Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 225 (2007) 1288–1313
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

The efficient implementation of a finite element, multi-resolution viscosity method for hyperbolic conservation laws

Marcus Calhoun-Lopez [a,1], Max D. Gunzburger [b,*,2]

[a] *Department of Mathematics, University of Maryland, College Park, MD 20742-4015, USA*
[b] *School of Computational Science, Florida State University, Tallahassee, FL 32306-4120, USA*

Received 29 November 2004; received in revised form 9 October 2006; accepted 25 January 2007
Available online 6 February 2007

Abstract

It is well known that the classic Galerkin finite element method is unstable when applied to hyperbolic conservation laws such as the Euler equations for compressible flows. Adding a diffusion term to the equations stabilizes the method but sacrifices too much accuracy to be of any practical use. An elegant solution developed in the context of spectral methods by Eitan Tadmor and coworkers is to add diffusion only to the high frequency modes of the solution and can lead to stabilization without sacrificing accuracy. We incorporate this idea into the finite element framework by using hierarchical functions as a multi-frequency basis. The result is a new finite element method for solving hyperbolic conservation laws. For this method, convergence for a one-dimensional scalar conservation law has previously been proved. Here, the method is described in detail, several issues connected with its efficient implementation are considered, and numerical results for several examples involving one- and two-dimensional hyperbolic conservation laws are provided. Several advantageous features of the method are discussed, including the ease for which discontinuities can be detected and artificial diffusion can be applied anisotropically and locally in physical as well as frequency space.

Keywords: Hyperbolic conservation laws; Finite element methods; Hierarchical bases; Multi-resolution viscosity 35L65; 65M60

1. Introduction

Naive, e.g., straightforward central difference or Galerkin, discretizations of hyperbolic conservation laws lead to unstable approximations. The most obvious stabilization approach is to add an artificial viscosity term to the conservation law but, as is well known, this leads to severe smearing of discontinuities and to low accuracy even in regions in which the solution is smooth. Of course, there have been many methods (finite difference, finite volume, finite element, spectral, kinetic, etc.) proposed for determining improved stabilized

* Corresponding author. Tel.: +1 850 644 7060; fax: +1 850 644 0098.
E-mail addresses: mcalhoun@math.umd.edu (M. Calhoun-Lopez), gunzburg@scs.fsu.edu (M.D. Gunzburger).
[1] Supported in part by the National Science Foundation under Grant No. DMS-0240049.
[2] Supported in part by the National Science Foundation under Grant No. DMS-0308845.

0021-9991/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2007.01.024

approximation solutions of hyperbolic conservation laws; see e.g. [2,8,11,14,15,18,19,22–25,28,29] for a small sample of the vast literature on the subject.

In this paper, a new finite element method for hyperbolic conservation laws is considered. The method is based on hierarchical basis functions and a scale-dependent artificial viscosity. Standard, nodal basis functions are all of the same scale, i.e., with respect to a given triangulation of a domain, their support is roughly equal. In contrast, hierarchical basis functions can be clustered into groups such that basis functions within a particular group are of a different scale from those within the other groups. The multi-scale nature of the hierarchical basis functions allows for the selective addition of viscosity only at the smallest scales, very much in the spirit of spectral viscosity methods. This flexibility allows for the stabilization of Galerkin finite element approximations and, at the same time, results in more accurate approximations both with respect to convergence rates in regions where the solution is smooth and the sharpness of the resolution of discontinuities in the solution. In [6], it was proved, for the case of a one-dimensional, periodic, scalar hyperbolic conservation law, that, under appropriate hypotheses, the approximate solution obtained using the new algorithm converges to the entropy solution of the conservation law. In this paper, the new algorithm is defined and several of its properties are discussed. Issues that arise in its efficient implementation are also considered and preliminary illustrative computational examples are provided.

1.1. Hyperbolic conservation laws

Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain. A general system of conservation laws has the form

$$\frac{\partial \mathbf{q}}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} \mathbf{f}_j(\mathbf{q}) = \mathbf{0} \quad \text{in } \Omega \times (0,\infty) \quad \text{and} \quad \mathbf{q}(\cdot,0) = \mathbf{g} \quad \text{in } \Omega \tag{1}$$

along with appropriate boundary conditions. Here, $\mathbf{q} : \Omega \times [0,\infty) \to \mathbb{R}^p$ denotes the vector-valued conserved variable, $\mathbf{f}_j : \mathbb{R}^p \to \mathbb{R}^p,\ j = 1,\ldots,d$, denote the d vector-valued flux functions, and $\mathbf{g} : \Omega \to \mathbb{R}^p$ denotes the given initial data. For $\widetilde{\mathbf{q}} \in \mathbb{R}^p$, let $\mathbf{A}_j(\widetilde{\mathbf{q}}) : \mathbb{R}^p \to \mathbb{R}^{p\times p}$ denote the $p \times p$ Jacobian matrix of $\mathbf{f}_j$, i.e., $\mathbf{A}_j(\widetilde{\mathbf{q}}) = \frac{\partial \mathbf{f}_j}{\partial \mathbf{q}}(\widetilde{\mathbf{q}})$. The system (1) is hyperbolic if for all solutions $\mathbf{q}$, any linear combination of $\{\mathbf{A}_j(\mathbf{q})\}_{j=1}^d$ has real eigenvalues with eigenvectors that span $\mathbb{R}^p$. The system (1) is strictly hyperbolic if the eigenvalues are distinct. See e.g. [11,9] for details.

The system (1) does not, in general, have a classical solution because of the spontaneous formation of discontinuities. Instead, one must look for a solution $\mathbf{q} \in L^\infty(\Omega \times (0,\infty); \mathbb{R}^p)$ which satisfies (1) in the distributional sense:

$$\int_0^\infty \int_\Omega \left[\mathbf{q} \cdot \frac{\partial \phi}{\partial t} + \sum_{j=1}^{d} \mathbf{f}_j(\mathbf{q}) \cdot \frac{\partial \phi}{\partial x_j} \right] \mathrm{d}\Omega\, \mathrm{d}t + \int_\Omega \mathbf{g} \cdot \phi(\cdot,0)\, \mathrm{d}\Omega = 0 \tag{2}$$

for all test functions $\phi \in C_0^\infty(\Omega \times [0,\infty); \mathbb{R}^p)$. It is clear that for a smooth enough solution, (1) and (2) are equivalent.

In the presence of discontinuities, solutions of the system (1) or of the weak formulation (2) are not uniquely determined. Additional conditions must be imposed to determine the unique, physically relevant solution. The second law of thermodynamics tells us that the entropy of the system should not decrease; satisfying this requirement suffices to allow one to obtain the unique, physically relevant entropy solution.

Let $\Phi, \{\Psi_j\}_{j=1}^d : \mathbb{R}^p \to \mathbb{R}$ be smooth functions; for (1), Φ is an entropy with entropy fluxes $\{\Psi_j\}_{j=1}^d$ if Φ is convex and $\nabla_{\mathbf{q}} \Phi^{\mathrm{T}} \frac{\partial \mathbf{f}_j}{\partial \mathbf{q}} = \nabla_{\mathbf{q}} \Psi_j$ in $\mathbb{R}^p$ for $1 \leqslant j \leqslant d$. A simple calculation gives us that if $\mathbf{q}$ is a smooth solution of (1), then $\Phi(\mathbf{q})$ satisfies a scalar conservation law with fluxes $\Psi_j(\mathbf{q})$:

$$\frac{\partial}{\partial t} \Phi(\mathbf{q}) + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} \Psi_j(\mathbf{q}) = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty). \tag{3}$$

In some instances, Φ can be interpreted as the negative of the physical entropy, so that (3) states that if $\mathbf{q}$ is a smooth solution of (1), then $\Phi \circ \mathbf{q}$ satisfies a conservation law with flux functions $\{\Psi_j \circ \mathbf{q}\}_{j=1}^d$.

For solutions with discontinuities, we impose the *entropy condition* on $\mathbf{q}$ that requires the physical entropy to be non-decreasing:

$$\frac{\partial}{\partial t}\Phi(\mathbf{q}) + \sum_{j=1}^{d} \frac{\partial}{\partial x_j}\Psi_j(\mathbf{q}) \leqslant 0 \tag{4}$$

for every entropy function Φ with entropy fluxes $\{\Psi\}_{j=1}^{d}$; (4) is an inequality in the distributional sense:

$$\int_0^\infty \int_\Omega \left[\Phi(\mathbf{q})\frac{\partial \phi}{\partial t} + \sum_{j=1}^{d} \Psi_j(\mathbf{q})\frac{\partial \phi}{\partial x_j}\right] \mathrm{d}\Omega\, \mathrm{d}t \geqslant 0 \tag{5}$$

for all $\phi \in C_0^\infty(\Omega \times (0,\infty))$, $\phi \geqslant 0$.

In (1), viscous effects are ignored. For the class of phenomena that are modeled by hyperbolic conservation laws, viscous effects are generally small, but they are present and play a role when sharp gradients (such as at shock waves) of the solution are present. An alternate and equivalent means of characterizing the unique, physically relevant solution of (1) is to let $\mathbf{q} = \lim_{\varepsilon\to 0}\mathbf{q}^\varepsilon$ a.e., where $\mathbf{q}^\varepsilon : \Omega \times [0,\infty) \to \mathbb{R}^p$ is the solution of the perturbed equation

$$\frac{\partial \mathbf{q}^\varepsilon}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j}\mathbf{f}_j(\mathbf{q}^\varepsilon) - \varepsilon\Delta\mathbf{q}^\varepsilon = \mathbf{0} \quad \text{in } \Omega\times(0,\infty) \quad \text{and} \quad \mathbf{q}^\varepsilon(\cdot,0) = \mathbf{g} \quad \text{in } \Omega \tag{6}$$

along with boundary conditions. In other words, the entropy solution is the limit of the *viscous solution* as the viscosity goes to zero.

1.2. Spectral viscosity methods

In [25], spectral viscosity (SV) methods were introduced as a scheme to obtain approximate solutions of the periodic Burgers' equation using Fourier spectral basis functions. The theory was further refined and extended in a series of papers [21,20,7,26,27,10,13]. Of particular importance to us are [13,20] in which Legendre polynomials are used. The variational formulation of the Legendre SV method is closest to our finite element formulation.

We present the most basic SV method, which uses Fourier spectral basis functions. Using standard notation from Fourier spectral methods, we define:

$$u_N = P_N u(x,t), \quad P_N u = \sum_{|k|\leqslant N} \hat{u}_k(t)\mathrm{e}^{\mathrm{i}k\pi x}, \quad \hat{u}_k(t) = \frac{1}{2}\int_{-1}^{+1} u(x,t)\mathrm{e}^{-\mathrm{i}k\pi x}\,\mathrm{d}\Omega.$$

We seek u_N such that

$$\frac{\partial u_N}{\partial t} + \frac{\partial}{\partial x}\left(P_N\frac{u_N^2}{2}\right) = \epsilon\frac{\partial}{\partial x}\left(Q_N\frac{\partial u_N}{\partial x}\right).$$

Q_N denotes the spectral viscosity operator defined as a convolution with the viscosity kernel $Q_N(x)$ so that

$$Q_N\frac{\partial u_N}{\partial x} = Q_N(x) * \frac{\partial u_N(x,t)}{\partial x} \quad \text{and} \quad Q_N(x) = \sum_{|k|\leqslant N}\widehat{Q}_k\mathrm{e}^{\mathrm{i}k\pi x}.$$

We choose $0 \leqslant \widehat{Q}_k \leqslant 1$ and $\widehat{Q}_k = 0$ for small $|k|$. It is easy to see the effect of Q_N if we write the diffusion term in Fourier space:

$$\epsilon\frac{\partial}{\partial x}\left(Q_N\frac{\partial u_N}{\partial x}\right) = -\epsilon\sum_{|k|\leqslant N}(k^2\pi^2\widehat{Q}_k\hat{u}_k\mathrm{e}^{\mathrm{i}k\pi x}).$$

Since $\widehat{Q}_k = 0$ for all but large $|k|$, Q_N eliminates the low frequency modes of u_N in the diffusion term. So, we see that the SV diffusion term is a compromise between not adding diffusion, which leads to instability, and adding full diffusion, which limits the convergence rate and smears out discontinuities in the solution. Ideally, one

would like to add diffusion only in the vicinity of a discontinuity. However, the global nature of the basis functions makes an adaptive viscosity kernel difficult to construct.

The SV solution u_N does not converge to the exact solution u at the optimal rate because of the poor convergence of $P_N u$. $P_N u$ is limited to first-order convergence in smooth regions and has O(1) Gibbs oscillations near a discontinuity. Post-processing u_N recovers spectral convergence. The post-processing scheme can be enhanced by knowing the locations of discontinuities, as in [10]. Because of the global nature of spectral basis functions, this edge detection task is a nontrivial one.

