Combining Wavelets with Finite Differences: Consistency Analysis

by
Sônia M . Gomes[1] and **Bertil Gustafsson**

Abstract

The method is an adaptive finite difference strategy for numerical solution of evolution partial differential equations. The principle is to represent the solution only through those point values indicated by the significant wavelet coefficients. Typically, few points are found in each time step, the grid being coarse in smooth regions, and refined close to irregularities. At each point, the spatial derivatives are discretized by uniform finite differences, using step size proportional to the point local scale. Eventually, required neighboring stencils are not present in the grid. In such case, the correspondig point values are approximated from coarser scales by using reconstruction operators defined by means of interpolating subdivision scheme. Our purpose in this paper is to analyse a generalization of the concept of truncation error, which is the familiar basis of the analysis of difference schemes. For this consistency analysis, we show that the adaptive finite difference scheme can also be formulated in terms of a collocation scheme for an adapted wavelet expansion of the solution. For this purpose, we first prove some results concerning the local behavior of the reconstruction operators, which stand for appropiate cone-like grids.

1 Introduction

The method to be analyzed here was introduced by Mats Holmström in his PhD thesis [9], and it was called the SPR method, for *sparse point representation.* The SPR method is an adaptive finite difference strategy for the numerical solution of evolution partial differential equations

$$u_t(x,t) = Lu(x,t),\ t > 0,\ x \in \Omega \tag{1}$$

augmented with initial and boundary conditions. The method combines the simplicity and accuracy of traditional finite difference schemes with the ability of wavelet coefficients in the characterization of local regularity of functions. In Holmström's case, the differential equation is of hyperbolic type, and the method was used to simulate some typical models in fluid dynamics, showing savings in CPU time when compared with the standard finite difference scheme.

The idea is to represent the functions by the point values corresponding to their significant wavelet coefficients. Typically, few points are found in each time step, the grid being coarse in smooth regions, and refined close to irregularities. At each point, the spatial derivatives are discretized by uniform finite differences, using

[1]IMECC-UNICAMP, Caixa Postal 6065, 13081-970 Campinas SP, Brasil

step size proportional to the point local scale. Eventually, required neighboring stencils are not present in the grid. In such case, the corresponding point values are approximated from coarser scales by using an interpolating subdivision scheme.

Our purpose in this paper is to analyze, on this adaptive finite difference strategy, a generalization of the concept of truncation error, which is the familiar basis of the analysis of difference schemes. This is a first and crucial step towards a complete error analysis, incorporating both time and space errors. For the consistency analysis, we show that, for a class of differential operators, this adaptive finite difference scheme can also be formulated in terms of a collocation scheme for an adapted wavelet expansion of the solution.

The paper is organized as follows. In Section 2, the basic aspects of the SPR method are summarized. In Section 3, required tools from wavelet analysis are described in detail. It seems that the appropriate context is the one of multiresolution analysis for point values on irregular grids, as defined by F. Plantevin in [10]. In our methodology, some concepts shall be used of the general approach suggested by Ami Harten [5], which is based on the concepts of conservative approximation schemes $\{\mathcal{D}^j, \mathcal{R}^j\}$, in terms of discretization and reconstruction operators. In Section 4 we prove the main results concerning the local behavior of the reconstruction operators, which stand for appropriate cone-like grids. Section 5 is dedicated to the discretization of derivatives in the multiresolution context, and to the analysis of the consistency of adaptive finite differences arising in the SPR scheme.

2 Some aspects of the SPR method

In a standard finite difference scheme, solutions are represented by vectors $U^n = U^{n,j}$ containing approximated node values at points $\mu = k2^{-j}$ on a uniform grid X^j, and time $t^n = n\Delta t$, i.e. $U^n_\mu \approx u(\mu, n\Delta t), \quad \mu \in X^j$. At the next time step, U^{n+1} is obtained by the application of a discrete evolution operator $\mathcal{L} = \mathcal{L}^{j,\Delta t}$, such that $U^{n+1} = \mathcal{L}U^n$. The action of $\mathcal{L}$ includes the discretization of the spatial differential operator L by a difference scheme L^j_{FD}, the discretization in time by means of some ODE solver, and the enforcement of boundary conditions.

For the purposes of this paper, we shall assume that this reference scheme is stable in the sense that $\mathcal{L}$ is a contraction

$$||\mathcal{L}U - \mathcal{L}W|| \leq ||U - W||, \tag{2}$$

for some appropriate norm $||.||$.

In the adaptive method, the goal is to use a finite difference model in a more efficient fashion, by taking into account local regularity information about the numerical solution. The first step in this direction is the representation of the numerical solution $U^n_a = U^{n,j}_a$ on a sparse grid $\Gamma^n = \Gamma^{j,n}$. Γ^n is expected to be coarse where the solution is smooth, and refined in regions of sharp variations. For the solution to evolve from U^n_a into U^{n+1}_a, three basic steps are undertaken.

1. **Refinement:** $U_a^{n+1/2} \leftarrow \mathcal{E} U_a^n$
2. **Evolution:** $\bar{U}_a^{n+1} \leftarrow \mathcal{L}_a U_a^{n+1/2}$
3. **Coarsening:** $U_a^{n+1} \leftarrow \mathcal{T}_\epsilon \bar{U}_a^{n+1}$

Refinement is done to prevent sharp gradients to move into coarse grid areas between time steps. Since the regions of smoothness or irregularities of the solution may change with time, Γ^n may not be convenient anymore at the next time step $t = t^{n+1}$. Therefore, before doing the time evolution, the representation of the solution should be extended to a grid $\Gamma^{n+1/2}$, which is expected to contain Γ^{n+1}. Then, a time evolution operator $\mathcal{L}_a$ is applied. The subscript a means that spatial derivatives at points $\mu \in \Gamma^{n+1/2}$ are performed by means of uniform finite differences, using a step size proportional to the point local scale. Finally, a thresholding operation $\mathcal{T}_\epsilon$ (coarsening) is applied, in order to eliminate from $\Gamma^{n+1/2}$ those points that are unnecessary for an accurate representation of U_a^{n+1}.

For the implementation of this adaptive algorithm, the basic tools come from wavelet analysis. For discretizations of a function on a grid $\Gamma \subset X^j$, the wavelet transform makes use of reconstruction operators, which are defined by means of an interpolating subdivision scheme adapted to Γ. The discrete information of a function f at Γ is organized by levels of resolution. At each level l, the wavelet coefficients d^l are defined in terms of local interpolation errors at the new points of the current level of discretization. In the construction of the adapted grids, the idea is to use the wavelet coefficients as indicators of local smoothness of the solution. This is the main ingredient in the definition of the thresholding operator $\mathcal{T}_\epsilon$: only those points corresponding to wavelet coefficients of magnitude greater than ϵ should be kept in the adapted grid Γ_ϵ. The interpolating subdivision scheme is also used to predict the solution at points not present in the grid. For instance, this is necessary in the extension operation, to obtain the approximate solution at points $\mu \in \Gamma^{n+1/2} \setminus \Gamma^n$, and also in the computations of the adapted finite differences, when some required stencil points are not in $\Gamma^{n+1/2}$. The details of such procedures shall be analyzed in the following sections.

⋆ Error analysis

Let $\tilde{U}_a^{n+1}$ be the extension of the adaptive solution U_a^{n+1} up to X^j, by means of the interpolating subdivision scheme. The error E^{n+1} shall be considered as the difference between the numerical solution given by the standard finite difference scheme based on X^j and $\tilde{U}_a^{n+1}$, i.e. $E^{n+1} = U^{n+1} - \tilde{U}_a^{n+1}$. Having in mind that $U^{n+1} = \mathcal{L} U^n$, the error E^{n+1} can be split into three terms

$$E^{n+1} = \left(\mathcal{L}U^n - \mathcal{L}\tilde{U}_a^n\right) + \left(\mathcal{L}\tilde{U}_a^n - \widetilde{\mathcal{L}_a U_a^{n+1/2}}\right) + \left(\widetilde{\mathcal{L}_a U_a^{n+1/2}} - \tilde{U}_a^{n+1}\right).$$

The first term is the difference of the application of the reference finite difference scheme $\mathcal{L}$ on different data. Therefore, its behavior depends on the stability properties of $\mathcal{L}$. The second term is the truncation error, which expresses the difference

in first extending and then applying the usual finite difference, instead of applying the adapted finite difference discretization and then extending. The last term corresponds to the thresholding error, since U_a^{n+1} is the result of the application of the thresholding operator on $\mathcal{L}_a U_a^{n+1/2}$.

