Indeed, we use m_j to control the order the finite differencing and n_j to control the order of the interpolation in the preparation for the wavelet mesh analysis.

Choosing proper values for switching between different orders in response to variations in n_j and m_j is a matter of some experimentation to obtain a suitable compromise between local smoothness and computational efficiency. In Table 4 we list the set of parameters used in the following for the solution of nonlinear wave problems, assuming that we allow for at most 18 levels of binary refinements, i.e., a total of at most 2^{18} intervals.

3 Numerical Examples

To verify the versatility and robustness of the scheme discussed in the above let us consider the solution of a few typical nonlinear wave problems, primarily inspired by problems in nonlinear fiber optics.

The semi-discrete finite-difference approximation is discretized in time using a 4th order fully explicit Runge-Kutta scheme with the time-step chosen adaptively by stability considerations. As the time-step is directly proportional to the smallest mesh-size, we find this to be sufficient to ensure that the error is controlled by spatial rather than temporal errors. A cautionary note is appropriate here as we are dealing with variable size, variable order finite difference schemes on general nonuniform grids, i.e., it is not unlikely that some of these operators are indeed inherently unstable. With a fixed grid and order of scheme such methods would naturally be unstable. However, the adaptive process may well act as a stabilizing averaging process that controls the few unstable eigenmodes through the inherent randomness of the process. Similar experiences have been discussed by Jameson (1998).

As a first example, let us demonstrate soliton collision in the nonlinear Schrödinger (NLS) equation,

$$\mathrm{i}\frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial x^2} + 2u|u|^2 = 0 \ , \tag{17}$$

being used as a standard model for describing pulse-propagation in weakly nonlinear fibers, see e.g. Newell and Moloney (1992). Here $u(x,t)$: $\mathsf{D}\times\mathsf{R}^+ \to \mathsf{C}$ signifies the envelope of the optical field. The NLS is fully integrable and possesses an infinite number of conserved quantities which can be used to test the accuracy of the computational framework. Throughout all computations we evaluate the conserved mass

$$N = \int |u|^2 \, \mathrm{d}x \ , \tag{18}$$

and momentum

$$M = \int (u\bar{u}_x - \bar{u}u_x) \, \mathrm{d}x \ , \tag{19}$$

and find them to be conserved to within 10^{-7}.

The initial conditions consist of two counter propagating NLS-solitons of the general form (Whitham (1974))

$$u(x,0) = \alpha \,\mathrm{sech}[\alpha x] \exp(i\beta x) \ ,$$

where (α, β) are constants related to the amplitude and direction of propagation of the soliton. Note in particular that the soliton becomes more compact with increasing amplitude α.

In Fig. 4 we illustrate the well known soliton collision scenario, showing perfect reconstruction of the solitons after the strongly nonlinear collision. The computation is performed with an average of 300 grid-points allowing up to 14 levels of refinement with the parameter for thresholding set to $C_R = 0.001$ in the wavelet optimized mesh generation scheme. We shall use this value in all subsequent computations.

As a second example, we consider the solution of the viscid Burgers equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2} \ , \tag{20}$$

with $u(x,t) : \mathsf{D} \times \mathsf{R}^+ \to \mathsf{R}$, as simple model to describe a variety of problems in gasdynamics, see e.g. Whitham (1974). For ν small but finite the equation can develop local regions with very steep gradients although the solution remains smooth.

We solve Eq.(20) on $\mathsf{D} = [0, 2\pi]$ with $\nu = 0.001$ and the initial conditions

$$u(x,0) = 1 + \cos(x) \ ,$$

which is expected to result in an initially steepening front and the development of a shock-like solution which propagates towards the right while decaying in amplitude.

To solve the problem we allow up to 14 levels in the mesh adaptation scheme, with an initial distribution of about 60 and about 200 grid points after the very steep gradient is formed.

In Fig. 4 we illustrate the computed solution showing the expected dynamics. The lack of oscillations of the solution around the very steep gradient is noteworthy and is a result of the combined mesh and order adaptation embedded in the present framework. The minimal spatial interval used throughout the calculation is $2\pi \cdot 2^{-14}$, i. e., one would need in the excess of 10000 grid points to recover a solution of similar quality using a uniform mesh.