2. Hierarchical finite element basis functions

The usual (nodal) basis functions used in finite element methods all have the same frequency. In order to have available multi-frequency basis functions, we use hierarchical basis functions; see [3,4,31–35]. In the elliptic partial differential equation setting, an early analysis of hierarchical basis functions, especially in one dimension, is given in [35]. For two dimensions, see [31]. A good overview of hierarchical basis functions can be found in [33].

First consider a polygonal domain Ω. Let $\mathscr{T}_0$ be a coarse grid approximation of Ω. The nth level triangulation $\mathscr{T}_n$ is obtained by subdividing the elements of $\mathscr{T}_{n-1}$. Let S^N be the space of continuous functions which are polynomials of degree p on the elements of $\mathscr{T}_N$. Let $\mathscr{N}_N \subseteq \overline{\Omega}$ be the nodes of the elements of $\mathscr{T}_N$. The nodal basis functions of S^N are defined by $\phi_i \in S^N$ such that $\phi_i(x_j) = \delta_{ij}$ for all $x_j \in \mathscr{N}_N$. It is well known that $S^N = \text{span}\{\phi_i\}_i$. The use of nodal bases leads to many nice numerical properties, such as sparse mass and stiffness matrices and the local assembly of matrices. However, we cannot use the nodal bases for our purposes because, as we noted earlier, the elements of $\{\phi_i\}_i$ all have the same frequency.

Let $\mathscr{N}_n$ denote the nodes of the n-th level triangulation $\mathscr{T}_n$, $\mathscr{S}^n$ denote the corresponding finite element space, and B^n denote the nodal basis of $\mathscr{S}^n$. The hierarchical basis functions are defined by

$$\psi_{n,i} \in B^n \text{ such that } \psi_{n,i}(x_j) = 0 \quad \forall x_j \in \mathscr{N}_{n-1}.$$

For $0 \leqslant n \leqslant N$, $\{\psi_{n,i}\}_{n,i} \subseteq S^N$ is a linearly independent set with the same dimension as S^N, so $S^N = \text{span}\{\psi_{n,i}\}_{n,i}$. See Fig. 1 for a comparison of the nodal and hierarchical bases for linear elements in one dimension. As can be seen from Fig. 1, $\psi_{n,i}$ is a low frequency function for small n and a high frequency function for large n.

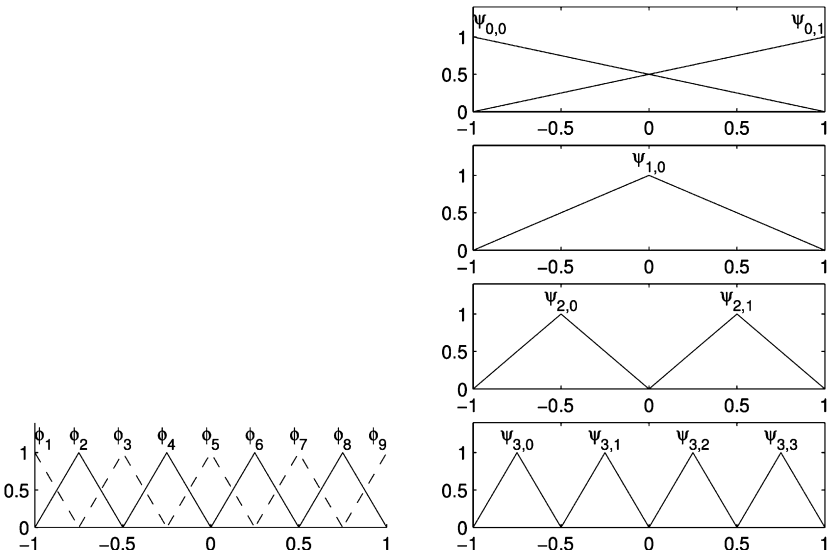

Fig. 1. Left: Nodal linear basis functions spanning a nine-dimensional finite element space. Right: Hierarchical linear basis functions spanning the same nine-dimensional finite element space; top to bottom: the 2 level 0 functions, the single level 1 function, the 2 level 2 functions, and the 4 level 3 function.

The strategy just outlined works for polynomials of any degree. For example, Fig. 2 displays a set of quadratic hierarchical basis functions and an alternate strategy for which linear hierarchical basis functions are used for $n < N$.

In order to determine $\mathscr{T}_{n+1}$ from $\mathscr{T}_n$, we must decide, for a given $T \in \mathscr{T}_n$, how many sub-elements to divide T into. For linear and quadratic rectangular-type elements in $\mathbb{R}^d$, the natural choice is 2^d sub-elements. For cubic elements, the natural choice is 3^d since the vertices of an element will then be a subset of the vertices of its parent. Here, we limit our attention to linear and quadratic basis functions.

For domains with curved boundaries, the situation is more complicated. Let Ω be our potentially complicated domain. One strategy is to use the hierarchical decomposition of some polygonal domain Ω' such that $\Omega \subseteq \Omega'$, as in [16]. Another strategy is to try and impose a hierarchical structure on an unstructured mesh, as in [4]. The hierarchical structure could also be imposed on the mapping of Ω to a polygonal domain.

2.1. Some properties of hierarchical finite element bases

A finite element discretization will result in a system of ordinary differential equations. After choosing an ODE solver to discretize time, we are left with a system of equations, possibly non-linear. Let J^{D} and J^{H} be the nodal and hierarchical Jacobian matrices, respectively, of the nonlinear discrete system. For simple (linear) problems, J could be the mass or stiffness matrix. Due to the small support of all nodal basis functions, J^{D} is much more sparse and easier to assemble than is J^{H}. *It is therefore best to work with the nodal basis as much as possible, then translate to the hierarchical basis only when needed.*

Let $\mathbf{x}^{\mathrm{D}}$ and $\mathbf{x}^{\mathrm{H}}$ be the nodal and hierarchical coefficients, respectively, of $u_N \in S^N$. Let S be the matrix such that

$$\mathbf{x}^{\mathrm{D}} = S\mathbf{x}^{\mathrm{H}}.$$

A simple but tedious calculation gives

$$J^{\mathrm{H}} = S^{\mathrm{T}} J^{\mathrm{D}} S \quad \text{and} \quad \mathbf{R}^{\mathrm{H}} = S^{\mathrm{T}}\mathbf{R}^{\mathrm{D}}.$$

The nonlinear solver requires the solution of

$$(J^{\mathrm{H}})^{-1}\mathbf{R}^{\mathrm{H}} = (S^{\mathrm{T}} J^{\mathrm{D}} S)^{-1} S^{\mathrm{T}}\mathbf{R}^{\mathrm{D}} = (J^{\mathrm{D}} S)^{-1}\mathbf{R}^{\mathrm{D}},$$

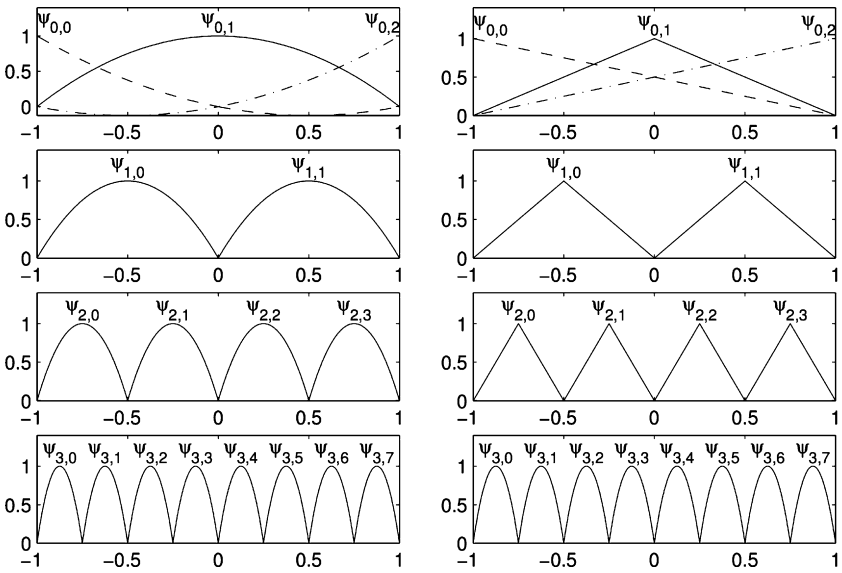

Fig. 2. Left: Quadratic hierarchical basis functions spanning a 17-dimensional, quadratic finite element space. Right: An alternate set of quadratic hierarchical basis functions spanning the same 17-dimensional, quadratic finite element space.

where $\mathbf{R}^{\mathrm{D}}$ and $\mathbf{R}^{\mathrm{H}}$ denote residual vectors in terms of the nodal and hierarchical bases, respectively. Thus, in an iterative linear solver, we need to calculate $J^{\mathrm{D}}S\mathbf{y}$ and perhaps $(J^{\mathrm{D}}S)^{\mathrm{T}}\mathbf{y}$ for some vector $\mathbf{y}$. Compared to J^{D}, S is not sparse, so we do not want to actually construct S. Making the linear solver efficient requires being able to calculate $S\mathbf{y}$, $S^{\mathrm{T}}\mathbf{y}$, and $S^{-1}\mathbf{y}$ quickly. Algorithms for this purpose are given in the next subsection.

2.1.1. Efficient calculation of transformations between bases

We now present algorithms for calculating $S\mathbf{y}$, $S^{\mathrm{T}}\mathbf{y}$, $S^{-1}\mathbf{y}$, and $S^{-\mathrm{T}}\mathbf{y}$. Here, i and j represent global node numbers.

Let $\mathbf{y}$ denote a vector representing a function $u_N \in S^N$ in the hierarchical basis. Then, $S\mathbf{y}$ represents u_N in the nodal basis. To calculate $S\mathbf{y}$, we evaluate u_N at the nodes; $\mathbf{y}$ is overwritten by $S\mathbf{y}$.

```
for level n = 0 to N − 1
  for element T ∈ 𝒯_n
    for x_i ∈ (𝒩_{n+1} \ 𝒩_n)
      for nodal basis function φ_j ∈ B^n with support in T
        y_i = y_i + y_j φ_j(x_i)
      end for φ_j
    end for x_i
  end for T
end for n
```

Now let $\mathbf{y}$ be a vector representing a function $u_N \in S^N$ in the standard nodal basis. Then, $S^{-1}\mathbf{y}$ represents u_N in the hierarchical basis. To calculate $S^{-1}\mathbf{y}$, at each level, we must use the coarser mesh to subtract away the global behavior of u_N; $\mathbf{y}$ is overwritten by $S^{-1}\mathbf{y}$.

```
for level n = N − 1 to 0
  for element T ∈ 𝒯_n
    for x_i ∈ (𝒩_{n+1} \ 𝒩_n)
      for nodal basis function φ_j ∈ B^n with support in T
        y_i = y_i − y_j φ_j(x_i)
      end for φ_j
    end for x_i
  end for T
end for n
```

Unlike S and S^{-1}, S^{T} does not have a clear geometric interpretation. As noted in [31], S can be represented as $S = S_0 S_1 \cdots S_{N-1} S_N$, where S_n represents the matrix multiplication in the outer loop of the $S\mathbf{y}$ algorithm. This gives us $S^{\mathrm{T}} = S_N^{\mathrm{T}} S_{N-1}^{\mathrm{T}} \cdots S_1^{\mathrm{T}} S_0^{\mathrm{T}}$.

```
for level n = N − 1 to 0
  for element T ∈ 𝒯_n
  for x_i ∈ (𝒩_{n+1} \ 𝒩_n) ⋂ T_c
      for nodal basis function φ_j ∈ B^n with support in T
        y_j = y_j + y_i φ_j(x_i)
      end for φ_j
    end for x_i
  end for T
end for n
```

Finally, we have $S^{-\mathrm{T}} = S_0^{-\mathrm{T}} S_1^{-\mathrm{T}} \cdots S_{N-1}^{-\mathrm{T}} S_N^{-\mathrm{T}}$.

```
for level n = 0 to N − 1
```

```
for element T ∈ 𝒯_n
 for x_i ∈ (𝒩_{n+1} \ 𝒩_n) ⋂ T_c
    for nodal basis function φ_j ∈ B^n with support in T
      y_j = y_j − y_i φ_j(x_i)
    end for φ_j
  end for x_i
 end for T
end for n
```

2.1.2. Condition numbers of standard matrices in nodal and hierarchical bases

Let h_N be the maximum mesh size of the finest triangulation $\mathscr{T}_N$. For a matrix J, let $\kappa(J)$ be the condition number. Let r be defined by $\kappa = \mathrm{O}(h_N^{-r})$. Let M and K be the mass and stiffness matrices, respectively, in the standard nodal basis. It is well know $\kappa(M) = \mathrm{O}(1)$ and $\kappa(K) = \mathrm{O}(h_N^{-2})$. In the hierarchical basis, the condition number of the stiffness matrix $K^{\mathrm{H}} = S^{\mathrm{T}}KS$ is smaller, while the condition number of the mass matrix $M^{\mathrm{H}} = S^{\mathrm{T}}MS$ is larger.