Assume that the initial error $||E^0||$, the thresholding and truncation errors can be bounded by a control parameter ϵ. Having in mind the stability hypothesis (2), it follows that

$$||E^{n+1}|| \quad = \quad C(n+1)\epsilon.$$

Usually, the control of the thresholding error is simple. However, regarding the truncation error, this is a more delicate task. The contributions to the truncation error may come from two sides. Despite the fact that $\tilde{U}_a^n$ has no wavelet components outside $\Gamma^{n+1/2}$, the action of $\mathcal{L}$ may activate some of them. Following the ideas proposed in [2], a careful refinement strategy can be performed in order to keep such kind of perturbation under control. However, inside $\Gamma^{n+1/2}$, another source of errors may arise if $\mathcal{L}$ is not calculated exactly. For instance, this could be the case if $\mathcal{L}$ is replaced by its adaptive version $\mathcal{L}_a$. In this paper we are mainly concerned with this kind of error, whose properties are associated with the adaptive finite difference strategy.

3 Multiresolution analysis on irregular grids

As explained before, our goal in this paper is to show how the the truncation error can be controlled, regarding possible perturbations coming from the application of the adaptive scheme $\mathcal{L}_a$ instead of the reference scheme $\mathcal{L}$. Before going to this point in the last two sections, we shall describe here the required tools from wavelet analysis in more detail.

⋆ Discrete aspect

In a multiscale framework, functions are represented at different scale levels. The main tool in a multiscale application is a transformation $WT = WT_L^j$ linking the highest scale level j to the lower ones. The input data form a vector f^j containing the most refined information about the function. After the transformation WT, the output contains the information f^L at the coarsest level, and multilevel information $d^l, L \leq l \leq j-1$, that keeps the details between a scale level l and the next upper level $l+1$. Schematicly, we have

$$f^j \stackrel{WT}{\rightarrow} (f^L, d^L, \ldots, d^{j-1}).$$

Usually, the discrete values f^j give local information of a function f associated with a certain grid Γ^j. Typically, they are produced by the action of local functionals. For instance, they can be point values or local weighted averages. In wavelet analysis,

the multilevel informations d^l are known as *wavelet coefficients*, and WT corresponds to the analysis part of the *wavelet transform*.

We shall refer to the mapping which associates the discrete values with a given function as the *discretization operator*

$$\mathcal{D}^j : f \to f^j.$$

In this paper, only discretization by point values $f^j_\mu = f(\mu), \mu \in \Gamma^j$ shall be considered. Another important question is how $f(x)$ can be predicted, at any location x, from the knowledge of f^j. This led us to define a *reconstruction operator* $\mathcal{R}^j$, which creates an approximation of $f(x)$ from the discrete values f^j

$$\mathcal{R}^j : f^j \to \mathcal{R}^j(x; f^j).$$

The approximation scheme, defined by the pair of operators $\{\mathcal{D}^j, \mathcal{R}^j\}$, is *conservative* provided the discretization of the reconstruction reproduces the input data. In the case of discretization by point values, this means that $\mathcal{R}^j(\mu; f^j) = f^j_\mu, \mu \in \Gamma^j$, so that $\mathcal{R}^j(x; f^j)$ is an interpolation operator.

The principle stands for having the information data f^j hierarchically organized by different levels of resolution. On this matter, we shall be concerned with a sequence $\{\Gamma^l\}_{L \le l \le j}$ of embedded grids $\Gamma^l \subset \Gamma^{l+1}$. It shall be assumed that $\Gamma^L = X^L$, and $\Gamma^l \subset X^l$, for $l > L$, where X^l is the dyadic uniform grid with grid size $h_l = 2^{-l}$, $X^l = \{\mu = kh_l,\ k \in \mathbf{Z}\}$. It shall also be assumed that Γ^l is constructed from Γ^{l-1} by including some new points of $X^l \setminus X^{l-1}$. Let the set of these new points be denoted by $\Lambda^{l-1} = \Gamma^l \setminus \Gamma^{l-1}$.

Given the approximation scheme $\{\mathcal{D}^j, \mathcal{R}^j\}$, the wavelet transform WT can be defined as follows. Since $f^j_\mu = f(\mu), \mu \in \Gamma^j$, then the discrete values at level $j-1$ are just obtained by restriction $f^{j-1}_\mu = f^j_\mu$ for $\mu \in \Gamma^{j-1}$. At this lower level, we can use $\mathcal{R}^{j-1}(\mu; f^{j-1})$ to estimate f^j_μ at the new points $\mu \in \Gamma^j$ that are not in Γ^{j-1}

$$f^j_\mu \approx \mathcal{R}^{j-1}(\mu; f^{j-1}).$$

The entries in d^{j-1} are then defined as the error in this approximation

$$d^{j-1}_\mu = f^j_\mu - \mathcal{R}^{j-1}(\mu; f^{j-1}),\ \mu \in \Lambda^{j-1}.$$

So defined, d^{j-1} measures the ability of the approximation scheme in estimating the discretization f^j from the knowledge of f^{j-1}.

With the knowledge of f^{j-1} and d^{j-1}, we recover $f^j_\mu,\ \mu \in \Lambda^{j-1}$ by the formula

$$f^j_\mu = d^{j-1}_\mu + \mathcal{R}^{j-1}(\mu; f^{j-1}).$$

⋆ The reconstruction operator

In the definition of $\mathcal{R}^j(x; f^j)$, the main ingredient is the interpolating prediction operator P^l_{l-1}.

Definition 3.1 Prediction operator: *Let* $s^{l-1} = (s^{l-1}(\mu))$ *and* $s^l = (s^l(\mu))$ *be sequences labeled by the points* $\mu \in X^{l-1}$ *and* $\mu \in X^l$*, respectively. The prediction operator* $P_{l-1}^l s^{l-1}$ *operating on a sequence* s^{l-1} *gives a sequence* s^l *such that* $s^l(\mu) = s^{l-1}(\mu)$*, if* $\mu \in X^{l-1}$*.*

At this point, a particular interest emerges regarding the predictions $P_{l-1}^l s^{l-1}$ by means of central polynomial Lagrange interpolation. The most simple example is given by central linear interpolation. Precisely, if $\mu = (2k+1)2^{-l}$, then $s^l(\mu)$ is obtained by interpolating $s^{l-1}(\mu^-)$ and $s^{l-1}(\mu^+)$, where $\mu^- = k2^{-l+1}$ and $\mu^+ = (k+1)2^{-l+1}$ are the closest neighbors of μ in X^{l-1}. Generally, if M is even, and the M interpolation points are chosen as close as possible to μ, we are in the classical central interactive subdivision scheme by Dubuc-Deslauriers [3]. Precisely, if $\mu = (2k+1)2^{-l} \in X^l \setminus X^{l-1}$, then $s^l(\mu)$ is given by the evaluation at μ of the polynomial that interpolates $s^{l-1}(\nu)$ at $\nu = (k-s)2^{-l+1}, -M/2 \le s \le M/2-1$. This scheme is also known to be convergent in the sense that, from any starting sequence s^0, there exists a continuous function $s(x)$ such that $s(\mu) = s^l(\mu),\ \mu \in X^l$.

For the definition of $\mathcal{R}^j(x; f^j)$, a simple modification is used for the classical subdivision scheme. Firstly, the definition is set for all dyadic points. Typically, this is done interactively, from bottom to top. On the first step, we set $\mathcal{R}^j(\nu; f^j) = f_\nu^j, \nu \in X^L$. Having $\mathcal{R}^j(\nu; f^j)$ defined for all $\nu \in X^{l-1}, l \ge L+1$, then the definition is extended to $X^l \setminus \Lambda^{l-1}$ by means of P_{l-1}^l, and it is set $\mathcal{R}^j(\nu; f^j) = f_\nu^j$ at $\nu \in \Lambda^{l-1}$. This procedure is summarized in the next definition.

Definition 3.2 Interpolating subdivision scheme for irregular grids:

Given $f_\mu^j, \mu \in \Gamma^j$*, define*

1. *For* $\nu \in X^L$

$$\mathcal{R}^j(\nu; f^j) = s^L(\nu) = f_\nu^j.$$

2. *For* $\nu \in X^l, L < l \le j$*,*

$$\mathcal{R}^j(\nu; f^j) = s^l(\nu) = \begin{cases} f_\nu^j & if\ \nu \in \Lambda^{l-1} \\ P_{l-1}^l s^{l-1}(\nu) & if\ \nu \in X^l \setminus \Lambda^{l-1}. \end{cases}$$

3. *For* $\nu \in X^l, l > j$*,*

$$\mathcal{R}^j(\nu; f^j) = s^l(\nu) = P_{l-1}^l s^{l-1}(\nu).$$

The second part consists in proving, by density and continuity arguments, that it makes sense to extend the definition for all $x \in \mathbf{R}$. Since for $\nu \in X^l, l > j$, the definition of $\mathcal{R}^j(\nu; f^j)$ proceeds as in a regular subdivision scheme, the definition of $\mathcal{R}^j(x; f^j)$ can be extended to all $x \in \mathbf{R}$, provided that the regular subdivision scheme is convergent.