To further understand the adaptive process and its solution dependent nature we show in Fig. 4 the temporal development of the order selection as well as the total number of active grid points as the solution evolves. As

expected, the scheme begins with a high order scheme on a few points and, as the solution steepens, enters a transient phase where a significant fraction of the solution is evolved using a lower order scheme at a finer grid. As expected this trend continues until the point of maximum steepness around $t = \pi/2$ after which the scheme slowly restores its high order nature and decreases the total number of points. This behavior is exactly as one would expect by comparing with the solution in Fig. 4.

As a third and more challenging problem, we return to the area of nonlinear optics and consider the solution of the extended derivative nonlinear Schrödinger equation (EDNLS), see e.g. Hesthaven et al. (1997),

$$\frac{\partial u}{\partial t} + \mathrm{i}\frac{\partial^2 u}{\partial x^2} + \alpha|u|^2\frac{\partial u}{\partial x} + \mathrm{i}|u|^4 u = 0 \ , \tag{21}$$

where $u(x,t) : \mathsf{D} \times \mathsf{R}^+ \to \mathsf{C}$ as for the NLS equation. Compared to Eq.(17), the EDNLS includes high-order effects accounting for nonlinear dispersion and higher order nonlinear Kerr effects.

Let us first consider the case of $\alpha = 0$, recognized as the critical nonlinear Schrödinger equation (CNLS), which admits finite-time collapsing solutions provided the initial condition has a mass, N, exceeding that of the soliton solution on the form $u(x,t) = \phi(x)\exp(-\mathrm{i}\lambda t)$ where $\lambda > 0$ and $|\phi(x)| = \sqrt{3\lambda}\,\mathrm{sech}[2\sqrt{\lambda}x]$ (Hesthaven et al. (1997)). For the soliton we have $N_0 \simeq 2.72...$

To accurately capture the finite time collapse we shall allow for 18 levels of adaptive refinement and order adaptation. The CNLS conserves mass and momentum similar to the NLS and we find them to be conserved to 10^{-6} in all computations until shortly before the late stage of the collapse renders the computation unstable unless further refinement is allowed.

In Fig. 4 we show the computed solution with the slightly supercritical initial conditions. The collapse is evident as is the ability to capture the dynamics of the very rapidly evolving solution using space-time adaptation. A detailed study of the purely real equivalent of the CNLS was performed by Berger and Kohn (1988) where results very similar to the ones shown here are found.

In a way similar to the Burgers equation, we illustrate in Fig. 4 the temporal dynamics of the adaptive process when solving the CNLS equation subject to supercritical initial conditions. The scheme essentially remains predominately 4th order accurate until shortly before the collapse where an increasing fraction of work is done at very fine grid with a second order scheme. Comparing with the development of the solution shown in Fig. 4 this is in line with the expected behavior. It is worth while emphasizing that even though the total number of points increases and the associated fraction of points solved at lower order, the actual fraction of the computational solution advanced at low order remains very small.

Let us finally also consider the general EDNLS with $\alpha = 1$, the dynamics of which was studied in great detail by Hesthaven et al. (1997) where it was shown that for the mass of the initial condition, N, exceeding the CNLS soliton mass as $N > 3/\sqrt{21}N_0$, one should expect a collapsing solution. However, due to the drift term the solution is slightly asymmetric and, furthermore, marginally stable oscillatory solutions were found for very small super/sub-critical initial conditions.

We shall consider a supercritical initial condition consisting of a perturbed CNLS soliton. As for the CNLS we allow 18 levels of refinement to capture the collapse dynamics. The mass and momentum are conserved to within 10^{-6} throughout the calculation.

In Fig. 4 we show the computed solution, reproducing those in Hesthaven et al. (1997), while significantly extending the time for the stable computation due to the full mesh/order adaptation ability. Without adaptivity the modeling of the collapsing solutions would be impractical at best and perhaps even impossible given the very high degree of refinement required to resolve the late-time dynamics of the solution.

In a way similar to the previous cases, we illustrate in Fig. 4 the dynamics of adaptive process, confirming the expected behavior.