Let us first consider some numerical experiments. We calculate the condition numbers of J, JS, and $S^{\mathrm{T}}JS$ for $J = K, M$ and for first and second degree piecewise polynomial bases. The results are given in Tables 1 and 2 and Fig. 3. The numerical evidence indicates that $\kappa(M^{\mathrm{H}}) = \kappa(K^{\mathrm{H}}) = \mathrm{O}(h_N^{-1})$.

We can prove that in one dimension, we cannot do better than $\kappa(M^{\mathrm{H}}) = \kappa(K^{\mathrm{H}}) = \mathrm{O}(h_N^{-1})$ for linear and quadratic basis functions. Let $H^1(U)$ and $H_0^1(U)$ be the standard notations for Sobolev spaces with inner product, norm, and semi-norm $(\cdot,\cdot)_1$, $\|\cdot\|_1$, and $|\cdot|_1$, respectively. Let $p \in \{1,2\}$ be the degree of the polynomial. For

$$v_N = \sum_{i=0}^{p+1} \beta_{0,i}\psi_{0,i} + \sum_{k=1}^{N}\sum_{i=0}^{2^{k+p-2}-1} \beta_{k,i}\psi_{k,i} \in S_N(U),$$

let us define the norm

$$\|v_N\|^2 = \sum_{i=0}^{p+1} |\beta_{0,i}|^2 + \sum_{k=1}^{N}\sum_{i=0}^{2^{k+p-2}-1} |\beta_{k,i}|^2.$$

For a lower bound of the condition number of the mass matrix, we will need to determine how small and how large $\frac{\|\cdot\|_0^2}{\|\cdot\|^2}$ can be. For $p = 1$, we have that

$$\|\psi_{1,0}\|^2 = 1 \quad \text{and} \quad \|\psi_{1,0}\|_0^2 = \frac{|U|}{3}.$$

For $p = 2$,

$$\|\psi_{0,1}\|^2 = 1 \quad \text{and} \quad \|\psi_{0,1}\|_0^2 = C|U|,$$

where C can be $\frac{8}{15}$ or $\frac{1}{3}$ depending on which quadratic hierarchical representation we use. We also have that

Table 1
Comparison of stiffness matrix condition numbers in one dimension

N	Linear bases						Quadratic bases					
	K		KS		$S^{\mathrm{T}}KS$		K		KS		$S^{\mathrm{T}}KS$	
	κ	r	κ	r	κ	r	κ	r	κ	r	κ	r
3	5.8e+0	–	3.2e+0	–	2.0e+0	–	3.3e+1	–	1.3e+1	–	8.0e+0	–
4	2.5e+1	2.1	9.4e+0	1.6	4.0e+0	1.0	1.4e+2	2.1	3.7e+1	1.6	1.6e+1	1.0
5	1.0e+2	2.0	2.7e+1	1.5	8.0e+0	1.0	5.5e+2	2.0	1.1e+2	1.5	3.2e+1	1.0
6	4.1e+2	2.0	7.7e+1	1.5	1.6e+1	1.0	2.2e+3	2.0	3.0e+2	1.5	6.4e+1	1.0
7	1.7e+3	2.0	2.2e+2	1.5	3.2e+1	1.0	8.9e+3	2.0	8.6e+2	1.5	1.3e+2	1.0
8	6.6e+3	2.0	6.2e+2	1.5	6.4e+1	1.0	3.5e+4	2.0	2.4e+3	1.5	2.6e+2	1.0
9	2.7e+4	2.0	1.7e+3	1.5	1.3e+2	1.0	1.4e+5	2.0	6.8e+3	1.5	5.1e+2	1.0

Table 2
Comparison of mass matrix condition numbers in one dimension

N	Linear bases						Quadratic bases					
	M		MS		$S^{\mathrm{T}}MS$		M		MS		$S^{T}KS$	
	κ	r	κ	r	κ	r	κ	r	κ	r	κ	r
3	3.9e+0	–	1.1e+1	–	3.0e+1	–	5.4e+0	–	1.7e+1	–	5.6e+1	–
4	3.9e+0	0.0	1.6e+1	0.6	6.8e+1	1.2	5.4e+0	0.0	2.5e+1	0.6	1.3e+2	1.2
5	4.0e+0	0.0	2.4e+1	0.6	1.6e+2	1.2	5.4e+0	0.0	3.7e+1	0.6	3.0e+2	1.2
6	4.0e+0	0.0	3.6e+1	0.6	3.4e+2	1.1	5.4e+0	0.0	5.5e+1	0.6	6.8e+2	1.1
7	4.0e+0	0.0	5.4e+1	0.6	7.5e+2	1.1	5.4e+0	0.0	8.1e+1	0.6	1.5e+3	1.1
8	4.0e+0	0.0	7.9e+1	0.6	1.6e+3	1.1	5.4e+0	0.0	1.2e+2	0.6	3.2e+3	1.1
9	4.0e+0	0.0	1.2e+2	0.5	3.5e+3	1.1	5.4e+0	0.0	1.7e+2	0.5	6.7e+3	1.1

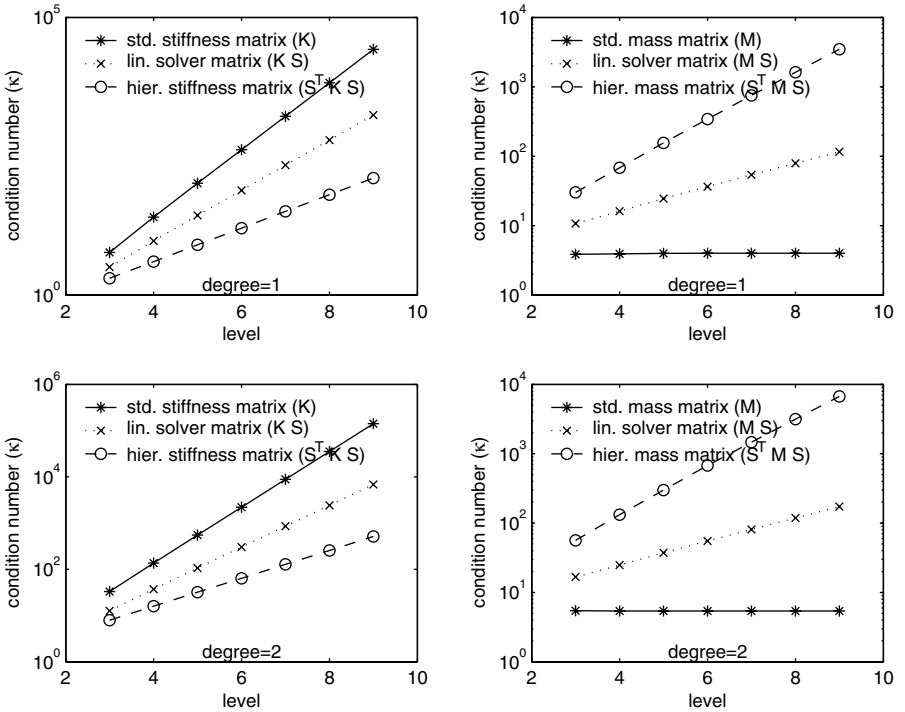

Fig. 3. Condition numbers of stiffness and mass matrices.

$$\|\psi_{N,0}\|^2 = 1 \quad \text{and} \quad \|\psi_{N,0}\|_0^2 = \begin{cases} \frac{|U|}{3}2^{-N+1}, & p = 1, \\ \frac{16}{15}\frac{|U|}{2}2^{-N}, & p = 2. \end{cases}$$

Therefore, for all N, there exist $v_N,\ u_N \in S^N(U) \cap H_0^1(U)$ such that

$$\frac{\|v_N\|_0^2}{\|v_N\|^2} = C_1 \quad \text{and} \quad \frac{\|u_N\|_0^2}{\|u_N\|^2} = C_2|U|2^{-N}.$$

For a lower bound of the condition number of the stiffness matrix, we will need to determine how small and how large $\frac{|\cdot|_1^2}{\|\cdot\|^2}$ can be. For $p = 1$, we have that

$$\|\psi_{1,0}\|^2 = 1 \quad \text{and} \quad |\psi_{1,0}|_1^2 = \frac{4}{|U|}.$$

For $p = 2$, we have that

$$\|\psi_{0,1}\|^2 = 1 \quad \text{and} \quad |\psi_{0,1}|_1^2 = \frac{C}{|U|},$$

where C can be $\frac{16}{3}$ or 4 depending on which quadratic hierarchical representation we use. We then have

$$\|\psi_{N,0}\|^2 = 1 \quad \text{and} \quad |\psi_{N,0}|_1^2 = \begin{cases} \frac{4}{|U|} 2^{N-1}, & p = 1, \\ \frac{16}{3} \frac{1}{|U|} 2^N, & p = 2. \end{cases}$$

Therefore, for all N, there exist $v_N, u_N \in S^N(U) \cap H_0^1(U)$ such that

$$\frac{|v_N|_1^2}{\|v_N\|^2} = C_3 \quad \text{and} \quad \frac{|u_N|_1^2}{\|u_N\|^2} = \frac{C_4}{|U|} 2^N.$$

From now on, C will denote a positive constant, not necessarily the same from line to line. Let $\lambda_{N,\min}$ and $\lambda_{N,\max}$ be the minimum and maximum eigenvalues, respectively, of a matrix.

Lemma 1. *In the hierarchical basis, the condition number of the mass matrix grows at least exponentially with the number of levels.*

Proof. Using the Rayleigh quotient, we have that

$$\lambda_{N,\min}(M^{\mathrm{H}}) = \inf_x \frac{x^{\mathrm{T}} M^{\mathrm{H}} x}{x^{\mathrm{T}} x} = \inf_{v_N \in S_N} \frac{\|v_N\|_0^2}{\|v_N\|^2} \leqslant C 2^{-N}$$

and

$$\lambda_{N,\max}(M^{\mathrm{H}}) = \sup_x \frac{x^{\mathrm{T}} M^{\mathrm{H}} x}{x^{\mathrm{T}} x} = \sup_{v_N \in S_N} \frac{\|v_N\|_0^2}{\|v_N\|^2} \geqslant C$$

so that

$$\kappa_N(M^{\mathrm{H}}) = \frac{\lambda_{N,\max}(M^{\mathrm{H}})}{\lambda_{N,\min}(M^{\mathrm{H}})} \geqslant C 2^N. \quad \square$$

Lemma 2. *In the hierarchical basis, the condition number of the stiffness matrix grows at least exponentially with the number of levels.*

Proof. Again, using the Rayleigh quotient, we have that

$$\lambda_{N,\min}(K^{\mathrm{H}}) = \inf_x \frac{x^{\mathrm{T}} K^{\mathrm{H}} x}{x^{\mathrm{T}} x} = \inf_{v_N \in S_N} \frac{|v_N|_1^2}{\|v_N\|^2} \leqslant C$$

and

$$\lambda_{N,\max}(K^{\mathrm{H}}) = \sup_x \frac{x^{\mathrm{T}} K^{\mathrm{H}} x}{x^{\mathrm{T}} x} = \sup_{v_N \in S_N} \frac{|v_N|_1^2}{\|v_N\|^2} \geqslant C 2^N$$

so that

$$\kappa_N(K^{\mathrm{H}}) = \frac{\lambda_{N,\max}(K^{\mathrm{H}})}{\lambda_{N,\ \min}(K^{\mathrm{H}})} \geqslant C 2^N. \quad \square$$

3. The finite element multi-resolution viscosity method

Assume we have a hierarchical sequence of partitions $\{\mathscr{T}_n\}_{n=0}^N$ of $U \subseteq \mathbb{R}^d$, which is an open bounded set. Let $\mathbf{S}^N$ be the space of continuous vector-valued functions whose components are in S^N. We seek an approximate solution to the hyperbolic conservation law (1) with appropriate boundary conditions.

Our finite element discretization of (1) is based on the weak formulation (2) and is defined as follows: seek $\mathbf{q}^N \in \mathbf{S}^N$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_U \mathbf{q}^N \cdot \mathbf{v}\,\mathrm{d}x + \int_U \sum_{j=1}^{d} \frac{\partial}{\partial x_j}\mathbf{f}_j(\mathbf{q}^N)\cdot\mathbf{v}\,\mathrm{d}x + \epsilon_N \sum_{i=1}^{p}\sum_{j,k=1}^{d}\int_U \frac{\partial}{\partial x_j}(Q_N^{j,k}q_i^N)\frac{\partial v_i}{\partial x_k}\,\mathrm{d}x - \epsilon_N \sum_{i=1}^{p}\sum_{j,k=1}^{d}\int_{\partial U}\frac{\partial}{\partial x_j}(Q_N^{j,k}q_i^N)v_i\widehat{n}_k\,\mathrm{d}s = 0 \quad \forall \mathbf{v}\in\mathbf{S}^N, \tag{7}$$

where $\hat{\mathbf{n}}$ denotes the unit normal to the boundary ∂U of U. As in the SV method, $Q_N^{j,k}$ reduces or eliminates the low frequency modes of a function:

$$Q_N^{j,k} : S^N \to S^N \quad \sum_{n=0}^{N}\sum_i \beta_{n,i}\psi_{n,i} \mapsto \sum_{n=0}^{N}\sum_i Q_{N;n,i}^{j,k}\beta_{n,i}\psi_{n,i}, \tag{8}$$

where

$$0 \leqslant Q_{N;n,i}^{j,k} \leqslant 1 \quad \text{and} \quad Q_{N;n,i}^{j,k} = \begin{cases} 0 & \text{for small } n \ (n \text{ near } 0), \\ 1 & \text{for large } n \ (n \text{ near } N). \end{cases}$$

To account for the boundary conditions that accompany (1), a subspace of S_N might need to be used (for essential boundary conditions) or the domain of integration in the boundary integral in (7) might not be all of ∂U (for natural boundary conditions).