As indicated by A. Harten [5], by choosing different prediction operators, this setting opens unlimited possibilities for the construction of interpolating multiresolution analyses. The prediction operators may even be nonlinear mappings. For instance, in the so called ENO schemes [6], the choice of the interpolation points is data dependent, and the purpose is to minimize oscillations close to points of sharp variation. However, it should be noted that the question of convergence is still an open problem for such nonlinear subdivision schemes.

⋆ Functional aspect

To fix ideas, let us restrict ourselves to the case of predictions by central Lagrange interpolation. From now on, the parameter M will always be connected with the degree $M-1$ of the polynomial interpolation used in the prediction operator. In this context, the reconstruction $\mathcal{R}^j(x; f^j)$ is a linear operator, and it can be expressed as

$$\mathcal{R}^j(x; f^j) = \sum_{\mu \in \Gamma^j} f^j_\mu \Phi^j_\mu(x),$$

where the function $\Phi^j_\mu(x)$ is the reconstruction $\mathcal{R}^j(x; \delta^j)$ of the delta-sequence $\delta^j_\nu = \delta(\mu - \nu)$. Therefore, Φ^j_μ satisfies the interpolatory property

$$\Phi^j_\mu(\nu) = \delta(\mu - \nu), \quad \nu \in \Gamma^j.$$

From the definition of the reconstruction operators, we conclude that

$$\mathcal{R}^j(x; f^j) = \mathcal{R}^{j+1}(x; \mathcal{D}^{j+1}\mathcal{R}^j(.; f^j)). \tag{3}$$

As a consequence, $\Phi^j_\mu(x)$ satisfies the scale relation

$$\Phi^j_\mu(x) = \sum_{\nu \in \Gamma^{j+1}} \Phi^j_\mu(\nu)\Phi^{j+1}_\nu(x).$$

Define

$$\mathcal{V}^j = span\{\Phi^j_\mu \ : \ \mu \in \Gamma^j\},$$

then it is clear that $\mathcal{V}^{j-1} \subset \mathcal{V}^j$. Let us also introduce the notation $\phi^j_\mu(x)$ and V^j for the case of uniform grid $\Gamma^j = X^j$. In this uniform setting, it is known that there is a basic function $\phi(x)$ such that $\phi^j_\mu(x) = \phi(2^j(x - \mu))$. From the definition of the reconstruction operator, it follows immediately that $\mathcal{V}^j \subset V^j$. Therefore, the functions in $\mathcal{V}^j$ inherit all regularity properties verified by the classical interpolatory scaling function $\phi(x)$.

Let $\mu \in \Lambda^{j-1}$. By definition, $\Phi_\mu^j(\nu) = 0$ for all $\nu \in \Gamma^j \setminus \{\mu\}$. It is easy to conclude from the bottom-to-top interpolation algorithm that $\Phi_\mu^j(\nu)$ also vanishes at all $\nu \in X^j \setminus \{\mu\}$. Therefore, for points μ in Λ^{j-1} the basic functions $\Phi_\mu^j(x)$, of the irregular case, and $\phi_\mu^j(x)$, of the uniform case, coincide.

Consider the wavelet subspace

$$\mathcal{W}^{j-1} = span\{\psi_\mu^{j-1}(x); \mu \in \Lambda^{j-1}\},$$

where $\psi_\mu^{j-1}(x) = \phi_\mu^j(x)$. Note that $\mathcal{W}^{j-1}$ is formed by those functions in $\mathcal{V}^j$ that vanish on Γ^{j-1}. Therefore, $\mathcal{V}^j$ is the direct sum of the spaces $\mathcal{V}^{j-1}$ and $\mathcal{W}^{j-1}$

$$\mathcal{V}^j = \mathcal{V}^{j-1} \oplus \mathcal{W}^{j-1}.$$

Consequently, the following two-level representation holds

$$\sum_{\mu \in \Gamma^j} f_\mu^j \Phi_\mu^j(x) \;=\; \sum_{\mu \in \Gamma^{j-1}} f_\mu^{j-1} \Phi_\mu^{j-1}(x) + \sum_{\mu \in \Lambda^{j-1}} d_\mu^{j-1} \psi_\mu^{j-1}(x). \tag{4}$$

Therefore, the difference of information between two consecutive levels is expressed by the wavelet terms

$$\sum_{\mu \in \Lambda^{j-1}} d_\mu^{j-1} \psi_\mu^{j-1}(x) = \mathcal{R}^j(x; f^j) - \mathcal{R}^{j-1}(x; f^{j-1}).$$

Interactively applying (4), we obtain the multilevel decomposition

$$\sum_{\mu \in \Gamma^j} f_\mu^j \Phi_\mu^j(x) = \sum_{\mu \in X^L} f_\mu^L \phi_\mu^L(x) + \sum_{l=L}^{j-1} \sum_{\mu \in \Lambda^l} d_\mu^l \psi_\mu^l(x). \tag{5}$$

Under this functional point of view, the transformations $f^j \leftrightarrow (f^L, d^L, \ldots, d^{j-1})$ correspond to the change between the basis $\{\Phi_\mu^j\}$ and $\{\phi_\mu^L\} \cup \{\psi_\mu^L\} \cup \ldots \cup \{\psi_\mu^{j-1}\}$. In the case of regular grids X^j, this interpolating wavelet approach was introduced by Donoho [4]. In the context of irregular grids, it was first considered in [10], and then extended in [8] to discretizations by cell averages on the interval.

Remark 3.1 *Note that the basic functions appearing on the right hand side of (5) correspond to the multiresolution analysis based on the uniform grid. This means that $\{\Phi_\mu^j(x)\}$ forms a basis for the subspace $\mathcal{V}^j \subset V^j$ that is spanned by $\{\phi_\mu^L(x)\}$ and the wavelets $\{\psi_\mu^l(x)\}$ associated with the points $\mu \in \Lambda^l, L \le l \le j-1$. It should be also noted that, for functions $f \notin \mathcal{V}^j$, the wavelet coefficients $d_\mu^l, \mu \in \Lambda^l$ appearing on the right hand side of (5), which are defined by means of the refinement scheme adapted to Γ^j, may be different from the corresponding wavelet coefficients $\tilde{d}_\mu^l$ produced by the multiresolution algorithm based on the full grid X^j. However, it is possible to have the same wavelet coefficients $d_\mu^l = \tilde{d}_\mu^l$, $\mu \in \Lambda^l$, provided that Γ^j satisfies an appropriate cone structure. This will be explained later, in Remark 4.1.*

⋆ **Sparse point representation**

In the above discussion, multiresolution analyses for point values on a pre-established irregular grid have been considered. The connection between the grid and the particular function under analysis has not been covered. However, in wavelet analysis this kind of irregular grid appears after the thresholding of wavelet coefficients. Therefore, different grids are usually associated to distinct functions. The terminology *sparse point representation* introduced in [9] is related to this kind of mapping between functions and grid point values. The main operations involved are described next.

- **Grid reduction:** $\Gamma^j \xrightarrow{\mathcal{T}_\epsilon} \Gamma_\epsilon$

  Suppose that a certain function f is represented by means of the node values on a certain grid Γ^j. The goal is to obtain a more economic representation, by means of node values in a subgrid $\Gamma_\epsilon \subset \Gamma^j$. This is done by using a thresholding operation $\mathcal{T}_\epsilon$. It acts after the analysis algorithm WT by simply removing from Γ^j those points whose corresponding wavelet contributions have magnitude less than ϵ.

- **Grid extension:** $\Gamma^j \xrightarrow{\mathcal{E}} \Gamma$

  There is another useful converse operation. From the representation of f by node values on a grid Γ^j, we may want another representation on an extended grid Γ containing Γ^j. Since in principle we may not have access to the exact node value at a new point $\nu \in \Gamma \setminus \Gamma^j$, it is set by the application of the bottom-to-top interpolation algorithm $\mathcal{R}(\nu; f^j)$ associated to Γ^j.

⋆ **A note about the notation**

In applications of this paper, there are several grids involved, with discretizations and reconstructions operators associated to them, and thus several multiresolution analyses. It is useful to keep a unified notation for most of the operations involved, independently of which sparse grid the discretization and reconstruction operators are associated with. For instance, we have already used the convenient notation $\tilde{f}^j$ to indicate the extension of f^j to the full grid X^j by the bottom-to-top interpolation scheme $\tilde{f}^j_\nu = \mathcal{R}^j(\nu; f^j),\ \nu \in X^j$. Similarly, the multiresolution transform, and the thresholding operators are persistently referred to as WT and $\mathcal{T}_\epsilon$, independently of the multiresolution analysis at hand. It is clear that always $\mathcal{T}_\epsilon f^j = \mathcal{T}_\epsilon \tilde{f}^j$.

4 Local behavior of the reconstruction operators

⋆ **Cone condition and local scale**

In wavelet analysis, the grids that satisfy a cone condition are of particular interest, since they arise in connection with significant wavelet coefficients of functions having local singularities.