4 Concluding Remarks

The ability to modify the local resolution and/or the order of the approximation in response to the dynamical evolution of the solution is not only a highly desirable property but often also an enabling technology for solving problems exhibiting highly localized dynamics. Indeed, for dynamic problems of realistic complexity this is conceivably the only practical computational approach.

To adapt the computational scheme in response to the evolution of the solution naturally requires access to some measure of local regularity of the solution in the computational domain. While the development of error estimators is an area of active research it remains a considerable challenge to develop such estimators for even simple nonlinear time-dependent problems.

With this reasoning it is perhaps more practical to seek a less detailed, yet robust and general procedure for analyzing the solution without introducing knowledge of the particular problem being considered, i.e., the problem essentially becomes one of signal analysis as the basis on which to adapt the mesh and the order of the scheme.

In this work we have demonstrated how a wavelet analysis of the time-dependent solution can yield excellent measures of the local regularity of the solution in a simple manner and enable the required adaptation in a

robust and accurate way. As we have demonstrated through the solution of a number of problems, primarily originating in nonlinear optics, such a high-order finite difference scheme with a wavelet optimized grid yields solutions of high fidelity even for nontrivial problems exhibiting finite-time collapse.

While we have illustrated the feasibility of using a fully space-time adaptive approach to solve problems of nonlinear optics, a number of issues need careful continued attention to make this into a tool suitable for large scale modeling efforts ongoing in the optical communications industry.

On the theoretical side we are currently unable to provide any justification for the observed robustness and stability of the scheme due to its nonlinear nature. We speculate that a probabilistic approach is the correct approach but are unaware of such efforts at present time.

However, given that we find the scheme to be highly robust over a large variation of parameters and problems, it is perhaps more pressing to address some issues of a more practical character. For the spatial adaptation and approximation, we are currently recovering a solution at the finest level prior to the wavelet analysis phase. This can clearly be improved on by considering some very simple interpolation schemes as it is very unlikely that a particular solution suddenly requires refinements at several levels, i.e., one could use only partial information about the wavelet coefficients.

A more severe limitation, however, of the current approach is the use of a fully explicit time advancing scheme. For the solution of realistic problems on a highly adapted mesh this clearly becomes prohibitive and we need to seek alternatives of which one can think of several, e.g., one can split the operators and use a semi-implicit treatment with a fully implicit advancement of the linear parts of the operator. Even though the operators are variable bandwidth due to order adaptation they are evaluated fast and a matrix-free Krylov method may work well for solving the linear system resulting from the semi-implicit approach.

A more appealing approach is perhaps to consider explicit/implicit Runge-Kutta schemes, (Archer et al. (1997)), allowing for a operator splitting as above or for treating the regions of the grid with high stiffness implicitly by using a diagonally implicit Runge-Kutta scheme and the remaining by an explicit Runge-Kutta approach.

With such improvements we anticipate that the methods discussed here will be competitive with FFT based split-step schemes for the solution of a variety of nonlinear wave problems. In particular in the area of high-speed optical communication does one find an increasing need for efficient, robust, and accurate alternatives to the classical split-step methods and we hope to report on such direct comparisons in the future.

Acknowledgment

The first author (IF) acknowledges the partial support of NSF grant DMS-9502142, while the second author (JSH) acknowledges the partial support of NSF grant DMS-0073923, and by the Alfred P. Sloan Foundation as a Sloan Research Fellow.

References

ASCHER, U. M., RUUTH, S. J., AND SPITERI, R. J. (1997), *Implicit-Explicit Runge-Kutta Methods for Time-Dependent Partial Differential Equations*, Appl. Numer. Math. **25**, pp. 151-167.

BABUŠKA, I., AND SURI, M. (1987), *The hp-Version of the Finite Element Method with Quasiuniform Meshes*, M^2AN, **21**, pp. 199-238.

BACRY, E., MALLAT, S., AND PAPANICOLAOU, G. (1992), *A Wavelet Based Space-Time Adaptive Numerical Method for Partial Differential Equations*, Math. Model. Num. Anal. **26**, pp. 703–734.