One recognizes that (7) is a weak formulation of

$$\frac{\partial \mathbf{q}}{\partial t} + \sum_{j=1}^{d}\frac{\partial}{\partial x_j}\mathbf{f}_j(\mathbf{q}) - \epsilon_N \sum_{j,k=1}^{d}\frac{\partial^2}{\partial x_j\,\partial x_k}(Q_N^{j,k}\mathbf{q}) = \mathbf{0}, \tag{9}$$

where $[Q_N^{j,k}\mathbf{q}]_i = Q_N^{j,k}q_i$ for $1 \leqslant i \leqslant p$. The dependence of $Q_N^{j,k}$ on j and k allows for the possibility of directional bias in the diffusion, which could reduce crosswind diffusion. As in the streamline diffusion method, this would probably require the use of entropy variables. We can simplify our formulation by using an isotropic diffusion term Q_N such that

$$Q_{N;n,i}^{j,k} = Q_{N;n,i}\delta_{j,k}. \tag{10}$$

Then, (7) and (9), respectively, reduce to

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_U \mathbf{q}\cdot\mathbf{v}\,\mathrm{d}x + \int_U \sum_{j=1}^{d}\frac{\partial}{\partial x_j}\mathbf{f}_j(\mathbf{q})\cdot\mathbf{v}\,\mathrm{d}x + \epsilon_N\int_U \nabla(Q_N\mathbf{q}) : \nabla\mathbf{v}\,\mathrm{d}x - \epsilon_N\int_{\partial U}\frac{\partial}{\partial n}(Q_N\mathbf{q})\cdot\mathbf{v}\,\mathrm{d}s = 0 \quad \forall\mathbf{v}\in\mathbf{S}^N \tag{11}$$

and

$$\frac{\partial\mathbf{q}}{\partial t} + \sum_{j=1}^{d}\frac{\partial}{\partial x_j}\mathbf{f}_j(\mathbf{q}) - \epsilon_N\Delta(Q_N\mathbf{q}) = \mathbf{0}. \tag{12}$$

Once one chooses a time discretization technique, one is left with a nonlinear system of equations to solve. The Jacobian of that system has the form

$$J^{\mathrm{H}} = \widetilde{J}^{\mathrm{H}} + K^{\mathrm{H}}Q,$$

where $\widetilde{J}^{\mathrm{H}}$ is the Jacobian of the time dependent and flux terms, and Q is a diagonal matrix whose nonzero elements are $\{Q_{N;n,i}^{j,k}\}$. For ease of notation, we ignore the boundary term. Translating to the nodal basis, we obtain

$$\begin{aligned} &J^{\mathrm{H}} = S^{\mathrm{T}}(\widetilde{J}^{\mathrm{D}}S + K^{\mathrm{H}}SQ), \quad \mathbf{R}^{\mathrm{H}} = S^{\mathrm{T}}\mathbf{R}^{\mathrm{D}}, \\ &(J^{\mathrm{H}})^{-1}\mathbf{R}^{\mathrm{H}} = (\widetilde{J}^{\mathrm{D}}S + K^{\mathrm{H}}SQ)^{-1}\mathbf{R}^{\mathrm{D}}. \end{aligned} \tag{13}$$

An iterative linear solver would then require, for a given vector $\mathbf{x}$, the calculation of

$$(\widetilde{J}^{\mathrm{D}}S + K^{\mathrm{H}}SQ)\mathbf{x} \quad \text{and} \quad (\widetilde{J}^{\mathrm{D}}S + K^{\mathrm{H}}SQ)^{\mathrm{T}}\mathbf{x}.$$

In the SV method, there is only one function, having global support, at a given frequency. In our finite element formulation, there are many functions, each having local support, at a given frequency. Hierarchical finite element bases offer two advantages: diffusion can be added locally and edge detection is trivial.

3.1. Adaptive diffusion

For large values of n, the hierarchical basis function $\psi_{n,i}$ has local support, so $Q^{i,k}_{N;n,i}$ only has a local effect. We can therefore add more diffusion near a discontinuity and less diffusion in the smooth regions. This should improve the accuracy of the method. As we are about to see, the size of $|\beta_{n,i}|$ can be used to determine if the support of $\psi_{n,i}$ resides in a smooth region or is near a discontinuity.

3.2. Edge detection

Using a hierarchical finite element basis, edge detection is a trivial task. Near a discontinuity, the high frequency hierarchical coefficients are order one. In a smooth region, they shrink exponentially. Fig. 4 displays the coefficients of a hierarchical decomposition of a piecewise smooth function with a discontinuity. We can easily determine the location of discontinuities by looking at the magnitude of the high frequency coefficients.

3.2.1. Edge detection for piecewise linear polynomials

Let us prove the behavior of the hierarchical coefficients for linear basis functions. As we can see in Fig. 5, the coefficient $\beta_{k+1,j}$ in the hierarchical finite element expansion can be calculated from the value of the function u at $x_{k+1,j}$ and the node points of the parent cell. For a uniform partition, the cell size at level k is $\Delta x_k = |U|2^{-k}$ and so $x_{k+1,j} - x_{k,i} = x_{k,i+1} - x_{k+1,j} = \Delta x_{k+1}$ and

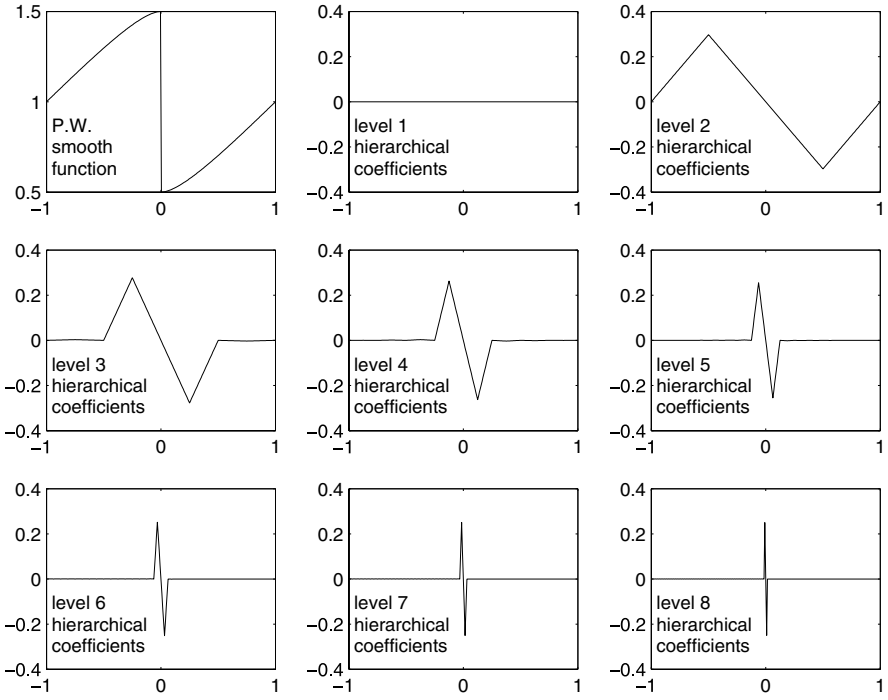

Fig. 4. Hierarchical decomposition of a piecewise smooth function with a discontinuity.

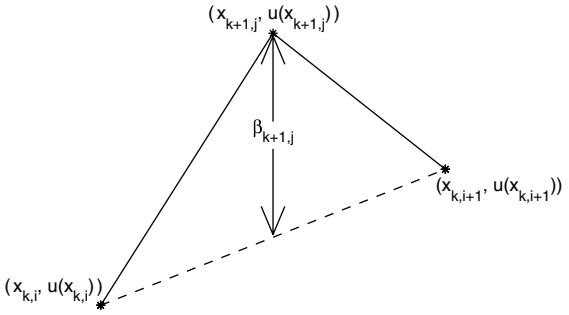

Fig. 5. Linear hierarchical coefficient from function values.

$$\beta_{k+1,j} = u(x_{k+1,j}) - \frac{u(x_{k,i+1}) + u(x_{k,i})}{2} = \frac{u(x_{k+1,j}) - u(x_{k,i})}{2} - \frac{u(x_{k,i+1}) - u(x_{k+1,j})}{2}. \tag{14}$$

Let $T_{k,i} = (x_{k,i}, x_{k,i+1})$.

If u is *discontinuous*, we assume that u has a discontinuity in $T_{k,i}$. At least one of the two terms in (14) has a relatively large value, and so $|\beta_{k+1,j}|$ will be the same order as the jump of u independent of k. If u is *continuously differentiable*, we now assume that $u \in C^1(\overline{T_{k,i}})$. Then, (14) can be written as

$$\beta_{k+1,j} = \frac{|U|}{4} 2^{-k} \left(\frac{u(x_{k+1,j}) - u(x_{k,i})}{x_{k+1,j} - x_{k,i}} - \frac{u(x_{k,i+1}) - u(x_{k+1,j})}{x_{k,i+1} - x_{k+1,j}} \right).$$

By the mean value theorem,

$$\beta_{k+1,j} = \frac{|U|}{4} 2^{-k} [u'(\tilde{x}_1) - u'(\tilde{x}_2)] \tag{15}$$

for some $\tilde{x}_1, \tilde{x}_2 \in T_{k,i}$. Therefore, $\beta_{k+1,j}$ is of order 2^{-k}:

$$|\beta_{k+1,j}| \leqslant |U| \|u'\|_{L^\infty(T_{k,i})} 2^{-k-1}.$$

If u is *twice continuously differentiable*, we assume that $u \in C^2(\overline{T_{k,i}})$. By the mean value theorem, (15) can be written as

$$\beta_{k+1,j} = -\frac{|U|}{4} 2^{-k} (\tilde{x}_2 - \tilde{x}_1) u''(\tilde{x})$$

for some $\tilde{x} \in T_{k,i}$. Therefore, $\beta_{k+1,j}$ is of order 4^{-k}:

$$|\beta_{k+1,j}| = \frac{|U|}{4} |\tilde{x}_2 - \tilde{x}_1| |u''(\tilde{x})| 2^{-k} \leqslant |U|^2 \|u''\|_{L^\infty(T_{k,i})} 4^{-k-1}.$$

3.2.2. Edge detection for piecewise quadratic polynomials

We can produce similar results for quadratic basis functions by using the relation

$$\beta_{k+1,j} = u(x_{k+1,j}) - \left(\frac{3}{8} u(x_{k,i-1}) + \frac{3}{4} u(x_{k,i}) - \frac{1}{8} u(x_{k,i+1}) \right) \tag{16}$$

$$= \frac{3}{8} [u(x_{k+1,j}) - u(x_{k,i-1})] - \frac{6}{8} [u(x_{k,i}) - u(x_{k+1,j})] + \frac{1}{8} [u(x_{k,i+1}) - u(x_{k+1,j})], \tag{17}$$

as in Fig. 6. We can further show that the hierarchical coefficients of a smooth function shrink more quickly in the quadratic case.