Definition 4.1 Cone condition: *Let $\rho > 0$ be an integer. Γ^j satisfies the ρ–cone condition if for $l \leq j-1$, and for $\gamma \in \Lambda^l$, all 2ρ points in X^l nearest to γ also belong to Γ^l.*

The cone condition ensures that the transition from the lower level to the highest one is not abrupt. If it is not otherwise stated, from now on we shall assume that Γ^j satisfies a cone condition with ρ at least equal to 1.

Remark 4.1 *Suppose that Γ^j satisfies a ρ-cone condition with $\rho \geq M/2$. Then, for a given point $\gamma \in \Lambda^l, l \leq j-1$, the M points in X^l closest to μ are also in Γ^l. Therefore, the computation of $\mathcal{R}^l(\gamma; f^l)$ interpolates the same values as in a regular subdivision scheme. This means that the wavelet coefficient d^l_γ coincides with the wavelet coefficient $\tilde{d}^l_\gamma$ produced by the uniform multiresolution analysis. Thus, the representation (5) for such cone like grids, may be interpreted as the result of a classical multiresolution representation of the function f followed by compression, i.e. after discarding the wavelet coefficients associated with $\nu \in X^{l+1} \setminus \{X^l \cup \Lambda^l\}$.*

On an irregular grid, the concept of local scale is also important.

Definition 4.2 Local scale: *With each point $\gamma \in \Gamma^j$ we associate the number $a^j(\gamma)$ defined as the distance between γ and its nearest neighbor in Γ^j, i.e.,*

$$a^j(\gamma) \;=\; \inf\{|\gamma - \mu| : \;\; \mu \in \Gamma^j, \mu \neq \gamma\}.$$

The results in the next lemma were proved in [10].

Lemma 4.1 *: Suppose that the grid Γ^j satisfies a cone condition with $\rho \geq 1$. Then*

1. *$2^{-j} \leq a^j(\gamma) \leq 2^{-L}, \gamma \in \Gamma^j$.*
2. *There exists n, $L \leq n \leq j$ such that $a^j(\gamma) = 2^{-n}$.*
3. *If $\gamma \in \Lambda^j$, then $a^{j+1}(\gamma) = 2^{-j-1}$.*
4. *If $\gamma \in \Gamma^j$, and $a^j(\gamma) = 2^{-n}$, $n \leq j$ then $\gamma \in \Gamma^n$.*

The behavior of the reconstruction $\mathcal{R}^j(x; f^j)$ close to a point $\mu \in \Gamma^j$ depends on its local scale. For instance, it was proved in [10] that if Γ^j is a cone-like grid with $\rho \geq M+1$, then $\Phi^j_\mu \in V^{n+1}$, where 2^{-n} is the local scale of μ. Without imposing such a restrictive assumption on the cone condition, some weaker local results shall be derived in the next lemmas, which are sufficient for our purposes in the next section.

Lemma 4.2 *Let $\mu \in \Gamma^j$ with $a^j(\mu) = 2^{-n}$, $n \leq j$. Then for $M = 2$*

$$\mathcal{R}^j(x; f^j) = \mathcal{R}^n(x; f^n)$$

for $x \in [\mu^-, \mu^+]$, where μ^- and μ^+ are the closest neighbors of μ in Γ^j.

Proof. Both $\mathcal{R}^j(x; f^j)$ and $\mathcal{R}^n(x; f^n)$ are continuous piecewise linear functions. The break points for $\mathcal{R}^j(x; f^j)$ is Γ^j, and for $\mathcal{R}^n(x, f^n)$ is Γ^n. By Lemma 4.1–4, $a^j(\mu) = 2^{-n}$ means that $\mu \in \Gamma^n$. In addition, the cone condition implies that both μ^- and μ^+ are in Γ^n. To see that, suppose that $\mu^- = (2k+1)2^{-l-1} \in \Lambda^l, l \geq n$. The cone condition implies that $\gamma = (k+1)2^{-l} \in \Gamma^l$. Since $\mu^- < \gamma$, and $|\mu^- - \gamma| = 2^{-l-1} < 2^{-n} = a^j(\mu)$, then $\mu^- < \gamma < \mu$. But this contradicts the fact that μ^- is the closest left neighbor of μ in Γ^j. With similar argument, it can also be proven that $\mu^+ \in \Gamma^n$. Since $\mathcal{R}^j(x; f^j)$ and $\mathcal{R}^n(x, f^n)$ coincide on X^n, and both do not have any other breakpoints along the interval $[\mu^-, \mu^+]$, then they must coincide there. □

Lemma 4.3 *Let $\mu \in \Gamma^j$ with $a^j(\mu) = 2^{-n}$, $n \leq j$. Then for $M = 4$*

$$\mathcal{R}^j(\nu; f^j) = \mathcal{R}^n(\nu; f^n)$$

for $\nu \in S^l = \{\mu + m2^{-l},\ |m| \leq 2\}$, and $n \leq l \leq j$.

Proof. Since $\mathcal{R}^j(x; f^j)$ and $\mathcal{R}^n(x; f^n)$ coincide on X^n, then the assertion is true for $l = n$. The proof advances by induction for $l \geq n+1$. For $m = \pm 2$, $\nu = \mu \pm 2^{-l+1} \in S^{l-1}$, we already know, from the previous step, that the result is true. For $m = \pm 1$, the corresponding points $\nu = \mu \pm 2^{-l}$ satisfy $|\nu - \mu| = 2^{-l} < 2^{-n}$. Thus, they are not in Γ^j (otherwise the local scale of μ would be $2^{-l} < 2^{-n}$). Consequently, $\mathcal{R}^j(\nu; f^j)$ and $\mathcal{R}^n(\nu; f^n)$ are obtained by cubic interpolation from their corresponding values at the 4 points in X^{l-1} closest to ν. For $\nu = \mu - 2^{-l}$, the interpolation points are $\mu - 2^{-l+2}, \mu - 2^{-l+1}, \mu$ and $\mu + 2^{-l+1}$. For $\nu = \mu + 2^{-l}$, the interpolation points are $\mu - 2^{-l+1}, \mu, \mu + 2^{-l+1}$, and $\mu + 2^{-l+2}$. In both cases $\mathcal{R}^j(x; f^j)$ and $\mathcal{R}^n(x; f^n)$ coincide at the interpolation points. Therefore they coincide at ν as well. □

Extending the result of Lemma 4.3 to higher order schemes, with $M \geq 6$, may be of interest. In the above analysis, the stencil points in S^l fall into one of the following three categories. There is a first step $l = n_0$ such that $\mathcal{R}^{n_0}(x; f^{n_0}) = \mathcal{R}^j(x; f^j)$, for $x \in S^{n_0}$, just because both reconstructions coincide over all X^{n_0}. For $l > n_0$, we

consider first those points $\nu = \mu + m2^{-l} \in S^l$ associated with even m, which are also in S^{l-1}. For them we already know, from the previous step, that the result is true. Using a local scale argument, the points in S^l corresponding to odd m result to be in $X^l \setminus \Gamma^j$. The induction argument works by checking that their interpolation points in X^{l-1} required for both interpolations also belong to the previous stencil set. According to Lemma 4.3, the choice $n_0 = n$ works well for $M = 4$. However, this parameter must be modified if we want to mimic the above arguments to consider higher orders $M \geq 6$, and wider stencils. For instance, problems may arise in the very first step, because the points $\nu = \mu \pm 3 \times 2^{-n-1}$ may be points in Γ^j. For $M = 6$ this difficulty may be avoided just by starting one level up, at $n_0 = n + 1$.

Lemma 4.4 *Let $\mu \in \Gamma^j$ with $a^j(\mu) = 2^{-n}$, $n < j$. Then for $M = 6$*

$$\mathcal{R}^j(\nu; f^j) = \mathcal{R}^{n+1}(\nu; f^{n+1})$$

for $\nu \in S^l = \{\mu + m2^{-l},\ |m| \leq 4\}$, and $n + 1 \leq l \leq j$.

Proof. As described above, it only remains to consider the stencil points in $S^l, l \geq n + 2$, associated with odd m, i.e., $\nu = \mu + m2^{-l}, m = \pm 1, \pm 3$. Note that $|\mu - \nu| \leq 3 \times 2^{-l} < 2^{2-l} \leq 2^{-n}$, which implies that $\nu \in X^l \setminus \Gamma^j$ (by the local scale argument). Therefore, $\mathcal{R}^j(\nu; f^j)$ and $\mathcal{R}^n(\nu, f^n)$ are both obtained by interpolation of degree 5 from their corresponding values at the 6 points of X^{l-1} centered around ν. Let us check that all such interpolation points are in the previous stencil set. For $\nu = \mu + (2k+1)2^{-l} \in S^l$ the associated interpolation points are $x = \mu + (k-s)2^{-l+1}, -3 \leq s \leq 2$. For $-2 \leq k \leq 1$, this gives the range $-4 \leq k - s \leq 4$, which means that $x \in S^{l-1}$. The result then follows as in the proof of Lemma 4.3. □

With increasing orders $M \geq 8$, the sets of interpolation points become wider. Thus the stencil points must also be increased. But then another difficulty appears. We cannot perform anymore the step of proving that a certain stencil point is not in Γ^j by just using the local scale argument. As we will see next, this problem may be solved by imposing an appropriate cone condition.