BERGER, M., AND KOHN, R. V. (1988), *A Rescaling Algorithm for the Numerical Studies of Blowing-Up Solutions*, Comm. Pure Appl. Math. **41**, pp. 841–863.

BEYLKIN, G. (1992), *On the Representation of Operators in Bases of Compactly Supported Wavelets*, SIAM J. Numer. Anal. **29**, pp. 1716-1740.

BEYLKIN, G., AND KEISER, J. M. (1997), *On the Adaptive Solution of Nonlinear Partial Differential Equations in Wavelet Bases*, J. Comput. Phys. **132**, pp. 233-259.

CAI, W., AND WANG, J. Z. (1996), *Adaptive Wavelet Collocation Methods for Initial-Boundary Value Problems of Nonlinear PDE*, SIAM J. Numer. Anal. **33**, pp. 937–970.

CAI, W., AND ZHANG, W. (1998), *An Adaptive Spline Wavelet ADI (SW-ADI) Method for Two-Dimensional Reaction-Diffusion Equations*, J. Comput. Phys. **139**, pp. 92-126.

DAUBECHIES, I. (1988), *Orthonormal Bases of Compactly Supported Wavelets*, Comm. Pure Appl. Math. **41**, pp. 909–996.

FORNBERG, B. (1998), *Calculation of Weights in Finite Difference Formulas*, SIAM Review **40**, pp. 685–691.

GOTTLIEB, D., AND HESTHAVEN, J. S. (2001), *Spectral Methods for Hyperbolic Problems*, J. Comput. Appl. Math. **128**, pp. 83-131.

GOTTLIEB, D., AND SHU, C. W. (1997), *On the Gibbs Phenomenon and its Resolution*, SIAM Review **39**, pp. 644-668.

HESTHAVEN, J. S., AND JAMESON, L. M. (1998), *A Wavelet Optimized Adaptive Multi-Domain Method.* J. Comput. Phys. **145**, pp. 280-296.

HESTHAVEN, J. S., JUUL RASMUSSEN, J., BERGÉ, L., AND WYLLER, J. (1997), *Numerical Studies of Localized Wavefields Governed by the Raman-Extended Derivative Nonlinear Schrödinger Equation*, J. Phys. A: Math. Gen. **30**, pp. 8207–8224.

JAMESON, L. (1993), *On the Wavelet Based Differentiation Matrix*, J. Sci. Comput. **8**, pp. 267-305.

JAMESON, L. (1996), *The Differentiation Matrix for Daubechies-Based Wavelets on the Interval*, SIAM J. Sci. Comput. **17**, pp. 498-516.

JAMESON, L. (1998), *A Wavelet-Optimized, Very High Order Adaptive Grid and Order Numerical Method*, SIAM J. Sci. Comp. **19**, pp. 1980–2013.

KATEHI, L. P. B., HARVEY, J. F., AND TENTZERIS, E. (1998), *Time-Domain Analysis Using Multiresolution Expansions.* In *Advanced in Computational Electrodynamics. The Finite-Difference Time-Domain Method.* A. Taflove (Eds.). Artech House, Boston.. pp. 111-162.

NEWELL, A. C., AND MOLONEY, J. V. (1992), *Nonlinear Optics.* Addison-Wesley Publishing Company, California.

SCHWAB, C. (1998), *p- and hp-Finite Element Methods. Theory and Applications in Solid and Fluid Mechanics.* Numerical Mathematics and Scientific Computation. Clarendon Press, Oxford.

VASILYEV, O. V., AND PAOLUCCI, S. (1996), *A Dynamically Adaptive Multilevel Wavelet Collocation Method for Solving Partial Differential Equations in a Finite Domain*, J. Comput. Phys. **125**, pp. 498-512.

WHITHAM, G. B. (1974), *Linear and Nonlinear Waves.* John Wiley & Sons. New York.

Figure Captions

Fig 1 Index structure of the mesh, directly reflecting the local resolution level.

Fig 2 Examples of wavelet optimized mesh generation and the effect of changing the threshold values for functions with different degrees of regularity.