If u is *three times continuously differentiable*, we can use an interpolation argument from [30]. Let

$$F(t) := u(t) - u_k(t) - RL(t),$$

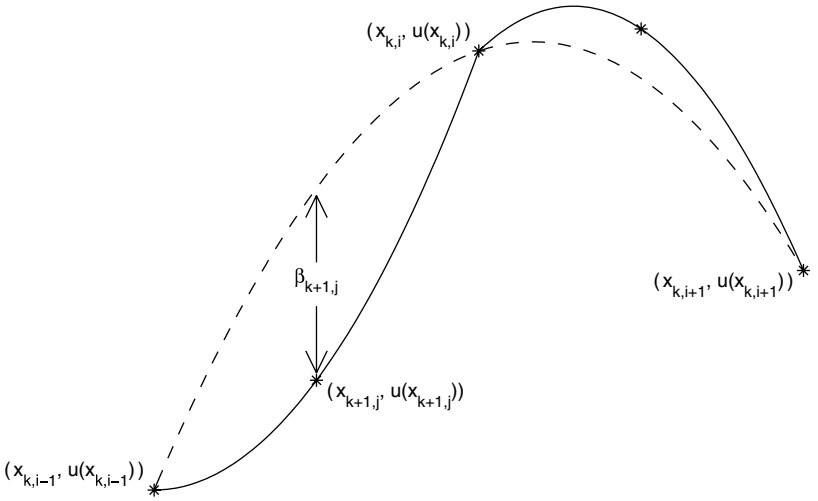

Fig. 6. Quadratic hierarchical coefficient from function values.

where $u_k \in S^k$ is the interpolant of u,

$$R := \frac{u(x_{k+1,j}) - u_k(x_{k+1,j})}{(x_{k+1,j} - x_{k,i-1})(x_{k+1,j} - x_{k,i})(x_{k+1,j} - x_{k,i+1})} = \frac{8}{3|U|^3} 8^k \beta_{k+1,j}$$

and

$$L(t) := (t - x_{k,i-1})(t - x_{k,i})(t - x_{k,i+1}).$$

Since $u_k(x) = u(x)$ for $x \in \{x_{k,i}, x_{k,i\pm1}\}$, we have $F(x_{k,i-1}) = F(x_{k,i}) = F(x_{k,i+1}) = 0$. From the construction of $F(t)$, we also have $F(x_{k+1,j}) = 0$. By repeated applications of Rolle's theorem, $F'''(t)$ has at least one zero $\eta \in (x_{k,i-1}, x_{k,i+1})$. u_k is a quadratic polynomial on $(x_{k,i-1}, x_{k,i+1})$, so $u_k'''(t) = 0$. Thus,

$$0 = F'''(\eta) = u'''(\eta) - u_k'''(\eta) - RL'''(\eta) = u'''(\eta) - 6R = u'''(\eta) - \frac{16}{|U|} 8^k \beta_{k+1,j}.$$

Therefore, $\beta_{k+1,j} = \frac{|U|}{16} 8^{-k} u'''(\eta)$ and

$$|\beta_{k+1,i}| \leqslant \frac{|U|^3}{2} 8^{-k-1} \|u'''\|_{L^\infty(x_{k,i-1}, x_{k,i+1})}. \tag{18}$$

4. Numerical results

In this section, we provide a few illustrative but preliminary computational experiments using hierarchical finite element methods for one- and two-dimensional hyperbolic conservations laws.

4.1. Model problems

We use four model problems to illustrate the implementation of the hierarchical finite element formulation to hyperbolic conservation laws.

4.1.1. Burgers' equation with a long-time steady state solution

The quintessential scalar conservation law in one dimension is Burgers' equation; we look at two different problems involving this equation. First, we consider Burgers' equation with Dirichlet boundary conditions:

$$\begin{cases} \frac{\partial q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q^2}{2}\right) = 0 \quad \text{on } (-1,+1)\times(0,\infty), \\ q(\pm 1, t) = \mp 1 \quad \text{for all } t \in (0,\infty), \\ q(x,0) = -\cos\left[\frac{\pi(x-1)}{2}\right]. \end{cases}$$

As $t \to \infty$,

$$q(x,t) \to \begin{cases} +1, & x < 0, \\ -1, & x > 0. \end{cases}$$

4.1.2. Periodic Burgers' equation problem

For the second model problem, we consider the periodic Burgers' equation:

$$\begin{cases} \frac{\partial q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q^2}{2}\right) = 0 \quad \text{on } (-1,+1)\times(0,T), \\ q(-1,t) = q(+1,t) \quad \text{for all } t \in (0,T), \\ q(x,0) = 1 + \frac{1}{2}\sin(\pi x). \end{cases} \tag{19}$$

It is shown in [9,17] that the entropy solution is given by

$$q(x,t) = \frac{x-y}{t},$$

where $y = y(x,t)$ minimizes

$$G(x,y,t) = \frac{t}{2}\left(\frac{x-y}{t}\right)^2 + \int q(y,0)\,\mathrm{d}y.$$

If x is not the point of discontinuity, then G will have only one global minimum, although it can have several local minima. The simplest strategy of solving for q is, for each x, to find a point y_0 near the global minimum of $G(x,y,t)$, then use any nonlinear solver to find the solution y of

$$-\frac{1}{t}(x-y) + q(y,0) = 0,$$

then set $q(x,t) = \frac{x-y}{t}$. See Fig. 7 for solutions at $t = 0.25$ and $t = 1$.

4.2. The Euler equations

An important example of a system of conservation laws is the Euler equations:

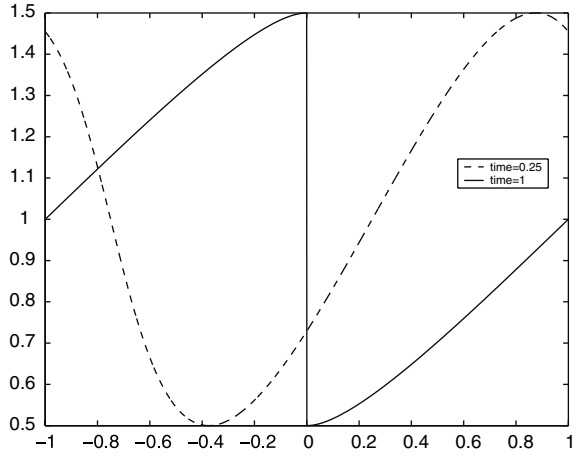

Fig. 7. Exact solutions of the periodic Burgers' equation at two different times.

$$\begin{cases} \frac{\partial \rho}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j}(\rho u_j) = 0, \\ \frac{\partial}{\partial t}(\rho u_i) + \sum_{j=1}^{d} \frac{\partial}{\partial x_j}(\rho u_i u_j) + \frac{\partial P}{\partial x_i} = 0 \quad \text{for } 1 \leqslant i \leqslant d, \\ \frac{\partial}{\partial t}(\rho e) + \sum_{j=1}^{d} \frac{\partial}{\partial x_j}(\rho e u_j + P u_j) = 0, \end{cases}$$

where ρ, $\mathbf{u} \in \mathbb{R}^d$, e, and P are the density, velocity, specific (per unit mass) total energy, and pressure, respectively, of the fluid. $\rho\mathbf{u}$ is the momentum per unit volume and ρe is the total energy per unit volume of the fluid. The Euler equations are a statement of conservation of mass, momentum, and energy.

The Euler equations in one space dimension consist of three equations with four unknowns. An equation of state is required to relate P, ρ, and the specific internal energy $\varepsilon = e - \frac{|\mathbf{u}|^2}{2}$. For a polytropic ideal gas, $\rho\varepsilon = \frac{P}{\gamma - 1}$, where γ is the ratio of specific heats. The value of γ for a diatomic gas, such as air (nitrogen and oxygen), is $\gamma = 1.4$. Thus, the equation of state we use is given by

$$P = (\gamma - 1)\left(\rho e - \rho \frac{|\mathbf{u}|^2}{2}\right).$$

Two other quantities will be used in our study of the Euler equations. The speed of sound in the fluid a and the Mach number $M = |\mathbf{u}|/a$. With the assumptions we have already made about the nature of the fluid, $a = \sqrt{\gamma P/\rho}$. The above definitions are taken from [1,11,18].

4.2.1. The shock tube problem

The third model problem we consider is the shock tube problem. Consider a tube, closed at both ends, which is divided into two sections. On the left is a gas at rest under high pressure; on the right is a gas under low pressure. At time $t = 0$, the divider is removed, and the gases in the two regions are allowed to move. If the properties of the gases in the two regions are initially constant, then this is a Riemann problem, and so at most four distinct regions of flows with constant properties should form. Separating the four regions are either an expansion wave, a contact discontinuity, or a shock discontinuity. In this particular case, all three types of separators are present.

Let us label the four regions as in Fig. 8. Let (ρ_i, u_i, P_i) be the variables in the ith region. (ρ_4, u_4, P_4) and (ρ_1, u_1, P_1) are know from the initial data. In [1], it is shown how the initial data can be used to calculate (ρ, u, P) in the entire domain. The contact discontinuity separates the two gases and moves with a velocity u_p. The shock moves to the right with speed W. $\frac{P_2}{P_1}$ can be calculated from the implicit equation

$$\frac{P_4}{P_1} = \frac{P_2}{P_1}\left\{1 - \frac{(\gamma_4 - 1)\left(\frac{a_1}{a_4}\right)\left(\frac{P_2}{P_1} - 1\right)}{\sqrt{2\gamma_1\left[2\gamma_1 + (\gamma_1 + 1)\left(\frac{P_2}{P_1} - 1\right)\right]}}\right\}^{\frac{-2\gamma_4}{\gamma_4 - 1}}.$$

We can now easily calculate the following:

$$u_3 = u_p, \quad u_2 = u_p, \quad P_3 = P_2, \quad P_2 = P_1\frac{P_2}{P_1},$$

$$\rho_3 = \rho_4\left(\frac{P_3}{P_4}\right)^{\frac{1}{\gamma_4}}, \quad \rho_2 = \rho_1\frac{1 + \left(\frac{\gamma_1+1}{\gamma_1-1}\right)\left(\frac{P_2}{P_1}\right)}{\frac{\gamma_1+1}{\gamma_1-1} + \frac{P_2}{P_1}},$$

$$u_p = \frac{a_1}{\gamma_1}\left(\frac{P_2}{P_1} - 1\right)\left(\frac{\frac{2\gamma_1}{\gamma_1+1}}{\frac{P_2}{P_1} + \frac{\gamma_1-1}{\gamma_1+1}}\right)^{\frac{1}{2}}, \quad W = a_1\sqrt{\left(\frac{\gamma_1 + 1}{2\gamma_1}\right)\left(\frac{P_2}{P_1} - 1\right) + 1}.$$

The left and right ends of the expansion wave travel with velocities $-a_4$ and $u_3 - a_3$, respectively. Inside the expansion wave,

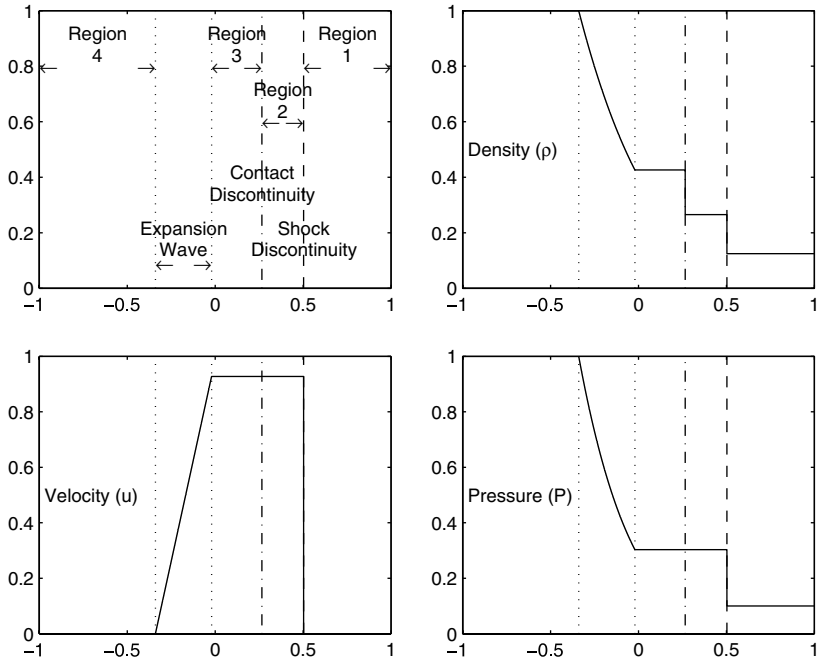

Fig. 8. Exact solution of the shock tube problem at $t = 0.287$.

$$u = \left(\frac{2}{\gamma_4 + 1}\right)\left(a_4 + \frac{x}{t}\right), \quad \rho = \rho_4\left[1 - \left(\frac{\gamma_4 - 1}{2}\right)\left(\frac{u}{a_4}\right)\right]^{\frac{2}{\gamma_4 - 1}},$$
$$P = P_4\left[1 - \left(\frac{\gamma_4 - 1}{2}\right)\left(\frac{u}{a_4}\right)\right]^{\frac{2\gamma_4}{\gamma_4 - 1}}.$$

Fig. 8 gives the exact solution of the shock tube problem at time $t = 0.287$ for the initial and boundary conditions:

$$(\rho, u, P)(x, 0) = \begin{cases} (1, 0, 1), & x < 0, \\ (0.125, 0, 0.1), & x > 0, \end{cases}$$
$$(\rho, u, P)(-1, t) = (1, 0, 1),$$
$$(\rho, u, P)(+1, t) = (0.125, 0, 0.1).$$

4.2.2. Supersonic flow over a wedge

The last model problem we study is that of a uniform, horizontal flow moving with Mach number $M_1 > 1$ that hits a wedge at an angle θ with the flow. As long as θ is not too large, a shock at an angleβ will form which is attached to the wedge. The flow will be uniform on either side of the shock.