Lemma 4.5 *Assume $M \geq 8$, and that Γ^j satisfies a cone condition with $\rho = \frac{M}{2} - 2$. Let $\mu \in \Gamma^j$ with $a^j(\mu) = 2^{-n}$, $n < j$. Then, for $\nu \in S^l = \{\mu + m2^{-l},\ |m| \leq M - 2\}$, and $n + 1 \leq l \leq j$*

$$\mathcal{R}^j(\nu; f^j) = \mathcal{R}^{n+1}(\nu; f^{n+1}).$$

Proof. According to the above comments, there are two things to prove. Firstly, the stencil points in $S^l, l \geq n + 2$, associated with odd $m = 2k + 1$ do not belong to Γ^j. Suppose this is not true for some $-M/2 + 1 \leq k \leq M/2 - 2$. The cone condition implies that $\mu + (k - s)2^{-l+1}, -M/2 + 2 \leq s \leq M/2 - 3$ also belong to Γ^j. By choosing $s = k - 1$ or $s = k + 1$, it follows that $\mu + 2^{-l+1}$ or $\mu - 2^{-l+1}$ are in Γ^j, which is a contradiction. Secondly, let us prove that all the interpolation points in X^{l-1} required to perform $\mathcal{R}^j(\nu; f^j)$ and $\mathcal{R}^{n+1}(\nu; f^{n+1})$ at $\nu = \mu + (2k + 1)2^{-l} \in S^l$ belong to the previous stencil set. Such interpolation points are $x = \mu + (k - s)2^{-l+1}, -M/2 \leq s \leq M/2 - 1$. For $-M/2 + 1 \leq k \leq M/2 - 2$, this gives the range $-M + 2 \leq k - s \leq M - 2$, which means that $x \in S^{l-1}$. The result then follows as in the previous cases. □

5 Discretization of derivatives

In this section, the focus is on dicretizations for the derivative $Lf = \frac{d}{dx}f$ based on the approximation scheme $\{\mathcal{D}^j, \mathcal{R}^j\}$. However, it is possible to extend the analysis to other differential operators as well. For instance, with slight modifications, the following arguments still hold for higher order derivatives, linear differential operators with variable coefficients, multilinear operators (e.g. $Lf = f\frac{d}{dx}f$).

Definition 5.1 Collocation scheme: *Let $M \geq 4$. Given the node values $f^j = \mathcal{D}^j f$ of a function f on the grid Γ^j, the collocation scheme $L^j f^j$ is defined as the node values of the derivative of the reconstruction $\mathcal{R}^j(x; f^j)$, i.e.,*

$$L^j f^j = \mathcal{D}^j[L(\mathcal{R}^j(.; f^j)]. \tag{6}$$

Recall that $\mathcal{R}^j(x; f^j)$, as a function in $\mathcal{V}^j \subset V^j$, may be represented as

$$\mathcal{R}^j(x; f^j) = \sum_{k \in Z} \tilde{f}_k^j \phi(2^j x - k),$$

where $\tilde{f}_k^j = \mathcal{R}^j(k2^{-j}; f^j)$ is the result of the application of the reconstruction algorithm up to the level j. Using this representation, one obtains at $\mu = m2^{-j} \in \Gamma^j$

$$\frac{d}{dx}\mathcal{R}^j(.; f^j)(\mu) = 2^j \sum_{k \in Z} \tilde{f}_k^j \phi'(m - k).$$

Based on this identity, $(L^j f^j)_\mu$ may be interpreted as the result of the action of a finite difference operator

$$(L^j f^j)_\mu = \left(L_{FD}^j \mathcal{R}^j(.; f^j)\right)(\mu), \quad \mu \in \Gamma^j, \tag{7}$$

where

$$(L_{FD}^j f)(x) = 2^j \sum_{k \in Z} f(x - k2^{-j})\beta(k), \tag{8}$$

with coefficients $\beta(k) = \phi'(k)$. This finite difference operator is closely connected with the underlying approximation scheme, which is exact for polynomials of degree less or equal $M - 1$. Then the associated finite difference operator also has order M. Since $\phi(x)$ is a symmetric function supported on the interval $[-M + 1, M - 1]$, the coefficients $\beta(n) = \phi'(n)$ end up being anti-symmetric, $\beta(n) = -\beta(-n)$, and $\beta(n) = 0, |n| \geq M - 1$.

It is clear that the formulation (7) – (8) also makes sense if L_{FD}^j is replaced by any other finite difference operator. For instance, for $M = 2$ we may use the second order central scheme with $\beta(-1) = -\beta(1) = 1/2$ and $\beta(n) = 0$ otherwise. For $M = 4$, the standard central difference operator has anti-symmetric coefficients $\beta(1) = 2/3, \beta(2) = -1/12$, and $\beta(n) = 0, n > 2$. This is also the same one associated with the collocation scheme (6), with cubic central interpolation. For $M \geq 6$ we may use a much shorter standard central finite difference scheme. This idea of

designing hybrid formulations, by replacing L^j_{FD} in formula (8) by a less costly finite difference operator, was suggested in [7].

⋆ Adaptive Formulation

M. Holmström [9] introduced the idea of implementing the finite difference scheme (7)–(8) in an adaptive fashion. The main ingredients are the bottom-to-top interpolation of degree $M-1$, combined with a short-length finite difference scheme of order M, but adapted to the local scale of each point.

Definition 5.2 Adaptive finite difference scheme $L^j_a = L^j_{a,\ell}$: *Let L^l_{FD} be a given finite difference operator based on the uniform grid X^l. For each point $\mu \in \Gamma^j$ with local scale $a^j(\mu) = 2^{-n}$ define*

$$(L^j_a f^j)_\mu = [L^{n+\ell}_{FD} \mathcal{R}^{n+\ell}(.; f^{n+\ell})](\mu). \tag{9}$$

According to this definition, the scheme proposed in [9] corresponds to $\ell = 0$. As we will see below, under certain circumstances, the adaptive scheme $L^j_a f$ may be equivalent to its non adaptive version $L^j f^j$, given in (8). However, in terms of operation count, the formulation (9) is more economic than (8), since it does not require the calculation of the stencils up to the most refined level, for those points $\mu \in \Gamma^j$ having big local scales $a^j(\mu) > 2^{-j}$. In the next statements, we shall assume that the finite difference schemes are associated with the interpolating subdivision schemes used in the reconstructions.

Theorem 5.1 *Assume that Γ^j satisfies the hypotheses of Lemmas 4.2 or 4.3, respectively for $M = 2$ or $M = 4$. Let $\mu \in \Gamma^j$ with $a^j(\mu) = 2^{-n}, n \leq j$. Then*

$$(L^j f^j)_\mu = (L^n f^n)_\mu.$$

That is, the adaptive finite difference scheme $L^j_a f^j$, with $\ell = 0$, coincides with its non adaptive version $L^j f^j$.

Proof. Let $\mu \in \Gamma^j$ with $a^j(\mu) = 2^{-n}, n < j$. For $M = 2$, Lemma 4.2 states that both $\mathcal{R}^n(x; f^n)$ and $\mathcal{R}^j(x; f^j)$ coincide on the interval $[\mu^-, \mu^+]$, where they are continuous piecewise linear, with break-points at μ^-, μ and μ^+. Let α^- and α^+ be their slopes on $[\mu^-, \mu]$ and $[\mu, \mu^+]$, respectively. Therefore

$$(L^j_a f^j)_\mu = [L^n_{FD} \mathcal{R}^n(.; f^n)](x)\mu = \frac{\alpha^+ + \alpha^-}{2} = [L^j_{FD} \mathcal{R}^j(.; f^j)](\mu) = (L^j f^j)_\mu.$$

Let us recall that for $M = 4$ the finite difference formulation and the collocation scheme are the same. This implies that

$$(L^j_a f^j)_\mu = [L^n_{FD} \mathcal{R}^n(.; f^n)](\mu) = \frac{d}{dx} \mathcal{R}^n(.; f^n)(\mu) = [L^j_{FD} \mathcal{R}^n(.; f^n)](\mu).$$

Lemma 4.3 assures that $\mathcal{R}^n(x; f^n)$ and $\mathcal{R}^j(x; f^j)$ coincide on the stencil points required for the evaluation of L^j_{FD} at μ. Thus, $[L^j_{FD}\mathcal{R}^n(.; f^n)](\mu) = [L^j_{FD}\mathcal{R}^j(.; f^j)](\mu)$, which gives

$$(L^j_a f^j)_\mu = [L^j_{FD}\mathcal{R}^j(.; f^j)](\mu) = (L^j f^j)_\mu.$$

□

Using the results of Lemmas 4.4 and 4.5, the following equivalence statement also holds.