Fig 3 Soliton collision in the nonlinear Schrödinger (NLS) equation. The absolute value of $u(x,t)$ is plotted at the time slots $t = 0, 0.2, 0.4, 0.5, 0.6, 0.8$, illustrating complete recovery of the initial conditions following a strongly nonlinear interaction.

Fig 4 Formation, propagation, and decay of solutions with local regions of very steep gradients obtained by solving Burgers equation with $\nu = 0.001$. The solution is shown at equidistant time intervals of length 0.25. On the right is shown the magnified plots with circles designating the grid points.

Fig 5 Illustration of the grid/order selection process as a function of time for solving Burgers equation as illustrated in Fig. 4. The open symbols signify the percentages of points advanced with a scheme of a given order while the full symbols measures the total number of active grid points.

Fig 6 Collapsing solution of the critical nonlinear Schrödinger equation with a slightly supercritical initial condition. The magnitude, $|u|$, is shown at times $t = 0, 0.0118, 0.0122$, in the full domain (left) as well as in a closeup (right), illustrating the highly adapted grid structure around the collapsing solution as it evolves.

Fig 7 Illustration of the grid/order selection process as a function of time for solving the critical nonlinear Schrödinger equation equation. The open symbols signify the percentages of points advanced with a scheme of a given order while the full symbols measures the total number of active grid points.

Fig 8 Collapsing solution of extended derivative nonlinear Schrödinger equation with a supercritical CNLS initial condition. The magnitude, $|u|$, is shown at times $t = 0, 0.6504, 0.6644, 0.6692, 0.6707$ in the full domain (left) as well as in a closeup (right), illustrating the highly adapted grid structure around the collapsing asymmetric solution as it evolves.

Fig 9 Illustration of the grid/order selection process as a function of time for solving the extended derivative nonlinear Schrödinger equation. The open symbols signify the percentages of points advanced with a scheme of a given order while the full symbols measures the total number of active grid points.

Figure 1

Figure 2

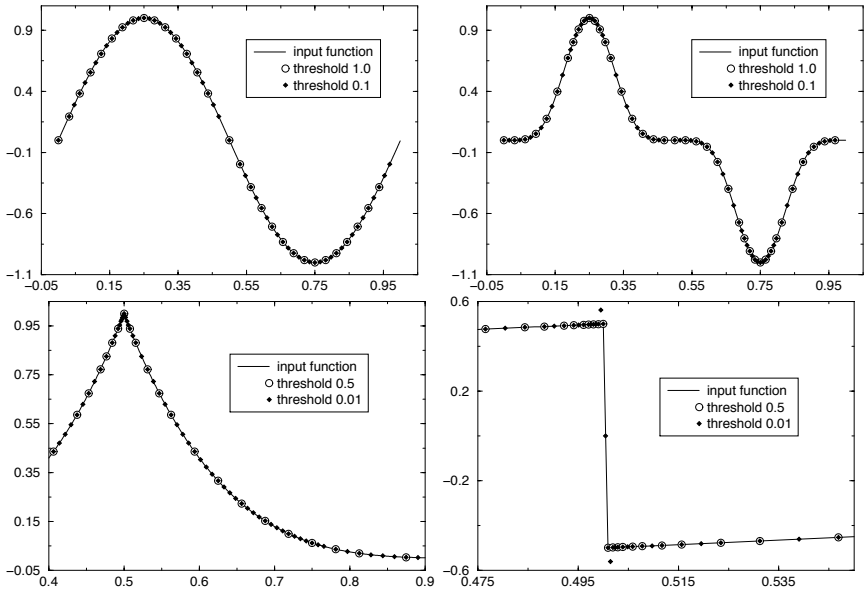
input function
threshold 1.0
threshold 0.1
input function
threshold 1.0
threshold 0.1
input function
threshold 0.5
threshold 0.01
input function
threshold 0.5
threshold 0.01