Let (ρ_i, u_i, P_i) denote the constant values on either side of the shock with $i = 1, 2$ denoting the left- and right-hand sides of the shock, respectively. In [1], it is shown that (ρ_1, u_1, P_1) and θ determines (ρ_2, u_2, P_2) and β. β can be determined from the implicit relation

$$\tan\theta = 2(\cot\beta)\left\{\frac{M_1^2 \sin^2\beta - 1}{M_1^2[\gamma + \cos(2\beta)] + 2}\right\}. \tag{20}$$

If θ is too large, then (20) does not have any real solutions. Physically, this corresponds to the shock forming in front of the wedge and becoming a detached shock. Otherwise, there can be two solutions of (20). If there is no additional downstream pressure, then the smaller value of β is the physically correct one. (ρ_2, u_2, P_2) are determined by γ and the normal component M_n of the Mach number relative to the shock:

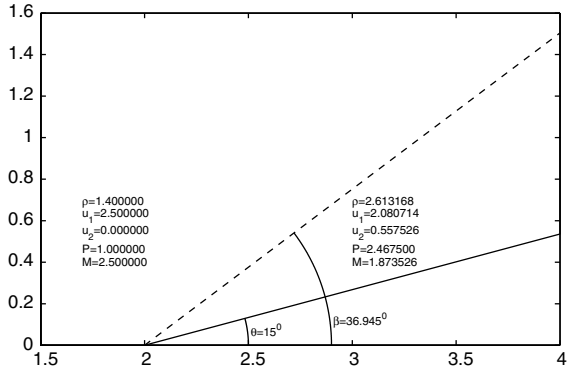

Fig. 9. Exact solution of supersonic flow over a wedge.

$$M_{n_1} = M_1 \sin\beta, \quad M_{n_2} = \sqrt{\frac{M_{n_1}^2 + \frac{2}{\gamma-1}}{\frac{2\gamma}{\gamma-1}M_{n_1}^2 - 1}}, \quad M_2 = \frac{M_{n_2}}{\sin(\beta-\theta)},$$

$$\rho_2 = \rho_1 \frac{(\gamma+1)M_{n_1}^2}{(\gamma-1)M_{n_1}^2 + 2}, \quad P_2 = P_1\left[1 + (\frac{2\gamma}{\gamma+1})(M_{n_1}^2 - 1)\right].$$

The direction of $\mathbf{u}$ is parallel to the wedge angle.

Fig. 9 displays the steady state solution to the wedge problem with $\theta = 15°$ and $M_1 = 2.5$. Note that on the left boundary we have the boundary conditions $\rho_1 = 1.4$ and $P_1 = 1$. $|\mathbf{u}|$ can be calculated from M, and $\mathbf{u}$ is in the x direction. On the top and bottom boundary, we have the boundary conditions

$$\mathbf{u} \cdot \hat{\mathbf{n}} = 0,$$

where $\hat{\mathbf{n}}$ is the outward unit normal. No boundary conditions are prescribed on the right boundary.

4.3. Set up of computational experiments

In the hierarchical finite element formulation (7), there are several free parameters that must be set.

Choosing ϵ_N. A natural choice for ϵ_N is $\epsilon_N = h_N$.

Choosing m so that $Q_{N;n,i} = 0$ *for* $n < m$. We were able to prove (see [6]) convergence when $Q_{N;n,i} = 1$ for $n > \frac{N}{2}$. Numerical experimentation indicates that one can add much less diffusion. We add diffusion to the finest one or two levels only: $m = N$ for one-dimensional problems and $m = N - 1$ for two-dimensional problems.

Choosing the form of Q_N. The spectral viscosity method yields better results if $Q_{N;n,i}$ is smooth with respect to the frequency number n. Spectral basis functions have many more levels than hierarchical finite element basis functions so that it is easier to achieve a smooth transition in the former case. Numerical evidence indicates that a sharp jump in $Q_{N;n,i}$ with respect to n does not adversely affect the hierarchical finite element calculations. $Q_{N;n,i} = 0$ for $n < m$ and then becomes 1 for $n \geqslant m$. In our two-dimensional calculations, we have chosen an isotropic diffusion term, as in (10). As we can see in (11), this greatly simplifies the formulation. As was mentioned in Section 3.1, one of the potential advantages of the hierarchical finite element approach is the ability to add diffusion only near a discontinuity. We have not yet explored the full exploitation of this advantage.

Choosing the quadratic hierarchical structure. As we saw in Fig. 2, there are two possible hierarchical structures for piecewise quadratic polynomials. In the first of these figures, all the basis functions are quadratic. In the second one, the low frequency basis functions are linear. The major difference between the two is that for the same Q_N and m, the linear low frequency structure will add more diffusion. For a piecewise quadratic function v, $Q_N v = \sum_{n,i} Q_{N;n,i}\beta_{n,i}\psi_{ni}$ with $Q_{N;n,i} = 0$ for small n. If all of the basis functions are quadratic, then in the smooth regions, $\beta_{n,i} = \mathrm{O}(8^{-n})$, as we showed in Section 3.2.2. If only the highest frequency basis functions are

quadratic, then for $n < N$, $\beta_{n,i} = \mathrm{O}(4^{-n})$ in the smooth regions, as we showed in Section 3.2.1. Therefore, $Q_N v$ will be larger for the linear low frequency basis functions. If diffusion is added to the finest level only (as we did in our one-dimensional calculations), then the hierarchical structure of the low frequencies do not have an effect.

Ordinary differential equation solver. Our finite element formulation turns the partial differential equations into a system of ordinary differential equations. We solved the system of ODEs using a third-order, strong stability-preserving (SSP) Runge–Kutta method found in [12]. The well known Courant, Friedrichs, and Lewy (CFL) condition restricts the size of the time step Δt. By trial and error, we choose a CFL number $\frac{\Delta t}{h_N} V_{\mathrm{m}} \leqslant 0.3$ for piecewise polynomials of degree 1 and a CFL number $\leqslant 0.1$ for degree 2. V_{m} is maximum wave propagation speed:

$$V_{\mathrm{m}} = \begin{cases} \max\{|q|\} & \text{for Burgers' equation,} \\ \max\{|u_j| + a\} & \text{for the Euler equations.} \end{cases}$$

Post-processing strategy. Because we are approximating a discontinuous solution with continuous piecewise polynomials, we see Gibbs oscillations near the discontinuity. A simple strategy to remove the oscillations is to set the coefficients of the hierarchical expansion to zero around the discontinuity. The question then becomes the location of the discontinuity. Let $\beta_{n+1,i}$ be a high frequency hierarchical coefficient. Let $\beta_{n,j}$ be the parent hierarchical coefficient, so the support of $\psi_{n+1,i}$ is a subset of $\psi_{n,j}$. If the solution is continuously differentiable in the region of the support of $\psi_{n,j}$, then we have that

$$\frac{\beta_{n,j}}{\beta_{n+1,i}} = \frac{C_1 2^{-n}}{C_2 2^{-n-1}} \approx 2.$$

Therefore, our strategy is: for a fixed (problem dependent) number of high frequencies, if a hierarchical coefficient is larger than half the value of its parent, then it is set to zero. The number of levels to post-process was found through experimentation: two for the steady state Burgers' equation and four for the periodic Burgers' equation.

Our simple strategy only affects the region around a discontinuity, but it has the disadvantage of smoothing across the discontinuity. Our solution therefore becomes more smeared.

We apply the post-processing strategy to solutions of the two variants of Burgers' equation, although it could also be applied to solutions of the Euler equations.

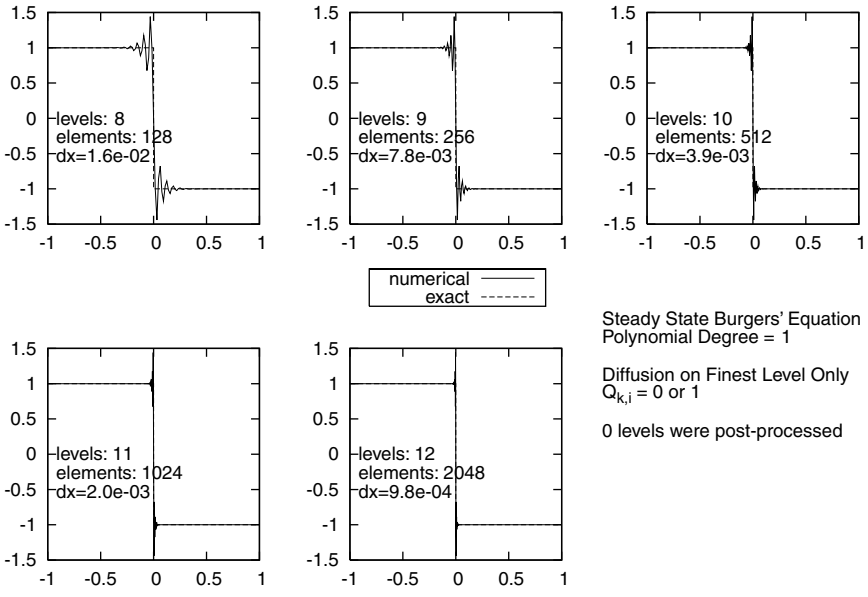

Fig. 10. Approximation of a steady state solution of Burgers' equations without any postprocessing.

Convergence rates. All our errors are given in the L^1 norm. Near a discontinuity, we are limited to how well a piecewise polynomial can approximate a solution. We are more interested in the convergence rates in the smooth regions. We therefore exclude a region around all the discontinuities in our error calculations. For the shock tube problem, the removed region has radius 0.05. For all other problems, the radius is 0.1.

For some regions where the solution is constant, the convergence rates are erratic because the L^1 error is almost zero; in these cases, the rates are omitted from the tables of results. In the domain of the steady state Burgers' equation and in regions 4 and 1 of the shock tube problem, the exact solution is constant, and the error is less than 10^{-13}.

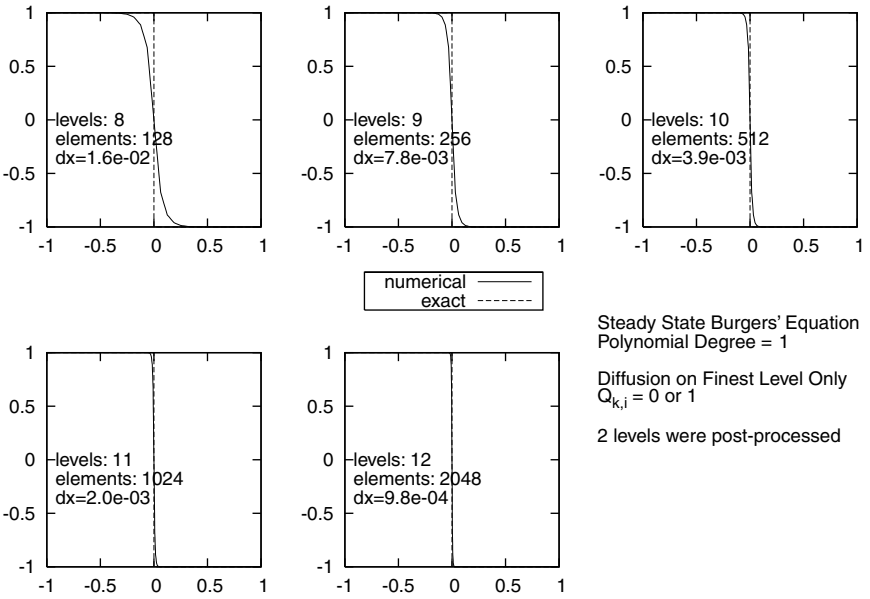

Fig. 11. Approximation of a steady state solution of Burgers' equations with simple postprocessing.

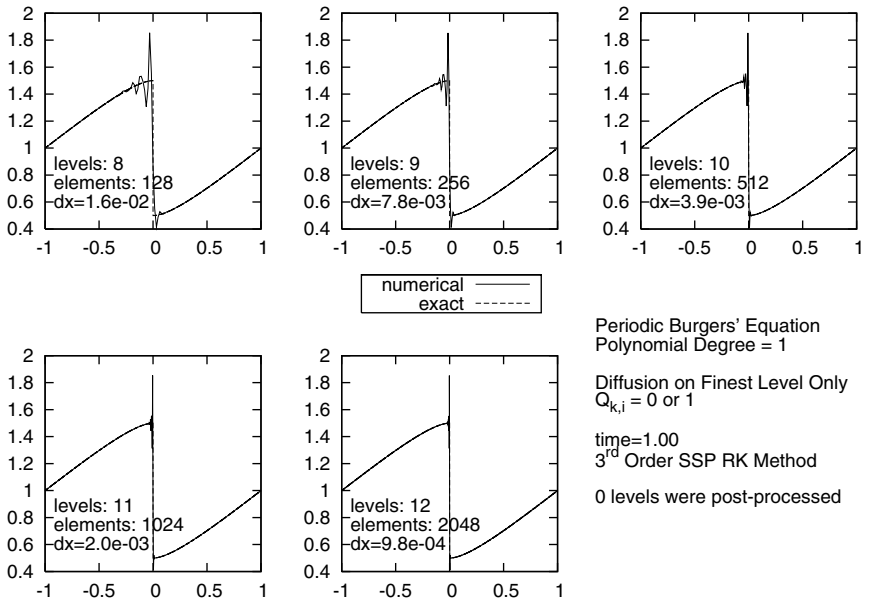

Fig. 12. Approximation of a periodic solution of Burgers' equations without any postprocessing.