Theorem 5.2 *Suppose that Γ^j satisfies the hypotheses of Lemmas 4.4 or 4.5, respectively for $M = 6$ or $M \geq 8$. Let $\mu \in \Gamma^j$ with $a^j(\mu) = 2^{-n}, n < j$. Then*

$$(L^j f^j)_\mu = (L^{n+1} f^{n+1})_\mu.$$

That is, the collocation scheme $L^j f^j$ coincides with its adaptive version $L^j_a f^j$, with $\ell = 1$.

Proof. According to formula (8), the evaluation of $\frac{d}{dx}\mathcal{R}^j(.; f^j)$ at $\mu = m2^{-j} \in \Gamma^j$ only requires the knowledge of the stencils $\tilde{f}^j_k = \mathcal{R}^j(\nu; f^j)$, for $\nu = \mu + k2^{-j}, 1 \leq |k| \leq M - 2$. As a consequence of Lemmas 4.4 – 4.5, we also have $\tilde{f}^j_k = \mathcal{R}^{n+1}(\nu; f^{n+1})$. This means that

$$(L^j f^j)_\mu = \frac{d}{dx}\mathcal{R}^j(.; f^j)(\mu) = \frac{d}{dx}\mathcal{R}^{n+1}(.; f^{n+1})(\mu) = (L^{n+1} f^{n+1})_\mu.$$

□

⋆ Consistency analysis

Let L^j_{FD} be a uniform finite difference operator with step 2^{-j}. Given a grid Γ^j and a function $f(x) = \mathcal{R}^j(x; f^j) \in \mathcal{V}^j$, consider

$$g(x) = [L^j_{FD}\mathcal{R}^j(.; f^j)](x).$$

What can be said about the wavelet content of g ?

$\mathcal{R}^j(x; f^j)$ only contains contributions of wavelets ψ^l_ν, $\nu \in \Gamma^j$. However, differentiation is an action that activates some other wavelets. Thus, components $\psi^l_\nu(x)$, with $\nu \in X^j \setminus \Gamma^j$, may be present in $g(x)$. For instance, consider the simple case where $f(x) = \psi^l_\mu(x)$ is the wavelet associated to some point $\mu \in \Lambda^l$. For this function, $\tilde{d}^l_\mu = 1$ is the only active wavelet coefficient. However, as detailed in the Appendix, there are several active wavelet coefficients in $[L^j_{FD}\psi^l_\mu](x)$, at all scales.

It would be helpful to introduce the functionals ψ^{*n}_γ that produce the wavelet coefficients in a uniform grid

$$\tilde{d}^n_\gamma = \left\langle \psi^{*n}_\gamma, f \right\rangle.$$

Having in mind the multilevel representation of $f \in \mathcal{V}^j$,

$$f(x) = \mathcal{R}^j(x; f^j) = \sum_{\mu \in \Gamma^L} f^L_\mu \phi^L_\mu(x) + \sum_{l=L}^{j-1} \sum_{\mu \in \Lambda^l} d^l_\mu \psi^l_\mu(x),$$

one obtains

$$< \psi^{*n}_\gamma, g >= \sum_{\mu \in \Gamma^L} f^L_\mu \mathcal{M}(\gamma, \mu) + \sum_{l=L}^{j-1} \sum_{\mu \in \Lambda^l} d^l_\mu \mathcal{M}(\gamma, \mu),$$

where

$$\mathcal{M}(\gamma, \mu) = \begin{cases} < \psi^{*n}_\gamma, L^j_{FD} \phi^L_\mu >, & \text{for } \mu \in X^L \\ \\ < \psi^{*n}_\gamma, L^j_{FD} \psi^l_\mu >, & \text{for } \mu \in \Lambda^l. \end{cases}$$

Some details about the entries $\mathcal{M}(\gamma, \mu)$ are given in the Appendix. For instance, if Γ^j satisfies a cone condition, then the significant contributions $\mathcal{M}(\gamma, \mu) \neq 0$ only come from few points μ in Γ^j close to γ. This means that if γ is a point far enough from Γ^j, then we may expect that $< \psi^{*n}_\gamma, g >$ is small. In fact, as it will be described in [1], following the ideas in [2], for a prescribed accuracy $\epsilon > 0$, it is possible to find an appropriate extended grid $\hat{\Gamma}^j$ such that the magnitude of the wavelet coefficients $| < \psi^{*n}_\gamma, g > |$ is less than ϵ, for all points γ outside $\hat{\Gamma}^j$.

Given such an extended grid $\hat{\Gamma}^j$, containing Γ^j, let $\hat{g}^j_\mu = [L^j_a \hat{f}^j](\mu)$, $\mu \in \hat{\Gamma}^j$ be the result of applying the adaptive difference scheme based on $\hat{\Gamma}^j$. We are interested in the error TE in approximating g by the interpolation of $\hat{g}^j$. That is,

$$TE = g(x) - \mathcal{R}^j(x; \hat{g}^j).$$

From now on, depending on M, let us assume that the original given grid Γ^j satisfies a cone condition compatible with the hypotheses of Theorems 5.1 or 5.2. Furthermore, suppose that in the extension process such a cone property is maintained in $\hat{\Gamma}^j$. As a consequence of the equivalence results in Theorems 5.1 and 5.2, g coincide at $\mu \in \hat{\Gamma}^j$ with the adaptive finite difference values $\hat{g}^j_\mu = [L^j_a \tilde{f}^j](\mu)$. This means that $TE(\gamma) = 0$ for all $\gamma \in \hat{\Gamma}^j$. As interpolation errors, $TE(\gamma)$ give the wavelet coefficients of g at points $\gamma \notin \hat{\Gamma}^j$. As a consequence of the previous considerations, it holds that $|TE(\gamma)| < \epsilon$ for $\gamma \in X^j \setminus \hat{\Gamma}^j$.

If we see the finite difference operator L^j_{FD} as playing the role of the derivative, then TE can be interpreted as *the difference between the derivative of the reconstruction and the reconstruction of the adaptive discrete derivative.* In this sense, we call TE the *truncation error.* However, if we want to be more specific, then we should note that $g(x) \in \mathcal{V}^j$ is just an approximation of the derivative of $\mathcal{R}^j(x; f^j)$. In fact, for $M \geq 4$, since the finite difference operator comes from the collocation scheme, $g(x)$ indeed coincides with $\frac{d}{dx}\mathcal{R}^j(x; f^j)$ at X^j.

These statements are summarized in the following theorem.

Theorem 5.3 *Consider a grid Γ^j satisfying a cone condition compatible with the hypotheses of Theorems 5.1 or 5.2, and let $f(x) = \mathcal{R}^j(x; f^j)$ be a function in $\mathcal{V}^j$.*

Given a precision $\epsilon > 0$, then there exists an extended grid $\hat{\Gamma}^j = \hat{\Gamma}^j(f, \epsilon)$ such that, for the adaptive finite discretization $L_a^j \hat{f}^j$ based on $\hat{\Gamma}^j$, the truncation error

$$TE = L_{FD}^j \mathcal{R}^j(x; f^j) - \mathcal{R}^j(x; L_a^j \hat{f}^j),$$

satisfies

$$|TE(\gamma)| < \epsilon, \ \forall \gamma \in X^j.$$

6 Appendix: Estimates for $\mathcal{M}(\gamma, \mu)$

$\star$ $\underline{M = 2}$

For $\mu \in X^{l+1} \setminus X^l$, let $\psi_\mu^l(x) = \phi_\mu^{l+1}(x)$ be its associated wavelet. It is a hat function in X^{l+1}, supported on the interval $[\mu^-, \mu^+]$, where μ^- and μ^+ are the two closest neighbors of μ in X^l. Let $g(x) = L_{FD}^j \psi_\mu^l(x) \in V^j$. It has the following node values at $x \in X^j$

$$g(x) = \begin{cases} 2^{l+1} & \mu^- < x < \mu \\ -2^{l+1} & \mu < x < \mu^+ \\ 2^l & x = \mu^- \\ -2^l & x = \mu^+ \\ 0 & \text{elsewhere.} \end{cases}$$

For $\gamma \in X^{n+1} \setminus X^n$, let γ^- and γ^+ be its closest neighbors in X^n. Consider the functional

$$\psi^*{}_\lambda(x) = \delta(x - \gamma) - \frac{\delta(x - \gamma^+) + \delta(x - \gamma^-)}{2},$$

which produces the error by linear interpolation

$$d_\gamma^n = \left\langle \psi_\gamma^*, f \right\rangle = f_\gamma - \frac{f_{\gamma^+} + f_{\gamma^-}}{2}.$$

Our purpose is to describe the elements

$$\mathcal{M}(\gamma, \mu) = \left\langle \psi_\gamma^*, g(x) \right\rangle.$$

It is clear that $\mathcal{M}(\gamma, \mu) = 0$ if $S = [\mu^-, \mu^+] \cap \{\gamma^-, \gamma, \gamma^+\} = \emptyset$. For instance, this is the case if $\gamma^+ < \mu^-$, or $\gamma^- > \mu^+$. Let us see what happens in the other cases.