Figure 3

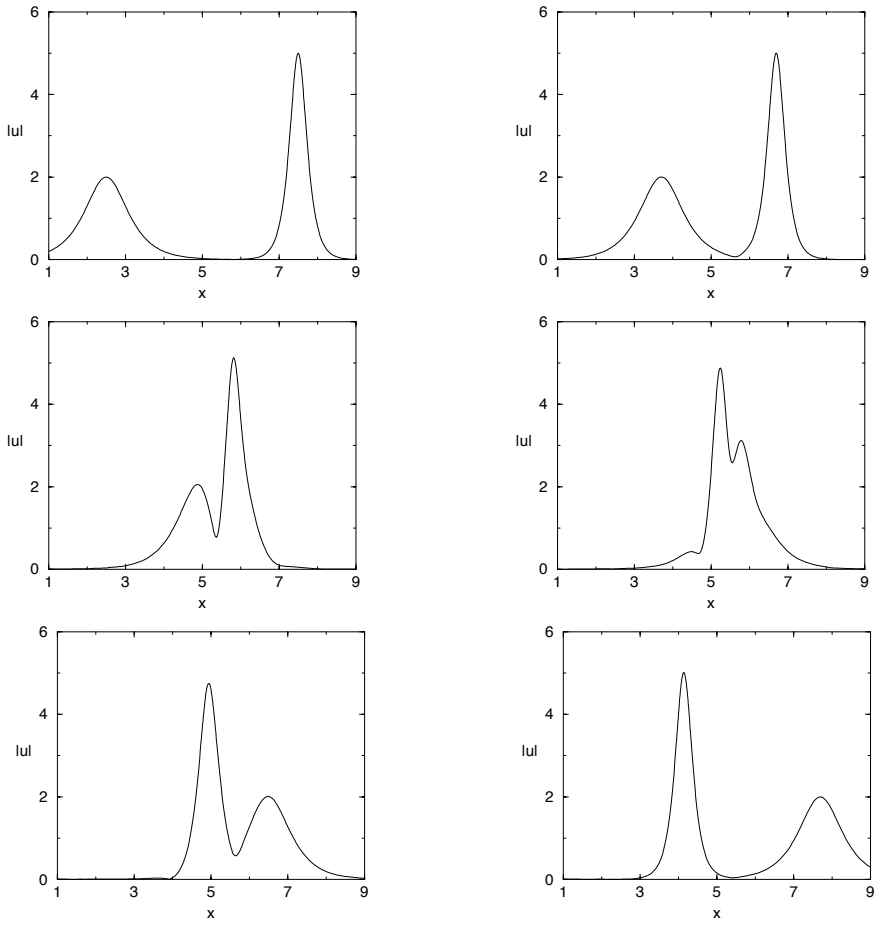

Figure 4

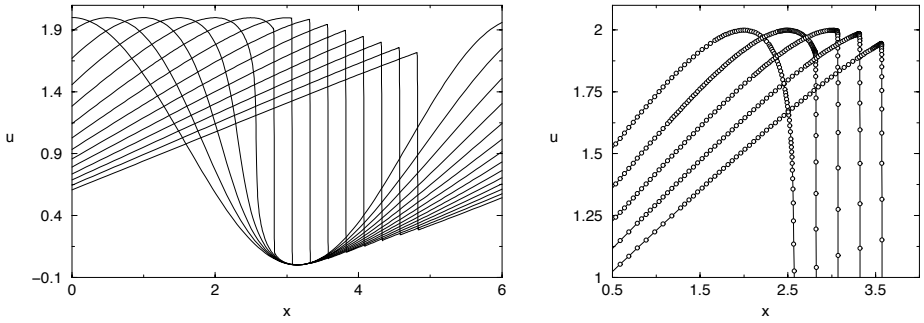

Figure 5

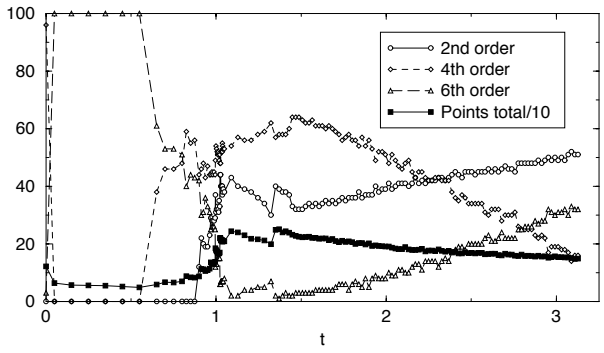
100
80
60
40
20
0
0
1
2
3
t
2nd order
4th order
6th order
Points total/10