For the shock tube problem, we observed different convergence rates in the five distinct regions depicted in Fig. 8. We therefore present the five different errors and convergence rates where appropriate. Similarly, for the wedge problem, we present the errors and convergence rates for the regions to the left and right of the shock.

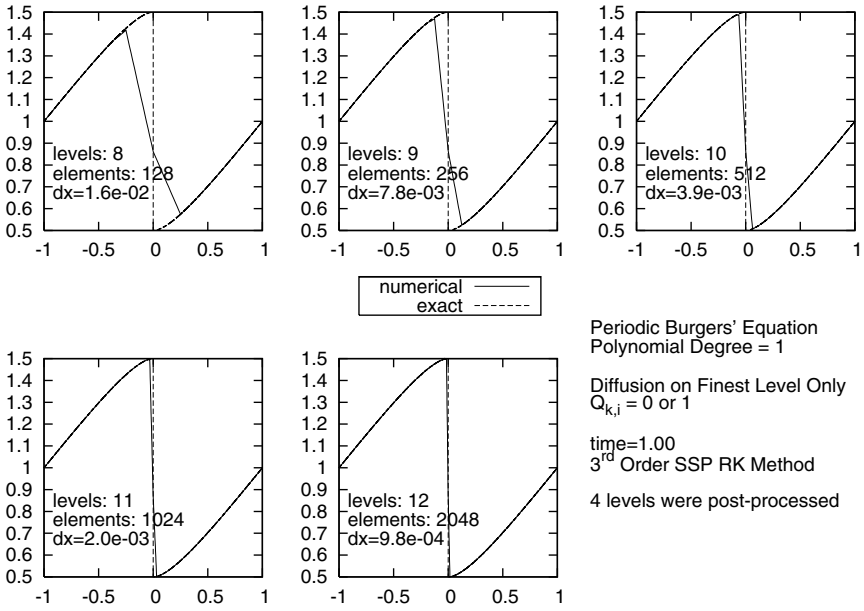

Fig. 13. Approximation of a periodic solution of Burgers' equations with simple postprocessing.

Table 3
Convergence rate for the periodic Burgers' equation with linear polynomials.

Levels	Without post-processing		With post-processing	
	L^1 error	Rate	L^1 error	Rate
8	4.2e−03	–	4.9e−02	–
9	3.5e−04	3.6	3.0e−03	4.0
10	1.3e−05	4.7	1.6e−05	7.5
11	7.7e−07	4.1	7.7e−07	4.4
12	1.9e−07	2.0	1.9e−07	2.0
13	4.8e−08	2.0	4.8e−08	2.0

Table 4
Convergence rate for the periodic Burgers' equation with quadratic polynomials

Levels	Without post-processing		With post-processing	
	L^1 error	Rate	L^1 error	Rate
8	2.7e−03	–	3.2e−02	–
9	2.5e−04	3.4	1.5e−03	4.4
10	4.2e−06	5.9	6.4e−06	7.9
11	6.4e−09	9.3	9.5e−09	9.4
12	4.8e−10	3.7	4.8e−10	4.3
13	6.2e−11	3.0	6.2e−11	3.0

4.4. Numerical simulations

In Figs. 10 and 11, we present results for the steady state Burgers' equation problem of Section 4.1.1. In smooth regions, the exact solution is constant; as a result, the L^1 error is almost zero and convergence rates

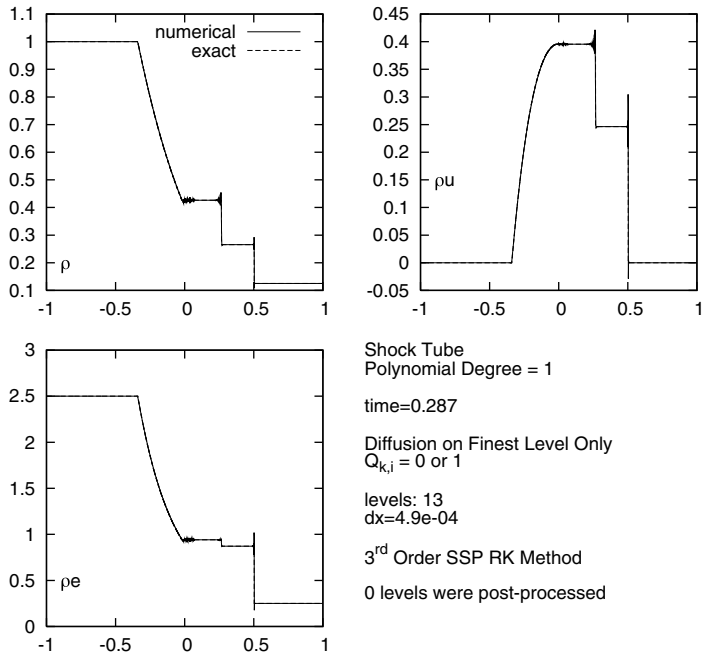

Fig. 14. Approximation of the solution of the shock-tube problem without any postprocessing.

Table 5
Convergence rate for the shock tube problem with linear polynomials

Levels	Region 4	Expansion Wv.		Region 3		Region 2		Region 1
	L^1 error	L^1 error	Rate	L^1 error	Rate	L^1 error	Rate	L^1 error
ρ								
8	1.3e−04	5.6e−03	–	3.7e−03	–	8.8e−04	–	1.1e−04
9	1.0e−05	2.9e−03	1.0	1.5e−03	1.3	1.2e−04	2.9	4.8e−06
10	2.8e−07	1.5e−03	1.0	8.3e−04	0.8	3.1e−05	2.0	2.1e−08
11	6.2e−10	7.4e−04	1.0	4.3e−04	0.9	5.9e−06	2.4	4.3e−13
12	8.9e−15	3.7e−04	1.0	1.8e−04	1.2	8.6e−07	2.8	2.3e−17
13	1.1e−15	1.9e−04	1.0	7.4e−05	1.3	2.9e−07	1.6	5.5e−17
ρu								
8	1.5e−04	3.4e−03	–	1.3e−03	–	1.8e−03	–	1.3e−04
9	1.2e−05	1.7e−03	1.0	6.1e−04	1.1	2.4e−04	2.9	4.7e−06
10	3.3e−07	8.6e−04	1.0	2.8e−04	1.1	4.2e−05	2.5	2.4e−08
11	7.4e−10	4.3e−04	1.0	8.9e−05	1.7	5.9e−06	2.8	4.4e−13
12	1.0e−14	2.2e−04	1.0	2.5e−05	1.8	1.0e−06	2.6	2.3e−18
13	4.9e−16	1.1e−04	1.0	8.7e−06	1.5	5.9e−07	0.8	2.5e−17
ρe								
8	4.4e−04	1.4e−02	–	6.1e−03	–	4.6e−03	–	3.3e−04
9	3.6e−05	7.4e−03	1.0	2.3e−03	1.4	6.0e−04	2.9	1.2e−05
10	9.7e−07	3.7e−03	1.0	1.3e−03	0.9	7.1e−05	3.1	6.3e−08
11	2.2e−09	1.9e−03	1.0	7.5e−04	0.8	4.7e−06	3.9	1.2e−12
12	3.0e−14	9.5e−04	1.0	3.5e−04	1.1	1.6e−06	1.5	1.7e−17
13	2.0e−15	4.7e−04	1.0	1.4e−04	1.3	1.5e−06	0.1	1.5e−16

are erratic. For this reason, we do not provide tables of the errors in the approximate solutions. The post-processing strategy removed the oscillations near the discontinuity.

In Figs. 12 and 13 and Tables 3 and 4, we present results for the periodic Burgers' equation problem of Section 4.1.2. In smooth regions, we see quasi-optimal convergence rates using linear and quadratic polyno-

Table 6
Convergence rate for the shock tube problem with quadratic polynomials.

Levels	Region 4	Expansion Wv.		Region 3		Region 2		Region 1
	L^1 error	L^1 error	Rate	L^1 error	Rate	L^1 error	Rate	L^1 error
ρ								
8	1.7e−05	2.3e−03	–	3.3e−03	–	5.7e−04	–	2.4e−05
9	7.3e−07	1.2e−03	1.0	2.3e−03	0.5	1.4e−04	2.0	1.2e−06
10	3.4e−09	5.9e−04	1.0	1.2e−03	0.9	4.4e−05	1.7	2.5e−08
11	2.0e−13	2.9e−04	1.0	5.1e−04	1.2	9.3e−06	2.3	4.3e−11
12	1.6e−14	1.5e−04	1.0	1.7e−04	1.6	1.6e−06	2.5	2.3e−16
13	1.5e−14	7.4e−05	1.0	3.3e−05	2.4	9.2e−07	0.8	4.5e−19
ρu								
8	2.0e−05	1.3e−03	–	8.6e−04	–	1.2e−03	–	2.5e−05
9	8.6e−07	6.6e−04	1.0	4.7e−04	0.9	2.2e−04	2.4	6.8e−07
10	4.0e−09	3.3e−04	1.0	1.9e−04	1.3	4.6e−05	2.2	2.7e−08
11	2.7e−13	1.7e−04	1.0	6.1e−05	1.6	8.9e−06	2.4	4.5e−11
12	1.3e−14	8.4e−05	1.0	1.7e−05	1.8	1.9e−06	2.2	1.6e−15
13	1.2e−14	4.2e−05	1.0	3.8e−06	2.2	1.0e−06	0.9	2.8e−15
ρe								
8	5.8e−05	5.8e−03	–	6.3e−03	–	2.9e−03	–	7.4e−05
9	2.5e−06	2.9e−03	1.0	4.2e−03	0.6	4.6e−04	2.7	3.2e−06
10	1.2e−08	1.5e−03	1.0	2.2e−03	0.9	4.5e−05	3.3	7.0e−08
11	6.9e−13	7.4e−04	1.0	9.8e−04	1.2	5.7e−06	3.0	1.2e−10
12	4.8e−14	3.7e−04	1.0	3.2e−04	1.6	2.5e−06	1.2	6.7e−16
13	4.2e−14	1.9e−04	1.0	6.2e−05	2.4	1.2e−06	1.0	5.8e−19

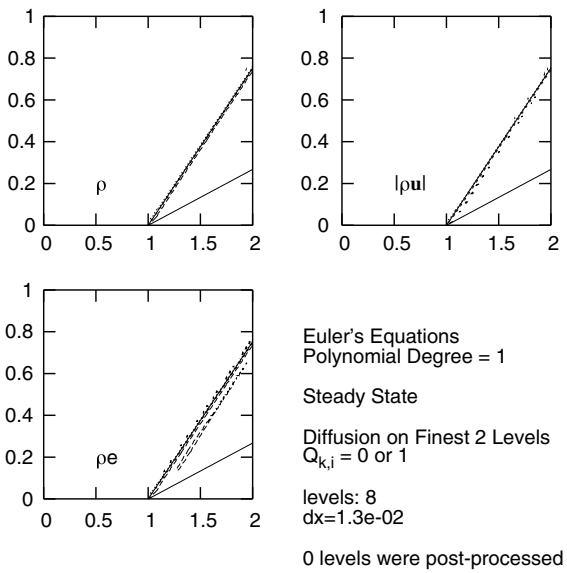

Fig. 15. Approximation of the solution of the wedge problem without any postprocessing.

mials and very sharp resolution of discontinuities. The post-processing strategy removed the oscillations near the discontinuity but resulted in considerable smearing of the discontinuity. More sophisticated post-processing strategies could remove oscillations while still preserving the sharpness at discontinuities.

In Fig. 14 and Tables 5 and 6, we present results for the shock tube problem of Section 4.2.1. In smooth regions where the solution is constant, the convergence rates are erratic, but the L^1 error is almost zero only in regions four and one. In regions three and two, the L^1 error is not as small as in regions four and one, but it is between 10 and 100 times smaller than the error in the expansion wave. The L^1 error in the expansion wave seems to be limited to first order. This is most likely an example of the phenomenon known as "downstream pollution" which similarly affects other numerical solvers of hyperbolic conservation laws. Note that the approximate solution could be postprocessed to remove oscillations.