1. $n = l$: The following three cases may occur.

   (a) $\mu = \gamma \Rightarrow S = \{\mu^-, \mu, \mu^+\}$

   $$\mathcal{M}(\gamma, \mu) = g(\mu) - \frac{g(\mu^+) + g(\mu^-)}{2} = 0.$$

(b) $\mu^+ = \gamma^- \Rightarrow S = \{\mu^+\}$

$$\mathcal{M}(\gamma, \mu) = -\frac{g(\mu^+)}{2} = 2^{l-1}.$$

(c) $\mu^- = \gamma^+ \Rightarrow S = \{\mu^-\}$,

$$\mathcal{M}(\gamma, \mu) = -\frac{g(\mu^-)}{2} = -2^{l-1}.$$

2. $n > l$: There are ten possibilities, half of them corresponding to the case $\gamma^+ \leq \mu$, and the other ones to $\gamma^- \geq \mu$. On the left side of μ one has:

(a) $\gamma^+ = \mu^- \Rightarrow S = \{\mu^-\}$,

$$\mathcal{M}(\gamma, \mu) = -\frac{g(\mu^-)}{2} = -2^{l-1}.$$

(b) $\gamma^- = \mu^-, \gamma^+ = \mu \Rightarrow S = \{\mu^-, \gamma, \mu\}$

$$\mathcal{M}(\gamma, \mu) = 2^{l+1} - 2^{l-1} = \frac{3}{2}2^l.$$

(c) $\gamma^- = \mu^-, \gamma^+ < \mu \Rightarrow S = \{\mu^-, \gamma, \gamma^+\}$

$$\mathcal{M}(\gamma, \mu) = 2^{l+1} - \frac{2^l + 2^{l+1}}{2} = 2^{l-1}.$$

(d) $\gamma^- > \mu^-, \gamma^+ = \mu \Rightarrow S = \{\gamma^-, \gamma, \mu\}$

$$\mathcal{M}(\gamma, \mu) = 2^{l+1} - \frac{2^{l+1}}{2} = 2^l.$$

(e) $\gamma^- > \mu^-, \gamma^+ < \mu \Rightarrow S = \{\gamma^-, \gamma, \gamma^+\}$. It follows that $g(\gamma^-) = g(\gamma) = g(\gamma^+)$, and thus

$$\mathcal{M}(\gamma, \mu) = 0.$$

Since $g(x)$ is anti-symmetric around $x = \mu$, the five cases $2.(a^{'})$ – $2.(e^{'})$, on the right side of μ, just have the opposite sign.

3. $n < l$: On the left side of γ one has:

(a) $\gamma^- = \mu^+ \Rightarrow S = \{\mu^+\}$,

$$\mathcal{M}(\gamma, \mu) = -\frac{g(\mu^+)}{2} = 2^{l-1}.$$

(b) $\mu^- = \gamma^-, \mu^+ = \gamma \Rightarrow S = \{\mu^-, \mu^+\}$

$$\mathcal{M}(\gamma, \mu) = -2^l - 2^{l-1} = -\frac{3}{2}2^l.$$

(c) $\mu^- = \gamma^-, \mu^+ < \gamma \Rightarrow S = \{\mu^-\}$

$$\mathcal{M}(\gamma, \mu) = -2^{l-1}.$$

(d) $\gamma^- < \mu^-, \mu^+ = \gamma \Rightarrow S = \{\mu^+\}$

$$\mathcal{M}(\gamma, \mu) = -2^{l}.$$

(e) $\gamma^- < \mu^-, \mu^+ < \gamma \Rightarrow S = \emptyset.$

$$\mathcal{M}(\gamma, \mu) = 0.$$

As in the previous case, there are five other reverse possibilities $3.(a^{'})$ – $3.(e^{'})$ associated to the case to $\mu^- \geq \gamma$.

In the analysis of the present paper, cone-like grids Γ^j are considered, with $\rho \geq 1$. For a given point $\gamma = (2k+1)2^{-n-1} \notin \Gamma^j$, a particular interest emerges in identifying those points in Γ^j interacting with γ, i.e. such that $\mathcal{M}(\gamma, \mu) \neq 0$. As described next, at each level l, at most two points μ and $\lambda \in \Lambda^l$ have this property.

- This is clear at level $l = n$, since only $\mu = \gamma^-$ or $\lambda = \gamma^+$ may interact with γ, and in both cases $|\mathcal{M}(\gamma, \mu)| = |\mathcal{M}(\gamma, \mu)| = 2^{l-1}$.

- The most significant contributions occur at level $l = n - 1$. This is the only situation when the case $2.(b)$ may occur. It may be an isolated occurrence, given the contribution $|\mathcal{M}(\gamma, \mu)| = \frac{3}{2}2^l$. But it may also occur simultaneously with the case $2.(a)$, with $\lambda = \mu - 2^{-l}$. In such a case the additional contribution is $|\mathcal{M}(\gamma, \lambda)| = 2^{l-1}$. A similar analysis holds for the case $2.(b^{'})$. The case $2.(a)$ may also be an isolated occurrence. The same is true for $2.(a^{'})$. However, they can not hold at the same time. In both cases the contribution is $|\mathcal{M}(\gamma, \mu)| = 2^{l-1}$. Since the cases $2.(c), (d), (c^{'})$ and $(d^{'})$ do not apply here, there is no other possibility at this level.

- For the lower levels $l \leq n-2$, let the interaction from the inside be considered first, with γ close to the lower extreme point, i.e. $\gamma^- = \mu^-$ (case $2.(c)$). If it occurs, then only one more interaction is possible. In fact, it may only occur from the outside, with $\gamma^- = \lambda^+$ (case $2.(a)$), where $\lambda = \mu - 2^{-l}$. The contributions of these two interactions coincide $|\mathcal{M}(\gamma, \mu)| = |\mathcal{M}(\gamma, \lambda)| = 2^{l-1}$. The same behavior occurs when the interaction is close to the upper extreme point μ^+ (case $2.(c^{'})$). When the interaction is close to μ, from its left side (case $2.(d)$), then no other interaction is possible. In this case, the contribution is $|\mathcal{M}(\gamma, \mu)| = 2^l$. The same is true on the right side of μ (case $2.(d^{'})$). The case $2.(a)$, as well as $2.(a^{'})$, may also be the only occurrence with $|\mathcal{M}(\gamma, \mu)| = 2^{l-1}$. They cannot occur simultaneously. The cases $2.(b)$ and $2.(b^{'})$ do not apply.

- The cone condition implies that μ^+ and μ^-, the two closest neighbors of μ in X^n, are in $\Gamma^n \subset \Gamma^j$. Consequently, the interval (γ^-, γ^+) cannot contain any point from Γ^j. Thus, among the possibilities discussed above, $3.(b)$ – $3.(d)$, and their reverse ones $3.(b')$ – $3.(d')$, are clearly excluded. Consequently, at each level $l > n$, only external interactions $\mu^+ = \gamma^-$ or $\lambda^- = \gamma^+$ are possible, and again $|\mathcal{M}(\gamma, \mu)| = |\mathcal{M}(\gamma, \lambda)| = 2^{l-1}$.

$\star$ $\underline{M \geq 4}$

Recall that $\phi(x)$ is a symmetric function supported on the interval $[-M+1, M-1]$. For $\mu \in X^{l+1} \setminus X^l$, the wavelet $\psi_\mu^l(x) = \phi_\mu^{l+1}(x) = \phi(2^{l+1}(x-\mu))$, is supported on the interval $[\mu-(M-1)2^{-l-1}, \mu+(M+1)2^{-l-1}]$. The function $g(x) = L_{FD}^j \psi_\mu^l(x) \in V^j$ shall be considered, where the finite difference operator comes from the collocation algorithm. This means that $g(x)$ is the approximation of $L\psi_\mu^l(x)$ in V^j by interpolation on the uniform grid X^j. That is, $L_{FD}^j \psi_\mu^l(x) = 2^{l+1}\phi'(2^{l+1}(x-\mu))$, for $x \in X^j$. For $M \geq 4$, the functions in V^j have smoothness at least in $C^{1+\alpha}, 0 < \alpha < 1$. Therefore, g has the same smoothness.