Figure 6

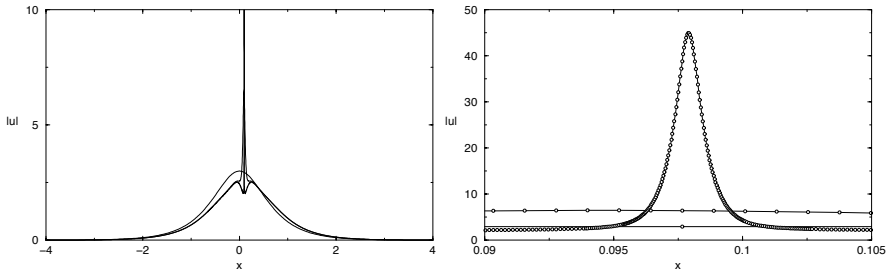

Figure 7

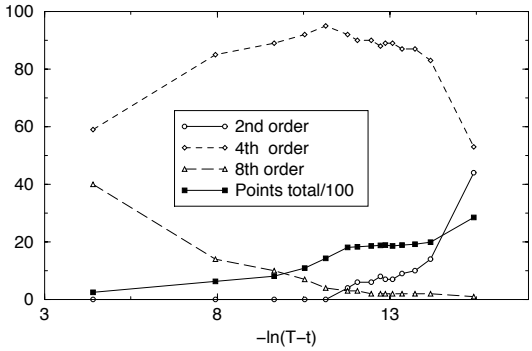
100
80
60
40
20
0
3
8
13
−ln(T−t)
2nd order
4th order
8th order
Points total/100

Figure 8

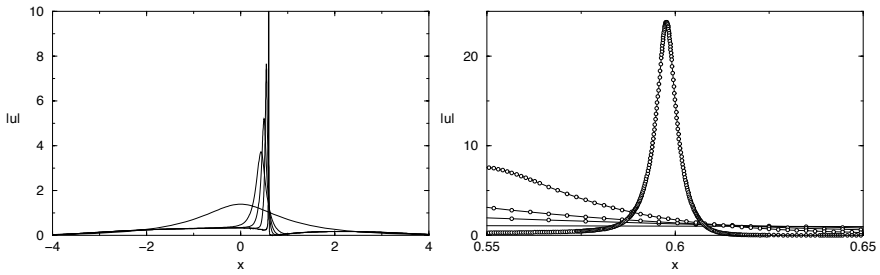

Figure 9

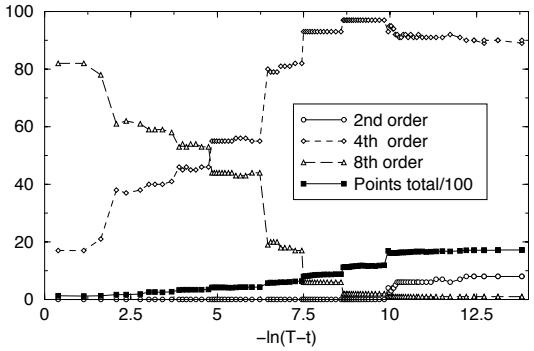
100
80
60
40
20
0
0
2.5
5
7.5
10
12.5
−ln(T−t)
2nd order
4th order
8th order
Points total/100

Table Captions

Table 1 Filter bank coefficients for D_4.

Table 2 Table used to guide the compromise between the order of the approximation and the level of refinement around location x_j.

Table 1

k	h_k	g_k
1	$(1+\sqrt{3})/4$	$(1+\sqrt{3})/4$
2	$(3+\sqrt{3})/4$	-$(3+\sqrt{3})/4$
3	$(3-\sqrt{3})/4$	$(3-\sqrt{3})/4$
4	$(1-\sqrt{3})/4$	-$(1-\sqrt{3})/4$

Table 2

order	n_j	m_j
1	18	–
2	16	18
4	12	16
8	8	8
12	4	4