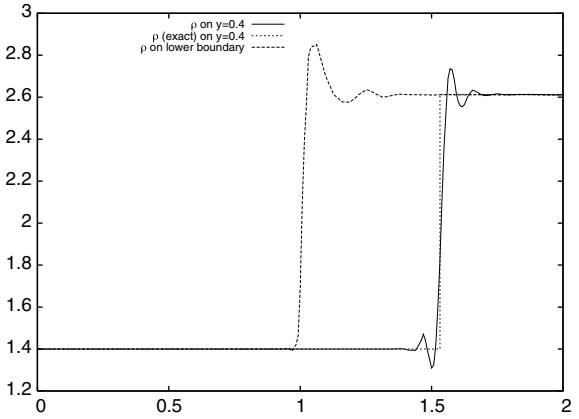

Fig. 16. Approximate density profiles for the wedge problem without any postprocessing.

Table 7
Convergence rate of wedge problem with linear polynomials.

Levels	L^1 error		L^1 error	
	Left region	Rate	Right region	Rate
ρ				
5	2.4e−02	–	7.8e−02	–
6	7.1e−03	1.8	5.3e−02	0.6
7	1.0e−03	2.8	3.3e−02	0.7
8	3.1e−05	5.1	1.8e−02	0.9
ρu_1				
5	4.9e−02	–	1.2e−01	–
6	1.5e−02	1.7	8.5e−02	0.5
7	2.1e−03	2.8	5.3e−02	0.7
8	6.5e−05	5.0	3.0e−02	0.8
ρu_2				
5	2.4e−02	–	9.1e−02	–
6	6.6e−03	1.8	6.2e−02	0.6
7	8.8e−04	2.9	3.8e−02	0.7
8	2.7e−05	5.0	2.1e−02	0.9
ρe				
5	1.1e−01	–	3.4e−01	–
6	3.2e−02	1.8	2.3e−01	0.6
7	4.5e−03	2.8	1.4e−01	0.7
8	1.4e−04	5.0	7.8e−02	0.9

In Figs. 15 and 16 and Table 7, we present results for the two-dimensional wedge problem of Section 4.2.2. To the right of the shock, we have first-order convergence. This is most likely another example of downstream pollution, which we saw in our solution of the shock tube problem. To the left of the shock, we see high-order convergence.

5. Concluding remarks

The initial results for this new method are promising. We have a stable finite element method which, in some cases, attains quasi-optimal convergence rates in smooth regions. We also have developed a theoretical foundation for understanding why this method works. These results, however, are preliminary. There are potential pitfalls awaiting in more complicated problems, but there is also untapped potential in the framework. In particular, we have yet to take full advantage of several important features of the new method, including

- the ease for which singularities in the solution can be detected (see Section 3.2);
- the ease for which one can apply the artificial viscosity only in the vicinity of singularities, i.e., *locally in physical space as well as in frequency space*;
- the ease for which one can apply the artificial diffusion in an anisotropic manner, e.g., in a streamwise direction.

In addition, as for other methods, e.g., the spectral viscosity method, post-processing of the approximate solution is necessary in order to recover full accuracy. So far, we have only applied simple and straightforward post-processing strategies; more sophisticated post-processing approaches will surely prove to be more effective. Fortunately, in this regard, we can avail ourselves of some of the post-processing strategies that have been developed for the spectral viscosity method.

We close by listing several *specific* directions for future research that are needed to further develop and validate the approach we introduced in this paper and to render that approach truly competitive with methods that have a much longer history.

Find an efficient adaptive diffusion form of Q_N. The operator $Q^{j,k}_{N;n,i}$ we used did not depend on i. It was done implicitly since for large n, $\beta_{n,i}$ (and so $Q^{j,k}_{N;n,i}\beta_{n,i}$) is small in the smooth regions and large near a discontinuity. An adaptive diffusion operator should improve the convergence rates in the smooth regions.

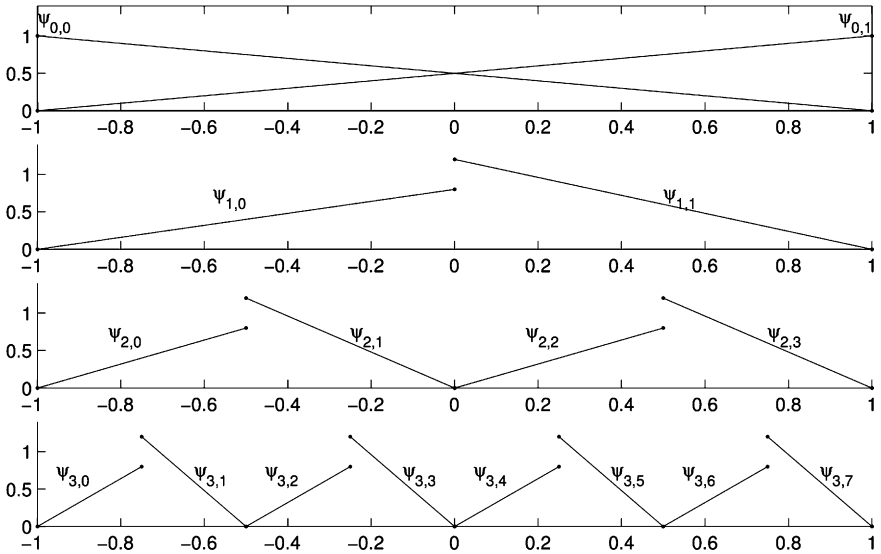

Fig. 17. Discontinuous linear hierarchical basis functions. Top to bottom: level 0, 1, 2, and 3 basis functions.

Implement an anisotropic diffusion term to minimize crosswind diffusion. The diffusion operator in our calculations was isotropic. In other words, our $Q^{j,k}_{N;n,i}$ is zero if $j \neq k$. We would like to implement an anisotropic diffusion term to minimize crosswind diffusion.

Implement adaptive grid refinement. In all of our calculations, we used uniform grids. Clearly, we would like to have a finer grid near a discontinuity. The ease of edge detection gives us a good way of knowing where more elements are needed. Hierarchical bases are well suited to adaptive refinement because adding another level does not affect the coefficient values of the previous levels.

Implement discontinuous basis functions. Multi-resolution viscosity ideas can be incorporated into the discontinuous Galerkin framework by using discontinuous hierarchical basis functions, as depicted in Fig. 17.

Impose boundary conditions weakly. By writing the flux term in our finite element formulation (7) as

$$-\int_U \sum_{j=1}^{d} \mathbf{f}_j(\mathbf{q}^N) \cdot \frac{\partial \mathbf{v}}{\partial x_j}\, \mathrm{d}x + \int_{\partial U} \sum_{j=1}^{d} \mathbf{f}_j(\mathbf{q}^N) \cdot \mathbf{v}\widehat{n}_j\, \mathrm{d}s$$

and replacing the resulting boundary integral with a suitable numerical flux, boundary conditions could be imposed weakly. This formulation would be more in keeping with finite volume methods.

Find a better post-processing strategy. Setting hierarchical high-frequency coefficients to zero near a discontinuity removes oscillations, but it also smoothes across the discontinuity, adding more diffusion. A better strategy would be to dampen the high frequency coefficient instead of setting them to zero. We would like to find an automatic process which finds what values to set the high frequency coefficients to near the discontinuity.

Run additional two- and three-dimensional test problems, especially on curved domains. The problems we have chosen are good initial test problems, but more realistic problems are required. We would like to implement this method on curved domains (such as airplane wings) and in three dimensions.

Perform comparisons with other numerical methods. We would like to perform rigorous comparisons with other methods to see if our method is competitive.

Acknowledgments

All of our numerical results were generated using the finite element library `deal.II` [5]. The authors thank Eitan Tadmor and Curt Ober for many helpful discussions.

References

[1] J. Anderson, Modern Compressible Flow with Historical Perspective, McGraw-Hill, New York, 1990.
[2] B. Azarenok, T. Tang, Second-order Godunov-type scheme for reactive flow calculations on moving meshes, J. Comput. Phys. 206 (2005) 48–80.
[3] R. Bank, Hierarchical bases and the finite element method, Acta Numerica, vol. 5, Cambridge University Press, Cambridge, 1996.
[4] R. Bank, J. Xu, An algorithm for coarsening unstructured meshes, Numer. Math. 73 (1996) 1–36.
[5] W. Bangerth, R. Hartmann, G. Kanschat, `deal.II` Differential equations analysis library, Technical Reference, IWR, Heidelberg. http://www.dealii.org.
[6] M. Calhoun-Lopez, M. Gunzburger, A finite element multi-resolution viscosity method for hyperbolic conservation laws, SIAM J. Numer. Anal. 43 (2005) 1988–2011.
[7] G. Chen, Q. Du, E. Tadmor, Spectral viscosity approximations to multidimensional scalar conservation laws, Math. Comp. 61 (1993) 629–643.
[8] B. Cockburn, G. Karniadakis, C.-W. Shu, The development of discontinuous Galerkin methods, Discontinuous Galerkin Methods Theory, Computation, and Applications, Springer, Berlin, 2000, pp. 3–50.
[9] L. Evans, Partial Differential Equations, AMS, Providence, 1998.
[10] A. Gelb, E. Tadmor, Enhanced spectral viscosity approximations for conservation laws, Appl. Numer. Math. 33 (2000) 3–21.
[11] E. Godlewski, P.-A. Raviart, Numerical Approximation of Hyperbolic Systems of Conservation Laws, Springer, New York, 1996.
[12] S. Gottlieb, C.-W. Shu, E. Tadmor, Strong stability-preserving high-order time discretization methods, SIAM Rev. 43 (2001) 89–112.
[13] B.-Y. Guo, H.-P. Ma, E. Tadmor, Spectral vanishing viscosity method for nonlinear conservation laws, SIAM J. Numer. Anal. 39 (2001) 1254–1268.
[14] C. Johnson, A. Szepessy, On the convergence of a finite element method for a nonlinear hyperbolic conservation law, Math. Comp. 49 (1987) 427–444.

[15] C. Johnson, A. Szepessy, P. Hansbo, On the convergence of shock-capturing streamline diffusion finite element methods for hyperbolic conservation laws, Math. Comp. 54 (1990) 107–129.
[16] R. Kornhuber, H. Yserentant, Multilevel methods for elliptic problems on domains not resolved by the coarse grid, Domain Decomposition Methods in Scientific and Engineering Computing, AMS, Providence, 1994, pp. 49–60.
[17] P. Lax, Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves, SIAM, Philadelphia, 1973.
[18] R. LeVeque, Numerical Methods for Conservation Laws, Birkhäuser, Basel, 1992.
[19] R. LeVeque, Finite Volume Methods for Hyperbolic Problems, Cambridge University Press, Cambridge, 2002.
[20] Y. Maday, S. Ould Kaber, E. Tadmor, Legendre pseudospectral viscosity method for nonlinear conservation laws, SIAM J. Numer. Anal. 30 (1993) 321–342.
[21] Y. Maday, E. Tadmor, Analysis of the spectral vanishing viscosity method for periodic conservation laws, SIAM J. Numer. Anal. 26 (1989) 854–870.
[22] B. Perthame, E. Tadmor, A kinetic equation with kinetic entropy functions for scalar conservation laws, Commun. Math. Phys. 136 (1991) 501–517.
[23] J. Stockie, J. Mackenzie, R. Russell, A moving mesh method for one-dimensional hyperbolic conservation laws, SIAM J. Sci. Comput. 22 (2000) 1791–1813.
[24] A. Szepessy, Convergence of a shock-capturing streamline diffusion finite element method for a scalar conservation law in two space dimensions, Math. Comp. 53 (1989) 527–545.
[25] E. Tadmor, Convergence of spectral methods for nonlinear conservation laws, SIAM J. Numer. Anal. 26 (1989) 30–44.
[26] E. Tadmor, Total variation and error estimates for spectral viscosity approximations, Math. Comp. 60 (1993) 245–256.
[27] E. Tadmor, Super-viscosity and spectral approximations of nonlinear conservation laws, Numerical Methods for Fluid Dynamics, vol. IV, Oxford, New York, 1993.
[28] E. Tadmor, Approximate solutions of nonlinear conservation laws and related equations, Recent Advances in Partial Differential Equations, AMS, Providence, 1998, pp. 325–368.
[29] H. Tang, T. Tang, Adaptive mesh methods for one- and two-dimensional hyperbolic conservation laws, SIAM J. Numer. Anal. 41 (2003) 487–515.
[30] C. Van Loan, Introduction to Scientific Computing, Prentice-Hall, Englewood Cliffs, NJ, 2000.
[31] H. Yserentant, On the multilevel splitting of finite element spaces, Numer. Math. 49 (1986) 379–412.
[32] H. Yserentant, Hierarchical bases give conjugate gradient type methods a multigrid speed of convergence, Appl. Math. Comp. 19 (1986) 347–358.
[33] H. Yserentant, Hierarchical bases, Proc. ICIAM 91, SIAM, Philadelphia, 1992, pp. 256–276.
[34] H. Yserentant, Coarse grid spaces for domains with a complicated boundary, Numer. Algor. 21 (1999) 387–392.
[35] O. Zienkiewicz, D. Kelly, J. Gago, I. Babuška, Hierarchical finite element approaches, error estimates and adaptive refinement, The Mathematics of Finite Elements and Applications, vol. IV, Academic Press, New York, 1982.