The functionals that produce the interpolation error have the form

$$d_\gamma^n = \left\langle \psi_\gamma^{*n}, f \right\rangle = f_\gamma^{n+1} - \sum_{s=1}^{M/2} h_s (f_{\gamma^{+s}}^n + f_{\gamma^{-s}}^n),$$

where $\gamma^{+s}, \gamma^{-s}, s = 1, \ldots, M/2$, are the M points in X^n closest to γ, and h_s are the coefficients coming from the central Lagrange polynomial interpolation. Thus, ψ_γ^{*n} is supported on $\{\gamma, \gamma^{\pm s}, s = 1, \ldots, M/2\}$.

The following estimate holds

$$\begin{aligned} |\mathcal{M}(\gamma, \mu)| &= \left| \left\langle \psi_\gamma^{*n}, L_{FD}^j \phi_\mu^l \right\rangle \right| \\ &\leq ||\psi_\gamma^{*n}||_{\ell^1} ||L_{FD}^j \phi_\mu^l||_{L^\infty} \\ &\lesssim 2^l. \end{aligned}$$

Here the notation $||\psi_\gamma^{*n}||_{\ell^1} = 1 + 2\sum_{s=1}^{M/2} |h_s|$ is introduced. The fact that the functionals ψ_γ^{*n} cancel polynomials p of degree less or equal $M-1$ may be used to get

$$\mathcal{M}(\gamma, \mu) = \left\langle \psi_\gamma^{*n}, L_{FD}^j \phi_\mu^l - p \right\rangle.$$

Therefore

$$\begin{aligned} |\mathcal{M}(\gamma, \mu)| &\leq ||\psi_\gamma^{*n}||_{\ell^1} \inf_p ||L_{FD}^j \phi_\mu^l - p||_{L^\infty(supp\psi_\gamma^{*n})} \\ &\lesssim 2^{-n\alpha} ||L_{FD}^j \phi_\mu^l||_{C^\alpha(supp\psi_\gamma^{*n})} \\ &\lesssim 2^{-n\alpha} 2^{l(1+\alpha)} \\ &= 2^{(l-n)\alpha} 2^l. \end{aligned}$$

If $l < n$, a gain in having more smoothness is noticed.

For a given $\gamma \notin \Gamma^j$, the goal is to identify those points in $\mu \in \Gamma^j$ interacting with γ , i.e. such that $\mathcal{M}(\gamma,\mu) \neq 0$. As described next, for $M = 4$, there are at most six points μ at each level Λ^l having this property.

$\star$ **Special case:** $M = 4$

For $\mu \in X^{l+1} \setminus X^l$, the wavelet $\psi^l_\mu(x) = \phi^{l+1}_\mu(x) = \phi(2^{l+1}(x-\mu))$ is supported on the interval $[\mu - 3\times 2^{-l-1}, \mu + 3\times 2^{-l-1}]$. The wavelet coefficient d^n_γ is produced by the functional

$$\left\langle \psi^{*n}_\gamma, f \right\rangle = f^{n+1}_\gamma - \left[h_1(f^n_{\gamma^{+1}} + f^n_{\gamma^{-1}}) + h_2(f^n_{\gamma^{+2}} + f^n_{\gamma^{-2}})\right],$$

where $h_1 = 9/16$, $h_2 = -1/16$. Therefore,

$$||\psi^{*n}_\gamma||_{\ell^1} = 18/16 + 2/16 + 1 = 2.25,$$

and

$$\begin{aligned}
\mathcal{M}(\gamma,\mu) &= 2^{l+1}\Big\{\phi'(2^{l+1}(\gamma-\mu)) \\
&- h_1\left(\phi'(2^{l+1}(\gamma-\mu+2^{-n-1})) + \phi'(2^{l+1}(\gamma-\mu-2^{-n-1}))\right) \\
&- h_2\left(\phi'(2^{l+1}(\gamma-\mu+3\times 2^{-n-1})) + \phi'(2^{l+1}(\gamma-\mu-3\times 2^{-n-1}))\right)\Big\}.
\end{aligned}$$

It is helpful to have in mind that $\phi'(x)$ is an anti-symmetric function vanishing for $|x| \geq 3$, with $\phi'(0) = 0, \phi'(1) = -2/3$, and $\phi'(2) = 1/12$.

- At level $l = n$, $\gamma - \mu = s2^{-l}$. Therefore

$$\begin{aligned}
\mathcal{M}(\gamma,\mu) &= 2^{l+1}\Big\{\phi'(2s) - h_1\left(\phi'(2s+1) + \phi'(2s-1)\right) \\
&- h_2\left(\phi'(2s+3) + \phi'(2s-3)\right)\Big\}.
\end{aligned}$$

  This implies that $\mathcal{M}(\gamma,\mu) = 0$ if $|s| > 2$, and

$$\mathcal{M}(\gamma,\mu) = \begin{cases} \pm 2^{l+1}\left\{\phi'(2) - (h_1 - h_2)\phi'(1)\right\} = \pm 2^l, & \text{if } s = \pm 1, \\ \pm 2^{l+1}h_2\phi'(1) = \pm\frac{2^{l-2}}{3}, & \text{if } s = \pm 2. \end{cases}$$

- For $l = n+1$, $\gamma - \mu = (2s+1)2^{-l-1}$, which gives

$$\begin{aligned}
\mathcal{M}(\gamma,\mu) &= 2^{l+1}\Big\{\phi'(2s+1) - h_1\left(\phi'(2s+3) + \phi'(2s-1)\right) \\
&- h_2\left(\phi'(2s+7) + \phi'(2s-5)\right)\Big\}.
\end{aligned}$$

  If $|2s+1| > 9$, then $\mathcal{M}(\gamma,\mu) = 0$. Since in the interval $(\gamma^{-1}, \gamma^{+1})$ there are no points from Γ^j, then those points μ such that $|2s+1| = 1$ should be excluded.

Therefore, the only cases that should be considered are $|2s+1| = 3, 5$ and 7, such that

$$\mathcal{M}(\gamma,\mu) = \begin{cases} \mp 2^{l+1} h_1 \phi'(1) \pm 3 \times 2^{l-2}, & \text{if, } 2s+1 = \pm 3, \\ \pm 2^{l+1} h_2 \phi'(1) = \pm \frac{2^{l-2}}{3}, & \text{if } 2s+1 = \pm 5, \\ \mp 2^{l+1} h_2 \phi'(1) = \mp \frac{2^{l-2}}{3}, & \text{if } 2s+1 = \pm 7. \end{cases}$$

- Let $l \geq n+2$. Since it is not possible to have $\mu \in (\gamma^-, \gamma^+)$, the only interactions is with the neighbors $\gamma^{\pm s}$. Therefore, the cases that should be considered here are $\mu = \gamma^{+1} + 2^{-l-1}$, $\mu = \gamma^{+2} \pm 2^{-l-1}$, and the reverse ones $\mu = \gamma^{-1} - 2^{-l-1}$, and $\mu = \gamma^{-2} \pm 2^{-l-1}$. Thus, the cases to be considered are $\gamma - \mu = (2s+1)2^{-l-1}$, with $2s+1 = \pm(1+2^{l-n}), \pm(1-3\times 2^{l-n})$ and $\pm(1+3\times 2^{l-n})$, that produce the same values for $\mathcal{M}(\gamma,\mu)$, as in the previous level $l = n+1$.

⋆ **Acknowledgment**

The research of this paper started while the first author was visiting the Department of Scientific Computing (TDB), at Uppsala University. She is grateful for the hospitality received during her stay there. Her research has also been partially supported by CNPq and FAPESP (Brazil).

References

[1] A. Cohen, S. M. Gomes, S.-M. Kaber and M. Postel. Adaptive multiresolution processing of evolution equations. In preparation.

[2] A. Cohen, S.-M. Kaber, S. Müller and M. Postel. Fully adaptive multiresolution finite volume schemes for conservation laws. To appear in *Math. Comp.*, 2001.

[3] G. Deslauriers and S. Dubuc. Symmetric iterative interpolation processes. *Construc. Approx.*. **5**: 49-68, 1989.

[4] D. Donoho. Interpolating wavelet transform. Technical Report 408, Dept. of Statistics, Stanford University,

[5] A. Harten. Multiresolution representation of data: a general framework. *SIAM J. Numer. Anal.* 33 (3): 1205 – 1256, 1996.

[6] A. Harten, B. Engquist, S. Osher, and S. Chakravarthy. Uniformly high order accurate essentially nonoscillatory schemes III. *J. Comput. Phys.*, 71: 231–303, 1987.

[7] L. Jameson. On the wavelet optimized finite difference method. Technical Report ICASE 94-9, NASA Langely Research Center, March 1994.

[8] M. K. Kaibara and S. M. Gomes. Multiresolution analysis of cell averages based on non-uniform grids. Preprint 1997.

[9] M. Holmström. Wavelet based methods for time dependent PDEs. PhD Thesis, University of Uppsala, 1997.

[10] F. Plantevin. Wavelets on Irregular Meshes. *Advances in Computational Mathematics* **4**: 293-329, 1995.